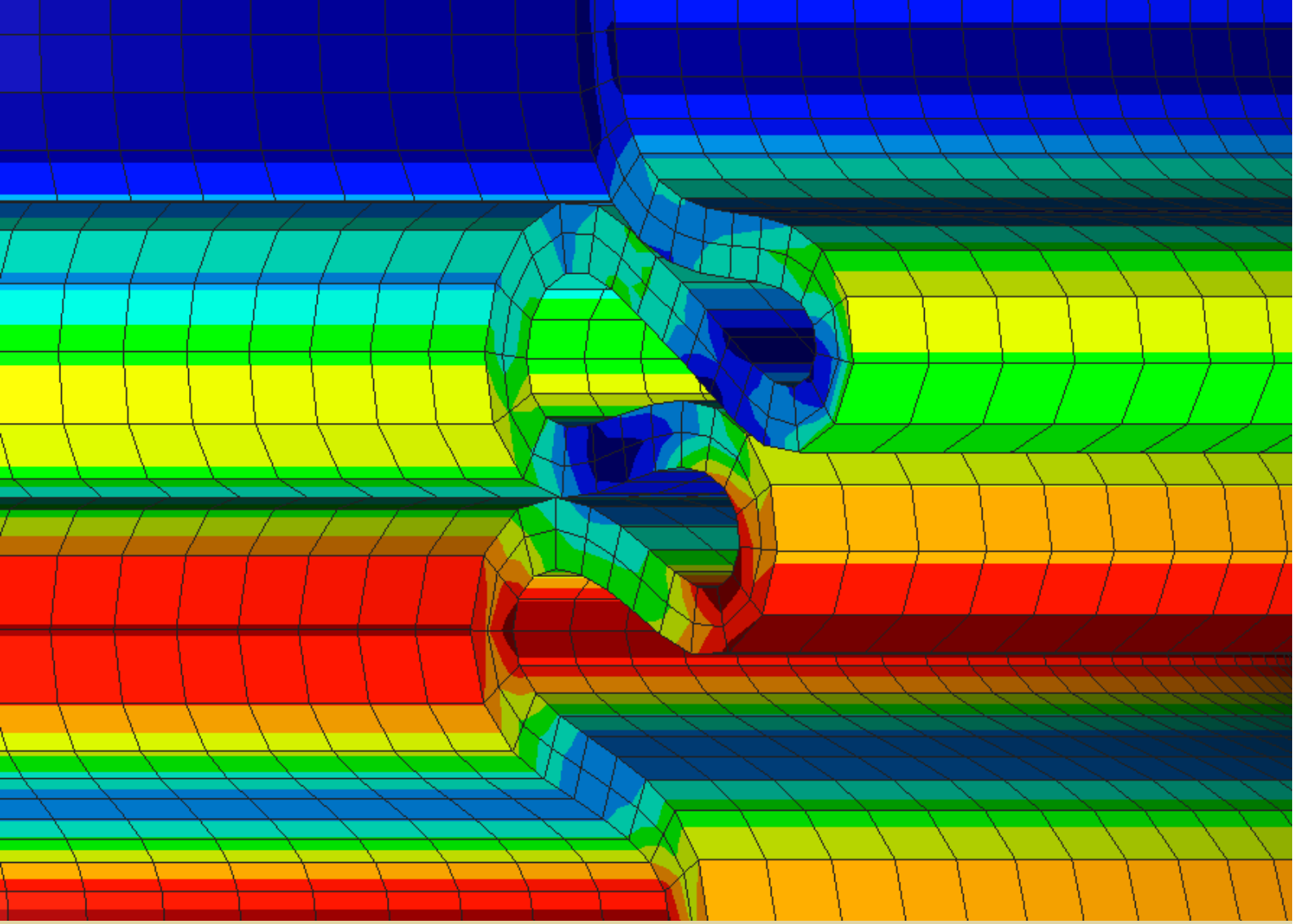


Diogo Mariano Simões Neto

NUMERICAL SIMULATION OF FRICTIONAL CONTACT PROBLEMS USING NAGATA PATCHES IN SURFACE SMOOTHING

Doctoral Thesis in Mechanical Engineering, specialization in Production Technologies, supervised by Professor Marta Cristina Cardoso de Oliveira and Professor Luís Filipe Martins Menezes, submitted to the Department of Mechanical Engineering, Faculty of Sciences and Technology of the University of Coimbra

September 2014

· U C ·

UNIVERSIDADE DE COIMBRA

· U  C ·

Numerical Simulation of Frictional Contact Problems using Nagata Patches in Surface Smoothing

A dissertation submitted to the Department of Mechanical Engineering of the University of Coimbra in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Mechanical Engineering

Diogo Mariano Simões Neto

September 2014

Scientific advisers:

Prof. **Marta Cristina Cardoso de Oliveira**, University of Coimbra, Portugal

Prof. **Luís Filipe Martins Menezes**, University of Coimbra, Portugal

Dissertation defence committee:

President	Prof. **José Valdemar Bidarra Fernandes** Full Professor, University of Coimbra, Portugal
Reviewer	Prof. **Pierre-Yves Manach** Full Professor, University of Southern Brittany, France
Reviewer	Prof. **José Manuel de Almeida César de Sá** Full Professor, University of Porto, Portugal
Reviewer	Prof. **Paulo António Firme Martins** Full Professor, University of Lisbon, Portugal
Reviewer	Prof. **José Luís de Carvalho Martins Alves** Assistant Professor, University of Minho, Portugal
Reviewer	Prof. **Amílcar Lopes Ramalho** Associate Professor, University of Coimbra, Portugal
Member	Prof. **Marta Cristina Cardoso de Oliveira** Assistant Professor, University of Coimbra, Portugal

To my family

"Everything must be made as simple as possible. But not simpler."
Albert Einstein

Acknowledgements

The research work present in this PhD thesis has been carried out in the last four years at the Department of Mechanical Engineering of the University of Coimbra. These four years of research were a very fruitful period in my engineering education, as well as in my personal development. The freedom in the investigation accompanied by a wise guidance of my scientific advisers allowed me to improve the knowledge in the field of computational contact mechanics.

I particularly wish to express my sincere gratitude to my scientific advisers Professor Marta Oliveira and Professor Luís Filipe Menezes, for the valuable guidance, support and consistent encouragement during my PhD thesis. I highly appreciate their outstanding scientific expertise and experience (shared with me), as well as the sincere friendship, trust and incessant care to provide me with the best possible working conditions. Their drive for scientific rigor have been indispensable to the successful and timely conclusion of this thesis and, at the same time, a great source of inspiration for me.

I would like to express my special thanks to Professor José Luís Alves for the fruitful discussions and critical remarks concerning the numerical implementation of the developed algorithms in DD3IMP finite element code, as well as for his friendship. His computational expertise, insightful comments and suggestions on the research direction were crucial for the success of this work.

I am grateful for the excellent support given within the Mechanical Engineering Centre of the University of Coimbra (CEMUC) chaired by Professor Valdemar Fernandes and the opportunity that was given to me to develop my research work, as well as the financial support to participate in some conferences. I also wish to thank all members of CEMUC for the friendly and scientifically inspiring atmosphere at the research centre. Additional thanks go to the Advanced Manufacturing Systems research group chaired by Professor Altino Loureiro, which has provided the computational resources necessary to run the numerical simulations.

I would like to thank my colleagues of office and friends Álvaro Mendes, André Pereira, André Rodrigues, Henrique Santos, João Martins, Luis Reis, Pedro Barros, Pedro Prates, Rui Leal and Vasco Simões for their friendship and the many inspiring (scientific and several non-scientific) discussions we had. I wish to thank all my friends for their encouragement, in particular the laughs of Rafael Raposo and the deep knowledge in computer programming of the new sportsman Ricardo Heleno. I cannot forget to thank my big French friend Jérémy Coër for his unconditional friendship, help and some useful scientific discussions about sheet metal forming process.

Finally, I like to thank my family for the love and support they gave me, in particular to my parents and my brother for their encouragement, personal sacrifice and understanding through all the years. Without them it would not been possible to finish this work.

The financial support of this work has been provided by the Portuguese Foundation for Science and Technology (FCT) through the PhD Grant SFRH/BD/69140/2010, which is greatly acknowledged.

Coimbra, September 2014

Resumo

Todos os movimentos no mundo envolvem contato e atrito, desde andar até conduzir um carro. A mecânica do contacto tem aplicação em muitos problemas de engenharia, incluindo a ligação de elementos estruturais com parafusos, projeto de engrenagens e rolamentos, estampagem ou forjamento, contato entre os pneus e a estrada, colisão de estruturas, bem como o desenvolvimento de próteses em engenharia biomédica. Devido à natureza não-linear e não-suave da mecânica do contato (área de contato desconhecida *a priori*), tais problemas são atualmente resolvidos usando o método dos elementos finitos no domínio da mecânica do contato computacional. No entanto, a maioria dos programas comerciais de elementos finitos atualmente disponíveis não é totalmente capaz de resolver problemas de contato com atrito, exigindo métodos numéricos mais eficientes e robustos. Portanto, o principal objetivo deste estudo é o desenvolvimento de algoritmos e métodos numéricos para aplicar na simulação numérica de problemas de contato com atrito entre corpos envolvendo grandes deformações. Os desenvolvimentos apresentados são implementados no programa de elementos finitos DD3IMP.

A formulação *quasi*-estática de problemas de contato com atrito entre corpos deformáveis envolvendo grandes deformações é primeiramente apresentada no âmbito da mecânica dos meios contínuos, seguindo o método clássico usado em mecânica dos sólidos. A descrição cinemática dos corpos deformáveis é apresentada adotando uma formulação Lagrangeana reatualizada. O comportamento mecânico dos corpos é descrito por uma lei constitutiva elastoplástica em conjunto com uma lei de plasticidade associada, permitindo modelar uma grande variedade de problemas de contacto envolvidos em aplicações industriais. O contacto com atrito entre os corpos é definido por duas condições: o princípio da impenetrabilidade e a lei de atrito de Coulomb, ambas impostas na interface de contato. O método do Lagrangeano aumentado é aplicado na resolução do problema de minimização com restrições resultantes das condições de contato e atrito, produzindo uma formulação mista envolvendo deslocamentos e forças de contato.

A discretização espacial dos corpos é realizada com elementos finitos sólidos isoparamétricos, enquanto a discretização da interface de contacto é realizado utilizando a técnica *Node-to-Segment*, impedindo os nós *slave* de penetrar na superfície *master*. A parte geométrica do elemento de contacto, definida por um nó *slave* e o segmento *master* mais próximo, é criada pelo algoritmo de deteção de contacto com base na seleção do ponto mais próximo na superfície *master*, obtido pela projeção normal do nó *slave*. No caso particular de contato entre um corpo deformável e um obstáculo rígido, a superfície rígida *master* pode ser descrita por parametrizações normalmente utilizadas em modelos CAD. No entanto, no caso geral de contato entre corpos deformáveis, a discretização espacial dos corpos com elementos finitos lineares origina uma representação da superfície *master* por facetas. Esta é a principal fonte de problemas na resolução de problemas de contato envolvendo grandes escorregamentos, uma vez que a distribuição dos vetor normais à superfície é descontínua. Assim, é proposto um método de suavização para descrever as superfícies de contacto *master* baseado na interpolação Nagata, que conduziu ao desenvolvimento do elemento de contacto *Node-to-Nagata*. A precisão do método de suavização das superfícies é avaliada através de geometrias simples. Os vetores normais nodais necessários para a interpolação Nagata são avaliados a partir da geometria CAD no caso de superfícies rígidas, enquanto no caso de corpos deformáveis são aproximados utilizando a média ponderada dos vetores normais das facetas vizinhas. Tanto os vetores de segundo membro como as matrizes residuais tangentes dos elementos de contacto são obtidas de forma coerente com o método de suavização da superfície, enquanto o método de Newton generalizado é utilizado para resolver o sistema de equações não lineares.

O método de suavização das superfícies e os elementos de contacto desenvolvidos são validados através de exemplos com soluções analíticas ou semi-analíticas conhecidas. Também são estudados outros problemas de contato mais complexos, incluindo o contato de um corpo deformável com obstáculos rígidos e o contato entre corpos deformáveis, contemplando fenómenos de auto-contato. A suavização da superfície *master* melhora a robustez dos métodos computacionais e reduz fortemente as oscilações na força de contato, associadas à descrição facetada da superfície de contato. Os resultados são comparados com soluções numéricas de outros autores e com resultados experimentais, demonstrando a precisão e o desempenho dos algoritmos implementados para problemas fortemente não-lineares.

Palavras-chave: Mecânica do Contacto, Método dos Elementos Finitos, Método do Lagrangeano Aumentado, Grandes Escorregamentos, Node-to-Segment, Suavização de Superfícies, Superfícies Nagata.

Abstract

All movements in the world involve contact and friction, from walking to car driving. The contact mechanics has application in many engineering problems, including the connection of structural members by bolts or screws, design of gears and bearings, sheet metal or bulk forming, rolling contact of car tyres, crash analysis of structures, as well as prosthetics in biomedical engineering. Due to the nonlinear and non-smooth nature of contact mechanics (contact area is not known *a priori*), such problems are currently solved using the finite element method within the field of computational contact mechanics. However, most of the commercial finite element software packages presently available are not entirely capable to solve frictional contact problems, demanding for efficient and robust methods. Therefore, the main objective of this study is the development of algorithms and numerical methods to apply in the numerical simulation of 3D frictional contact problems between bodies undergoing large deformations. The presented original developments are implemented in the in-house finite element code DD3IMP.

The formulation of *quasi*-static frictional contact problems between bodies undergoing large deformations is firstly presented in the framework of the continuum mechanics, following the classical scheme used in solid mechanics. The kinematic description of the deformable bodies is presented adopting an updated Lagrangian formulation. The mechanical behaviour of the bodies is described by an elastoplastic constitutive law in conjunction with an associated flow rule, allowing to model a wide variety of contact problems arising in industrial applications. The frictional contact between the bodies is established by means of two conditions: the principle of impenetrability and the Coulomb's friction law, both imposed to the contact interface. The augmented Lagrangian method is applied for solving the constrained minimization incremental problem resulting from the frictional contact inequalities, yielding a mixed functional involving both displacements and contact forces.

The spatial discretization of the bodies is performed with isoparametric solid finite elements, while the discretization of the contact interface is carried out using the classical

Node-to-Segment technique, preventing the slave nodes from penetrating on the master surface. The geometrical part of the contact elements, defined by a slave node and the closest master segment, is created by the contact search algorithm based on the selection of the closest point on the master surface, defined by the normal projection of the slave node. In the particular case of contact between a deformable body and a rigid obstacle, the master rigid surface can be described by smooth parameterizations typically used in CAD models. However, in the general case of contact between deformable bodies, the spatial discretization of both bodies with low order finite elements yields a piecewise bilinear representation of the master surface. This is the central source of problems in solving contact problems involving large sliding, since it leads to the discontinuity of the surface normal vector field. Thus, a surface smoothing procedure based on the Nagata patch interpolation is proposed to describe the master contact surfaces, which led to the development of the Node-to-Nagata contact element. The accuracy of the surface smoothing method using Nagata patches is evaluated by means of simple geometries. The nodal normal vectors required for the Nagata interpolation are evaluated from the CAD geometry in case of rigid master surfaces, while in case of deformable bodies they are approximated using the weighted average of the normal vectors of the neighbouring facets. The residual vectors and tangent matrices of the contact elements are derived coherently with the surface smoothing approach, while the generalized Newton method is used for solving the nonlinear system of equations.

The developed surface smoothing method and corresponding contact elements are validated through standard numerical examples with known analytical or semi-analytical solutions. More advanced frictional contact problems are studied, covering the contact of a deformable body with rigid obstacles and the contact between deformable bodies, including self-contact phenomena. The smoothing of the master surface improves the robustness of the computational methods and reduces strongly the non-physical oscillations in the contact force introduced by the traditional faceted description of the contact surface. The presented results are compared with numerical solutions obtained by other authors and experimental results, demonstrating the accuracy and performance of the implemented algorithms for highly nonlinear problems.

Keywords: Contact Mechanics, Finite Element Method, Augmented Lagrangian Method, Large Sliding, Node-to-Segment Discretization, Surface Smoothing, Nagata Patches.

Résumé

Tous les mouvements dans le monde impliquent du contact et du frottement, depuis le simple fait de marcher à celui de conduire une voiture. La mécanique du contact trouve ses applications dans de nombreux problèmes d'ingénierie, comprenant par exemple la liaison d'éléments structurants par vis/écrou, la conception des engrenages et des roulements, l'emboutissage ou le forgeage, le contact entre les pneus et la route, l'analyse de la collision des structures, ou encore le développement de prothèses en ingénierie biomédicale. En raison de la nature non linéaire et non lisse du contact mécanique (la surface de contact exacte est a priori inconnue), de tels problèmes sont actuellement résolus en utilisant la méthode des éléments finis, dans le domaine du calcul scientifique de la mécanique du contact. Cependant, la majorité des programmes de calcul par éléments finis commerciaux, actuellement disponibles, n'est pas pleinement en mesure de résoudre les problèmes de contact avec frottement, qui exigent des méthodes numériques plus efficaces et plus robustes. Par conséquent, le principal objectif de cette étude est le développement d'algorithmes et de méthodes numériques, qui permettront de réaliser la simulation numérique de problèmes de contact avec frottement entre des corps impliquant de grandes déformations. Les développements présentés sont implémentés dans le code de calcul par éléments finis DD3IMP.

La formulation de problèmes de contact *quasi*-statique avec frottement entre corps déformables en grandes déformations est dans, un premier temps, présentée dans le cadre de la mécanique des milieux continus, en suivant le schéma classique utilisé en mécanique du solide. La description cinématique des corps déformables est présentée en adoptant une formulation Lagrangienne réactualisée. Le comportement mécanique des corps est décrit par une loi de comportement élasto-plastique, en conjonction avec une loi d'écoulement plastique associée, permettant de modéliser une grande variété des problèmes de contact survenants dans des applications industrielles. Le contact avec frottement établi entre les corps est défini par deux conditions : le principe de non-pénétration et la loi de frottement de Coulomb, toutes deux imposés à l'interface de contact. La méthode du Lagrangien augmenté est utilisée dans la résolution du problème de minimisation avec restriction

résultantes des conditions du contact et de friction, qui produisent une formulation mixte impliquant les déplacements et les forces de contact.

La discrétisation spatiale des corps est faite avec des éléments finis solides isoparamétriques, tandis que la discrétisation de l'interface de contact est réalisée en utilisant la technique du *Node-to-Segment*, empêchant ainsi aux nœuds esclaves de pénétrer la surface maître. La partie géométrique de l'élément de contact, définie par un nœud esclave et le segment principal le plus proche, est créée par l'algorithme de détection du contact sur la base de la sélection du point le plus proche de la surface maître, obtenue par la projection normale du nœud esclave. Dans le cas particulier du contact entre un corps déformable et un obstacle rigide, la surface rigide maître peut être décrite par des paramétrisations couramment utilisés dans les modèles de CAO. Toutefois, dans le cas général d'un contact entre corps déformables, la discrétisation spatiale des corps avec des éléments finis linéaires engendre une représentation bilinéaire de la surface maître par facettes. Il s'agit là de la principale source de difficultés dans la résolution des problèmes de contact impliquant de grands glissements, étant donné qu'elle conduit à une distribution discontinue du champ des vecteurs normaux à la surface. Ainsi, il est proposé une méthode de lissage pour décrire les surfaces de contact maître basée sur l'interpolation Nagata, qui a conduit à l'élaboration d'un élément de contact *Node-to-Nagata*. La précision de cette méthode de lissage des surfaces est évaluée au travers de géométries simples. Les vecteurs normaux nodaux nécessaires à l'interpolation Nagata sont évalués à partir de la géométrie CAO dans le cas de surfaces rigides, alors que dans le cas des corps déformables ils sont estimés à l'aide de la moyenne pondérée des normales des facettes voisines. Les vecteurs résiduels et les matrices tangentes des éléments de contact sont dérivés de manière cohérente avec la méthode de lissage de surface, tandis que la méthode de Newton généralisée est utilisée pour la résolution du système d'équations non linéaires.

La méthode de lissage des surfaces et les éléments de contact développés sont d'abord validés à travers des exemples numériques standards, dont les solutions analytiques ou semi-analytiques sont connues. D'autres problèmes de contact frottant plus évolués sont ensuite étudiés, en traitant le contact d'un corps déformable avec des obstacles rigides et le contact entre corps déformables, incluant les phénomènes d'auto-contact. Le lissage de la surface maître améliore la robustesse des méthodes de calcul et réduit fortement les oscillations non physiques de la force de contact, induite par la description traditionnelle facettisée de la surface de contact. Les résultats présentés sont comparés aux solutions numériques obtenues par d'autres auteurs ainsi qu'à des résultats expérimentaux, démontrant alors la précision et les performances des algorithmes mis en œuvre pour des problèmes fortement non linéaires.

Mots-Clés: Mécanique du Contact, Méthode des éléments finis, Méthode du Lagrangien augmenté, Grand glissements, Discrétisation Node-to-Segment, Lissage de surfaces, Patch de Nagata.

Contents

List of Figures

List of Tables

Nomenclature

The symbols used in this thesis are separated into Greek and Roman category. Throughout this text, preference is given to the use of compact tensor notation, where no indices are used to represent mathematical entities.

Scalar (zero order tensor) – lower case italic Latin and Greek letters

Vector (first order tensor) – lower case bold Latin and Greek letters

Second order tensor – upper case bold Latin and Greek letters

Operations

δ_{ij}	Kronecker delta
$\lvert a \rvert$	Absolute value of a scalar
$\lVert \mathbf{a} \rVert$	Euclidean norm of a vector
$\mathbf{A}^{-1}$	Inverse of a second order tensor
$\mathbf{A}^{\mathrm{T}}$	Transpose of a second order tensor
$\mathbf{A}^{-\mathrm{T}}$	Transpose of the inverse of a second order tensor
$\mathbf{a} \cdot \mathbf{b}$	Inner or scalar product of two vectors
$\mathbf{a} \times \mathbf{b}$	Cross or vector product of two vectors
$\mathbf{a} \otimes \mathbf{b}$	Tensor or dyadic product of two vectors
$\mathbf{A} : \mathbf{B}$	Double contraction of two second order tensors
$\det(\mathbf{A})$	Determinant of a second order tensor
$\mathrm{cond}(\bullet)$	Condition number of a matrix
$\arg\min(\bullet)$	Minimum value of the argument function
$\mathrm{int}(\bullet)$	Integer part of a scalar
$\min(\bullet)$	Minimum of a function
$\max(\bullet)$	Maximum of a function

$\text{sign}(\bullet)$	Sign function
$\text{proj}_{\Re^-}(\bullet)$	Projection in the negative half-line
$\text{div}(\bullet)$	Divergence operator
$\langle\bullet\rangle$	Macaulay brackets
$\nabla(\bullet)$	Gradient operator
$\delta(\bullet), \Delta(\bullet)$	First variations
$\Delta\delta(\bullet)$	Second variation

Greek symbols

α_0, α_1	Angle between the nodal normal vectors and the straight line
α_i	Angle between two adjacent edges of a facet
β	Central angle in the circular arc
γ_0, γ_1	Angle between the nodal normal vectors and the normal vector of the curve
$\Gamma \equiv \partial\Omega$	Boundary in the current configuration
$\Gamma^0 \equiv \partial\Omega^0$	Boundary in the reference configuration
Γ_σ	Neumann boundary conditions
Γ_c	Potential contact surface
Γ_u	Dirichlet boundary conditions
$\overline{\Gamma}_c^{1*}$	Slip active contact zone
$\overline{\Gamma}_c^{1\bullet}$	Stick active contact zone
δ_{local}	Deviation between local and global coordinates in the interpolation
δ_r	Radial error of the interpolation
δ_n	Normal vector error of the interpolation
δ_{shape}	Shape error of the interpolation
ε	Generic penalty parameter
ε_n	Penalty parameter in the normal direction
ε_t	Penalty parameter in the tangential direction
$\overline{\varepsilon}^p$	Equivalent plastic strain
ε^r	Threshold value used in the projection algorithm
$\boldsymbol{\zeta} = (\zeta^1, \zeta^2)$	Local coordinates of the slave surface
η, ζ	Local coordinates of the Nagata patch
θ	Error in the nodal normal vector approximation
$d\lambda$	plastic multiplier

$\boldsymbol{\lambda}$	Lagrange multiplier vector
$\hat{\boldsymbol{\lambda}}$	Augmented Lagrange multiplier vector
μ	Coulomb friction coefficient
ν	Poisson's ratio
$\boldsymbol{\nu}$	Unit normal vector of the slave surface
$\boldsymbol{\xi} = (\xi^1, \xi^2)$	Local coordinates of the master surface
Π^{contact}	Contact contribution to the global functional
Π^{ext}	Contribution of the external loads to the global functional
Π^{int}	Contribution of the internal stresses to the global functional
Π^{PM}	Contact contribution of the penalty method to the global functional
ρ	Projection scaling factor
$\bar{\sigma}$	Equivalent stress
σ_0	Initial yield stress
$\boldsymbol{\sigma}$	Cauchy stress tensor
$\boldsymbol{\sigma}_{\text{b}}$	Back stress tensor
$\boldsymbol{\tau}$	Kirchhoff stress tensor
$\boldsymbol{\tau}_1^{\text{m}}, \boldsymbol{\tau}_2^{\text{m}}$	Covariant basis vectors for the master surface
$\boldsymbol{\tau}_1^{\text{s}}, \boldsymbol{\tau}_2^{\text{s}}$	Covariant basis vectors for the slave surface
$\boldsymbol{\tau}^{1\text{m}}, \boldsymbol{\tau}^{2\text{m}}$	Contravariant basis vectors for the master surface
φ	Angle between the normal vector of the Nagata curve and the vector $\mathbf{d}$
$\boldsymbol{\varphi}$	Mapping between reference and current configuration
ψ^+	Indicator function of the positive half-line
$\partial\psi^+$	Sub-differential of the indicator function ψ^+
$\psi^*_{C(p_{\text{n}})}$	Conjugate function of the disk indicator function $\psi_{C(p_{\text{n}})}$
$\partial\psi^*_{C(p_{\text{n}})}$	Sub-differential of the indicator function $\psi^*_{C(p_{\text{n}})}$
$\boldsymbol{\psi}$	Mapping between parametric space and current configuration
$\boldsymbol{\psi}_0$	Mapping between parametric space and reference configuration
Ω	Current (deformed) domain of the body
Ω^0	Reference (undeformed) domain of the body

Roman symbols

a	Cosine of the angle between the two nodal normal vectors
$\mathcal{A}^i$	Parametric domain of contact surface
$b_{\alpha\beta}$	Second fundamental form of the master surface (curvature tensor)

$\mathbf{B}$	Left Cauchy–Green tensor
$\mathcal{B}$	Generic deformable body
$\mathbf{c}$ / $\mathbf{c}'$	Curvature parameter in the original/modified Nagata interpolation
$C(p_{\mathrm{n}})$	Convex disk of radius $\mu\lvert p_{\mathrm{n}}\rvert$
$\mathbf{C}$	Right Cauchy–Green tensor
$\mathbf{C}(t)$	NURBS curve
$\mathscr{C}(\xi)$	Nagata curve
$\mathscr{C}_{\xi}$	First derivative of the Nagata curve
$\mathbf{d}$	Vector joining the end points of the edge
E	Young's modulus
$\mathbf{E}$	Green–Lagrange strain tensor
f	Yield function
$\mathbf{f}_{\mathrm{v}}$	Volume forces vector
$\mathbf{F}$	Deformation gradient
$\mathbf{F}^{\mathrm{e}}$	Elastic deformation gradient
$\mathbf{F}^{\mathrm{p}}$	Plastic deformation gradient
$\mathbf{F}_{u}^{\mathrm{s}}$	Residual vector related with structural elements
$\mathbf{F}_{u}^{\mathrm{c}}, \mathbf{F}_{\lambda}^{\mathrm{c}}$	Residual vector related with contact elements
$\mathbf{F}^{\mathrm{c}}$	Frictional contact operator
$\mathbf{F}^{\mathrm{equi}}$	Equilibrium sub-operator
$\mathbf{F}^{\mathrm{ext}}$	External forces vector
$\mathbf{F}^{\mathrm{int}}$	Internal forces vector
$\mathbf{F}^{\mathrm{proj}}$	Nonlinear system of equations for the normal projection
$\mathbf{F}^{\mathrm{suppl}}$	Supplementary sub-operator
$\mathbf{F}^{*}$	Mixed contact operator
$\mathcal{F}$	System of nonlinear equations
g_{n}	Normal gap function
$\mathbf{g}_{\mathrm{t}}$	Tangential relative slip
h	Isotropic hardening
$h+1$	Number of control points in the NURBS curve
h_{pol}	Finite element length in the poloidal direction of the torus
h_{tor}	Finite element length in the toroidal direction of the torus
$\mathbf{I}$	Second order identity tensor
$\mathbb{I}$	Time domain of interest
J	Determinant of $\mathbf{F}$
$\mathbf{J}^{*}$	Generalized Jacobian of the mixed contact operator
$\mathbf{K}^{\mathrm{s}}$	Structural tangent matrix

l	Degree of the NURBS curve
$l_{\mathrm{n}}, l_{\mathrm{t}}$	Augmented Lagrangian functionals
$\mathbf{L}, \mathbf{U}$	Lower and upper triangular matrix in LU decomposition
$\mathcal{L}$	Lagrangian functional
$\mathcal{L}^{\mathrm{a}}$	Augmented Lagrangian functional
$m+1$	Number of control points in the v direction of the NURBS surface
$m_{\alpha\beta}$	First fundamental form of the master surface (metric tensor)
$n+1$	Number of control points in the u direction of the NURBS surface
n_{f}	Number of facets surrounding a node
n_{gd}	Number of grid divisions in each parametric direction
n_{m}	Number of master nodes
n_{n}	Number of closest master nodes
n_{s}	Number of slave nodes
$\mathbf{n}$	Unit normal vector to the master surface
$\mathbf{n}_{\mathrm{analytical}}$	Unit normal vector of the analytical surface
$\mathbf{n}_{\mathrm{approx}}$	Approximated unit normal vector in the node
$\mathbf{n}_{\mathrm{Nagata}}$	Unit normal vector of the interpolated surface
$\mathbf{n}_{\mathrm{NURBS}}$	Unit normal vector of the NURBS surface
$\mathbf{n}_i^{\mathrm{facet}}$	Unit normal vector of the facet
N	Total number of increments
Ne	Number of structural finite elements
Nn	Number of nodes
N_i	Element shape functions
$N_{i,l}$	Normalized B-spline basis functions of degree l
p, q	Degree of the NURBS surface in each direction
p_{n}	Contact pressure
$\hat{p}_{\mathrm{n}}$	Regularized contact pressure at solution
$\mathrm{P}_0, \mathrm{P}_1$	End points of the Nagata curve
$\mathbf{P}$	First Piola–Kirchhoff stress tensor
$\mathbf{P}_i$	Position vector of the control points (curve)
$\mathbf{P}_{i,j}$	Position vector of the control points (surface)
$\boldsymbol{\mathscr{P}}^{\mathrm{t}}(\eta,\zeta)$	Triangular Nagata patch
$\boldsymbol{\mathscr{P}}^{\mathrm{q}}(\eta,\zeta)$	Quadrilateral Nagata patch
$\boldsymbol{\mathscr{P}}_\eta^{\mathrm{t}}, \boldsymbol{\mathscr{P}}_\zeta^{\mathrm{t}}$	First order partial derivatives of the triangular Nagata patch

$\mathscr{P}_{\eta}^{\mathrm{q}}, \mathscr{P}_{\zeta}^{\mathrm{q}}$	First order partial derivatives of the quadrilateral Nagata patch
r	Radius of the circular arc, cylinder, sphere and minor radius of torus
r_{m}	Maximum ratio between maximum and minimum finite element side length
$r_{\min}$	Strategy to adjust automatically the increment size
r_{t}	Ratio between the two finite element sides
R	Major radius of the torus
R_1	Residuum of the Taylor series
$\mathbf{R}$	Rotation tensor in polar decomposition
$\Re^2$	Two-dimensional Euclidean space
$\Re^3$	Three-dimensional Euclidean space
$\mathbf{s}$	Solution vector of the normal projection algorithm
S	Number of contact elements
$\mathbf{S}$	Second Piola–Kirchhoff stress tensor
$\mathbf{S}(u,v)$	NURBS surface
t	Generic time (instant)
$\mathbf{t}$	Cauchy stress vector, tangential slip direction
$\mathbf{t}_{\mathrm{t}}$	Frictional force
$\mathbf{T}$	Knot vector of the NURBS curve
$\mathbf{u}$	Displacement vector field
$\delta\mathbf{u}$	Virtual displacement field
$\mathbf{U}$	Right stretch tensor
$\mathbf{U},\mathbf{V}$	Knot vector of the NURBS surface in each direction
$\mathcal{U}$	Space of vectors in the 3D Euclidean space
dv	Infinitesimal volume in the current configuration
$\mathbf{v}$	Test function or virtual function vector
dV	Infinitesimal volume in the reference configuration
$\mathbf{V}$	Left stretch tensor
$\mathcal{V}$	Space of virtual displacements
w_i	Weight of the control points (curve)
$w_{i,j}$	Weight of the control points (surface)
$\mathbf{x}$	Position vector in current configuration
$d\mathbf{x}$	Infinitesimal vector in the current configuration
$\mathbf{x}^{\mathrm{s}}$	Position vector of a generic slave point/node in current configuration
$\mathbf{x}^{\mathrm{m}}$	Position vector of a generic master point/node in current configuration
$\mathbf{X}$	Position vector in reference configuration
$d\mathbf{X}$	Infinitesimal vector in the reference configuration

Acronyms

2D	Two Dimensional
3D	Three Dimensional
ALM	Augmented Lagrangian Method
CAD	Computer-Aided Design
CAE	Computer-Aided Engineering
CAM	Computer-Aided Manufacturing
DD3IMP	Deep Drawing 3D IMPlicit finite element code
DSS	Direct Sparse Solver
Euro NCAP	European New Car Assessment Programme
FEA	Finite Element Analysis
FEM	Finite Element Method
FI	Full Integration
GNM	Generalized Newton Method
HITA	HIerarchy-Territory Algorithm
IGES	Initial Graphics Exchange Specification
KKT	Karush–Kuhn–Tucker
MKL	Math Kernel Library
MWA	Mean Weighted by Angle
MWAAC	Mean Weighted by Angle and Area of Circumscribed circle
MWAAT	Mean Weighted by Areas of Adjacent Triangles
MWE	Mean Weighted Equal
MWELR	Mean Weighted by Square Root of Edge Length Reciprocals
MWSELR	Mean Weighted by Sine and Edge Length Reciprocals
NTN	Node-to-Node
NTS	Node-to-Segment
NURBS	Non-Uniform Rational Basis Spline
PM	Penalty Method
SRI	Selective Reduced Integration
STEP	STandard for the Exchange of Product model data
STS	Segment-to-Segment
URI	Uniform Reduced Integration

Chapter 1

Introduction

This chapter introduces the overall framework of frictional contact problems and their resolution with the finite element method, emphasizing the industrial interest and their broad application in several fields of mechanical engineering. The objectives of the research developed in this thesis are presented, which are defined based in the current challenges of computational contact mechanics. In order to facilitate the readability and comprehension of the thesis, the structure and the main topics addressed in each chapter are also presented.

1.1. Motivation

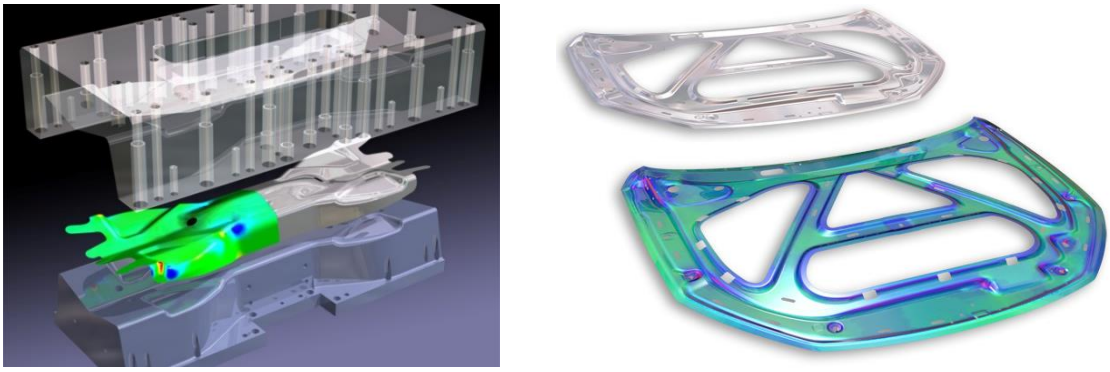
Any mechanical load results from the contact interaction between two separate bodies or parts of a single body. From the engineering point of view, the contact interactions can be intentional, such as in the mechanical systems of sheet metal forming, or undesired, as in a car crash. Indeed, virtually all structural and mechanical systems comprise the contact between deformable components. Despite the importance of contact mechanics in several engineering applications, contact effects are rarely taken into account in structural analysis due to the complexity of the contact phenomena (multi-physical nature). In fact, contact problems are inherently nonlinear because the contact surface on which the loads are transferred from one body to another is *a priori* unknown. Furthermore, the complexity increases if the friction effect is taken into account at the interface, which is very difficult to model accurately, due to its dependence on several factors (contact pressure, surface roughness, temperature, etc.). Thus, although contact always involves friction phenomena, in some situations it is neglected to simplify the analysis.

The history of contact mechanics, which began hundreds of years ago in ancient Egypt with the movement of large stone blocks, can be divided into three distinct stages [Zhong 93]. In the first stage, the contact bodies were assumed rigid, while the analysis was restricted to only total contact forces. The Newton's third law and the Coulomb's friction law can be stated as the two principal contribution in this stage. The second stage is defined by the need to know the contact pressure and frictional shear stress distribution across the contacting surfaces (local phenomena). The development of the mechanical science and increase of engineering activity has allowed to achieve this requirement. The key point of this second stage is dictated by the work carried out by Hertz [Hertz 82], establishing the analytical solution for frictionless contact between two ellipsoidal bodies. This successful treatment of a static contact problem in elasticity represents a milestone in the field of modern contact mechanics. Many others researchers (mainly mathematicians) studied contact problems using different shapes and under different circumstances, in order to overcome the restrictions of the Hertz theory, such as pure elasticity, frictionless and small deformations. Nevertheless, these approaches are obviously very restrictive and can be only applied to a few simple problems. The third stage began with the application of numerical methods to solve contact problems, where the finite element method has been the most widely used. Complicated frictional contact problems coming from industrial needs, involving complex geometries, boundary conditions and nonlinear materials, can be solved with the finite element method. The first important problem solved in the field of computational contact mechanics deals with the static equilibrium of a linearly elastic body in frictionless contact with a rigid foundation, called Signorini's problem. The existence and uniqueness of the solution of the Signorini's problem is proven in the book of Kikuchi and Oden [Kikuchi 88], where a consistent description of the variational inequality approach to contact problems is presented. The extension of this approach to the case of frictional contact problems with large deformation and, finally, to contact between deformable bodies has been developed over the last few decades. Nevertheless, the current industrial need of a fast and accurate resolution of contact problems, which may include friction, wear, adhesion, large deformations, large sliding and nonlinear material behaviour, is still a big challenge for computational contact mechanics.

1.2. Computational contact mechanics

Since many engineering problems require the description of the mechanical interaction across interfaces, the application of contact mechanics is broad, covering important industrial problems in the mechanical and civil engineering, as well as relevant applications in biomechanics, such as joint prosthesis systems. However, the closed form solution of these problems is generally unknown. Therefore, the contact conditions were often

modelled in the past using simple boundary conditions. Nowadays, with the fast development of computing power, the finite element method can be successfully used to simulate some complex problems involving contact. Nevertheless, the demand for reliable and fast numerical simulations is still a challenge for computational contact mechanics. This is related with the highly nonlinear and non-smooth behaviour of the contact problems, making its numerical treatment extremely difficult. The big importance of the frictional phenomena and the rapid improvement of modern computer technology motivates the development of efficient and robust numerical algorithms to improve the accuracy of the numerical simulations.

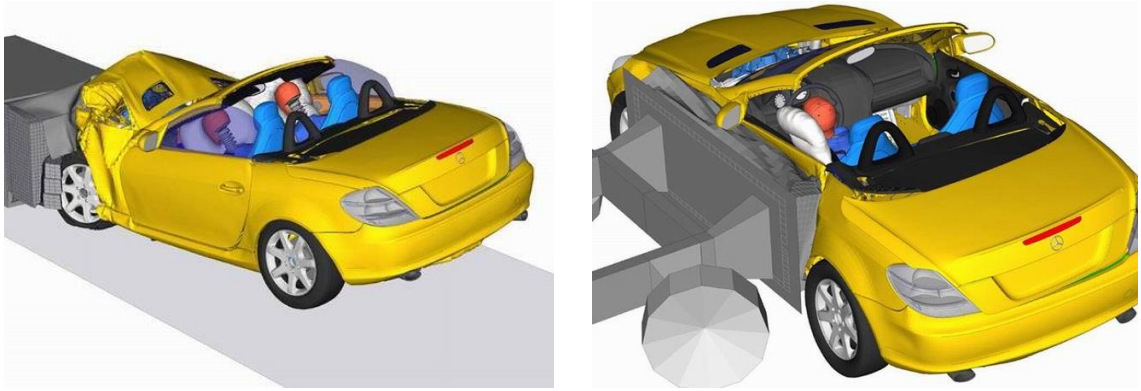


(a) (b)

Figure 1.1. Sheet metal forming: (a) double-attached structural component (Courtesy of Rockford Toolcraft); (b) engine hood inner (Courtesy of Volkswagen Autoeuropa).

The application of contact mechanics in civil engineering covers problems such as bearing of steel constructions, connection of structural members by bolts or the relative slip between concrete and reinforcing steel. Concerning its application in mechanical engineering, the main problems addressed are the design of gears and bearings, metal forming processes, development of pneumatic tires and crash analysis of cars, among others. While the numerical analysis of gears and bearings is carried out considering linear elasticity, metal forming processes (sheet metal or bulk forming) involve large deformations of the sheet/bulk and, therefore, the adoption of an inelastic constitutive model is mandatory. Nevertheless, the forming tools are typically considered rigid, which simplify the numerical simulation of the contact problem. The numerical simulation of two examples of sheet metal forming is presented in Figure 1.1, both applied in the automotive industry. Due to the increasing complexity of the components and the strong reduction of development periods in the automotive industry, traditional tool design methods based on a trial-and-error are gradually replaced by the use of numerical simulation with the finite element method [Tekkaya 00]. Currently, the AutoForm™ is the commercial finite element code (static-implicit) most commonly used in the automotive industry to simulate sheet metal forming processes, which was created from a research project at ETH Zurich. Both

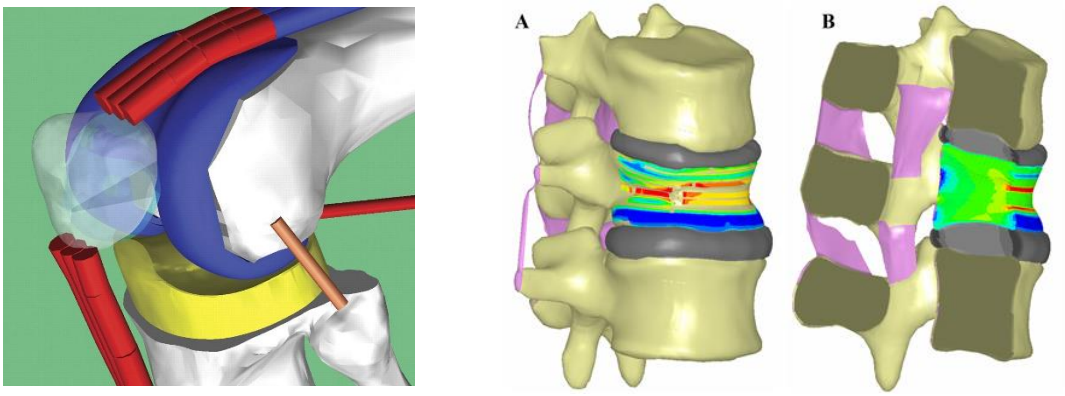
results presented in Figure 1.1 were obtained with this software. The rolling contact of car tyres is also a very challenging problem, requiring a fine discretization to accurately describe the complex tyre structure and the treads. Besides, the frictional contact between the tyres and the road surface should include other multi-physical characteristics related with the heating of the tyre (thermo-mechanical coupling) or its noise production [Wriggers 06].



(a) (b)

Figure 1.2. Crash of a car against a deformable barrier (Courtesy of Daimler Chrysler AG): (a) frontal impact; (b) lateral impact.

The analysis of cars crash have a significant industrial importance in automotive engineering, because its numerical simulation can reduce the development time and the final costs of the new cars. In fact, in order to fulfil the safety regulations and consumer protection tests, the experimental crashworthiness assessment involves significant costs in the automotive industry. Some preliminary experimental tests performed during the conception of the new cars can be replaced by numerical simulations. The crash of a car against a deformable barrier following the Euro NCAP norms is shown in Figure 1.2, analysing the difference between frontal (initial speed of 64 km/h) and lateral impact using numerical simulation results. However, car crash simulations using the finite element method are one of the most challenging and complex contact problems [Wriggers 06]. Due to the large quantity of components involved in the numerical simulation, a sufficiently refined model comprises some millions of finite elements, increasing significantly the computational cost. Moreover, the model need to account for large deformations, inelastic constitutive material models, damage models, dynamic effects and multiple contact surfaces including self-contact (see Figure 1.2). The incorporation of all these complex phenomena in the finite element model demands a deep knowledge in mechanical engineering, numerical methods and computational methods from the programming point of view. The LS-DYNA™ general purpose finite element code (dynamic-explicit) is often adopted for crash simulations, which was used to obtain the results presented in Figure 1.2.



(a) (b)

Figure 1.3. Biomechanics applications for human joints: (a) knee implants [Morra 08]; (b) contact between lumbar vertebral bodies (Courtesy of Brigham Young University).

Other important applications of contact mechanics are related to biomechanics and biomedical engineering. The design of prosthetics for human joints, such as knee implants, tibia-femoral joins or facet joints for lumbar vertebral bodies are problems with a large impact in the health and well-being of the society, particularly in the patient's life. Figure 1.3 presents the application of computational contact mechanics to solve this type of problems, using the finite element method. The accurate numerical simulation of this particular type of contact problems requires the incorporation of complicated nonlinear material models in the analysis, taking into account the large sliding between the contacting bodies and large deformation. The validation of the numerical models with experimental results is much more difficult in this field than in the previously presented industrial examples. Nevertheless, the importance of this field in people life motivates the present intense research in this topic. Currently, the LifeMOD™ software is used by orthopaedic surgeons to plan specific surgeries in a computer-based environment, leading to improved patient outcomes that replace knees, hips and other joints [Morra 08]. This dedicated software is a validated musculoskeletal modelling system developed by LifeModeler (California, USA).

As mentioned above, several important engineering applications comprise contact and friction. In addition to the frictional contact phenomena at the interface, the behaviour of each body must be properly taken into account to simulate advanced industrial applications. In the context of material behaviour, the main difficulties arise from large inelastic deformations, which are necessarily induced in metal forming simulations. When the mechanical response and the thermal conduction interacts in the contact area, the thermo-mechanical coupling need to be considered within contact analysis, such as in the hot stamping process. Hence, the computational contact mechanics is an interdisciplinary area.

The weak form of the nonlinear boundary value problem is the basis on which the finite element method includes inequality constraints due to the impenetrability and friction conditions. Thus, the development of a variational principle leads to a variational inequality instead of a classic variational equality [Kikuchi 88]. Nevertheless, the assumption of *a priori* known contact surface allows to replace the variational inequality by a variational equality with an additional contact term, which depends on the method chosen to enforce the contact constraints. The most known and widely used methods in contact mechanics are the penalty, the Lagrange multiplier and the augmented Lagrangian, which are typically implemented in modern commercial finite element codes. The penalty method is simple to implement and presents an intuitive physical meaning. However, the contact conditions are only exactly fulfilled using an infinite value for the penalty parameter. The Lagrange multiplier method satisfies the contact conditions exactly by introducing extra degrees of freedom, called Lagrange multipliers, leading to an unconstrained saddle point problem. The augmented Lagrangian method takes advantage of these two cited methods, allowing to fulfil the contact constraints with a finite value of the penalty parameter. Besides, its application with the Uzawa's algorithm (nested update of dual variables) leads to a problem where the unknowns are only the displacements. The complete list of regularization methods adopted in the numerical simulation of contact problems can be found in [Wriggers 06] and [Laursen 02].

The treatment of contact problems in the framework of the finite element method and implicit time integration comprises the following steps: (i) contact detection; (ii) creation of contact elements and its incorporation in the general nonlinear problem, through the residual vectors and tangent matrices and (iii) resolution of the resulting problem [Yastrebov 13]. Although the contact detection is the step preceding all others, it is strongly connected with the discretization of the contact interface and, consequently, with the type of contact elements used. The contact discretization defines the structure of the contact elements transferring efforts from one contacting surface to another. The most commonly used discretization procedures are the Node-to-Node, the Node-to-Segment and the Segment-to-Segment. The simplest is the Node-to-Node discretization, which does not allow any finite sliding and introduces restrictions on mesh generation (conforming meshes). On the other hand, the Node-to-Segment discretization is valid for non-conforming meshes, large deformation and large sliding. Recently the Segment-to-Segment discretization has been efficiently coupled with the mortar method for nonconforming meshes, leading to a consistent formulation of the frictional contact problem for large sliding and large deformations.

The Node-to-Segment discretization is adopted in the present work, which is associated with the master-slave approach. This means that the impenetrability and friction conditions are enforced in the nodes of the contact surface assigned as slave, preventing the slave nodes from penetrating on the contact master surface. The detection step determines the

contacting pairs on the discretized surfaces, i.e. finds for each slave node the closest point on the master surface. The increasing demand for solutions of large scale contact problems entailed the development of efficient contact detection techniques, since this can be a bottleneck in the computation performance of the resolution methods for contact problems. Although contact detection is a purely algorithmic task, the determination of the closest point on the discretized master surface may fail when the master surface is not smooth, which occurs when the contact surface is discretized with low order finite elements. The smoothing of the discretized contact surface improves the accuracy and stability of the numerical procedure. However, "research work for a better and simpler way to represent smooth surfaces is still needed" [Wriggers 06].

As mentioned above, in the resolution step, the inequality constraints resulting from the frictional contact are introduced in the objective energy functional as additional terms by means of the penalty, Lagrange multiplier or augmented Lagrangian method. The set of unknowns in the final problem are determined by the resolution method. The penalty method leads to an unconstrained problem, where the nodal displacement are the only unknown variables. On the other hand, the augmented Lagrangian method and the Lagrange multiplier method lead to a saddle point problem. The objective function is minimized changing primal variables (displacement) and maximized through the dual variables (Lagrange multipliers representing the contact forces) [Pietrzak 97]. Note that the positivity of the Lagrange multipliers has to be satisfied in case of the Lagrange multiplier method. The augmented Lagrangian approach converges to the exact solution for a finite value of the penalty coefficient, allowing to reduce the optimization problem with inequality constraints to a fully unconstrained problem. In both methods, the nonlinear system of equations resulting from the application of the finite element method includes nodal displacements and nodal contact forces. Thus, the coupled augmented Lagrangian method leads to a mixed system of equations with simultaneous update of primal and dual variables [Alart 91]. The generalized Newton method, which allows to deal with non-smooth functions, is the numerical scheme usually adopted to solve these problems in computational contact mechanics.

1.3. Aims and objectives of the work

Despite the progresses made in computational contact mechanics during the last years, the accurate and efficient numerical simulation of contact problems is still extremely difficult. Contact problems are characterized by geometric and material discontinuity at the interface, in contrast with the usual continuity assumptions in classical solid mechanics. Therefore, these problems are inherently highly nonlinear and non-smooth, which are the main difficulties arising in its numerical solution, together with the heavy computational

costs associated with contact detection. In order to improve the accuracy, robustness and performance of the numerical simulations, the main objective of this thesis is the development and implementation of computational methods for the 3D treatment of contact problems with and without friction. The contact interaction is considered in the context of *quasi*-static nonlinear solid mechanics applications, solved with the finite element method adopting an implicit time integration scheme. Moreover, the large sliding between the curved contacting surfaces is taken into account, while the friction effect is governed by the Coulomb friction law. The integration of the contact inequality constraints (impenetrability and friction conditions) into the initial mechanical problem is performed with the augmented Lagrangian method, which enables the exact representation of contact constraints.

Since the algorithms and methods developed in this thesis are aiming at very general multipurpose applications, all types of contact problems are addressed: (i) contact between deformable and rigid bodies; (ii) contact between deformable bodies and (iii) self-contact. The material behaviour of the deformable bodies is assumed nonlinear and described by elastoplastic constitutive laws. The main difficulties arising in the solution of large sliding contact problems, based on a Node-to-Segment contact discretization, are motivated by the traditional piecewise bilinear description of the contact surfaces. The discontinuity of the surface normal vector between segments leads to non-physical oscillations in the contact force and severe convergence problems, particularly for strongly curved master surfaces. In order to overcome the above mentioned problems, the purpose of the present work is to develop a general 3D surface smoothing method using the Nagata patch interpolation. Its applicability to unstructured surface finite element meshes (local support) is the main feature of the proposed smoothing method over the existing ones. The contact detection algorithm, required for determining the contacting pairs on the discretized surfaces, is divided in two phases: global and local search. The developed algorithm must be simultaneously accurate and efficient in order to quickly identify all potential contact zones. Therefore, a global contact search algorithm is specifically developed for each type of contact problem, in order to take advantage of its particular features, i.e. contact with rigid surfaces, contact between deformable bodies and self-contact.

All methods and algorithms developed in this thesis are implemented in the in-house finite element code DD3IMP, which has been specifically developed to simulate sheet metal forming processes. Consequently, another important objective of this thesis is the extension of the finite element code capability to deal with frictional contact problems between deformable bodies, including self-contact phenomena. These developments enlarges the range of application of DD3IMP finite element code, allowing facing new challenges of computational contact mechanics.

1.4. Outline of the thesis

The contents of the thesis is organized into seven chapters, including three appendices. In order to facilitate its readability and comprehension, this section presents a brief summary of the topics covered in each chapter and appendix.

Chapter 2 presents the basic concepts and the governing equations of continuum solid mechanics, followed by the formulation of the frictional contact problems in the continuum setting. The constraints related with the impenetrability and friction conditions (Coulomb law) are stated. The resulting variational inequality is converted in a variational equality using the classical resolution methods, namely penalty, Lagrange multiplier and augmented Lagrangian. The advantages and drawbacks of each one are illustrated using a simple example.

Chapter 3 is devoted to the numerical aspects of frictional contact mechanics, introducing the principles of the finite element method and a brief description of the main features of the finite element code adopted in this study. The discretization of the contact interface with the Node-to-Segment method is addressed in this chapter, using both the single-pass and the two-pass strategies. The proposed contact search algorithm, responsible for the location of the contact zones, is decomposed in two steps: global local search. Some details of its implementation are discussed, in particular the global contact search algorithm. The weakness of the piecewise bilinear surfaces representation combined with the local search procedure based in the normal projection is highlighted using a simple example.

Chapter 4 presents the proposed 3D contact surface smoothing method based in the Nagata patch interpolation. The interpolation method is reviewed and its shortcomings in the surface smoothing application are shown, accompanied by proposed improvements. The accuracy of the surface smoothing method is assessed by means of the shape and normal vector error distributions, which is evaluated in simple geometries defined by analytical functions. The proposed algorithm to evaluate the nodal normal vectors using the information contained in the IGES file is presented, where the rigid surface geometry is represented by trimmed NURBS surfaces. In case of deformable contact surfaces, the nodal normal vectors are approximated using the weighted average of the normal vectors of facets adjacent to the node. Six different weighting factors are presented and compared between them in terms of accuracy. The influence of the nodal normal vectors on the interpolation accuracy is briefly discussed.

Chapter 5 provides the formulation of the contact elements, developed in the framework of the Node-to-Segment contact discretization using the Nagata patches in the description of the master surface, called Node-to-Nagata contact elements. The residual vectors and tangent matrices of the developed contact elements are derived for the augmented Lagrangian method and presented for all contact status. The particular case of

contact between a deformable body and a rigid obstacle is firstly addressed, leading to a mixed system of equations with nodal displacements and contact forces as unknowns. The nonzero pattern of the resulting global tangent matrix is briefly discussed. The reduced system of equations is derived by eliminating the Lagrange multipliers, obtaining a problem with nodal displacements as the only unknowns. The connection between the slave node and the master ones in case of contact between deformable bodies is shown in the definition of the contact elements. Thus, the nonzero pattern of the global tangent matrix is dictated by the sliding of the slave nodes over the master patches, motivating the creation of the presented multi-face contact elements.

Chapter 6 contains the numerical examples of some frictional contact problems, which range from simple examples with known analytical solution, to complex problems with relevant industrial interest. The incorporation of the elastic deformation of the forming tools into the numerical model is shown through the sheet metal forming of a cylindrical cup, comparing numerical and experimental results. The accuracy, robustness and performance of the proposed 3D contact surface smoothing method is validated by means of its comparison with the traditional piecewise bilinear finite element mesh representation.

Chapter 7 summarizes the main issues addressed in the thesis, emphasising the main contributions brought with the present work. Some recommendations for future work in the computational contact mechanics field are suggested, taking advantage of the developed software, as well as the 3D contact surface smoothing method.

Appendix A gives the list of publications performed in the scope of the thesis, summarising the results of the scientific research carried out in the framework of this dissertation. Although some parts of these publications are contained in this thesis, the papers are not referenced in the text.

Appendix B provides a summary of the Initial Graphics Exchange Specification (IGES) file format, which is widely used to transfer information between CAD and CAE software packages. The file organisation is described, focusing in the information required for the mathematical definition of the surface model geometry.

Appendix C introduces the basic equations necessary to evaluate the projection of a generic point onto a NURBS surface. The solution of the resulting nonlinear system of equations is obtained with the Newton–Raphson method. The closed form expressions necessary to evaluate the first- and second-order partial derivatives, in a generic point of a NURBS surface, are also presented.

Chapter 2

Formulation of Contact Problems and Resolution Methods

The purpose of this chapter is to introduce the general concepts and governing equations for solving frictional contact problems between deformable bodies. The principles of nonlinear solid mechanics problems are reviewed, given particular attention to large displacements. The kinematical relations, the constitutive models and the static equilibrium equations in the variational form are described. The second part of the chapter contains the formulation of the frictional contact problem, presented in the framework of continuum mechanics, following the scheme previously used in solid mechanics. The contact kinematic and static variables, the formulation of frictional contact conditions (impenetrability and frictional constraints), as well as the governing variational principle are specified. This section is followed by a brief description of the procedures usually applied to incorporate the contact and friction constraints into the variational principle, providing a variational equality for a known contact area. The penalty method, the Lagrange multipliers method and the augmented Lagrangian method are presented, which are the most important and widely used methods for the enforcement of contact constraints. A simple contact problem is solved using all these methods, in order to highlight the advantages and drawbacks of each of them.

2.1. Continuum solid mechanics

This section contains the general concepts and principles of the continuum solid mechanics given particular attention to large displacement of solids. Both thermal and inertial effects are neglected in the presented formulation. The kinematical relations, the

constitutive models and the equilibrium equations in the variational form are briefly described. The underlying theory of continuum mechanics is not treated in depth detail, which is strictly limited to the requirements associated with the finite element formulation of the contact problem between deformable bodies. Further information concerning the continuum solid mechanics can be found in following monographs and textbooks [Gurtin 81], [Simo 98], [Holzapfel 00], [Lemaitre 90], [Belytschko 00], [Malvern 69], [Bonet 97], among other.

2.1.1. Kinematics of deformation

The motion and deformation of a body are described by means of kinematic quantities, which are required within the constitutive models and the weak formulation of balance laws. In the continuum framework, a deformable body $\mathcal{B}$ is described in a formal way by a set of continuously distributed points, which occupy a region in the three-dimensional Euclidean space $\Re^3$. The set of points defining the body $\mathcal{B}$ is denoted by the open set $\Omega \subset \Re^3$ and its boundary $\partial\Omega$ is subdivided into two non-overlapping regions Γ_u and Γ_σ, where Dirichlet (prescribed motions) and Neumann (prescribed tractions) boundary conditions are specified, where the following conditions satisfied:

$$\begin{aligned} \Gamma_u \cup \Gamma_\sigma &= \partial\Omega, \\ \Gamma_u \cap \Gamma_\sigma &= \varnothing. \end{aligned} \tag{2.1}$$

In mechanical problems comprising large deformations, it is necessary to distinguish the reference configuration (denoted as Ω^0) from the current configuration at a given time $t \in \mathbb{I}$, where $\mathbb{I} = [0, T]$ denotes the time domain of interest. The current (deformed) configuration is obtained applying a configuration mapping $\boldsymbol{\varphi}$ to Ω^0 (see Figure 2.1). As before, the boundary $\partial\Omega$ is decomposed into two non-overlapping subdomains, where motions and tractions are prescribed, which are denoted by Γ_u^0 and Γ_σ^0, respectively. The points of the body in the reference configuration Ω^0 are denoted by the position vector $\mathbf{X}$, while in the current configuration at time t are denoted as $\mathbf{x}(\mathbf{X}, t)$, such that $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{X}, t)$, as shown in Figure 2.1. This equation defines a curve in $\Re^3$, which describes the path of a point during the time domain of interest, $t \in \mathbb{I}$.

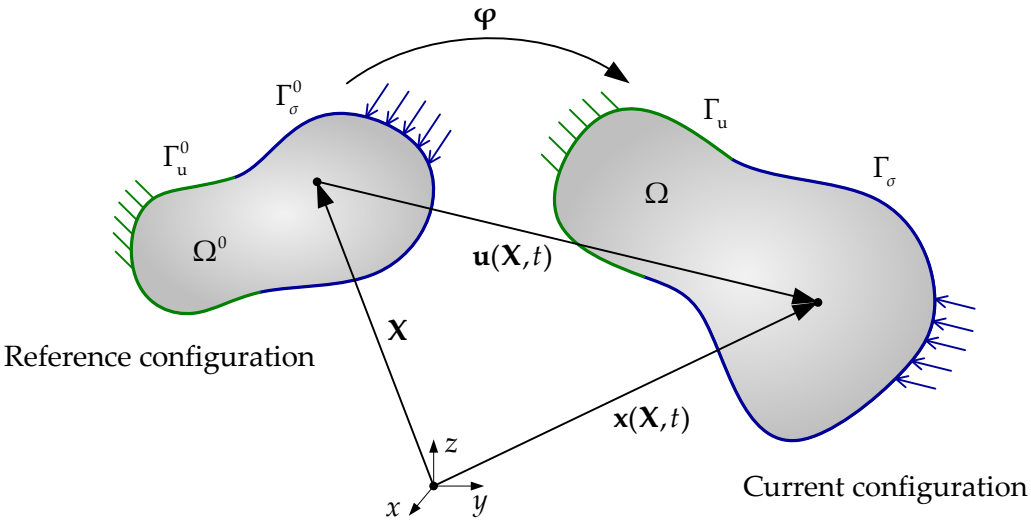


Figure 2.1. Reference and current configuration of a body $\mathcal{B}$ undergoing large deformations.

The Lagrangian description of the problem is the most frequent choice in solid mechanics. Thus, the body motion can be described relatively to a fixed spatial frame defined by orthonormal Cartesian basis vectors, see Figure 2.1. The displacement vector $\mathbf{u}(\mathbf{X},t)$ of a material point is defined as the difference between its position vector in the current and reference configuration, which is given by:

$$\mathbf{u}(\mathbf{X},t)=\boldsymbol{\varphi}(\mathbf{X},t)-\mathbf{X}=\mathbf{x}(\mathbf{X},t)-\mathbf{X}. \tag{2.2}$$

In order to describe the deformation process locally, the deformation of an infinitesimal material fibre $d\mathbf{X}$ from the reference to the current configuration is given by:

$$d\mathbf{x}=\mathbf{F}d\mathbf{X}, \tag{2.3}$$

where $\mathbf{F}$ is deformation gradient, which is the fundamental quantity used to measure strain in large deformations, being defined using the displacement field as:

$$\mathbf{F}=\frac{\partial\boldsymbol{\varphi}(\mathbf{X},t)}{\partial\mathbf{X}}=\frac{\partial\mathbf{x}(\mathbf{X},t)}{\partial\mathbf{X}}=\frac{\partial\mathbf{u}(\mathbf{X},t)}{\partial\mathbf{X}}+\frac{\partial\mathbf{X}}{\partial\mathbf{X}}=\frac{\partial\mathbf{u}(\mathbf{X},t)}{\partial\mathbf{X}}+\mathbf{I}, \tag{2.4}$$

where $\mathbf{I}$ is the second order identity tensor. The deformation gradient is used to quantify local volume change for a point $\mathbf{X}$ in Ω^0. Defining a reference volume dV at a point $\mathbf{X}$ and denoting this volume in the current configuration by dv, then these two volumes are related by:

$$dv=JdV, \tag{2.5}$$

where

$$J = \det \mathbf{F}. \tag{2.6}$$

The quantity J is the determinant of the deformation gradient (also called Jacobian), which gives the volume change and should be positive $J > 0$ to exclude self-penetration of points of the body. Furthermore, this condition ensure that the inverse of deformation gradient $\mathbf{F}$ defined in (2.4) exists and is unique.

According to the polar decomposition theorem [Malvern 69], the deformation gradient $\mathbf{F}$ defined in (2.4) can be decomposed into a stretching and a rotational part:

$$\mathbf{F} = \mathbf{RU} = \mathbf{VR}, \tag{2.7}$$

where $\mathbf{R}$ represents a rotation tensor $(\mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}} \text{ and } \det \mathbf{R} = 1)$, while $\mathbf{U}$ is the right symmetric stretch tensor and $\mathbf{V}$ is the left symmetric stretch tensor. The decompositions in (2.7) are called right and left polar decompositions of $\mathbf{F}$, respectively. Figure 2.2 illustrates the polar decomposition theorem and the interpretation of $\mathbf{R}$ and $\mathbf{U}/\mathbf{V}$ as a rigid body rotation and a stretch, respectively. In fact, the polar decomposition is crucial within the definition of the strain measures to distinguish pure stretches from rigid body rotations, since the last ones do not contribute for changes of the body geometry.

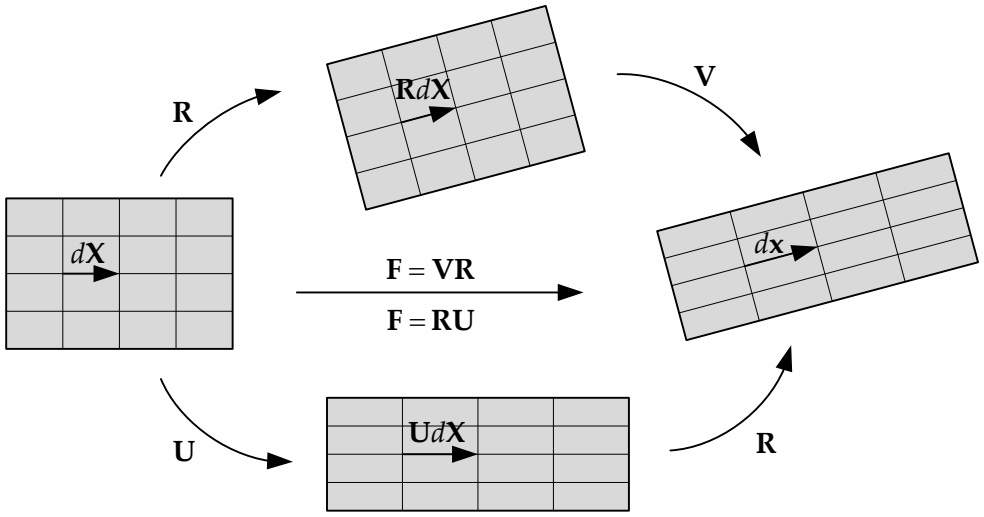


Figure 2.2. Schematic representation of the polar decomposition theorem.

Since $\mathbf{F}$ is not objective (depending on a rigid body motion), it is convenient also to define the right and left Cauchy–Green strain tensors $\mathbf{C}$ and $\mathbf{B}$, respectively, which are given by:

$$\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = \mathbf{U}^2 \quad \text{and} \quad \mathbf{B} = \mathbf{F}\mathbf{F}^{\mathrm{T}} = \mathbf{V}^2. \tag{2.8}$$

The right Cauchy–Green tensor $\mathbf{C}$ refers to the reference configuration Ω^0, while the left Cauchy–Green tensor is related to the current configuration. Since the strain measure defined by the right Cauchy–Green tensor is not zero at the initial state $(\mathbf{F}=\mathbf{I}\Rightarrow\mathbf{C}=\mathbf{I})$, it is convenient to introduce the Green–Lagrange strain tensor, defined as:

$$\mathbf{E}=\frac{1}{2}(\mathbf{C}-\mathbf{I})=\frac{1}{2}(\mathbf{F}^{\mathrm{T}}\mathbf{F}-\mathbf{I}), \tag{2.9}$$

which is the most important strain measure, regarding the Lagrangian approach [Belytschko 00].

2.1.2. Stress measures

Several stress measures can be used in the context of large deformations, being its definition dependent on the configurations where the loads are applied. In the current configuration, the stress state is described by the Cauchy stress tensor $\boldsymbol{\sigma}$, where the tensor components represent the external forces acting on the current configuration per unit of the current area at a given point $\mathbf{x}$. Considering a given plane passing through the point, the traction (force per unit area) acting on this plane is given by:

$$\mathbf{t}=\boldsymbol{\sigma}\mathbf{n}, \tag{2.10}$$

where $\mathbf{n}$ is the unit normal vector to the plane under analysis and $\mathbf{t}$ is called the Cauchy stress vector. The second stress measure is the Kirchhoff stress tensor, which is related to the Cauchy stress tensor via:

$$\boldsymbol{\tau}=J\boldsymbol{\sigma}, \tag{2.11}$$

where J is the Jacobian of the transformation defined in (2.6). Due to the symmetry of the Cauchy stress tensor $(\boldsymbol{\sigma}=\boldsymbol{\sigma}^{\mathrm{T}})$, Kirchhoff stress tensor is also symmetric. Another important stress measures are the first and second Piola–Kirchhoff stress tensors, $\mathbf{P}$ and $\mathbf{S}$, respectively given by:

$$\mathbf{P}=J\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}, \tag{2.12}$$

$$\mathbf{S}=J\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{T}}. \tag{2.13}$$

where $\mathbf{F}$ is the deformation gradient defined in (2.4). While in the current configuration the stress state is described by the Cauchy stress tensor $\boldsymbol{\sigma}$, the first Piola–Kirchhoff stress tensor $\mathbf{P}$ defines the area vector in the reference configuration to the corresponding force vector in the current configuration. Moreover, it should be noted that $\mathbf{P}$ is not a symmetric tensor in general. The second Piola–Kirchhoff stress tensor uses both the area and the force

vectors measured in the reference configuration, leading a tensor always symmetric. Nevertheless, the second Piola–Kirchhoff stress tensor $\mathbf{S}$ does not represent a stress that can be interpreted physically in terms of surface tractions as the Cauchy stress.

2.1.3. Constitutive laws

Since different materials behave differently under the same external load, the constitutive law characterizes the material response of a specific body. In order to describe the material behaviour it is necessary to establish a relationship between the strains and stresses, being usual practise to consider the stresses as dependent variables [Bertram 08]. Indeed, the theory of constitutive laws is rather complicated, thus the constitutive models are usually classified in two groups: phenomenological models based on experimental observations, often at macroscopic scales [Belytschko 00], and atomistic models motivated by atomic interactions [Mishin 10], [Sauer 07]. Due to their widely practical application and numerical simplicity, only phenomenological models are addressed in this thesis, given particular attention on the elastoplastic constitutive laws for large strains. Many other constitutive laws exist for miscellaneous applications. However, the focus of this thesis is on contact interaction rather than constitutive modelling. Indeed, constitutive laws for the bodies coming into contact can be arbitrary, not affecting the main formulation of contact problems [Wriggers 06].

In many engineering applications involving small strains and rotations, the response of the material may be considered to be linearly elastic (perfectly reversible and path-independent). The corresponding constitutive equation is often referred to as the generalized Hooke's law, which can incorporate fully anisotropic material response. The extension of the small strain linear elasticity to the case of finite strain can be carried out in different ways and different constitutive relations can be established. The constitutive models for large strain elasticity are typically categorized in two classes: hypoelasticity and hyperelasticity. The hypoelastic material laws are formulated in rate form, which relate the rate of stress to the rate of deformation. In order to satisfy the principle of material frame indifference, the stress rate should be objective and should be related to an objective measure of the deformation rate. Due to the properties of the rate of deformation tensor, a hypoelastic material do not strictly reflects the path independence of elasticity. However, if the elastic strains are small, the behaviour is close enough to path independent. Thus, the principal use of hypoelastic constitutive relations is in the representation of the elastic response of elastoplastic constitutive relations, where the elastic deformations are small [Khan 95]. On the other hand, hyperelastic or Green materials are characterized by the existence of a stored energy function. Thus, the work done by a hyperelastic material is independent of the deformation path [Ogden 84]. This description is valid for many

materials, such as rubbers, foams and biological tissues, which exhibit a nonlinear stress-strain behaviour undergoing finite deformations and an almost incompressible response.

2.1.3.1. Elastoplastic behaviour

Many materials widely used in mechanical applications present a nonlinear behaviour. A large class of nonlinear materials can be described by the assumption of elastoplastic behaviour, referring examples of these materials, metallic ones, such as steels and aluminium alloys. The fundamental idea behind the phenomenological approach to elastoplasticity is that the deformation gradient $\mathbf{F}$ is multiplicatively decomposed in two components:

$$\mathbf{F} = \mathbf{F}^{e}\mathbf{F}^{p}, \tag{2.14}$$

where $\mathbf{F}^{e}$ and $\mathbf{F}^{p}$ denotes the elastic and plastic deformation gradients, respectively [Lee 69], [Simo 88].

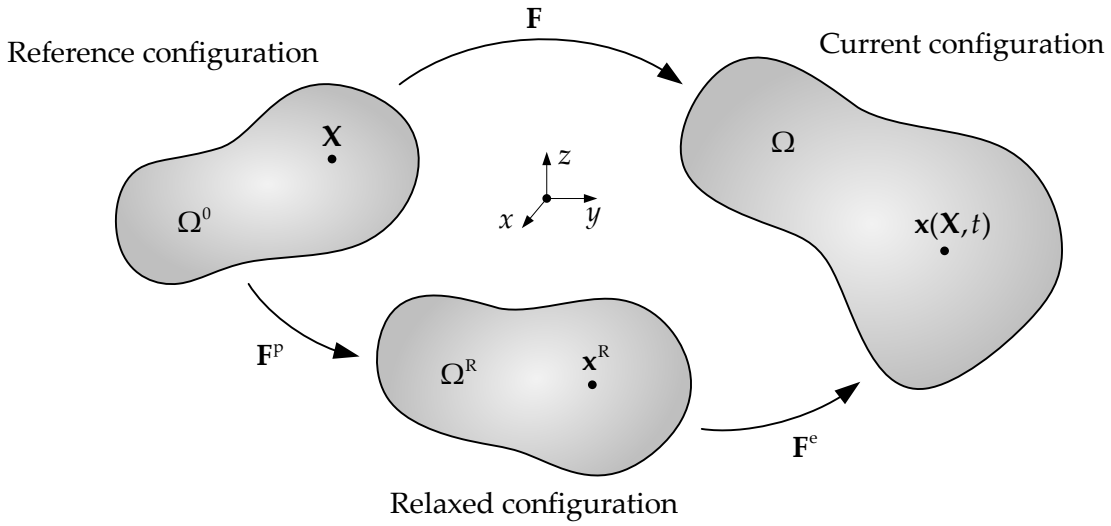


Figure 2.3. Schematic representation of the multiplicative decomposition theorem.

The multiplicative decomposition assumes the existence of an intermediate configuration (stress free state), as shown in Figure 2.3. This relaxed configuration is obtained considering that the deformed body has undergone a purely plastic deformation, thus nonconforming configuration with respect to neighbouring material points. Since an arbitrary rigid body rotation can be selected for the intermediate configuration, it is in general not uniquely defined. In order to overcome the uniqueness problem, by convention, all the rigid body rotation is lumped into the plastic deformation gradient $\mathbf{F}^{p}$, such that the elastic deformation gradient $\mathbf{F}^{e}$ includes stretch only [Dunne 05]. Since the elastoplastic behaviour is a path-dependent deformation process, it is commonly analysed using an

incremental procedure. Thus, variables such as displacements, plastic strains, and external loads are written in an incremental form. This requires the use of rate-type measures of deformation, which are derived from the multiplicative decomposition of the deformation gradient (2.14).

The elastoplastic material behaviour is defined by three properties: (i) a yield function; (ii) a flow rule and (iii) a hardening law. The yield function establishes the state of multi-axial stress corresponding to the start of plastic flow, defined in the general form by:

$$f(\boldsymbol{\sigma},\boldsymbol{\sigma}_{\mathrm{b}},\bar{\varepsilon}^{\mathrm{p}}) = F(\boldsymbol{\sigma}-\boldsymbol{\sigma}_{\mathrm{b}}) - \sigma_{\mathrm{Y}}(\bar{\varepsilon}^{\mathrm{p}}), \tag{2.15}$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor and $\boldsymbol{\sigma}_{\mathrm{b}}$ is called back stress tensor, which is related to kinematic hardening. The evolution of the yield stress during plastic flow is defined by the hardening law σ_{Y}, which depends on the cumulative equivalent plastic strain $\bar{\varepsilon}^{\mathrm{p}}$. Mixed hardening models combine isotropic and kinematic hardening, which results into an evolution of the yield surface by simultaneous translation and expansion. A points is in the elastic domain if the stress state corresponds to $f<0$. When that point is located on the boundary of the yield surface $(f=0)$ plastic deformation can occur depending on the loading condition, whereas $f>0$ is inadmissible. The irreversibility of the plastic flow process is expressed by a flow rule. In most metals, an associated flow rule can be considered, where the increment of plastic strain occurs in the normal direction to the yield surface at the load point. This direction is given by the gradient of the yield function (2.15) with respect to the stress tensor:

$$d\dot{\boldsymbol{\varepsilon}}^{\mathrm{p}} = d\lambda \frac{\partial f(\boldsymbol{\sigma},\boldsymbol{\sigma}_{\mathrm{b}})}{\partial \boldsymbol{\sigma}}, \tag{2.16}$$

where $d\lambda$ is a scalar which determines the size of the plastic strain increment. The model is completed by specifying the loading/unloading conditions for elastoplasticity. These may be given in terms of the classical Karush–Kuhn–Tucker (KKT) conditions for inequality constraints in optimization:

$$f \le 0, \quad d\lambda \ge 0, \quad d\lambda f = 0, \tag{2.17}$$

which ensure that the stress point is not outside the yield surface, that the magnitude of the plastic strain rate is always positive, and that plasticity only occurs when the stress point is on the yield surface [Laursen 02]. The integration of the constitutive law at each material point allows to determine the increments of stress state and the equivalent plastic strain. For further details readers are addressed to see [Alves 03].

2.1.4. Equilibrium equations and variational principle

The problems addressed in this thesis are restricted to *quasi*-static conditions, which is an appropriate approximation when the inertial forces are negligible when compared to the internal and applied forces. The equilibrium equations in the strong form for the finite deformation boundary value problem, which describe the balance of linear and angular momentum at any given time $t \in \mathbb{I}$, are given by:

$$\begin{cases} \operatorname{div}(\boldsymbol{\sigma}) + \mathbf{f}_v = \mathbf{0} \quad \text{in } \Omega, \\ \mathbf{t} = \bar{\mathbf{t}} \quad \text{at } \Gamma_\sigma, \\ \mathbf{u} = \bar{\mathbf{u}} \quad \text{at } \Gamma_u, \end{cases} \tag{2.18}$$

where $\boldsymbol{\sigma}$ is the Cauchy stress tensor and $\mathbf{f}_v$ is a vector of volume forces. The notation $\bar{\mathbf{t}}$ represents a prescribed Cauchy traction (Neumann boundary conditions) and $\bar{\mathbf{u}}$ denotes a prescribed displacement (Dirichlet boundary conditions), as shown schematically in Figure 2.1. The divergence operator in (2.18) is to be interpreted as being with respect to the current coordinates $\mathbf{x}$, since the equilibrium equations are formulated in the current configuration of the body $\mathbf{x} \in \Omega$.

The weak formulation (also called variational form) of the momentum balance is indispensable to approximate the strong form by finite elements procedures. If the equilibrium equation $(2.18)_1$ is satisfied in each point of the volume Ω, then the dot product of this equation with any arbitrary vector function $\mathbf{v}$ (also called test function or virtual function) integrated over the body volume is given by:

$$\int_\Omega (\operatorname{div}(\boldsymbol{\sigma}) + \mathbf{f}_v) \cdot \mathbf{v} \, d\Omega = 0, \tag{2.19}$$

where the first term can be integrated by parts if the test function is sufficiently smooth, $\mathbf{v} \in C^1$. Thus, using the divergence theorem, we obtain:

$$\int_\Omega \boldsymbol{\sigma} : \nabla \mathbf{v} \, d\Omega - \int_\Omega \mathbf{f}_v \cdot \mathbf{v} \, d\Omega - \int_\Gamma (\boldsymbol{\sigma}\mathbf{n}) \cdot \mathbf{v} \, d\Gamma = 0. \tag{2.20}$$

where $\mathbf{n}$ is an outward unit normal vector at the surface body Γ. The stress vector $\boldsymbol{\sigma}\mathbf{n}$ is non-null on the surface where the stress vector is prescribed (Γ_σ) and on the surface where the displacement is prescribed (Γ_u). Replacing the test functions by arbitrary test displacements (also called virtual displacements) $\delta\mathbf{u}$, the weak form in (2.20) results in the balance of virtual work formulated in the current configuration:

$$\int_\Omega \boldsymbol{\sigma} : \delta\nabla \mathbf{u} \, d\Omega - \int_\Omega \mathbf{f}_v \cdot \delta\mathbf{u} \, d\Omega - \int_{\Gamma_\sigma} \bar{\mathbf{t}} \cdot \delta\mathbf{u} \, d\Gamma = 0. \tag{2.21}$$

where the first term represents the internal virtual work contribution, while the second and third terms together denote the virtual work of external forces, which is divided into two categories: volume forces $\mathbf{f}_v$ and surface forces $\bar{\mathbf{t}}$. The equivalence between the strong and the weak formulations establishes that the solution of the finite deformation boundary value problem defined in the strong form (2.18) also satisfy the weak formulation presented in (2.21). Furthermore, the weak formulation poses weaker differentiability requirements to the solution functions $\mathbf{u}$ than the strong formulation. Thus, the following solution and weighting spaces can be defined by:

$$\begin{aligned} \mathcal{U} &= \left\{ \mathbf{u} \in H^1(\Omega) \,\middle|\, \mathbf{u} = \bar{\mathbf{u}} \quad \text{on } \Gamma_u \right\}, \\ \mathcal{V} &= \left\{ \delta\mathbf{u} \in H^1(\Omega) \,\middle|\, \delta\mathbf{u} = 0 \quad \text{on } \Gamma_u \right\}, \end{aligned} \tag{2.22}$$

where $H^1(\Omega)$ denotes the Sobolev space of functions with square integrable values and first derivatives. The solution space $\mathcal{U}$ may in general depend on the time due to a possible time dependency of the Dirichlet boundary conditions, while the weighting space $\mathcal{V}$ does not depend on the time [Popp 09].

The equivalent expression for the principle virtual work in the reference configuration is given by:

$$\int_{\Omega^0} \mathbf{P} : \delta\nabla\mathbf{u} d\Omega^0 - \int_{\Omega^0} \bar{\mathbf{f}}_v \cdot \delta\mathbf{u} d\Omega^0 - \int_{\Gamma_\sigma^0} \bar{\mathbf{t}}_0 \cdot \delta\mathbf{u} d\Gamma^0 = 0, \tag{2.23}$$

where $\mathbf{P}$ is the first Piola–Kirchhoff stress field, $\bar{\mathbf{f}}_v = J\mathbf{f}_v$ is the reference volume body force and $\bar{\mathbf{t}}_0$ is a prescribed boundary traction in reference area.

2.2. Continuum contact mechanics

In standard nonlinear solid mechanics, as considered in Section 2.1, both the displacements and external forces prescribed on the body boundary are known *a priori*, as well the areas over which they are applied. The extension of this framework to potential contact between multiple bodies requires an additional contact boundary, which is unknown *a priori* and can change continuously over time. Therefore, the contact areas, contact forces and motions of associated boundaries are unknown in advance, which are determined as part of the solution. Hence, besides the nonlinear kinematic relations (large deformations) and nonlinear material behaviour, the contact interaction introduces a new type of nonlinearity into the continuum mechanics problem formulation.

This section presents the continuum based formulation of frictional contact problems, being the contact governing equations developed in the continuum form, suitable for large deformation/large sliding cases. This formulation involves two fundamental conditions: the

principle of impenetrability of one body by another and the law of friction on their common interface. Since these contact conditions are imposed at the interface between bodies, a rigorous description of the geometry of the contacting surfaces is mandatory for the continuum contact problem [Konyukhov 05]. The main notions and concepts of contact mechanics, including contact kinematics and statics, formulation of contact conditions as well the governing variational principle are presented. The basis of the presented formulation was first established by Laursen and Simo [Laursen 93], [Laursen 94] and Klarbring [Klarbring 95]. For further details, the interested reader is referred to classical textbook [Kikuchi 88] or more recent monographs on computational contact mechanics [Laursen 02] and [Wriggers 06].

2.2.1. Kinematic and static variables

From the computational point of view, contact problems can be classified in two important categories: (i) contact of a deformable body against a rigid obstacle (often referred as Signorini's problem) and (ii) two deformable bodies undergoing contact interaction. In fact, the contact problem involving a deformable body with a rigid obstacle is easily obtained as a particular case of the two body contact problem. Moreover, contact problems comprising self-contact (body contacting itself) and contact involving multiple bodies also represent special cases, which are characterized by presenting multiple separated contact zones. However, without any loss of generality, all mathematical fundamentals concerning contact kinematics and contact constraints can be formulated considering that the contact occurs at a single contact zone between two deformable bodies. Thus, attention is restricted to contact between two deformable bodies undergoing large deformations, with a single contact zone between them, as shown schematically in Figure 2.4. The details associated with contact interaction are introduced into the framework established in Section 2.1.1, for the nonlinear solid mechanics without contact. The bodies in the reference configuration are represented by the open sets $\Omega^{01} \subset \Re^3$ and $\Omega^{02} \subset \Re^3$, as shown in Figure 2.4. Their motion is expressed by the mappings $\boldsymbol{\varphi}^1$ and $\boldsymbol{\varphi}^2$, which cause physical contact between the bodies and produce contact forces during some portion of the time interval $\mathbb{I} = (0, T]$. The open sets $\Omega^1 \subset \Re^3$ and $\Omega^2 \subset \Re^3$ represent the bodies in the current configuration, which are obtained by application of the deformation mappings to open sets in the reference configuration. For each deformable body *i*, its boundary surface $\partial\Omega^i$ is subdivided in three subdomains: Γ_{u}^i where motions are prescribed (Dirichlet boundaries), Γ_{σ}^i where tractions are prescribed (Neumann boundaries) and Γ_{c}^i represents the potential contact surface. These three disjoint subsets satisfy the following conditions:

$$\begin{aligned} \Gamma_{\mathrm{u}}^{i} \cup \Gamma_{\sigma}^{i} \cup \Gamma_{\mathrm{c}}^{i} &= \partial\Omega^{i}, \\ \Gamma_{\mathrm{u}}^{i} \cap \Gamma_{\sigma}^{i} = \Gamma_{\mathrm{u}}^{i} \cap \Gamma_{\mathrm{c}}^{i} &= \Gamma_{\sigma}^{i} \cap \Gamma_{\mathrm{c}}^{i} = \varnothing, \end{aligned} \tag{2.24}$$

for each of the bodies *i*.

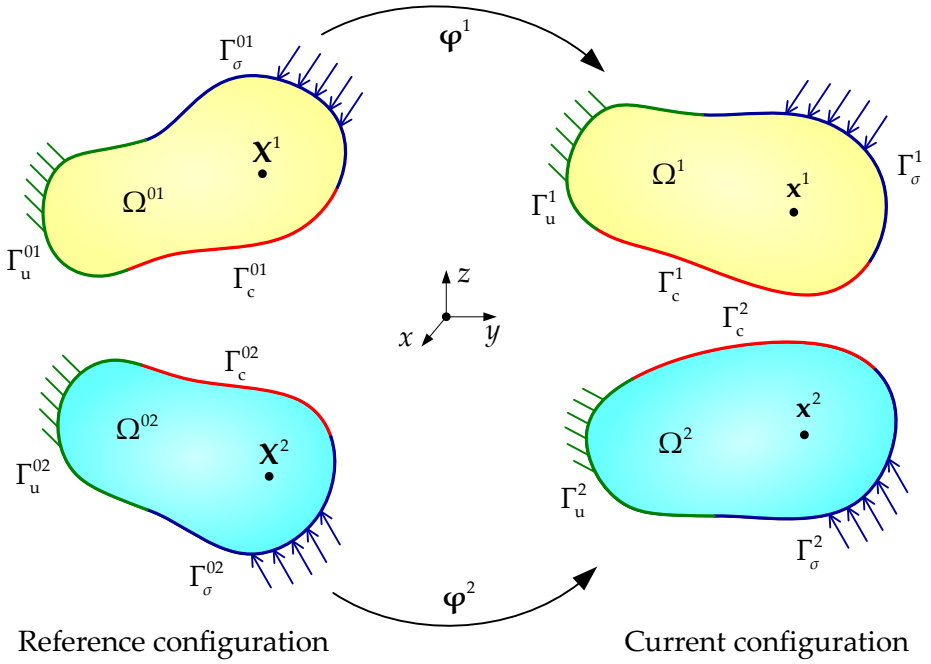


Figure 2.4. Basic notation for the two body large deformation contact problem.

According to the notation adopted in Section 2.1.1, for each deformable body *i*, the vector $\mathbf{X}^i \in \Omega^{0i}$ denotes the position of the material point in the reference configuration and the vector $\mathbf{x}^i \in \Omega^i$ in the current configuration (see Figure 2.4). Since two distinct points in the initial configuration can occupy the same position in the current configuration, contact conditions are formulated with respect to the current configuration. All points on the boundary $\partial\Omega$ where contact occurs are included in the potential contact surface Γ_c^i of each body *i*, location where the contact constraints are defined. Moreover, the potential contact surface can be divided into two nonintersecting sets: active contact surface $\overline{\Gamma}_\mathrm{c}^i \subset \Gamma_\mathrm{c}^i$ (points in contact) and inactive contact surface $\Gamma_\mathrm{c}^i \setminus \overline{\Gamma}_\mathrm{c}^i$ (points not in contact). Nevertheless, from the contact problems definition, the active contact surface $\overline{\Gamma}_\mathrm{c}^i$ is unknown *a priori* (can change over time), being determined as part of the nonlinear solution process.

In order to simplify the notation employed in the governing equations, the following shorthand notation is adopted [Yastrebov 13]:

- 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: $\partial\Omega \supset \Gamma_\sigma = \Gamma_\sigma^1 \cup \Gamma_\sigma^2$;
- 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 $\Gamma_c^2 \subset \partial\Omega^2$.

In order to define the kinematics of the body surfaces and contact constraints, it is useful to consider a parameterisation of the contact surfaces by means of a local coordinate system, shown schematically in Figure 2.5 for body 2. The contact surfaces are considered 2D smooth surfaces in Euclidean space $\Re^3$. Thus, all points on the contact surfaces can be parameterized by mappings ψ^i, such that $\psi^i : \mathcal{A}^i \to \Re^3$. The parametric domain $\mathcal{A}^i$ defines the set of points located in the potential contact surface mapped to $\Re^2$, being the parameterisation of the contact surfaces obtained through $\Gamma_c^{0i} = \psi_0^i(\mathcal{A}^i)$ and $\Gamma_c^i = \psi^i(\mathcal{A}^i)$. Considering any point in the reference configuration $\mathbf{X}^2 \in \Gamma_c^{02}$ on the potential contact surface of body 2, its vector position in the reference and current configuration can be obtained respectively by $\mathbf{X}^2 \in \psi_0^2(\xi)$ and $\mathbf{x}^2 \in \psi^2(\xi)$ for some point $\xi \in \mathcal{A}^2$, as shown in Figure 2.5. Furthermore, it is assumed that the parameterization ensures sufficient smoothness, such that the required derivatives can be evaluated.

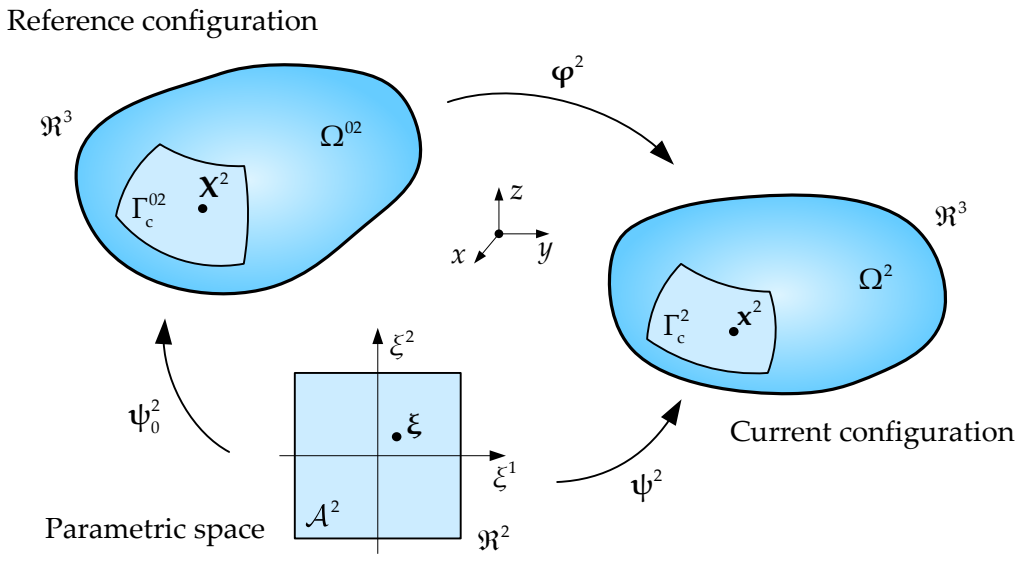


Figure 2.5. Parameterization of the potential contact surface for body 2.

2.2.1.1. Gap function and tangential relative sliding

In the framework of the continuum contact problem the classification of the bodies as master or slave is somewhat arbitrary in general, although the choice becomes important

in the discrete setting, as will be discussed in Section 3.3. In the following analysis the bodies 1 and 2 are conveniently referred as the slave and master body, respectively. Accordingly, the contact surface Γ_c^1 is denoted as the slave surface and Γ_c^2 is denoted as the master surface. In order to distinguish a specific point on the potential contact surface from a general point $\mathbf{x}^i \in \Omega^i$ (in the current configuration), the slave points are denoted by $\mathbf{x}^s \subset \mathbf{x}^1 \in \Gamma_c^1$ and master points by $\mathbf{x}^m \subset \mathbf{x}^2 \in \Gamma_c^2$.

The potential contact surfaces (slave and master) are parameterized in order to create a local surface coordinate attached to each contact surface, being the basis formed by two surface tangent vectors and an associated surface normal vector. The surface tangent vectors are given as partial derivatives of the configuration mapping ψ^i with respect to the surface parameterization. Hence, the tangential vector field (covariant basis vectors), defined in the current configuration for the master contact surface is denoted by:

$$\boldsymbol{\tau}_\alpha^m(\xi) \equiv \boldsymbol{\tau}_\alpha^m(\psi^2(\xi)) = \frac{\partial \psi^2(\xi)}{\partial \xi^\alpha}, \quad \alpha = 1,2, \tag{2.25}$$

where $\boldsymbol{\xi} = (\xi^1, \xi^2)$ defines the local parameterization of the master surface via convective coordinates, as illustrated in Figure 2.5. The surface normal vector can be computed from the tangent basis using the vector product formula, given by:

$$\mathbf{n}(\xi) \equiv \mathbf{n}(\psi^2(\xi)) = \frac{\boldsymbol{\tau}_1^m(\xi) \times \boldsymbol{\tau}_2^m(\xi)}{\left\| \boldsymbol{\tau}_1^m(\xi) \times \boldsymbol{\tau}_2^m(\xi) \right\|}, \tag{2.26}$$

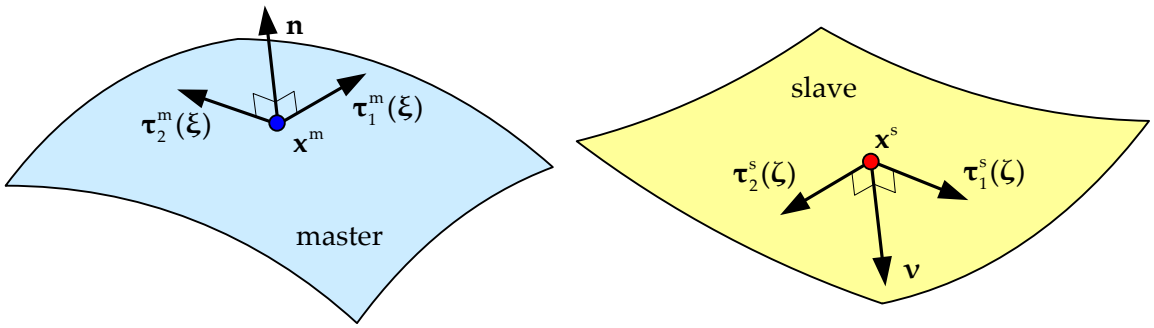
where $\times$ denotes the cross product between two vectors and $\|\bullet\|$ represents the Euclidean norm of a vector. The local surface coordinate system attached to the master surface is shown in Figure 2.6 (a). Concerning the slave surface, the tangential basis vectors (covariant basis vectors) in the current configuration are given by:

$$\boldsymbol{\tau}_\alpha^s(\zeta) \equiv \boldsymbol{\tau}_\alpha^s(\psi^1(\zeta)) = \frac{\partial \psi^1(\zeta)}{\partial \zeta^\alpha}, \quad \alpha = 1,2, \tag{2.27}$$

where the parameterization of the master surface is defined by $\boldsymbol{\zeta} = (\zeta^1, \zeta^2)$, which is obtained analogous to the master surface shown in Figure 2.5. Therefore, the surface normal vector is determined by:

$$\boldsymbol{\nu}(\zeta) \equiv \boldsymbol{\nu}(\psi^1(\zeta)) = \frac{\boldsymbol{\tau}_1^s(\zeta) \times \boldsymbol{\tau}_2^s(\zeta)}{\left\| \boldsymbol{\tau}_1^s(\zeta) \times \boldsymbol{\tau}_2^s(\zeta) \right\|}. \tag{2.28}$$

In both (2.26) and (2.28), it is assumed that the tangent vectors present proper orientations, such that the contact surfaces are oriented with positive directions pointing outside the bodies, as shown in Figure 2.6.



(a) (b)

Figure 2.6. Definition of the local coordinate system in the contact surfaces: (a) master surface; (b) slave surface.

The kinematic variables used to measure the relative motion of two bodies coming into contact are the normal gap function and the tangential relative sliding. All quantities related with contact need to be evaluated in the current configuration, even though the boundary value problem may still be formulated with respect to the reference configuration (2.23). The contact interaction between bodies at each instant t establishes a one-to-one correspondence between each point of the slave surface and its counterpart point on the master surface. Thus, for each slave point $\mathbf{x}^{\mathrm{s}} \in \Gamma_{\mathrm{c}}^{1}$, the corresponding master point $\mathbf{x}^{\mathrm{m}} \in \Gamma_{\mathrm{c}}^{2}$ is identified for the purposes of contact kinematic relations, even if the bodies are not in contact. Among different alternatives, typically the contact point $\mathbf{x}^{\mathrm{m}} \in \Gamma_{\mathrm{c}}^{2}$ is determined according to the closest point projection of the slave point $\mathbf{x}^{\mathrm{s}}$ to the master surface Γ_{c}^{2}, defined as:

$$\bar{\mathbf{x}}^{\mathrm{m}}(\mathbf{x}^{\mathrm{s}}) = \arg\min_{\mathbf{x}^{\mathrm{m}} \in \Gamma_{\mathrm{c}}^{2}} \left\| \mathbf{x}^{\mathrm{s}} - \mathbf{x}^{\mathrm{m}}(\xi) \right\|, \tag{2.29}$$

where the terminology $\bar{\mathbf{x}}^{\mathrm{m}}(\mathbf{x}^{\mathrm{s}})$ indicates a one-to-one correspondence between points $\mathbf{x}^{\mathrm{s}}$ and $\bar{\mathbf{x}}^{\mathrm{m}}$, resulting from the minimization problem presented in (2.29). Under certain restrictions (at least locally convex region), the closest point problem (2.29) results in the orthogonal projection of the slave point onto the master surface, where the normal gap function is defined as:

$$g_{\mathrm{n}}(\mathbf{x}^{\mathrm{s}}) = (\mathbf{x}^{\mathrm{s}} - \bar{\mathbf{x}}^{\mathrm{m}}) \cdot \bar{\mathbf{n}}, \tag{2.30}$$

where $\mathbf{n}$ denotes the outward unit normal vector to Γ_{c}^{2} at the point $\bar{\mathbf{x}}^{\mathrm{m}}$. Figure 2.7 presents the orthogonal projection of a given slave point $\mathbf{x}^{\mathrm{s}}$ onto the current master surface Γ_{c}^{2}. The resulting projection point and other quantities evaluated at the solution point are denoted by a bar over the quantity, as illustrated in Figure 2.7. The normal gap

function g_n can be expressed as the signed distance between the points $\mathbf{x}^s$ and $\bar{\mathbf{x}}^m$, measured in the normal direction to the master surface.

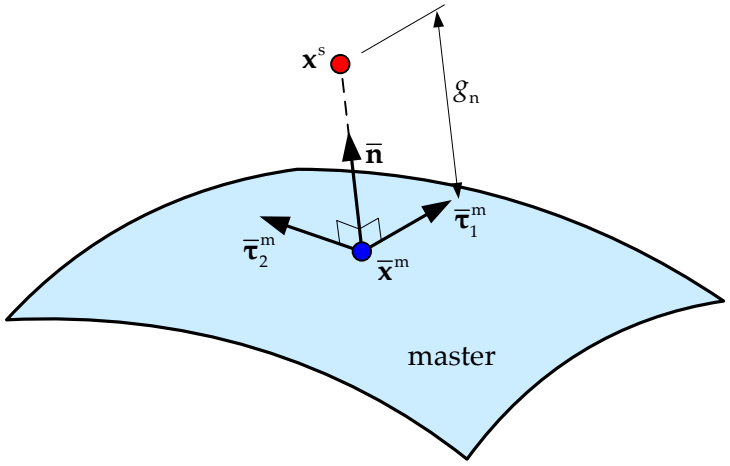


Figure 2.7. Orthogonal projection of the slave point $\mathbf{x}^s$ onto the master surface.

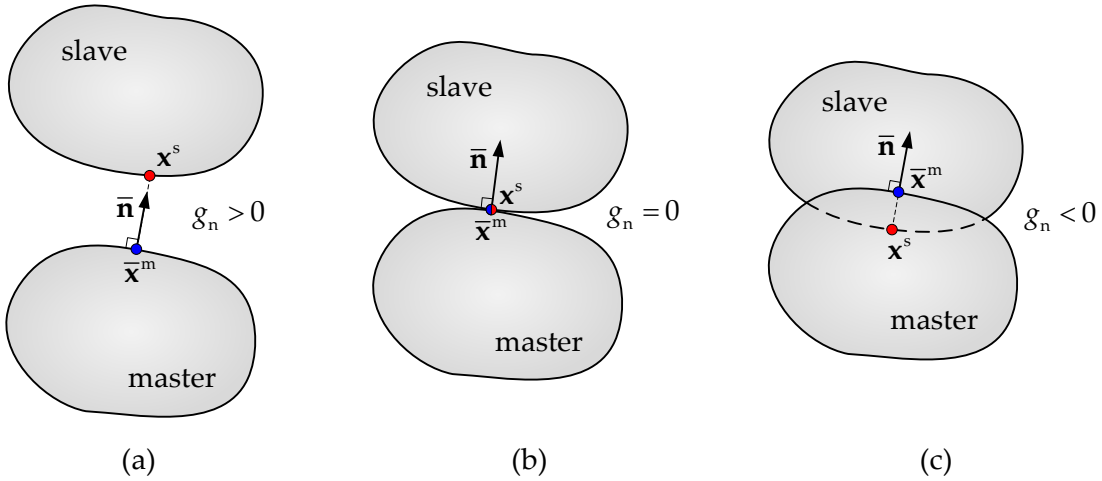


Figure 2.8. Three different geometrical conditions of the slave point with respect to the master body: (a) separated; (b) contact; (c) penetration.

Note that applying the definition of normal gap function presented in (2.30), this quantity is positive if the slave point is outside of the solid enveloped by the master surface (see Figure 2.7), otherwise it will be negative. The three possible geometrical situations of a slave point with respect to the master body are schematically indicated in Figure 2.8. Despite the simple definition of the normal gap function presented in (2.30), its evaluation for each point of the slave surface, as a distance to the closest point of the master surface, is not a trivial task. Indeed, there is an ambiguity in the definition of the normal gap between

contacting surfaces. Thus, three problems can arise from such definition [Wriggers 06], [Yastrebov 13]:

- Asymmetry of contact surfaces (see Figure 2.9 (a));
- Non-uniqueness of the closest point (see Figure 2.9 (b));
- Non-existence of the orthogonal projection point (see Figure 2.9 (c)).

All these problems affect the geometrical description of the contact. The asymmetry of the normal gap function results in an asymmetric treatment of contact surfaces, resulting from the so-called master-slave approach. If instead of looking for the closest master point to the slave point (2.29), inverting the problem and searches for the slave point closest to the master point, the obtained solutions may not coincide, as shown in Figure 2.9 (a). Therefore, the one-to-one equitable correspondence between surface points does not exist.

Another difficulty with the definition of the gap function given in (2.30) is that the projection point $\bar{\mathbf{x}}^{m}$ may not be unique, which means that the gap function is locally non-differentiable. Figure 2.9 (b) shows an example of a slave point containing two admissible solutions for the orthogonal projection. The non-uniqueness of the closest point leads to non-physical discontinuities in the sliding path, which are not allowed for a rigorous description of the frictional contact problem. In practice, this discontinuity is not of major consequence, because in the next step the point gets closer to one side and the discontinuity disappears. The uniqueness of the closest point projection associated to the curvature of the surface has been discussed by Heergaard and Curnier [Heegaard 96] and Konyukhov [Konyukhov 08]. A very complete overview of the closest point projection, its mathematical properties (convex analysis) and the fundamental problems of existence and uniqueness of the solution can be found in [Konyukhov 08].

The non-smoothness of the contacting surface (master) represents another cause for the arising of multiple closest points and generates blind angles in the orthogonal projection domain, as shown in the example of Figure 2.9 (c). Moreover, the discontinuity in the normal vector field leads to oscillations and possible divergence of the numerical solution. In fact, the smoothness of the contacting surface is not a sufficient but a necessary condition for the existence of the orthogonal projection point [Yastrebov 13]. However, typically the surfaces in the finite element framework are only piecewise smooth due to the discretization, leading to a discontinuity in the normal vector field. This problem will be discussed in more detail in Chapter 4.

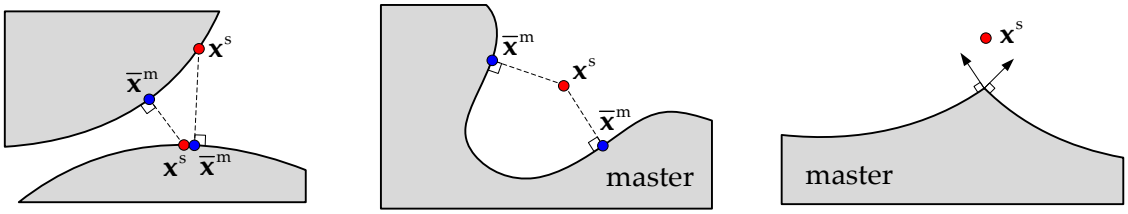


Figure 2.9. Geometry related problems: (a) asymmetry of the closest point definition; (b) non-uniqueness of the closest point; (c) nonexistence of the orthogonal projection point.

The change of the closest point projection describes the tangential relative sliding between surfaces, which is necessary for modelling friction effects. The definition of the contact relative velocity is more complicated than the normal gap function, since it requires taking into account both the motion of the slave point and the master surface, to define the relative motion between the two surfaces. In order to express the tangential relative sliding, a simple 1D example is shown in Figure 2.10, where the relative motion of a point A over a straight segment BC is presented [Yastrebov 13]. The absolute velocities of the points A, B and C are denoted by v^{A}, v^{B} and v^{C}, respectively. The objective is to express the relative tangential velocity of the point A' (projection of the point A) on the segment BC.

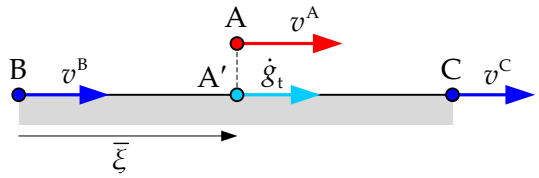


Figure 2.10. Definition of relative tangential velocity (adapted from [Yastrebov 13]).

Considering that the segment BC is parameterized with $\xi \in [0,1]$, then the coordinate of point A' can be expressed as:

$$x^{\mathrm{A}'} = \bar{\xi} x^{\mathrm{C}} + (1-\bar{\xi}) x^{\mathrm{B}}, \tag{2.31}$$

where x^{B} and x^{C} denotes the coordinates of points B and C, respectively. Since the problem is 1D, the points A and A' have the same coordinate $x^{\mathrm{A}} = x^{\mathrm{A}'}$, being the parameter value $\bar{\xi}$ given by:

$$\bar{\xi} = \frac{x^{\mathrm{A}} - x^{\mathrm{B}}}{x^{\mathrm{C}} - x^{\mathrm{B}}}. \tag{2.32}$$

Then, the absolute velocity of the projection point can be evaluated as:

$$v^{\mathrm{A}'} = \underbrace{\bar{\xi} v^{\mathrm{C}} + (1-\bar{\xi}) v^{\mathrm{B}}}_{\frac{\partial x^{\mathrm{A}'}}{\partial t}} + \underbrace{(x^{\mathrm{C}} - x^{\mathrm{B}})\dot{\bar{\xi}}}_{\frac{\partial x^{\mathrm{A}'}}{\partial \xi}\frac{\partial \bar{\xi}}{\partial t}}, \tag{2.33}$$

which is composed by two components, related with the motion of each body. Therefore, the relative tangential velocity $\dot{g}_{\mathrm{t}}$ is obtained subtracting from (2.33) the absolute velocity of the projection point for a fixed parameter $\bar{\xi}$, being given by:

$$\dot{g}_{\mathrm{t}} = \frac{\partial x^{\mathrm{A}'}}{\partial \xi}\dot{\bar{\xi}}, \tag{2.34}$$

where the parameter $\dot{\bar{\xi}}$ can be determined using the derivative of Eq. (2.31). Indeed, the relative tangential velocity (2.34) is the derivative of the projection point coordinate $x^{\mathrm{A}'}$ with the changing field of the segment BC. Note that the tangential relative sliding velocity is not the time derivative of any function vector.

In the general case of contact between two bodies, the tangential relative sliding velocity is given as:

$$\dot{\mathbf{g}}_{\mathrm{t}} = \bar{\boldsymbol{\tau}}_{\alpha}^{\mathrm{m}} \dot{\bar{\xi}}^{\alpha}, \quad \alpha = 1,2, \tag{2.35}$$

where $\bar{\boldsymbol{\tau}}_{\alpha}^{\mathrm{m}}$ are the covariant tangential basis vectors defined in (2.27) and $\bar{\boldsymbol{\xi}} = (\bar{\xi}^1, \bar{\xi}^2)$ are the convective coordinates of the master surface at the projection point. The time derivative of $\bar{\xi}^{\alpha}$ can be computed from the relation:

$$(\mathbf{x}^{\mathrm{s}} - \bar{\mathbf{x}}^{\mathrm{m}}) \cdot \bar{\boldsymbol{\tau}}_{\alpha}^{\mathrm{m}} = 0, \quad \alpha = 1,2, \tag{2.36}$$

which is valid at the slave contact point, since the difference $\mathbf{x}^{\mathrm{s}} - \bar{\mathbf{x}}^{\mathrm{m}}$ is orthogonal to the tangent vectors of the master surface at the projection point $\bar{\mathbf{x}}^{\mathrm{m}}$, as shown in Figure 2.7. After some algebraic operations, starting with the time derivative of (2.36), an expression for $\dot{\bar{\xi}}^{\alpha}$ in terms of the material velocities is obtained:

$$\dot{\bar{\xi}}^{\alpha} = \left[\bar{m}_{\alpha\beta} - g_{\mathrm{n}} \bar{b}_{\alpha\beta} \right]^{-1} \left((\dot{\mathbf{x}}^{\mathrm{s}} - \dot{\bar{\mathbf{x}}}^{\mathrm{m}}) \cdot \bar{\boldsymbol{\tau}}_{\beta}^{\mathrm{m}} + g_{\mathrm{n}} \bar{\mathbf{n}} \cdot \frac{\partial \dot{\bar{\mathbf{x}}}^{\mathrm{m}}}{\partial \xi^{\beta}} \right), \quad \alpha, \beta = 1,2, \tag{2.37}$$

where $\bar{m}_{\alpha\beta}$ and $\bar{b}_{\alpha\beta}$ denote the components of the first and second covariant fundamental surface tensors, evaluated at the solution point $\bar{\boldsymbol{\xi}}$, which are defined by:

$$\begin{aligned} m_{\alpha\beta} &= \boldsymbol{\tau}_{\alpha}^{\mathrm{m}} \cdot \boldsymbol{\tau}_{\beta}^{\mathrm{m}}, \quad \alpha, \beta = 1,2, \\ b_{\alpha\beta} &= \mathbf{n} \cdot \frac{\partial \boldsymbol{\tau}_{\alpha}^{\mathrm{m}}}{\partial \xi^{\beta}}, \quad \alpha, \beta = 1,2. \end{aligned} \tag{2.38}$$

In the particular case of perfect sliding, assuming that $\mathbf{x}^{\mathrm{s}}$ and $\bar{\mathbf{x}}^{\mathrm{m}}$ are exactly coincident for some interval of time, i.e. the normal gap is assumed to be relatively small $g_{\mathrm{n}} = \dot{g}_{\mathrm{n}} = 0$, then Eq. (2.37) simplifies to:

$$\dot{\bar{\xi}}^{\alpha} = (\dot{\mathbf{x}}^{\mathrm{s}} - \dot{\bar{\mathbf{x}}}^{\mathrm{m}}) \cdot \bar{\boldsymbol{\tau}}^{\alpha \mathrm{m}}, \quad \alpha = 1,2, \tag{2.39}$$

where $\boldsymbol{\tau}^{\alpha \mathrm{m}} = m^{\alpha\beta} \boldsymbol{\tau}_{\beta}^{\mathrm{m}}$ represents the contravariant components of the metric tensor and the relationship $(m_{\alpha\beta})^{-1} = m^{\alpha\beta}$ is obtained through the conjugacy property of the metric tensor defined by $m_{\alpha\beta} m^{\alpha\beta} = \delta_{\alpha\beta}$, representing the Kronecker delta. This simplification allows recovering the interpretation of $\dot{\bar{\xi}}^{\alpha}$ as the components of the tangential relative velocity between $\mathbf{x}^{\mathrm{s}}$ and $\bar{\mathbf{x}}^{\mathrm{m}}$ during active sliding.

In order to prevent one body from penetrating other one, a contact pressure arises at the contact zone. The static variables used to model the frictional contact interactions are called contact forces or stresses, which are necessary for establishing the global and local contact principles. Since adhesive stresses will not be allowed in the contact interface, the contact pressure is assumed negative in compression and zero in inactive contact zones. In continuum mechanics, the contact traction is expressed by the Cauchy stress vector at the surface of the body, as defined in (2.10). Thus, the contact traction vector is defined for the slave body surface in the current configuration as:

$$\mathbf{t} = \boldsymbol{\sigma} \boldsymbol{\nu}, \tag{2.40}$$

where $\boldsymbol{\nu}$ is the outward unit normal vector at the slave surface and $\boldsymbol{\sigma}$ denotes the Cauchy stress tensor. Similar to the kinematic measures (2.30) and (2.35), the contact traction can be decomposed into normal and tangential components. Since the normal vectors of slave and master surfaces are equal and opposite $\mathbf{n} = -\boldsymbol{\nu}$ in the active contact zone, then the contact pressure can be written as:

$$p_{\mathrm{n}} = \mathbf{t} \cdot \boldsymbol{\nu} = (\boldsymbol{\sigma} \mathbf{n}) \cdot \mathbf{n}, \tag{2.41}$$

which is assumed negative in compression. The tangential components of the Cauchy stress vector (frictional traction) are defined as:

$$\mathbf{t}_{\mathrm{t}} = \mathbf{t} - p_{\mathrm{n}} \boldsymbol{\nu} = (\mathbf{I} - \boldsymbol{\nu} \otimes \boldsymbol{\nu}) \mathbf{t} = (\mathbf{I} - \mathbf{n} \otimes \mathbf{n}) \mathbf{t}, \tag{2.42}$$

which physically represents the frictional force exerted by the master surface on the slave point. Besides, the contact traction vector satisfies the action–reaction principle in the contact point, defining that the contact force exerted by the slave point on the master surface equal and opposite. Note that the tangential stress is zero in the case of frictionless contact $\mathbf{t}_{\mathrm{t}} = \mathbf{0}$.

2.2.1.2. First and second order variations of kinematic quantities

The variation of the geometrical contact quantities are needed for solving the contact contribution in the virtual work balance in the current configuration (2.21). This section contains the first and second order variations of both the normal gap function g_{n} and the tangential relative slip $\mathbf{g}_{\mathrm{t}}$, which are essential for the linearization of the frictional contact formulation. Detailed discussion and rigorous mathematical deduction of the presented variations can be found in [Laursen 93], [Pietrzak 97] and [Yastrebov 13]. Considering that the orthogonal projection point, defined in the Euclidean space $\bar{\mathbf{x}}^{\mathrm{m}}$, is defined in the parametric domain $\mathcal{A}^2$ by the convective coordinates $\bar{\boldsymbol{\xi}} = (\bar{\xi}^1, \bar{\xi}^2)$, then $\bar{\mathbf{x}}^{\mathrm{m}} = \boldsymbol{\psi}^2(\bar{\boldsymbol{\xi}}) \in \Gamma_{\mathrm{c}}^2$. This point on the master surface can change its geometrical position due to the slave body deformation, since it results from the orthogonal projection solution, being not attached to any material particle of the master surface. Therefore, variations of $\bar{\mathbf{x}}^{\mathrm{m}}$ result both from deformation of the master/slave body and from relative movement of the projection due to the change of the reference point $\bar{\boldsymbol{\xi}}$. The first and second variations with respect to the time variable t are denoted by δ and Δ, respectively.

The first variation of the normal gap is given by:

$$\delta g_{\mathrm{n}} = (\delta \mathbf{x}^{\mathrm{s}} - \delta \bar{\mathbf{x}}^{\mathrm{m}}) \cdot \bar{\mathbf{n}}, \tag{2.43}$$

which is obtained through the derivative of the normal gap function (2.30) with respect to time and posteriorly multiplied by the master surface normal vector. The first variation of the tangential relative slip is obtained in a similar way, but the derivative with respect to time is multiplied by the tangent base vector. Thus, the first variation of the tangential relative slip can be stated as the time derivative presented in (2.35), obtaining the result:

$$\delta \mathbf{g}_{\mathrm{t}} = \delta \bar{\xi}^{\alpha} \bar{\boldsymbol{\tau}}_{\alpha}^{\mathrm{m}}, \quad \alpha = 1,2, \tag{2.44}$$

where the first variation of $\bar{\xi}^{\alpha}$ is defined by:

$$\delta \bar{\xi}^{\alpha} = \left[\bar{m}_{\alpha\beta} - g_{\mathrm{n}} \bar{b}_{\alpha\beta} \right]^{-1} \left((\delta \mathbf{x}^{\mathrm{s}} - \delta \bar{\mathbf{x}}^{\mathrm{m}}) \cdot \bar{\boldsymbol{\tau}}_{\beta}^{\mathrm{m}} + g_{\mathrm{n}} \bar{\mathbf{v}} \cdot \frac{\partial (\delta \bar{\mathbf{x}}^{\mathrm{m}})}{\partial \xi^{\beta}} \right), \quad \alpha, \beta = 1,2. \tag{2.45}$$

This relation derives simply from (2.37) by replacing the velocities by the test displacements, which is the classical expression originally obtained in [Laursen 93].

The second order variation of the normal gap is computed after a reasonable amount of algebra, being given by:

$$\Delta\delta g_{\mathrm{n}} = -\bar{\mathbf{v}}\cdot\left(\frac{\partial(\delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\alpha}}\Delta\bar{\xi}^{\alpha}+\delta\bar{\xi}^{\alpha}\frac{\partial(\Delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\alpha}}+\delta\bar{\xi}^{\alpha}\frac{\partial^{2}\bar{\mathbf{x}}^{\mathrm{m}}}{\partial\xi^{\alpha}\partial\xi^{\beta}}\Delta\bar{\xi}^{\beta}\right)+$$
$$+g_{\mathrm{n}}\bar{\mathbf{v}}\cdot\left(\frac{\partial(\delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\alpha}}+\frac{\partial^{2}\bar{\mathbf{x}}^{\mathrm{m}}}{\partial\xi^{\alpha}\partial\xi^{\gamma}}\delta\bar{\xi}^{\gamma}\right)m^{\alpha\beta}\left(\frac{\partial(\Delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\beta}}+\frac{\partial^{2}\bar{\mathbf{x}}^{\mathrm{m}}}{\partial\xi^{\beta}\partial\xi^{\sigma}}\Delta\bar{\xi}^{\sigma}\right)\cdot\bar{\mathbf{v}}, \tag{2.46}$$
$$\alpha,\beta,\gamma,\sigma=1,2,$$

where $\Delta\bar{\xi}^{\alpha}$ is defined as given in (2.45) with the substitution of Δ for δ. The expression (2.46) has been also obtained in [Laursen 93]. Sometimes the normal gap is assumed to be relatively small $g_{\mathrm{n}}\approx 0$, providing a simpler expression for the second order variation of the normal gap, which is defined only by the first term of Eq. (2.46). Finally, the second order variation of the surface parameter is given by:

$$\Delta\delta\bar{\xi}^{\alpha} = \left[\bar{m}_{\alpha\beta}-g_{\mathrm{n}}\bar{b}_{\alpha\beta}\right]^{-1}\left\{-\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\beta}\cdot\left(\delta\bar{\xi}^{\sigma}\frac{\partial(\Delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\sigma}}+\frac{\partial(\delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\gamma}}\Delta\bar{\xi}^{\gamma}\right)-\right.$$
$$-\delta\bar{\xi}^{\gamma}\left(\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\beta}\cdot\frac{\partial\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\gamma}}{\partial\xi^{\sigma}}-g_{\mathrm{n}}\bar{\mathbf{v}}\cdot\frac{\partial^{2}\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\beta}}{\partial\xi^{\gamma}\partial\xi^{\sigma}}\right)\Delta\bar{\xi}^{\sigma}-\delta\bar{\xi}^{\gamma}\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\gamma}\cdot\left(\frac{\partial(\Delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\beta}}+\frac{\partial\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\beta}}{\partial\xi^{\sigma}}\Delta\bar{\xi}^{\sigma}\right)-$$
$$-\left(\frac{\partial(\delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\beta}}+\frac{\partial\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\beta}}{\partial\xi^{\gamma}}\delta\bar{\xi}^{\gamma}\right)\cdot\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\sigma}\Delta\bar{\xi}^{\sigma}+g_{\mathrm{n}}\bar{\mathbf{v}}\cdot\left(\frac{\partial^{2}(\delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\beta}\partial\xi^{\gamma}}\Delta\bar{\xi}^{\gamma}+\delta\bar{\xi}^{\sigma}\frac{\partial^{2}(\Delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\beta}\partial\xi^{\sigma}}\right)+ \tag{2.47}$$
$$\left.+(\delta\mathbf{x}^{\mathrm{s}}-\delta\bar{\mathbf{x}}^{\mathrm{m}})\cdot\left(\frac{\partial(\Delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\beta}}+\frac{\partial\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\beta}}{\partial\xi^{\sigma}}\Delta\bar{\xi}^{\sigma}\right)+\left(\frac{\partial(\delta\bar{\mathbf{x}}^{\mathrm{m}})}{\partial\xi^{\beta}}+\frac{\partial\bar{\boldsymbol{\tau}}^{\mathrm{m}}_{\beta}}{\partial\xi^{\gamma}}\delta\bar{\xi}^{\gamma}\right)\cdot(\Delta\mathbf{x}^{\mathrm{s}}-\Delta\bar{\mathbf{x}}^{\mathrm{m}})\right\},$$
$$\alpha,\beta,\gamma,\sigma=1,2,$$

which is coincident with the one obtained in [Laursen 93]. The approximation for small gaps $g_{\mathrm{n}}\approx 0$ leads to a simpler form, by neglecting third derivative of the master surface vector included in (2.47).

2.2.2. Contact and friction laws

The mechanical formulation of frictional contact problems involves additional constraints related with impenetrability and friction conditions. These constraints are expressed considering relationships between the previously presented kinematic and static variables, such that the contact traction in the normal and in the tangential directions is coupled with the normal distance and the tangential slip velocity, respectively.

2.2.2.1. Unilateral contact law

The unilateral contact condition defines the physical requirements of impenetrability and compressive interaction between the bodies. Due to the sign convention chosen for the gap function defined in (2.30), a positive value $g_{\mathrm{n}}>0$ defines points not in contact (see

Figure 2.8 (a)), while a negative value $g_n < 0$ denotes penetration between bodies (see Figure 2.8 (c)), which is physically not admissible. Therefore, the normal contact conditions on the contact interface can be formulated using the classical set of Karush–Kuhn–Tucker (KKT) conditions for optimality, being stated as:

$$g_n \geq 0, \quad p_n \leq 0, \quad p_n g_n = 0, \tag{2.48}$$

which must hold for all points on the slave contact surface. These conditions are commonly called Hertz–Signorini–Moreau conditions for frictionless contact. The first condition in (2.48) simply states the geometric impenetrability condition, whereas the second condition one refers that the contact pressure must be compressive (no adhesive stresses are allowed) in the contact zone. The last condition in (2.48) states the complementarity condition, which forces the gap to be zero if compressive tractions occur and the pressure to be zero if the gap function is positive. Moreover, the unilateral contact conditions can be decomposed into two parts: active and inactive contact zones, defined as:

$$\begin{cases} g_n = 0, & p_n < 0, \quad \text{at } \bar{\Gamma}_c^1 \\ g_n > 0, & p_n = 0, \quad \text{at } \Gamma_c^1 \setminus \bar{\Gamma}_c^1 \end{cases}, \tag{2.49}$$

where the two possible geometrical situations denoted as contact and gap are illustrated in Figure 2.11.

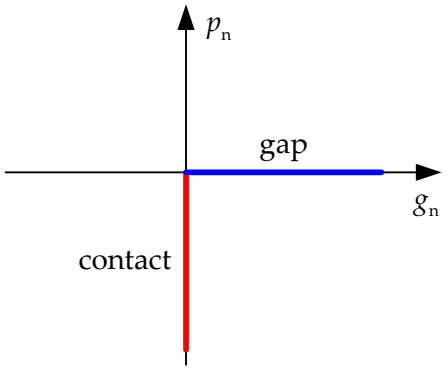


Figure 2.11. Unilateral contact law defined by the Karush–Kuhn–Tucker conditions.

The set of inequality conditions stated in (2.48) leads to a non-smooth contact law for the normal contact pressure, but also the relation between the contact pressure and the normal gap is multivalued at $g_n = 0$, as shown in Figure 2.11. This means that the contact law is non-differentiable and can take an infinite number of values at the origin. This singularity can be physically interpreted as a result from the fact that the contact pressure is a reaction force, which cannot be calculated by the unilateral contact law (2.48). It is a result of the equilibrium between the bodies in contact. This difficulty usually arises also in optimization problems subjected to inequality constraints. Thus, standard solution

techniques from optimization theory can be adapted for contact mechanics. In spite of these peculiarities, the unilateral contact law (2.48) can be expressed by sub-gradients of non-differentiable convex potentials, following Moreau's works on convex analysis. This formalism applied to contact problems is stated in the works of Alart and Curnier [Alart 91] and Heegaard and Curnier [Heegaard 93].

Since the present analysis is restricted to a macroscopic description of normal contact compliance, the contact constraints defined in (2.48) for the contact interface are based on a purely geometrical perspective. Consequently, the micromechanical behaviour is not taken into account, which in general depends upon material parameters such as hardness and on the geometrical parameter such as surface roughness [Wriggers 06]. Nevertheless, some contact problems require the knowledge of the micromechanical interaction for a proper treatment of the physical phenomena. Hence, some constitutive equations for normal contact have been developed based on experiments, which take into account the micro deformation of the contact bodies [Yastrebov 12].

2.2.2.2. Contact friction law

In this study, the contact with friction response is restricted to simple dry friction, which is formulated through the classical non-associated Coulomb's friction law. This law establishes that the frictional force at the contact interface depends on the contact pressure, which is only known at the solution (cf. Section 2.2.2.1). Thus, the constraints imposed by the friction law are solution dependent, leading to additional difficulties in the formulation of the frictional contact problem. The Coulomb's non-associated friction law can be described by the following three conditions:

$$\left\|\mathbf{t}_{\mathrm{t}}\right\| - \mu\left|p_{\mathrm{n}}\right| \le 0, \quad \mathbf{t}_{\mathrm{t}} + \mu\left|p_{\mathrm{n}}\right|\frac{\dot{\mathbf{g}}_{\mathrm{t}}}{\left\|\dot{\mathbf{g}}_{\mathrm{t}}\right\|} = \mathbf{0}, \quad \left\|\dot{\mathbf{g}}_{\mathrm{t}}\right\|\left(\left\|\mathbf{t}_{\mathrm{t}}\right\| - \mu\left|p_{\mathrm{n}}\right|\right) = 0, \tag{2.50}$$

where μ is the friction coefficient, which depends of the material properties of the bodies which define the interface. This coefficient is assumed constant in the classical Coulomb law. The first condition in (2.50), usually referred as Coulomb friction condition, imposes that the magnitude of the frictional force does not exceed a threshold value, defined as the product of the friction coefficient by the contact pressure modulus (see Figure 2.12). The second condition in (2.50), also referred as slip rule, defines that the frictional force vector is collinear with the tangential relative sliding velocity. Besides, the frictional force arising at the slave point is opposite to its slip direction. The last equation in (2.50) is a complementarity condition, which distinguish two different contact situations, usually referred as stick and slip statuses. If the frictional force has not reached the Coulomb threshold, the contact point in contact is not allowed to move in the tangential direction (null velocity), assigned a commonly called stick status. On the other hand, when the

tangential traction reaches the Coulomb limit, the contact point moves in the tangential direction of the contact interface, the point assigned with the slip status. The conditions (2.50) can be represented graphically, as shown in Figure 2.12, where stick and slip statuses are depicted.

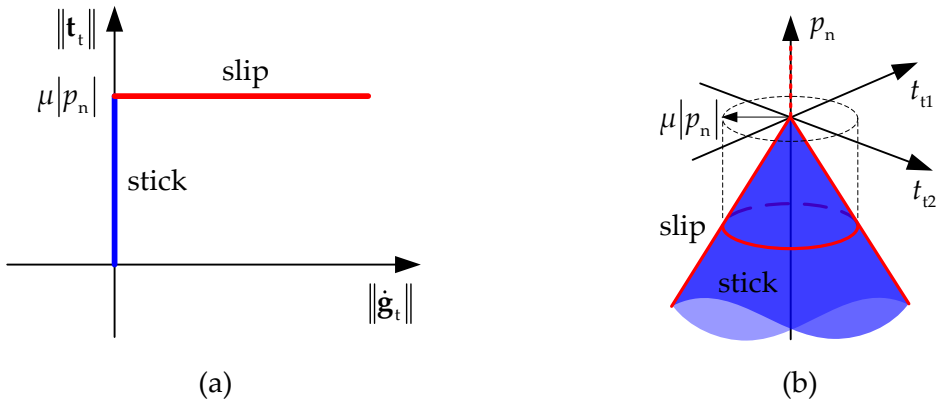


Figure 2.12. Graphical representation of Coulomb's frictional conditions: (a) relation between the norm of the tangential velocity and the norm of the frictional force vector; (b) relation between the contact pressure and the components of the frictional force vector (Coulomb's frictional cone).

Such as in the unilateral contact law, the Coulomb's friction law also yields a non-smooth functional at the onset of sliding (see Figure 2.12 (a)), creating difficulties from a numerical point of view. However, the friction law can also be expressed by means of convex analysis in the form of subgradients of non-differentiable, convex quasi-potentials as presented by Alart and Curnier [Alart 91] and Pietrzak and Curnier [Pietrzak 99]. The cone depicted in Figure 2.12 (b) is called the Coulomb's cone, which relates the contact pressure and the components of the contact tangential stress vector, defined in the first condition of (2.50). Thus, any admissible contact stress vector corresponds to a unique point either in the interior of the cone (stick status) or on its closure (slip status). The change of position in the interior of Coulomb's cone does not result in relative tangential displacements, while the relative sliding implies that the point is located on Coulomb's cone surface. The analogy of the non-associated Coulomb's friction law with the simple elastoplasticity formulation obtained for the rigid–perfectly plastic material model has been investigated by several authors, in the development of numerical algorithms for friction [Curnier 84], [Simo 98]. The slip condition given by the first condition in (2.50) may be understood as the yield condition in rigid plasticity. The contact tractions correspond to the stress tensor in the plasticity theory and the relative sliding velocity is analogous to the plastic strain rate. This analogy is performed considering both associated and non-associated plasticity in the study carried out by [Michalowski 78].

All frictional contact conditions discussed previously, comprising both the unilateral contact law (Section 2.2.2.1) and the Coulomb's friction law (Section 2.2.2.2), can be summarized into three distinct zones:

$$\begin{cases} g_n = 0, \quad \dot{\mathbf{g}}_t = \mathbf{0}, \quad p_n < 0, \quad \|\mathbf{t}_t\| < \mu |p_n|, & \text{in the stick zone } \Gamma_c^{1\bullet} \\ g_n = 0, \quad p_n < 0, \quad \mathbf{t}_t + \mu |p_n| \dfrac{\dot{\mathbf{g}}_t}{\|\dot{\mathbf{g}}_t\|} = \mathbf{0}, & \text{in the slip zone } \Gamma_c^{1*} \\ g_n > 0, \quad p_n = 0, \quad \mathbf{t}_t = \mathbf{0}, & \text{in the inactive contact zone } \Gamma_c^1 \setminus \bar{\Gamma}_c^1 \end{cases}, \tag{2.51}$$

The active contact zone $\bar{\Gamma}_c^1$ is split into a stick $\Gamma_c^{1\bullet}$ and a slip Γ_c^{1*} zone, such that $\Gamma_c^{1\bullet} \cup \Gamma_c^{1*} = \bar{\Gamma}_c^1$ and $\Gamma_c^{1\bullet} \cap \Gamma_c^{1*} = \varnothing$. On the other hand, the inactive contact zone does not presents any imposed boundary condition, as expressed in (2.51). In fact, the contact nonlinearities result from the presence of unknown *a priori* active contact zones, within which it is necessary to distinguish between stick and slip contact status.

2.2.3. Variational formulation of the contact problem

The variational formulation of the frictional contact problem involving two bodies undergoing large deformations is presented in this section, following the framework established in Section 2.1, for the nonlinear solid mechanics without contact interaction. Both bodies must satisfy the balance of virtual work, expressed in the current configuration by (2.21) for each body. Then, the global variational principle for the two body system is obtained by adding the two weak forms (master and slave bodies), which is expressed as:

$$\int_\Omega \boldsymbol{\sigma} : \delta\nabla\mathbf{u} d\Omega - \int_\Omega \mathbf{f}_v \cdot \delta\mathbf{u} d\Omega - \int_{\partial\Omega} (\boldsymbol{\sigma}\mathbf{v}) \cdot \delta\mathbf{u} d\Gamma = 0, \tag{2.52}$$

where the contact contribution to the variational principle is included in the last term of (2.52), where the Cauchy stress vector is non-null only in the active contact zones ($\bar{\Gamma}_c^1 \in \Gamma_c^1$ and $\bar{\Gamma}_c^2 \in \Gamma_c^2$), i.e. on the surfaces where the stress vector is prescribed (Γ_σ) and on the surfaces where the displacement is prescribed (Γ_u). Therefore, the virtual work associated to the surface forces acting on the bodies can be expressed as:

$$\int_{\partial\Omega} (\boldsymbol{\sigma}\mathbf{v}) \cdot \delta\mathbf{u} d\Gamma = \int_{\bar{\Gamma}_c^1} (\boldsymbol{\sigma}\mathbf{v}) \cdot \delta\mathbf{x}^s d\bar{\Gamma}_c^1 + \int_{\bar{\Gamma}_c^2} (\boldsymbol{\sigma}\mathbf{n}) \cdot \delta\bar{\mathbf{x}}^m d\bar{\Gamma}_c^2 + \int_{\Gamma_\sigma} \bar{\mathbf{t}} \cdot \delta\mathbf{u} d\Gamma_\sigma, \tag{2.53}$$

where $\bar{\mathbf{t}}$ is a prescribed traction (Neumann boundary conditions), $\mathbf{v}$ is the slave surface normal vector at $\mathbf{x}^s \in \bar{\Gamma}_c^1$ and $\mathbf{n}$ is the master surface normal vector at $\bar{\mathbf{x}}^m \in \bar{\Gamma}_c^2$. Since the displacements are prescribed in the Dirichlet boundaries, $\delta\mathbf{u} = 0$ at Γ_u, this term does not appears in the virtual work form presented in (2.53). This expression for virtual work due to the contact forces includes two integrals, one over each contact surface. Nevertheless,

due to the balance of linear momentum at the contact interface, the contact traction vector satisfies the action–reaction principle in the contact point, given by:

$$\mathbf{t}^{\mathrm{s}} d\bar{\Gamma}_{\mathrm{c}}^{1} = (\boldsymbol{\sigma}\boldsymbol{\nu}) d\bar{\Gamma}_{\mathrm{c}}^{1} = -(\boldsymbol{\sigma}\mathbf{n}) d\bar{\Gamma}_{\mathrm{c}}^{2} = -\mathbf{t}^{\mathrm{m}} d\bar{\Gamma}_{\mathrm{c}}^{2}, \tag{2.54}$$

which establishes that the differential contact force applied to the slave surface $\mathbf{t}^{\mathrm{s}} d\bar{\Gamma}_{\mathrm{c}}^{1}$ is equal and opposite to the one produced on the master surface $\mathbf{t}^{\mathrm{m}} d\bar{\Gamma}_{\mathrm{c}}^{2}$. Therefore, the two integrals on the contact surfaces can be replaced by a single integral over the slave surface, expressed by:

$$\int_{\bar{\Gamma}_{\mathrm{c}}^{1}} (\boldsymbol{\sigma}\boldsymbol{\nu}) \cdot \delta\mathbf{x}^{\mathrm{s}} d\bar{\Gamma}_{\mathrm{c}}^{1} + \int_{\bar{\Gamma}_{\mathrm{c}}^{2}} (\boldsymbol{\sigma}\mathbf{n}) \cdot \delta\bar{\mathbf{x}}^{\mathrm{m}} d\bar{\Gamma}_{\mathrm{c}}^{2} = \int_{\bar{\Gamma}_{\mathrm{c}}^{1}} (\boldsymbol{\sigma}\boldsymbol{\nu}) \cdot \delta(\mathbf{x}^{\mathrm{s}} - \bar{\mathbf{x}}^{\mathrm{m}}) d\bar{\Gamma}_{\mathrm{c}}^{1}, \tag{2.55}$$

where the surface stress vector acting on the slave surface can be decomposed into normal and tangential components (see (2.41) and (2.42)) to obtain:

$$\int_{\bar{\Gamma}_{\mathrm{c}}^{1}} (\boldsymbol{\sigma}\boldsymbol{\nu}) \cdot \delta(\mathbf{x}^{\mathrm{s}} - \bar{\mathbf{x}}^{\mathrm{m}}) d\bar{\Gamma}_{\mathrm{c}}^{1} = \int_{\bar{\Gamma}_{\mathrm{c}}^{1}} \mathbf{t}^{\mathrm{s}} \cdot \delta(\mathbf{x}^{\mathrm{s}} - \bar{\mathbf{x}}^{\mathrm{m}}) d\bar{\Gamma}_{\mathrm{c}}^{1} = \int_{\bar{\Gamma}_{\mathrm{c}}^{1}} (p_{\mathrm{n}}\mathbf{n} + \mathbf{t}_{\mathrm{t}}^{\mathrm{s}}) \cdot \delta(\mathbf{x}^{\mathrm{s}} - \bar{\mathbf{x}}^{\mathrm{m}}) d\bar{\Gamma}_{\mathrm{c}}^{1}. \tag{2.56}$$

This expression for the contact contribution to the virtual work balance of the system can be expressed considering the first variations of the contact kinematic variable defined in (2.43) and (2.44). Therefore, the virtual work balance for the frictional contact problem of the two body system may be summarized as:

$$\int_{\Omega} \boldsymbol{\sigma} : \delta\nabla\mathbf{u} d\Omega - \int_{\Omega} \mathbf{f}_{\mathrm{v}} \cdot \delta\mathbf{u} d\Omega - \int_{\Gamma_{\sigma}} \bar{\mathbf{t}} \cdot \delta\mathbf{u} d\Gamma_{\sigma} + \int_{\bar{\Gamma}_{\mathrm{c}}^{1}} (p_{\mathrm{n}} \delta g_{\mathrm{n}} + \mathbf{t}_{\mathrm{t}}^{\mathrm{m}} \cdot \delta\mathbf{g}_{\mathrm{t}}) d\bar{\Gamma}_{\mathrm{c}}^{1} = 0, \tag{2.57}$$

where the contact pressures p_{n} and the frictional traction vector $\mathbf{t}_{\mathrm{t}}^{\mathrm{m}} = -\mathbf{t}_{\mathrm{t}}^{\mathrm{s}}$ have definitions that depend on the regularisation method adopted, as discussed in the next section. Since when a contact point slides, the movement is in a direction opposite to the arising frictional force, the friction process is dissipative (virtual work of frictional forces presents the same sign (positive) as the virtual work of internal forces) and hence the solution becomes path-dependent. On the other hand, the normal contact contribution is negative because the contact pressure is assumed to be non-positive $p_{\mathrm{n}} \leq 0$ and δg_{n}.

2.3. Contact constraint enforcement methods

This section describes different procedures that can be applied to incorporate the contact constraints into the variational principle. The formulation of a variational principle for contact problems can results in a variational inequality subjected to geometrical constraints, leading to a nonlinear optimization problem under constraints defined as

inequalities [Kikuchi 88], [Duvaut 76], [Barboteu 03]. However, assuming that the active contact zone is known, the variational inequality can be replaced by a variational equality (2.57), resulting in an unconstrained or partly unconstrained optimization problem [Wriggers 06]. This allows applying in contact mechanics methods well known from optimization theory [Luenberger 08]. Due to the requirement of a known contact zone, the classical optimization method has to be considered in combination with an active set strategy [Luenberger 08]. The definition of the active contact zone is dictated by the contact constraints given as inequalities, which should be checked and updated in each solution step. The existence and uniqueness of solution for the small deformation frictionless contact problem, using variational inequalities is discussed in [Kikuchi 88]. Besides, for elastostatic problems, the solution of variational inequality is given by the minimum of total potential energy function subject to contact constraints. Examples highlighting the non-uniqueness of the solution or non-existence are discussed in [Klarbring 90] and [Martins 94].

For a *quasi*-static motion the equilibrium equation of a conservative system follows from the principle of stationary potential energy, dictating that at equilibrium the system has stationary potential energy. Therefore, the variational equality (2.57) accompanied by restrictions on the virtual displacements (2.22) and subject to the contact constraints (2.48) and (2.50) can be expressed simply as:

$$\delta\Pi^{\text{int}}(\mathbf{u}) - \delta\Pi^{\text{ext}}(\mathbf{u}) + \delta\Pi^{\text{contact}}(\mathbf{u},\boldsymbol{\lambda}) = 0, \tag{2.58}$$

where $\delta\Pi^{\text{int}}$ denotes the virtual work due to the internal stresses and $\delta\Pi^{\text{ext}}$ indicates the virtual work due to applied external loads, while the notation $\delta\Pi^{\text{contact}}$ represents the virtual work due to the contact forces on the active contact zones. From an optimization perspective, the contact contributions are indicated generically in a system functional, which is defined for the two body contact problem as:

$$\Pi^{\text{int}}(\mathbf{u}) - \Pi^{\text{ext}}(\mathbf{u}) + \Pi^{\text{contact}}(\mathbf{u},\boldsymbol{\lambda}), \tag{2.59}$$

where Π^{contact} indicates the contact contribution to the global functional, which is defined according to the procedure used for the enforcement of contact constraints. Depending on the method adopted to enforce the contact constraints, the contact contribution can include the dependent variable $\boldsymbol{\lambda}$, which physically represents the contact force. The enforcement of the variational equality (2.57) corresponds to a minimization of the total potential energy (2.59), leading to the equilibrium configuration of the contact problem.

Among the most popular and widely used resolution methods in contact mechanics are those inspired from optimization theory:

- The penalty method;
- The barrier method;
- The Lagrange multiplier method;

- The perturbed Lagrangian method;
- The augmented Lagrangian method;
- The Nitsche method;

and others, which can be found in [Wriggers 06] and references in it. In this thesis the attention is restricted to the three most important ones: penalty, Lagrange multipliers and augmented Lagrangian methods. Each one presents its own advantages and drawbacks, which are discussed in detail in the following.

2.3.1. Penalty method

The penalty method is one of the most widely used methods for treating contact problems, both in commercial and scientific finite element codes [Yastrebov 13]. The basic idea behind this method is to remove the contact constraints explicitly from the variational formulation by means of a penalization of the constraint violations. The magnitude of the penalization increases according to how severely the constraint is violated [Kikuchi 88]. Therefore, it is known to be simple and can be physical interpreted as a series of springs in the contact interface with zero initial length, as represented in Figure 2.13.

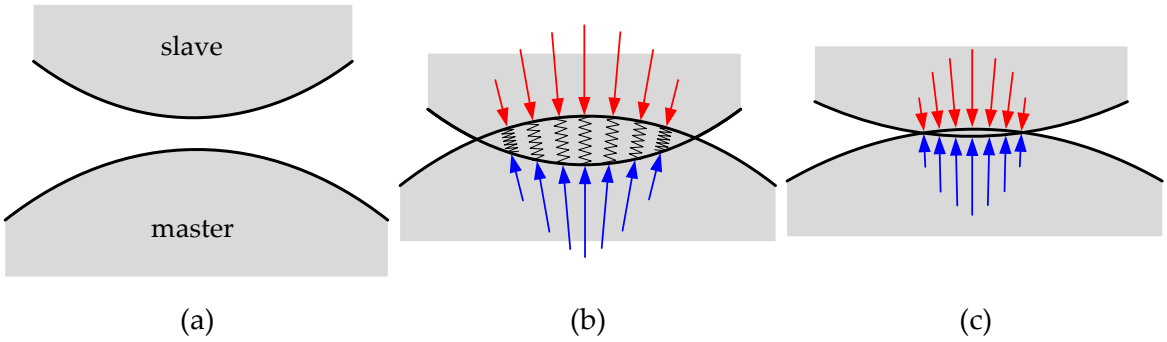


Figure 2.13. Physical interpretation of the penalty method: (a) initial configuration; (b) configuration after penetration; (c) equilibrium state [Yastrebov 13].

Since in the penalty method the contact constraints are explicitly removed from the variational formulation, the contact problem is formulated as an unconstrained optimization problem, where the unknowns are only the displacement variables. This particular advantage is replicated in the contact contribution of the penalty method to the system functional (2.59), defined for the frictionless contact problem as follows:

$$\Pi^{\mathrm{PM}}(\mathbf{u}) = \frac{1}{2}\int_{\Gamma_c^1} \varepsilon_{\mathrm{n}} \langle -g_{\mathrm{n}} \rangle^2 d\Gamma_c^1, \tag{2.60}$$

where $\varepsilon_{\mathrm{n}} > 0$ is the penalty parameter and the notation $\langle \bullet \rangle$ is used to denote the Macaulay bracket, which simply gives the positive part of its operand, being defined as:

$$\langle x \rangle := \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}. \tag{2.61}$$

Note that the energy of the penalty term in (2.60) exhibits the same structure as the potential energy of a simple spring, so the penalty parameter ε_n can be interpreted as the spring stiffness in the contact interface. Moreover, due to the definition of the Macaulay brackets, the penalty term (2.60) is positive only in the active contact zone, which is defined through the normal gap function in (2.49). Thus, the integral concerning the contact contribution of the penalty method (2.60) can be defined over the full potential contact zone.

The solution of the contact problem is obtained by minimizing (2.59), making the functional stationary with respect to variations of displacements $\mathbf{u}$. For the frictionless contact problem, the variation of (2.60) yields:

$$\delta\Pi^{PM}(\mathbf{u}) = \int_{\Gamma_c^1} -\varepsilon_n \langle -g_n \rangle \delta g_n d\Gamma_c^1, \tag{2.62}$$

which is the contact contribution to the balance of virtual work (compare with the last term in (2.57)). The penalty regularization dictates that the non-penetration condition stated in (2.48) is only approximately fulfilled (see Figure 2.13). The contact pressure is assumed to be a continuous function dependent of the penetration, given by:

$$p_n(g_n) = -\varepsilon_n \langle -g_n \rangle = \begin{cases} \varepsilon_n g_n & g_n \leq 0 \\ 0 & g_n > 0 \end{cases}, \tag{2.63}$$

which is contained within the frictionless contact integral (2.62), describing the work of the contact pressure on virtual penetration. The regularization of the unilateral contact law is illustrated schematically in Figure 2.14 (a), where the contact condition is strictly fulfilled for non-negative gaps (compare with Figure 2.11). On the other hand, the linear relation between contact pressure and the gap function dictates that the contact arises only for negative gap values, allowing penetration of the bodies. Therefore, the impenetrability condition (2.48) is only perfectly represented using an infinite value for the penalty parameter, which is a practical impossibility in real calculations. In fact, large numbers for the penalty parameter will lead to an ill-conditioned numerical problem, i.e. high condition number of the matrix that is obtained in the numerical solution of the problem. Nevertheless, the penalty method is widely used in practice, typically due to its easy implementation.

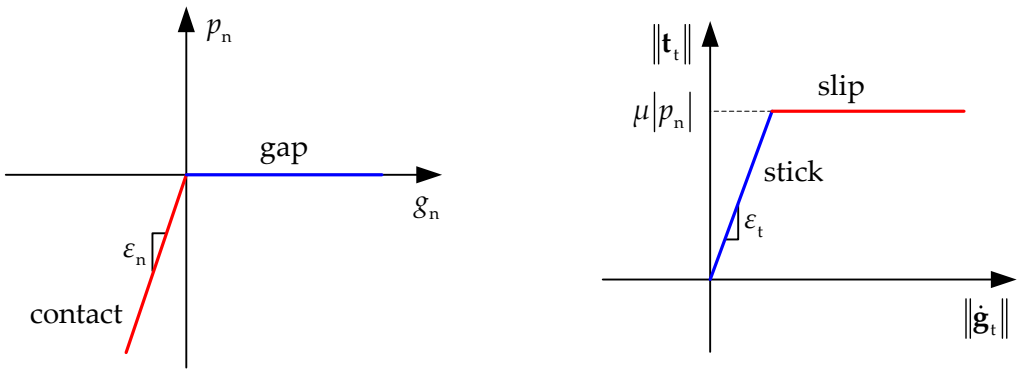


Figure 2.14. Application of the penalty method to the frictional contact problem: (a) regularized unilateral contact law; (b) regularized Coulomb's friction law.

The introduction of friction into the contact problem is accomplished by a new frictional term in the virtual work due to the contact (2.57). Such as in the case of the unilateral contact law, the classical Coulomb's friction law (2.50) can be approximately fulfilled using a penalty function. Then, the regularization of the Coulomb friction law considers the frictional force as a function of the tangential sliding limited by the Coulomb's cone surface, such as:

$$\mathbf{t}_t = \begin{cases} \varepsilon_t \mathbf{g}_t & \text{if } \varepsilon_t \|\mathbf{g}_t\| < \mu |p_n| \quad \textit{stick} \\ \mu |p_n| \dfrac{\mathbf{g}_t}{\|\mathbf{g}_t\|} & \text{if } \varepsilon_t \|\mathbf{g}_t\| \geq \mu |p_n| \quad \textit{slip} \end{cases}, \tag{2.64}$$

where ε_t denotes the tangential penalty parameter, which is not necessarily equal to the normal penalty parameter ε_n. The undesirable (from a numerical point of view) multivalued nature of the Coulomb friction law (see Figure 2.12) is removed through the penalty regularization, providing the frictional force single-valued function of the tangential displacement, as shown schematically in Figure 2.14 (b). The perfect representation of the Coulomb law is only recovered using an infinite value for the penalty parameter, as for the unilateral contact law (see (2.63)).

The penalty method allows some tangential movement at the contact interface for the stick status (see Figure 2.14 (b)), situation prohibited in the unregularized Coulomb law (2.50). Therefore, if the selected penalty parameter value is too low, this can produce non-physical behaviour at the interface. Nevertheless, there is significant experimental evidence indicating that the sharp transition between slick and slip friction (Figure 2.12 (a)) does not exist in many systems due to elastic deformation of the contact surface asperities [Courtney-Pratt 57]. Hence, in some cases the penalty regularization can be used with the purpose to better model the friction behaviour [Laursen 02].

Since the penalty term is only added for active constraints, normal and tangential conditions are formulated separately, distinguishing between stick and slip conditions. In the case of slip status the frictional force is evaluated through the contact pressure and the slip direction. Nevertheless, when only stick status occurs in the contact interface there is no need to distinguish between the normal and tangential directions, dictating equal penalty parameters for both directions [Wriggers 06]. Therefore, the variational equality for frictional contact using the penalty method is generally defined as:

$$\begin{aligned}\delta\Pi^{\text{int}}(\mathbf{u})-\delta\Pi^{\text{ext}}(\mathbf{u})+\int_{\bar{\Gamma}_c^{1*}}-\varepsilon_n\langle -g_n\rangle(\delta g_n-\mu\frac{\mathbf{g}_t}{\|\mathbf{g}_t\|}\cdot\delta\mathbf{g}_t)d\bar{\Gamma}_c^1+\\+\int_{\bar{\Gamma}_c^{1\bullet}}-\varepsilon_n\langle -g_n\rangle\delta g_n+\varepsilon_t\mathbf{g}_t\cdot\delta\mathbf{g}_t d\bar{\Gamma}_c^1=0,\end{aligned} \tag{2.65}$$

where $\bar{\Gamma}_c^{1*}$ and $\bar{\Gamma}_c^{1\bullet}$ are the slip and the stick active contact zones on the master surface, respectively.

2.3.2. Lagrange multiplier method

The Lagrange multiplier method can also be used to include constraints in the balance of virtual work (2.57), producing a variational equality for the contact problem. This method is commonly used in optimization problems to find the maximum/minimum of a functional subjected to equality constraints. The idea behind the Lagrange multiplier method is to introduce a vector of additional unknowns $\boldsymbol{\lambda}$, called Lagrange multipliers, replacing the original constrained optimization problem by a saddle point of the functional, called Lagrangian $\mathcal{L}(\mathbf{u},\boldsymbol{\lambda})$.

Regarding the frictionless contact problem, taking into account the inequality constraints (2.48), the constrained minimization contact problem can be formulated as:

$$\min_{\mathbf{u}\in\boldsymbol{\mathcal{V}},g_n\geq 0}\left\{\Pi^{\text{int,ext}}(\mathbf{u})\right\}\quad\Leftrightarrow\quad\exists\mathbf{u}^*:\quad\forall\mathbf{u}\in\boldsymbol{\mathcal{V}},g_n(\mathbf{u})\geq 0\text{ on }\bar{\Gamma}_c:\Pi^{\text{int,ext}}(\mathbf{u}^*)\leq\Pi^{\text{int,ext}}(\mathbf{u})\quad\text{in }\Omega, \tag{2.66}$$

where $\Pi^{\text{int,ext}}$ denotes the total energy of the mechanical system (sum of internal the external energy). This problem, subject to inequalities related with the unilateral contact law, can be replaced by a stationary point problem of the following Lagrangian:

$$\mathcal{L}(\mathbf{u},\lambda_n)=\Pi^{\text{int}}(\mathbf{u})-\Pi^{\text{ext}}(\mathbf{u})+\int_{\bar{\Gamma}_c^1}\lambda_n g_n(\mathbf{u})d\bar{\Gamma}_c^1, \tag{2.67}$$

where λ_n represents a continuous set of values (Lagrange multipliers) on the active contact zone, which is interpreted as the contact pressure on the interface, necessary to prevent interpenetration of the bodies. The stationary condition for the Lagrangian can be formulated as the variation of the Lagrangian (2.67), defined by:

$$\delta\mathcal{L}(\mathbf{u},\lambda_{\mathrm{n}}) = \delta\Pi^{\mathrm{int}}(\mathbf{u}) - \delta\Pi^{\mathrm{ext}}(\mathbf{u}) + \int_{\bar{\Gamma}_{\mathrm{c}}^{1}} g_{\mathrm{n}}(\mathbf{u})\delta\lambda_{\mathrm{n}} + \lambda_{\mathrm{n}}\delta g_{\mathrm{n}}(\mathbf{u}) d\bar{\Gamma}_{\mathrm{c}}^{1} = 0, \tag{2.68}$$

where the first term in the integral describes the enforcement of the constraints, while the second term is associated with the virtual work of the Lagrange multipliers along the variation of the gap function [Kikuchi 88]. Indeed, for zero gap value the integral due to contact in (2.68) is quite similar to the frictionless part of the contact integral in (2.57). Since a new unknown (Lagrange multiplier) is included in the Lagrangian, the solution of the mixed variational formulation $(\mathbf{u}^{*},\lambda_{\mathrm{n}}^{*})$ is obtained such that:

$$\forall \mathbf{u} \in \boldsymbol{\mathcal{V}}, \lambda_{\mathrm{n}} \leq 0: \quad \mathcal{L}(\mathbf{u}^{*},\lambda_{\mathrm{n}}) \leq \mathcal{L}(\mathbf{u}^{*},\lambda_{\mathrm{n}}^{*}) \leq \mathcal{L}(\mathbf{u},\lambda_{\mathrm{n}}^{*}), \tag{2.69}$$

which can be shortly formulated as an unconstrained saddle point problem, denoted as:

$$\min_{\mathbf{u}\in\boldsymbol{\mathcal{V}}} \max_{\lambda_{\mathrm{n}}\leq 0} \mathcal{L}(\mathbf{u},\lambda_{\mathrm{n}}), \tag{2.70}$$

meaning that the solution point corresponds to a minimum of $\mathcal{L}(\mathbf{u},\lambda_{\mathrm{n}})$ with respect to the displacements $\mathbf{u}$ and a maximum of $\mathcal{L}(\mathbf{u},\lambda_{\mathrm{n}})$ with respect to the Lagrange multipliers. Note that the Lagrange multiplier method does not convert a minimization problem with inequality constraints to a fully unconstrained one because the constraint $\lambda_{\mathrm{n}} \leq 0$ has still to be fulfilled [Yastrebov 13]. The Lagrange multiplier method allows an exact fulfilment of the contact constrains (2.48) by the introduction of additional degrees of freedom (Lagrange multipliers), which dictates an increase of the computational effort in comparison with the penalty method.

In case of frictional contact it is convenient to separate the stick and slip contact conditions. As for the penalty method, in case of stick status there is no need to distinguish tangential and normal directions [Wriggers 06]. In that case, the Lagrangian of the energy balance becomes:

$$\mathcal{L}(\mathbf{u},\boldsymbol{\lambda}) = \Pi^{\mathrm{int}}(\mathbf{u}) - \Pi^{\mathrm{ext}}(\mathbf{u}) + \int_{\bar{\Gamma}_{\mathrm{c}}^{1\bullet}} \boldsymbol{\lambda}\cdot\mathbf{g}(\mathbf{u}) d\bar{\Gamma}_{\mathrm{c}}^{1}, \tag{2.71}$$

where the Lagrange multiplier vector $\boldsymbol{\lambda}$ represents the contact force vector. The equilibrium of the two body system is defined by the variation on the Lagrangian (2.71), which takes the following form:

$$\delta\mathcal{L}(\mathbf{u},\boldsymbol{\lambda}) = \delta\Pi^{\mathrm{int}}(\mathbf{u}) - \delta\Pi^{\mathrm{ext}}(\mathbf{u}) + \int_{\bar{\Gamma}_{\mathrm{c}}^{1\bullet}} \mathbf{g}(\mathbf{u})\cdot\delta\boldsymbol{\lambda} + \boldsymbol{\lambda}\cdot\delta\mathbf{g}(\mathbf{u}) d\bar{\Gamma}_{\mathrm{c}}^{1} = 0. \tag{2.72}$$

In case of slip status, considering the Coulomb's friction law, the frictional force is completely expressed by the contact pressure and the sliding direction (see (2.50)).

Therefore, its corresponding Lagrange multiplier vector $\boldsymbol{\lambda}_{\mathrm{t}}$ is not required, contrarily to the case of stick status where it is included in the contact force vector $\boldsymbol{\lambda}$. Hence, the virtual work balance using the Lagrange multiplier method for the slip status takes the form:

$$\delta\mathcal{L}(\mathbf{u},\lambda_{\mathrm{n}},\boldsymbol{\lambda}_{\mathrm{t}})=\delta\Pi^{\mathrm{int}}(\mathbf{u})-\delta\Pi^{\mathrm{ext}}(\mathbf{u})+\int_{\bar{\Gamma}_{\mathrm{c}}^{1*}} g_{\mathrm{n}}(\mathbf{u})\delta\lambda_{\mathrm{n}}+\lambda_{\mathrm{n}}\delta g_{\mathrm{n}}(\mathbf{u})+\mu\left|\lambda_{\mathrm{n}}\right|\frac{\dot{\mathbf{g}}_{\mathrm{t}}}{\left\|\dot{\mathbf{g}}_{\mathrm{t}}\right\|}\cdot\delta\mathbf{g}_{\mathrm{t}}(\mathbf{u})d\bar{\Gamma}_{\mathrm{c}}^{1}, \quad (2.73)$$

which becomes (2.68) in case of null friction coefficient. Finally, the variational equality of frictional contact problem defined through the Lagrange multiplier method for the active contact zone is expressed as:

$$\begin{aligned}\delta\mathcal{L}(\mathbf{u},\lambda_{\mathrm{n}},\boldsymbol{\lambda}_{\mathrm{t}})=\delta\Pi^{\mathrm{int}}(\mathbf{u})-\delta\Pi^{\mathrm{ext}}(\mathbf{u})+\int_{\bar{\Gamma}_{\mathrm{c}}^{1\bullet}}\mathbf{g}(\mathbf{u})\cdot\delta\boldsymbol{\lambda}+\boldsymbol{\lambda}\cdot\delta\mathbf{g}(\mathbf{u})d\bar{\Gamma}_{\mathrm{c}}^{1}+\\ +\int_{\bar{\Gamma}_{\mathrm{c}}^{1*}} g_{\mathrm{n}}(\mathbf{u})\delta\lambda_{\mathrm{n}}+\lambda_{\mathrm{n}}\delta g_{\mathrm{n}}(\mathbf{u})+\mu\left|\lambda_{\mathrm{n}}\right|\frac{\dot{\mathbf{g}}_{\mathrm{t}}}{\left\|\dot{\mathbf{g}}_{\mathrm{t}}\right\|}\cdot\delta\mathbf{g}_{\mathrm{t}}(\mathbf{u})d\bar{\Gamma}_{\mathrm{c}}^{1}=0,\end{aligned} \quad (2.74)$$

where the constraint $\lambda_{\mathrm{n}}\leq 0$ is still remaining. Due to this inequality constraint contained in the Lagrange multiplier method, it should be coupled with an active set strategy [Luenberger 08]. In fact, the contact virtual work defined in (2.74) is only integrated in the points where the contact constraint is active, which are in general unknown *a priori*. Therefore, the solution of (2.74) involves an iterative procedure to determine the active contact zone [Laursen 02].

2.3.3. Augmented Lagrangian method

Another approach widely used to regularize the non-differentiable normal contact (2.48) and friction terms (2.50) is provided by the augmented Lagrangian approach, originally proposed by Hestenes [Hestenes 69] and Powell [Powell 69] to solve constrained optimization problems. The application of this method to frictional contact problems involving large displacements was firstly presented by Alart and Curnier [Alart 91]. The main idea of this approach is to combine the advantages of both the penalty and the Lagrange multiplier methods and avoid their drawbacks, allowing an exact enforcement of contact constraints for a finite value of the penalty parameter. The augmented Lagrangian method can be expressed as a Lagrange multiplier formulation regularized by penalty functions [Yastrebov 13]. It yields a C^1-differentiable energy functional, which is described in detail in [Pietrzak 97], being the saddle point solution problem (minimize primal variables and maximize dual variables) fully unconstrained. This particular advantage is not verified in the Lagrange multiplier method since the condition of non-positivity of the Lagrange multipliers must be satisfied, making the augmented Lagrangian method better for practical application [Pietrzak 99].

The application of the augmented Lagrangian method to deal with frictionless contact problems, which are subject to inequality constraints (2.48), leads to the following unconstrained functional:

$$\mathcal{L}^{a}(\mathbf{u},\lambda_{n}) = \Pi^{int}(\mathbf{u}) - \Pi^{ext}(\mathbf{u}) + \int_{\Gamma_c^1} \frac{1}{2\varepsilon_n} \langle -(\lambda_n + \varepsilon_n g_n(\mathbf{u})) \rangle^2 - \frac{1}{2\varepsilon_n} \lambda_n{}^2 d\Gamma_c^1, \tag{2.75}$$

which can be denoted in the expanded form, using Macaulay bracket definition (2.61), as:

$$\mathcal{L}^{a}(\mathbf{u},\lambda_{n}) = \Pi^{int}(\mathbf{u}) - \Pi^{ext}(\mathbf{u}) + \begin{cases} \int_{\bar{\Gamma}_c^1} \lambda_n g_n(\mathbf{u}) + \frac{1}{2}\varepsilon_n g_n{}^2(\mathbf{u}) d\Gamma_c^1, & \lambda_n + \varepsilon_n g_n(\mathbf{u}) \le 0 \\ \int_{\Gamma_c^1 \setminus \bar{\Gamma}_c^1} -\frac{1}{2\varepsilon_n} \lambda_n{}^2 d\Gamma_c^1, & \lambda_n + \varepsilon_n g_n(\mathbf{u}) > 0 \end{cases} \tag{2.76}$$

Note that the Lagrange multiplier λ_n is not restricted to be non-positive as in case of the Lagrange multiplier method. Comparing (2.76) with (2.60) and (2.67) becomes clear that the Augmented Lagrangian approach is a combination of the Lagrange multiplier and the penalty method, containing both a Lagrange multiplier and a penalty contribution. Moreover, in the case of an exact satisfaction of the contact constraints the penalty term vanishes, leading to the same solution as its Lagrange multiplier counterpart. On the other hand, when λ_n is set to zero, the penalty functional in (2.60) is recovered.

The implementation of the augmented Lagrangian method to solve contact problems can be carried out in two different ways. The first approach considers the augmented Lagrangian method with Uzawa's algorithm, referred as nested augmented Lagrangian algorithm, which was firstly reported by Simo and Laursen in [Simo 92]. In that case, the value of the Lagrange multipliers is assumed given, being removed from the unknowns. Therefore, the system to be solved is composed only by displacements, presenting the same dimension as the problem without contact. Nevertheless, this leads to a double loop algorithm, where the Lagrange multiplier is held constant only during an iteration loop to solve the weak form in the inner loop, where Lagrange multiplier is updated in the outer loop. The rate of convergence of this method, where primal and dual variables are updated independently, is linear [Powell 69]. The advantage is that the resulting functional is smooth enough to apply a standard Newton's technique, in order to obtain the solution problem. The second approach is the coupled augmented Lagrangian method, which has been undertaken by Alart and Curnier [Alart 91]. In that case the Lagrange multipliers are retained as independent variables in the coupled problem, increasing the computational effort. Similar to the Lagrange multiplier method, a vector of additional unknowns is introduced, which physically represent the contact forces. Then, a min-max problem is obtained, for which the solution is attained by minimizing the functional primal variables and maximizing the dual ones. However, the new functional is not sufficiently smooth to

apply a standard Newton's technique, being proposed a generalized Newton method to ensure the convergence within the non-smooth potential [Alart 91] [Alart 97].

The application of the augmented Lagrangian method to solve frictional contact problems is given in detail by Pietrzak in [Pietrzak 97] and [Pietrzak 99]. The formulation is based in the formalism used by Alart and Curnier [Alart 91] and Heegaard and Curnier [Heegaard 93], which follows the Moreau's convex analysis. Following the cited authors, the Hertz–Signorini–Moreau conditions for frictionless contact (2.48) can be rewritten as a single sub-differential inclusion:

$$p_{\mathrm{n}} \in \partial\psi^{+}(g_{\mathrm{n}}), \tag{2.77}$$

where ψ^{+} denotes the indicator function of the positive half-line and $\partial\psi^{+}$ represents its sub-differential [Heegaard 93]. Similarly, the conditions arising from the Coulomb's friction law (2.50) can be reformulated as a sub-differential inclusion, being the frictional contact force vector expressed as:

$$\mathbf{t}_{\mathrm{t}} \in \partial\psi^{*}_{C(p_{\mathrm{n}})}(\dot{\mathbf{g}}_{\mathrm{t}}), \tag{2.78}$$

where $\psi^{*}_{C(p_{\mathrm{n}})}$ denotes the conjugate function of the disk indicator function $\psi_{C(p_{\mathrm{n}})}$, while $C(p_{\mathrm{n}})$ represents the convex disk of radius $\mu\left|p_{\mathrm{n}}\right|$ centred at the origin (section of the Coulomb's cone), which is function of the unknown contact pressure. Besides, in the incremental *quasi*-static analysis the tangential relative sliding velocity in (2.78) is replaced by increments. The variational problem stated in (2.59) can be formulated including the framework of sub-differential inclusions, being the frictional contact contribution defined by:

$$\Pi^{\mathrm{contact}}(\mathbf{u},\boldsymbol{\lambda}) = \int_{\Gamma_{\mathrm{c}}^{1}} \left\{ \psi^{+}(g_{\mathrm{n}}) + \psi^{*}_{C(p_{\mathrm{n}})}(\dot{\mathbf{g}}_{\mathrm{t}}) \right\} d\Gamma_{\mathrm{c}}^{1}, \tag{2.79}$$

which is a non-differentiable energy due to contact interaction. The regularization of the indicator functions (also called *quasi*-potentials) included in (2.79) by the augmented Lagrangian method is presented in [Pietrzak 97]. Following that study, the augmented Lagrangian functional is constructed from (2.59) using the regularized frictional contact form (2.79), being expressed as:

$$\mathcal{L}^{\mathrm{a}}(\mathbf{u},\lambda_{\mathrm{n}},\boldsymbol{\lambda}_{\mathrm{t}},\hat{p}_{\mathrm{n}}) = \Pi^{\mathrm{int}}(\mathbf{u}) - \Pi^{\mathrm{ext}}(\mathbf{u}) + \int_{\Gamma_{\mathrm{c}}^{1}} l_{\mathrm{n}}(g_{\mathrm{n}},\lambda_{\mathrm{n}}) + l_{\mathrm{t}}(\mathbf{g}_{\mathrm{t}},\boldsymbol{\lambda}_{\mathrm{t}},\hat{p}_{\mathrm{n}}) d\Gamma_{\mathrm{c}}^{1}, \tag{2.80}$$

where l_{n} and l_{t} denotes the augmented Lagrangians related to normal and frictional contact, which represent the regularized functions (2.77) and (2.78), respectively. As in the Lagrange multiplier method, λ_{n} and $\boldsymbol{\lambda}_{\mathrm{t}}$ are the Lagrange multipliers representing the

contact pressure and the frictional contact force vector, respectively. Note that the contact pressure at solution is required in the definition of the augmented Lagrangian l_t, due to the non-associated character of the Coulomb's friction law (the contact pressure is independent of the friction force but the friction force depends on the contact pressure). Hence, $\hat{p}_n = p_n + \varepsilon_n g_n$ denotes a regularized contact pressure and at solution $\hat{p}_n = p_n$ holds.

In order to formulate the virtual work principle, the augmented Lagrangian functionals l_n and l_t need to be explicitly expressed. Following [Pietrzak 99], the augmented Lagrangian functional related with the unilateral contact law takes the following form:

$$l_n(g_n, \lambda_n) = \begin{cases} g_n \hat{\lambda}_n - \dfrac{\varepsilon_n}{2} g_n{}^2, & \hat{\lambda}_n \le 0, \quad \textit{contact} \\ -\dfrac{1}{2\varepsilon_n} \lambda_n{}^2, & \hat{\lambda}_n > 0, \quad \textit{non-contact} \end{cases}, \tag{2.81}$$

where the augmented Lagrange multiplier is denoted by a hat:

$$\hat{\lambda}_n = \lambda_n + \varepsilon_n g_n. \tag{2.82}$$

The augmented Lagrangian functional due to Coulomb's friction law is written as:

$$l_t(\mathbf{g}_t, \boldsymbol{\lambda}_t, \hat{p}_n) = \begin{cases} \hat{\boldsymbol{\lambda}}_t \cdot \mathbf{g}_t - \dfrac{\varepsilon_t}{2} \mathbf{g}_t \cdot \mathbf{g}_t, & \left\| \hat{\boldsymbol{\lambda}}_t \right\| \le -\mu \hat{p}_n, \quad \textit{stick} \\ -\dfrac{1}{2\varepsilon_t} (\boldsymbol{\lambda}_t \cdot \boldsymbol{\lambda}_t + 2\mu \hat{p}_n \left\| \hat{\boldsymbol{\lambda}}_t \right\| + \mu^2 \hat{p}_n{}^2), & \left\| \hat{\boldsymbol{\lambda}}_t \right\| > -\mu \hat{p}_n, \quad \textit{slip} \\ -\dfrac{1}{2\varepsilon_t} \boldsymbol{\lambda}_t \cdot \boldsymbol{\lambda}_t, & \hat{p}_n > 0, \quad \textit{non-contact} \end{cases} \tag{2.83}$$

where $\hat{p}_n$ is the regularized contact pressure at solution. Note that the tangential regularized functional is extended to the non-contact domain $\hat{p}_n > 0$, covering all possible contact statuses. This fact results in the prolongation of the Coulomb's cone for positive values of the normal contact pressure, as represented in Figure 2.12 (b) by a dashed red line. The augmented Lagrange multiplier used in (2.83) is defined as:

$$\hat{\boldsymbol{\lambda}}_t = \boldsymbol{\lambda}_t + \varepsilon_t \mathbf{g}_t. \tag{2.84}$$

The variation of the integral denoting the contact contribution to the energy of the system (2.80) leads to:

$$\begin{aligned}\delta\Pi^{\mathrm{ALM}}(\mathbf{u},\lambda_{\mathrm{n}},\boldsymbol{\lambda}_{\mathrm{t}},\hat{p}_{\mathrm{n}}) &= \int_{\Gamma_{\mathrm{c}}^{1}} \delta l_{\mathrm{n}}(g_{\mathrm{n}},\lambda_{\mathrm{n}}) + \delta l_{\mathrm{t}}(\mathbf{g}_{\mathrm{t}},\boldsymbol{\lambda}_{\mathrm{t}},\hat{p}_{\mathrm{n}})\, d\Gamma_{\mathrm{c}}^{1} \\ &= \int_{\Gamma_{\mathrm{c}}^{1}} \frac{\partial l_{\mathrm{n}}}{\partial g_{\mathrm{n}}}\delta g_{\mathrm{n}} + \frac{\partial l_{\mathrm{n}}}{\partial \lambda_{\mathrm{n}}}\delta\lambda_{\mathrm{n}} + \frac{\partial l_{\mathrm{t}}}{\partial \mathbf{g}_{\mathrm{t}}}\cdot\delta\mathbf{g}_{\mathrm{t}} + \frac{\partial l_{\mathrm{t}}}{\partial \boldsymbol{\lambda}_{\mathrm{t}}}\cdot\delta\boldsymbol{\lambda}_{\mathrm{t}}\, d\Gamma_{\mathrm{c}}^{1},\end{aligned} \tag{2.85}$$

where the contact pressure $\hat{p}_{\mathrm{n}}$ is not subjected to the variation since it is assumed to be the known contact pressure at solution. The integral over the master surface in (2.85) represents the augmented virtual work developed by contact and friction forces at the contact interface. In order to evaluate the derivatives contained in (2.85) it is useful to divide the contact zone into three (stick, slip and non-contact), being each derivative expressed by:

$$\frac{\partial l_{\mathrm{n}}(g_{\mathrm{n}},\lambda_{\mathrm{n}})}{\partial g_{\mathrm{n}}} = \begin{cases} \hat{\lambda}_{\mathrm{n}}, & \hat{\lambda}_{\mathrm{n}} \le 0, \quad \textit{contact} \\ 0, & \hat{\lambda}_{\mathrm{n}} > 0, \quad \textit{non-contact} \end{cases}, \tag{2.86}$$

$$\frac{\partial l_{\mathrm{n}}(g_{\mathrm{n}},\lambda_{\mathrm{n}})}{\partial \lambda_{\mathrm{n}}} = \begin{cases} g_{\mathrm{n}}, & \hat{\lambda}_{\mathrm{n}} \le 0, \quad \textit{contact} \\ -\dfrac{1}{\varepsilon_{\mathrm{n}}}\lambda_{\mathrm{n}}, & \hat{\lambda}_{\mathrm{n}} > 0, \quad \textit{non-contact} \end{cases}, \tag{2.87}$$

$$\frac{\partial l_{\mathrm{t}}(\mathbf{g}_{\mathrm{t}},\boldsymbol{\lambda}_{\mathrm{t}})}{\partial \mathbf{g}_{\mathrm{t}}} = \begin{cases} \hat{\boldsymbol{\lambda}}_{\mathrm{t}}, & \left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| \le -\mu\hat{p}_{\mathrm{n}}, \quad \textit{stick} \\ -\mu\hat{p}_{\mathrm{n}}\dfrac{\hat{\boldsymbol{\lambda}}_{\mathrm{t}}}{\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\|}, & \left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| > -\mu\hat{p}_{\mathrm{n}}, \quad \textit{slip} \\ 0, & \hat{p}_{\mathrm{n}} > 0, \quad \textit{non-contact} \end{cases}, \tag{2.88}$$

$$\frac{\partial l_{\mathrm{t}}(\mathbf{g}_{\mathrm{t}},\boldsymbol{\lambda}_{\mathrm{t}})}{\partial \boldsymbol{\lambda}_{\mathrm{t}}} = \begin{cases} \mathbf{g}_{\mathrm{t}}, & \left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| \le -\mu\hat{p}_{\mathrm{n}}, \quad \textit{stick} \\ -\dfrac{1}{\varepsilon_{\mathrm{t}}}\left(\boldsymbol{\lambda}_{\mathrm{t}} + \mu\hat{p}_{\mathrm{n}}\dfrac{\hat{\boldsymbol{\lambda}}_{\mathrm{t}}}{\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\|}\right), & \left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| > -\mu\hat{p}_{\mathrm{n}}, \quad \textit{slip} \\ -\dfrac{1}{\varepsilon_{\mathrm{t}}}\boldsymbol{\lambda}_{\mathrm{t}}, & \hat{p}_{\mathrm{n}} > 0, \quad \textit{non-contact} \end{cases}. \tag{2.89}$$

Considering the derivatives (2.86)–(2.89), the virtual work balance of the frictional contact problem, using the augmented Lagrangian method to incorporate the contact integral, takes the following form:

$$\delta\mathcal{L}^{\mathrm{a}}(\mathbf{u},\lambda_{\mathrm{n}},\boldsymbol{\lambda}_{\mathrm{t}},\hat{p}_{\mathrm{n}}) = \delta\Pi^{\mathrm{int}}(\mathbf{u}) - \delta\Pi^{\mathrm{ext}}(\mathbf{u}) + \int_{\bar{\Gamma}_{\mathrm{c}}^{1\bullet}} \hat{\lambda}_{\mathrm{n}}\delta g_{\mathrm{n}} + g_{\mathrm{n}}\delta\lambda_{\mathrm{n}} + \hat{\boldsymbol{\lambda}}_{\mathrm{t}}\cdot\delta\mathbf{g}_{\mathrm{t}} + \mathbf{g}_{\mathrm{t}}\cdot\delta\boldsymbol{\lambda}_{\mathrm{t}}\, d\Gamma_{\mathrm{c}}^{1} +$$
$$+ \int_{\bar{\Gamma}_{\mathrm{c}}^{1*}} \hat{\lambda}_{\mathrm{n}}\delta g_{\mathrm{n}} + g_{\mathrm{n}}\delta\lambda_{\mathrm{n}} - \mu\hat{p}_{\mathrm{n}}\frac{\hat{\boldsymbol{\lambda}}_{\mathrm{t}}}{\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\|}\cdot\delta\mathbf{g}_{\mathrm{t}} - \frac{1}{\varepsilon_{\mathrm{t}}}\left(\boldsymbol{\lambda}_{\mathrm{t}} + \mu\hat{p}_{\mathrm{n}}\frac{\hat{\boldsymbol{\lambda}}_{\mathrm{t}}}{\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\|}\right)\cdot\delta\boldsymbol{\lambda}_{\mathrm{t}}\, d\Gamma_{\mathrm{c}}^{1} + \tag{2.90}$$
$$+ \int_{\Gamma_{\mathrm{c}}^{1}\setminus\bar{\Gamma}_{\mathrm{c}}^{1}} -\frac{1}{\varepsilon_{\mathrm{n}}}\lambda_{\mathrm{n}}\delta\lambda_{\mathrm{n}} - \frac{1}{\varepsilon_{\mathrm{t}}}\boldsymbol{\lambda}_{\mathrm{t}}\cdot\delta\boldsymbol{\lambda}_{\mathrm{t}}\, d\Gamma_{\mathrm{c}}^{1} = 0,$$

where $\bar{\Gamma}_{\mathrm{c}}^{1\bullet}$ denoted the stick zone, $\bar{\Gamma}_{\mathrm{c}}^{1*}$ denoted the slip zone and $\bar{\Gamma}_{\mathrm{c}}^{1}$ represents the active contact zone, being a non-contact zone defined by $\Gamma_{\mathrm{c}}^{1}\setminus\bar{\Gamma}_{\mathrm{c}}^{1}$. Note that the only difference between the augmented Lagrangian virtual work principle (2.90) and the standard non-smooth virtual work principle (2.57) is in the contact terms. Indeed, the augmented Lagrangian regularization does not change the solid mechanics part of the virtual work principle, thus it is independent of the choice for the constitutive material law [Pietrzak 99]. The augmented Lagrangian method is the one selected throughout this thesis for the frictional constraint enforcement, presenting several advantages highlighted in the following simple example.

2.3.4. Example of a spring in contact with a rigid wall

The purpose of this section is to compare the previously presented resolution methods using a simple example of minimization under constraints. The penalty, Lagrange multipliers and augmented Lagrangian methods are compared with respect to the accuracy of the solutions and efficiency of their numerical implementation. The selected example involves a spring subjected to a prescribed displacement on its left end, while the right end is constrained by a rigid wall (example adapted from [Yastrebov 13] and [Laursen 02]). Figure 2.15 depicts both the initial and final configuration of the system under analysis, being the displacement u on the right end the unknown of the problem. The internal energy of the spring with one degree of freedom is defined by:

$$\Pi^{\mathrm{int}}(u) = \frac{1}{2}E(u-\bar{u})^2 = (u-2)^2, \tag{2.91}$$

where $E = 2\,\mathrm{N/mm}$ is the spring stiffness and $\bar{u} = 2\,\mathrm{mm}$ denotes the prescribed displacement on its left end. In the absence of obstacle, the equilibrium solution occurs when the energy is minimal or the variation of the energy is zero:

$$\min_{u}\left\{\Pi^{\mathrm{int}}(u)\right\} \quad \text{or} \quad \delta\Pi^{\mathrm{int}}(u) = 0, \tag{2.92}$$

where the variation of (2.91) leads to the following equation:

$$\delta\Pi^{\text{int}}(u) = 2(u-2)\delta u = 0 \quad \Leftrightarrow \quad u = 2\,\text{mm}. \tag{2.93}$$

The solution for the displacement is equal to the one imposed on the left end of the spring, i.e. rigid body motion occurs. Nevertheless, the introduction of the rigid wall restricts the displacement on the right end of the spring. The gap function g is given by:

$$g(u) = g_0 - u = 1 - u \geq 0 \quad \Leftrightarrow \quad u \leq 1\,\text{mm}, \tag{2.94}$$

where $g_0 = 1\,\text{mm}$ is initial distance of the spring right end to the wall (see Figure 2.15 (a)), which obviously restricts the solution domain of the contact problem.

(a) (b)

Figure 2.15. One degree of freedom contact problem example: (a) initial configuration; (b) deformed configuration due to the contact with a rigid wall.

2.3.4.1. Penalty method

The introduction of the rigid wall into the initial unconstrained minimization problem dictates a new solution (constrained), being the problem reformulated using the penalty method. In the framework of variational equalities and the penalty method, the constrained minimization problem can be rewritten as a simple minimization problem:

$$\min_{u \leq 1} \left\{ \Pi^{\text{int}}(u) \right\} \quad \rightarrow \quad \min_{u} \left\{ \Pi^{\text{int}}(u) + \Pi^{\text{PM}}(u) \right\}, \tag{2.95}$$

where Π^{PM} is the penalty term due to the violation of the contact constraints. Using the definition given in (2.60), the functional related with the contact is expressed for this example as:

$$\Pi^{\text{PM}}(u) = \frac{1}{2}\varepsilon\langle -g \rangle^2 = \frac{1}{2}\varepsilon\langle -(1-u) \rangle^2, \tag{2.96}$$

where ε is the penalty parameter and the gap function is defined in (2.94). Using the energy of the spring (2.91) together with the penalty term (2.96), the solution of the contact problem is expressed by the minimization problem:

$$\min_u \left\{ \Pi^{\text{int}}(u) + \Pi^{\text{PM}}(u) \right\} \quad \Leftrightarrow \quad \min_u \left\{ (u-2)^2 + \frac{1}{2}\varepsilon \langle -(1-u) \rangle^2 \right\}, \tag{2.97}$$

with solution given by:

$$\frac{\partial \left\{ \Pi^{\text{int}}(u) + \Pi^{\text{PM}}(u) \right\}}{\partial u} = 2(u-2) + \varepsilon(u-1) = 0 \quad \Leftrightarrow \quad u = \frac{4+\varepsilon}{2+\varepsilon}. \tag{2.98}$$

Since the functional (2.97) is valid for any displacement, the problem can be rewritten as a variational problem similar to (2.62), expressed as:

$$\delta\Pi^{\text{int}}(u) + \delta\Pi^{\text{PM}}(u) = 0, \tag{2.99}$$

where the required variations are evaluated from (2.91) and (2.96) using the relationship $\delta g = -\delta u$ (see (2.94)), being the solution given by:

$$2(u-2)\delta u - \varepsilon(1-u)(\delta u) = 0 \quad \Leftrightarrow \quad u = \frac{4+\varepsilon}{2+\varepsilon}, \tag{2.100}$$

where the obtained solution is the same that was calculated in (2.98), which is dependent of the value defined for the penalty parameter.

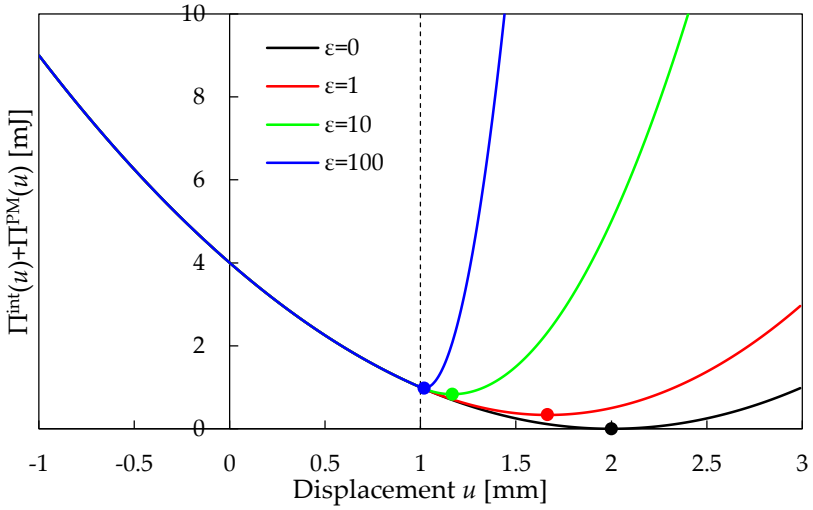


Figure 2.16. Penalized energy functional for different values of penalty parameter and corresponding solution (hollow points).

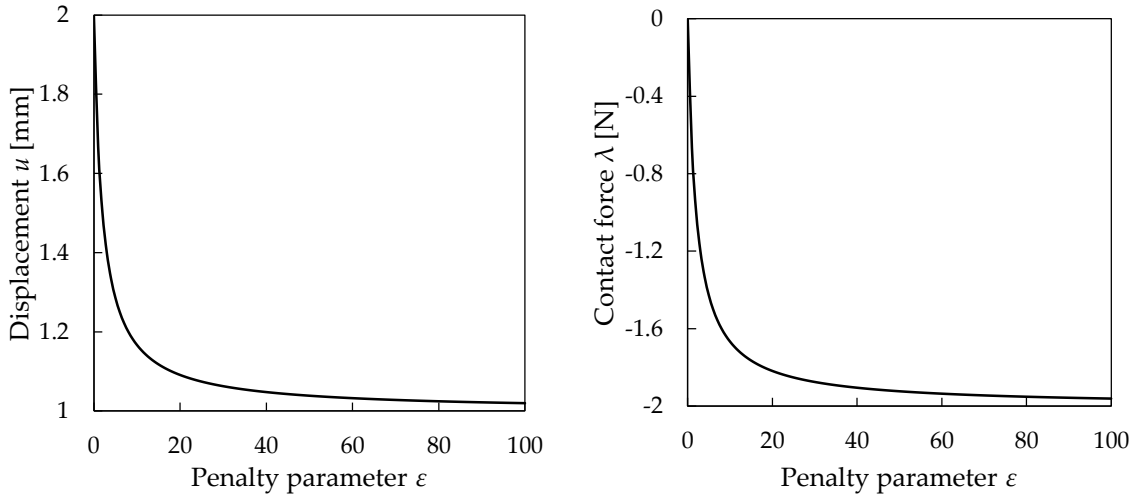
The penalized energy functional of the constrained spring system is presented in Figure 2.16, which is composed by the energy of the spring alone (2.91) together with the penalty term (2.96). The resulting functional is smooth and the violations of the impenetrability constraint (2.94) are penalised by large energies, where the minimum of the energy function

(denoted by hollow points) converges to the solution for increasing penalty parameter. In fact, using a zero value for the penalty parameter, the solution of the initial unconstrained problem (2.93) is recovered. The main disadvantage of the penalty method is also highlighted in Figure 2.16, through the dependence of the solution on the penalty parameter value.

The solution for the displacement of the right end of the spring is expressed in (2.100) as function of the penalty parameter. Its evolution for different values of the penalty parameter is shown in Figure 2.17 (a), highlighting the violation of the contact constraint (2.94), particularly for low values of penalty parameter. Therefore, the penalty solution is only an approximation of the correct enforcement of the constraint condition, being the exact solution obtained in the limit case of infinite value for penalty parameter. The contact force λ in the right end of the spring, which is constant along all spring, is evaluated using (2.91) being given by:

$$\lambda = E(u-\overline{u}) = 2(u-2) = 2\left(\frac{4+\varepsilon}{2+\varepsilon}-2\right) = -\frac{2\varepsilon}{2+\varepsilon}, \tag{2.101}$$

where the minus signal means compressive forces on the spring. The force is plotted in Figure 2.17 (b) for different values of the penalty parameter. Low values for the penalty parameter leads to a large deviation of the calculated force from the one obtained with the exact enforcement of the impenetrability constraint.



(a) (b)

Figure 2.17. Dependency of the penalty parameter on the spring system solution: (a) displacement; (b) contact force.

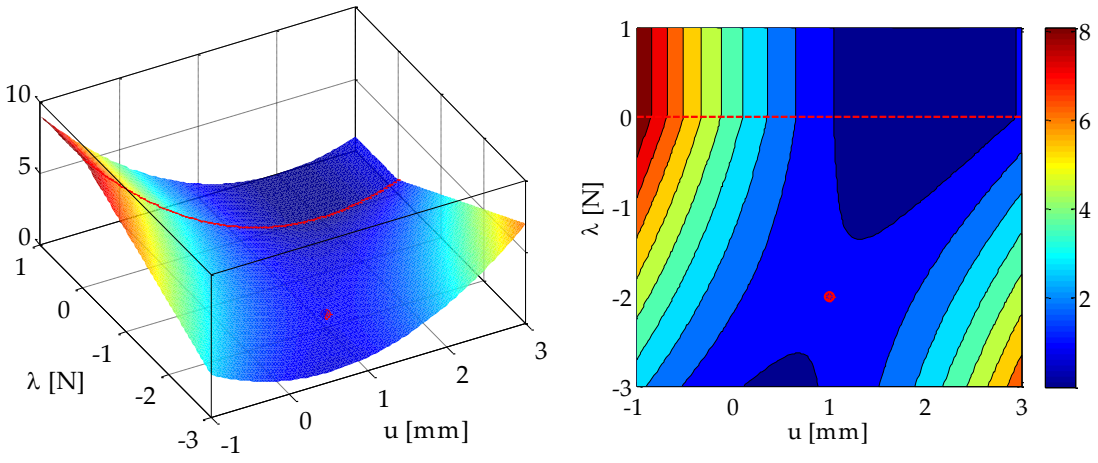
2.3.4.2. Lagrange multiplier method

This section deals with the application of the Lagrange multiplier method to solve the same contact problem (Figure 2.15). The energy functional (2.91) and the contact constraint (2.94) are introduced in the Lagrangian expression (2.67), leading to the following form for the corresponding Lagrangian functional:

$$\mathcal{L}(u,\lambda) = \Pi^{\text{int}}(u) + \lambda g = (u-2)^2 + \lambda(1-u), \quad \lambda \leq 0, \tag{2.102}$$

where λ is the contact force arising at the right end of the spring. The solution represents a stationary point with respect to both displacement u and contact force λ, as shown in Figure 2.18, where the Lagrangian functional is represented. The saddle point is indicated by a red point in the figure, being located in $u = 1$ and $\lambda = -2$, defining the exact solution of the constrained minimization problem. The position of the stationary point is easily identified using isolines in the contour of the Lagrangian functional, as shown in Figure 2.18 (b).

The drawback associated to the Lagrange multiplier method is evident in Figure 2.18, where the lack of smoothness is observed in the line $\lambda = 0$ (red dashed line), which separates contact and non-contact zones. The active set strategy adopted in this example defines that the constraint is active for non-positive contact force values leading to the functional (2.102), while the original functional (2.91) is recovered when the constraint is inactive (non-contact zone). In this simple example only one constraint is present to be enforced. Nevertheless, in general the Lagrangian functional (2.102) is only applied where the contact constraint is active, being the determination of the active constraints a nontrivial task, which in general involves an iterative procedure [Laursen 02].



(a) (b)

Figure 2.18. Lagrangian functional corresponding to the spring in contact with a rigid wall: (a) surface with a saddle point solution; (b) contour of the functional values and stationary point.

The solution of the constrained minimization problem is obtained through the variation of the Lagrangian (2.102), leading to a mixed variational formulation given by:

$$\delta\mathcal{L}(u,\lambda)=\begin{bmatrix}2(u-2)-\lambda\\ 1-u\end{bmatrix}^{\mathrm{T}}\begin{Bmatrix}\delta u\\ \delta\lambda\end{Bmatrix}=0,\quad \lambda\le 0, \tag{2.103}$$

where the solution constitutes a saddle point of the functional (2.102). Since δu and $\delta\lambda$ are arbitrary, the solution point is defined as:

$$\begin{cases}2u-4-\lambda=0\\ 1-u=0\end{cases}\quad\Leftrightarrow\quad\begin{bmatrix}2 & -1\\ -1 & 0\end{bmatrix}\begin{Bmatrix}u\\ \lambda\end{Bmatrix}=\begin{Bmatrix}4\\ -1\end{Bmatrix}, \tag{2.104}$$

where the solution of the system of equations is $u=1$ and $\lambda=-2$. Since λ is negative, the solution of the problem is $u=1$, otherwise the contact contribution has to be excluded from consideration and the obtained equation is $2u=4$, with solution expressed in (2.93). The Lagrange multiplier method allows a strict satisfaction of the contact constraint (2.94) increasing the global system size, being the number of additional unknowns identical to the number of constraints (one in this example).

Note that for a spring with nonlinear behaviour, the resulting system of equations (2.104) becomes nonlinear, and its numerical solution can be obtained with the Newton–Raphson method. In that case, the linearization of the system of equations is required for a given iteration in order to compute the incremental solution. The active set strategy should be integrated in the convergence loop in order to check the constraints and update the active/inactive set according with the possible constraint violations [Luenberger 08]. The active set strategy based on the check of the violation of the geometric impenetrability condition $g\ge 0$ is commonly adopted due to its robustness [Yastrebov 13]. However, its simplest implementation leads to two nested iteration loops, an inner loop solving the nonlinear finite element formulation and the outer loop for solving the correct active set [Brunssen 07], [Hartmann 08]. An efficient alternative is the primal–dual active set strategy, which is based on a reformulation of the Hertz–Signorini–Moreau conditions (2.48) as one equivalent equality constraint. This strategy allows integrate all nonlinearities (finite deformations, nonlinear material behaviour and the search for the correct active set) into one single nonlinear iteration scheme [Hüeber 08], [Popp 09], [Popp 12]. In fact, the Lagrange multiplier formulation based in this primal–dual active set strategy is analogous to the coupled augmented Lagrangian method stated in [Alart 91].

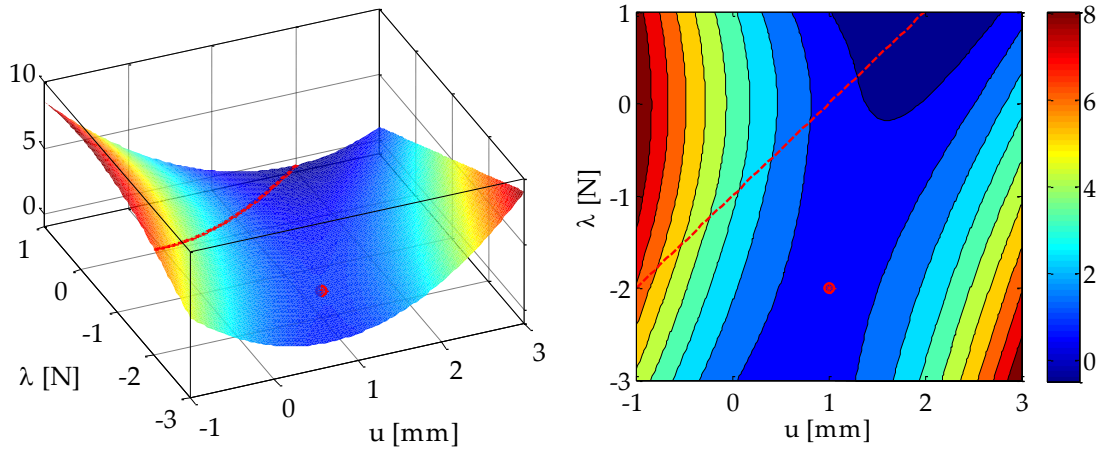
2.3.4.3. Augmented Lagrangian method

The purpose of this section is to solve the simple example considered in the previous sections: the spring in contact with a rigid wall (Figure 2.15), adopting the augmented Lagrangian method. The spring is subjected to a constraint due to the contact with a rigid

wall, which is defined in (2.94). Considering the linear behaviour of the spring with the potential given by (2.91), the augmented Lagrangian (2.76) takes the form:

$$\mathcal{L}^{a}(u,\lambda)=\begin{cases}(u-2)^2+\lambda(1-u)+\dfrac{1}{2}\varepsilon(1-u)^2, & \lambda+\varepsilon(1-u)\le 0\\ (u-2)^2-\dfrac{1}{2\varepsilon}\lambda^2, & \lambda+\varepsilon(1-u)>0\end{cases}, \tag{2.105}$$

which is defined by two different potentials for contact and non-contact zones. The representation of the augmented Lagrangian functional (2.105) using a unitary penalty parameter value is presented in Figure 2.19. The selected value for the penalty parameter is small in order to keep a reasonable scale (similar to the one used in Figure 2.18) for plotting the potentials. Moreover, this selection allows highlight that, in contrast to the penalty method, the augmented Lagrangian approach achieves the exact solution for finite values of penalty parameter. By means of a simple visual observation, the augmented Lagrangian functional seems to be smooth in the transition between contact and non-contact zones (red dashed line in Figure 2.19), which is defined by $\lambda+\varepsilon(1-u)=0$. The value of the penalty parameter dictates the slop of the line, which always cross the point $(u=1,\lambda=0)$.



(a) (b)

Figure 2.19. Augmented Lagrangian functional for the spring in contact with a rigid wall, using a unitary penalty parameter value: (a) surface with a saddle point solution; (b) contour of the functional values and stationary point.

The solution of the contact problem corresponds to the saddle point of the augmented Lagrangian functional, indicated by a red point in Figure 2.19, being its position independent of the penalty parameter. The stationary point is obtained by rendering (2.105) stationary with respect to both the primal and dual variables, expressed by the variation of the augmented Lagrangian, given by:

$$\delta\mathcal{L}^{\mathrm{a}}(u,\lambda)=\begin{cases}\begin{bmatrix}2(u-2)-\lambda-\varepsilon(1-u)\\(1-u)\end{bmatrix}^{\mathrm{T}}\begin{Bmatrix}\delta u\\\delta\lambda\end{Bmatrix}=0, & \lambda+\varepsilon(1-u)\le 0\\ \begin{bmatrix}2(u-2)\\-\dfrac{1}{\varepsilon}\lambda\end{bmatrix}^{\mathrm{T}}\begin{Bmatrix}\delta u\\\delta\lambda\end{Bmatrix}=0, & \lambda+\varepsilon(1-u)>0\end{cases}, \tag{2.106}$$

which represents the variational principle of virtual work. The gradients of the augmented Lagrangian denoted by $\nabla_u\mathcal{L}^{\mathrm{a}}(u,\lambda)$ and $\nabla_\lambda\mathcal{L}^{\mathrm{a}}(u,\lambda)$ are represented in Figure 2.20, which are defined by the first and second line of system (2.106), respectively. It is possible to observe that both gradients loose the smoothness across the contact–non-contact interface $\lambda+\varepsilon(1-u)=0$. Indeed, the first variation of the augmented Lagrangian (2.106) (balance of virtual work) is continuous for any value of u and λ, nevertheless its derivative is not continuous in the transition between contact and non-contact zones (red dashed line), as shown in Figure 2.20. This discontinuity is particularly undesirable when the spring presents a nonlinear behaviour, since (2.106) becomes a nonlinear system of equations, which should be linearized for applying the Newton–Raphson method. This discontinuity in the Jacobian matrix motivates the usage of the generalized Newton's method [Alart 91].

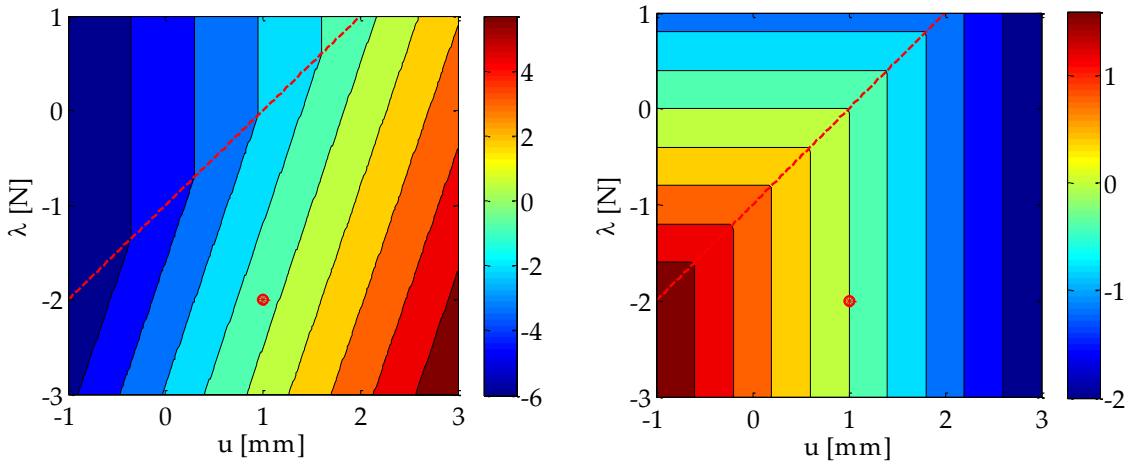


Figure 2.20. Gradients of the augmented Lagrangian functional using unitary penalty parameter: (a) gradient with respect to displacement; (b) gradient with respect to Lagrange multiplier (force).

Supposing that the system (2.106) is nonlinear, the required second variation of (2.105) is given as follows:

$$\Delta\delta\mathcal{L}^{\mathrm{a}}(u,\lambda)=\begin{cases}\begin{bmatrix}\Delta u\\ \Delta\lambda\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}2+\varepsilon & -1\\ -1 & 0\end{bmatrix}\begin{Bmatrix}\delta u\\ \delta\lambda\end{Bmatrix}=0, & \lambda+\varepsilon(1-u)\le 0\\ \begin{bmatrix}\Delta u\\ \Delta\lambda\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}2 & 0\\ 0 & -1/\varepsilon\end{bmatrix}\begin{Bmatrix}\delta u\\ \delta\lambda\end{Bmatrix}=0, & \lambda+\varepsilon(1-u)>0\end{cases}, \tag{2.107}$$

where the number two in the first element of the matrices represents the spring stiffness E. Thus, for a linear material an appropriate choice of the penalty parameter is of the order of its stiffness. In the case of a nonlinear material and nonlinear constraints, it is suggested to increase the penalty parameter value during iterations in order to improve the convergence of the augmented Lagrangian method [Yastrebov 12]. Different techniques for the penalty update have been proposed by Mijar and Arora [Mijar 04]. Nevertheless, very high values can lead to ill-conditioning of the matrix. The matrices in (2.107) are the tangent matrices involved in the incremental solution procedure. Thus in case of non-contact, the eigenvalues of the tangent matrix are expressed by $\lambda_{\max}=2$ and $\lambda_{\min}=-1/\varepsilon$, while the condition number is given by:

$$\mathrm{cond}(\mathbf{K})=\frac{|\lambda_{\max}|}{|\lambda_{\min}|}=\frac{2}{|-1/\varepsilon|}=2\varepsilon, \tag{2.108}$$

where $\mathbf{K}$ denotes the tangent matrix for non-contact status. For a high penalty parameter and high stiffness coefficient the condition number of the tangent matrix becomes very high (ill-conditioned), affecting negatively the precision of the solution.

Chapter 3

General Finite Element Framework

This chapter is devoted to the numerical aspects of contact mechanics, namely the numerical treatment of the frictional contact problem with the finite element method. The implicit time integration scheme and the spatial discretization of the deformable bodies with solid finite elements are briefly addressed, since these are the general features of the finite element code adopted to perform the implementation of the developed algorithms. The solution of the resulting nonlinear system of equations using the generalized Newton method is briefly discussed, as well as its applicability to contact problems regularized through the augmented Lagrangian method. The discretization of the contact interface using the Node-to-Segment technique is presented, comparing the classical single-pass with the two-pass contact algorithm. The particular application of the Node-to-Segment technique for contact problems involving rigid obstacles is also addressed. The accurate identification of the contact zones is very important in the numerical simulation of contact problems and, at the same time, a computationally expensive task. The presented contact search algorithm is decomposed in two steps: global local search. The global contact search algorithm developed to deal with contact between a deformable body and several rigid obstacles is presented in detail, which takes into account the topology of the contact surface discretization. Its extension to handle the contact between separated deformable bodies is described, including the particular case of self-contact problems, where the master-slave discretization is unknown. The local contact search algorithm is based in the normal projection of the slave node on the discretized master surface. It is identical for both types of contact, i.e. contact against a rigid obstacle or contact between deformable bodies. A simple example is presented to show the influence of the contact surface discretization on the accuracy and stability of the local contact search algorithm.

3.1. Finite element code DD3IMP

This section contains a brief description of the in-house finite element code DD3IMP, which has been continuously developed at the Mechanical Engineering Centre of the University of Coimbra over several years. Although the purpose of its development has been the three-dimensional numerical simulation of sheet metal forming processes [Menezes 00], nowadays its range of application is much broader [Oliveira 08], whereby the presented work provides a significant contribution. Since all algorithms developed throughout this thesis will be implemented within the DD3IMP code, some details concerning the organization of this finite element code and relevant algorithms are addressed.

Typically the inertial terms can be neglected in sheet metal forming processes, thus the finite element code was developed within the framework of the *quasi*-static nonlinear analysis. The mechanical model considers large elastoplastic strains and rotations and assumes that the elastic strains are negligibly small. The evolution of the deformation process is described by an updated Lagrangian scheme, where the reference configuration is always the last converged configuration. The elastic behaviour is assumed to be isotropic, while the plastic behaviour can be described by several anisotropic yield criteria: (i) Hill'48 [Hill 48]; (ii) Barlat'91 [Barlat 91]; (iii) Drucker +L; (iv) Karafillis & Boyce'93 [Karafillis 93]; (v) Cazacu & Barlat'01 [Cazacu 01] and (vi) Cazacu & Plunkett'06 [Cazacu 06]. Besides, different isotropic and kinematic work hardening laws are available to describe the evolution of the yield surface with plastic work [Oliveira 07]: Swift and Voce laws for isotropic hardening, which can be combined with a nonlinear kinematic hardening law [Lemaitre 90].

The contact is established between a deformable body and the forming tools, which are assumed rigid (frictional Signorini's problem). The geometry of the forming tools is modelled by parametric Bézier patches and the friction is described by a classical Coulomb's law [Heege 96]. The frictional contact problem is regularized through the augmented Lagrangian method [Alart 91], leading to a mixed system of equations, comprising both displacements and contact forces as unknowns. The principal feature of DD3IMP is the adoption of a single iterative loop to solve simultaneously the nonlinearities related with the mechanical behaviour (large deformations and elastoplastic material behaviour) and the contact with friction [Menezes 00]. In order to reduce the required computational time, some high performance computing techniques have been incorporated to take advantage of multi-core processors, namely OpenMP directives in the most time consuming branches of the code [Menezes 11]. The global framework of the finite element environment is schematically represented in Figure 3.1, in which the contact component are highlighted.

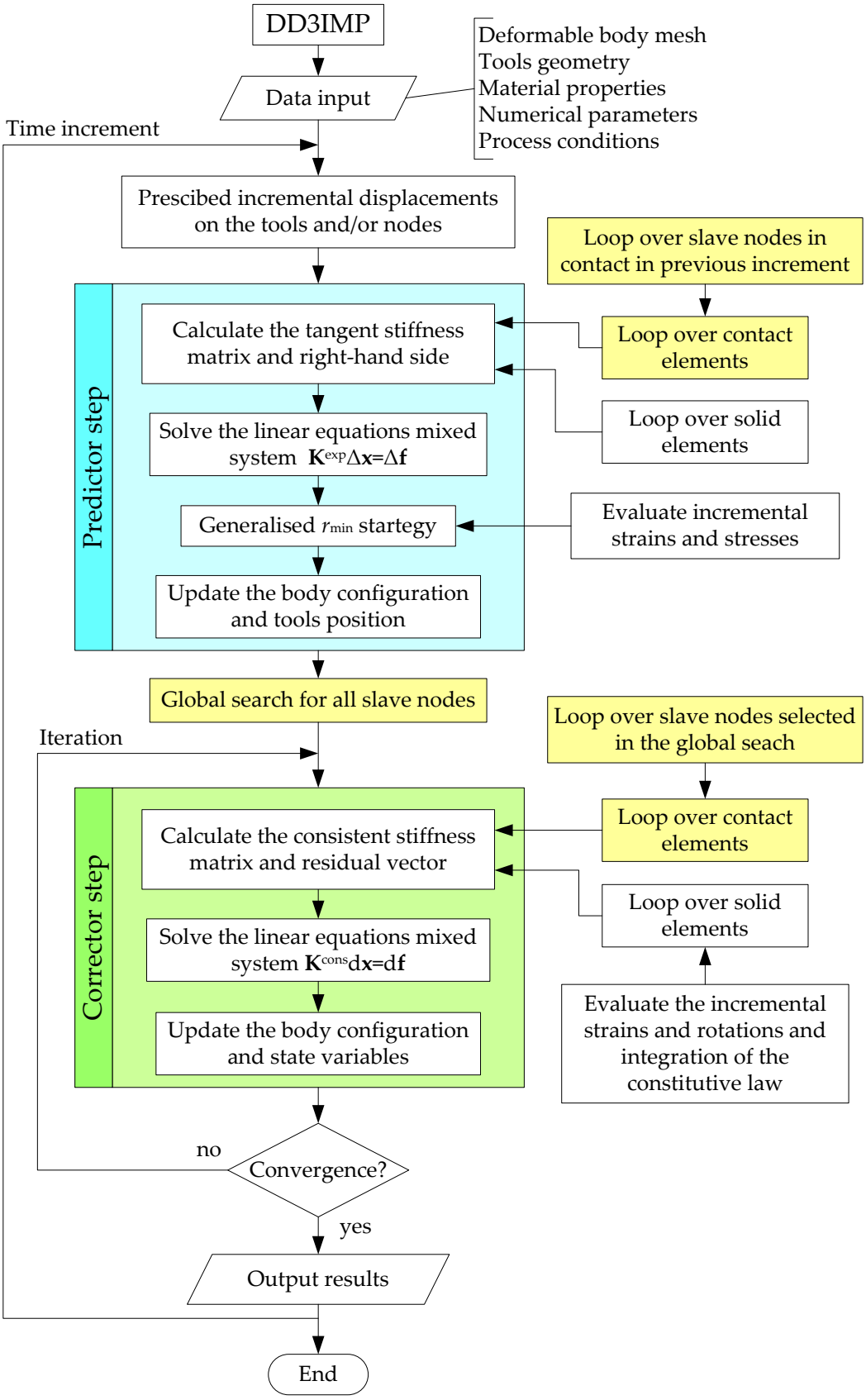


Figure 3.1. Global framework of finite element code DD3IMP.

3.1.1. Time integration

The materials that present a mechanical behaviour dependent of the deformation history are called path-dependent, where the metals are included due to its plastic behaviour at large strains. Moreover, the contact conditions dictate a continuous change of the contact interfaces during the time interval of interest T, imposing different boundary conditions at each instant. When the friction phenomena is taken into account, the solution becomes path-dependent due to the dissipative effect of the friction process. Therefore, a temporal discretization is necessary to capture all these time dependent effects, where the time interval of interest is subdivided into a set of subintervals such as:

$$\mathbb{I} = [0, T] = \bigcup_{n=0}^{N-1} [t_n, t_{n+1}], \tag{3.1}$$

where n is an index on the time increment and N is the total number of increments. While it is commonly assumed that $t_0 = 0$ and $t_N = T$, in general it is not imposed that all time increments have the same magnitude. Indeed, when dealing with a *quasi*-static formulation of contact problems (rigid-deformable), the time increment can be replaced by the incremental displacement of the rigid surface that controls the process.

The time integration method adopted in the finite element code DD3IMP is a fully implicit scheme, which ensures the static equilibrium conditions at each instant by means of an iterative Newton–Raphson procedure. Nevertheless, it is well known that the implicit methods suffer from convergence problems, particularly due to the frictional contact constraints that lead to strong nonlinearities. In order to overcome some of these problems, each time increment is divided into two steps: the Predictor step, which determines the explicit trial solution, and the Corrector step, where this solution is iteratively corrected by an implicit method, as highlighted in Figure 3.1. The explicit approach is applied to calculate an approximate first solution for the incremental displacements, stresses and contact forces. Then, in order to improve the convergence during the iterative stage, the time increment size (defined by the user) is adjusted using a generalization of the $r_{\min}$ strategy proposed by [Yamada 68]. The restrictions on the increment size are mainly imposed by the incremental strains, rotations and stresses obtained in the explicit solution (for more details see [Oliveira 04]). The explicit integration method used in the predictor step guarantees that all calculated variables are proportional to the increment size. Thus, the solution for the adjusted increment size is easily obtained by the application of a dimensionless coefficient to the explicit solution.

The explicit solution resulting from the predictor step does not satisfies the variational principle due to all nonlinearities involved in the problem. Therefore, this trial solution is successively corrected using an implicit method until a satisfactory equilibrium state of the deformable body is achieved. The implicit time integration methods are in general

unconditionally stable, allowing for relatively large time step sizes in comparison with explicit schemes. However, the implementation of implicit methods is more challenging, since it requires the linearization of the nonlinear system of equations, which includes the evaluation of consistent stiffness matrices [Belytschko 00]. The Newton–Raphson method is employed to solve the nonlinear set of equations derived from the *quasi*-static implicit formulation, where the required initial solution is obtained from the explicit solution, as schematically shown in Figure 3.1. Note that the frictional contact problem is solved in DD3IMP with the augmented Lagrangian method, thus the system of equations is mixed, i.e. comprises both displacement and contact forces as unknowns. Since the implicit finite element formulation requires the solution of a linear system of equations in each iteration (see Figure 3.1), the computation time depends quadratically on the number of degrees of freedom [Tekkaya 09].

3.1.2. Spatial integration

The basic idea of the finite element method is to replace the original continuous problem (infinitely dimensional) by a similar finite dimensional problem. This process is called discretization, where the entity discretized is the variational form (2.57). The first step comprises the geometrical approximation of the bodies' domain, on which the boundary value problem is defined. Thus, the continuum domain is discretized by Ne finite elements, such as:

$$\Omega \approx \Omega^{\mathrm{h}} = \bigcup_{i=1}^{Ne} \Omega_i^{\mathrm{h}}, \tag{3.2}$$

where Ω^{h} represents the discrete domain of Ω and Ω_i^{h} denotes the domain of a generic finite element. The operator $\bigcup$ stands for the addition operation between all elements. Figure 3.2 presents the spatial discretization of a continuous deformable body, which is defined by a set of finite elements, each one composed by a set of nodes, indicated by dots. All degrees of freedom in the discrete system are associated with these nodes. The shape of each element is completely characterized by the coordinates of the nodes attached to it and the associated shape functions. Note that for a finite number of nodes, some points of the continuous body have no counterpart in the discretized geometry and vice versa, as shown in Figure 3.2. This difference occurs only in the boundary, which is a very important location for accurately imposing boundary conditions, particularly during the contact treatment.

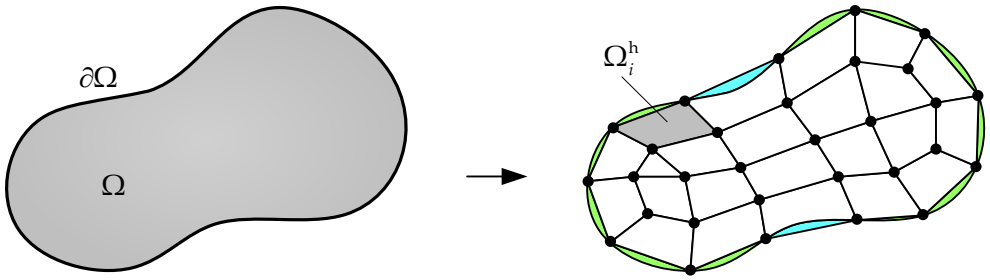


Figure 3.2. Continuous body and its discretized representation using a finite element mesh composed by finite elements and nodes.

For the spatial discretization of the deformable body, standard three-dimensional isoparametric finite elements are employed. Based on the isoparametric concept, the same shape function is applied to interpolate both the geometry and the independent field variables (deformations or stresses), such as:

$$\mathbf{x} \approx \mathbf{x}^h = \sum_{i=1}^{Nn} N_i(\boldsymbol{\xi}) \mathbf{x}_i^h, \quad \text{and} \quad \mathbf{u} \approx \mathbf{u}^h = \sum_{i=1}^{Nn} N_i(\boldsymbol{\xi}) \mathbf{u}_i^h, \tag{3.3}$$

where Nn defines the number of nodes of the finite element, N_i represent the element shape functions (Lagrange polynomials), while $\mathbf{x}_i^h$ and $\mathbf{u}_i^h$ denote the unknown nodal coordinates and displacements, respectively. The shape functions $N_i(\boldsymbol{\xi})$ are defined with respect to the reference element geometry or parameter space, commonly denoted as natural coordinates $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$ for three-dimensional problems (see Figure 3.3). For general shape functions, see [Zienkiewicz 00a]. The use of shape functions (3.3) in the finite element method introduces restrictions on the solution and weighting spaces defined in (2.22). In the discrete setting, these spaces only contains a finite number of solution and weighting functions, respectively, which is expressed mathematically in terms of finite dimensional subspaces $\mathcal{U}^h \subset \mathcal{U}$ and $\mathcal{V}^h \subset \mathcal{V}$.

After the division of the body domain (volume of integration) into a finite number of solid elements, the integration over each element domain is performed approximated via numerical integration procedures. The Gauss quadrature method is used to integrate all field variables over the solid element domain. This method is easily employed in the reference configuration (canonical domain) of the finite element, depicted in Figure 3.3 (left). For every finite element, the coordinates of each point defined in the Euclidean space are related with the local coordinates of the element through the shape functions (3.3), as shown schematically in Figure 3.3. Then, the elemental contributions are sorted into global vectors based on the assembly operator, which manages the position of each local vector quantity into the global vector, such as:

$$\int_{\Omega}(\cdot)d\Omega \approx \mathop{\mathbf{A}}_{i=1}^{Nel}\int_{\Omega_i^h}(\cdot)d\Omega_i^h, \tag{3.4}$$

where $\mathbf{A}$ is the standard finite element assembly operator.

Currently, the finite element library of DD3IMP is composed by tetrahedral and hexahedral solid elements, namely: (i) 4-node linear tetrahedral; (ii) 8-node tri-linear hexahedral; (iii) 10-node quadratic tetrahedral; (iii) 20-node serendipity hexahedral; (iv) 10-node quadratic tetrahedral and (v) 27-node tri-quadratic hexahedral. Tetrahedral elements are geometrically more adaptable and easier to handle in automatic meshing of complex shapes than hexahedral elements [Tekkaya 09]. Nevertheless, standard linear tetrahedral elements are overly stiff, very much sensitive to mesh distortion and the plastic incompressibility constraint results in volumetric locking. In fact, since a tetrahedral element is geometrically a degenerated hexahedral element, more tetrahedral elements are required to achieve the same level of accuracy as in hexahedral elements [Benzley 95]. The tri-linear hexahedral isoparametric elements, when associated with a full integration scheme, present a deficient behaviour in elastoplastic problems [Nagtegaal 74]. The element stiffness increases causing the occurrence of artificial hydrostatic stresses, which leads to a complete deterioration of the solution. This effect can be eliminated using a selective reduced integration method, where a reduced integration is used only in particular terms of the stiffness matrix. Regarding the method currently implemented in DD3IMP, the hydrostatic component of the velocity field gradient is considered constant in the whole element (calculated at its central point), replacing its evaluation in each integration point [Hughes 80]. On the other hand, the uniform reduced integration applied to all terms of the stiffness matrix prevents volumetric locking in nearly incompressible cases, nevertheless introduces spurious zero-energy deformation modes that lead to hourglassing. Presently three integration methods are available in DD3IMP code: (i) Full Integration (FI); (ii) Uniform Reduced Integration (URI) and (iii) Selective Reduced Integration (SRI). The elemental stiffness matrix and the nodal force vector definition that results from the discretization of the linearized principle of virtual velocities can be found in [Oliveira 08] for each integration method.

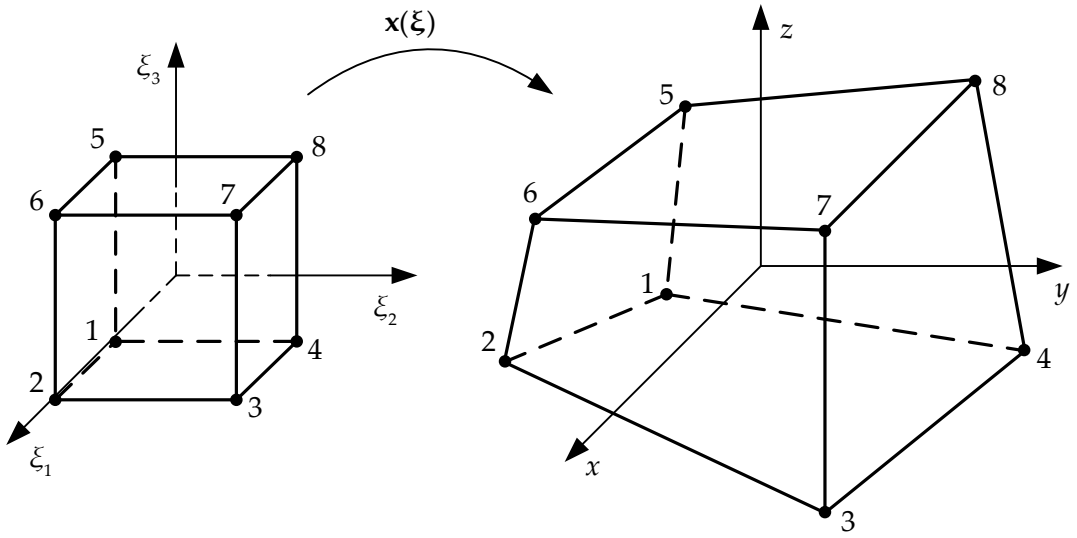


Figure 3.3. Isoparametric 8-node tri-linear hexahedral solid element.

3.2. Generalized Newton method

The finite element spatial discretization transforms the original continuum boundary value problem, expressed by the principle of virtual work (2.57), into a nonlinear system of algebraic equations. The nonlinearities presented in the continuum solid mechanics are caused by geometric (large deformation of bodies) and material (nonlinear constitutive laws) contributions. Moreover, the frictional contact phenomena is governed by nonlinear and non-smooth laws, resulting in even more difficulties from the numerical point of view. Since these three sources of nonlinearities are embedded in the discrete system of equations, an iterative scheme has to be employed.

The Newton's method (also called Newton–Raphson) is a very efficient algorithm to solve the nonlinear problems occurring in the finite element method, since it exhibits quadratic convergence near the solution. The main idea of Newton–Raphson method is to replace the nonlinear problem by a series of linear problems, which are directly solvable by standard methods of linear algebra. This idea can be easily defined using a one-dimensional example, focussing on the linearization process. Considering that the nonlinear equation can be written in the following form:

$$f(\bar{x}) = 0, \tag{3.5}$$

where f is a continuous scalar function with continuous first derivatives, as shown in Figure 3.4, and $\bar{x}$ is the unknown solution. Under this assumption, the nonlinear function can be expressed by a Taylor's series as:

$$f(x) = f(x_0) + \left.\frac{df}{dx}\right|_{x_0} (x - x_0) + \frac{1}{2}\left.\frac{d^2 f}{dx^2}\right|_{x_0} (x - x_0)^2 + \cdots, \tag{3.6}$$

where x_0 is an given initial guess of the solution. The Newton–Raphson method is based in the first terms of the Taylor series, neglecting higher order terms, obtaining a linear equation from (3.6), which is the basis of the iterative procedure. The algorithm is generally expressed in iteration i by:

$$\left.\frac{df}{dx}\right|_{x_i} \Delta x_{i+1} = -f(x_i) \quad \text{with update} \quad x_{i+1} = x_i + \Delta x_{i+1}, \tag{3.7}$$

where the derivative of f is evaluated in the last know solution. Iterations continue until the condition (3.5) is fulfilled at least with a given precision $\left|f(x_i)\right| \le \varepsilon$ sufficiently close to zero. The associated convergence behaviour is depicted in Figure 3.4.

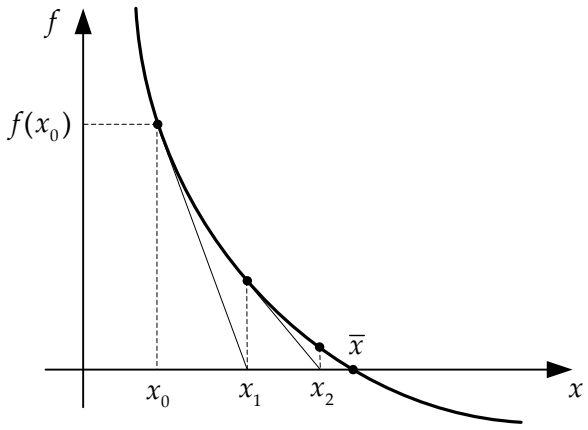


Figure 3.4. Geometrical interpretation of the Newton–Raphson method.

During the development of the Newton–Raphson method from the Taylor series (3.6), all terms contained in the residuum are neglected, supposing that they are negligibly small. Nevertheless, the residuum of the Taylor series R_1 plays an important role in the convergence behaviour, which is expressed in the Lagrange form by:

$$R_1(x_i) = \frac{1}{2}\left.\frac{d^2 f}{dx^2}\right|_{\xi_i} (\bar{x} - x_i)^2, \tag{3.8}$$

where $\xi_i \in]\bar{x}, x_i[$. Applying this definition in the complete Taylor series (3.6) and adding (3.7), the following equality can be rewritten:

$$\varepsilon_{i+1} = -\frac{1}{2}\left[\left.\frac{df}{dx}\right|_{x_i}\right]^{-1}\left.\frac{d^2 f}{dx^2}\right|_{\xi_i}(\varepsilon_i)^2, \tag{3.9}$$

where the discrepancy between x_i and the exact solution $\bar{x}$ is denoted by $\varepsilon_i = \bar{x} - x_i$. According to (3.9), if the second derivate of the function is zero everywhere, then the function is linear and the algorithm converges in one iteration. Moreover, the error is roughly proportional to the square of the previous error, allowing rewrite (3.9) to yield:

$$\left|\varepsilon_{i+1}\right| \le M\left|\varepsilon_i\right|^2, \quad \text{such that} \quad \frac{1}{2}\left|\left.\frac{d^2 f}{dx^2}\right|_{\xi_i}\right|\left|\left[\left.\frac{df}{dx}\right|_{x_i}\right]^{-1}\right| \le M < \infty, \tag{3.10}$$

where M is the least upper bound of the variable coefficient of $\left|\varepsilon_i\right|^2$. If the starting point x_0 is sufficiently close to the solution $\bar{x}$, expressed by the condition $M\left|\varepsilon_0\right| < 1$, then the convergence is quadratic. Although the Newton–Raphson method is often very efficient, there are situations where it performs poorly, for example if:

$$\left.\frac{\partial f}{\partial x}\right|_{x_i} \xrightarrow{x_i \to \bar{x}} 0 \quad \text{or} \quad \left|\left.\frac{\partial^2 f}{\partial x^2}\right|_{x_i}\right| \xrightarrow{x_i \to \bar{x}} \infty, \tag{3.11}$$

then the convergence, if it presents any, is not quadratic [Yastrebov 13].

The above presented one-dimensional Newton–Raphson method is generalized to a multidimensional space, where Lipschitz continuity of the first derivative is required for convergence, as is given by the Kantorovich theorem [Kantorovich 48]. The set of general nonlinear equations to be solved can be stated as:

$$\boldsymbol{\mathcal{F}}(\mathbf{x}) = \mathbf{0}, \tag{3.12}$$

where $\boldsymbol{\mathcal{F}}$ represent a system of nonlinear algebraic equations and $\mathbf{x}$ denotes the set of unknown variables. The expression for the solution increment in the framework of the Newton–Raphson method is written as:

$$\left.\frac{\partial \boldsymbol{\mathcal{F}}}{\partial \mathbf{x}}\right|_{\mathbf{x}_i} \Delta\mathbf{x}_i = -\boldsymbol{\mathcal{F}}(\mathbf{x}_i) \quad \text{with update} \quad \mathbf{x}_{i+1} = \mathbf{x}_i + \Delta\mathbf{x}_i, \tag{3.13}$$

where the lower index i denotes the iteration number. Thus, this linear system of equations has to be solved in each iteration. The directional derivatives defined in (3.13) lead to a square matrix, which is called the tangent matrix for nonlinear problems and Jacobian matrix of $\boldsymbol{\mathcal{F}}$ in optimization theory, denoted as:

$$\mathbf{K}(\mathbf{x}_i) = \left.\frac{\partial \boldsymbol{\mathcal{F}}}{\partial \mathbf{x}}\right|_{\mathbf{x}_i} . \tag{3.14}$$

The inverse of this matrix, required to solve the linear equation system (3.13), exists if and only if its determinant is not zero. Another important characteristic of the tangent matrix is its condition number, already introduced in (2.108), and defined by:

$$\text{cond}(\mathbf{K}) = \frac{|\lambda_{\max}|}{|\lambda_{\min}|}, \tag{3.15}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are respectively the maximal and the minimal eigenvalues of the matrix $\mathbf{K}$. In numerical analysis, the obtained solution is different from the exact analytical solution due to two factors: (i) precision required by the user and (ii) finite number of digits in computer data types. In fact, higher values for the condition number leads to lower number of correctly evaluated digits in the solution. Therefore, ill–conditioning of the matrix (high condition number) results in loss of accuracy and may also result even in divergence of the iterative scheme [Yastrebov 13].

Each iteration of the Newton–Raphson method requires the evaluation of a new tangent matrix, which can be an expensive task, characterizing the main drawback of this iterative method. Indeed, the procedures associated with the tangent matrix calculation present the major contribution to the computational time in the implicit finite element method. In order to alleviate this difficulty, alternatives to the Newton–Raphson nonlinear solution procedure are available. The procedure in which the tangent matrix is computed only in the first iteration is known as modified Newton method, where the tangent matrix is inverted only once being used in all iterations. Another approach is to approximate the tangent matrix by a secant, which is computed from the known states of the previous iterations, establishing the main idea of a *quasi*-Newton method. Typically these methods yield less expensive iterations, nevertheless the convergence rate close to the solution is slower, namely linear and super linear in the modified Newton method and *quasi*-Newton method, respectively [Luenberger 08].

The Newton–Raphson method provides quadratic rate of convergence near the solution if the conditions of convexity and smoothness are fulfilled. However, as shown in Chapter 2, the virtual work functional for contact problems is non-smooth everywhere due to the geometrical restrictions on the displacement field. In fact, the virtual work principle resulting from the augmented Lagrangian method is piecewise smooth with first derivative discontinuities across the gap–contact status lines, as shown in Figure 2.20. Due to this lack of differentiability, most studies based on the augmented Lagrangian in contact mechanics are applied with the Uzawa's algorithm [Simo 92], leading to stable procedures with poor convergence rates due to the alternate treatment of the primal and the dual variables. On

the other hand, the simultaneous treatment of both variables by the Newton–Raphson method converges much faster. Thus, the extension of the Newton–Raphson method to the class of non-differentiable problems arising from contact mechanics was investigated by Alart and Curnier [Alart 91], [Alart 97] and Heegaard and Curnier [Heegaard 93], developing the generalized Newton method (GNM).

The main idea of the generalized Newton method is to split into two parts the system of nonlinear equations (2.90), i.e. a differentiable structural part $\mathcal{F}^{\mathrm{s}}$ and a non-differentiable contact part $\mathcal{F}^{\mathrm{c}}$, such that:

$$\mathcal{F}(\mathbf{u},\boldsymbol{\lambda}) = \mathcal{F}^{\mathrm{s}}(\mathbf{u}) + \mathcal{F}^{\mathrm{c}}(\mathbf{u},\boldsymbol{\lambda}) = \mathbf{0}, \tag{3.16}$$

where $\mathcal{F}^{\mathrm{s}}$ represents the virtual work of the two body system in absence of contact and $\mathcal{F}^{\mathrm{c}}$ denotes the virtual work due to the contact forces. Then, the generalized Newton method is stated as:

$$\begin{bmatrix} \nabla_{\mathbf{u}}\mathcal{F}^{\mathrm{s}}(\mathbf{u}) + \nabla_{\mathbf{u}}\mathcal{F}^{\mathrm{c}}(\mathbf{u},\boldsymbol{\lambda}) \\ \nabla_{\boldsymbol{\lambda}}\mathcal{F}^{\mathrm{c}}(\mathbf{u},\boldsymbol{\lambda}) \end{bmatrix}\Bigg\|_{\mathbf{u}_i,\boldsymbol{\lambda}_i} \begin{Bmatrix} \Delta\mathbf{u}_i \\ \Delta\boldsymbol{\lambda}_i \end{Bmatrix} = -\left\{ \mathcal{F}^{\mathrm{s}}(\mathbf{u}_i) + \mathcal{F}^{\mathrm{c}}(\mathbf{u}_i,\boldsymbol{\lambda}_i) \right\}, \tag{3.17}$$

where i is the iteration index and $\nabla_{\mathbf{u}}\mathcal{F}^{\mathrm{s}}$ denotes the tangent matrix of the contacting bodies. The sub-gradients $\nabla_{\mathbf{u}}\mathcal{F}^{\mathrm{c}}$ and $\nabla_{\boldsymbol{\lambda}}\mathcal{F}^{\mathrm{c}}$ are components of the generalized Jacobians for primal and dual variables:

$$\nabla_{\mathbf{u}}\mathcal{F}^{\mathrm{c}}(\mathbf{u},\boldsymbol{\lambda}) \in \partial_{\mathbf{u}}\mathcal{F}^{\mathrm{c}}(\mathbf{u},\boldsymbol{\lambda}), \quad \nabla_{\boldsymbol{\lambda}}\mathcal{F}^{\mathrm{c}}(\mathbf{u},\boldsymbol{\lambda}) \in \partial_{\boldsymbol{\lambda}}\mathcal{F}^{\mathrm{c}}(\mathbf{u},\boldsymbol{\lambda}). \tag{3.18}$$

In practice, the generalized Jacobian is evaluated as a classical Jacobian at the current interaction for the contact status of each node. Thus, for each contact status a different Jacobian matrix is derived. In the general case of frictional contact, the resulting tangent matrices contained in (3.17) are non-symmetric, non-positive definite and present zero values on the diagonal [Alart 95b]. The convergence properties of the augmented Lagrangian method for frictional contact problem were investigated in detail by [Alart 91], [Alart 97]. The results shown a good convergence for frictionless contact, both in case of small and large slip. In case of frictional problems, some conditions are imposed in the selection of the penalty parameter value to avoid infinite cycling, ensuring convergence of the GNM.

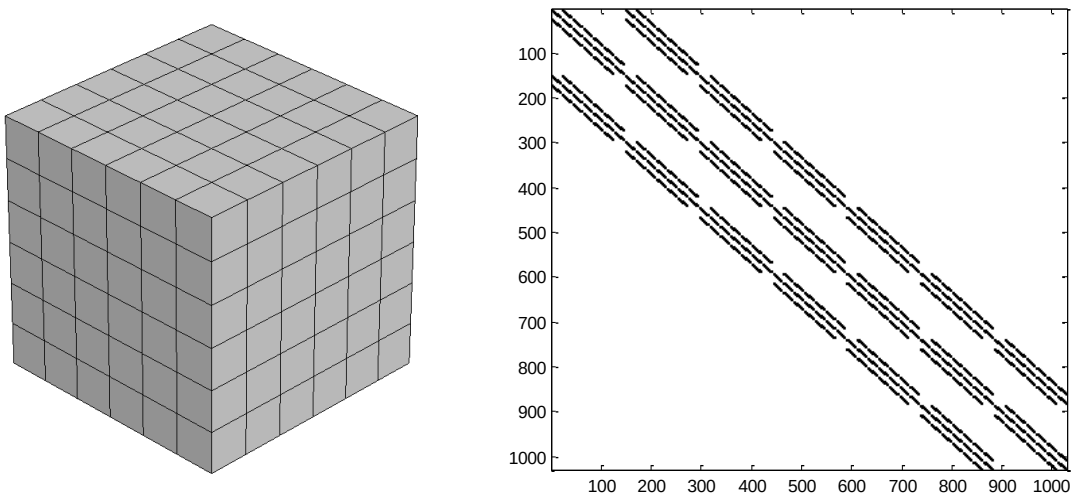
3.2.1. Systems of linear equations

The solution of nonlinear problems (3.16) using the Newton–Raphson method requires the solution of a linear system of equations (3.17) in each iteration step. Therefore, the solution technique employed for solving the linear systems of equations has a significant

effect on the efficiency of the nonlinear finite element solution. In case of large finite element models, very large finite element linear equation systems arise in the iterative solution, which often represent the computationally most expensive part of the overall solution process, within each time step [Wriggers 08]. The structure of the linear system of equations presented in (3.17) can be expressed mathematically in the following usual notation:

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{3.19}$$

where $\mathbf{A}$ denotes the square system matrix, $\mathbf{b}$ is referenced as the right-hand side vector and $\mathbf{x}$ contains the solution vector. The finite element discretization dictates that $\mathbf{A}$ is a sparse matrix (most of the entries are zero) with an evident band structure, which results from the locality of the shape functions. The sparse matrices can be divided in two types: structured and unstructured. The sparse matrix is called structured when the nonzero entries create a regular pattern, typically along a small number of diagonals. On the other hand, if the entries are irregularly located then the matrix is classified as unstructured, which is usually obtained in the finite element discretization of complex geometries [Saad 03].



(a) (b)

Figure 3.5. Example of sparse matrix associated with the finite element discretization: (a) discretized unitary cube; (b) sparse matrix pattern (black dots represent nonzero entries).

The concept of sparse matrices associated with a simple finite element system is presented in Figure 3.5. The three-dimensional cube depicted in Figure 3.5 (a) is discretized with 8-node hexahedral elements, where each edge of the cube is subdivided into 6 finite elements, leading to a model composed by 216 finite elements. The number of nodes is 343, which results into a system with 1,029 degrees of freedom (three for each node). The pattern

of the resultant sparse matrix is illustrated in Figure 3.5 (b), which is highly influenced by the sequence of the numeration assigned to the nodes. Nevertheless, the matrix pattern is always symmetric, and in this case is defined by three principal diagonals. In the present example, the matrix that defines the linear system of equations (3.19) is composed by 1,029 rows/columns. However, the number of nonzero entries (black dots in Figure 3.5 (b)) is only 61,731 in a total of 1,058,841 entries, which correspond to approximately 6%.

There are many different methods to solve sparse linear systems of equation (3.19), which can be classified as direct and iterative methods. The direct methods are based on an LU decomposition of the underlying system matrix, leading two triangular matrices easily invertible. The main advantage of direct solvers is that the solution is always achieved, allowing to solve ill-conditioned and non-positive definite systems of equations, as long as round-off does not affect the solution [Saad 03]. Besides, they try to reduce the computational cost minimizing the fill-in (entries that change from zero to a nonzero value during factorization process), using a renumbering strategy that exchange rows and columns. However, its application is usually limited by the system size, since for very large systems the solution becomes inefficient and the memory requirements increase due to the fill-in created during the factorization and Gaussian elimination. Nonetheless, modern sparse solvers have less memory requirements, allowing to solve problems of several million of unknowns [Schenk 04], [Gould 07].

Iterative solvers are advantageous when large systems of equations have to be solved, since the memory requirement and the total number of operations is less, when compared to direct solvers. However, preconditioning techniques are a prerequisite for the success of the iterative solvers, since its rate of convergence is largely influenced by the condition number, defined in (3.15). The main idea of the preconditioning is to convert the linear system (3.19) into an equivalent system with a lower condition number. The preconditioned iterative solution of sparse linear systems is usually achieved through so-called Krylov subspace methods, where the conjugate gradient method and the generalized minimal residual method are the most popular [Saad 03]. Concerning the preconditioners, its adequate selection is very important for an efficient iterative solver. The Jacobi and the Gauss–Seidel preconditioners are the simplest ones, which are derived from the iterative methods with the same designation. A broad class of preconditioners is based on incomplete factorizations of the coefficient matrix, leading the so-called incomplete LU factorization techniques, which are the most popular and efficient to solve large sparse linear systems [Alart 95b], [Chow 97]. However, the selection and evaluation of a good preconditioner for an iterative method can be computationally more expensive than using a direct method [Gould 07] .

Although the direct methods usually need more operations and larger memory requirements than iterative ones, a direct solver is selected throughout this thesis for solving large sparse linear systems. The Direct Sparse Solver (DSS) from Intel® Math Kernel

Library (Intel® MKL) is the adopted as alternative to the conjugate gradient squared iterative solver previously implemented in DD3IMP [Alves 03], [Menezes 11]. This modern library is highly optimized for scientific and engineering applications that require solving large problems. Moreover, it is optimized for the latest Intel processors, including processors with multiple cores, which are currently standard in personal computers (PC's).

3.2.1.1. Direct Sparse Solver (DSS)

This section deals with the principal phases required to employ the MKL library for solving large sparse linear systems through a direct method. The basic ingredient to obtain the solution of the system (3.19) using a direct method is the factorization of the matrix $\mathbf{A}$ into the product of lower and upper triangular matrices $\mathbf{L}$ and $\mathbf{U}$, usually referred as LU decomposition. Then, the solution of the original problem can be rewritten as follows:

$$\mathbf{Ax} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{LUx} = \mathbf{b} \quad \Leftrightarrow \quad \mathbf{L}(\mathbf{Ux}) = \mathbf{b}. \tag{3.20}$$

This leads to the following two-step strategy for obtaining the solution of the original system of equations:

$$\mathbf{Ly} = \mathbf{b} \quad \text{and} \quad \mathbf{Ux} = \mathbf{y}, \tag{3.21}$$

where the solution of the first and second system is referenced as forward elimination and backward substitution, respectively. When dealing with the solution of sparse systems of equations, the fill-in and the reordering becomes important in terms of solution efficiency. Fill-in denotes the situation of $\mathbf{L}$ having nonzero entries in positions where the original sparse matrix $\mathbf{A}$ contains zero values. This situation leads to an inefficient computation during the forward and backward solver phases (3.21), since the sparsity of the original matrix is lost after the factorization process. Therefore, the rows and columns of $\mathbf{A}$ are permuted based in its nonzero pattern (reordering) in order to reduce the fill-in when $\mathbf{L}$ is computed in the factoring process. Another important aspect to take into account in sparse matrices is the storage format. Since it is computationally more efficient to store only the nonzero elements of a sparse matrix, the basic idea is storing all nonzero entries into a linear array and provide auxiliary arrays to describe its locations in the original matrix [Intel 14].

The Direct Sparse Solver (DSS) interface supported by the Intel MKL is a group of user-callable routines that are used for solving sparse systems of linear equations using the general scheme previously described. The typical invoking order of the DSS interface routines is depicted in Figure 3.6, where the solving process is divided into six phases. In addition, the names of the routines corresponding to each phase are presented in the box in the figure. The first phase (`dss_create`) initializes the solver and creates the basic data structures necessary for the solver, using double precision by default. The purpose of the

second phase (`dss_define_structure`) is to define the locations of the nonzero entries of the matrix, i.e. the sparse matrix pattern. The first step of this phase is to define the general nonzero structure of the matrix: symmetrically structured, symmetric and non-symmetric. The tangent matrices arising from the finite element method applied to nonlinear problems are enclosed in the symmetrically structured group, where the pattern of nonzero entries is symmetric (see Figure 3.5 (b)). The storing of the nonzero entries of the sparse matrix into a linear array is performed walking across each row in order (row-major order). The location of the nonzero entries is performed by means of two arrays, one provides the column number containing the entry, while the other gives the location of the first nonzero entry within each row. The next phase of the DSS interface (`dss_reorder`) is the permutation of rows and columns at the original matrix in order to minimize the fill-in during the factorization phase. The factorization phase (`dss_factor_real`) computes the LU factorization of the sparse matrix as stated in (3.20), where the nonzero values of $\mathbf{A}$ are given in a new linear array. In the following phase (`dss_solve_real`), the solution vector of the system of equations (3.19) is computed based on the factorization computed in the previous phase, performing the forward and backward substitutions (3.21) for a given right-hand side. The last phase (`dss_delete`) deletes all data structures created during the solving process. For more details about each routine see the Intel MKL reference manual [Intel 14].

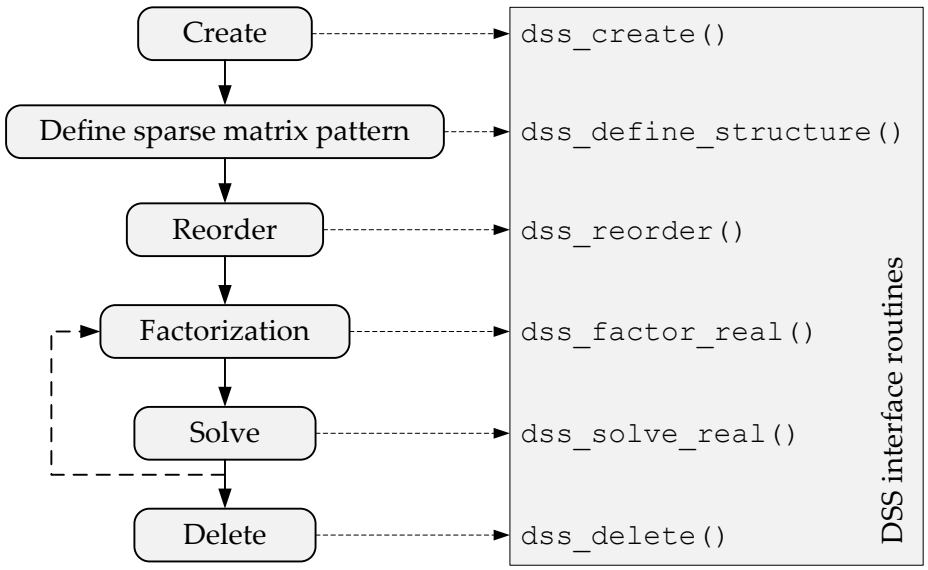


Figure 3.6. Typical sequence for invoking DSS interface routines.

The solution vector for a single system of equations with a single right-hand side is obtained by invoking once each Intel MKL DSS interface routine, using the sequence presented in Figure 3.6. However, solution vectors for multiple right-hand sides are required in some situations. Consequently, the routine that provides the system solution

(`dss_solve_real`) is invoked as many times as the number right-hand sides, keeping the LU factorization performed in the previous phase. Moreover, in certain applications different matrices with the same nonzero pattern are necessary, such as the finite element method in nonlinear solid mechanics. In this case, the phase related with the LU factorization (`dss_factor_real`) is repeated for each matrix, as represented by the dashed line in Figure 3.6.

A simple numerical example is selected to evaluate the performance of the direct solver adopted, which is applied to solve the linear systems of equations arising from finite element analysis. This numerical analysis is performed with the finite element code DD3IMP, running on a personal computer equipped with an Intel Core™ i7–2600K Quad-Core processor (3.4 GHz), 8.0 GB RAM and the Windows 7 Professional (64-bits platform) operating system. The cube illustrated in Figure 3.5 (a), which presents an edge length of 1 mm, is compressed 0.1 mm by means of imposing Dirichlet boundary conditions on the upper surface. Despite the elastoplastic material behaviour assumed for the cube, a single increment is used to impose the boundary conditions. Moreover, symmetry conditions are applied on the lower surface and on the two perpendicular lateral surfaces to avoid rigid body motion. The discretization of the cube is performed with 8-node hexahedral elements, using a different number of subdivisions along the cube edges, as presented in Table 3.1. The displacement of each node (three components) are the degrees of freedom (dof's) for this nonlinear *quasi*-static problem. The two principal performance parameters analysed are the elapsed time and the memory requirements, which obviously increase with the finite element mesh refinement.

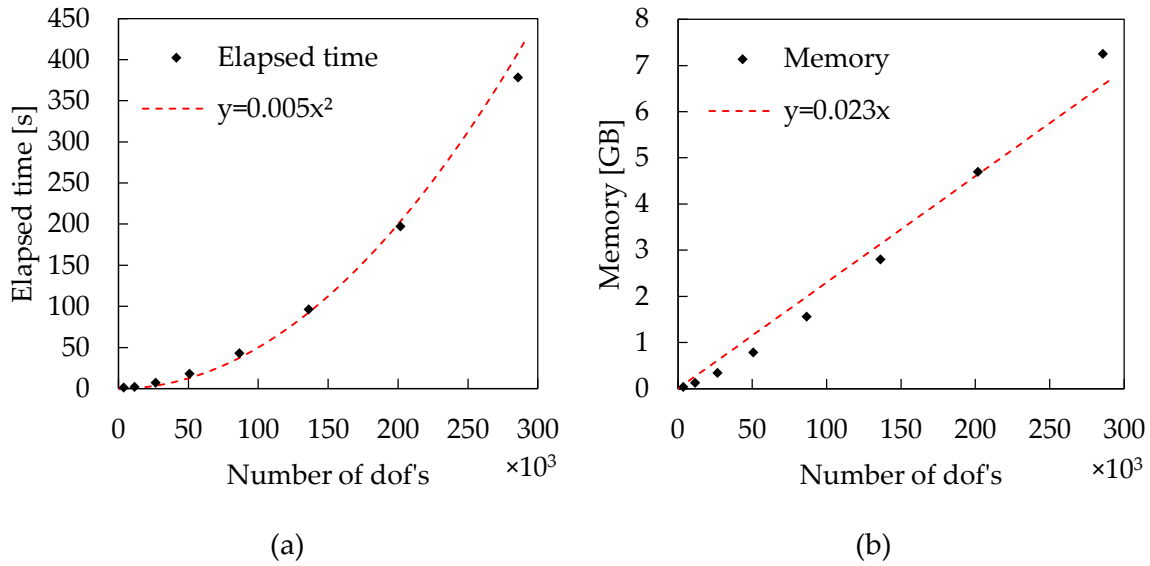
Table 3.1. Main characteristics of the finite element mesh of the cube depicted in Figure 3.5 (a) using n finite elements in each edge subdivision.

n	Nº elements	Nº nodes	Nº dof's	Nº iterations
10	1,000	1,331	3,630	8
15	3,375	4,096	11,520	7
20	8,000	9,261	26,460	7
25	15,625	17,576	50,700	7
30	27,000	29,791	86,490	7
35	42,875	46,656	136,080	7
40	64,000	68,921	201,720	7
45	91,125	97,336	285,660	7

It is well know that in case of a *quasi*-static implicit finite element formulation the computational time depends quadratically on the number of degrees of freedom, when a direct solver is employed [Tekkaya 09]. This relationship is also observed in the present

example, as shown in Figure 3.7 (a), where the elapsed time is approximately quadratic with the number of dof's. For the case of the finer mesh (91,125 finite elements), the elapsed time required to accomplish the results is about 6 minutes. The number of Newton iterations necessary to attain the equilibrium of the body is shown in Table 3.1, which results from the nonlinearities introduced by both the large deformations and the elastoplastic material behaviour. The convergence criterion used in DD3IMP is based in the ratio between the norm of the residual force vector and the contact force norm, where the threshold value was taken as 0.001, for this problem. All finite element meshes need seven iterations to attain this threshold value, except the coarse mesh that requires eight iterations.

The memory consumed during the numerical simulation is shown in Figure 3.7 (b) for each mesh analysed. The necessary memory increases approximately linearly with the number of degrees of freedom. Indeed, most of this memory is used to store the sparse matrix, where the number of nonzero entries increases linearly with the number of dof's (structured mesh). In the case of a simple geometry discretized with a structured mesh (see Figure 3.5 (a)) the number of nonzero entries in each line of the matrix is roughly constant (similar structure), as observed in Figure 3.5 (b). The memory requirement for the finer mesh is higher than 7 GB, which is almost the computer capacity.



(a) (b)

Figure 3.7. Influence of the finite element mesh refinement on the: (a) elapsed time; (b) memory requirement.

3.3. Contact discretization

As shown in the previous section, in general, the spatial discretization of the contacting bodies using a finite element mesh originates differences between the continuous and the discretized boundary geometry, as highlighted in Figure 3.2. In fact, the surface of the finite element mesh is an approximation of the real surface of the body:

$$\partial\Omega \approx \partial\Omega^{h} = \bigcup_{i=1}^{Ns} \partial\Omega_i^{h}, \tag{3.22}$$

where $\partial\Omega$ represents the continuous surface of the body and $\partial\Omega^{h}$ denotes its discretized counterpart, which is defined by the reunion of each finite element surface boundary. Since the finite element discretization produces only piecewise smooth contacting surfaces (see Figure 3.2), some mathematical and numerical difficulties arise in the contact treatment, due to the inaccurate contact surface description.

The contact discretization dictates the structure of the contact elements transferring efforts from one contact surface to the other [Yastrebov 13]. Contact elements can be geometrically interpreted as "bridge elements" between separated surfaces, which share components (nodes or segments) of both surfaces. Typically, the discretization of the contact interface is classified into three main types: (i) Node-to-Node (NTN); (ii) Node-to-Segment (NTS) and (iii) Segment-to-Segment (STS).

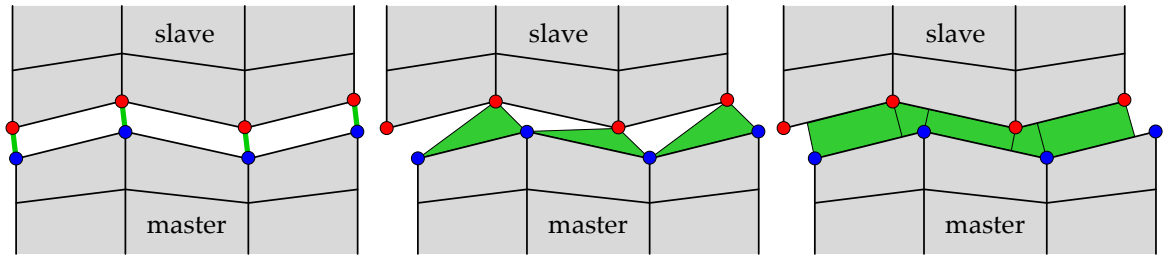


Figure 3.8. Schematic illustration of different types of contact discretization: (a) Node-to-Node; (b) Node-to-Segment; (c) Segment-to-Segment.

The simplest discretization technique is the so-called Node-to-Node, which incorporates the contact constraints directly on the nodal pairs [Francavilla 75], as shown in Figure 3.8 (a). Therefore, this contact discretization can be established only in case of small deformations, not allowing the finite sliding. Furthermore, restrictions on mesh generation are introduced due to the requirement of guaranteeing conforming meshes along the contacting surfaces, i.e. each node of one contacting surface has a single corresponding node on the opposite surface (see Figure 3.8 (a)). On the other hand, due to the imposed conforming contact interface, the NTN discretization passes the contact patch

test proposed by Taylor and Papadopoulos [Taylor 91] (uniform pressure transferred correctly through the contact interface).

For the general case of contact including large deformation and large sliding, the so-called Node-to-Segment discretization technique is quite popular and widely used [Hallquist 85], [Wriggers 90]. This contact discretization is valid for non-conforming meshes (nodes at the contact interface located at dissimilar positions), which can arise from the sliding of the contact interface or when the finite element meshes are generated independently in each body. The contact pairs are composed by a node of the slave surface and a corresponding segment of the master surface, as shown in Figure 3.8 (b), where the contact segments are lines (in two-dimensional problems) joining adjacent master nodes [Simo 85b]. However, in case of non-conforming meshes the NTS discretization fails the contact patch test [Taylor 91] (uniform contact pressure incorrectly transferred across the contact interface).

Alternative methods based in the Segment-to-Segment contact discretization have been first proposed for two-dimensional problems [Simo 85b], [Zavarise 98]. Recently, this discretization technique has been successfully coupled with the mortar method, which was originally developed in the context of domain decomposition techniques for non-conforming meshes [Bernardi 90], [Wohlmuth 01]. The extension of the mortar method to contact problems between elastic bodies was formulated by Belgacem et al. [Belgacem 98] and its application for two-dimensional frictional contact problems subjected to small deformations was firstly established by McDevitt and Laursen [McDevitt 00]. The extension of the STS contact discretization using the mortar method to three-dimensional contact problems involving large deformation and large sliding has been developed during the last few years [Puso 04a], [Puso 04b], [Yang 08], [Puso 08]. Nowadays, the application of mortar methods in computational contact mechanics is still one of the most active research topics [Laursen 12]. The principal property of the mortar method is the incorporation of the contact constrains in a weak form (see Figure 3.8 (c)), allowing an accurate exchange of the contact forces between interface surfaces. Therefore, this technique is stable and passes the contact patch test for non-conforming meshes [Puso 04a]. Moreover, using the dual spaces for the Lagrange multipliers it is possible to eliminate the Lagrange multipliers from the set of linear equations, thus the unknowns are only the nodal displacements [Wohlmuth 00], [Popp 09]. Although the mortar-based STS discretization is more robust and accurate than the classic NTS counterpart, its implementation presents a great challenge. Indeed, for the general three-dimensional case, the mortar interface coupling and numerical integration for arbitrary surface meshes is a complex task due to the surface segmentation.

3.3.1. Node-to-Segment

Despite the drawbacks associated with the Node-to-Segment (NTS) discretization technique, it is the most widely used technique for large deformation contact problems due to its simplicity and flexibility [Zavarise 09b]. For these reasons, it is often implemented in commercial finite element codes being also the contact discretization technique adopted throughout this dissertation. The first step of the NTS discretization comprises the selection of one contacting surface as slave and the other one as master, leading to an asymmetry in the contact problem because the contact surfaces are treated differently. The impenetrability conditions are enforced only in a discrete number of points on the contact slave surface, preventing the slave nodes from penetrating on the contact master surface [Hallquist 85]. However, the master nodes are allowed to penetrate into the slave surface, as depicted in Figure 3.9 (a). Each contact element (not structural) is composed by a slave node and the closest segment (element edge/facet) on the master surface (see Figure 3.9 (b)), which is selected through the orthogonal projection of the slave node. Nevertheless, particularly when the NTS approach is applied with low order finite elements [Crisfield 00], the identification of the master segment is either ambiguous or impossible, which may result in slow convergence or even in divergence of the numerical solution. Some strategies have been specially developed to deal with such problems in 2D frictionless contact problems [Zavarise 09b].

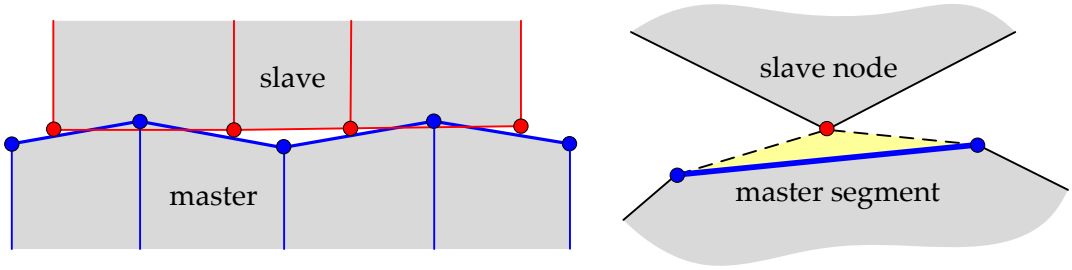


Figure 3.9. Node-to-Segment contact discretization: (a) penetration conditions unchecked in master nodes; (b) contact element composed by a slave node and a master segment.

3.3.1.1. Selection of master and slave surfaces

The proper selection of master and slave surfaces is fundamental to the success of the NTS contact discretization. After identifying the pair of surfaces that will interact, one surface is assigned as master and the other one as slave, leading to an asymmetric contact treatment, as shown in Figure 3.9 (a). The selection should be carried out taking into account that the slave nodes cannot penetrate into the master surfaces, but the master nodes can

penetrate the slave surface. Hence, the main guidelines for choosing the master and slave surfaces are listed below:

- The contact surface that presents the coarse mesh should be the master surface, while the surface with the fine mesh is the slave surface;
- When the stiffness between the contacting bodies is different, the contact surface of the stiffer body should be the master surface and the other should be assigned as slave surface;
- When the contact occurs between a deformable body and a rigid obstacle, the surfaces of the rigid obstacle must be specified as master surfaces;
- In the case of contact between a convex surface with a flat or concave surface, the master surface should be the flat/concave surface;
- If one body slides over another with a contact surface considerably larger, the larger surface should be the master surface, in order to minimize the contact status changes.

In order to highlight the importance of the master/slave surface selection, the contact patch test example is presented. The original patch test [Taylor 91] is modified in this example, where two elastic cubes with identical geometry (each edge with 10 mm) and the same material properties ($E = 100\,\text{MPa}$ and $\nu = 0.3$) are pressed against each other considering frictionless contact. The finite element discretization of each cube using 8-node hexahedral elements is depicted in Figure 3.10 (a), showing that the meshes do not coincide at the contact interface. The bottom surface of the lower cube is constrained against vertical displacements and the four lateral surfaces of the cubes are constrained against displacements in its normal direction. The uniform pressure is imposed by applying a vertical displacement of 1 mm on the top surface of the upper cube.

The distribution of the vertical stress component in the cubes is depicted in Figure 3.10 (b), employing the NTS contact discretization with the upper cube (finer mesh) defined as master. The obtained results do not satisfy the contact patch test due to the non-conforming meshes at the contact interface. Since the finer mesh is assigned as master (incorrect choice), the penetration of some master nodes into the slave cube can be considered excessive (see Figure 3.10 (b)), which leads to high deviations in the predicted contact stress, i.e. inaccurate transmission of constant normal stresses between two contacting surfaces. On the other hand, by exchanging the master and slave surfaces definition, the resulting distribution of vertical stress is shown in Figure 3.10 (c). The noise in the contact stress is reduced considerably when the coarse mesh is assigned as master surface, leading to a contact surface approximately flat after loading. Although the mesh refinement of the slave surface improves the accuracy, only matching meshes at the contact interface (NTN contact discretization) allows the complete elimination of the inaccurate transmission of stresses.

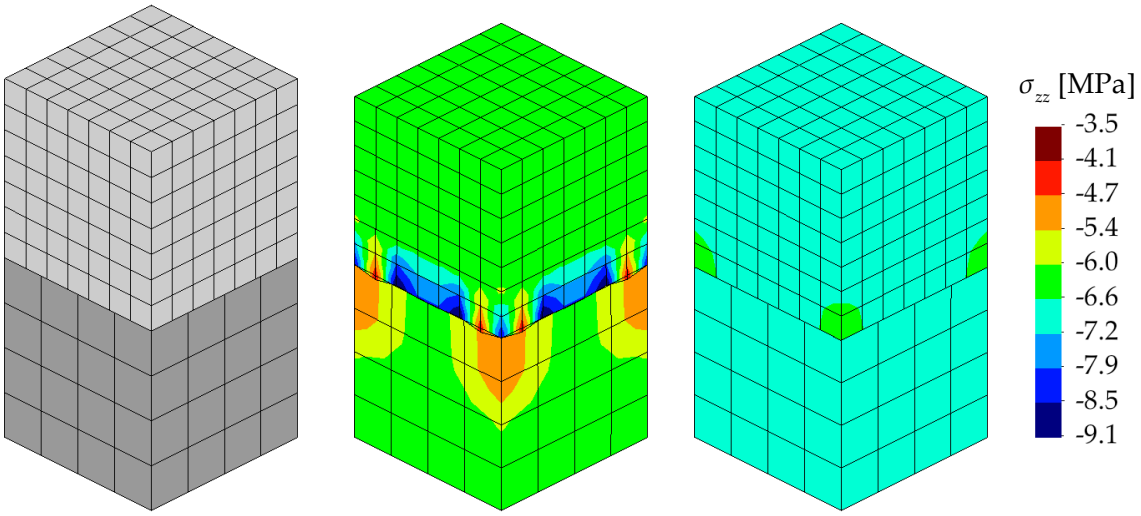


Figure 3.10. Distribution of the vertical stress component for different choices of the master and slave surfaces in the NTS contact discretization: (a) finite element mesh of contacting cubes; (b) upper cube defined as master and lower as slave; (c) upper cube defined as slave and lower as master.

The value of the vertical stress component (compressive) and the corresponding contact force applied on the upper cube can be exactly calculated performing the patch test using a conforming mesh at the contact interface. In this case, the obtained vertical stress is uniform in both cubes being the compressive contact stress 6.9 MPa and the contact force 690.0 N. Nevertheless, when the cube discretized with the finer mesh is wrongly defined as master (Figure 3.10 (b)), the predicted contact force is 646.9 N, which is a consequence of the local penetration of the cubes. On the other hand, the proper selection of the master surface (Figure 3.10 (c)) leads to a contact force value of 689.7 N, which is very close to the exact value. Recently, a modification of the NTS discretization has been proposed by Zavarise and De Lorenzis [Zavarise 09a], which passes the contact patch test in 2D frictionless contact problems using the penalty method to enforce the contact constrains. The basic idea of this algorithm is to create two virtual slave nodes located at the quarter points of each slave segment, improving the contact contribution to the stiffness matrix and to the internal force vector.

3.3.1.2. Two-pass contact

The master-slave formulation used in the NTS discretization is inherently asymmetric (see Figure 3.9 (a)), which is not in accordance with the physical observation of contact problems. Besides, as pointed by Taylor and Papadopoulos [Taylor 91], the single-pass NTS algorithm does not satisfy the contact patch test, which in some circumstances can yield unsatisfactory results, as shown in Figure 3.10. Therefore, the two-pass contact (also called

symmetric contact) approach was developed to try to overcome this problem. The main idea of the two-pass approach is the definition of the each contact surface as master and slave simultaneously, performing a double definition of the contact pair, exchanging the master and slave surfaces. This approach precludes penetration of the slave nodes into the master segments (first pass), while the master nodes are restricted from penetrating on the slave segments, when the master and slave surfaces are exchanged in the second pass. Therefore, this strategy allows eliminating the geometric asymmetry by reversing the role of master and slave surfaces and repeating the same process performed in the single-pass algorithm. Since the number of contact elements created is higher, the two-pass NTS algorithm is less efficient than the single-pass contact approach.

The NTS discretization associated with the two-pass contact algorithm passes the contact patch test in 2D and in some 3D mesh configurations (with sufficient symmetry) for low order finite elements [Taylor 91]. Nevertheless, this discretization technique based on the two-pass contact algorithm has the recognised deficiency of locking, due to the over-constrained system of equations [Puso 04a]. Indeed, if any two nodes on both contact surfaces have identical locations, the corresponding contact constraint is created in duplicate during the two-pass algorithm, which results in a rank deficient matrix (linearly dependent rows and columns) [El-Abbasi 01]. This situation can lead to some numerical difficulties, such as singularities and zero pivots, which can be avoided applying a search algorithm to detect and remove these duplications. On the other hand, the smoothing of the master surface can alleviate the locking problems, due to the continuous change of the surface normal vector [Puso 04b].

In order to evaluate the robustness of the adopted direct solver (Intel MKL DSS), briefly described in Section 3.2.1.1, a simple contact problem comprising 2 finite elements is performed using the two-pass NTS algorithm. The same geometry, material properties and the boundary conditions of the last example are considered (Figure 3.10), but each cube is discretized with a single element, leading to matching meshes at the contact interface. Thus, the selection of the master and slave surfaces is completely arbitrary. The augmented Lagrange method, described in Section 2.3.3, is applied for handling the inequality constraints due to the contact, using a penalty parameter value of 60. The single-pass NTS algorithm leads to a problem involving 12 dof's, 8 representing the nodal displacements (only vertical components) and 4 for the nodal contact forces, which are evaluated only in the slave surface. The square matrix of the linear system of equations (3.19) arising in the first iteration is given by:

$$\mathbf{A} = \begin{bmatrix} 254.5 & -60 & 59.2 & 59.2 & 0 & 0 & 23.6 & 0 & -1 & 0 & 0 & 0 \\ -60 & 254.5 & 0 & 0 & 59.2 & 59.2 & 0 & 23.6 & 1 & 0 & 0 & 0 \\ 59.2 & 0 & 254.5 & 23.6 & -60 & 0 & 59.2 & 0 & 0 & -1 & 0 & 0 \\ 59.2 & 0 & 23.6 & 254.5 & 0 & -60 & 59.2 & 0 & 0 & 0 & -1 & 0 \\ 0 & 59.2 & -60 & 0 & 254.5 & 23.6 & 0 & 59.2 & 0 & 1 & 0 & 0 \\ 0 & 59.2 & 0 & -60 & 23.6 & 254.5 & 0 & 59.2 & 0 & 0 & 1 & 0 \\ 23.6 & 0 & 59.2 & 59.2 & 0 & 0 & 254.5 & -60 & 0 & 0 & 0 & -1 \\ 0 & 23.6 & 0 & 0 & 59.2 & 59.2 & -60 & 254.5 & 0 & 0 & 0 & 1 \\ -60 & 60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -60 & 0 & 60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -60 & 0 & 60 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -60 & 60 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3.23)$$

where the first 8 rows represent the displacements and the last 4 denote contact forces. The obtained matrix is full rank, i.e. all rows are linearly independent. Besides, the determinant of the matrix presented in (3.23) is 2.03e+17 and its condition number is 1.796e+3. Therefore, obtaining the solution for the linear system of equations (3.19) with the matrix (3.23) should be stable and accurate, whatever the numerical method adopted.

On the other hand, the application of the two-pass NTS algorithm in the same example leads to a problem with 16 dof's, 8 representing to the nodal displacements and 8 denoting the nodal contact forces (contact forces evaluated on both slave and master surfaces). The final stiffness matrix for the first iteration is given by:

$$\mathbf{A} = \begin{bmatrix} 314.5 & -120 & 59.2 & 59.2 & 0 & 0 & 23.6 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -120 & 314.5 & 0 & 0 & 59.2 & 59.2 & 0 & 23.6 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 59.2 & 0 & 314.5 & 23.6 & -120 & 0 & 59.2 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 59.2 & 0 & 23.6 & 314.5 & 0 & -120 & 59.2 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 59.2 & -120 & 0 & 314.5 & 23.6 & 0 & 59.2 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 59.2 & 0 & -120 & 23.6 & 314.5 & 0 & 59.2 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 23.6 & 0 & 59.2 & 59.2 & 0 & 0 & 314.5 & -120 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 23.6 & 0 & 0 & 59.2 & 59.2 & -120 & 314.5 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 60 & -60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -60 & 60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 60 & 0 & -60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 60 & 0 & -60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -60 & 0 & 60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -60 & 0 & 60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 60 & -60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -60 & 60 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3.24)$$

where the last 8 rows are related with the nodal contact forces. Since the two-pass algorithm is applied with conforming meshes at the contact interface, some rows are linearly dependent, as observed in matrix presented in (3.24). Indeed, the rank of this matrix is 12 (rank deficient matrix), where the four pairs of linearly dependent rows are 9-10, 11-13, 12-

14 and 15-16. Thus, the condition number is infinite and the matrix determinant is zero (the matrix is not invertible). Therefore, the linear system (3.19) has either no solution or an infinite number of solutions. Nevertheless, the usage of the direct solver from the Intel MKL allows to solve this system of equations without visible numerical problems, providing the solution obtained with the single-pass NTS algorithm (correct solution). This indicates the robustness of the adopted solver, which overcomes efficiently the principal drawback associated with the two-pass NTS algorithm, i.e. locking behaviour due to the over-constraint. Since every node of the contact interface acts as slave node in the two-pass algorithm, static variables are evaluated on both contacting surfaces, which can make the interpretation of some results (nodal contact forces) difficult.

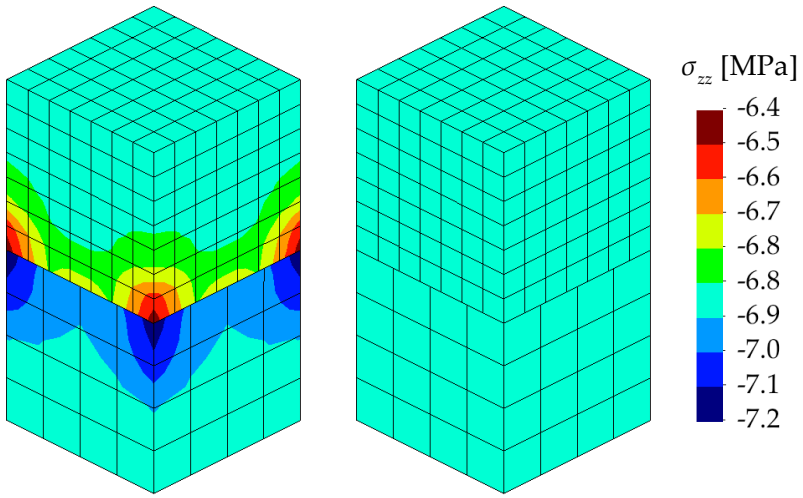


Figure 3.11. Distribution of the vertical stress component in the contact patch test using: (a) the single-pass NTS algorithm with the lower cube as master; (b) the two-pass NTS algorithm.

The comparison between the single-pass and the two-pass NTS contact algorithm is presented in Figure 3.11 for the contact patch test, with the discretization employed in the previous section. For the single-pass NTS contact algorithm, the cube with the coarse mesh is defined as master, while the selection of the master and slave surfaces in the two-pass algorithm is arbitrary due to its exchange in the second pass. The distribution of the vertical stress component obtained with the single-pass algorithm is presented in Figure 3.11 (a), which are the same results shown in Figure 3.10 (c) using a different range of values. In fact, the single-pass algorithm does not satisfies the contact patch test. On the other hand, the two-pass contact algorithm exactly transmits the constant normal stresses between the contacting surfaces, as shown in Figure 3.11 (b), thus solving the contact patch test. The patch test is passed since the expected contact surface is horizontal (flat). In case of refined

finite element meshes on both contact surfaces, the improvements obtained with the two-pass contact algorithm are insignificant while the required computational cost increases.

The NTS contact discretisation enables to use non-conforming meshes at the contact interface, allowing to take into account large sliding of the surfaces during the deformation process [Hallquist 85]. However, the two-pass contact algorithm does not satisfy the contact patch test for some contact surface mesh configurations [Puso 04b]. The patch test example previously presented is repeated using an unstructured mesh at the contact interface (non-conforming). The contact surface mesh of the upper and lower cube is shown in Figure 3.12 (a) and (b), respectively. The distribution of the vertical stress component obtained with the single-pass algorithm is depicted in Figure 3.12 (c), where the noise in the contact stress is within a range slightly higher than the one obtained with the structured mesh (see Figure 3.11). Although the contact stress clearly converges to the exact value using the two-pass NTS contact algorithm, it does not passes the contact patch test using this interface discretization, as illustrated in Figure 3.12 (c). The vertical component of stress (compressive) ranges between 6.84 and 6.96 MPa and the obtained contact force is 690.0 N, which coincides with the exact value.

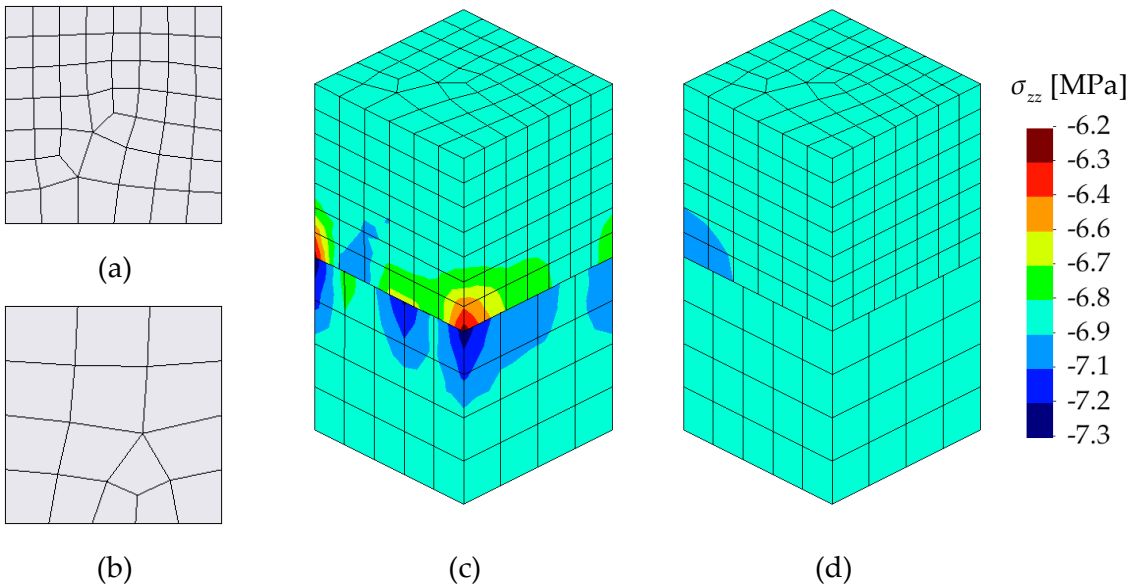


Figure 3.12. Distribution of the vertical stress component in the contact patch test adopting an unstructured mesh: (a) upper surface mesh; (b) lower surface mesh; (c) single-pass NTS discretization defining the lower cube as master; (d) two-pass NTS discretization.

In specific situations the identification of the contact pairs can be a big challenge, such as in self-contact problems, where the identification in advance of the individual contacting areas is very difficult (or impossible) due to large folding of the body. Moreover, the distinction between master and slave surfaces is not clear. In these cases is convenient to

apply the two-pass NTS contact algorithm because the classification of the master and slave surfaces is arbitrary. The two-pass algorithm also is particularly appropriate for problems in which both contact surfaces present very coarse meshes, since the contact constraints are enforced in more locations than in the single-pass contact algorithm.

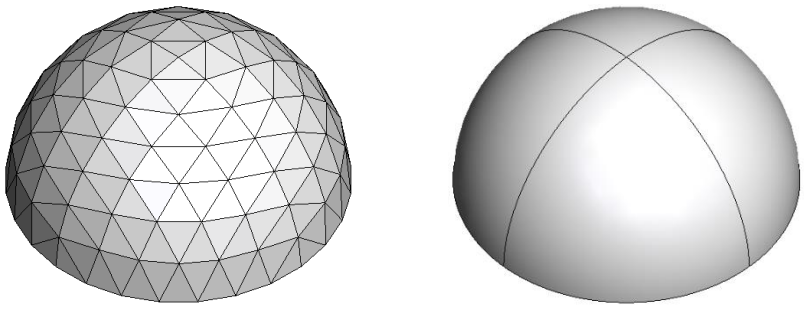
3.3.2. Rigid contact surfaces

The finite element analysis of contact problems can be classified in two distinct categories: (i) contact between discretized deformable bodies and (ii) contact between one discretized deformable body and a rigid surface. In fact, several engineering problems can be included in the second category, such as: metal forming processes, tyre on road, rubber seals and indentation tests. In general, it is considered that the contact occurs between two bodies with significantly different stiffness, i.e. the stiffer body is approximated by a rigid surface while the other body is modelled as deformable. This simplification is very useful since there are no calculations over the rigid body (no additional degrees of freedom involved), leading to a contact problem computationally more efficient. As mentioned before, the rigid body is always classified as the master surface, providing a better convergence of the numerical scheme. Indeed, the contact forces arise only on the slave surface due to the violation of geometrical constraints and friction law conditions, but they are not transmitted through the contact interface to the master surface [Yastrebov 13].

The single-pass Node-to-Segment contact discretization is typically employed in the contact of a deformable body with a rigid obstacle, where the geometrical penetration is evaluated at the nodes of the discretized body (slave) against a rigid surface. Since the discretization of the master surface is not required, this contact discretization technique is usually called the Node-to-Surface. Several approaches have been developed to handle with the contact surface description of rigid bodies, which can be divided in three groups: (i) analytical functions; (ii) finite element meshes (Figure 3.13 (a)) and (iii) parametric patches (Figure 3.13 (b)).

The first description scheme is restricted to simple geometries where the contact surface is composed by an assembly of simple analytical shapes (planes, cylinders, spheres and tori) [Santos 95]. This method is usually adopted in the axisymmetric contact problems, due to the exact description of the contact geometry and associated good convergence, dictated by the smooth surface. On the other hand, for arbitrarily-shaped contact surfaces, the finite element mesh scheme is commonly employed due to its ability and simplicity. However, this approximation can lead to large errors in the geometry, requiring an extremely fine mesh in curved regions to attain a sufficiently accurate surface description [Hama 08], [Hachani 13]. Typically, in order to use the minimum number of finite elements required for a proper description, small elements are used in curved areas and large elements are applied in flat regions. Nevertheless, the discontinuities in the contact surface field (faceted

surfaces) introduced by the finite element discretization (see Figure 3.13 (a)) leads to artificial oscillations in the contact force when sliding is significant, generating convergence problems in the iterative procedure [Hachani 13]. These problems related with the discontinuity of the surface normal vector field can be avoided using parametric patches, which can be obtained directly from the Computer-Aided Design (CAD) model. This surface description method allows creating complex geometries ensuring a high level of continuity with a small number of patches, as shown in Figure 3.13 (b), leading to a more robust behaviour of the iterative solution algorithms [Hansson 90]. Various parameterizations, usually used in CAD models, can be used to define 3D rigid contact surfaces, such as Bézier patches [Heege 96], Hermite patches [Heegaard 96] and trimmed NURBS patches [Shim 00]. However, they are characterized by high order interpolation and complex algorithms, which leads to high computational cost in the contact treatment [Wang 97]. Furthermore, CAD models are known to be plagued by geometrical or topological errors and inconsistencies (gaps/overlaps between abutting surfaces), which result from the lack of a robust solution to the surface intersection problem [Farouki 99]. Therefore, before using the surface model within a finite element environment it is always necessary to perform some laborious manual intervention such as geometry repair, clean-up and preparation [Zhu 13]. Since this manual treatment is incompatible with complex models involving hundreds of patches, the faceted finite element mesh is still the contact surface description method most used in commercial finite element codes, where the treatment required for this case is less laborious and more automatic.



(a) (b)

Figure 3.13. Description of a rigid contact surface (half spherical shell) using: (a) finite element mesh; (b) parametric patches.

The parametric patches also can be employed to smooth the discretized rigid contact surfaces, improving both the accuracy of the surface representation and the convergence behaviour [Hama 08]. Nevertheless, this strategy is more frequently used to smooth the master surface in contact problems between deformable bodies (subject discussed in Chapter 4), where the adoption of a CAD model to describe deformable bodies is impossible. Since the smooth parametric patches are created based on the information from

several adjacent finite elements, its application is frequently restricted to structured surface meshes composed by regular quadrilateral elements (surrounded by eight neighbouring elements). In fact, the extension of the surface smoothing method to arbitrary surface meshes involve several difficulties [Puso 02], [Belytschko 02], presenting an actual research topic in computational contact mechanics [Wriggers 06].

At the beginning of this work, the finite element code DD3IMP was limited to frictional contact problems involving a deformable body with various rigid obstacles, since it was specifically developed for sheet metal forming [Oliveira 08]. The rigid contact surfaces were described by Bézier patches and the information of each patch was obtained directly from the Euclid Styler CAD package using the standard STEP file format. However, this strategy involves some limitations due to the inherent simplicity of Bézier patches, particularly when employed to describe very complex geometries. Indeed, it is necessary to perform some careful operations on the CAD model, such as surfaces division, in order to attain a proper geometrical definition with patches presenting a reduced degree (less or equal to six).

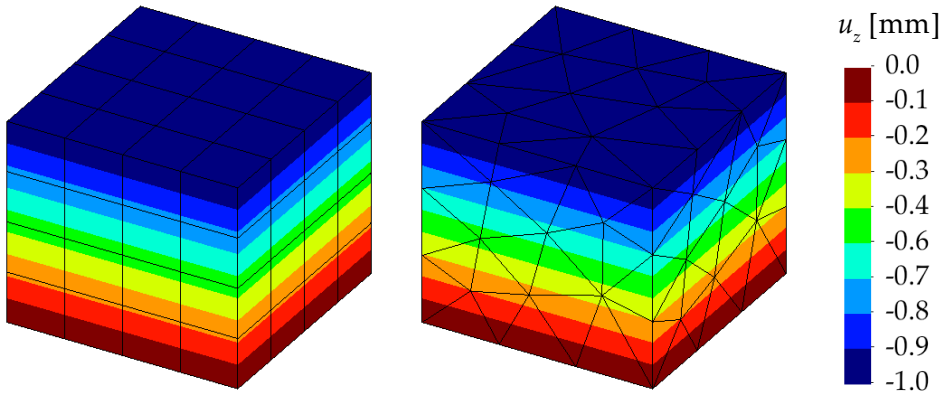


Figure 3.14. Vertical displacement distribution using linear finite elements to describe the cube: (a) 8-node tri-linear hexahedral; (b) 4-node linear tetrahedral.

A simple example of a frictionless contact between a cube (deformable body) and a rigid flat surface is presented. Since the rigid surface is flat, any surface description method provides the same level of accuracy (exact geometrical definition). The deformable body is discretized with both linear and quadratic finite elements, in order to evaluate the effectiveness of the NTS contact discretization for higher order elements. The dimension of each cube edge is 10 mm and the material properties are $E = 100\,\text{MPa}$ and $\nu = 0.3$ (similar to the previous section). The lower surface and the two perpendicular lateral surfaces of the cube are subject to symmetry conditions. The rigid contact surface is pressed against the upper surface of the cube through a vertical displacement of 1 mm (single

increment). The predicted vertical displacement distribution in the cube using both 8-node hexahedral and 4-node tetrahedral low order finite elements, combined with SRI, is shown in Figure 3.14. The obtained numerical results are identical for both types of finite elements and the contact surface is kept horizontal flat with all slave nodes in contact with the rigid surface.

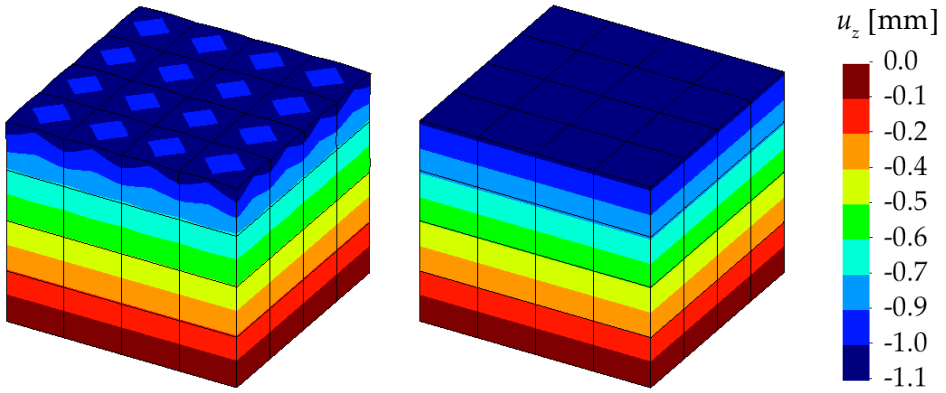


Figure 3.15. Vertical displacement distribution using quadratic finite elements to describe the cube: (a) 20-node Serendipity hexahedral; (b) 27-node tri-quadratic hexahedral.

In general, the NTS contact discretization is used in conjunction with lower order finite elements [Crisfield 00]. This selection is motived by the simplicity of the formulation and due to the non-uniform distribution of nodal forces associated with higher order shape functions in presence of a uniform pressure (1/4/1 distribution for a line with quadratic shape functions). This behaviour present in the quadratic elements is particularly adverse in contact between two deformable bodies, leading to some difficulties to pass the contact patch test [El-Abbasi 01], [Crisfield 00]. Figure 3.15 presents the vertical displacement distribution in the cube predicted by both the 20-node serendipity hexahedral and the 27-node tri-quadratic hexahedral, which are combined with FI. The Serendipity hexahedral finite element leads to an irregular distortion of the mesh close to the contacting surface, as shown in Figure 3.15 (a). The resulting deformed contact surface is not flat, presenting some penetration into the rigid surface. Indeed, only the mid-side nodes have the vertical displacement imposed by the rigid surface, where the nodal contact forces arise. Since the impenetrability conditions are enforced only at the nodes, this type of element is inappropriate for contact problems, when the NTS contact discretization is adopted. On the other hand, the tri-quadratic hexahedral finite elements deliver the same numerical results previously obtained with the linear elements (hexahedral and tetrahedral), as can be see comparing Figure 3.14 and Figure 3.15 (b). However, the application of a uniform pressure (present example) leads to a non-uniform distribution of the nodal contact forces, creating some difficulties when they are transmitted to the master surface in the case of contact

between two deformable bodies. Besides, the computational cost associated with the quadratic elements is considerably higher than using linear elements.

3.4. Contact search algorithm

The contact search is the step preceding the creation of the contact elements, which aims to determine the contacting pairs on the discretized surfaces. Consequently, the contact search algorithm is strongly connected with the contact surface discretization scheme, as well as the type of contact, i.e. contact between two separated bodies or self-contact (unknown master slave pair). Moreover, its robustness determines the accuracy of the entire resolution scheme, highlighting its importance to the numerical treatment of contact problems. Nowadays, the complexity demanded for the computational contact problems includes large deformation, large sliding and self-contact, which entailed the development of efficient contact search algorithms for large size models [Yang 08]. Since the contact areas can change considerably during the deformation process, the contact search has to be performed in every time step of the numerical simulation, presenting a bottleneck for an efficient treatment of contact problems [Yastrebov 13]. On the other hand, in case of small deformations (Node-to-Node discretization) the contact pairs of nodes are established at the beginning and do not change during the solution steps (no slip occurs), requiring only one execution of the search procedure [Alart 91]. Nevertheless, the contact search algorithms presented in this dissertation are specifically developed for the NTS discretization, which allows to deal with large deformation and large sliding problems.

The contact search algorithms are typically decomposed into two distinct phases [Wriggers 06]: (i) global search and (ii) local search. The global contact search procedure is related to purely geometrical considerations and its purpose is to determine potential contacting bodies/surfaces. A hierarchical structure is created to find out which contact surfaces are able to come into contact in a given time step [Zhong 93]. The aim of the local contact search process is to find for each slave node the closest point on the associated master segment, defining the contact element [Areias 04]. The contact search in the implicit integration scheme requires that the possible penetration has to be known at the beginning of the time step, in order to incorporate the contact residual vector and the stiffness matrix in the resolution stage [Aragón 13].

3.4.1. Global contact search

Several different methods have been developed and applied to identify all possible candidate contact partners. In case of NTS discretization (Section 3.3.1), each contact element is composed by a slave node and a master segment, which is the closest element determined based on the normal projection of the slave node on it. The simplest search

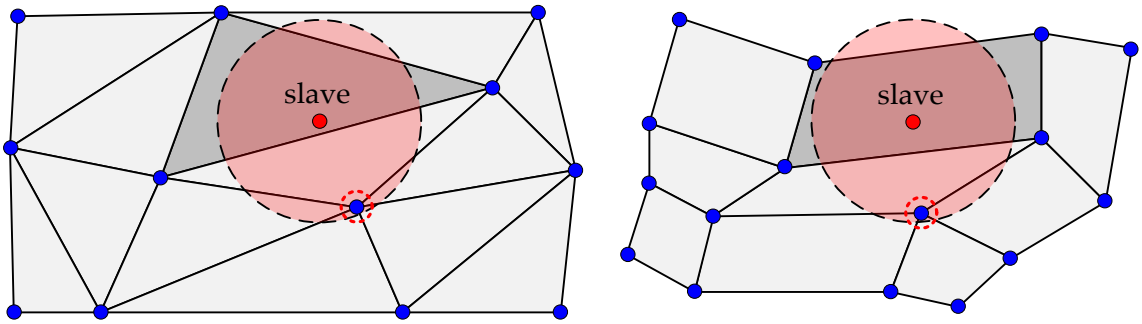
algorithm is the all-to-all detection, sometimes referred to as the "brute force" approach. In that case, all slave nodes are projected on all master segments in order to determine for each slave node the closest master segment with existent projection. However, the number of operations required in this approach is of the order $O(n_s n_m)$ for a two surface contact problem, where n_s and n_m are the numbers of slave nodes and master segments, respectively. Since the evaluation of the projection point requires the solution of a nonlinear system of equations for any master surface in 3D (except surfaces formed by faceted triangular finite elements), this approach leads to unsupportable computational cost for most situations [Yastrebov 13]. Nevertheless, the global search is highly dependent of the contact surface description, namely the regular or irregular finite element mesh discretization. In fact, the global search is quite straightforward and efficient using a regular mesh for the contact definition [Santos 95], while the irregular discretization of the contact surfaces (general case) leads to complex and less efficient contact search algorithms. The most efficient global contact search algorithms are the Hierarchy-Territory algorithm (HITA) [Zhong 89], [Zhong 93], the typical bucket sorting algorithm [Benson 90] and the position code algorithm [Oldenburg 94].

In the contact search algorithm proposed by Zhong [Zhong 93] the contact system is decomposed into four contact hierarchies: bodies, surfaces, segments and nodes. The search procedure is based in the definition of a territory for each contact element (cubic box containing the element). The territories at the same hierarchical level are compared to detect common territory, where the lower levels of hierarchy are tested. A node and a segment are a candidate pair for contact when the distance between the node and the segment is sufficiently small. The main idea of the bucket sorting algorithm proposed by Benson and Hallquist [Benson 90] is the subdivision of the space into equally cells or buckets, assigning each node/segment of the contact surface to a bucket number. Then, the list of segments in each bucket is used to find the candidate master segments for each slave node, allowing to reduce locally the area of search (inside the bucket). In the position code algorithm, developed by Oldenburg and Nilsson [Oldenburg 94], each master segment is checked for the presence of slave nodes situated within the segment territory, which is defined by the smallest cubic box holding the contact territory. The algorithm to detect the contact nodes within the segment territories is based on sorting and searching in one dimension. The mapping from three dimensions to one dimension is achieved by the definition of a discrete position code, which is the number of the box resulting from the division of the space into cubic boxes. The study conducted by Oldenburg and Nilsson [Oldenburg 94] concluded that the HITA and the position code algorithms are superior in terms of computational efficiency. While the bucket sorting algorithm performs sorting and searching in three dimensions in a nested manner, the position code reduces the three-dimensional space into a one dimensional searching problem, hence decreasing the number of operations involved, due to the binary search procedure used [Fujun 00].

The finite element analysis of contact problems is usually divided in two categories: contact between deformable bodies and contact of a deformable body with a rigid surface, both considered in this study. Therefore, two distinct global contact search algorithms are developed and presented in detail, which follow the structure previously implemented in DD3IMP [Oliveira 03]. They are based on the concept of contact hierarchies suggested by Zhong [Zhong 93], exploring the particular features of the contact surface description method adopted. Besides, the global search algorithm is extended to deal with contact of one surface with itself (self-contact), which is more complicated than the standard contact detection between two different surfaces.

3.4.1.1. Global search between a deformable body and a rigid surface

The starting point for the global search algorithm is the creation of contact pairs, each one containing a slave surface and one or more master surfaces. This step is performed by the user at the beginning, where the nodes of the deformable body on the slave surface are defined as slave nodes, while all rigid surfaces involved in the contact problem are automatically specified as master surfaces. Therefore, each slave node is associated with a set of rigid surfaces, allowing the relative sliding of the slave node over the master surfaces (Node-to-Surface contact discretization). The surface description method adopted throughout this thesis to describe the rigid contact surfaces is the finite element mesh. The surface mesh can be composed either by triangular or quadrilateral finite elements, as well as a combination of both. Besides, unstructured meshes are allowed, which are more suitable to deal with complex surface geometries. In fact, they are more attractive from the user point of view and are usually adopted in the automatic mesh generation algorithms, which are essential to deal with industrial problems. The amount of finite elements required to accurately define each rigid contact surface is highly variable, depending both on the dimension and the geometric complexity of the model. The number of finite elements needed in complex industrial problems can reach several thousand, as it will be shown in a numerical example presented in Chapter 6. Furthermore, the adoption of an unstructured mesh significantly influences the performance of the contact detection algorithm, leading to more effort in the global search due to the irregular distribution of the finite elements.



(a) (b)

Figure 3.16. Examples of failure of the global search algorithm based on the closest master node: (a) triangular master mesh; (b) quadrilateral master mesh.

The proposed global contact search algorithm starts with the evaluation of the distance between each slave node and every master node of the associated rigid surfaces. The procedure is divided into two stages: (i) searching the closest master node for each slave node and (ii) selection of the master segments that have the closest master node as one of their vertices using the mesh topology. However, the assumption that the slave node is in contact with the master segment sharing the closest master node is not always correct, as shown in Figure 3.16. Thus, this detection method may fail for unstructured meshes, allowing the slave node to penetrate master segments not attached to the closest master node. Figure 3.16 presents two simple examples of flat rigid contact surfaces described by irregular meshes, highlighting a slave node that does not contact a master segment connected to the closest master node.

Although detection methods based on the search of the closest master node and adjacent segments have been widely used [Benson 90], they are known for not being robust. Thus, the algorithm proposed in [Benson 90] is improved to overcome such limitation. The suggested algorithm performs the detection of a set of closest nodes instead of the closest node, where the amount of selected master nodes is dictated by the topology of the surface mesh (dictated by the presence of distorted finite elements). The number of closest master nodes for each slave node is given by:

$$n_{\mathrm{n}} = 5 + \mathrm{int}\left(\frac{r_{\mathrm{m}}{}^{2}}{65}\right), \tag{3.25}$$

where r_{m} is the maximum value of the ratio between the maximum and minimum finite element side length of the master surface. The problem illustrated in Figure 3.16 occurs mainly in unstructured surface meshes. In such cases, the number of master nodes provided by (3.25) will increase due to the presence of highly distorted finite elements and subsequent heterogeneous disposition. The minimum number of master nodes selected for each slave node is 5 (also used in structured meshes), while the maximum is limited to 50

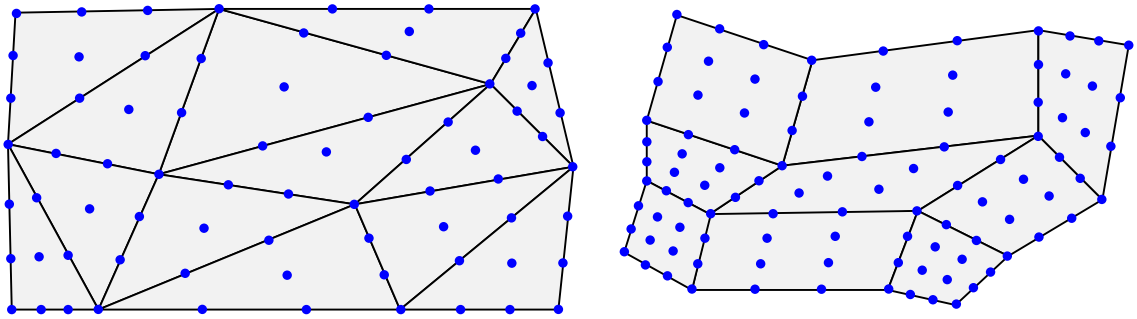
in order to control the computational cost of the contact search. The second stage of this procedure is the identification of the finite elements surrounding the closest master nodes, easily defined through the inverse mesh connectivity. However, since the resulting number of finite elements is extremely dependent on the mesh distortion, it can be computationally inefficient to directly apply the local contact search algorithm to all this preselected surfaces.

The second step of the global search algorithm consists in creating a grid of points on each surface finite element, which is posteriorly used to reduce the number of elements tested in the local search. The grid is obtained through the parameterization of the finite element by the natural coordinates in both directions. The number of uniform grid divisions (number of grid points less one) in each parametric direction is defined as:

$$n_{\mathrm{gd}} = \max\left\{2, \mathrm{int}\left(\frac{r_{\mathrm{m}}}{10}\right)\right\}, \tag{3.26}$$

where r_{m} is the maximum ratio between the maximum and minimum finite element side length of the contact surface discretization. The minimum number of grid divisions in each parametric direction is established as two, which represents 9 grid points for quadrilateral finite elements and 6 for triangular elements. Since the grid dimension increases quadratically with the number of divisions, the maximum value allowed in each direction is limited to 6 to control the computational cost of the global contact search. The grid of points created on each finite element for the master surfaces shown in Figure 3.16 is presented in Figure 3.17, for three divisions in each parametric direction. The number of grid points is particularly important for highly distorted finite elements and near of abrupt transitions between large and small elements. The grid is created only once, at the beginning of the numerical simulation, and posteriorly the coordinates of the points are updated appropriately in each increment. The expressions presented in (3.25) and (3.26) are the result of an extensive study using different structured and unstructured contact surface meshes with different levels of complexity, allowing to obtain a robust algorithm.

The last step comprises the selection of the ten closest finite elements for each slave node through the distance between the slave node and the grid of points created on the finite elements selected in the first step of the proposed algorithm. Besides the ten finite elements, the local coordinates of the closest grid point on the master segment are stored, which will be used in the local contact search. The sorting of the selected finite elements is not important since all of them will be tested during the local contact search procedure.



(a) (b)

Figure 3.17. Examples of grid of points created on the master surface: (a) triangular master mesh; (b) quadrilateral master mesh.

3.4.1.2. Global search between deformable bodies

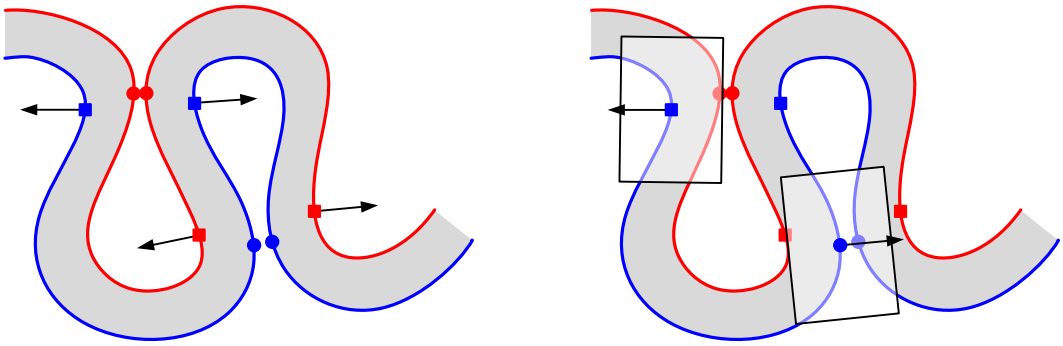
In case of contact problems between separated deformable bodies, the surface portions of the bodies which come into contact are defined by the user at the beginning. The contacting surface pairs are created based on the expected interaction between the bodies. The adoption of the single-pass NTS algorithm to deal with the contact surface discretization requires, for each contact pair, the identification of one contact surface as slave and the other one as master, following the guidelines presented in Section 3.3.1.1. On the other hand, for the two-pass NTS algorithm the classification of the master and slave surfaces is arbitrary due to its exchange in the second pass. In any case, the nodes located on the slave surface are defined as slave nodes, as in the case of a deformable body in contact against a rigid surface. The discretized master contact surface is defined from the assembly of solid element faces looking outside the body, which are parameterized by two natural coordinates. For deformable bodies discretized with linear solid finite elements, the contact faces are defined by quadrilateral elements in meshes composed by 8-node hexahedral elements and triangular elements in meshes composed by 4-node tetrahedral elements.

Since the accuracy of the numerical solution is highly influenced by the quality of the finite element mesh, typically the discretization of the deformable body does not involves distorted finite elements. In fact, the extremely distorted finite elements are allowed in the description of the rigid contact surfaces because no calculations over the rigid body are performed (master nodes do not have any degrees of freedom). Therefore, the global search algorithm previously presented is modified to perform more efficiently the contact detection between two deformable contact surfaces. Indeed, the algorithm is based on the search of the closest master node and subsequent selection of the master elements that share the master node, as proposed in [Benson 90]. As already mentioned, this strategy sometimes fails in the presence of distorted finite elements, as shown in Figure 3.16. Nevertheless, an occasional failure of an isolated node is usually not catastrophic and the two-pass contact

algorithm diminishes the effects of a contact search failure [Benson 90], [Yastrebov 13]. Besides, the adoption of regular meshes allows to obtain a reliable solution, reducing the failure risk to almost zero. The global contact search algorithm evaluates the distance between each slave node and every master node in order to choose the master node with the minimum distance (see Figure 3.16). Then, all finite elements neighbouring to the previously determined master node are used in the local contact search algorithm. Note that the overall cost of the global search doubles using the two-pass NTS contact algorithm, due to the exchange of the master and slave surfaces.

For some mechanical problems the *a priori* assignment of the contact pairs is impossible or presents a big challenge, for example in the analysis of self-contact problems. Although the portions of one surface folded over onto itself are unknown *a priori*, the contact of a surface with itself should be detected during the contact search procedure. Therefore, the above described global search algorithm is enhanced to deal with this class of contact problems. The bucket sorting method is frequently used to detect this type of contact, which is known as the single surface contact algorithm [Benson 90]. Nevertheless, the global search algorithm developed in the present study for problems with unknown *a priori* master-slave discretization is based in geometric informations (i.e. curvature criterion) [Yang 08b], given particular attention to self-contact problems. Since the contact pair is specified by a single contact surface instead of two different surfaces, only the two-pass NTS contact formulation is allowed in self-contact problems. Thus, each node of the surface acts as a slave node and every finite element face composing the surface is defined as master surface.

The self-contact problem occurs generally in the post-buckling of thin walled structures, for which one or two dimensions are much smaller than the others. This leads to some difficulties in the contact search due to the contact with the reverse side, as shown in Figure 3.18 (a). The schematic example selected to highlight the difficulties is composed by a thin solid with two sided contact zones (red and blue in Figure 3.18), which can be treated independently. In fact, adopting the previously described contact search algorithm, it is predicted that the square points are penetrating under the surface with the same colour due to the proximity and opposite surface normal vectors. The solution proposed by Benson and Hallquist [Benson 90] for this problem consists in creating an additional history variable to keep track of which side of the contact surface has been penetrated (slave node is coming from the *exterior* or the *interior*). However, the solution adopted in this study takes into account the normal vectors at the nodes to determine potential contacting surfaces, defining a maximal detection distance smaller than the wall thickness (see Figure 3.18 (b)). The evaluation of the normal vector for each contact node is discussed in detail in Chapter 4.



(a) (b)

Figure 3.18. Example of self-contact in a thin walled structure: (a) difficulties in the contact with the reverse side; (b) proposed global search applied in 2 points.

The first step of the derived global contact search involves the determination of the normal vector (outward normal at the surface of the solid) for each contact node located on the (single) contact surface. Note that these normal vectors will also be used in the surface smoothing method presented in next chapter. Since adjacent nodes present approximately the same normal vector orientation, the comparison between normal vectors of different nodes allows eliminating from the analysis the nodes immediately adjacent. Moreover, since the outward normal vectors of the contacting surfaces at the contact point are opposed (pointing towards each other), only the nodes with opposite normal vectors are considered. Therefore, the normal vector of each contact node is compared with the normal vector of all other contact nodes in order to exclude from the global search the nodes with an angle between normal vectors inferior to 120°. The next step comprises the evaluation of the distance to the previously selected contact nodes in order to identify the closest node. This node should be located within a cylinder with its axis coincident with the normal vector of the contact node and its centre on the node, as shown in Figure 3.18 (b). The diameter of the cylinder is dictated by the biggest diagonal of the contact finite element mesh and the height is given by the thickness of the structure. These dimensions allow to avoid the selection of nodes in the reverse side and guarantees that nodes coming from the exterior direction are checked accurately. However, the maximum step size have to be appropriately selected to avoid penetrations larger than the wall thickness, defining the principal drawback of this strategy [Benson 90]. Nevertheless, the availability of a reliable contact detection method for self-contact problems extends significantly the capacities of DD3IMP finite element code, as shown through the numerical examples presented in Chapter 6.

Since the contact surface geometry can change considerably during the simulation of contact problems which include large sliding (both in case of known *a priori* and unknown master slave surface), the contact search is required almost in each time step. Although the implicit finite element codes allow to use large time steps, the contact search task can represent a significant computational cost. The approach proposed by Benson and

Hallquist [Benson 90] to reduce the computational time associated with contact detection consists in two stages: (i) search for the closest master node to each slave node only once (first time step); and for the following steps (ii) selection of the closest master node among the nodes belonging to the master segments determined in the previous step. However, this technique is not very general and sometimes fails [Yastrebov 13]. Moreover, cannot be used for self-contact problems because it traces the slave nodes that are sliding on the contact surface. Therefore, the global contact above described search algorithm is repeated in each time step, i.e. the set of master segments selected for each slave node are updated at each time step, following always the same procedure. This global search algorithm always works but it is clearly also more expensive.

3.4.2. Local contact search

The purpose of the local contact search algorithm is to find the point on the master surface closest to each slave node. In order to reduce the computational time associated to the contact detection, the local search is performed only within the set of master segments identified in the global contact search. However, the computational cost of the local search is typically higher than the one required by the global search procedure, because the local search is repeated in each iteration of every time step. The methods commonly adopted in the local search are the closest point projection algorithm [Hallquist 85], [Konyukhov 08], the pinball algorithm proposed by Belytschko and Neal [Belytschko 91] and the inside–outside algorithm suggested by Wang and Nakamachi [Wang 97]. For the closest point projection procedure, the minimum distance between each slave node and the master surface is calculated based on the normal projection of the node onto the surface. The Newton–Raphson method is typically used to solve this problem and find the contact point coordinates. On the other hand, the pinball algorithm is very efficient when combined with the penalty method, since it is based on simple checks that eliminate any iterative procedure. However, some inaccuracies concerning the real geometry of the contacting bodies are introduced, since the penetration between their surfaces is assumed as the interpenetration of two spherical balls. The inside–outside algorithm is based in the status of the projected point of the slave node along the mesh normal direction. Only two states are allowed, i.e. the projection point is located either inside or outside the master segment. This algorithm is fast, robust and does not requires any iterative procedure to perform the search (closed-form expression), when faceted finite elements are used in the contact surfaces description [Wang 97]. The method adopted in this study to perform the local search is the normal projection algorithm [Hallquist 85], [Konyukhov 08], since it is more accurate and adequate for smooth contact surfaces, which are the purpose of this dissertation.

3.4.2.1. Closest point projection

The normal gap previously defined in (2.30) is strongly connected with the closest point projection used in the local contact detection procedure. The value of penetration is measured as the closest distance from the slave node $\mathbf{x}^{\mathrm{s}}$ to the master surface $\mathbf{x}^{\mathrm{m}}$, which is parameterized by the surface coordinates $\boldsymbol{\xi} = (\xi^1, \xi^2)$, as shown in Figure 2.5. This leads to the following minimization problem:

$$d = \left\| \mathbf{x}^{\mathrm{s}} - \mathbf{x}^{\mathrm{m}}(\xi^1, \xi^2) \right\| \rightarrow \min, \tag{3.27}$$

which is frequently solved numerically using the Newton–Raphson method. Some difficulties arise in this procedure related to the uniqueness and existence of the closest point projection (see Figure 2.9), which are analysed by Konyukhov and Schweizerhof [Konyukhov 08]. These problems arise mainly when the contact surfaces are modelled by low order finite elements, leading to a piecewise bilinear representation of the surfaces [Hallquist 85], [Heegaard 96]. In fact, the smoothness of the master surface is necessary condition, although a not sufficient for the existence of the normal projection point [Yastrebov 13].

The aim of the closest point projection is to find, for each slave node, the point belonging to the contact master surface that is closest to the slave node, as shown in Figure 2.7. The coordinates of a generic slave node $\mathbf{x}^{\mathrm{s}}$ can be correlated with a vector describing any point on the master surface $\mathbf{x}^{\mathrm{m}}$ through the normal gap value, as given by the following equation:

$$\mathbf{F}^{\mathrm{proj}}(\xi^1, \xi^2, g_{\mathrm{n}}) = \mathbf{x}^{\mathrm{m}}(\xi^1, \xi^2) + g_{\mathrm{n}} \mathbf{n}(\xi^1, \xi^2) - \mathbf{x}^{\mathrm{s}}, \tag{3.28}$$

where the unit normal vector of the master surface is defined in (2.28) and the value of the normal gap function (2.30) is the third coordinate of the surface coordinate system [Konyukhov 05]. The solution of the minimization problem (3.27) is obtained from the solution of $\mathbf{F}^{\mathrm{proj}} = \mathbf{0}$ and provides the coordinates of the contact point, while assuring that the vector connecting this point with the slave node is collinear with the normal vector (see Figure 2.7 and Eq. (3.28)). The Newton–Raphson method is used for solving the nonlinear system of equations, which can be summarized as follows:

$$\mathbf{s}_{i+1} = \mathbf{s}_i - \left[\nabla \mathbf{F}^{\mathrm{proj}}(\mathbf{s}_i) \right]^{-1} \mathbf{F}^{\mathrm{proj}}(\mathbf{s}_i), \tag{3.29}$$

where $\mathbf{s}_i = \left[\xi^1, \xi^2, g_{\mathrm{n}} \right]_i^{\mathrm{T}}$ contains the solution vector at iteration i. Note that the solution yields simultaneously the normal gap g_{n} and the local coordinates of the contact point on the master segment $\bar{\boldsymbol{\xi}} = (\bar{\xi}^1, \bar{\xi}^2)$ [Heege 96]. In case of contact with a rigid master surface, the initial solution is given by the closest grid point found in the global contact search (see

Section 3.4.1.1). On the other hand, in case of contact between deformable bodies, the midpoint of the master contact segment is the initial guess selected for the iterative procedure. The maximum number of iterations allowed is limited to 10 in order to reduce the computational cost. Nevertheless, the normal projection expressed by Eq. (3.28) is applied to each of the ten master segments selected in the global contact search to determine the correct segment. In case of multiple solutions the algorithm selects the segment with minimum normal distance [Oliveira 03].

In order to employ the Newton–Raphson method to solve (3.28), it is necessary to determine the Jacobian matrix of the system of equations, at any point, which is defined as follows:

$$\nabla \mathbf{F}^{\text{proj}} = \left[\frac{\partial \mathbf{F}^{\text{proj}}}{\partial \xi^1}, \frac{\partial \mathbf{F}^{\text{proj}}}{\partial \xi^1}, \frac{\partial \mathbf{F}^{\text{proj}}}{\partial g_{\text{n}}} \right], \tag{3.30}$$

where covariant basis vectors $\boldsymbol{\tau}_{\alpha}^{\text{m}}(\boldsymbol{\xi})$ defined in (2.27) and the gradient of the normal vector, with respect to the local coordinates, are involved in the formulation of the Jacobian matrix. The derivatives of the unit normal vector can be calculated directly using the Weingarten formula [Heege 96], [Konyukhov 05]:

$$\frac{\partial \mathbf{n}}{\partial \xi^{\alpha}} = -b_{\alpha\beta}(\xi^1, \xi^2) m^{\beta\gamma}(\xi^1, \xi^2) \boldsymbol{\tau}_{\gamma}^{\text{m}}(\xi^1, \xi^2), \quad \alpha, \beta, \gamma = 1, 2, \tag{3.31}$$

where $b_{\alpha\beta}$ denotes the covariant components of the symmetric curvature tensor and $m^{\beta\gamma}$ expresses the contravariant components of the metric tensor, both defined in (2.38). The full properties of the master surface are defined by two fundamental tensors: the metrics tensor (first fundamental tensor) and the curvature tensor (second fundamental tensor) [Konyukhov 08].

The finite element approximation of the master surface leads to a non-smooth surface representation between finite elements, which leads to greater difficulties to solve large slip contact problems [Pietrzak 97]. This problem arises mostly in the bilinear parametric representations, but also in high order finite elements. In fact, this situation leads several difficulties in finding the projection of the slave node on the master segment. Each master segment presents its "normal projection" zone, as shown in Figure 3.19, where the slave node can have at least one projection onto the master surface [Yastrebov 13]. However, sometimes the assembly of the "normal projection" zones does not fill the neighbouring space completely, creating *deadzones* where no normal projection exists. If several projections are found (see Figure 3.19), the projection point with minimum normal gap is selected to create a contact element [Heege 96]. Nevertheless, when no projection is found, as shown in Figure 3.19, serious numerical problems may arise [Zavarise 09b]. Two types of blind spots can be distinguished: internal and external. Slave nodes situated in external

blind spots are not detected before they penetrate under the master surface, as shown in Figure 3.19 (a). On the other hand, in the presence of internal blind spots the contact is predicted correctly (before penetration), but if the slave node penetrates under the master surface it is not detected, as shown in Figure 3.19 (b).

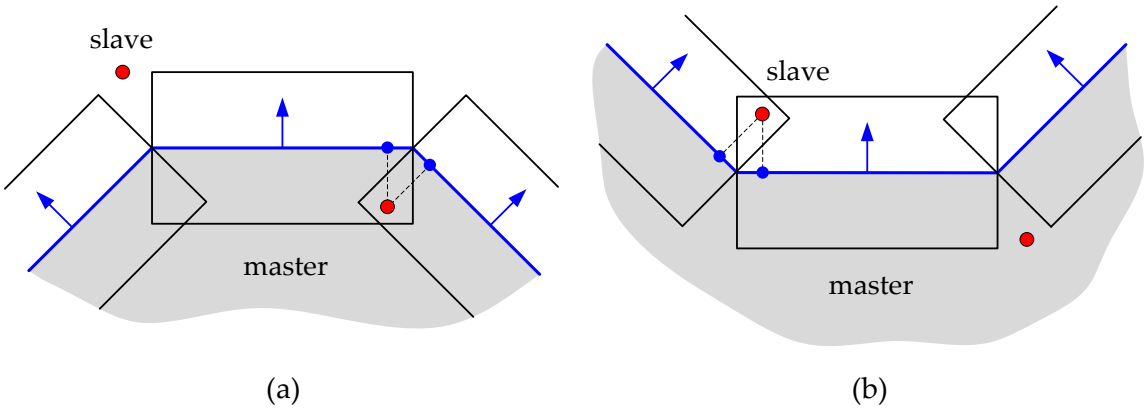


Figure 3.19. Example of a slave node near a sharp corner/valley: (a) convex master surface; (b) concave master surface.

In brief, the finite element discretization of the master surface leads to discontinuities in its normal and tangential vector fields, which cause serious convergence difficulties in the solution of contact problems involving large sliding. The approaches adopted to overcome the above mentioned type of problems (blind spots) are typically divided in two groups: (i) smoothing the master surface (ii) extension of the master segments parameterization outside the standard domain. In the present study the first approach is adopted, avoiding the discontinuities introduced by the discretization procedure and providing a smooth description of the master contact surface. The necessary surface interpolation method is analysed in Chapter 4.

Although the frictionless contact problem only requires the evaluation of the normal gap function (2.30) and surface normal direction to determine the contact force, the frictional contact formulation also needs the tangential relative sliding (2.35) to calculate the frictional (tangential) force. This is due to the nature of the friction force that is path-dependent, thus requiring an incremental update procedure. The change of the closest point projection describes the tangential sliding between the contacting surfaces. In the case of an incremental solution of *quasi*-static frictional contact problems, the tangential relative sliding velocity (2.35) can be replaced directly by the corresponding tangential slip increment [Heege 96]. In order to build the sliding path, the coordinates of the solution point at the last converged configuration are stored and used in the current time step as input parameters. The history variables adopted in the mapping are the global coordinates of the solution point $\mathbf{x}^{m}(\overline{\xi}^{1},\overline{\xi}^{2})$ instead the local coordinates $(\overline{\xi}^{1},\overline{\xi}^{2})$, avoiding the problems related with the sliding of a slave node over several master segments [Krstulović-

Opara 02], [Wriggers 01]. Therefore, the tangential slip increment of a slave node is evaluated using the knowledge of its location relative to the master surface in the previous time step.

The coordinates of the closest point projection in the current time step are defined by ${}^{n+1}\mathbf{x}^{\mathrm{m}}({}^{n+1}\overline{\xi}^1, {}^{n+1}\overline{\xi}^2)$, while in the last converged configuration are denoted as ${}^{n}\mathbf{x}^{\mathrm{m}}({}^{n}\overline{\xi}^1, {}^{n}\overline{\xi}^2)$. Note that the position of the master surface is updated in case of rigid surfaces and it can move and deform in case of deformable master body. The global coordinates of the closest point in the last converged configuration mapped into the current time step are written using the tilde symbol ${}^{n}\tilde{\mathbf{x}}^{\mathrm{m}}({}^{n}\overline{\xi}^1, {}^{n}\overline{\xi}^2)$. The basic idea is to store the local coordinates of the projection point obtained in the last time step to evaluate its global coordinates in the current time step. Thus, the slip increment ${}^{n+1}\mathbf{g}$ is simply defined as the vector connecting the solution point in the last converged configuration mapped into the current time step and the slave node in the current time step, which can be defined as follows:

$$ {}^{n+1}\mathbf{g} = {}^{n+1}\mathbf{x}^{\mathrm{s}} - {}^{n}\tilde{\mathbf{x}}^{\mathrm{m}}({}^{n}\overline{\xi}^1, {}^{n}\overline{\xi}^2), \tag{3.32} $$

where the position of the slave node tends to the master surface during the interactive procedure in order to eliminate the penetration, thus converging to ${}^{n+1}\mathbf{x}^{\mathrm{m}}({}^{n+1}\overline{\xi}^1, {}^{n+1}\overline{\xi}^2)$. However, in case of curved contact surfaces, the slip increment vector is not lying in the tangential plane of the master surface at the current solution point. Hence, the tangential slip increment vector within the current time step is defined as:

$$ {}^{n+1}\mathbf{g}_{\mathrm{t}} = {}^{n+1}\mathbf{g} - ({}^{n+1}\mathbf{g} \cdot {}^{n+1}\mathbf{n})\,{}^{n+1}\mathbf{n}, \tag{3.33} $$

where ${}^{n+1}\mathbf{n}$ denotes the unit normal vector of the master surface at the closest point, evaluated in the current time step. Since the normal vector is calculated through the projection point algorithm (2.28) and it is updated in each equilibrium iteration, the local frame system (normal vector and tangent plane) changes within the iterative procedure. Note that the tangential slip increment vector does not converges to zero at the solution, as happen for the normal gap function. In fact, tangential slip vector provides the direction of the friction force required for the contact slip status.

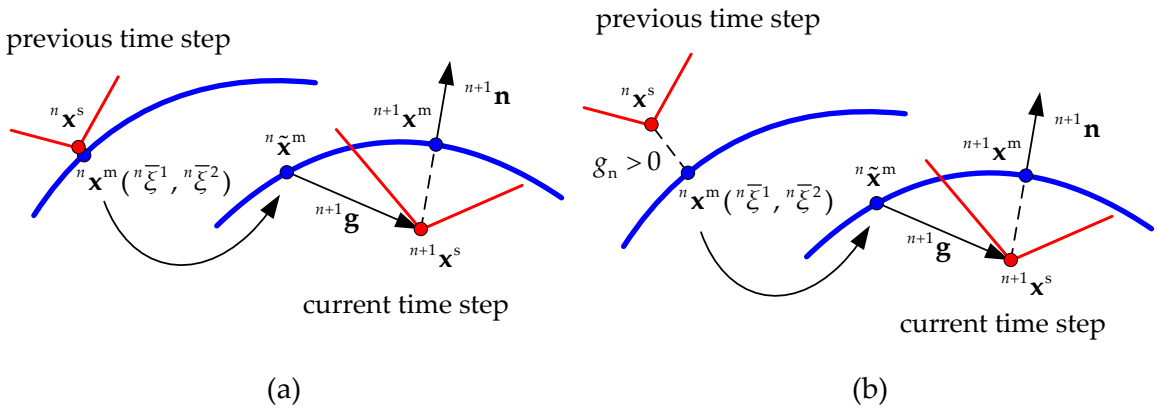


Figure 3.20. Definition of the slip increment vector for: (a) slave node in contact in the previous time step; (b) slave node not in contact in the previous time step.

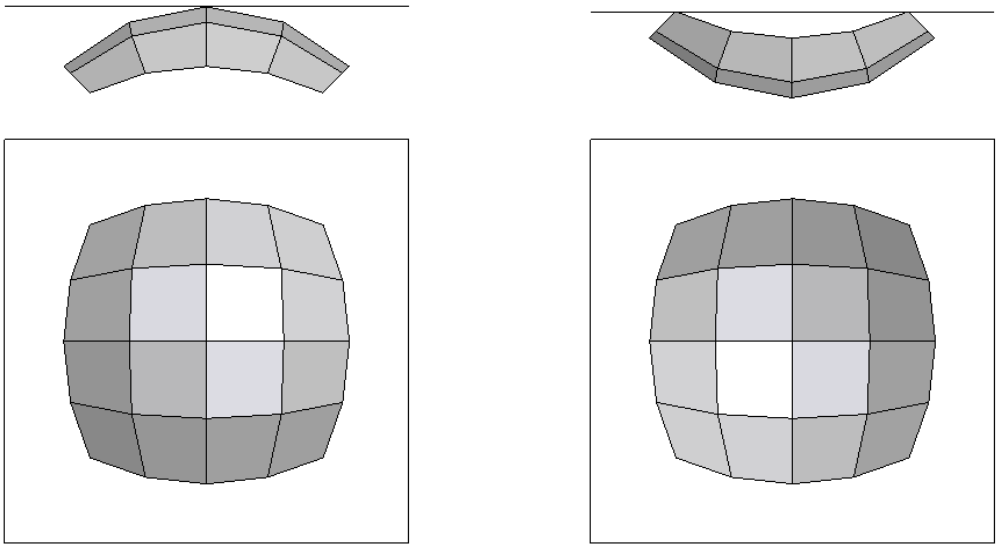
The major problems in the definition of the tangential slip increment vector arise when a slave node comes into contact during the actual time step [Heege 96]. Figure 3.20 presents a scheme illustrating the definition of the slip increment vector for the two possible contact status (contact and gap) at the previous time step. If the slave node is in contact at the previous time step, the slip vector can be determined easily from the slip increment of the slave node relatively to the master surface (Figure 3.20 (a)). On the other hand, when the slave node comes into contact during the current time step (Figure 3.20 (b)), the slip increment vector is defined using the normal projection of the slave node obtained in the previous equilibrium step. This procedure was proposed by [Heege 96] for the contact of a deformable body with a rigid surface and it is also adopted for contact between deformable bodies involving large sliding [Wriggers 01].

3.4.3. Example of normal projection in a spherical surface

The purpose of this section is to highlight the importance of the surface description method used in the numerical simulation of frictional contact problems involving large sliding. The discretization of the master surfaces using low order finite elements leads to sudden changes in the surface normal field, which can cause convergence problems in the solution procedure. Moreover, some blind spots arise in the normal projection of the slave nodes on the faceted master segments (see Figure 3.19), causing severe difficulties in the local search detection procedure. Therefore, a simple example is selected to evaluate the impact of using a non-smooth master surface on the efficiency of the normal projection method, and consequently in the local search procedure.

Two surfaces are involved in the proposed example, a flat surface representing the slave surface and a spherical surface describing the master surface. Besides, two different configurations are analysed, i.e. the master surface can be either convex or concave in relation to the slave surface. In both configurations the master surface is discretized by 16

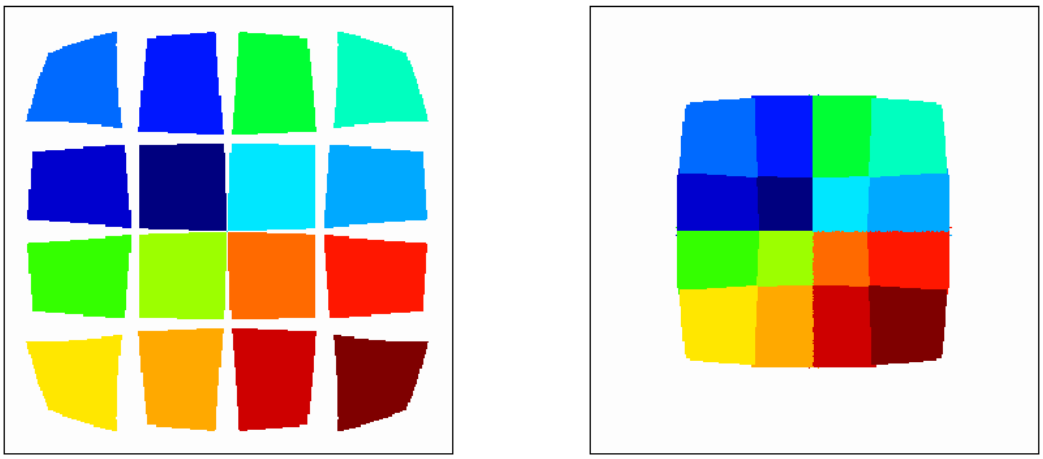
bilinear quadrilateral finite elements, as shown in Figure 3.21. On the other hand, a fine grid of points (300 divisions in each direction) is created over the square flat surface, producing a total of 90,601 points. Then, the aforementioned normal projection algorithm is applied to each of these points in order to calculate its closest point on the master surface. The relative position of the two surfaces is presented in Figure 3.21, where two views of each configuration are shown.



(a) (b)

Figure 3.21. Configuration of the problem composed by a flat surface and a spherical surface (lateral and top views): (a) convex surface; (b) concave surface.

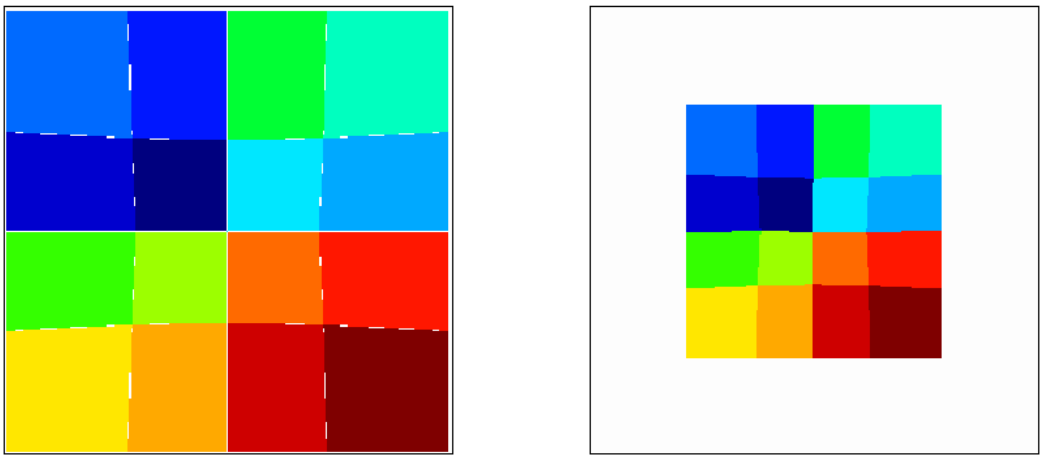
The normal projection of a slave point on the curved master surface described by bilinear quadrilateral finite elements (Figure 3.21) may have multiple solutions (concave surface) or no solution (convex surface) near the common edges of the master finite elements. Figure 3.22 shows the colour map denoting the finite elements on which the slave points are projected with smallest normal gap. Some *deadzones* (white colour) arise in the case of convex surface, which are larger for points located more distant to the surface, due to the pyramidal shape of the blind spots, as illustrated previously in Figure 3.19 (a). Therefore, the points of the slave surface located within the white regions shown in Figure 3.22 do not have any normal projection with any master finite element. On the other hand, the normal projection of the slave points on the concave surface does not comprise any external blind spot, as shown in Figure 3.19 (b), providing a continuous projection field (Figure 3.22 (b)).



(a) (b)

Figure 3.22. Areas of flat surface with normal projection on the non-smoothed spherical surface: (a) convex surface; (b) concave surface. Each colour denotes a different finite element.

The blind spots in the normal projection, created by the finite element discretization of the master surfaces with linear elements (Figure 3.22), can be avoided through the surface smoothing, ensuring a continuous projection on the master surface. This strategy is adopted in the present study to improve the accuracy of the local search detection and avoid convergence problems in large sliding contact problems. The next chapter is entirely dedicated to surface smoothing with Nagata patches. Thus, in order to highlight the advantages of adopting this approach, the finite element discretization of the spherical surface presented in Figure 3.21 is smoothed with Nagata patches. After that, the same contact search procedure performed for the faceted surface description is carried out for the smoothed spherical surface, in order to compare the results of the normal projection. The colour map denoting the patches on which the slave points are projected with smallest normal gap is presented in Figure 3.23. The blind spots observed in Figure 3.22 (a) for the convex surface modelled by faceted elements are strongly reduced using a smoothed master surface. In fact, the areas of the slave surface where no normal projection exist (white colour) are now located in a very narrow range near the edges of the patches. Since the configuration with the concave surface (Figure 3.21 (b)) does not contain any external blind spot, the normal projection field is continuous. Although not shown here, the zones with multiple projections are also drastically reduced. Also, the pattern is slightly different from the one obtained with the non-smoothed spherical surface, as observed by comparing Figure 3.22 (b) with Figure 3.23 (b). Note that both configurations of the selected example originate a square pattern of the normal projection when the spherical surface is smooth, as shown in Figure 3.23. This is also a consequence of the smoothing procedure applied to the surface.



(a) (b)

Figure 3.23. Areas of flat surface with normal projection on the smoothed spherical surface: (a) convex surface; (b) concave surface. Each colour denotes a different patch.

Chapter 4

Surface Smoothing with Nagata Patches

This chapter presents the developed contact surface smoothing method, which is based in the Nagata patch interpolation. The local support of the adopted interpolation method allows to handle both structured and unstructured surface finite element meshes, as well as combining triangular and quadrilateral Nagata patches. The Nagata patch interpolation method is reviewed, showing its shortcomings in terms of surface smoothing application. The combination of the original formulation with a modified one is proposed to obtain a stable and accurate surface smoothing method. The accuracy of the surface smoothing method is evaluated using simple geometries defined by analytical functions (arc, cylinder, sphere and torus), by means of the shape and normal vector error distributions. When the contact surface is assumed rigid, the nodal normal vectors required for the interpolation are calculated using the information contained in the available IGES file, where the surface geometry is represented by trimmed NURBS surfaces. On the other hand, in case of deformable contact surfaces, the nodal normal vectors are approximated using the weighted average of the normal vectors of facets adjacent to the node. Six different weighting factors are presented and their accuracy is assessed using simple and complex geometries. The influence of the nodal normal vectors on the interpolation accuracy is briefly discussed.

4.1. Contact smoothing procedures

The finite element discretization of the contact surfaces may cause two types of difficulties in the solution of contact problems between deformable bodies involving large sliding. The first difficult is associated with the local searching procedure based on the

normal projection method, which can provide multiple solutions and blind spots in the projection domain due to the C^0 continuity of the surface, as discussed in Section 3.4.2. The second difficulty is created by the sudden changes in the surface normal field between adjacent master elements, making the contact surface only piecewise differentiable. Therefore, combining the Node-to-Segment contact discretization technique with faceted master surface description leads to non-physical jumps in the contact forces when a slave node slides over two adjacent master elements. Moreover, these jumps induced by spatial discretization can cause convergence problems and possible divergence of the numerical solution [Wriggers 06]. In order to overcome the above mentioned drawbacks, various surface smoothing procedures have been developed, trying to achieve a smooth description of the master surface, which would render a robust local detection procedure and a reliable convergence of the iterative solution algorithms [Yastrebov 13].

The idea behind the contact smoothing procedures is to define an accurate and continuous contact master surface using higher order interpolations, while the bulk is discretized with the classical linear finite elements. This procedure ensures a continuous projection on the master surface, improving the accuracy obtained in the evaluation of the normal gap function, which is strongly connected with the closest point definition. When the contact surface is itself deformable, the smoothing method needs to be applied in each iteration of a time step in order to evaluate accurately the kinematic contact variables (Section 2.2.1). Therefore, the computational time required by the surface smoothing algorithm should be considerably lower than the entirely time necessary to perform an iteration. The surface smoothing methods adopted in 2D contact problems are typically based in Hermite [Padmanabhan 01], Bézier [Wriggers 01], Spline [El-Abbasi 01], [Al-Dojayli 02] or NURBS [Stadler 03] interpolation of the mesh, attaining at least C^1 continuity in the resulting smoothed master surface. In fact, the technique proposed by Stadler et al. [Stadler 03] allows obtaining an arbitrary level of continuity in the master surface representation.

However, the straightforward extension of such interpolation methods to 3D contact problems is limited to structured quadrilateral finite element meshes [Pietrzak 97], [Pietrzak 99]. In fact, the development of smoothing techniques for deformable contact surfaces discretized with an arbitrary mesh topology is significantly more complex, presenting an actual challenge and research topic in computational contact mechanics [Wriggers 06]. Therefore, only two different approaches are currently available to address this issue. One is the surface smoothing method proposed by Puso and Laursen [Puso 02], which uses Gregory patches to interpolate both structured and unstructured meshes of quadrilateral elements. However, this method only ensures G^1 continuity at the patch boundaries of the surface representation, i.e. the direction of the tangent vector is continuous (not its magnitude). On the other hand, this smoothed contact algorithm is roughly twelve times computationally more expensive than the non-smoothed

implementation. The other approach, proposed by Krstulović-Opara et al. [Krstulović-Opara 02], emploies quartic triangular Bézier patches in the surface smoothing of tetrahedral finite element meshes, using the nodes and the centroid of the exterior triangular elements. This method allows achieving *quasi*-C^1 continuity surfaces, producing a smooth transition between edges of triangular elements except at the nodes. However, since this smoothed contact surface passes through the centroids of the finite elements, which are not necessarily points lying on the original surface, the obtained geometrical accuracy can be unsatisfactory.

Alternative techniques to the classic surface smoothing method have been adopted, such as the subdivision scheme presented by Stadler and Holzapfel [Stadler 04] , which can be applied for both quadrilateral and triangular meshes of arbitrary topology. However, this technique requires special treatment of the nodes where the mesh is unstructured. Another alternative is the smoothing procedure based on a meshfree technique, as proposed by Chamoret et al. [Chamoret 04], which allows deal with hybrid surface meshes (tetrahedral and hexahedral elements). This approach generates a smooth contact surface using a least-squares approximation. A similar approach was suggested by Belytschko et al. [Belytschko 02], where the smoothing is performed implicitly by constructing smooth signed gap functions for the bodies. However, the contact surface provided by these methods does not pass through the nodes of the master surface discretization, which introduces some imprecisions in the geometry of the contacting bodies [Qian 14].

In this dissertation, a general 3D surface smoothing method using the Nagata patch interpolation [Nagata 05] is developed to describe both rigid and deformable contact master surfaces. The idea behind the Nagata patch is the local quadratic interpolation using only the position and surface normal vectors at the nodes of the surface finite element mesh. This interpolation method has been recently successfully applied in the smoothed description of forming tools (rigid contact surfaces) involved in diverse metal forming processes [Hama 08], [Hachani 13]. The Nagata patch formulation is described in the following, including the proposed modifications and improvements. The accuracy of the surface smoothing procedure is evaluated both for analytical and approximated nodal normal vectors, which are required for the interpolation.

4.2. Nagata patch interpolation

The Nagata patch interpolation was recently proposed by Nagata [Nagata 05] for interpolating discretized surfaces in order to recover the original geometry with good accuracy. Its central idea is the quadratic interpolation, requiring only the position and normal vectors at the nodes of the surface mesh. Moreover, it can be applied to general finite element meshes with arbitrary topology. Although the formulation can account for

discontinuity of normal vectors, sharp edges and singular points [Nagata 05], in this study the interpolation method is restricted to smooth contact surfaces.

Figure 4.1 presents a 2D interpolation example applied to an edge, as shown in where $\mathbf{x}_0$ and $\mathbf{x}_1$ denote the position vectors of the edge ends (element nodes). The interpolation of this edge is replaced by a Nagata curve in the form:

$$\mathscr{C}(\xi) = \mathbf{x}_0 + (\mathbf{d} - \mathbf{c})\xi + \mathbf{c}\xi^2, \tag{4.1}$$

where ξ is the local coordinate that satisfies the condition $0 \le \xi \le 1$, and $\mathbf{d} = \mathbf{x}_1 - \mathbf{x}_0$ is the vector joining the end points of the edge. The coefficient vector $\mathbf{c}$, called the curvature parameter, adds the curvature to the edge. Assuming that the Nagata curve (4.1) is orthogonal to the unit normal vectors $\mathbf{n}_0$ and $\mathbf{n}_1$, the curvature parameter $\mathbf{c}$ is determined minimizing its norm, as follows:

$$\mathbf{c}(\mathbf{d}, \mathbf{n}_0, \mathbf{n}_1) = \begin{cases} \dfrac{[\mathbf{n}_0, \mathbf{n}_1]}{1-a^2} \begin{bmatrix} 1 & -a \\ -a & 1 \end{bmatrix} \begin{Bmatrix} \mathbf{n}_0 \cdot \mathbf{d} \\ -\mathbf{n}_1 \cdot \mathbf{d} \end{Bmatrix} & (a \neq \pm 1) \\ \dfrac{[\mathbf{n}_0, \pm \mathbf{n}_0]}{2} \begin{Bmatrix} \mathbf{n}_0 \cdot \mathbf{d} \\ \mp \mathbf{n}_0 \cdot \mathbf{d} \end{Bmatrix} = \mathbf{0} & (a = \pm 1) \end{cases}, \tag{4.2}$$

where $a = \mathbf{n}_0 \cdot \mathbf{n}_1$ denotes the cosine of the angle between the two normal vectors and $[\cdot, \cdot]$ represents a matrix composed by two vectors. When the normal vectors are parallel $(a = \pm 1)$, the curvature parameter vanishes and the Nagata curve degenerates into a linear segment. The first derivative of function defining the Nagata curve is given by:

$$\frac{d\mathscr{C}(\xi)}{d\xi} \equiv \mathscr{C}_\xi(\xi) = \mathbf{d} + (2\xi - 1)\mathbf{c}, \tag{4.3}$$

which is orthogonal to the normal vectors $\mathbf{n}_0$ and $\mathbf{n}_1$ at the end points P_0 and P_1, i.e. satisfies the imposed boundary conditions for $a \neq \pm 1$ (Figure 4.1). The derivative of the curve provides the tangential direction, which is used to evaluate the normal direction at each point on the Nagata curve (always coplanar). Hence, the normal direction of the plane containing the curve is required to calculate the normal vector of a Nagata curve. Since both the curvature parameter (4.2) and the vector joining the end points $\mathbf{d}$ are contained in the plane, its normal direction is obtained by the cross product of these vectors. Therefore, the unit normal vector to a Nagata curve is given by:

$$\mathbf{n}_{\text{Nagata}}(\xi) = \pm \frac{(\mathbf{c} \times \mathbf{d}) \times \mathscr{C}_\xi(\xi)}{\left\| (\mathbf{c} \times \mathbf{d}) \times \mathscr{C}_\xi(\xi) \right\|}, \tag{4.4}$$

where the signal $\pm$ provides the orientation of the normal vector. Note that if the Nagata curve degenerates into a linear segment it is impossible to define a normal vector.

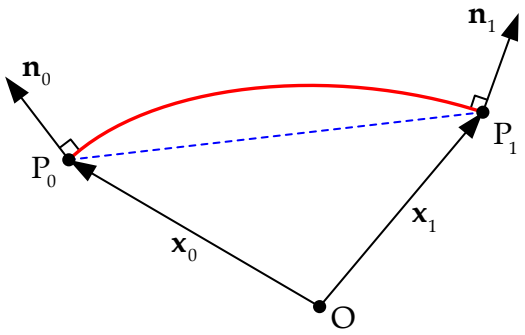


Figure 4.1. Nagata interpolation of a curved segment.

The above interpolation algorithm expressed by Eqs. (4.1) and (4.2) is the basis to apply the Nagata interpolation to general *n*-sided patches, such as triangular and quadrilateral patches. The idea is to first interpolate independently each edge of the finite element through the quadratic curve (4.1). Afterwards, the Nagata patch is defined by its trace on the quadratic curves. In the following section the algorithm is described for triangular and quadrilateral patches, which are the simplest and most commonly used by the mesh generator codes to describe surfaces of arbitrary geometry [Nagata 05].

4.2.1. Triangular and quadrilateral patches

In the case of a triangular finite element, schematically presented in Figure 4.2 (a), the Nagata patch is defined by interpolating each edge of the element using (4.1) and then the interior is filled. The input data necessary in the vertices v_1, v_2 and v_3 are the position vectors $\mathbf{x}_{00}$, $\mathbf{x}_{10}$ and $\mathbf{x}_{11}$, and the unit normal vectors $\mathbf{n}_{00}$, $\mathbf{n}_{10}$ and $\mathbf{n}_{11}$ (see Figure 4.2 (a)). In case of a triangular patch, the interpolated surface is approximated by the following quadratic polynomial:

$$\mathscr{P}^t(\eta,\zeta)=\mathbf{c}_{00}+\mathbf{c}_{10}\eta+\mathbf{c}_{01}\zeta+\mathbf{c}_{11}\eta\zeta+\mathbf{c}_{20}\eta^2+\mathbf{c}_{02}\zeta^2, \tag{4.5}$$

where the patch domain is defined in the local coordinates η and ζ satisfying the condition $0\le\zeta\le\eta\le 1$, as shown in Figure 4.2 (b). The coefficient vectors of the triangular Nagata patch are given by:

$$\begin{aligned}
\mathbf{c}_{00} &= \mathbf{x}_{00},\\
\mathbf{c}_{10} &= \mathbf{x}_{10}-\mathbf{x}_{00}-\mathbf{c}_1,\\
\mathbf{c}_{01} &= \mathbf{x}_{11}-\mathbf{x}_{10}+\mathbf{c}_1-\mathbf{c}_3,\\
\mathbf{c}_{11} &= \mathbf{c}_3-\mathbf{c}_1-\mathbf{c}_2,\\
\mathbf{c}_{20} &= \mathbf{c}_1,\\
\mathbf{c}_{02} &= \mathbf{c}_2,
\end{aligned} \tag{4.6}$$

where $\mathbf{c}_1$, $\mathbf{c}_2$ and $\mathbf{c}_3$ are the curvature parameter defined for the finite element edges $(\mathbf{x}_{00}, \mathbf{x}_{10})$, $(\mathbf{x}_{10}, \mathbf{x}_{11})$ and $(\mathbf{x}_{00}, \mathbf{x}_{11})$, respectively. Thus, these vectors can be determined by applying Eq. (4.2) considering:

$$\begin{aligned} \mathbf{c}_1 &\equiv \mathbf{c}(\mathbf{d}_1, \mathbf{n}_{00}, \mathbf{n}_{10}), \\ \mathbf{c}_2 &\equiv \mathbf{c}(\mathbf{d}_2, \mathbf{n}_{10}, \mathbf{n}_{11}), \\ \mathbf{c}_3 &\equiv \mathbf{c}(\mathbf{d}_3, \mathbf{n}_{00}, \mathbf{n}_{11}). \end{aligned} \tag{4.7}$$

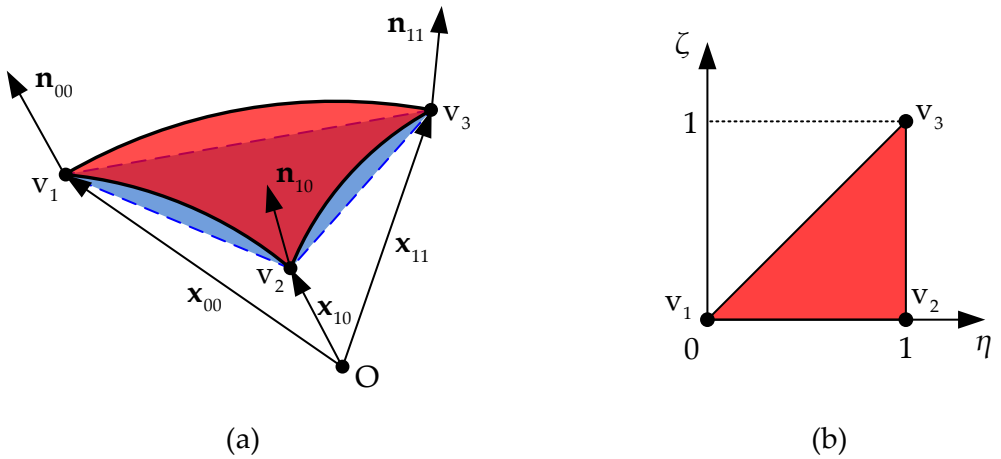
Note that replacing $\mathbf{c}_1$, $\mathbf{c}_2$ and $\mathbf{c}_3$ by zero vectors in (4.6), the faceted triangular element description is recovered as a particular case due to the linear interpolation. The partial derivatives of the triangular Nagata patch (4.5) are given by the following expressions:

$$\begin{aligned} \frac{\partial \mathscr{P}^{\mathrm{t}}(\eta,\zeta)}{\partial \eta} &\equiv \mathscr{P}^{\mathrm{t}}_{\eta}(\eta,\zeta) = \mathbf{c}_{10} + \mathbf{c}_{11}\zeta + 2\mathbf{c}_{20}\eta, \\ \frac{\partial \mathscr{P}^{\mathrm{t}}(\eta,\zeta)}{\partial \zeta} &\equiv \mathscr{P}^{\mathrm{t}}_{\zeta}(\eta,\zeta) = \mathbf{c}_{01} + \mathbf{c}_{11}\eta + 2\mathbf{c}_{02}\zeta, \end{aligned} \tag{4.8}$$

which are required to evaluate the normal vector of a point on the patch. The outward unit normal vector of any point on the patch is defined by cross product of the derivatives, as follows:

$$\mathbf{n}_{\mathrm{Nagata}}(\eta,\zeta) = \frac{\mathscr{P}^{\mathrm{t}}_{\eta}(\eta,\zeta) \times \mathscr{P}^{\mathrm{t}}_{\zeta}(\eta,\zeta)}{\left\| \mathscr{P}^{\mathrm{t}}_{\eta}(\eta,\zeta) \times \mathscr{P}^{\mathrm{t}}_{\zeta}(\eta,\zeta) \right\|}, \tag{4.9}$$

where the orientation is dictated by the ordering of the nodes that define the finite element. In fact, the orientation of the nodal normal vectors has no influence on the patch orientation, while its shape is defined by the direction of the nodal normal vectors.



(a) (b)

Figure 4.2. Triangular Nagata patch interpolation: (a) sketch of the patch with normal vectors; (b) patch domain defined in the local coordinates.

The quadrilateral Nagata patch, schematically illustrated in Figure 4.3 (a), is obtained in a similar way as the triangular patch. The necessary input data for the vertices v_1, v_2, v_3 and v_4 are the position vectors $\mathbf{x}_{00}$, $\mathbf{x}_{10}$, $\mathbf{x}_{11}$ and $\mathbf{x}_{01}$, and the unit normal vectors $\mathbf{n}_{00}$, $\mathbf{n}_{10}$, $\mathbf{n}_{11}$ and $\mathbf{n}_{01}$ (see Figure 4.3 (a)). The nodes defining the finite element do not need to be coplanar. The position vector of any point of the quadrilateral Nagata patch is given by:

$$\mathscr{P}^{\mathrm{q}}(\eta,\zeta)=\mathbf{c}_{00}+\mathbf{c}_{10}\eta+\mathbf{c}_{01}\zeta+\mathbf{c}_{11}\eta\zeta+\mathbf{c}_{20}\eta^2+\mathbf{c}_{02}\zeta^2+\mathbf{c}_{21}\eta^2\zeta+\mathbf{c}_{12}\eta\zeta^2, \tag{4.10}$$

where the patch domain in the local coordinates η and ζ is defined as $0\le\eta,\zeta\le 1$, as shown in Figure 4.3 (b). The coefficient vectors for the quadrilateral Nagata patch are given by:

$$\begin{aligned}
\mathbf{c}_{00}&=\mathbf{x}_{00},\\
\mathbf{c}_{10}&=\mathbf{x}_{10}-\mathbf{x}_{00}-\mathbf{c}_1,\\
\mathbf{c}_{01}&=\mathbf{x}_{01}-\mathbf{x}_{00}-\mathbf{c}_4,\\
\mathbf{c}_{11}&=\mathbf{x}_{11}-\mathbf{x}_{10}-\mathbf{x}_{01}+\mathbf{x}_{00}+\mathbf{c}_1-\mathbf{c}_2-\mathbf{c}_3+\mathbf{c}_4,\\
\mathbf{c}_{20}&=\mathbf{c}_1,\\
\mathbf{c}_{02}&=\mathbf{c}_4,\\
\mathbf{c}_{21}&=\mathbf{c}_3-\mathbf{c}_1,\\
\mathbf{c}_{12}&=\mathbf{c}_2-\mathbf{c}_4,
\end{aligned} \tag{4.11}$$

where $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$ and $\mathbf{c}_4$ are the curvature parameters defined for the finite element edges $(\mathbf{x}_{00},\mathbf{x}_{10})$, $(\mathbf{x}_{10},\mathbf{x}_{11})$, $(\mathbf{x}_{01},\mathbf{x}_{11})$ and $(\mathbf{x}_{00},\mathbf{x}_{01})$, respectively. These vectors are determined by applying Eq. (4.2) to each of these edges considering:

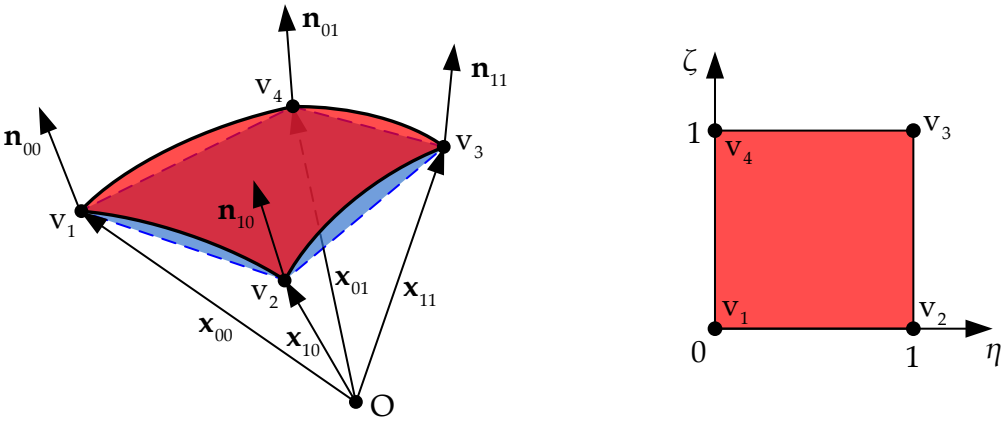
$$\begin{aligned}
\mathbf{c}_1&\equiv\mathbf{c}(\mathbf{d}_1,\mathbf{n}_{00},\mathbf{n}_{10}),\\
\mathbf{c}_2&\equiv\mathbf{c}(\mathbf{d}_2,\mathbf{n}_{10},\mathbf{n}_{11}),\\
\mathbf{c}_3&\equiv\mathbf{c}(\mathbf{d}_3,\mathbf{n}_{01},\mathbf{n}_{11}),\\
\mathbf{c}_4&\equiv\mathbf{c}(\mathbf{d}_4,\mathbf{n}_{00},\mathbf{n}_{01}).
\end{aligned} \tag{4.12}$$

The bilinear parametric representations of the quadrilateral finite elements is recovered as a particular case when the curvature parameters defined in (4.12) are set as zero vectors. The partial derivatives of the quadrilateral Nagata patch (4.10), relative to the local coordinates, are given by the following expressions:

$$\begin{aligned}
\frac{\partial\mathscr{P}^{\mathrm{q}}(\eta,\zeta)}{\partial\eta}&\equiv\mathscr{P}^{\mathrm{q}}_{\eta}(\eta,\zeta)=\mathbf{c}_{10}+\mathbf{c}_{11}\zeta+2\mathbf{c}_{20}\eta+2\mathbf{c}_{21}\eta\zeta+\mathbf{c}_{12}\zeta^2,\\
\frac{\partial\mathscr{P}^{\mathrm{q}}(\eta,\zeta)}{\partial\zeta}&\equiv\mathscr{P}^{\mathrm{q}}_{\zeta}(\eta,\zeta)=\mathbf{c}_{01}+\mathbf{c}_{11}\eta+2\mathbf{c}_{02}\zeta+\mathbf{c}_{21}\eta^2+2\mathbf{c}_{12}\eta\zeta,
\end{aligned} \tag{4.13}$$

which are used to evaluate the outward unit normal vector of any point of the patch, defined by cross product of the derivatives as follows:

$$\mathbf{n}_{\text{Nagata}}(\eta,\zeta)=\frac{\boldsymbol{\mathscr{P}}_{\eta}^{\text{q}}(\eta,\zeta)\times\boldsymbol{\mathscr{P}}_{\zeta}^{\text{q}}(\eta,\zeta)}{\left\|\boldsymbol{\mathscr{P}}_{\eta}^{\text{q}}(\eta,\zeta)\times\boldsymbol{\mathscr{P}}_{\zeta}^{\text{q}}(\eta,\zeta)\right\|}. \tag{4.14}$$



(a) (b)

Figure 4.3. Quadrilateral Nagata patch interpolation: (a) sketch of the patch with normal vectors; (b) patch domain defined in the local coordinates.

4.2.2. Shortcomings and improvements

In order to properly apply the Nagata patch interpolation to smooth discretized contact surfaces, the interpolation algorithm must be simple, efficient and stable. However, the adopted local interpolation method has a stability problem that strongly affects the geometrical accuracy. This drawback arises from the definition of the curvature parameter (4.2), which cannot be calculated when the two normal vectors are parallel (zero denominator) and it is set to be zero, recovering the linear interpolation. Although the problem is solved mathematically, when the denominator is nearly zero, the Nagata interpolation can lead to very sharp surfaces with inverted orientation, which are prohibited for the contact surface description. In fact, due to the quadratic formulation of the interpolation method, a single Nagata curve has no ability to describe a curve with an inflection point, generating a curve very inflated with inversion of orientation near to the node. In order to overcome this stability problem, Boschiroli et al. [Boschiroli 11] proposed to increase the domain of linear interpolation near the singularity, using use a threshold value in the denominator of the curvature parameter definition (4.2). However, since the threshold value is connected with the angle between the normal vectors at the end points of the edge, its value must be set quite high to obtain a stable surface smoothing approach

[Boschiroli 11]. On the other hand, high threshold values lead to linear interpolation in several cases and degradation of the accuracy provided by the quadratic interpolation. A similar approach was first explored in this study to overcome these problems, which is based in the geometric relationships between normal vectors. Nevertheless, it requires the definition of various threshold values by the user, which is unattractive for general application.

In the present study, a modified Nagata patch local interpolation proposed by Sekine and Obikawa [Sekine 10] is analysed and compared with the original formulation. This modified Nagata patch interpolation method, developed for application in the tool path generation, avoids the degradation of geometrical accuracy and inversion of orientation induced from asymmetry. The idea behind this improvement is the creation of an interpolation technique less dependent on the directions of the normal vectors [Sekine 10]. However, the second-order interpolation, the local support and the simplicity are guaranteed in the modified surface representation, preserving its distinctive features.

As shown by [Sekine 10], the curvature parameter of the original formulation, previously defined in (4.2), which is responsible to provide the curvature to each edge of the finite element, can be rewrite in the following form for $a \neq \pm 1$:

$$\mathbf{c}(\mathbf{d}, \mathbf{n}_0, \mathbf{n}_1) = \frac{1}{1+a}(k_0\mathbf{n}_0 + k_1\mathbf{n}_1), \tag{4.15}$$

where the parameters k_0 and k_1 are defined by:

$$\begin{aligned} k_0 &= \frac{\lambda_0 - a\lambda_1}{1-a}, \\ k_1 &= \frac{\lambda_1 - a\lambda_0}{1-a}, \end{aligned} \tag{4.16}$$

and the parameters λ_0 and λ_1 are given by:

$$\begin{aligned} \lambda_0 &= \mathbf{n}_0 \cdot \mathbf{d}, \\ \lambda_1 &= -\mathbf{n}_1 \cdot \mathbf{d}. \end{aligned} \tag{4.17}$$

On the other hand, the curvature parameter $\mathbf{c}'$ of the modified Nagata patch interpolation proposed by [Sekine 10] is given by:

$$\mathbf{c}'(\mathbf{d}, \mathbf{n}_0, \mathbf{n}_1) = \frac{k_0 + k_1}{2(1+a)}(\mathbf{n}_0 + \mathbf{n}_1), \tag{4.18}$$

which is slightly different from the one (4.15) involved in the original formulation. While the Nagata patch interpolation method proposed by [Nagata 05] provides a surface that is orthogonal to the normal vectors given at the nodes (imposed boundary conditions), the modified formulation [Sekine 10] does not satisfy these boundary conditions.

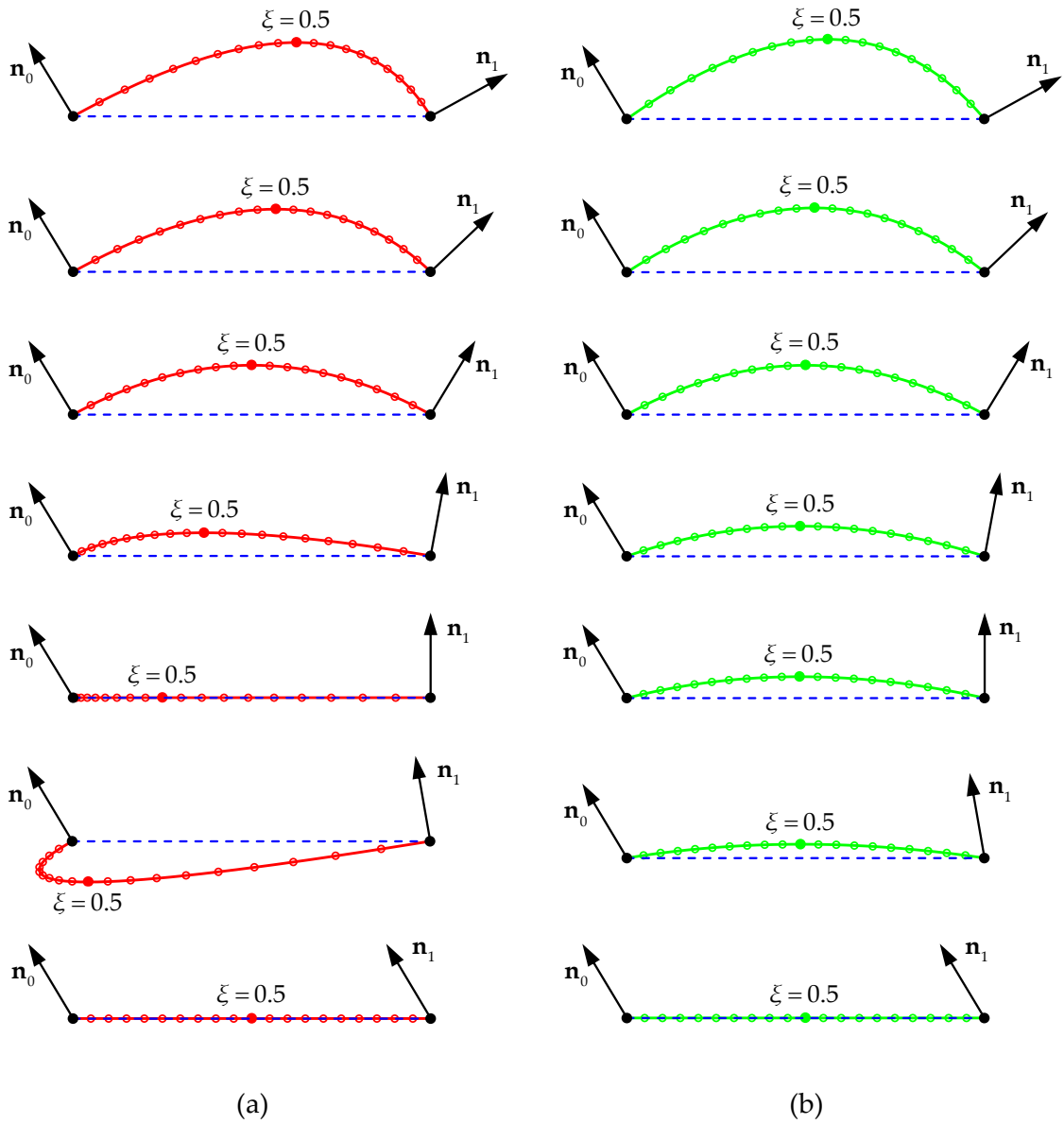


Figure 4.4. Nagata patch interpolation applied to an edge: (a) original formulation; (b) modified formulation.

Considering the simple case of a curve interpolation, the tangent vector of the Nagata curve obtained with modified curvature parameter (4.18) is not orthogonal to the normal vectors defined at the nodes. In fact, the dot product of the Nagata curve derivative defined in (4.3) with the unit normal vectors at the nodes is given by:

$$\mathscr{C}'_{\xi}(\xi=0)\cdot\mathbf{n}_0=\mathscr{C}'_{\xi}(\xi=1)\cdot\mathbf{n}_1=\frac{(\mathbf{n}_0+\mathbf{n}_1)\cdot\mathbf{d}}{2}, \tag{4.19}$$

which is zero (fulfilment of the boundary conditions) only for symmetric distribution of the nodal normal vectors. Figure 4.4 presents the comparison between original and modified Nagata interpolation applied to an edge, providing a qualitative analysis of the

interpolation behaviour. In this example, the angle between the normal vector $\mathbf{n}_0$ and the vector joining the end points of the edge $\mathbf{d}$ is 120°, while the angle between the normal vector $\mathbf{n}_1$ and $\mathbf{d}$ ranges from 30° to 120° (descending path in Figure 4.4).

The boundary conditions imposed by the nodal normal vectors are exactly fulfilled using the original Nagata interpolation for $a \neq \pm 1$. On the other hand, the modified interpolation only guarantees the orthogonality of the curve with the nodal normal vectors in case of symmetric interpolation (4.19), for which both formulations are equivalent (third frame in Figure 4.4). However, the original interpolation formulation can generate unrealistic curves with inversion of orientation, as shown in the penult frame of Figure 4.4. Moreover, the relative position of the midpoint of the curve $(\xi = 0.5)$, which is indicated by a solid point in Figure 4.4, is much more dependent of the normal vectors orientation in the original formulation. The angles between the nodal normal vectors and the straight line that connects the nodes are defined by:

$$\begin{aligned} \alpha_0 &= \cos^{-1}\left(\frac{-\mathbf{n}_0 \cdot \mathbf{d}}{\|\mathbf{d}\|}\right) = \cos^{-1}\left(-\lambda_0 / \|\mathbf{d}\|\right), \\ \alpha_1 &= \cos^{-1}\left(\frac{\mathbf{n}_1 \cdot \mathbf{d}}{\|\mathbf{d}\|}\right) = \cos^{-1}\left(-\lambda_1 / \|\mathbf{d}\|\right), \end{aligned} \tag{4.20}$$

which are illustrated in Figure 4.5 (a). These angles are posteriorly used to quantify the deviation in the fulfilment of the boundary conditions in the modified Nagata formulation. The comparison between the original (4.15) and the modified (4.18) Nagata interpolations highlights that the curvature parameter $\mathbf{c}'$ is not as sensitive as $\mathbf{c}$ to the differences between the angles α_0 and α_1 [Sekine 10]. Indeed, the direction of the coefficient vector in the original interpolation is obtained by the weighted average of the nodal normal vectors, while in the modified interpolation is given only by the average. In addition, both interpolation formulations have the same form if $\lambda_0 = \lambda_1$.

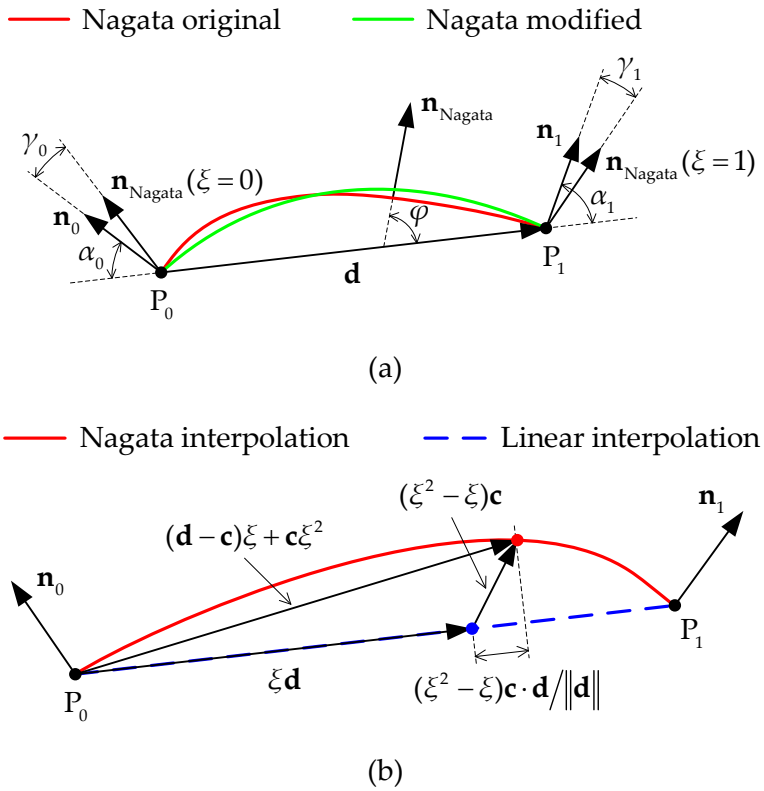


Figure 4.5. Definition of some variables used in the comparison between: (a) original and modified Nagata interpolation; (b) linear and Nagata interpolation.

The effect of the modification applied to the Nagata interpolation is examined in detail through the curve interpolation example with unitary distance between nodes and angles of the nodal normal vectors $\alpha_0 = 60°$ and $\alpha_1 = 80°$ (fourth frame in Figure 4.4). The angle φ between the normal vector of the Nagata curve and the vector $\mathbf{d}$ joining the nodes (see Figure 4.5 (a)) is evaluated for any local coordinate using (4.4) with the curvature parameter defined by (4.15) for the original and by (4.18) for the modified Nagata. Its distribution is presented in Figure 4.6 (a) for the original and modified Nagata interpolation, as well as for the linear interpolation. The higher difference between both Nagata interpolations occurs at the nodes, where the angle obtained with the modified interpolation is lower in node P_0 and higher at node P_1. In fact, the angle distribution provided by the modified interpolation is approximately linear in the parametric domain of the curve. Note that the normal vector of the Nagata curve defined by the original formulation satisfies the imposed boundary conditions at the nodes.

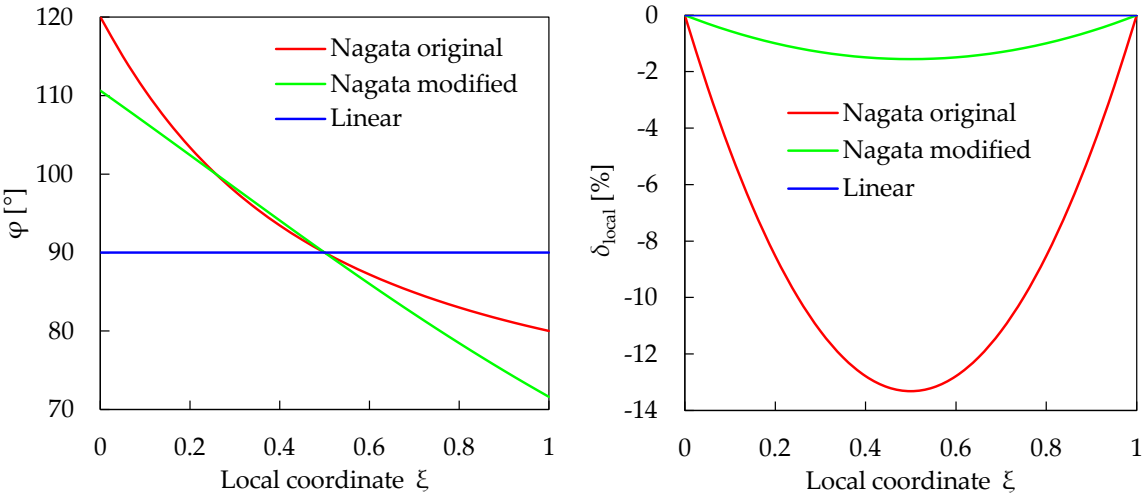


Figure 4.6. Comparison between original and modified Nagata curve interpolation for $\alpha_0 = 60°$ and $\alpha_1 = 80°$: (a) distribution of the angle defined by the normal vector; (b) deviation in the local coordinates to the Euclidean space.

The second parameter studied is related with the connection between local coordinates and Euclidean space, which is particularly important in the description of contact surfaces for deformable bodies. The linear interpolation is taken as reference since it establish a linear relationship between local and global coordinates. Figure 4.5 (b) presents the position vector of a generic point lying on the curve obtained by linear and Nagata interpolation with the same local coordinate. The direction of the vector joining the point of the linear interpolation with the point of the Nagata curve indicates the level of nonlinearity in the relationship between local and global coordinates. Therefore, the deviation is defined by the projection of this vector on the vector joining the nodes normalized by its length, which is given by:

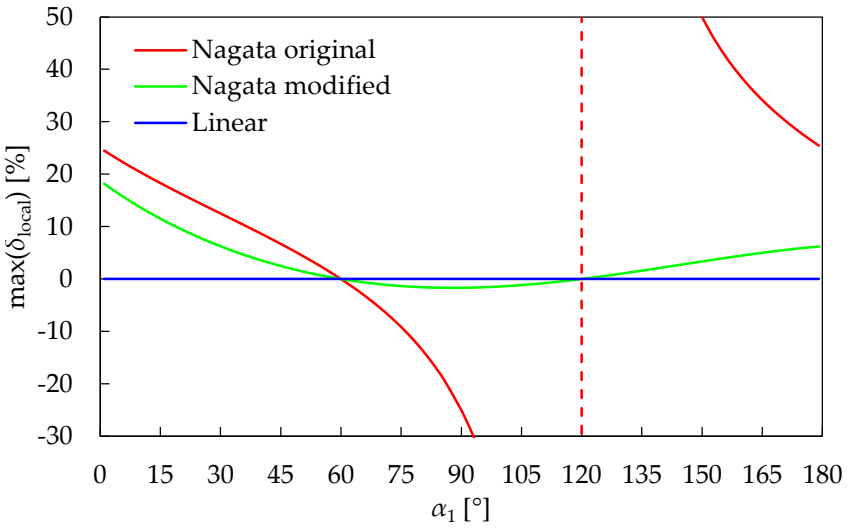
$$\delta_{\text{local}}(\xi) = \left\{ (\xi^2 - \xi)\mathbf{c} \cdot \frac{\mathbf{d}}{\|\mathbf{d}\|} \right\} \Big/ \|\mathbf{d}\| = \frac{(\xi^2 - \xi)\mathbf{c} \cdot \mathbf{d}}{\|\mathbf{d}\|^2}. \tag{4.21}$$

This deviation can be easily observed through the non-uniform distribution of points in the Euclidean space, which are uniformly created in the local coordinate system (see Figure 4.4 (a)). The maximum value is always obtained for the midpoint of the curve, defined by:

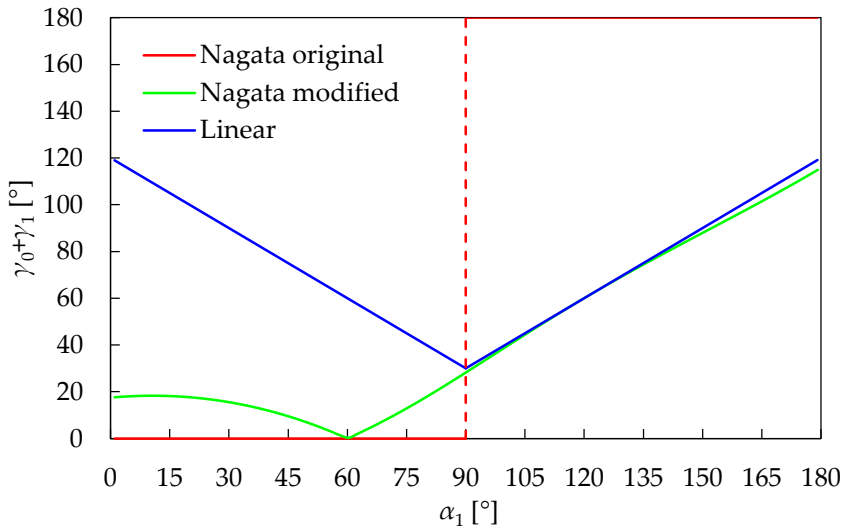
$$\max(\delta_{\text{local}}) = \delta_{\text{local}}(\xi = 0.5) = -\frac{\mathbf{c} \cdot \mathbf{d}}{4\|\mathbf{d}\|^2}. \tag{4.22}$$

The distribution of the deviation between local and global coordinates is presented in Figure 4.6 (b) for the same example previously analysed (fourth frame in Figure 4.4),

comparing the original interpolation formulation with the modification proposed by [Sekine 10]. The maximum deviation is reduced from approximately −13% to −1.5% adopting the modified Nagata interpolation. The negative sign means that the asymmetry of the curve leads to more points closer to the node P_0, as shown in Figure 4.4. This figure also highlights that this improvement lead to a more uniform distribution of the points created on the Nagata curves when the modified formulation is adopted.



(a)



(b)

Figure 4.7. Comparison between original and modified Nagata interpolation for $\alpha_0 = 60°$: (a) maximum deviation between local and global coordinates; (b) violation of the imposed boundary conditions.

The maximum deviation between local and global coordinates is shown in Figure 4.7 (a) for the same example with $\alpha_0 = 60°$ but with a variable α_1 angle, where the original and modified Nagata interpolations are compared. The difference between both formulations is higher when the normal vectors are almost parallel ($\alpha_1 = 120°$), but the deviation is always significantly inferior for the modified interpolation. Moreover, the original Nagata interpolation leads to an asymptote, which is related with the singular point in the definition of the curvature parameter (4.2) and results the inversion of the curve orientation for adjacent angles. On the other hand, the main drawback of the modified Nagata interpolation is the non-fulfilment of the boundary conditions (normal vectors) imposed at the nodes. In fact, the normal vectors defined as input are different from the ones of the modified Nagata curve at the nodes, as shown in Figure 4.5 (a) and Figure 4.6 (a). In order to evaluate the influence of the asymmetry between the normal vectors at the both ends, this difference in the nodal normal vectors is evaluated in this study through the angle between the imposed normal vector and the normal vector of the interpolated curve, given by the following expressions:

$$\begin{aligned} \gamma_0 &= \cos^{-1}(\mathbf{n}_0 \cdot \mathbf{n}_{\text{Nagata}}(\xi = 0)), \\ \gamma_1 &= \cos^{-1}(\mathbf{n}_1 \cdot \mathbf{n}_{\text{Nagata}}(\xi = 1)), \end{aligned} \tag{4.23}$$

where the normal vectors to the Nagata curve are evaluated using (4.4) with the curvature parameter defined by (4.15) for the original and by (4.18) for the modified Nagata.

Considering the previous example, the violation of the boundary conditions imposed by the nodal normal vectors is presented in Figure 4.7 (b) for both Nagata interpolation formulations, as well as for the linear interpolation. The violation is expressed through the sum of the two angles evaluated using (4.23), which are shown in Figure 4.5 (a). The original Nagata formulation satisfies the imposed boundary conditions for $0° \leq \alpha_1 < 90°$, while for angles higher than 90° the angular error is $\gamma_0 + \gamma_1 = 180°$, due to the inversion of the curve orientation in one of the nodes. On the other hand, the modified interpolation leads to a error distribution that is zero only for symmetric interpolation $\alpha_0 = \alpha_1 = 60°$, while for angles higher than 90° the angular error is similar to the one obtained with the linear interpolation (inferior to Nagata original). Although the original Nagata formulation fulfils the boundary conditions (nodal normal vectors) for a wider range of α_1, this quadratic interpolation can generate unrealistic curves with large deviation between local and global coordinates (Figure 4.7). On the other hand, usually the modified formulation does not satisfies the boundary conditions, but the interpolated curve is much more stable presenting a reduced deviation between local and global coordinates.

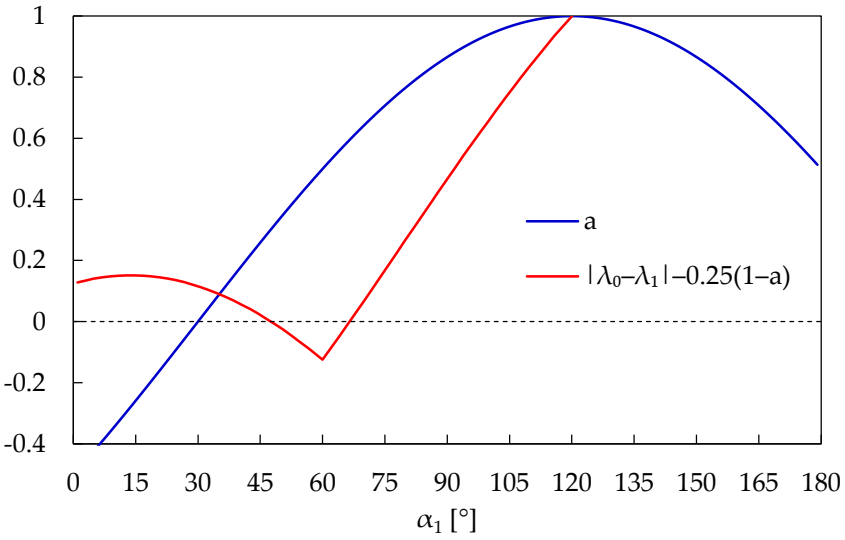


Figure 4.8. Evolution of the functions defining the condition (4.24) for $\alpha_0 = 60°$.

In order to explore the advantages of each interpolation method, the strategy proposed in the present study consists in combining both Nagata formulations, as well as the linear interpolation. This procedure allows to obtain an accurate and stable surface smoothing method. As previously mentioned, the original Nagata interpolation only provides stable curves for nodal normal vectors approximately symmetric. Consequently, the application of the original Nagata interpolation for surface smoothing is restricted to the region defined by the following condition:

$$\left|\lambda_0 - \lambda_1\right| - \frac{1}{4}(1-a)\left\|\mathbf{d}\right\| \le 0, \tag{4.24}$$

where the parameters λ_0 and λ_1 are defined in (4.17). This condition is always fulfilled for $\lambda_0 = \lambda_1$, which represents the perfect symmetry in the interpolation. Moreover, the range of the nodal normal vectors orientation for satisfy the condition (4.24) increases for larger angles between them, i.e. the range of application of the original Nagata interpolation is higher for lower values of *a*. Figure 4.8 presents the evolution of the function that establishes the interpolation condition (4.24) for the example previously analysed ($\alpha_0 = 60°$). In this case, according to (4.24), the original interpolation can only be adopted for an orientation of the nodal normal vector within the range $47.5° < \alpha_1 < 66.5°$. This range is independent from the distance between the nodes since this variable is also included in the definition of the first term in condition (4.24), through (4.17). This feature allows to apply condition (4.24) for identical models since it is invariant under scaling, rotation and translation.

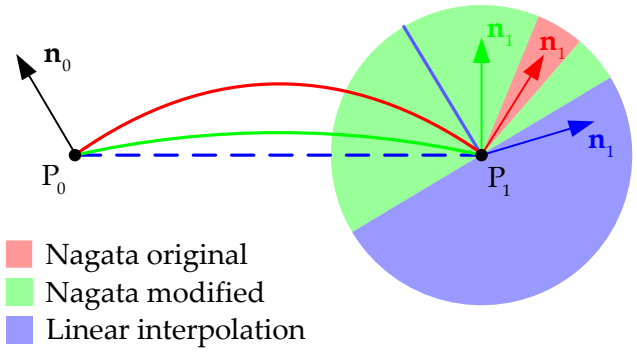


Figure 4.9. Schematic representation of the interpolation method selected for the surface smoothing procedure for $\alpha_0 = 60°$.

The selection of the interpolation method used during the surface smoothing (original or modified Nagata interpolation or linear interpolation) is shown in Figure 4.9 for the same example $(\alpha_0 = 60°)$. The selection criteria is based in the angle between the nodal normal vectors $\mathbf{n}_0$ and $\mathbf{n}_1$ (right-hand side in Figure 4.9). When the angle between the nodal normal vectors is higher than 90° $(a < 0)$ the interpolation adopted is linear to avoid generating curves with large curvature. Besides, due to the definition of the curvature parameter (4.2), the linear interpolation is also used when the two nodal normal vectors are parallel $(a = 1)$. The original Nagata interpolation [Nagata 05] is applied when the condition (4.24) is satisfied. Otherwise, the modified formulation [Sekine 10] is adopted, as shown in Figure 4.9.

The general algorithm adopted in the selection of the interpolation method for the surface smoothing procedure is presented in Table 4.1. As previously discussed, it is based in the relative orientation of the two nodal normal vectors. This strategy allows to combine the accuracy of the original Nagata interpolation in the fulfilment of the boundary conditions with the stability of the modified interpolation, taking advantage of both formulations. Moreover, the linear interpolation is used for parallel nodal normal vectors , as well as when the angle between them is higher than 90°, which is physically inadmissible for a discretized contact surface due to the inversion of the surface normal vector orientation in a single finite element. The adoption of quadratic interpolation in such situations provides curves with high curvature.

Table 4.1. Outline of the algorithm adopted in the selection of the interpolation method.

If $a<0$ or $a=1$ **Then**
Linear interpolation, curvature parameter $\mathbf{c}=\mathbf{0}$
Elseif $\left\|\lambda_0-\lambda_1\right\|-\frac{1}{4}(1-a)\left\|\mathbf{d}\right\|<0$ **Then**
Nagata original, curvature parameter evaluated from (4.15)
Else
Nagata modified, curvature parameter evaluated from (4.18)
Endif

4.3. Interpolation accuracy

As mentioned previously, the surface smoothing procedures allow to obtain a more accurate evaluation of the kinematic contact variables, while the discontinuities in the surface normal vector are eliminated. This section highlights the geometric improvements obtained with the smoothing method, when compared with the traditional finite element description. The accuracy of the surface representation is evaluated in this study through two distinct types of error: shape of the interpolated surface and deviation in the surface normal vector [Hama 08], [Batailly 13]. The first one dictates the accuracy in the computation of the normal gap function (2.30), while the second one is related with the arising of artificial oscillations in the contact force for large sliding contact problems. Besides, the influence of the surface discretization on the geometrical accuracy is analysed. In this section only geometries defined by analytical functions are studied, since the required nodal normal vectors for the Nagata interpolation can be easily evaluated from the analytical function.

In some simple axisymmetric geometries, such as cylinders and spheres, the 3D surface geometry can be converted into a 2D circular arc. Thus, the shape accuracy of the interpolation is evaluated by means of the radial error. Considering a circular arc of radius r, the radial error associated with the interpolation is defined by:

$$\delta_{\mathrm{r}}(\eta,\zeta)=\frac{(\mathscr{P}(\eta,\zeta)-\mathbf{o})\cdot\mathbf{n}_{\mathrm{analytical}}-r}{r}, \tag{4.25}$$

where $\mathscr{P}(\eta,\zeta)$ denotes the position vector of a generic point on the interpolated surface, $\mathbf{o}$ is the position vector of the circle centre and $\mathbf{n}_{\mathrm{analytical}}$ is the unit normal vector of the circular arc defined by the analytic function. The radial error represents the dimensionless distance between the interpolated curve and the circular arc defined by the analytical function, measured in the radial direction (see Figure 4.10 (a)). In order to evaluate the

geometric error in a generic point on the interpolated surface, the shape error is defined in a similar way as:

$$\delta_{\text{shape}}(\eta,\zeta) = (\mathscr{P}(\eta,\zeta) - \mathbf{x}_{\text{P}'}) \cdot \mathbf{n}_{\text{analytical}}, \tag{4.26}$$

where $\mathbf{x}_{\text{P}'}$ denotes the position vector of the normal projection of $\mathscr{P}(\eta,\zeta)$ on the analytical geometry, as shown in Figure 4.10 (b). The shape error defines the absolute signed distance between two points, measured in the normal direction of the analytical geometry. Figure 4.10 (b) presents this error evaluated in two distinct points of a geometry, where the left- and right-hand point present negative (inside interpolation) and positive (outside interpolation) shape error, respectively. The second type of error studied is the normal vector error being its modulus defined by:

$$\left|\delta_{\text{n}}(\eta,\zeta)\right| = \cos^{-1}(\mathbf{n}_{\text{Nagata}}(\eta,\zeta) \cdot \mathbf{n}_{\text{analytical}})\,[^\circ], \tag{4.27}$$

where $\mathbf{n}_{\text{Nagata}}$ is the unit normal vector of the interpolated surface, which is defined by expressions (4.9) and (4.14) for triangular and quadrilateral patches, respectively. The modulus of the normal vector error expresses the angle between the normal vector of the interpolated surface and the analytical normal vector at the projection point (see Figure 4.10). This error is directly associated with the discontinuity of the normal vector orientation in the contact surface, which is a key point for the robustness of the solution procedure adopted for solving contact problems. In fact, if the error in the normal vector is zero in the transition between neighbouring patches, the contact surface is smooth.

The traditional approach to model contact surfaces using bilinear finite elements can be considered a particular case of the proposed surface smoothing method. The adopted Nagata patch interpolation allows to create patches without recovering their curvature. Thus, in this study the bilinear facets are defined through the Nagata patch interpolation, setting to zero the value to the curvature parameter (4.2). This strategy allows to compare both surface description methods, using the same numerical methods and algorithms involved in the contact treatment.

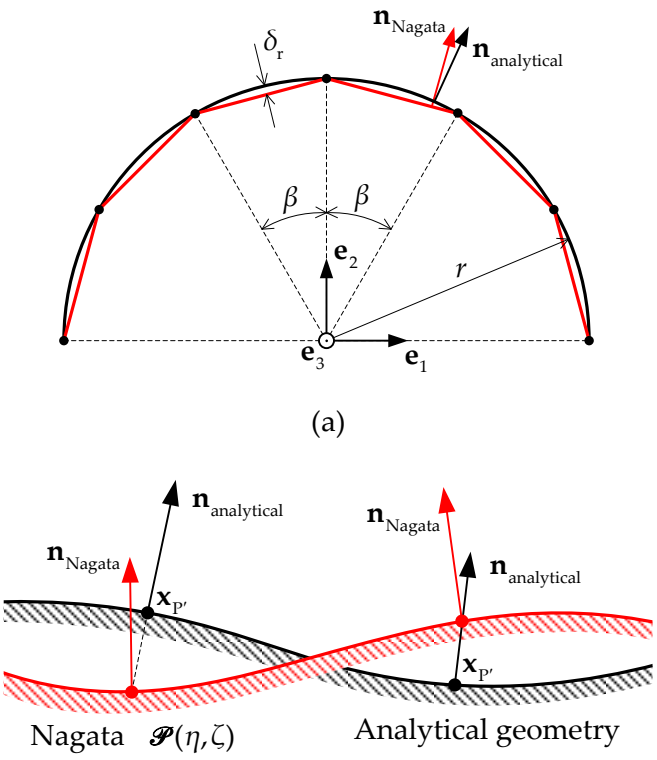


Figure 4.10. Evaluation of the accuracy in the surface interpolation: (a) radial error in a circular arc; (b) normal vector error evaluated in two generic points.

4.3.1. Circular arc

In order to measure the accuracy of each surface description method, this section is dedicated to the analysis of a 2D geometry, the circular arc (Figure 4.10 (a)). Although it is a simple geometry, constant radius fillet surfaces and surfaces of revolution are frequently found in the description of contact surfaces [Lin 01]. Both the radial and the normal vector error distributions are calculated for both contact description methods analysed (linear and Nagata interpolation). However, the definition of the normal vector error is improved for this 2D example, which is given by:

$$\delta_n(\xi) = \text{sign}(\delta_n(\xi))\left|\delta_n(\xi)\right|,$$
$$\text{sign}(\delta_n(\xi)) = \begin{cases} -1 & \text{if} \quad (\mathbf{n}_{\text{Nagata}} \times \mathbf{n}_{\text{analytical}}) \cdot \mathbf{e}_3 < 0 \\ 1 & \text{if} \quad (\mathbf{n}_{\text{Nagata}} \times \mathbf{n}_{\text{analytical}}) \cdot \mathbf{e}_3 > 0 \end{cases}, \tag{4.28}$$

where $\text{sign}(\bullet)$ denotes the signal convention chosen for δ_n and $\mathbf{e}_3$ is the off-plane unit base vector shown in Figure 4.10. The signed definition of the normal vector error allows to highlight the discontinuities in the normal vector field introduced by the linear interpolation.

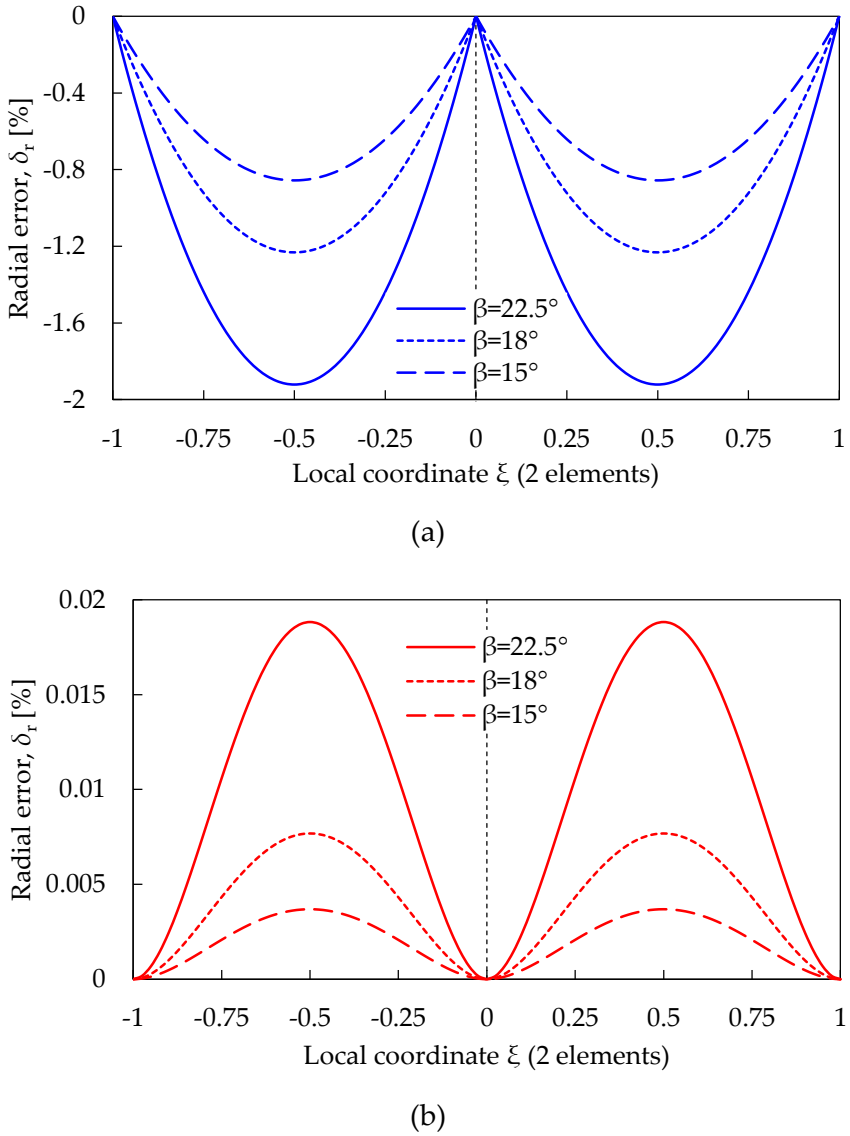
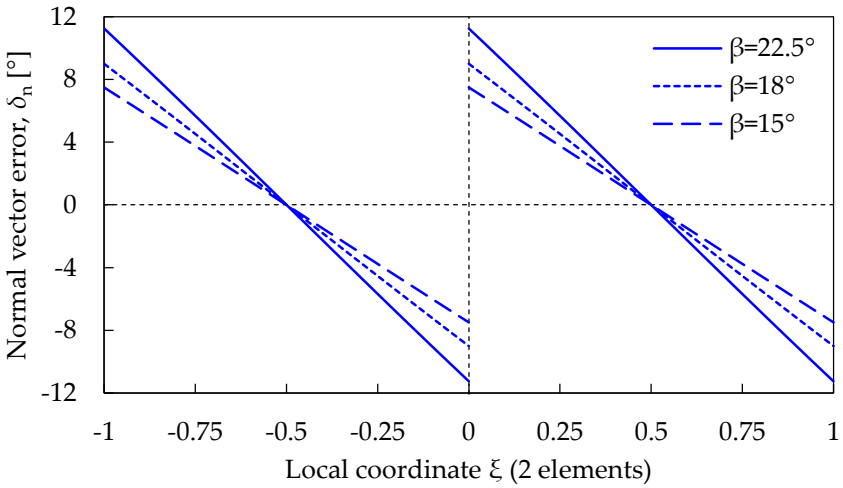


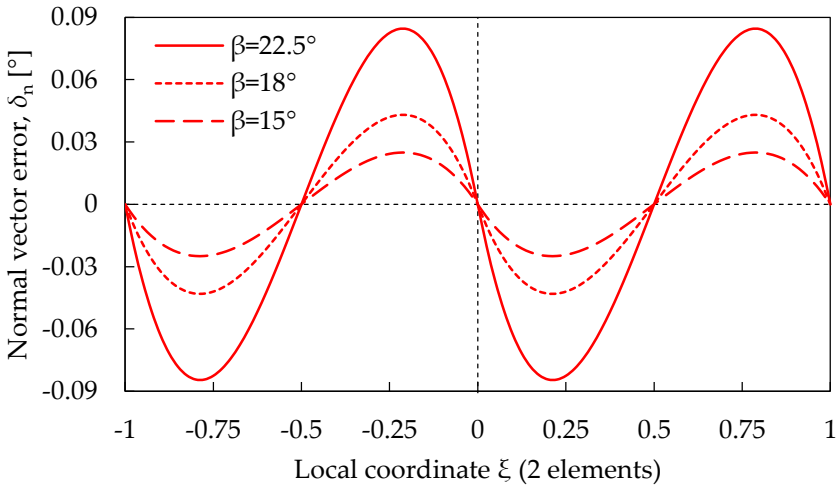
Figure 4.11. Radial error distribution in a circular arc defined by two elements: (a) linear interpolation; (b) Nagata interpolation.

The radial error distribution in a circular arc discretized by two neighbouring finite elements is presented in Figure 4.11 for each interpolation method. The central angle β indicated in Figure 4.10 (a) is the parameter selected to quantity the mesh refinement, which gives the length of each discretized circular arc. Three different values of central angle are analysed, corresponding to the division of half circumference into 8, 10 and 12 elements. As expected, for both interpolation methods, the radial error decreases with the mesh refinement, as shown in Figure 4.11. Besides, the linear interpolation provides a negative value of radial error (curve inside the circular arc), while the Nagata interpolation leads to a curve outside the circular arc (positive radial error). Although the evolution of the radial error is roughly similar for both interpolation methods, the maximum value attained is completely different, although is always located at the midpoint of each interpolated curve.

The error evolution is represented with respect to the local coordinate $0 \leq \xi \leq 1$, although the domain of the left-hand curve is dropped one unity to the negative direction. Note that Figure 4.11 (a) and (b) present different scales since the order of magnitude of the results is not comparable.



(a)



(b)

Figure 4.12. Normal vector error distribution in a circular arc defined by two elements: (a) linear interpolation; (b) Nagata interpolation.

The normal vector error distribution in two neighbouring finite elements of the discretized circular arc is presented in Figure 4.12, for each interpolation method. Such as for the radial error, the normal vector error decreases with the increasing of the number of elements used to describe the circular arc (mesh refinement). The sudden changes in the normal vector created by the linear interpolation can be observed through the discontinuity of the normal vector error distribution across element boundaries, as illustrated in Figure

4.12 (a). In fact, the normal vector error is zero where the radial error attains its maximum value (negative), while the maximum value of the normal vector error occurs at the nodes, and corresponds to half value of the arc central angle. On the other hand, the Nagata interpolation assures G^1 continuity in the resulting curve, since the normal vector error distribution is continuous, as shown in Figure 4.12 (b). Besides, the error is zero at the nodes and in the locations where the radial error attains its maximum value (midpoint). The scales used in Figure 4.12 (b) and (b) are different because the order of magnitude of the results is not comparable.

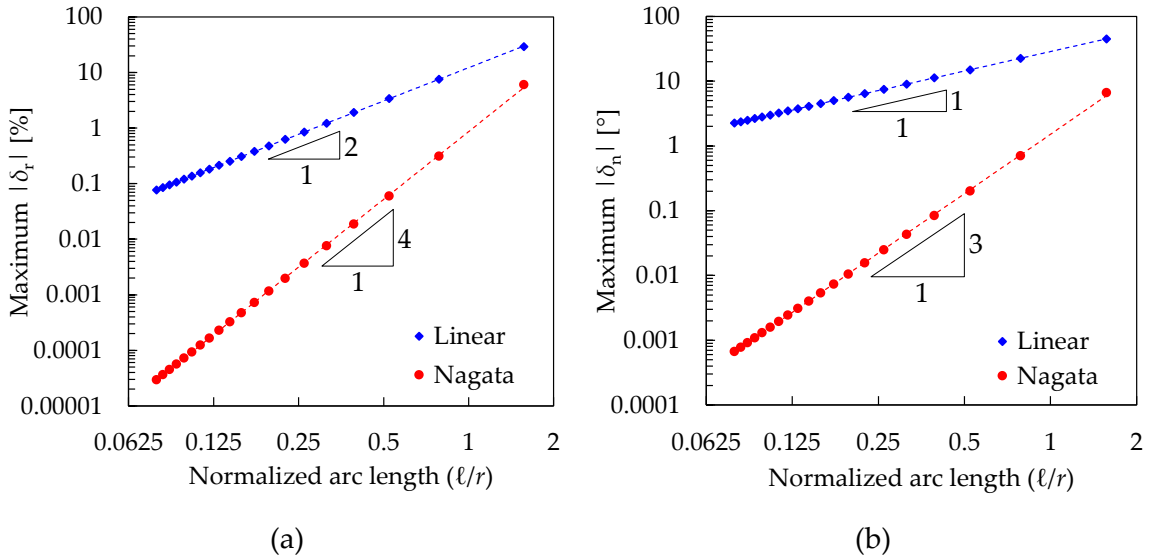


Figure 4.13. Comparison between linear and Nagata interpolation accuracy applied to a circular arc: (a) maximum radial error modulus; (b) maximum normal vector error modulus.

In order to study the influence of the discretization on the interpolation error, different lengths of circular arc are studied. The range considered for the normalized arc length is from 1.571 until 0.079, which corresponds to dividing a quarter of circle from 1 to 20 equal elements, respectively. The normalized arc length expresses the division of the arc length ℓ by the radius r of the circular arc. Figure 4.13 (a) presents the maximum norm of radial error in function of the normalized arc length (ℓ / r), for both the linear and the Nagata interpolation. The maximum value of the radial error decreases with the decrease of the normalized arc length, converging to the analytical geometry. The order of convergence in the radial error provided by the linear interpolation is quadratic, while when applying Nagata interpolation the convergence rate is quartic [Nagata 05]. Figure 4.13 (b) shows the maximum normal vector error modulus as function of the normalized arc length, for both surface description methods. The error decreases linearly when the linear interpolation is adopted, while the Nagata interpolation method provides a cubic convergence rate.

Besides, the undesirable discontinuity of the normal vector field created by the linear interpolation (see Figure 4.12 (a)), the convergence rate of the normal vector error is only linear with the mesh refinement, which produces substantial oscillations in the contact force and associated convergence problems. In fact, for the wide range of normalized arc length analysed in the present study, the maximum normal vector error modulus is always larger than 2° for the linear interpolation, while the Nagata interpolation gives an error always inferior to 1°, with the exception of the coarser mesh with only one element (Figure 4.13 (b)).

4.3.2. Cylinder

The cylindrical geometry is obtained from a circumference extruded along the normal direction of the plane containing it. Therefore, this geometry only presents curvature in the circumferential direction, which is identical to the circular arc, previously studied. In order to evaluate the influence of the surface mesh topology, both structured and unstructured discretizations are studied. In fact, one of the principal advantages of the proposed surface smoothing procedure is its applicability to irregular meshes. Firstly, the cylinder is discretized using a structured coarse mesh composed by both triangular and quadrilateral finite elements (Figure 4.14). The circumferential direction is divided into 8 elements, while the axial direction is discretized by 4 elements. The topology of the triangular elements is obtained from the division of a quadrilateral element through its diagonal.

Figure 4.14 presents the accuracy of each interpolation method applied to a cylindrical surface, which is evaluated through the radial and normal vector error distributions. Since the adopted discretization can be obtained from a 2D profile, both interpolation error distributions are constant along the axial direction. Therefore, its maximum value is independent of the number of elements in the axial direction. This means that the number of triangular finite elements required to describe the cylindrical surface with the same level of error is twice than the number of quadrilateral elements (see Figure 4.14). Moreover, in each cross section of the cylindrical surface, the error distributions are exactly the same to the ones obtained in the circular arc (see Figure 4.11 and Figure 4.12 for the radial and normal vector error, respectively). The maximum value of radial error in the faceted surface representation is −7.7%, while the surface description with Nagata patches leads a maximum value of 0.32%, as shown in Figure 4.14 (a) and (b). Regarding the modulus of the normal vector error, its maximum value for the faceted surface description is 22.5° and only 0.71° in case of surface smoothed with Nagata patches, as shown in Figure 4.14 (c) and (d).

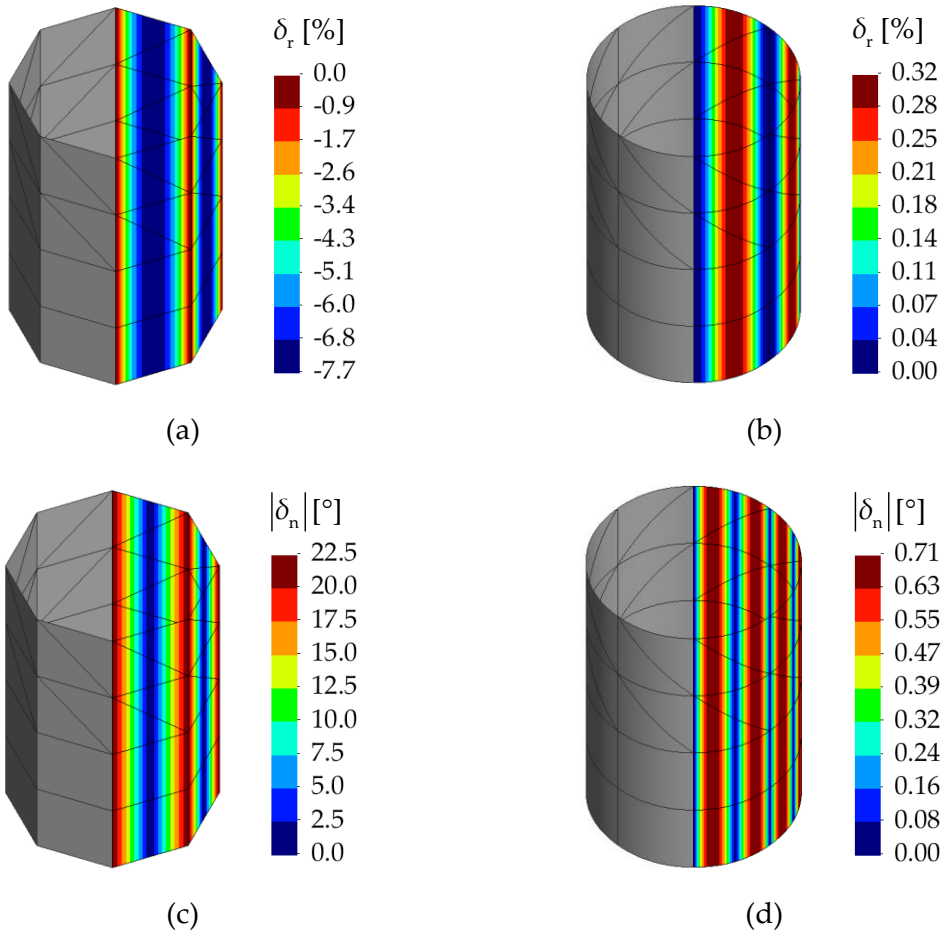


Figure 4.14. Structured discretization of the cylindrical surface with triangular and quadrilateral finite elements: (a) radial error in the faceted elements; (b) radial error in the Nagata patches; (c) normal vector error in the faceted elements; (d) normal vector error in the Nagata patches.

The accuracy of each surface representation method in the description of a cylindrical surface discretized with an unstructured mesh of triangular elements is shown in Figure 4.15. The adopted finite element mesh is composed by 118 nodes and 196 triangular elements (Figure 4.15 (a)). The radial and normal vector error distributions in the faceted finite element description are shown in Figure 4.15 (b) and (c), respectively. The largest (negative) radial error occurs at the middle of the edges aligned with the circumferential direction, while the maximum value of normal vector error is located at the nodes and edges aligned with the axial direction. On the other hand, the smoothing method with Nagata patches provides a much more accurate surface representation (Figure 4.15 (d)). The radial and normal vector error distributions are presented in Figure 4.15 (e) and (f), respectively. The maximum value of radial error occurs in the same locations for both interpolation methods, while the normal vector error is zero at the nodes due to the boundary condition imposed at the nodal normal vectors for the Nagata interpolation. Both analysed errors

decrease more than an order of magnitude when the proposed smoothing procedure is applied.

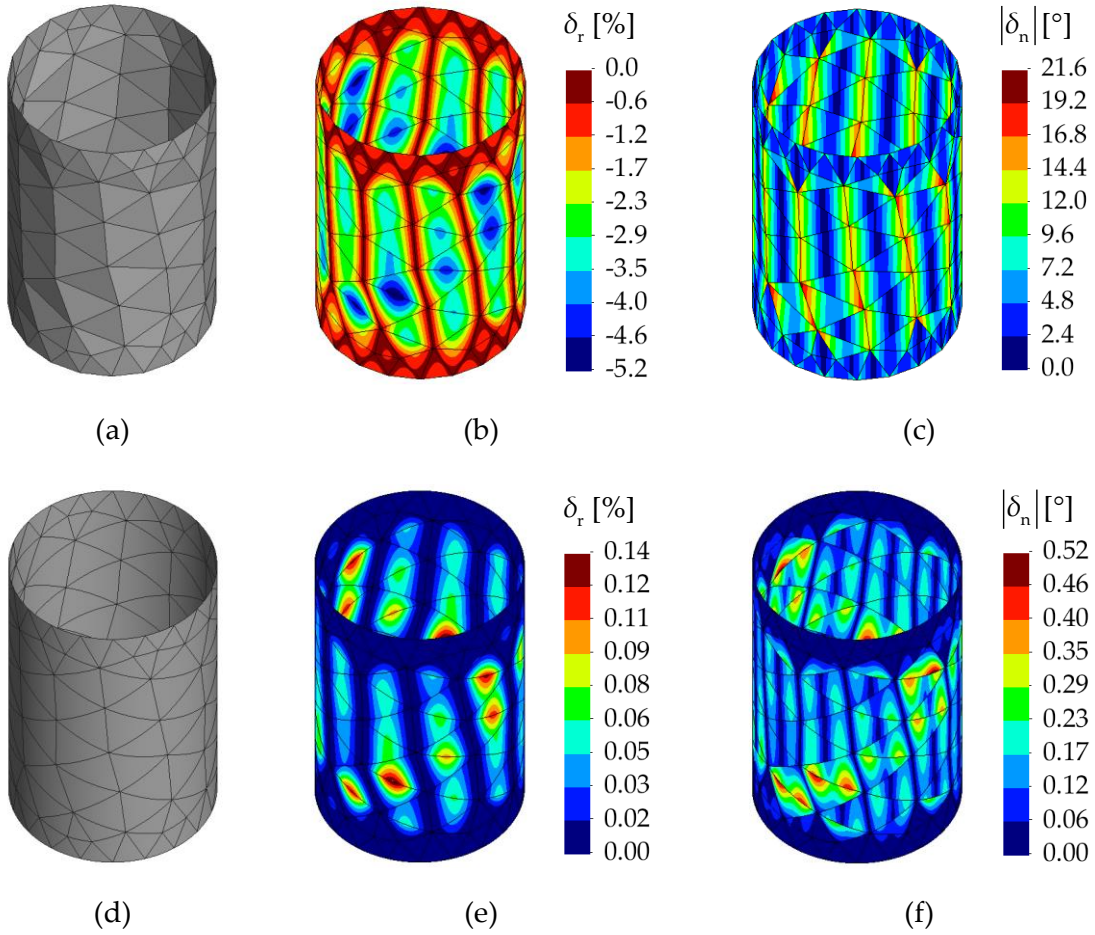


Figure 4.15. Unstructured discretization of the cylindrical surface using triangular elements: (a) faceted mesh; (b) radial error in the faceted elements; (c) normal vector error in the faceted elements; (d) Nagata patches; (e) radial error in the Nagata patches; (f) normal vector error in the Nagata patches.

The second unstructured finite element mesh is composed by 148 nodes and 128 quadrilateral elements, shown in Figure 4.16 (a). The radial and normal vector error distributions in the faceted surface description are shown in Figure 4.16 (b) and (c), respectively. As for the triangular elements, the largest radial error (negative) arises at the middle of the edges aligned with the circumferential direction, while the maximum value of normal vector error is located at the nodes and edges aligned with the axial direction. The radial and normal vector error distributions obtained with the smoothing method are presented in Figure 4.16 (e) and (f), respectively. The largest values of radial error occur in the central zone of the largest patches, while the largest errors in the normal vector are observed along the edges of the Nagata patches. Despite the coarse refinement for the

adopted unstructured meshes, significant improvements are obtained using Nagata patches in the smoothing procedure, as shown in Figure 4.15 and Figure 4.16.

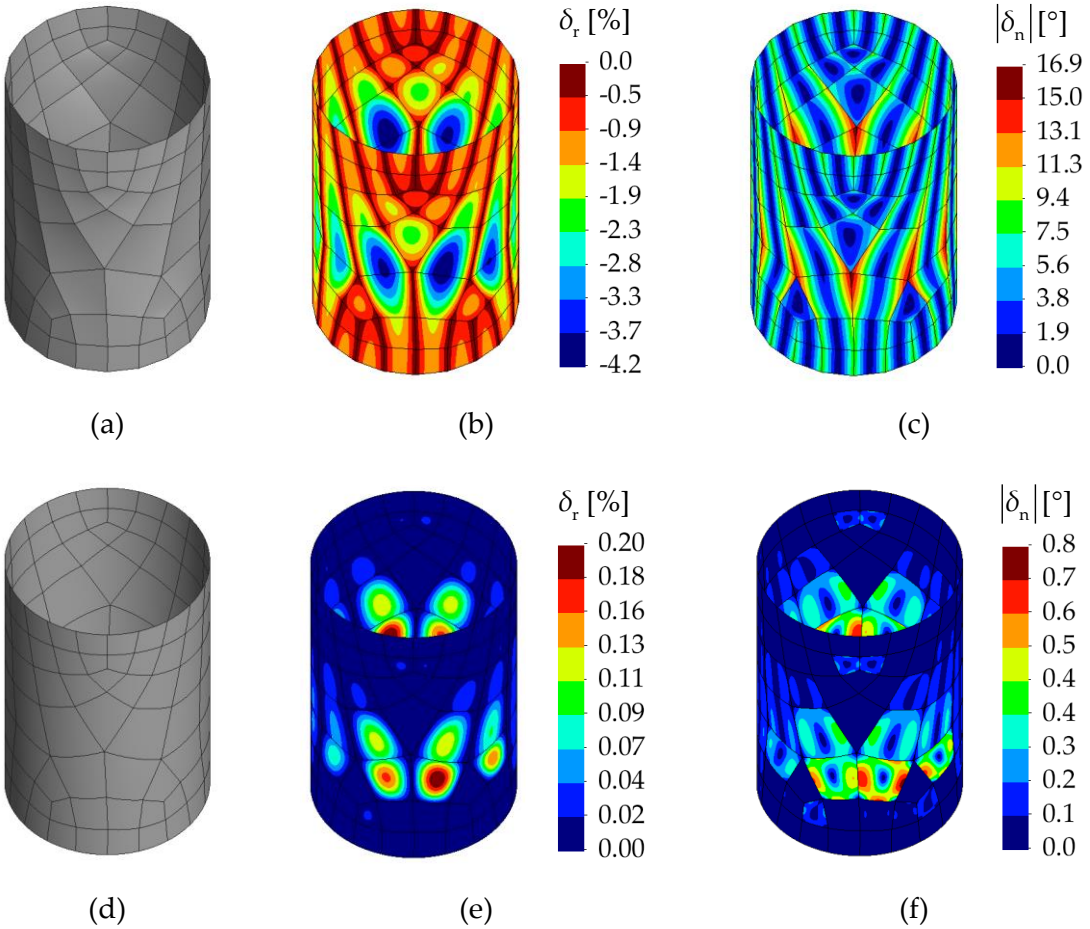


Figure 4.16. Unstructured discretization of the cylindrical surface using quadrilateral elements: (a) faceted mesh; (b) radial error in the faceted elements; (c) normal vector error in the faceted elements; (d) Nagata patches; (e) radial error in the Nagata patches; (f) normal vector error in the Nagata patches.

4.3.3. Sphere

The spherical surface (typical smooth surface) of unit radius is selected to evaluate the accuracy of both surface description methods. Both structured and unstructured meshes are studied, as well as meshes composed either by triangular or quadrilateral finite elements. Concerning the structured discretizations, two distinct meshes are used in the description of the spherical surface, which are composed by triangular and quadrilateral finite elements. The main features of each discretization are presented in Table 4.2, namely the number of nodes and elements. The average element area was evaluated for both meshes, obtained from the division of the spherical surface area by the number of finite

elements. In case of fine meshes, it is 0.025 and 0.033 for the triangular mesh and the quadrilateral mesh, respectively. Since the surface mesh is structured, the average value is close to the maximum value, indicated in Table 4.2.

Table 4.2. Main characteristics of the structured meshes used to describe the spherical surface.

	Nº nodes	Nº elements		Max element area	
		Triangles	Quadrilaterals	Triangles	Quadrilaterals
Coarse mesh	22	16	12	0.433	0.460
Fine mesh	322	256	192	0.029	0.038

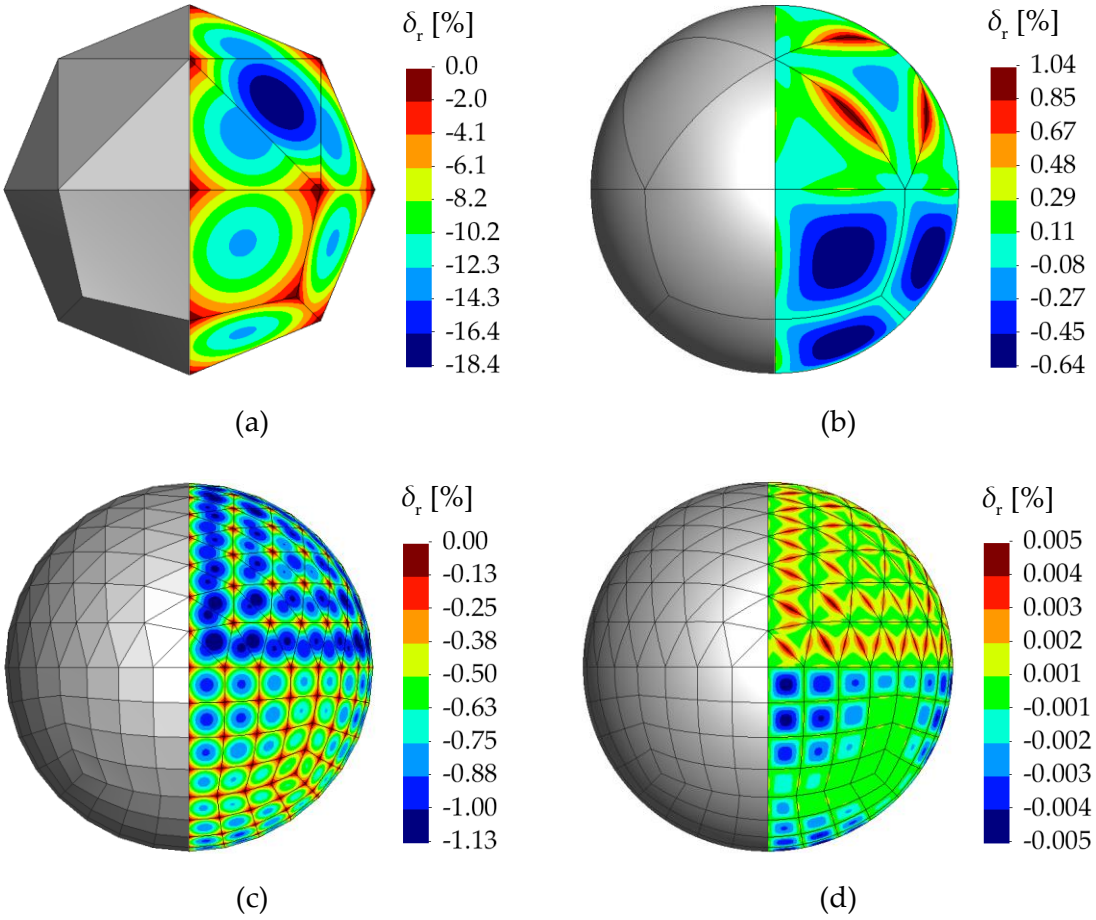


Figure 4.17. Radial error distribution in the spherical surface described by triangular and quadrilateral finite elements: (a) faceted coarse mesh; (b) Nagata patches coarse mesh; (c) faceted fine mesh; (d) Nagata patches fine mesh.

Figure 4.17 presents the spherical surface discretized by traditional bilinear finite elements (left-hand side) and smoothed with Nagata patches (left-hand side). The radial error distribution in the coarse and fine discretizations described by faceted finite elements and Nagata patches is presented in Figure 4.17 (a) and (c) and Figure 4.17 (b) and (d), respectively. A quick qualitative error analysis shows that the faceted approximation provides poor accuracy when compared with the smoothed model. The radial error is negative when the faceted finite element mesh is employed, either using triangular or quadrilateral finite elements. Besides, the largest values of error (negative) occur in the central zone of the faceted elements with biggest area (Figure 4.17 (a) and (c)). On the other hand, the application of the Nagata interpolation to smooth the surface leads to an approximated geometry where the radial error tends to be positive in the triangular patches and negative in the quadrilateral patches (Figure 4.17 (b) and (d)). The Nagata interpolation produces a geometry with largest radial error (positive) along the edges of triangular patches and (negative) at central zone of the largest quadrilateral patches.

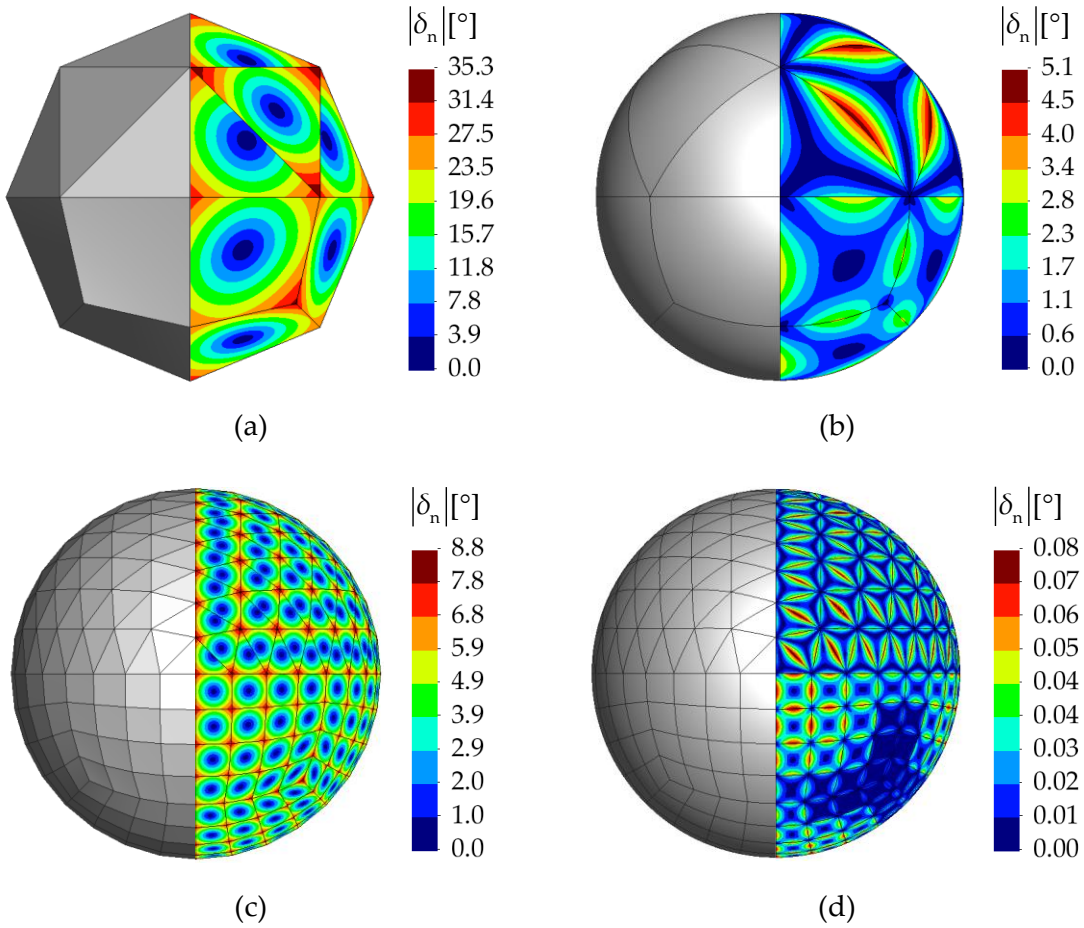


Figure 4.18. Normal vector error distribution in the spherical surface described by triangular and quadrilateral finite elements: (a) faceted coarse mesh; (b) Nagata patches coarse mesh; (c) faceted fine mesh; (d) Nagata patches fine mesh.

The normal vector error distribution in the spherical surface is presented in Figure 4.18 for both surface description methods, comparing the coarse and fine meshes. Considering the surface description with the faceted finite elements, the maximum value of error occurs always at the nodes of the mesh (Figure 4.18 (a) and (c)). On the other hand, the adoption of Nagata patches in the smoothing procedure leads to an interpolated surface with maximum error located along the edges, as shown in Figure 4.18 (b) and (d). Besides, the nodes do not present normal vector error since the Nagata interpolation uses the nodal normal vectors as input parameter. This means that the smoothing method ensures G^1 continuity in the nodes. However, the discrepancy of normal vectors at the patch boundaries can be observed, which is particularly evident in the coarse mesh (Figure 4.18 (b)). Therefore, the contact surface described by the proposed smoothing procedure does not satisfies the G^1 continuity at patch boundaries.

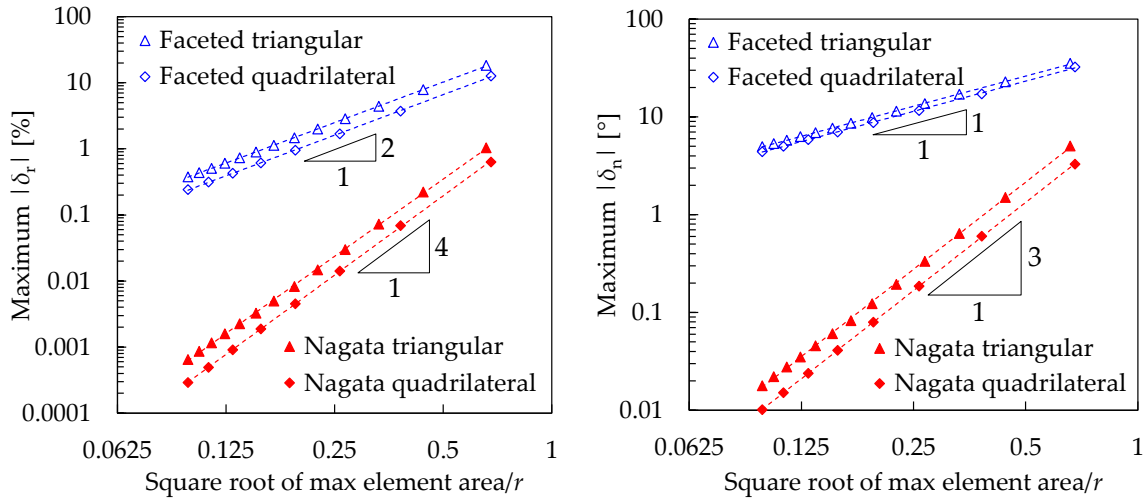


Figure 4.19. Comparison between faceted and Nagata patch interpolation accuracy in the description of a spherical surface: (a) maximum radial error modulus; (b) maximum normal vector error modulus.

The effect of discretization (mesh refinement) in the interpolation accuracy is presented in Figure 4.19 (a) through the maximum value of radial error, evaluated using structured meshes. The comparison between the two surface description methods shows that the maximum radial error modulus is significantly lower when the smoothing method is applied. Indeed, the maximum value of radial error in the Nagata interpolation decreases quartically with the square root of the maximum element area normalized by the sphere radius. On the other hand, the convergence rate in the faceted surface description is only quadratic, as shown in Figure 4.19 (a). Moreover, for both surface description methods, the discretization with quadrilateral finite elements provides a maximum value of radial error

always inferior to the one obtained with triangular elements, considering the same finite element area.

The convergence rate of the maximum normal vector error modulus attained in the description of the spherical surface is presented in Figure 4.19 (b), for each surface description method. Adopting the faceted surface description, the maximum error value decreases linearly with the normalized square root of maximum element area, while the smoothing method with Nagata patches provides a cubic convergence rate. Thus, the spherical surface discretization presented in Figure 4.18 (d), which is composed by 448 finite elements (see Table 4.2), the maximum normal vector error is about 100 times smaller when using the Nagata patch. Besides, this ratio becomes larger with the mesh refinement, as shown in Figure 4.19 (b). Therefore, the presented cubic convergence rate in the normal vector error allows considering that the smoothing method with Nagata patches ensures *quasi*-G^1 continuity in the patch boundaries. In fact, for the analysed range of element area, the maximum normal vector error modulus is always larger than 4° for the bilinear interpolation, while the Nagata patch interpolation provides a maximum error always inferior to 5°, as shown Figure 4.19 (b). Also, as previously shown for the maximum radial error, the maximum value of normal vector error in the quadrilateral finite element typology is always inferior to the one obtained by the triangular typology.

Two distinct unstructured meshes are applied in the description of the spherical surface, one composed by triangular finite elements (Figure 4.20 (a)) and the other with quadrilateral finite elements (Figure 4.21 (a)). The maximum element area of each discretization is 0.131 and 0.152 for the triangular and quadrilateral finite element mesh, respectively. The distribution of both analysed errors in the description of the spherical surface using an unstructured discretization composed by triangular elements is presented in Figure 4.20. The maximum value of radial error in the faceted surface description is located in the finite elements with largest area (Figure 4.20 (b)). Concerning the normal vector error, its maximum is situated in the nodes (Figure 4.20 (c)). On the other hand, the radial and normal vector error distributions obtained with the smoothing method are presented in Figure 4.20 (e) and (f), respectively. The largest value for both errors is observed along the patch edges with largest length, regardless of its orientation. Therefore, the location of the maximum value of error is independent from the discretization topology (structured or unstructured mesh). Besides, a significant improvement of the surface representation is achieved using the smoothing method with Nagata patches, where the maximum radial error is about 45 times smaller and the maximum normal vector error is approximately 20 times smaller, as shown in Figure 4.20.

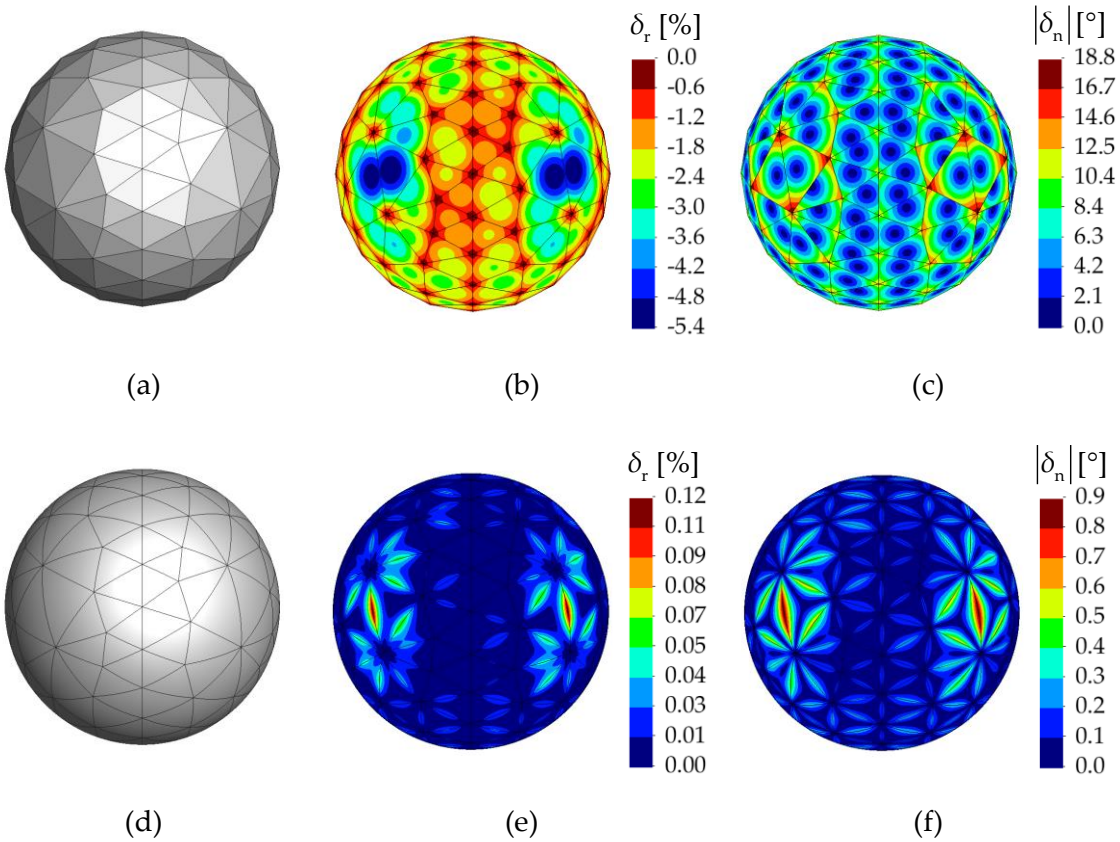


Figure 4.20. Unstructured discretization of spherical surface using triangular elements: (a) faceted mesh; (b) radial error in the faceted elements; (c) normal vector error in the faceted elements; (d) Nagata patches; (e) radial error in the Nagata patches; (f) normal vector error in the Nagata patches.

The accuracy of each surface representation method applied to the sphere discretized with quadrilateral elements is presented in Figure 4.21, which is evaluated through the radial error and the normal vector error. Concerning the surface description with bilinear quadrilateral finite elements, the maximum value of both errors is located in the same positions than in the triangular elements, i.e. the radial error occurs in the central area of the largest finite elements and normal vector error at the nodes (see Figure 4.21 (b) and (c)). The application of the smoothing method using Nagata patches leads to a shift of the position where the maximum value of error occurs. The maximum radial error is located in the large patches (negative) as well as in distorted patches (positive), while the normal vector error arises in the edges of the previously mentioned patches, as shown in Figure 4.21 (e) and (f). A significant improvement of the surface representation is achieved using the smoothing method with Nagata patches, obtaining a maximum radial error about 55 times smaller and a maximum normal vector error about 35 times smaller.

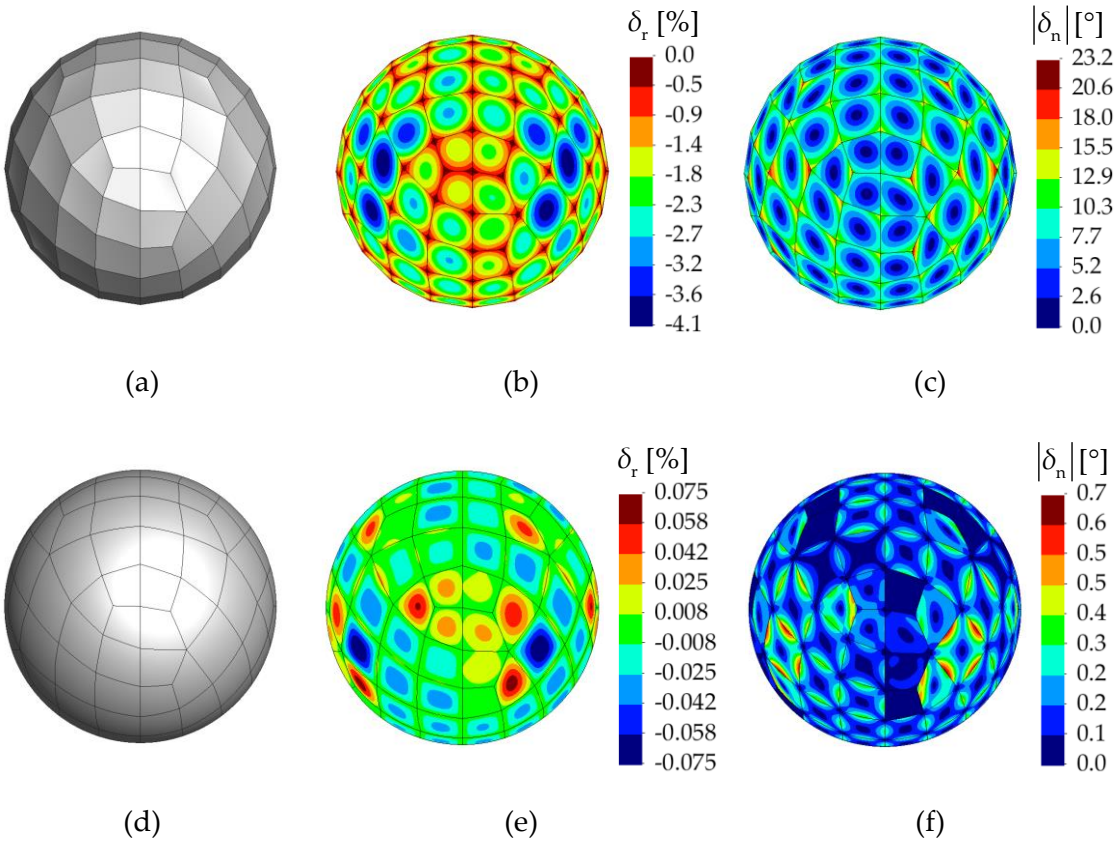
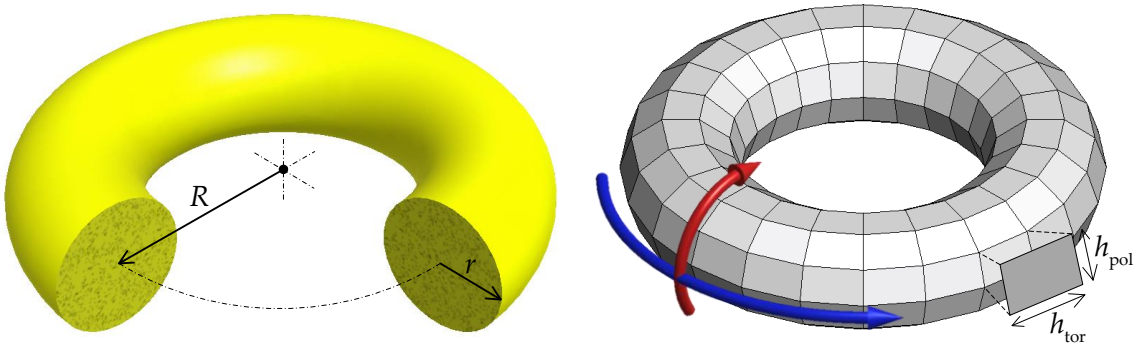


Figure 4.21. Unstructured discretization of spherical surface using quadrilateral elements: (a) faceted mesh; (b) radial error in the faceted elements; (c) normal vector error in the faceted elements; (d) Nagata patches; (e) radial error in the Nagata patches; (f) normal vector error in the Nagata patches.

4.3.4. Torus

The ring torus is the second closed surface geometry adopted in the present study to evaluate the accuracy of the Nagata patch interpolation. This geometry is defined by two dimensions (Figure 4.22 (a)), the major radius R and the minor radius r, where the ratio between them gives the aspect ratio of the torus R/r. Moreover, it is characterized by having regions of elliptic, parabolic and hyperbolic points, which make it more complex than the sphere (elliptic surface). Three different aspect ratios are studied and compared, i.e. three distinct values of major radius $(R=2,4,6)$ are tested while the minor radius $(r=1)$ is kept constant. Furthermore, several structured discretizations are performed for each aspect ratio of the torus, as well as unstructured, to examine the influence of mesh distortion.



(a) (b)

Figure 4.22. Torus geometry: (a) main dimensions; (b) poloidal (red arrow) and toroidal (blue arrow) directions and characteristic finite element dimensions.

The accuracy of the surface description is strongly influenced by the finite element size in the poloidal and toroidal directions (see Figure 4.22 (b)), which should be adjusted taking into account the aspect ratio of the torus. Therefore, the ratio between the two principal dimensions of the finite elements is an important parameter for the mesh generation, which is defined by:

$$r_t = h_{pol} / h_{tor}\,, \tag{4.29}$$

where h_{pol} and h_{tor} denotes the finite element length in poloidal and toroidal directions, as shown in Figure 4.22 (b). Since the finite element length in the toroidal direction increases with the distance to the centre, the ratio (4.29) is evaluated in the farthest element to the centre, in order to capture the maximum element area. The range of the ratio (4.29) considered in this study is situated between approximately 0.5 and 2.0, i.e. the number of finite elements in the toroidal direction ranges from 16 to 64 while the number of elements in the poloidal is set as 8, as shown in Figure 4.23. This value is elected based on the previous study performed for the cylinder (see Figure 4.14). This means that the coarse and fine meshes have the finite elements aligned in the toroidal and poloidal directions, respectively. The comparison between the two surface description methods applied in the torus with $R=4$ and $r=1$ is presented in Figure 4.23, where the left half of the torus is described by bilinear finite elements and the right half is smoothed with Nagata patches.

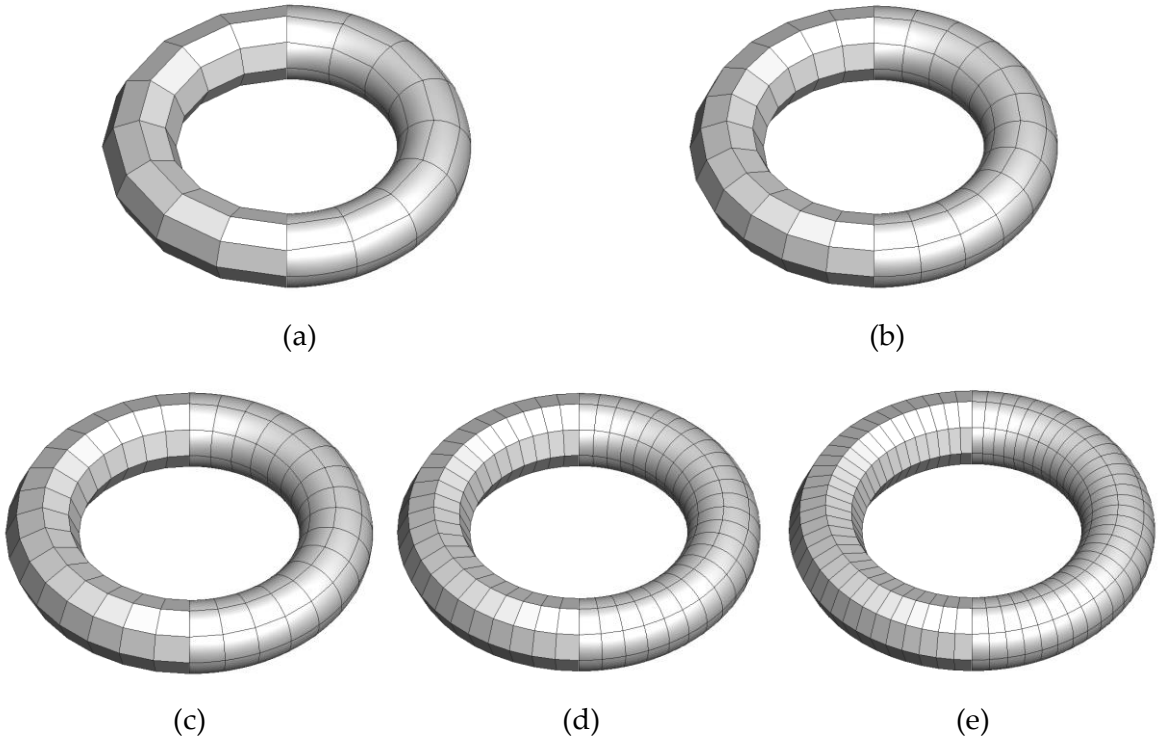


Figure 4.23. Comparison between faceted (left-hand) and smoothed (right-hand) surface descriptions of the torus with $R=4$ and $r=1$, considering different meshes: (a) $r_t=0.49$; (b) $r_t=0.73$; (c) $r_t=0.98$; (d) $r_t=1.46$; (e) $r_t=1.95$.

The comparison of the radial error distribution achieved by each surface description method in the torus geometry is presented in Figure 4.24, for three different structured discretizations. As expected, the error distribution becomes constant in the toroidal direction due to the mesh refinement adopted in this direction. Since the torus geometry contains both hyperbolic and elliptic points, the radial error range in the faceted surface description includes positive and negative values (Figure 4.24 (a)). The same behaviour can be observed in coarse discretizations smoothed with Nagata patches (Figure 4.24 (d)). For both surface description methods, the maximum (positive) value of radial error occurs in the hyperbolic region, while the maximum (negative) value arises at the elliptic points. Moreover, in case of finite element discretizations with low value of r_t, both limits of the radial error range are located in the middle of the Nagata patches, as shown in Figure 4.24 (d). The maximum value of radial error is at least 15 times smaller in the smoothed surface than in the traditional bilinear interpolation with finite elements.

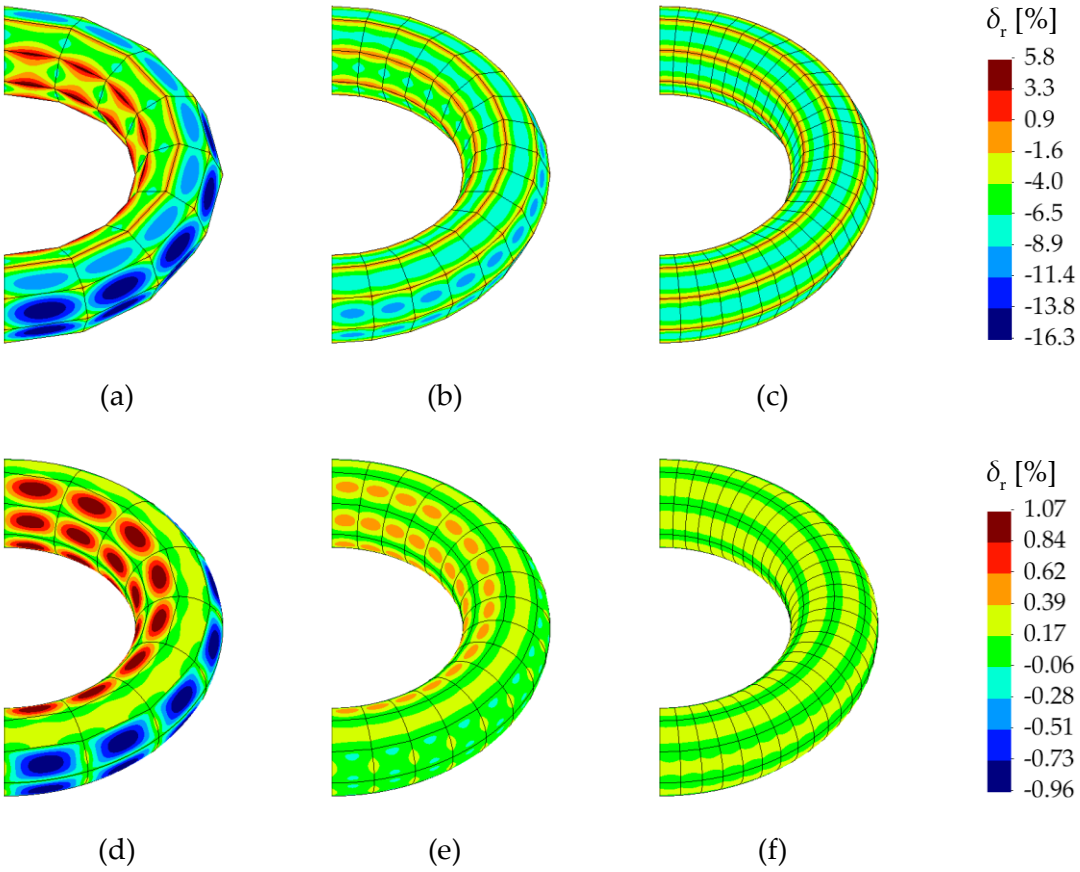


Figure 4.24. Radial error distribution in the toroidal surface for different discretizations: (a) faceted elements $r_t = 0.49$; (b) faceted elements $r_t = 0.98$; (c) faceted elements $r_t = 1.95$; (d) Nagata patches $r_t = 0.49$; (e) Nagata patches $r_t = 0.98$; (f) Nagata patches $r_t = 1.95$.

The normal vector error distribution in the torus geometry is presented in Figure 4.25 for both surface description methods, evaluating the influence of the mesh discretization in the toroidal direction. Regarding the surface description using bilinear finite elements, the largest value of normal vector error is observed along the edges aligned in the toroidal direction, as shown in Figure 4.25 (a)–(c). Furthermore, increasing the number of elements in the poloidal direction leads to a slight decrease of the maximum error value. In fact, the normal vector error distribution in the discretization presented in Figure 4.25 (c) is similar to the one obtained in the cylindrical surface shown in Figure 4.14 (c) (same number of elements in the circumferential direction). The adoption of the Nagata patches in the surface description leads to an important improvement in the surface normal vector definition. The largest value of normal vector error in the smoothed surface occurs in the patch edges aligned in the toroidal direction, located in the elliptic region of the torus. Besides, the maximum normal vector error decreases significantly by means of the mesh refinement in the toroidal direction, as shown in Figure 4.25 (d)–(f).

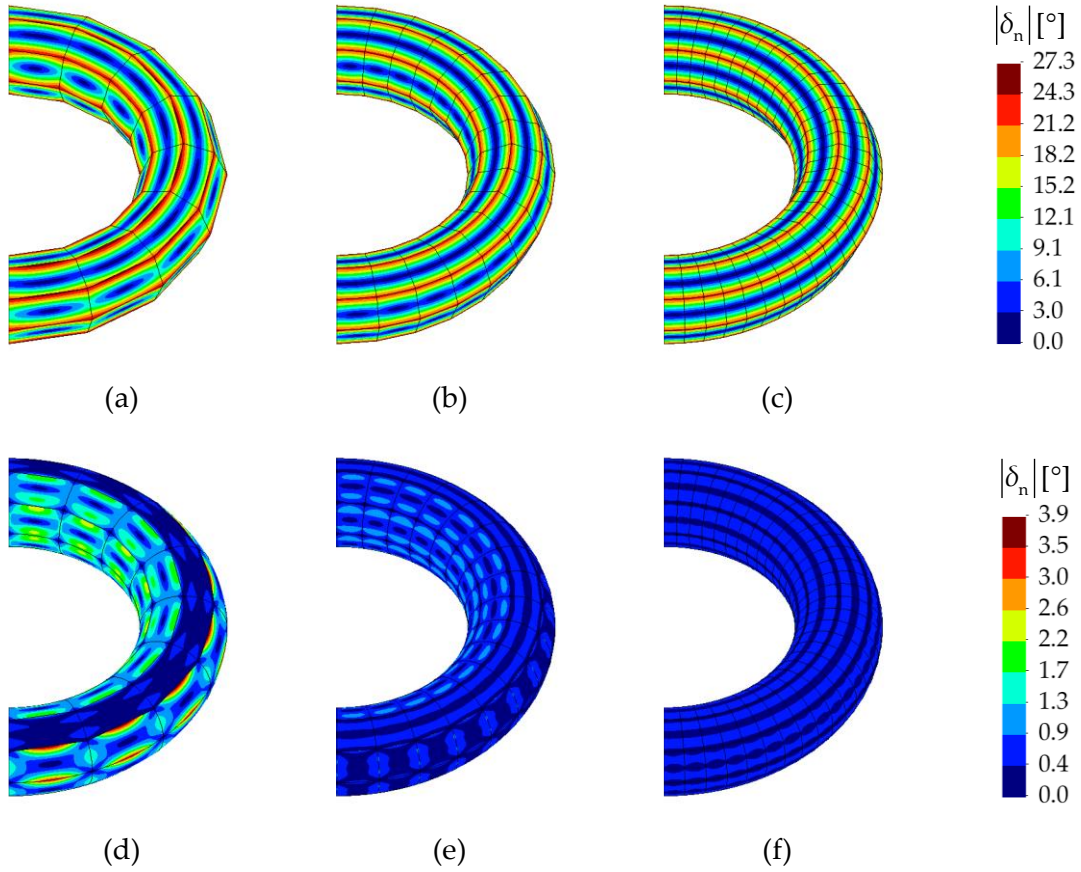


Figure 4.25. Normal vector error distribution in the toroidal surface for different discretizations: (a) faceted elements $r_t = 0.49$; (b) faceted elements $r_t = 0.98$; (c) faceted elements $r_t = 1.95$; (d) Nagata patches $r_t = 0.49$; (e) Nagata patches $r_t = 0.98$; (f) Nagata patches $r_t = 1.95$.

The range of radial and normal vector error in the toroidal surface described by Nagata patches is presented in Figure 4.26 (a) and (b), respectively, for the five discretizations shown in Figure 4.23. As seen above, both studied errors decrease with the mesh refinement in the toroidal direction. However, the rate of convergence is not linear, as highlighted in Figure 4.26. In fact, the error range decreases quickly when the r_t ratio is inferior to one, while for higher values the convergence is very slow. Moreover, the negative component of the radial error arising in the torus vanishes with the increase of r_t ratio, as shown in Figure 4.26 (a), converging to the cylinder error presented in Figure 4.14 (b) (same number of elements in the circumferential direction). The same behaviour can be observed in the normal vector error, which also converges to the error range of the cylindrical surface (see Figure 4.14 (d)). Therefore, in order to attain a good accuracy in the smoothed surface representation using the lowest possible total number of elements, the ratio r_t should be approximately one, i.e. finite elements with square topology.

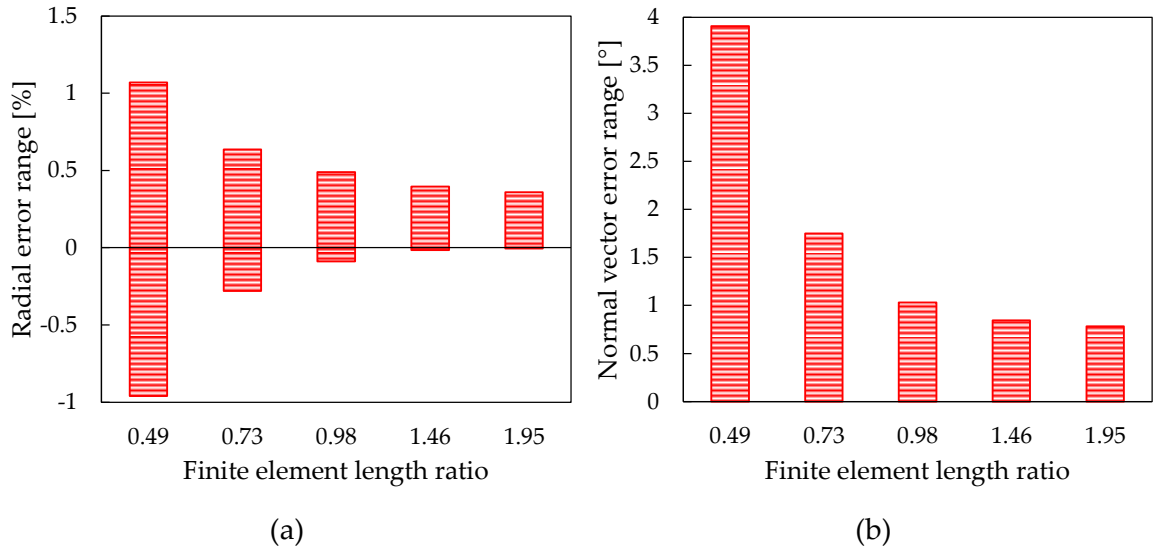


Figure 4.26. Accuracy in the smoothed surface description of the torus for different ratio values of finite element length: (a) radial error range; (b) normal vector error range.

The influence of the torus aspect ratio in the accuracy of the surface description method with Nagata patches is also evaluated using three different geometries $(R = 2, 4, 6)$, which are presented in Figure 4.27. Each half of the geometry is discretized with triangular (left half) and quadrilateral (right half) finite elements to compare both Nagata patches. Since the triangular elements are obtained from the division of the quadrilateral elements in two triangular elements (see Figure 4.27), it is required twice the number of triangular elements to describe the same geometry. As previously analysed, the equilibrium between the number of patches and the smoothed surface accuracy is given for an r_t ratio close to 1.0. Thus, this is the ratio adopted in the following discretizations. The radial error distribution in each torus geometry is presented in Figure 4.27. In case of quadrilateral patches, the maximum (positive) value of radial error occurs in the hyperbolic region of the torus. On the other hand, in case of triangular patches the extremum values (minimum and maximum) of radial error occur near the parabolic line of the torus. Despite the different locations, the maximum value of radial error is approximately the same for both finite element typologies (triangular and quadrilateral patches). The negative component of radial error in the torus with lower aspect ratio, described by quadrilateral patches, vanishes when using triangular Nagata patches (see Figure 4.27 (b)). However, the maximum error value is only slightly affected by the torus aspect ratio.

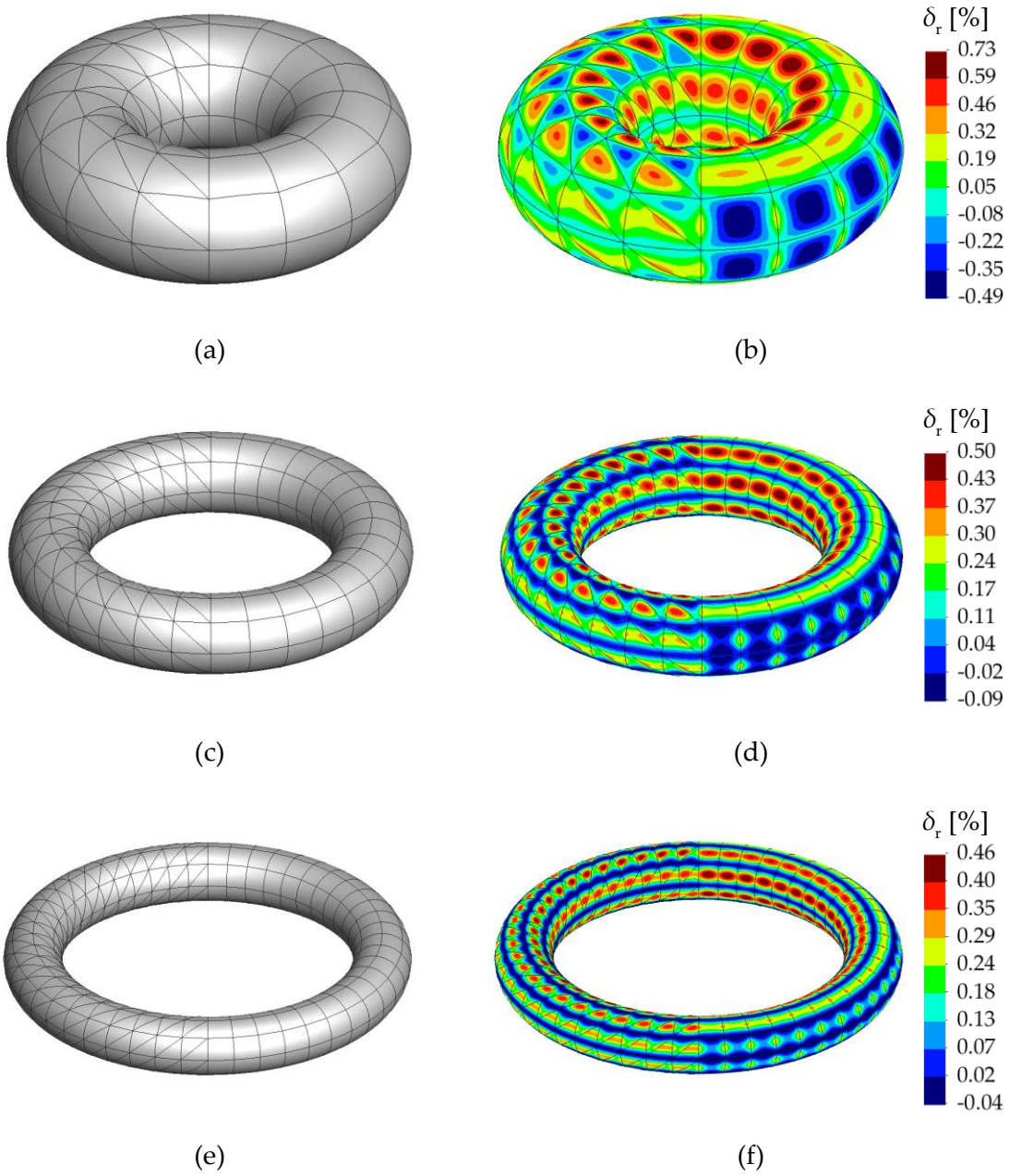


Figure 4.27. Radial error distribution in the torus surface discretized with triangular and quadrilateral finite elements for different values of major radius: (a) discretization for $R=2$; (b) error distribution for $R=2$; (c) discretization for $R=4$; (d) error distribution for $R=4$; (e) discretization for $R=6$; (f) error distribution for $R=6$.

The normal vector error distribution in the torus geometry is presented in Figure 4.28, where the accuracy of triangular and quadrilateral Nagata patches is compared for different aspect ratios. Concerning the surface description using triangular patches, the maximum value of normal vector error occurs along the patch edges unaligned with the two main directions and it is located in the hyperbolic zone (see Figure 4.28). On the other hand, the error distribution in the quadrilateral patches is considerably different since the patches are

aligned with the two principal directions of the torus. In fact, the location of the largest value of error switch from the elliptic to the hyperbolic surface when the aspect ratio of torus increases. Besides, it is somewhat affected by the torus aspect ratio, i.e. the maximum error value decreases with the aspect ratio, as shown in Figure 4.28. Nevertheless, the maximum value of error is always lower in the quadrilateral patches than in the triangular ones.

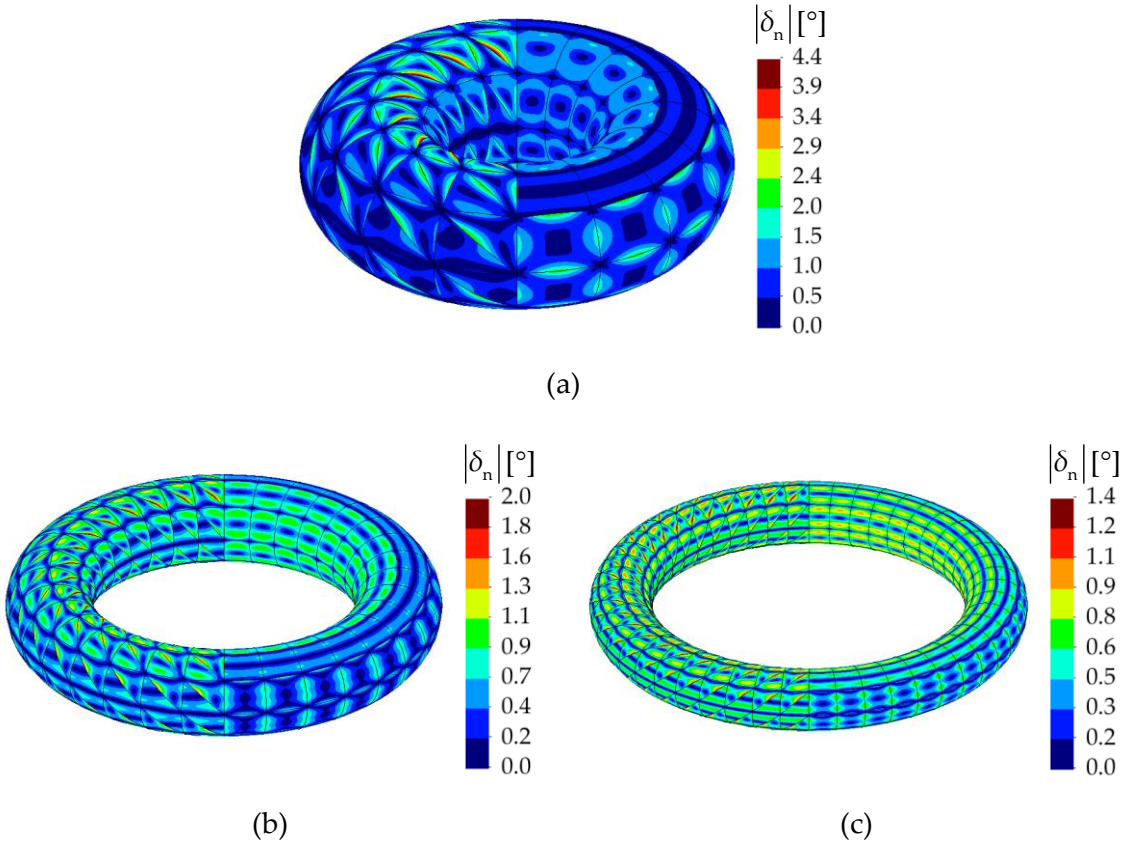
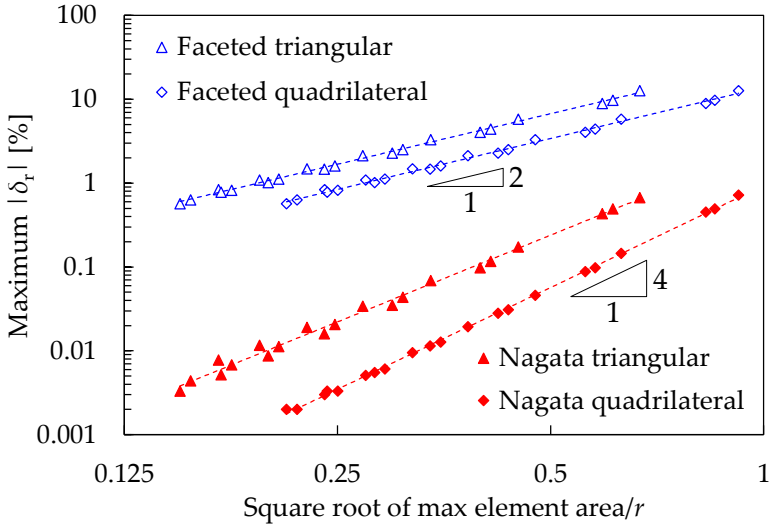


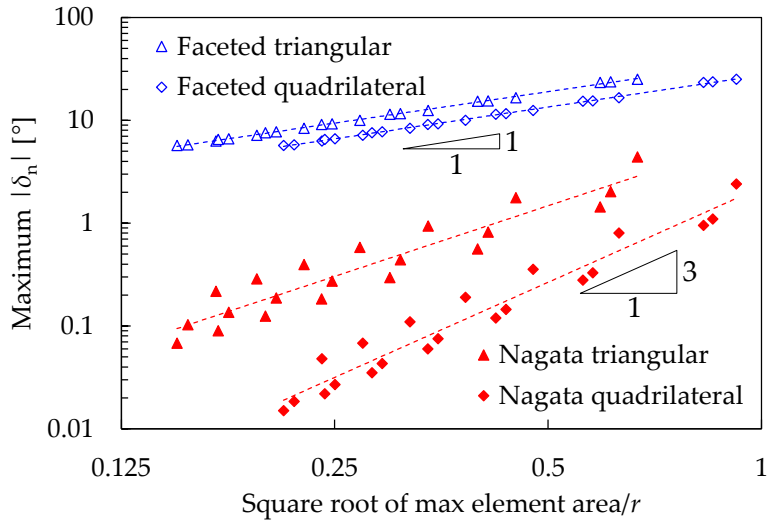
Figure 4.28. Normal vector error distribution in the torus surface discretized with triangular and quadrilateral finite elements for different values of major radius: (a) $R=2$; (b) $R=4$; (c) $R=6$.

The influence of mesh refinement on the surface description accuracy is presented in Figure 4.29 (a) through the maximum value of radial error in the torus geometry, considering three different aspect ratios. The number of finite elements in the poloidal direction ranges from 8 to 32, while the number of elements in the toroidal direction is defined in order to achieve a value of r_t ratio close to one. Comparing both surface description methods, the maximum value of radial error is significantly lower using the smoothing method (always inferior to 0.75%). Besides, the convergence in the faceted surface description is only quadratic, while the Nagata interpolation exhibits an order of convergence approximately quartic, with the square root of the maximum element area normalized by the minor radius, as shown in Figure 4.29 (a). In addition, the maximum

value of the radial error in the quadrilateral patches is always inferior to the one obtained with triangular patches of similar area.



(a)



(b)

Figure 4.29. Comparison between faceted and Nagata patch interpolation accuracy in the description of the torus: (a) maximum radial error modulus; (b) maximum normal vector error modulus.

The convergence rate of the maximum normal vector error modulus in the surface description of the torus is presented in Figure 4.29 (b), for three different aspect ratios. Considering the bilinear finite element representation, the maximum error value decreases linearly with the square root of the maximum element area normalized by the minor radius. Although the quadrilateral finite elements provide a normal vector error inferior to the one

obtained with triangular elements, the maximum value of error is always superior to 5° for the linear interpolation (see Figure 4.29 (b)). On the other hand, the description of the torus using Nagata patches leads to a significant improvement in the surface normal vector definition. However, the aspect ratio of the torus has a large impact on the maximum value of the normal vector error, as shown through the dispersion of the points around the trend line in Figure 4.29 (b). The convergence rate of the maximum normal vector error using triangular patches in the surface representation is approximately quadratic, while the application of quadrilateral patches gives a convergence rate near cubic. In fact, the quadrilateral Nagata patches are clearly advantageous in the description of the torus surface, both in terms of radial and normal vector errors (Figure 4.29).

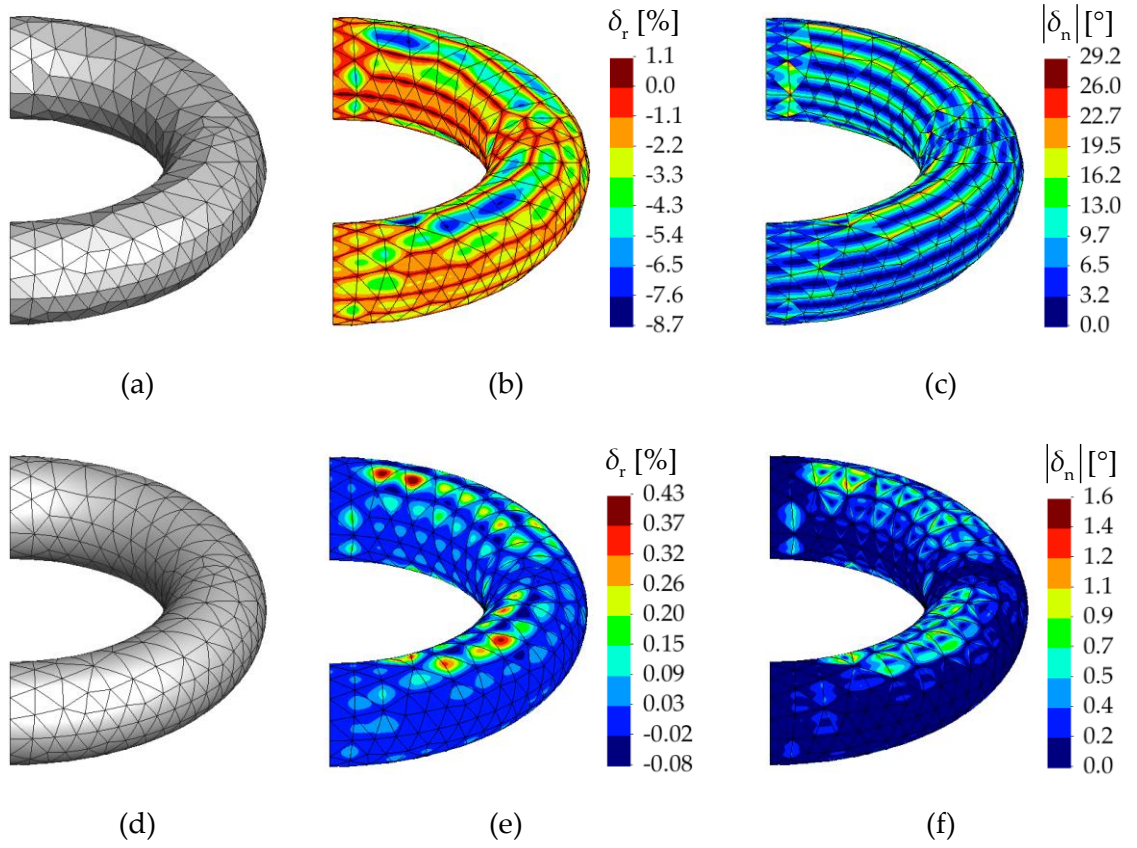


Figure 4.30. Unstructured discretization of the torus using triangular elements: (a) faceted mesh; (b) radial error in the faceted elements; (c) normal vector error in the faceted elements; (d) Nagata patches; (e) radial error in the Nagata patches; (f) normal vector error in the Nagata patches.

The distribution of both analysed errors in the description of the torus with $R=4$ and $r=1$ using an unstructured discretization composed by triangular elements is presented in Figure 4.30. The maximum value of radial error is located in the finite elements with largest area, where it is negative in the faceted surface description (Figure 4.30 (b)) and

positive for the smoothing method (Figure 4.30 (e)). However, the maximum value of radial error is about 20 times smaller when using Nagata patches. Concerning the normal vector error, its maximum occurs in the edges aligned with the toroidal direction for bilinear interpolation (Figure 4.30 (c)) and in the patch edges aligned with the poloidal direction for Nagata patches (Figure 4.30 (f)). The maximum value of normal vector error is 18 smaller in the smoothed surface than in the traditional bilinear interpolation.

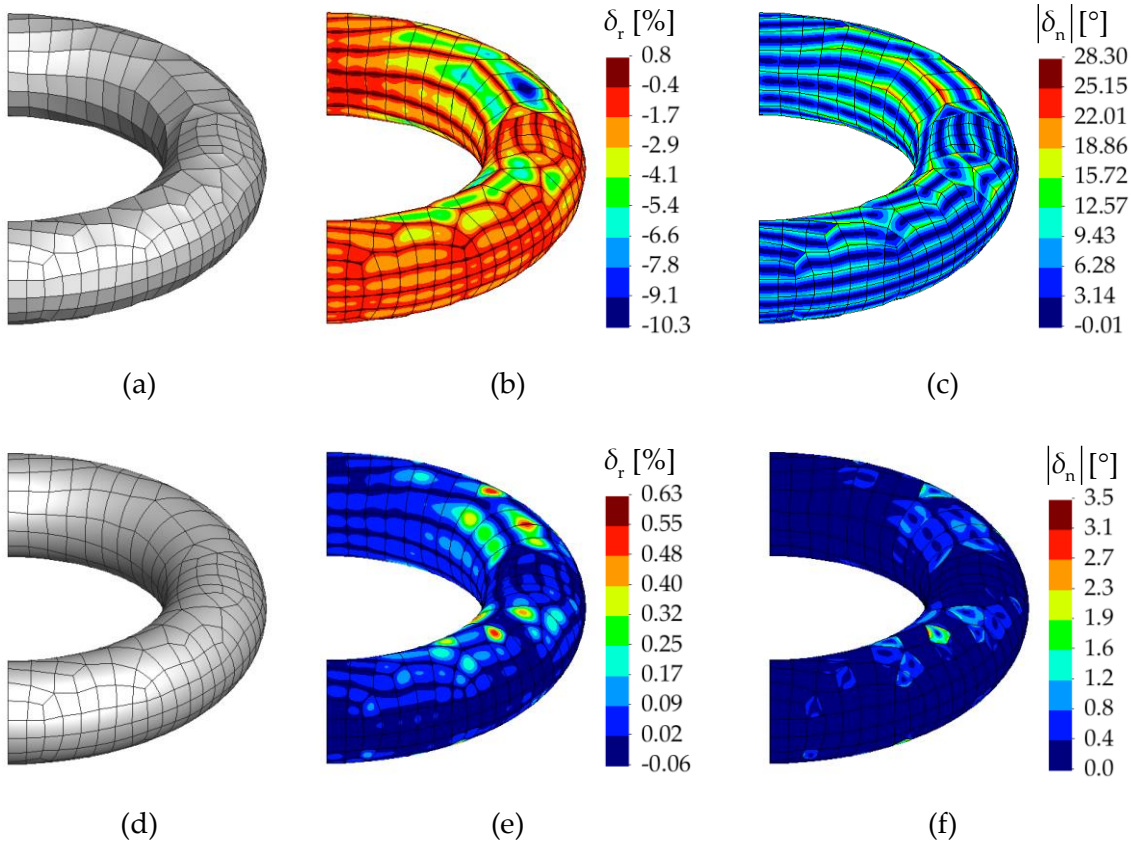


Figure 4.31. Unstructured discretization of the torus using quadrilateral elements: (a) faceted mesh; (b) radial error in the faceted elements; (c) normal vector error in the faceted elements; (d) Nagata patches; (e) radial error in the Nagata patches; (f) normal vector error in the Nagata patches.

The accuracy of each surface description method in the unstructured discretization of the torus using quadrilateral finite elements is presented in Figure 4.31. Adopting the bilinear interpolation, the maximum value of radial error arises in the finite element with largest area, while the error in the normal vector is located in the edges of these elements aligned with the poloidal direction. The surface smoothing method leads to an improvement in the accuracy of the torus geometry. The maximum value of radial error is about 15 times smaller using the Nagata patch and the maximum value of normal vector error is approximately 10 times smaller, as shown in Figure 4.31. The maximum radial error

occurs both in the patches with largest area and in distorted patches. In case of normal vector error, its maximum arises mainly in distorted patches.

4.4. Normal vectors evaluated from the CAD geometry

The Nagata patch interpolation algorithm requires the knowledge of the surface normal vector in each node of the mesh, as highlighted in the definition of the curvature parameter (4.2). Typically, when a surface is discretized, the finite element mesh file only contains the coordinates of the nodes and the finite element connectivity. Moreover, the analytical evaluation of the nodal normal vectors is restricted to simple geometries, as the ones previously considered in Section 4.3. In the particular case of the frictional contact problems involving a deformable body with various rigid obstacles (e.g. sheet metal forming), the finite element mesh of the rigid surface is generated from a model created in a CAD software package. Thus, the information available in the CAD model can be used to evaluate the nodal normal vectors. There are many different neutral file formats for data exchange between CAD and CAE software packages. The two most important and powerful neutral formats are the Initial Graphics Exchange Specification (IGES) and the STandard for the Exchange of Product model data (STEP). It is commonly stated that the IGES file format is preferable for exporting 3D surface models, while the STEP file format should be adopted for transferring 3D solid models [Basu 95], [Bhandarkar 00].

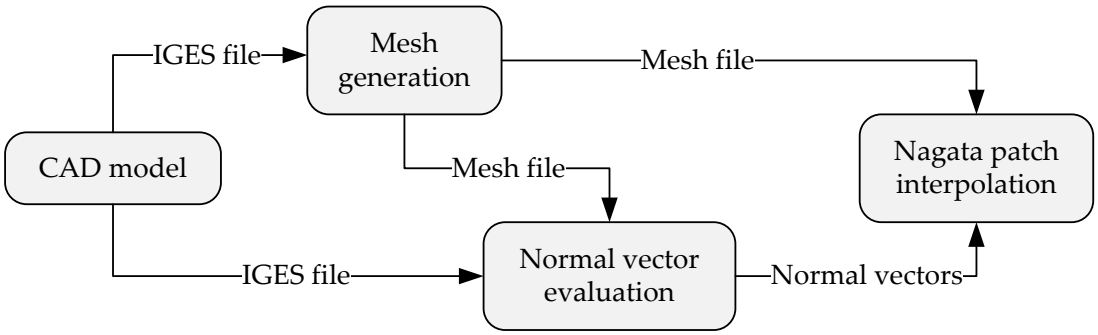


Figure 4.32. Procedure followed to evaluate the surface normal vectors from the CAD model information.

The proposed methodology to evaluate the nodal normal vectors uses the IGES file format, which contains all the information required for the mathematical definition of the surface geometry and it is organised in a structured manner, following a standard specification [IGES 96]. The geometry is represented in the form of trimmed Non Uniform Rational B-Spline (NURBS) surfaces, which result from the intersection between adjacent basis NURBS surfaces (i.e. surfaces with a rectangular parametric mapping). In fact, the trimmed NURBS surfaces are increasingly used in CAD and computer graphics due to the

high flexibility to represent complex models [Shim 00], [Litke 01], [Wang 04]. However, as discussed in Section 3.3.2, most of these CAD models are composed by trimmed NURBS surfaces that contain many overlaps and gaps [Zhu 13]. Therefore, in order to use the model in the mesh generation it is necessary to reconstruct the model into a watertight surface. The flowchart presented in Figure 4.32 shows the surface smoothing procedure using Nagata patches with normal vectors evaluated from the information contained in the CAD model. The procedure can be divided in three steps: (i) finite element mesh generation; (ii) nodal normal vectors evaluation and (iii) Nagata patch interpolation.

4.4.1. Trimmed NURBS surfaces

A trimmed NURBS surface is a basis NURBS surface bounded by a set of properly ordered trimming curves, which arise from the intersection with neighbouring surfaces or from the created holes, as schematically shown in Figure 4.33. Although the trimming curves can be of any form, when dealing with NURBS entities it is recommended to also represent them in NURBS form [Piegl 95]. Besides, the trimming curves lying within the parametric space of the untrimmed surface, as shown in Figure 4.33 (b), the boundary of a trimmed region is defined by a closed loop of trimming curves. If the loop is oriented counter-clockwise, the area within the loop is retained and the outside area is discarded. On the other hand, when the loop is oriented clockwise, the area within the loop is discarded and the outside is retained. Therefore, in order to define the geometry of a trimmed NURBS surface, this section contains a brief description of NURBS curves and surfaces, which are the two main ingredients used in the description of trimmed NURBS surfaces.

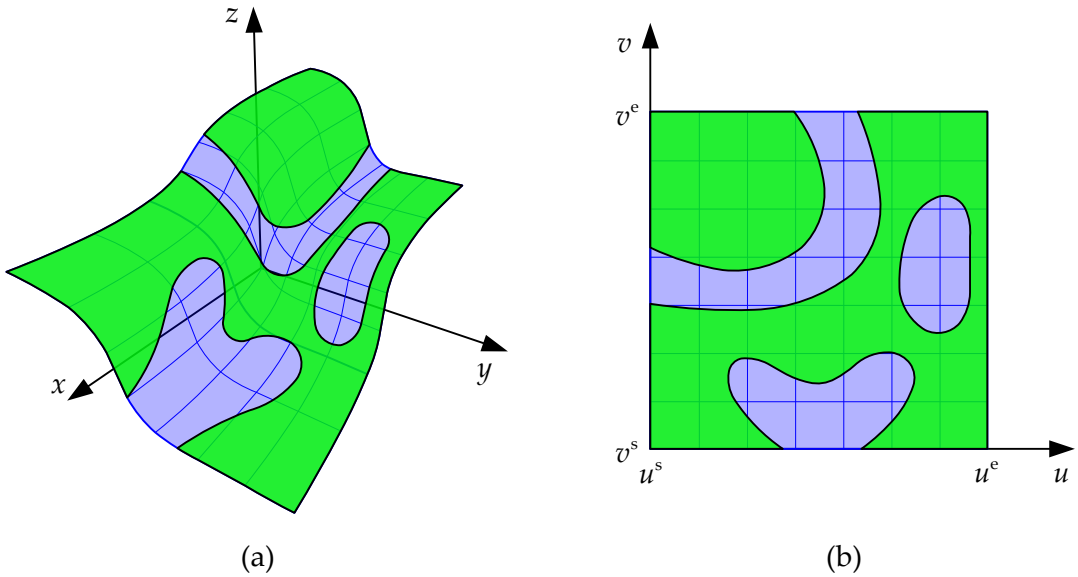


Figure 4.33. Representation of a trimmed NURBS surface in the: (a) Euclidean space; (b) parametric domain.

4.4.1.1. Definition of NURBS curves

A NURBS curve is defined by a set of weighted control points and a knot vector. It is constructed from B-spline basis functions (see Figure 4.34 (a)) using the weighted position vector of the control points as coefficients of the basis functions [Piegl 91]. They allow for an exact construction of conic sections such as circles and ellipses [Piegl 97]. The piecewise linear interpolation of the control points defines the so-called control net, as shown in Figure 4.34 (b). The shape of the curve is dictated mainly by the location of the control points. However, the curve can be pulled or pushed towards each control point through its associated weight. Mathematically, a NURBS curve of degree l is a piecewise rational function defined by a set of $h+1$ control points along with their weights, expressed by [Piegl 97]:

$$\mathbf{C}(t)=\frac{\sum_{i=0}^{h} N_{i,l}(t) w_i \mathbf{P}_i}{\sum_{i=0}^{h} N_{i,l}(t) w_i}, \quad t^{\mathrm{s}} \le t \le t^{\mathrm{e}}, \tag{4.30}$$

where w_i are the so-called weights, $\mathbf{P}_i$ are the position vectors of the control points and $N_{i,l}(t)$ are the normalized B-spline basis functions of degree l. Typically, the B-spline basis functions are defined recursively starting with zero order basis function $(l=0)$. Thus, the *i*th B-spline basis function of degree l is defined by:

$$\begin{aligned} N_{i,0}(t) &= \begin{cases} 1 & \text{if } t_i \le t < t_{i+1} \\ 0 & \text{otherwise} \end{cases}, \\ N_{i,l}(t) &= \frac{t-t_i}{t_{i+l}-t_i} N_{i,l-1}(t) + \frac{t_{i+l+1}-t}{t_{i+l+1}-t_{i+1}} N_{i+1,l-1}(t), \end{aligned} \tag{4.31}$$

which is referred to as the Cox–de Boor recursion formula [Cox 72], [de Boor 72]. Note that $N_{i,0}(t)$ is a step function, equal to zero everywhere except on the half-open interval $t \in [t_i, t_{i+1})$. Throughout the calculation of functions (4.31), ratios of the form 0/0 are defined as zero by convention.

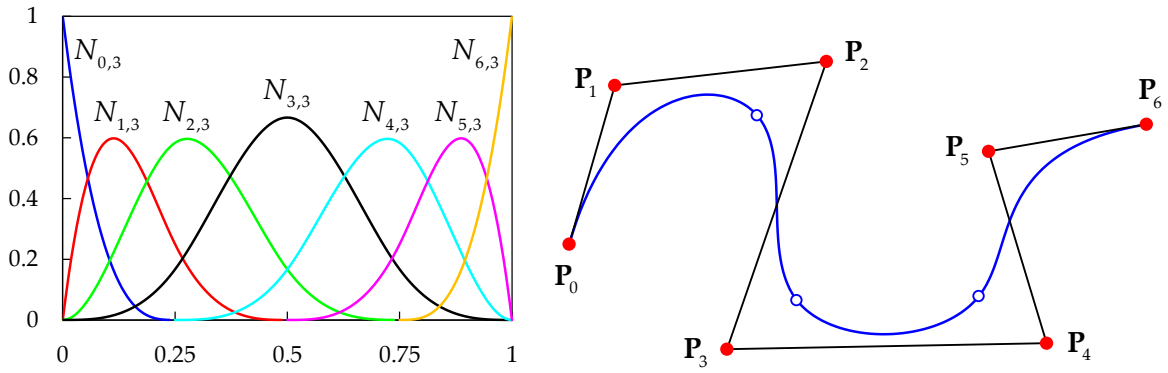
The computation of the set of basis functions (4.31) requires the specification of a knot vector [Piegl 97], which is a non-decreasing sequence of coordinates in the parameter space, defined in one dimension by:

$$\mathbf{T}=\left\{\underbrace{t^{\mathrm{s}},\ldots,t^{\mathrm{s}}}_{l+1}, t_{l+1},\ldots,t_{g-l-1}, \underbrace{t^{\mathrm{e}},\ldots,t^{\mathrm{e}}}_{l+1}\right\}, \tag{4.32}$$

where t_i are the so-called knots composing the knot vector of dimension $g+1$. The curve degree l, the number of knots $g+1$ and the number of control points $h+1$ are related

through the relationship $g = h + l + 1$. The knot vector is referred as uniform when the knots are equally spaced in the parametric space, otherwise it is defined as non-uniform. Since consecutive knots can have the same value (repeated knots), a knot vector is assumed to be open if its first and last knots are repeated with multiplicity $l+1$, which is usually found in the curves/surfaces of CAD models, and thus it is implicit in the definition presented in (4.32). The NURBS curves defined with such knot vectors start and end in a control point, while the interpolated curve is tangential at the endpoints to the first and last legs of the control net (see Figure 4.34 (b)). Moreover, in most practical applications the values of the end knots are $t^{\mathrm{s}} = 0$ and $t^{\mathrm{e}} = 1$, i.e. the knot vector is defined within the unitary parametric domain, which contains the NURBS curve (4.30).

An example of a cubic NURBS curve is presented in Figure 4.34 (b), using the B-spline basis functions shown in Figure 4.34 (a), which are defined for the open, non-uniform knot vector $\mathbf{T} = \{0, 0, 0, 0, 0.25, 0.5, 0.75, 1, 1, 1, 1\}$. Note that the interpolated curve passes at the first and last control points due to the fact that the knot vector is open. The three hollow points represented on the NURBS curve of Figure 4.34 (b) denote the division of the parametric space into four intervals, which are defined by the number of nonzero knot spans in the knot vector. Therefore, this means that each control point has a local support on the NURBS curve, i.e. a specific control point only influences the curve intervals where it is active (see Figure 4.34 (a)). B-spline and Bézier curves are particular cases (simplifications) of NURBS curves. Indeed, the B-spline curve is obtained from the NURBS curve (4.30) using equal weights for all control points. Also, the Bézier curve is a B-spline for which the non-periodic knot vector contains only one nonzero knot span and $h = l$. In fact, the degree of a Bézier curve is given by the number of control points less one, while the local support of the B-spline vanishes (change of one control point affects the whole curve).



(a) (b)

Figure 4.34. Example of a NURBS curve with equal weights: (a) cubic B-spline basis functions for open, non-uniform knot vector; (b) cubic NURBS curve with location of control points (red dots).

4.4.1.2. Definition of NURBS surfaces

A NURBS surface is obtained as the tensor product of two NURBS curves. The detailed mathematical description can be found in the literature, e.g. [Piegl 97]. Hence, a NURBS surface of deegre p in the u direction and degree q in the v direction has the following form:

$$\mathbf{S}(u,v)=\frac{\sum_{i=0}^{n}\sum_{j=0}^{m}N_{i,p}(u)N_{j,q}(v)w_{i,j}\mathbf{P}_{i,j}}{\sum_{i=0}^{n}\sum_{j=0}^{m}N_{i,p}(u)N_{j,q}(v)w_{i,j}},\quad u^{\mathrm{s}}\le u\le u^{\mathrm{e}}\quad\text{and}\quad v^{\mathrm{s}}\le v\le v^{\mathrm{e}}, \tag{4.33}$$

where $\mathbf{P}_{i,j}$ are the position vectors of the control points that form a bidirectional control net (see Figure 4.35 (a)) and $w_{i,j}$ are the weights of the control points. The number of control points is defined as $n+1$ in the u direction and $m+1$ the v direction. The $N_{i,p}$ and $N_{j,q}$ are the B-spline basis functions (4.31) of degree p and q, respectively, defined on the following knot vectors:

$$\begin{aligned}\mathbf{U}&=\left\{\underbrace{u^{\mathrm{s}},\ldots,u^{\mathrm{s}}}_{p+1},u_{p+1},\ldots,u_{r-p-1},\underbrace{u^{\mathrm{e}},\ldots,u^{\mathrm{e}}}_{p+1}\right\},\\ \mathbf{V}&=\left\{\underbrace{v^{\mathrm{s}},\ldots,v^{\mathrm{s}}}_{q+1},v_{q+1},\ldots,v_{s-q-1},\underbrace{v^{\mathrm{e}},\ldots,v^{\mathrm{e}}}_{q+1}\right\},\end{aligned} \tag{4.34}$$

where $r=n+p+1$ and $s=m+q+1$ express the size of each knot vector less one. Open knot vectors are commonly used in the definition of NURBS surfaces, i.e. the first and last knots are repeated with multiplicity $p+1$ and $q+1$, respectively, as represented in (4.34).

Figure 4.35 presents an example of a NURBS surface using equal weights for all control points. The control net is composed by 20 control points (red dots in Figure 4.35 (a)), which are distributed into a rectangular grid with 5 and 4 points in each direction. As in the case of a NURBS curve, the surface passes at the four corner control points $\mathbf{P}_{0,0}$, $\mathbf{P}_{4,0}$, $\mathbf{P}_{0,3}$, $\mathbf{P}_{4,3}$ since the two knot vectors adopted in the surface definition are open.

Currently, NURBS is the most general parametric surface description method adopted in CAD systems, since it allows a mathematical representation of both analytic and free-form shapes. However, NURBS parameterizations are not amenable to local refinement because the control points must lie topologically in a rectangular grid (see Figure 4.35 (a)). Thus, a large number of control points have no significant geometric information, serving only to satisfy topological constraints [Sederberg 04]. In order to overcome these weaknesses, the T-splines surfaces were recently developed, which are a generalization of NURBS surfaces. They can contain areas with different levels of detail, allowing a

significant reduction in the number of superfluous control points [Sederberg 04]. Its implementation and adoption in the commercial CAD software packages is currently restricted to some companies (Autodesk).

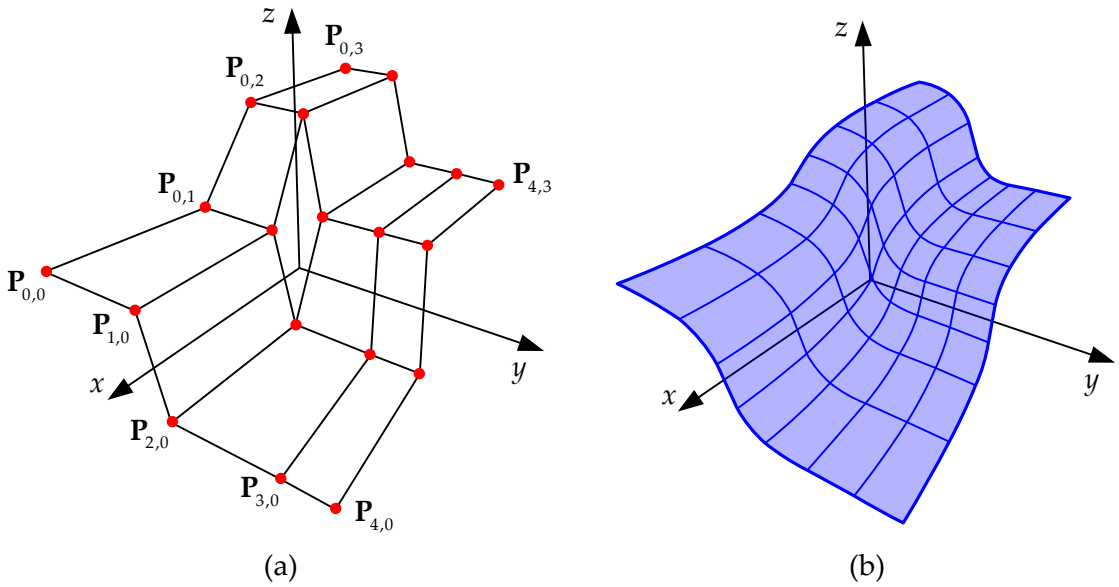


Figure 4.35. Example of a NURBS surface: (a) control points denoted by red dots (forming a control net); (b) NURBS surface.

4.4.2. IGES file structure

The purpose of this section is to provide a direct link between the mathematical description of trimmed NURBS surfaces, presented in Section 4.4.1, and the information contained within the IGES file format. In the present study, all information required to define each trimmed NURBS surface is extracted from which the neutral IGES file, an American National Standard (ANS) format, is actually the most widely used format for exchanging product data among all important CAD/CAM/CAE systems [Shim 00], [Lin 04], [Hughes 05]. This neutral file supports three types of data exchange formats: (i) fixed line length ASCII; (ii) compressed ASCII and (iii) binary [IGES 96]. Although CAD systems usually preserve transcript files in a binary format, the most commonly used format for data transfer is the fixed line length ASCII, where the entire file is divided into lines of 80 characters [Basu 95]. This format is partitioned into five sequentially numbered main sections, organized in the following order:

- Start section (S);
- Global section (G);
- Directory entry section (D);
- Parameter data section (P);
- Terminate section (T).

The columns from 1 to 72 comprise the main information, which varies according to the file current section. The file section is identified in column 73 through the letter (S, G, D, P or T), previously specified in brackets. The remainder columns (74-80) contain the line numbering for each section.

The start section (S) usually contains, in human readable free-form, comments of the sender. The global section (G) presents the principal format rules adopted the file, such as the parameter delimiter and the record terminator used in the subsequent sections, as well as the units of the model. The default symbols for the delimiter and the terminator parameters are the comma "," and the semi-comma ";", respectively. In the present application, the information presented in this section is ignored, except the delimiter and the terminator parameters. The directory entry section (D) contains two consecutive lines to define each entity. This section provides the pointer to the parameter data section (P), where the information regarding each entity is contained. Indeed, two important informations are extracted from the directory entry section (D), namely the number of starting line and the total number of lines spent in the parameter data section (P) to describe the entity. The parameter data section (P) always starts with the identification of the entity number, followed by the complete information about the parameters associated with the entity. Finally, the terminate section (T) comprises a single line describing the number of lines used in each of the previous four sections [IGES 96], [Basu 95].

The IGES file of a simple geometry composed by two NURBS surfaces is presented in Appendix B. Despite the simplicity of the selected geometry (one quarter of cylinder and an eighth of sphere), the IGES file of such surface model contains 150 lines. The directory entry and the parameter data sections are defined in 144 lines, highlighting the importance of these two sections in the IGES file structure. The data contained in the IGES file are defined as an organized collection of entities, each one denoted by a distinct number. Indeed, the IGES 5.3 [IGES 96] describes about 88 different entities, which are categorized as geometric and non-geometric. The geometric entities define the physical shape of a model including points, curves, surfaces, solids and relations between entities. On the other hand, the non-geometric entities are needed for graphical purposes, providing attributes of entities such as colour properties and luminous intensity [Bhandarkar 00].

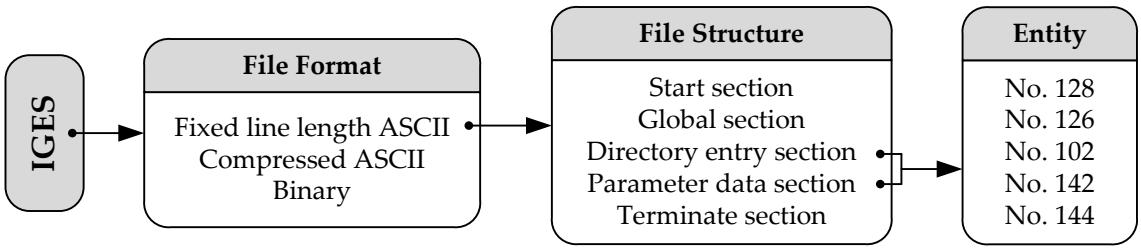


Figure 4.36. IGES file format and its division into sections, highlighting entities related with trimmed NURBS surfaces.

Table 4.3. Geometric entities required to define trimmed NURBS surfaces [IGES 96].

Entity type	Number
Rational B-Spline surface	128
Rational B-Spline curve	126
Composite curve	102
Curve on a parametric surface	142
Trimmed parametric surface	144

In the present study, only some geometric entities are required to obtain the information necessary to define the trimmed NURBS surfaces composing the CAD model. Table 4.3 presents a brief description of the required five geometric entities, as well as their identification number in the IGES specification. The pathway followed to identify all parameters related with trimmed NURBS surfaces definition is schematically shown in Figure 4.36. All parameters involved in the mathematical definition of both NURBS curves and surfaces (see Section 4.4.1) are defined through the information contained in the entities specified in Table 4.3. In order to know the data arrangement of these geometric entities within the parameter data section, a detailed description of each one is presented in the following section.

4.4.2.1. Curve and surface geometric entities

Since a trimmed NURBS surface is defined by a basis NURBS surface and a set of trimming NURBS curves, these two geometries are the principal elements under analysis. The information given by the rational B-Spline surface entity (No. 128) corresponds to the basis NURBS surface definition, previously expressed by (4.33) and (4.34). In order to easily identify the parameters involved in the NURBS surface definition, the information contained in the parameter data section associated with this entity presents the following specification:

$$128, n, m, p, q, u^{\mathrm{c}}, v^{\mathrm{c}}, pr, u^{\mathrm{p}}, v^{\mathrm{p}}, \overbrace{u_0, u_1, \ldots, u_r}^{\mathbf{U}}, \overbrace{v_0, v_1, \ldots, v_s}^{\mathbf{V}}, \overbrace{w_{0,0}, w_{1,0}, \ldots, w_{n,m}}^{w_{i,j}},$$
$$\underbrace{P^{\mathrm{x}}_{0,0}, P^{\mathrm{y}}_{0,0}, P^{\mathrm{z}}_{0,0}, P^{\mathrm{x}}_{1,0}, P^{\mathrm{y}}_{1,0}, P^{\mathrm{z}}_{1,0}, \ldots, P^{\mathrm{x}}_{n,m}, P^{\mathrm{y}}_{n,m}, P^{\mathrm{z}}_{n,m}}_{\mathbf{P}_{i,j}}, u^{\mathrm{s}}, u^{\mathrm{e}}, v^{\mathrm{s}}, v^{\mathrm{e}};$$

Note that some numerical parameters present in the above specification are not directly included in the mathematical definition of a basis NURBS surface (see (4.33) and (4.34)). This parameters arrangement can be observed in the sample IGES file of a model containing two NURBS surfaces, shown in Appendix B, as well as a brief description of each one of these parameters. For more detailed information please refer to IGES 5.3 [IGES 96]. The

rational B-Spline curve entity (No. 126) contains the information related with the NURBS curves. The specification used in the parameter data section for this entity is the following:

$$126,h,l,c^{\mathrm{p}},t^{\mathrm{c}},pr,t^{\mathrm{p}},\overbrace{t_0,t_1,\ldots,t_g}^{\mathbf{T}},\overbrace{w_0,w_1,\ldots,w_h}^{w_i},$$
$$\underbrace{P_0^{\mathrm{x}},P_0^{\mathrm{y}},P_0^{\mathrm{z}},P_1^{\mathrm{x}},P_1^{\mathrm{y}},P_1^{\mathrm{z}},\ldots,P_h^{\mathrm{x}},P_h^{\mathrm{y}},P_h^{\mathrm{z}}}_{\mathbf{P}_i},t^{\mathrm{s}},t^{\mathrm{e}},n^{\mathrm{x}},n^{\mathrm{y}},n^{\mathrm{z}};$$

The data arrangement used in the rational B-Spline curve entity (No. 126) is analogous to the one adopted in the entity No. 128. In the same way, the numerical parameters involved in (4.30) and (4.32), which define a general NURBS curve, can also be easily identified in the above specification. A brief description of the parameters used to define the rational B-Spline curve entity is presented in Appendix B, which can be easily identified in the parameter data section of the provided sample IGES file. All NURBS curves composing the model are represented following the above specification. However, when the curve is used to define the boundary of a trimmed NURBS surface, it is presented twice in the parameter data section. Thus, the same geometric curve is also specified in the parametric domain of the basis NURBS surface trimmed by the curve. The main difference to the previous specification is the domain where the control points are defined. While in the above specification they are defined in the Euclidean space $(P_i^{\mathrm{x}},P_i^{\mathrm{y}},P_i^{\mathrm{z}})$, in the alternative specification they are defined in the parametric domain of the basis NURBS surface, as follows:

$$126,h,l,c^{\mathrm{p}},t^{\mathrm{c}},pr,t^{\mathrm{p}},\overbrace{t_0,t_1,\ldots,t_g}^{\mathbf{T}},\overbrace{w_0,w_1,\ldots,w_h}^{w_i},$$
$$\underbrace{P_0^{\mathrm{u}},P_0^{\mathrm{v}},P_0^{\mathrm{w}},P_1^{\mathrm{u}},P_1^{\mathrm{v}},P_1^{\mathrm{w}},\ldots,P_h^{\mathrm{u}},P_h^{\mathrm{v}},P_h^{\mathrm{w}}}_{\mathbf{P}_i},t^{\mathrm{s}},t^{\mathrm{e}},n^{\mathrm{x}},n^{\mathrm{y}},n^{\mathrm{z}};$$

Since the parametric domain of the surface is a 2D space, the third coordinate of each control point is always zero, $P_i^{\mathrm{w}}=0.0$. Note that the curve degree, as well as the number of control points is not obligatorily the same in both specifications. In fact, the last one is useful to define the parametric domain of the trimmed NURBS surface.

The other three entities specified in Table 4.3 contain additional information to define the relationships between surfaces and curves, i.e., topological information [IGES 96]. The composite curve entity (No. 102) is an assembly of the individual simple curves that result in a continuous closed curve. Indeed, this entity is simply an ordered list of curves, where the terminate point of each simple curve is the start point of the subsequent curve. The specification adopted to indicate this entity in the parameter data section is the following:

$$102,n_{\mathrm{sc}},sc_1,sc_2,\ldots,sc_{n_{\mathrm{sc}}};$$

where n_{sc} denotes the number of simple curves that constitute the composite curve entity. The parameters sc_i denote the pointers to each i simple curve entity (No. 126), i.e. the corresponding line number in the parameter data section (values indicated in the columns 66-72 of IGES file). In fact, the pointers used in the parameter data section correspond to the line number in the directory entry section associated to the entity. Since the rational B-Spline curve entity (No. 126) is expressed in two different domains, the composite curve entity is also defined in two distinct ways. The description of each parameter involved in the composite curve entity is presented in Appendix B.

The curve on a parametric surface entity (No. 142) associates a composite curve with a surface and classifies the curve as lying on the surface. The specification used in the IGES format file for this entity is the following:

$$142, wc, se, cc_{uv}, cc_{xyz}, rep;$$

The surface entity (No. 128) on which the curve lies is identified through the pointer denoted by se. Since the composite curve entity (No. 102) can be defined in two different ways, the closed curve lying on the surface is identified by two pointers. Hence, the pointer cc_{uv} is used to express the composite curve in the parametric domain of the surface (u,v), while the pointer cc_{xyz} is applied to indicate the curve in the Euclidean space. Such as for the other entities, Appendix B presents the description of the parameters employed to define this entity.

Finally, the trimmed parametric surface entity (No. 144) contains information about the basis NURBS surface, as well as the set of trimming curves that define the boundary of the trimmed surface. The specification adopted in the IGES format to represent this entity is the following:

$$144, se, tr, n_{cci}, cps^{ob}, cps_1^{ib}, cps_2^{ib}, \ldots, cps_{n_{cci}}^{ib};$$

The pointer se indicates the basis NURBS surface entity (No. 128) and the numeric parameter tr specifies the surface domain. The value $tr = 0$ means that the domain of the trimmed surface is equal to the one of the basis NURBS surface, i.e. the surface is not trimmed. On the other hand, if the surface domain is defined by a set of trimming curves, the parameter takes the value $tr = 1$. However, according to our practical experience with IGES format, both trimmed surfaces as well as supposedly untrimmed ones are represented as trimmed entities. For more details about the parameters used to define the trimmed parametric surface entity (No. 144) see Appendix B.

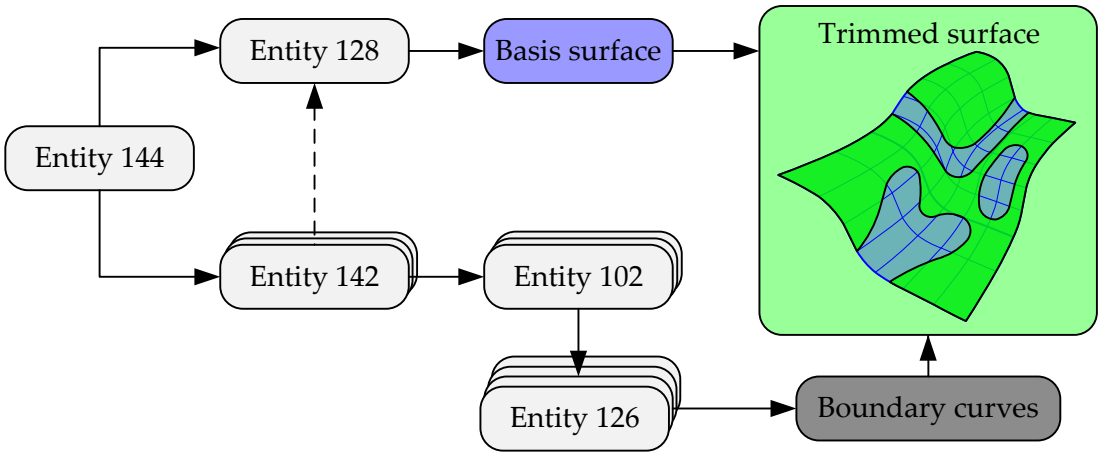


Figure 4.37. Procedure to define a trimmed NURBS surface through the entities contained in the IGES file.

The five geometry entities previously presented are connected through hierarchical relationships, as shown in Figure 4.37. This figure presents the sequential procedure performed to extract the information necessary to define each trimmed NURBS surface contained in the IGES file format. The trimmed parametric surface (No. 144) is the first entity to be identified, which holds the highest level of the hierarchy. In order to perform the data extraction from the IGES file, some routines have been developed to read the necessary entities defined in the IGES file. Since some geometric entities need large memory storage (see the rational B-Spline surface entity in Appendix B), the information about their size is essential for an efficient memory allocation. Therefore, before data extraction routine, the IGES file is read to evaluate the size of all arrays involved in the definition of trimmed NURBS surfaces. Nevertheless, the arrays used to store the IGES file information are deallocated after the evaluation of the normal vector in each node.

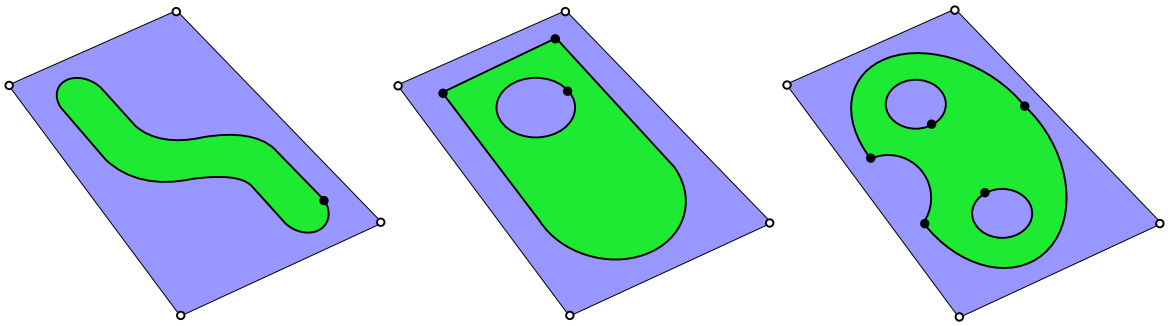
4.4.3. Normal vector evaluation

The purpose of this section is to present the method adopted for the evaluation of the surface normal vector in each node of the finite element mesh. The surface model is provided to the mesh generator using the IGES file format, where the surface discretization process creates nodes on the trimmed NURBS surfaces that compose the CAD model. Hence, the information available in the IGES file, previously used in the mesh generation, is posteriorly used to determine the nodal normal vectors. The proposed algorithm involves the following three steps: (i) surface global search; (ii) local search and (iii) normal vector evaluation. Typically, the surface geometry created by CAD systems is composed by several trimmed NURBS surfaces, reaching several hundreds of surfaces in complex models. However, each node of the surface mesh is generated over only one NURBS surface, which is unknown *a priori*. Thus, the first step of the proposed algorithm comprises

the identification of the surface on which the node is laying. After that, the coordinates of that node are evaluated in the parametric domain of the NURBS surface, where the surface normal vector can be easily calculated through the surface partial derivatives.

The aim of the global search is ordering all surfaces composing the model according to their distance to the node under analysis. Since the model can be composed by several trimmed NURBS surfaces, the local search procedure is posteriorly applied to a restricted set of surfaces selected in the global search, improving the computational performance of the proposed algorithm. The application of the global search procedure requires, for each mesh node, a loop over all trimmed NURBS surfaces composing the model. Therefore, in the global search only simple quantities (distance between points) are evaluated, avoiding the computationally more expensive calculations based on the NURBS surfaces definition [Piegl 97]. Some points of the trimmed NURBS surfaces are selected in order to evaluate these distances. In the proposed algorithm, the vertices of each trimmed NURBS surface are selected to represent a simple approximation of the surface boundary.

Although the IGES file format does not include any entity comprising the coordinates of the vertices of the surface boundaries, they can be evaluated through the end points of the trimming curves. Each trimmed NURBS surface is identified in the IGES file by means of the entity No. 144, which contains a pointer to the curve on the parametric surface entity (No. 142) defining the outer boundary surface. This closed curve is expressed by the composite curve entity (No. 102), which is an ordered list of simple NURBS curves (No. 126). This procedure is schematically shown in Figure 4.37. Besides, all curves start in the first control point and end in the last one, due to the open knot vector. Therefore, the starting point of each curve is extracted from the information available in the entity No. 126 of the IGES specification. Note that the number of extracted points is equal to the number of trimming NURBS curves.



(a) (b) (c)

Figure 4.38. Example of trimmed NURBS surfaces with outer boundary composed by few trimming curves: (a) one curve; (b) two curves; (c) three curves.

In some particular cases this approach can be inefficient due to the reduced number of points in the definition of the outer surface boundary. Figure 4.38 presents three examples of trimmed NURBS surfaces containing a reduced amount of trimming NURBS curves in the outer boundary definition. The problem is more evident in surfaces trimmed by only one curve (see Figure 4.38 (a)), for which the trimmed surface domain cannot be correctly approximated using only one vertex. Therefore, in the present study, a specific methodology is proposed when the outer boundary of the trimmed surface is defined with less than four trimming curves. In that case, instead of using the trimmed surface vertices, the four vertices of the basis NURBS surface are employed in the global search (represented in Figure 4.38 by the hollow points). The same approach is applied for the case of untrimmed NURBS surfaces.

Typically the CAD geometry is composed by many surfaces with a large difference in size (area), particularly in complex models [Oliveira 03]. Since the distance between surface vertices is larger for surfaces with big area than for small surfaces, the adoption of only the vertices can lead to some problems in the global search procedure. Therefore, the centroid of each trimmed NURBS surface is also evaluated using the outer surface vertices previously defined. Following this strategy, the first step of the global search algorithm comprises the evaluation of the vertices and centroid coordinates for each trimmed NURBS surface. Then, the Euclidean distance between the mesh node and this set of points is calculated to create an ordered list of all trimmed NURBS surfaces for each mesh node. The surfaces are ordered by increasing distance to the node, where each one is characterized by its nearest vertex point.

The local search procedure is performed to find both the correct trimmed NURBS surface and the local coordinates of the node defined in the surface domain. Since the local search is computationally more expensive, it is applied according to the ordered surfaces list previously determined. The coordinates of each mesh node in the parametric domain of the trimmed NURBS surface are obtained from the normal projection of the mesh node onto the surface [Piegl 97], [Ma 03], [Konyukhov 08]. The main feature of the projection algorithm used in the present application is the very small distance between the node and its projected point on the surface. Furthermore, the C^2 continuity is assured inside each NURBS surfaces, as well as between adjacent trimmed NURBS surfaces since the surfaces models are usually smooth and without sharp edges. Therefore, the Newton–Raphson method is used to find the normal projection point. The projection of a generic point onto a NURBS surface is presented in Appendix C, which is the base of the local search procedure.

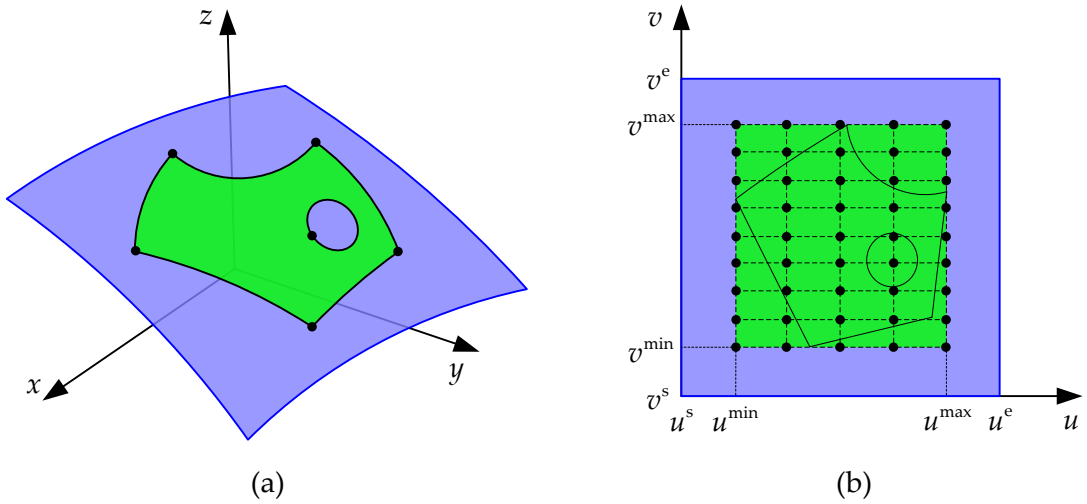


Figure 4.39. Example of trimmed NURBS surface: (a) vertices of the trimmed NURBS surface; (b) grid of points defined in the smallest rectangular domain containing the trimmed surface.

The convergence rate of the Newton–Raphson method is strongly influenced by the surface degree in both parametric directions, as well as by the initial solution selected for the iterative procedure [Ma 03]. The strategy followed in this study to improve the convergence rate is based in a good selection of the initial solution for the iterative method. Thus, a uniform grid of points is created over each trimmed NURBS surface in order to use the nearest grid point as initial solution. The number of points created in each parametric direction is related with the number of surface control points through the following expression:

$$n_{gp} = 2n_{cp} + 1, \tag{4.35}$$

where n_{cp} denotes the number of control points in the direction under analysis. In order to control the computational cost of the local search, the maximum number of grid points generated in each direction is limited to 17. Since all nodes are located within the trimmed surface domain, the grid of points is generated considering only the smallest rectangular domain covering the trimmed surface domain, as shown in Figure 4.39. This domain is defined by the vertices of the trimming curves forming the outer boundary surface (evaluated in the parametric domain of the basis NURBS surface), as shown in Figure 4.39 (b). Note that for the particular cases of trimmed surfaces with less than four trimming curves (see Figure 4.38), the same methodology adopted for the global search is employed, i.e. the domain of the basis NURBS surface is used. Thus, the values that delimit the surface domain are defined by the conditions $u \in [u^{min}, u^{max}]$ and $v \in [v^{min}, v^{max}]$, where the maximum and minimum local coordinate are obtained from the vertices of the trimming curves (see Figure 4.39).

The evaluation of the normal projection point on the NURBS surface results into a nonlinear system of equations, which is solved with the Newton–Raphson method, as presented in Appendix C. Considering the position vector $\mathbf{x}_{\mathrm{P}}$ of a generic node, which is projected onto a NURBS surface expressed by (4.33), the adopted convergence criterion is based on the simultaneously fulfilment of the three following conditions:

$$\begin{cases} \left\| \mathbf{S}(u_i, v_i) - \mathbf{x}_{\mathrm{P}} \right\| \le \varepsilon^{\mathrm{r}} \\ u^{\min} \le u_i \le u^{\max} \\ v^{\min} \le v_i \le v^{\max} \end{cases}, \tag{4.36}$$

where ε^{r} denotes the predefined threshold value for the distance between the node and the projection point, defined by the local coordinates u_i and u_i. Since all mesh nodes are generated over a trimmed NURBS surface, this distance should be very small at the solution. The last two conditions in (4.36) limit the admissible domain for the local coordinates (u, v), which is restricted to a simple rectangular domain (see Figure 4.39 (b)), reducing the possibility of projecting the node outside of the trimmed NURBS surface domain. This approximated domain can be adopted because the first condition in (4.36) ensures that the node lies on the basis NURBS surface. In fact, sometimes the trimmed curves are not really laying on the NURBS surface, but the mesh nodes are generated on the curves. Therefore, both the threshold value in (4.36) and the threshold values of the convergence criterion within the Newton–Raphson method (Appendix C) are influenced by the imperfections in the input CAD data.

The local search procedure involves the projection of the node onto each trimmed NURBS surface, since all part of the ordered list provided by the global search. However, when the Newton–Raphson method converges for a specific trimmed surface, within the maximum allowed number of iterations, that surface is selected and the local search process is terminated. Therefore, for each mesh node a single solution is obtained, i.e. the first surface of the ordered list that contains a projection point fulfilling the conditions presented in (4.36). The unit normal vector of a generic node can be evaluated based on the cross product of the first derivatives, given by:

$$\mathbf{n}_{\mathrm{NURBS}} = \frac{\mathbf{S}_u(u,v) \times \mathbf{S}_v(u,v)}{\left\| \mathbf{S}_u(u,v) \times \mathbf{S}_v(u,v) \right\|}, \tag{4.37}$$

where the partial derivatives of the NURBS surface should be evaluated in the projection point, defined in the local coordinates of the surface, as shown in Appendix C.

4.5. Normal vectors evaluated from the surface mesh

The algorithm developed to evaluate the nodal normal vectors using the IGES file of the surface is restricted to geometries for which the CAD model is available (see Section 4.4). In the general case of contact between two deformable bodies, the CAD model (IGES file) of each body is available only for the undeformed geometry (initial configuration). Indeed, the shape of the contact surface changes due to the interaction between the bodies. Therefore, the finite element discretization of the contact surfaces is the only information available to evaluate the required nodal normal vectors. Two different approaches have been developed to approximate the surface normal vectors in the nodes of a finite element mesh. The first approach is based on the fitting of a smooth parametric surface using the neighbouring mesh nodes and subsequent evaluation of the normal vector using the generated surface [Todd 86], [Meek 00], [OuYang 05]. However, the accuracy of this approach is highly dependent on the distribution of the neighbouring nodes position, as well as on the degree of the interpolated surface. On the other hand, the surface normal vector at a given node can be approximated using the weighted average of the normal vectors of facets adjacent to the node, which is the principal idea of the second approach [Page 02], [Jin 05], [Ubach 13]. This last approach is the one adopted in the present study due to low computational cost and simplicity.

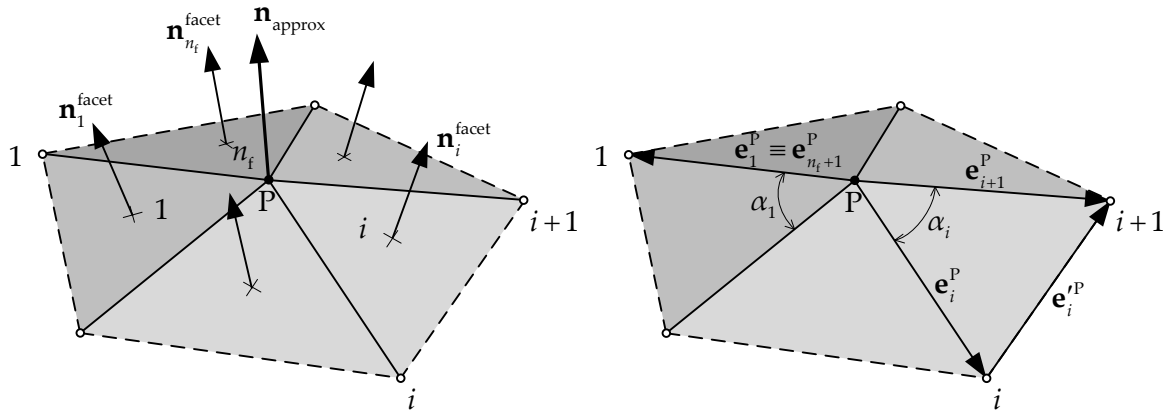


Figure 4.40. Schematic representation of the nodal normal vector evaluated through the normal vectors of the surrounding facets, including the notation adopted.

Due to the Node-to-Segment contact discretization technique adopted (see Section 3.3), only the finite element mesh of the master surface is smoothed, requiring the evaluation of its nodal normal vectors, in order to apply the Nagata patch interpolation. The normal vector in each node of the mesh is calculated by a weighted sum of the normal vectors of the facets defined by the reciprocal edges of the node. Figure 4.40 presents the notation used to define the reciprocal edges, as well as the normal vectors of each facet, which are

evaluated by the cross product of the two reciprocal edges $\mathbf{n}_i^{\text{facet}} = \mathbf{e}_i^{\text{P}} \times \mathbf{e}_{i+1}^{\text{P}} \big/ \left\| \mathbf{e}_i^{\text{P}} \times \mathbf{e}_{i+1}^{\text{P}} \right\|$. All the methods presented in this section share the concept of weighting adjacent facet normal vectors, but they differ substantially in the weighting factor adopted [Jin 05], [Ubach 13]. Various averages have been used, which take into account topological information provided by the finite element mesh.

The normal vector at a generic node of a finite element mesh surrounded by n_{f} facets is obtained from the weighted sum of the normal vector of each facet, expressed by:

$$\mathbf{n}_{\text{approx}} = \frac{\sum_{i=1}^{n_{\text{f}}} w_i \mathbf{n}_i^{\text{facet}}}{\left\| \sum_{i=1}^{n_{\text{f}}} w_i \mathbf{n}_i^{\text{facet}} \right\|}, \tag{4.38}$$

where $\mathbf{n}_{\text{approx}}$ is the approximated unit normal vector of the unknown surface at a generic node P. The unit normal vector of the *i*th finite element (facet) surrounding the node P is denoted by $\mathbf{n}_i^{\text{facet}}$, while its weight in the average is indicated by w_i. The graphical representation of (4.38) is illustrated in Figure 4.40 for a node surrounded by 5 triangular elements. When quadrilateral finite elements (generally non-coplanar) are adopted in the surface description, the normal vector of each facet required for (4.38) is evaluated using the two edges that share the node.

4.5.1. Weighting factors

As previously presented in (4.38), the nodal normal vector is defined as a weighted average of the normal vectors of the neighbouring facets (finite elements). Several weighting factors have been developed taking into account different surface properties. Since the discretized contact master surface can provide a wide range of shapes, the accuracy of the most important weighting factors is evaluated and compared.

The first algorithm was introduced by Gouraud [Gouraud 71], which will be referred in the present study as the mean weighted equally (MWE) algorithm because the weighting factor of each facet is given by:

$$w_i^{\text{MWE}} = 1, \tag{4.39}$$

where each adjacent facet contributes equally to the nodal normal vector evaluation. The second algorithm presented, which was proposed by Thürmer and Wüthrich [Thürmer 98], uses the incident angle of each facet as the weight. Defining α_i as the angle between the two edges of *i*th facet (plane), which are denoted by the vectors $\mathbf{e}_i^{\text{P}}$ and $\mathbf{e}_{i+1}^{\text{P}}$ (see Figure 4.40), the weighting factor of each facet is expressed as:

$$w_i^{\mathrm{MWA}} = \alpha_i, \tag{4.40}$$

which will be referred as the mean weighted by angle (MWA) algorithm. It assumes that the nodal normal vector is only defined locally, independent of the shape or length of the adjacent facets. The next three algorithms were developed by Max [Max 99]. The first one, referred as the mean weighted by areas of adjacent triangles (MWAAT), incorporates the area of the triangle formed by the two edges incident on the node (whether the facet is triangular or not). Thus, this algorithm assigns larger weights to facets with larger area, wherein the the weighting factor is given by:

$$w_i^{\mathrm{MWAAT}} = \left\|\mathbf{e}_i^{\mathrm{P}}\right\|\left\|\mathbf{e}_{i+1}^{\mathrm{P}}\right\|\sin(\alpha_i) = \left\|\mathbf{e}_i^{\mathrm{P}} \times \mathbf{e}_{i+1}^{\mathrm{P}}\right\|, \tag{4.41}$$

where $\mathbf{e}_i^{\mathrm{P}}$ and $\mathbf{e}_{i+1}^{\mathrm{P}}$ denote the vectors representing the edges incident on the node, as schematically represented in Figure 4.40. Max [Max 99] also proposes another weighting factor, referred as the mean weighted by edge length reciprocals (MWELR), expressed by:

$$w_i^{\mathrm{MWELR}} = \frac{1}{\left\|\mathbf{e}_i^{\mathrm{P}}\right\|\left\|\mathbf{e}_{i+1}^{\mathrm{P}}\right\|}, \tag{4.42}$$

which assigns larger weights to smaller edges. Finally, the third weighting factor proposed by Max [Max 99] considers that the surface fitting the nodes is a sphere. The algorithm is referred as the mean weighted by sine and edge length reciprocals (MWSELR), wherein the weighting factor of each facet is expressed as:

$$w_i^{\mathrm{MWSELR}} = \frac{\sin(\alpha_i)}{\left\|\mathbf{e}_i^{\mathrm{P}}\right\|\left\|\mathbf{e}_{i+1}^{\mathrm{P}}\right\|}, \tag{4.43}$$

which takes into account the differences in size of the adjacent edges, as well as the angle between them. Note that this weighting factor is obtained by adding the sine factor to (4.42), assigning larger weights to smaller edges and angles between edges close to 90°. Besides, the weights provided by (4.43) yield the exact normal vector if the discretized surface is a sphere. The last weighting factor presented in this study was recently proposed by Ubach et al. [Ubach 13] and combines the properties of different factors. The weighting factor of each facet is defined by:

$$w_i^{\mathrm{MWAAC}} = \left(\frac{\sin(\alpha_i)}{\left\|\mathbf{e}_i'^{\mathrm{P}}\right\|}\right)^2 \alpha_i, \tag{4.44}$$

where $\mathbf{e}_i'^P$ denotes the vector representing the edge opposite to the node, as represented in Figure 4.40. This algorithm is referred as the mean weighted by angle and area of circumscribed circle (MWAAC).

The accuracy of the above mentioned algorithms is typically evaluated using surface models with known analytical normal vectors. Moreover, it is assumed that the selected surface is smooth with high level of continuity and the discretization is carried out in order to have a relatively fine irregular mesh. Under these assumptions, the previously described algorithms are suitable and present a good level of accuracy, particularly the ones using the last two weighting factors defined by (4.43) and (4.44) [Ubach 13]. However, the contact surfaces under analysis are usually composed by an assembly of several flat and curved surface regions. In fact, considering the contact of a deformable body against a rigid obstacle, the rigid surfaces are discretized using much more elements in the curved zones than in the flat areas. This discrepancy in the finite element size at the boundary between these two types of surfaces leads to a deterioration of the accuracy attained in the normal vector approximation, particularly for coarse meshes. Therefore, in order to reduce the nodal normal vector approximation error in those regions, a methodology is proposed to adjust the normal vector in such nodes.

4.5.2. Adjusted normal vectors for special edge effect

The approximated nodal normal vector provided by the weighted average (4.38) may lead to inaccurate results in some locations, particularly in the intersection between two surface geometries and in the surface boundaries. Therefore, these two distinct situations should be identified *a priori* in the finite element mesh, to enable the adjustment of the nodal normal vector in order to reduce the error.

4.5.2.1. Intersection between flat and curved geometries

The intersection between flat and curved geometries is the first situation studied. The central idea of the proposed methodology is based in the comparison between the normal vector of each facet and the approximated nodal normal vector evaluated from (4.38). The first step is to identify all finite elements generated on flat surfaces and then enforce the normal vector at the mesh nodes to be equal to the flat surface normal vector. Therefore, when at least a node belonging to a finite element presents an approximated normal vector given by (4.38) identical to the normal vector of the own element, then all nodes of such element must present the normal vector of the own element.

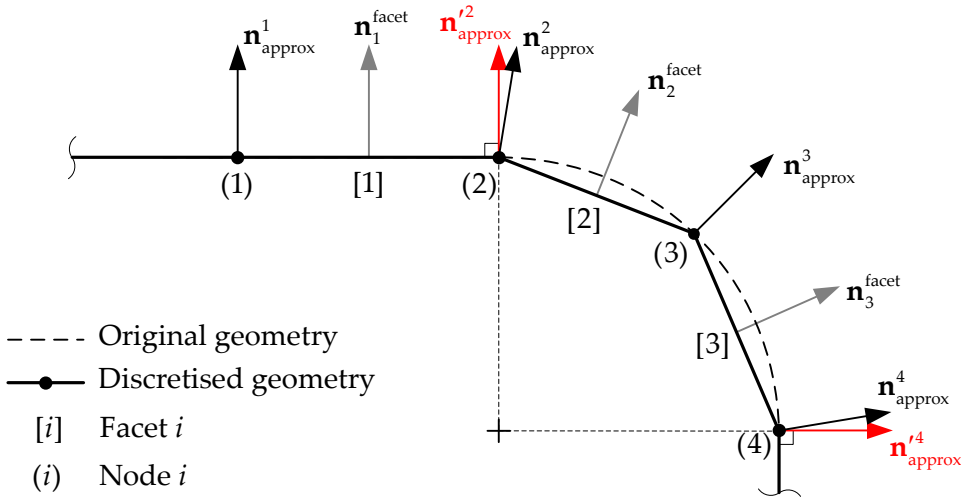


Figure 4.41. Adjusted normal vectors in the intersection between flat and curved regions (node 2 and 4 with red arrow).

The example of a 2D geometry composed by two transition zones between flat and curved surfaces is presented in Figure 4.41. This geometry is representative of several cross sections of 3D surface geometries usually involved in contact mechanics. The geometry is discretized with linear finite elements as shown in Figure 4.41, adopting a very coarse mesh to highlight the improvement obtained with the proposed correction method. The nodal normal vectors shown in the figure (denoted by $\mathbf{n}^i_{approx}$) are obtained from the application of the MWE algorithm with weighting factor given by (4.39). Considering only the nodes 2 and 4, which are located in the intersection between flat and curved regions, the angle between the normal vector defined by (4.38) and the analytical one is 11.25°. In fact, regardless of the normal vector approximation algorithm adopted, the normal vector in these nodes is always influenced by the normal vector of the finite elements 2 and 3 (weighted average), respectively. Since the normal vector of facet 1 and the one approximated in node 1 are identical (see Figure 4.41), then the proposed methodology imposes that all nodes of such finite element present the normal vector of the facet, changing the normal vector at the node 2, denoted by $\mathbf{n}'^2_{approx}$ after correction. The same procedure is carried out for the node 4, which is also located between a flat and curved surface. Indeed, the application of the proposed correction methodology to this simple example provides approximated nodal normal vectors identical to the analytical ones.

4.5.2.2. Symmetry planes

The second situation studied is the correction of the normal vectors for the nodes located in symmetry planes. The application of the finite element method to solve contact problems takes advantage of the model symmetries (geometry and load) in order to explore

efficiently the available computational resources and reduce the required computational time. However, the symmetry planes are new free boundaries of the model, where the accuracy of the nodal normal vector approximation is affected by the reduced amount of information available. In fact, it is assumed that only half model is used to calculate the normal vector in nodes belonging to the symmetry plane, reducing the number of facets involved in the average (4.38).

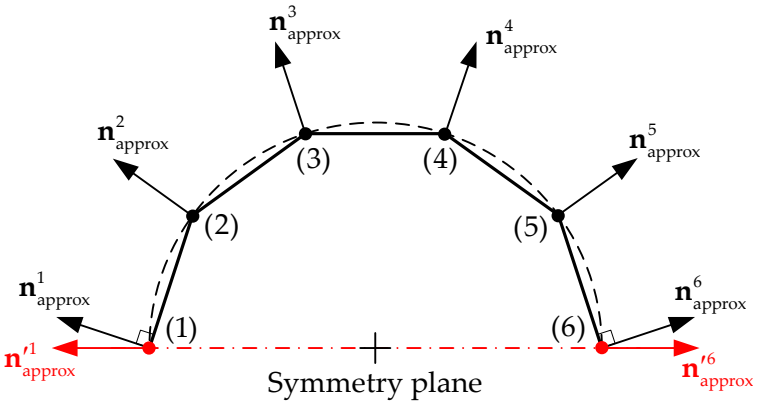


Figure 4.42. Adjusted normal vectors in the symmetry plane (node 1 and 6 with red arrow).

A simple example of a circular arc is presented in Figure 4.42, where only half model is analysed due to symmetry conditions, which is discretized with five linear finite elements (six nodes). Since the nodes located in the symmetry plane are shared by a single finite element, the approximated nodal normal vector provided by (4.38) is the same of the adjacent facet, as shown for nodes 1 and 6 in Figure 4.42. Therefore, the normal vectors of the nodes located in the symmetry planes are adjusted using the information available for the symmetry plane. Thus, the normal vector of each node is firstly evaluated using the weighted average (4.38) and posteriorly adjusted by means of its projection onto the symmetry plane. After that, it is normalized by its magnitude in order to obtain a unitary normal vector.

4.5.3. Accuracy of the approximated nodal normal vector

In order to evaluate the accuracy of each weighting factor (see Section 4.5.1) used in the approximated nodal normal vector given by (4.38), simple and complex geometries are analysed. The error in the normal vector approximation is defined in each node of the mesh, using the following expression:

$$\theta = \cos^{-1}(\mathbf{n}_{approx} \cdot \mathbf{n}_{analytical})\,[^\circ], \tag{4.45}$$

where $\mathbf{n}_{approx}$ is the approximated nodal normal vector provided by (4.38) and $\mathbf{n}_{analytical}$ is the unit normal vector evaluated from the analytical function (or CAD model). This error is evaluated only in the mesh nodes, denoting the angle between the analytical and the approximated normal vectors, expressed in degrees.

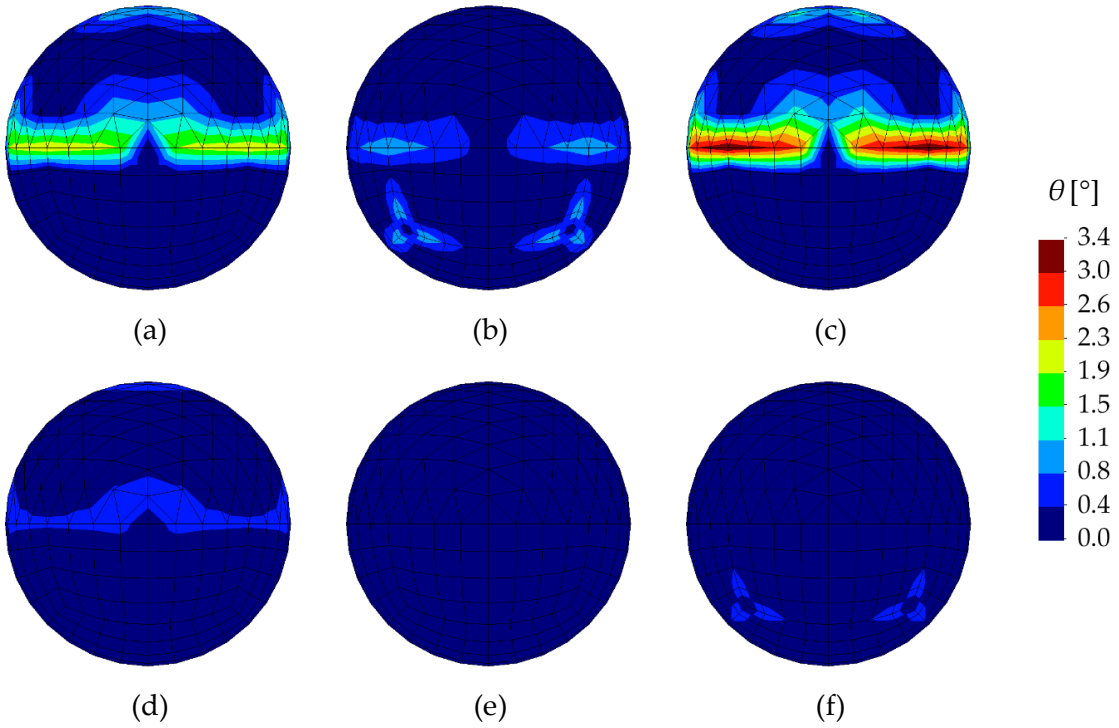


Figure 4.43. Error of approximated nodal normal vectors for triangular and quadrilateral finite elements, using different weighting factors (spherical surface): (a) MWE; (b) MWA; (c) MWAAT; (d) MWELR; (e) MWSELR; (f) MWAAC.

The first geometry analysed is the spherical surface, which is discretized with both triangular and quadrilateral finite elements, using the structured mesh presented in Figure 4.17 (c). The error is calculated in each node of the finite element mesh using the six different weighting factors previously presented in Section 4.5.1, obtaining the error distribution presented in Figure 4.43. The maximum value of error occurs in the transition between triangular and quadrilateral finite elements, except for the last weighting factor considered. This is most evident when the weighting factor is based in the areas of adjacent triangles (MWAAT), where the largest error value is about 3.4°. On the other hand, the weighting factor MWSELR gives the exact nodal normal vector because it was developed from the assumption of spherical surface [Max 99]. Comparing the six weighting factors studied, the last two ones (MWSELR and MWAAC) provide better approximations for the nodal normal vectors, as shown in Figure 4.43 (e) and (f).

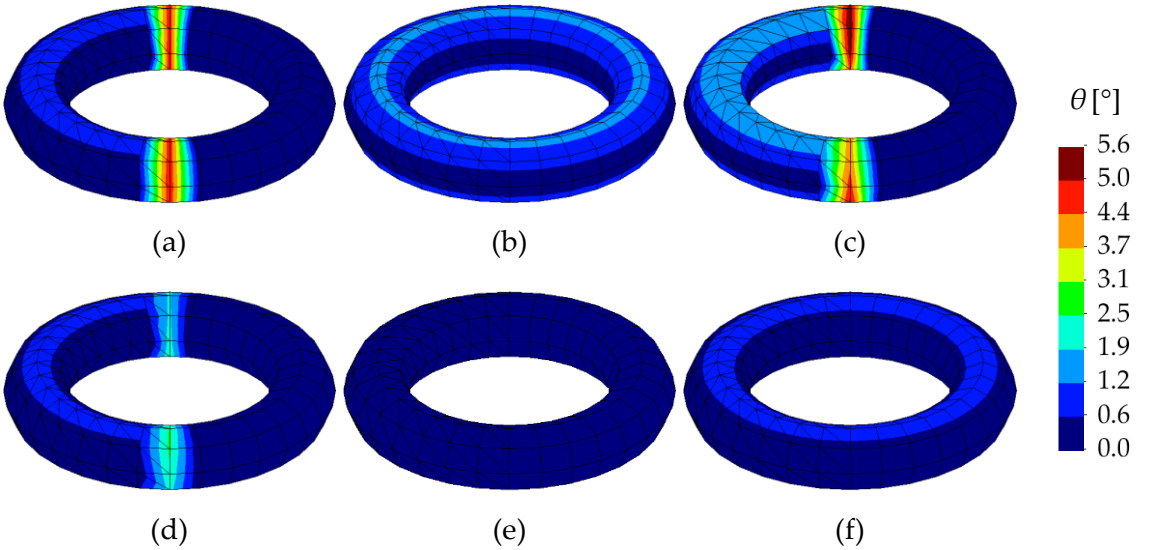


Figure 4.44. Error of approximated nodal normal vectors for triangular and quadrilateral finite elements, using different weighting factors (toroidal surface): (a) MWE; (b) MWA; (c) MWAAT; (d) MWELR; (e) MWSELR; (f) MWAAC.

The second geometry selected to evaluate the accuracy of the weighting factors is the torus with dimensions $R=4$ and $r=1$. Since this geometry presents all kinds of Gaussian curvatures, it is more illustrative of complex surface geometries. The surface is discretized using a structured mesh composed by both triangular and quadrilateral finite elements, as shown in Figure 4.27 (c).

The error in the normal vector approximation, evaluated in each node of the finite element mesh, is presented in Figure 4.44 for the various weighting factors analysed. When the MWE, MWAAT and MWELR weighting factors are adopted, the nodes located between triangular and quadrilateral finite elements comprise the largest value of angular error. The same behaviour was observed for the spherical surface shown in Figure 4.43. Although the triangular mesh (left half) comprises more finite elements that the quadrilateral mesh (right half), the error of the nodal normal vector approximation is typically inferior in the right half of the torus. This is associated with the distribution of the quadrilateral elements, which are aligned with the poloidal and toroidal directions. The weighting factor MWSELR gives the lowest value of error in the normal vector approximation, which is inferior to 0.5°, as shown in Figure 4.44 (e). On the other hand, the weighting factor leading to the worst approximation of the normal vector in case of triangular meshes is the MWAAT (see Figure 4.44 (c)).

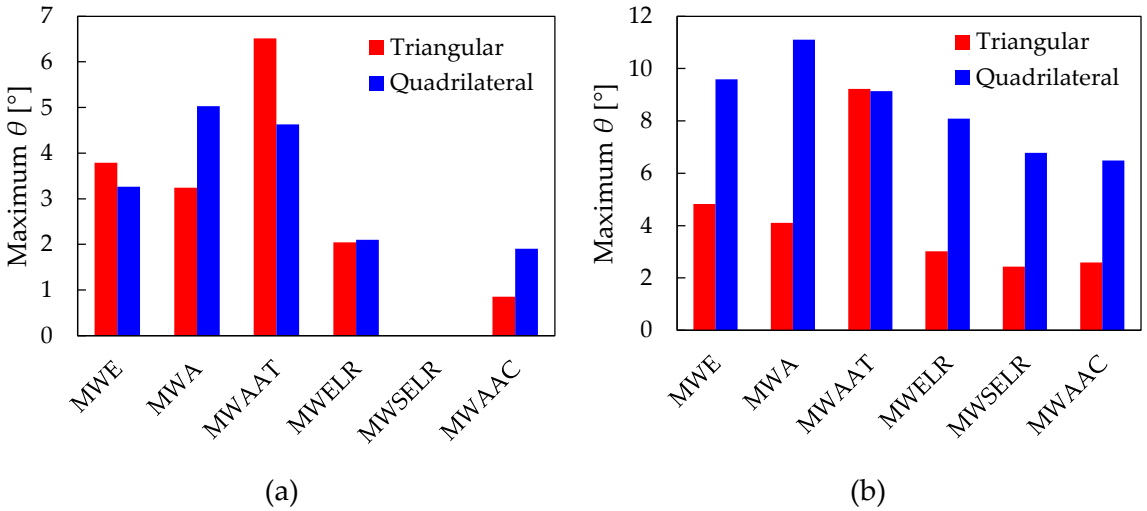
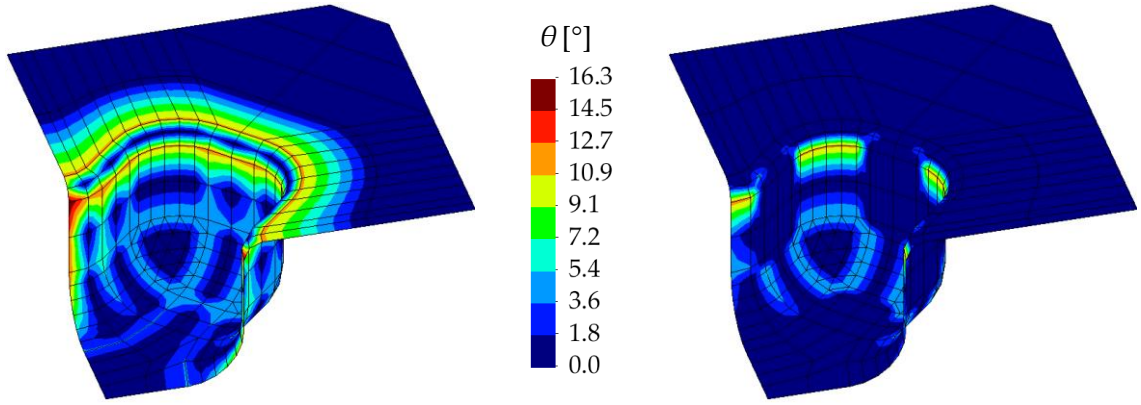


Figure 4.45. Maximum error in the normal vector approximation using different weighting factors for unstructured meshes: (a) spherical surface; (b) toroidal surface.

The accuracy of the approximated nodal normal vector expressed by (4.38) is also evaluated using unstructured meshes of the same geometries (sphere and torus). The discretizations adopted for the spherical surface are the ones previously used in Section 4.3.3, shown in Figure 4.20 (a) and Figure 4.21 (a) for triangular and quadrilateral finite elements, respectively. Concerning the torus geometry, the discretizations presented in Figure 4.30 (a) and Figure 4.31 (a) are used to evaluate the error in the triangular and quadrilateral finite element meshes, respectively. The maximum value of error in the nodal normal vector approximation is presented in Figure 4.45 for different weighting factors, applied to unstructured meshes. Regarding the spherical surface, the adoption of the MWSELR weighting factor gives the analytical nodal normal vectors (zero error), both for triangular and quadrilateral finite elements, as expected. In case of the toroidal surface, the last two weighting factors (MWSELR and MWAAC) provide the best approximation, as shown in Figure 4.45 (b). For both geometries studied, the worst approximation for the triangular mesh is obtained by the MWAAT weighting factor (see Figure 4.45). When the surfaces are discretized by quadrilateral finite elements, the maximum error in the normal vector approximation is achieved by the weighting factor based in the incident angle of the facet (MWA), as shown in Figure 4.45. Note that these conclusions, obtained with unstructured surface meshes, are also true for the structured meshes (see Figure 4.43 and Figure 4.44). Although the maximum value of error is globally lower for the toroidal surface described by triangular elements (see Figure 4.45 (b)), this direct comparison cannot be performed. Indeed, the number of finite elements is substantially higher in the mesh composed by triangles (1256 and 856 for the triangular and quadrilateral finite elements, respectively).



(a) (b)

Figure 4.46. Error in the nodal normal vector approximation of the cross tool geometry using the MWE weighting factor: (a) without adjusted normal vectors; (b) with adjusted normal vectors.

The geometry of the cross tool die used in sheet metal forming simulation is the last example selected to evaluate the accuracy of the nodal normal vector approximation. The analytical nodal normal vectors required to calculate the approximation error (4.45) are obtained from the IGES file, using the procedure described in Section 4.4. Due to geometric symmetry, only one quarter of the surface model is considered, which is discretized by a structured finite element mesh composed by 273 nodes and 196 finite elements (triangular and quadrilateral).

The principal purpose of this example is to investigate the influence of the proposed methodology to adjust the normal vectors in some nodes, presented in Section 4.5.2. The distribution of the normal vector approximation error obtained with the MWE weighting factor is presented in Figure 4.46, without and with adjusted normal vectors. The error observed in the intersection between flat and curved surfaces vanishes using the adjusted normal vectors, similar to the 2D example presented in Figure 4.41. In fact, the adjustment procedure guarantees that all nodes belonging to a flat surface present the analytical normal vector. Besides, the approximation error in the nodes located in the symmetry planes is reduced substantially. Indeed, for this example the number of nodes with error in the normal vector approximation is reduced to about by half using the adjustment normal vectors procedures proposed (see Figure 4.46).

The comparison between the different weighting factors application to the cross tool die geometry is presented in Figure 4.47, highlighting the influence of the proposed methodology to adjust the normal vectors. The angular error of the nodal normal vector approximation is represented in the histogram as function of cumulative distribution of mesh nodes. The application of the correction methodologies described in Section 4.5.2 to adjust the normal vectors improves significantly the accuracy of the nodal normal vector approximation, reducing both the maximum error value and the amount of nodes with

non-null error (see Figure 4.47). Since the discretization adopted for the cross tool geometry uses large finite elements to describe the flat surfaces, the MWAAT weighting factor provides the best approximation for the normal vector approximation. Moreover, the description of cylindrical surfaces with low number of elements in the axial direction (linear) also contributes to the accurate normal vectors estimative when the MWAAT weighting factor is adopted. Nevertheless, the examples analysed in this section show that the accuracy of each weighting factor is highly dependent of the nature of the surface geometry as well as the topology of the discretization.

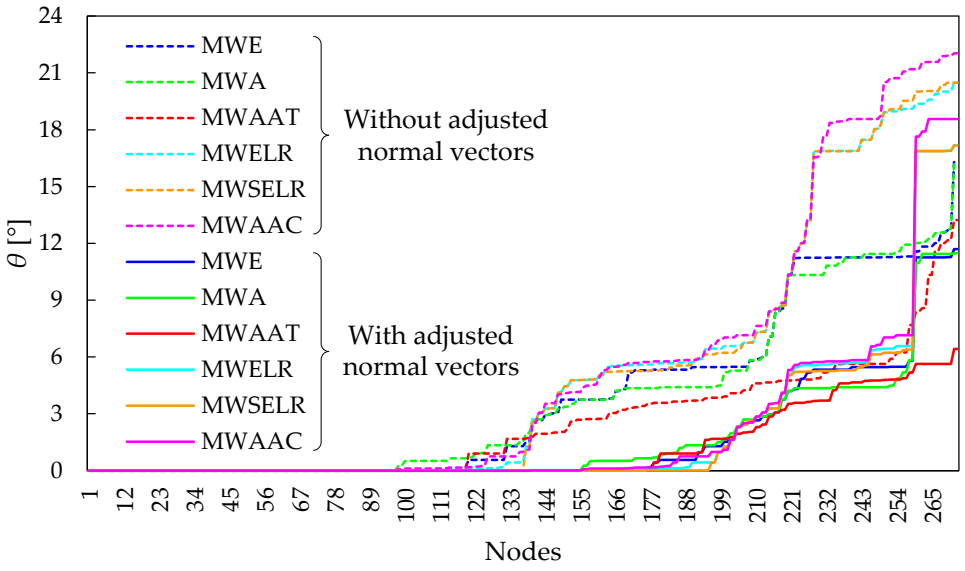


Figure 4.47. Influence of the adjustment of the normal vectors in the nodal normal vector approximation error for different weighting factors (cross tool geometry).

4.6. Effect of the normal vectors on the interpolation accuracy

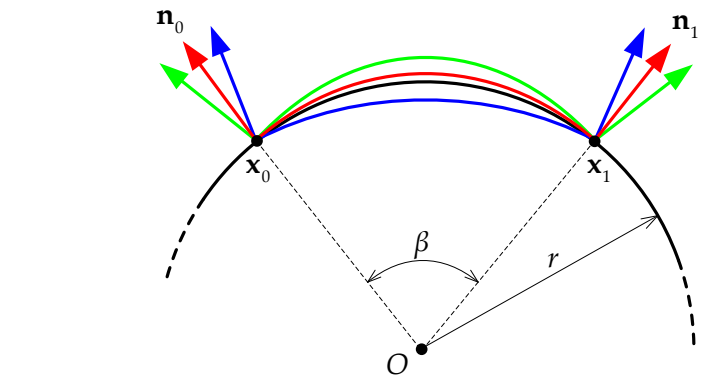
The application of the Nagata patch interpolation in the smoothing of the contact surface requires the knowledge of the surface normal vector in each node of the mesh. Moreover, the interpolation accuracy is influenced by the adopted nodal normal vectors, which dictate the boundary conditions of the patches at the nodes. Note that the modified Nagata patch interpolation proposed by [Sekine 10] is less dependent of the nodal normal vectors than the original formulation developed by [Nagata 05], as shown in the example of Figure 4.4. In fact, for a specific surface discretization, the nodal normal vectors are the only variables that influences the interpolation accuracy. The present section quantifies the impact of the error in the nodal normal vectors on the Nagata interpolation error. The simplest geometry previously analysed in Section 4.3.1 using the normal vectors evaluated

from the analytic function (circular arc) is the first geometry studied. Since the Nagata patch interpolation is completely local (each facet can be interpolated independently), the interpolation of each edge is the base to generate the interior of the Nagata patch, which highlights the importance of the circular arc geometry.

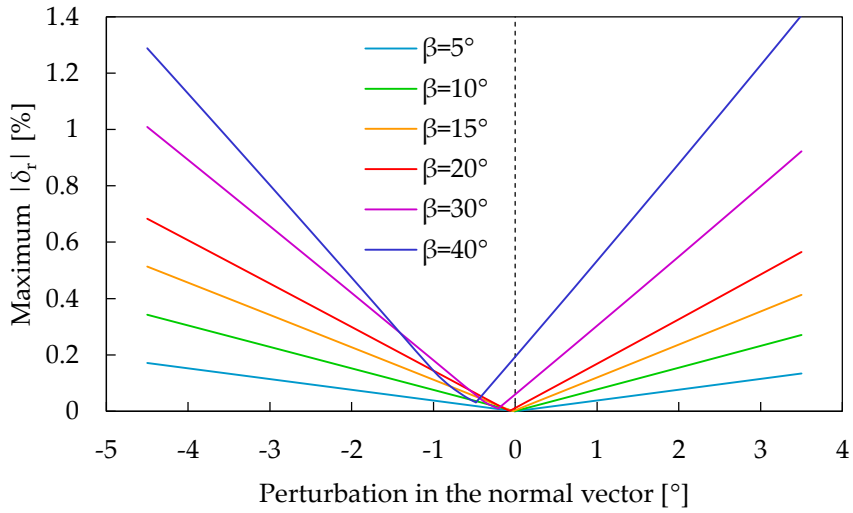
4.6.1. Circular arc

Since the accuracy of the Nagata interpolation is directly related with the discretization of the circular arc (see the convergence rate in Figure 4.13), different discretized arc lengths are used to quantity the radial error attained by the Nagata interpolation. The length of the linear segments resulting from the discretization can be defined by means of the central angle β indicated in Figure 4.48 (a). Six different central angles are used in the present analysis (5°, 10°, 15°, 20°, 30° and 40°), which are selected taking into account the typical arc length range obtained during the discretisation of curved contact surfaces. A small perturbation is introduced in the nodal normal vectors evaluated using the analytical function, in order to investigate its influence on the Nagata interpolation accuracy. Besides, the imposed perturbations are always applied symmetrically, i.e. the perturbation applied to the normal vector $\mathbf{n}_0$ presents the same amplitude and opposite orientation to the one applied to the normal vector $\mathbf{n}_1$. This symmetry is shown in the example presented in Figure 4.48 (a). The perturbation is defined as positive when the angle between both nodal normal vectors increases (green arrows in Figure 4.48 (a)), otherwise the perturbation angle is defined as negative (blue arrows in Figure 4.48 (a)).

The interpolation accuracy is evaluated in this section through the maximum radial error modulus of the Nagata interpolation. The effect of the induced perturbation in the maximum radial error modulus is presented in Figure 4.48 (b) for the six predefined central angles. The selected range for the perturbation angle is between −4.5° and 3.5°, where the trend between the maximum value of radial error obtained and the perturbation is approximately linear (see Figure 4.48 (b)). In fact, for positive values of perturbation the maximum value of radial error increases linearly with the perturbation amplitude. Since the radial error in the Nagata interpolation of a circular arc using analytical normal vectors is positive (see Figure 4.11(b)), the increase of the perturbation angle in the negative direction leads to a transition of the radial error from positive to negative values. This transition is identified through the lower value of radial error modulus, which occurs for perturbation angles slightly negative, as shown in Figure 4.48 (b). Indeed, the maximum value of radial error obtained with the Nagata interpolation using the analytical nodal normal vectors is not the lowest possible. This means that the radial error can be minimized by losing accuracy in the normal to the circular arc at the end points.



(a)



(b)

Figure 4.48. Influence of the nodal normal vectors on the interpolation accuracy: (a) schematic representation of a circular arc with central angle β; (b) maximum radial error in function of the angular perturbation induced in the normal vectors.

4.6.2. 3D geometries

The analysis of the Nagata interpolation accuracy as function of the induced perturbation angle, previously presented for the circular arc, can be extended for studying 3D geometries. In fact, each curved edge of the Nagata patches is generated from the interpolation of the linear edge using (4.1). While the normal vectors used in the circular arc interpolation were obtained from the analytic function by adding a perturbation, the nodal normal vectors used in the Nagata patch interpolation of 3D geometries are calculated by means of the approximation methods presented in Section 4.5, comparing all weighting factors studied. Indeed, if the analytical evaluation of the nodal normal vectors is impossible and the CAD geometry (IGES file) is unavailable, then the normal vector is

calculated using only the information concerning the surface discretization. This will be always the case for contact surfaces of deformable bodies.

The spherical and toroidal surfaces are analysed first considering both structured and unstructured surface meshes. The radial error distribution in the spherical surface described by Nagata patches is presented in Figure 4.49, using the nodal normal vectors obtained with different weighting factors. The radial error range is directly influenced by the accuracy of the approximation for the nodal normal vectors, which is shown in Figure 4.43. In fact, higher values of error in the normal vector approximation yields higher values of radial error in the Nagata interpolation. Moreover, due to the local behaviour of the adopted interpolation method, the maximum values of radial error are located in the vicinity of the nodes with higher value of error in the normal vector approximation (compare Figure 4.43 and Figure 4.49). Since the MWSELR weighting factor provides the analytical nodal normal vector in case of spherical surfaces, the resulting radial error distribution is identical to the one previously shown in Figure 4.17 (d). Note that the scale used in Figure 4.49 for the radial error is higher in order to take into account the increase of the error range motivated by the approximated normal vectors.

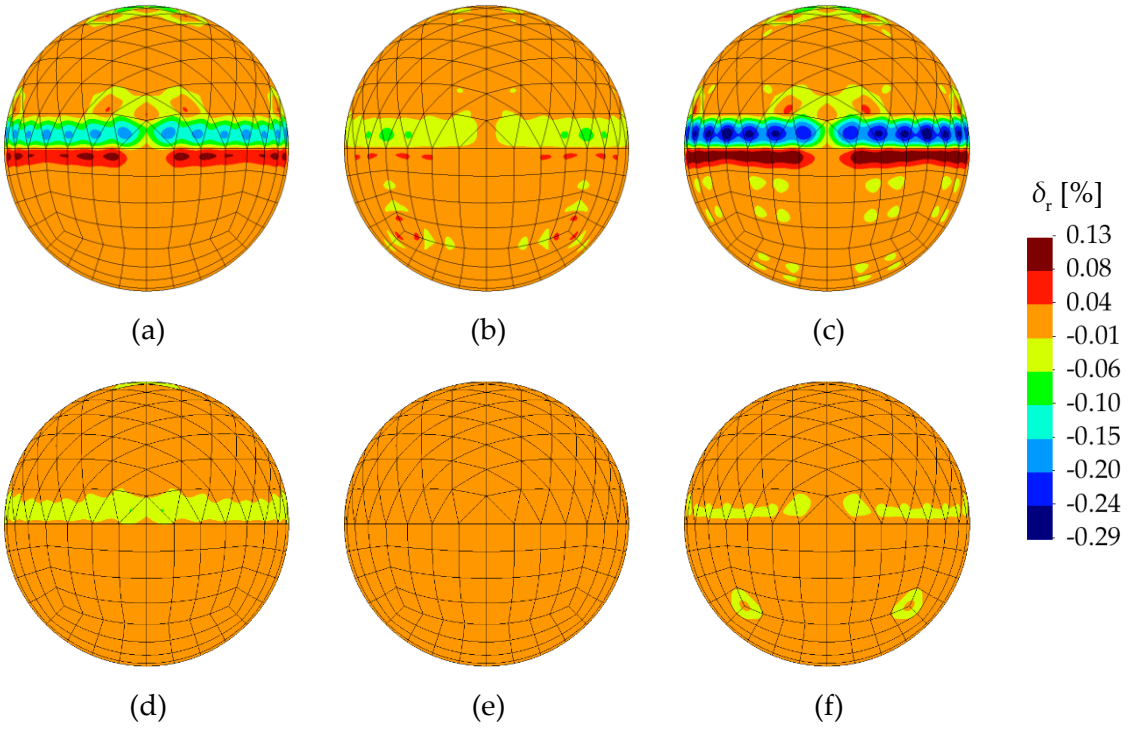


Figure 4.49. Radial error distribution in the discretized spherical surface with triangular and quadrilateral finite elements, using approximated nodal normal vectors: (a) MWE; (b) MWA; (c) MWAAT; (d) MWELR; (e) MWSELR; (f) MWAAC.

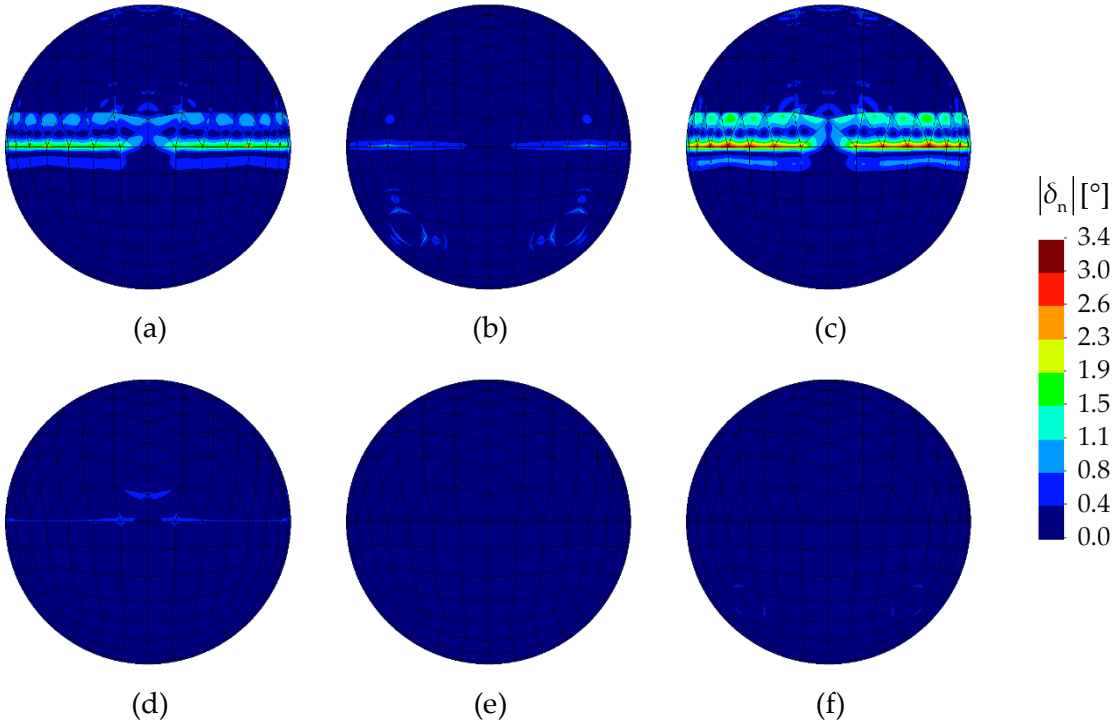


Figure 4.50. Normal vector error distribution in the discretized spherical surface with triangular and quadrilateral finite elements, using approximated nodal normal vectors: (a) MWE; (b) MWA; (c) MWAAT; (d) MWELR; (e) MWSELR; (f) MWAAC.

The use of approximated nodal normal vectors leads to the normal vector error distribution in the Nagata patches used to describe the sphere presented in Figure 4.50. The error in the approximated nodal normal vectors given by each weighting factor (see Figure 4.43) is directly reproduced in the Nagata interpolation. Indeed, the weighting factors presenting worst estimates lead to the largest values for the normal vector error. Moreover, the maximum value of normal vector error is located in the nodes, which is dictated by the error in the evaluation of the nodal normal vectors.

The radial error distribution in the toroidal surface described by Nagata patches is presented in Figure 4.51 for the different weighting factors adopted in the approximation of the nodal normal vectors. The weighting factors that provide a larger error for the nodes located between triangular and quadrilateral finite elements (see Figure 4.44) are the worst to use in the Nagata interpolation, as shown in Figure 4.51. In fact, the radial error is higher in the patches shared by such nodes. On the other hand, since the weighting factors MWSELR and MWAAC provide the best approximations for the nodal normal vectors, the radial error in the Nagata patches attains the lowest values. The radial error distribution for the same discretization, using analytical nodal normal vectors, is presented in Figure 4.27 (d), showing that it gives a more accurate interpolation. Concerning the normal vector error, the higher values are also located in the transition between triangular and quadrilateral finite elements.

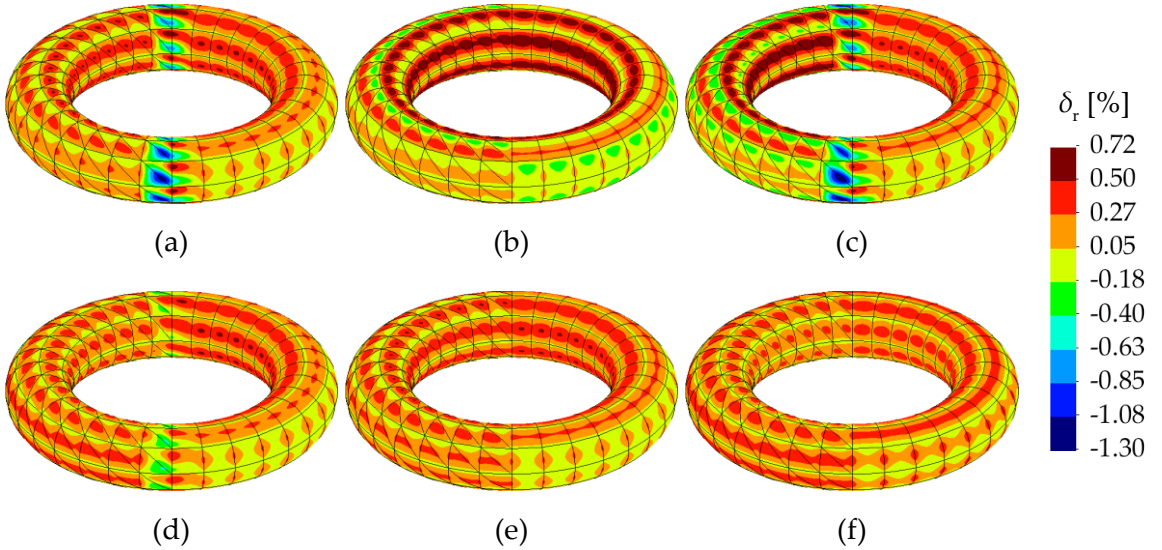


Figure 4.51. Radial error distribution in the discretized torus surface with triangular and quadrilateral finite elements, using approximated nodal normal vectors: (a) MWE; (b) MWA; (c) MWAAT; (d) MWELR; (e) MWSELR; (f) MWAAC.

Since the error in the approximation of the nodal normal vector increases when unstructured meshes are adopted, the Nagata patch interpolation accuracy is also evaluated for these cases. The unstructured triangular mesh of the spherical surface shown in Figure 4.20 (a) is the one selected to evaluate the effect of the normal vectors estimate on the interpolation accuracy. The radial error range in the Nagata patch interpolation of the spherical surface is presented in Figure 4.52 (a), comparing the approximated nodal normal vectors with the analytical ones (IGES). The comparison of the maximum error in the normal vector approximation, presented in Figure 4.45 (a), with the radial error range highlights the direct correlation between them. The best (worst) weighting factor used in the normal vector evaluation provides the most accurate (inaccurate) interpolation.

Concerning the second geometry analysed, the unstructured triangular mesh adopted for the torus surface is shown in Figure 4.30 (a). The radial error range in the Nagata patch interpolation using approximated nodal normal vectors is illustrated in Figure 4.52 (b). The worst weighting factor to evaluate the nodal normal vector is the MWAAT, defined by (4.41), as shown through the maximum error value presented in Figure 4.45 (b). Accordingly, the application of such weighting factor in the nodal normal vector evaluation leads to the interpolation with higher radial error range. In fact, for both surface geometries studied with unstructured meshes (sphere and torus), this weighting factor provides the worst approximation to describe the surface using Nagata patches, as shown in Figure 4.52. However, the range of the radial error attained with the smoothing method when adopting this worst weighting factor is always significantly inferior to the one obtained with the faceted description of the surfaces. Can be seen by comparing this range with the one shown in Figure 4.20 (b) and Figure 4.30 (b) for the sphere and torus geometry, respectively. Since

typically the unstructured meshes are generated using triangular finite elements, only results obtained with triangular elements are presented.

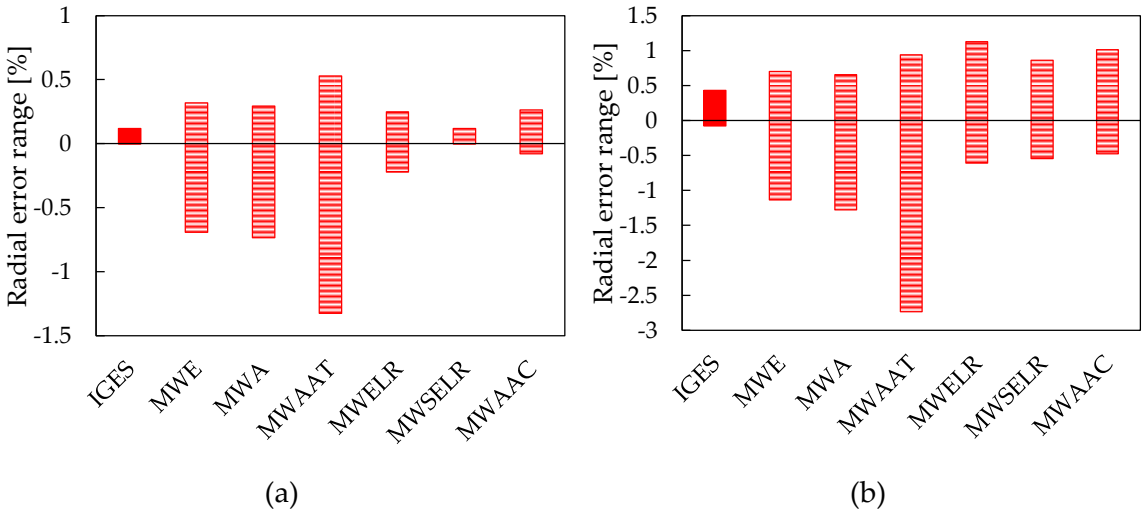


Figure 4.52. Radial error range in the Nagata patch interpolation for unstructured meshes composed by triangular elements: (a) spherical surface; (b) toroidal surface.

The last geometry studied is the cross tool, which involves the assembly of different simple geometries (planes, cylinders, spheres and tori). The effect of the nodal normal vectors estimate in the accuracy of the Nagata patches interpolation is analysed only for the MWE weighting factor, since it is the simplest. The shape error distribution in the patches, which is defined by (4.26), is presented in Figure 4.53 (a) and (b) using analytic and approximated nodal normal vectors, respectively. The error in the normal vector approximation is presented in Figure 4.46 (b) for the selected weighting factor with adjusted normal vectors. Considering the nodal normal vectors evaluated from the CAD geometry using the IGES file, the maximum value of shape error is inferior to 45 μm (Figure 4.53 (a)). In fact, the maximum (positive and negative) values of shape error occur in the toroidal surfaces in hyperbolic and elliptic surface points. On the other hand, the adoption of nodal normal vectors evaluated using only the mesh information yields higher error in the patches shared by the nodes with inaccurate normal vectors, as shown in Figure 4.53 (b). The range of the shape error is dictated by the patches located in the cylindrical surface of the vertical wall, and it is about 7 times higher than the one obtained using analytic nodal normal vectors. Since the normal vectors are adjusted in the nodes belonging to the flat surfaces (see Section 4.5.2.1), the shape error in such surfaces is zero.

The normal vector error distribution in the Nagata patches used to describe the cross tool geometry is presented in Figure 4.54 (a) and (b), for the interpolation obtained using analytic and approximated nodal normal vectors, respectively. Considering the nodal normal vectors evaluated from the CAD geometry, the normal vector error occurs mainly

in the patches located in the fillet created between the flat horizontal surface and the vertical wall. On the other hand, the adoption of approximated nodal normal vectors (MWE weighting factor) increases both the maximum value of the normal vector error and the amount of patches with higher error value, as shown in Figure 4.54. Note that the maximum error value attained in the approximation of the normal vectors (see Figure 4.46 (b)) is higher than one obtained in the Nagata patches, which is located near the nodes. This is related with the adoption of the modified Nagata interpolation [Sekine 10], which does not satisfies the boundary conditions, as illustrated in Figure 4.4.

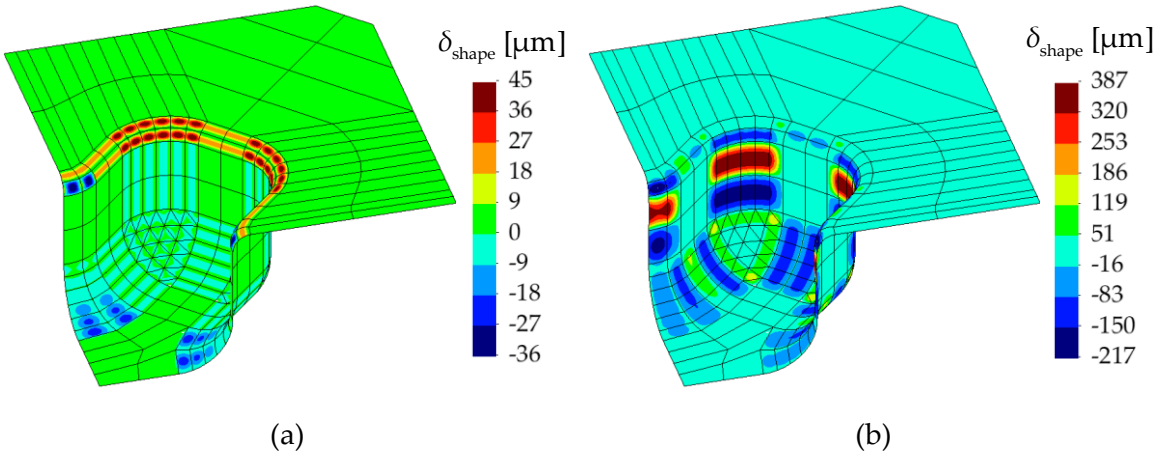


Figure 4.53. Shape error distribution in the cross tool geometry using Nagata patches with nodal normal vectors evaluated from: (a) CAD model; (b) surface mesh using the MWE weighting factor.

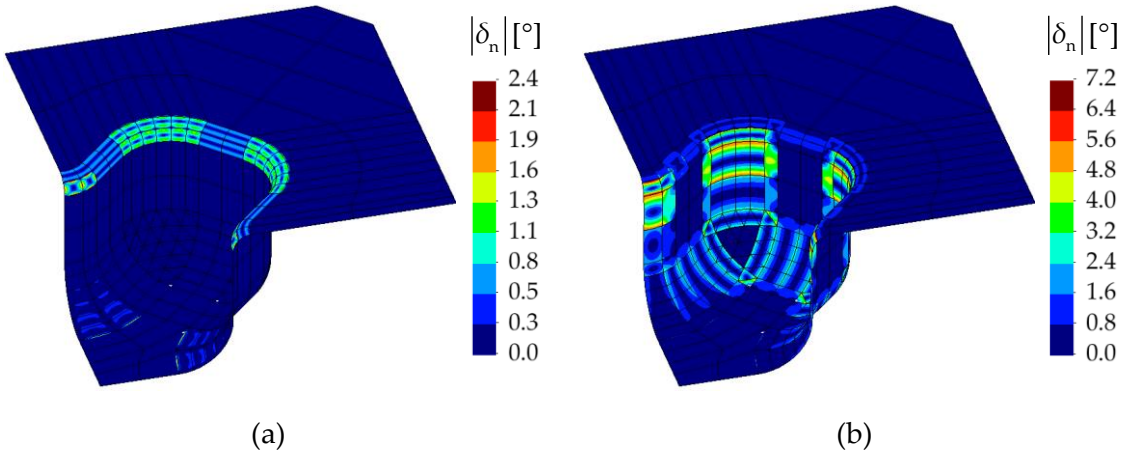


Figure 4.54. Normal vector error distribution in the cross tool geometry using Nagata patches with nodal normal vectors evaluated from: (a) CAD model; (b) surface mesh using the MWE weighting factor.

Chapter 5

Node-to-Nagata Contact Elements

This chapter contains the formulation of the contact elements developed in the framework of the Node-to-Segment contact discretization. Due to the geometrical description of the master surface with Nagata patches, they are called Node-to-Nagata contact elements. Since the contact with a rigid surface does not requires the transmission of the contact forces from the slave nodes to the rigid surface, the simplified formulation of the contact element is firstly presented. The residual vectors and tangent matrices of the developed contact element are derived for the augmented Lagrangian method, for all contact statuses. The contribution of the contact elements to the virtual work leads to a mixed system of equations with nodal displacements and contact forces as unknowns. The nonzero pattern of the resulting global tangent matrix is briefly discussed. The reduced system of equations is derived from the mixed system, obtaining a system of equations involving only nodal displacements as unknowns. On the other hand, the contact between deformable bodies requires the transmission of the contact forces arising in the slave nodes to the master nodes, composing the contact element. This connection is replicated both in the residual vectors and in the tangent matrices of the contact elements, which are presented for triangular and quadrilateral master Nagata patches. The effect of the large sliding on the nonzero pattern of the global tangent matrix is addressed. In order to improve the computational time, the multi-face contact elements are presented and discussed.

5.1. Frictional contact with curved rigid surfaces

The frictional contact between a deformable body (slave) and a rigid obstacle (master) is analysed using the Node-to-Surface contact discretization technique, previously

presented in Section 3.3.2. The kinematic contact variables (normal gap function and tangential relative sliding) are evaluated at the nodes of the slave surface (slave nodes), while the master surface is described by Nagata patches (see the surface smoothing method presented in Chapter 4). The interaction between the slave nodes and the Nagata patches, composing the master surface, is defined by means of the contact elements, which connect these two components. In the case of contact with an obstacle, the geometrical part of each contact element is composed by a slave node and the associated Nagata patch (determined based on the closest point on the master surface). Since the augmented Lagrangian method is adopted to formulate the frictional contact problem [Alart 91], each contact element is complemented by an artificial node to store the contact force (Lagrange multipliers), as illustrated in Figure 5.1. Note that the contact forces arise only on the slave nodes due to the violation of geometric and friction law constraints, but are not transferred to the master surface (the rigid obstacle has no degrees of freedom associated).

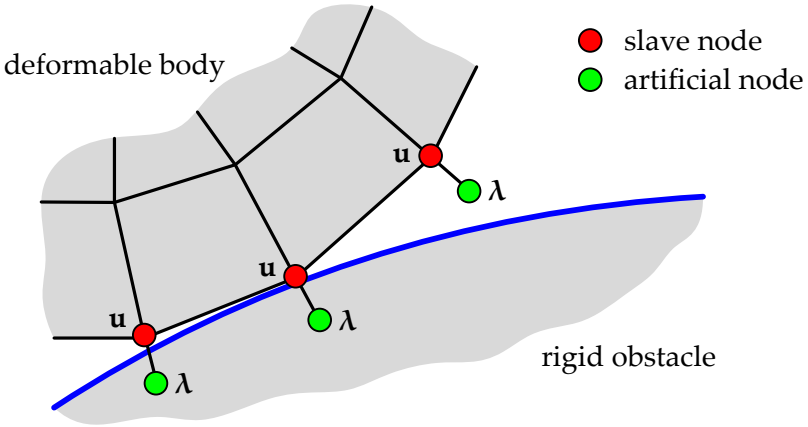


Figure 5.1. Definition of the contact elements in the slave surface of the discretized deformable body (contact force defined in the artificial node).

The virtual work balance for the frictional contact problem, derived from the variation of the augmented Lagrangian functional, was given in (2.90) for continuous problems. The contact surfaces and the related frictional contact forces are incorporated in the integral weak form, by evaluating the work of the contact forces, for given virtual displacements within the contact interface. The integrals related with the contact contribution are evaluated in the master surface, which is composed by a set of Nagata patches in the finite element model. Adopting the augmented Lagrangian method to handle the inequality constraints, all contact elements (active and inactive) contribute to the virtual work of the system. This particular feature ensures the smoothness of the energy potential and the continuity of the virtual work functional (see Figure 2.19 for a spring in contact with a rigid wall) [Pietrzak 97].

Following the finite element discretization of the deformable body and the Node-to-Surface contact interface discretization, the virtual work of the discretized contact problem is obtained from (2.90) by replacing the exact fields by their approximations. In the case of *quasi*-static frictional contact problems, the resulting problem is a system of nonlinear algebraic equations with nodal displacements and contact forces as unknowns [Parisch 89], [Klarbring 92], [Laursen 95]. Concerning the assembly of all elemental contributions to the virtual work balance equation, the present study is focused on the contact elements, while the contribution of the structural finite elements is not addressed. In fact, the routines required to perform the assembly of the structural finite elements can be found, for example, in book of Zienkiewicz and Taylor [Zienkiewicz 00a], [Zienkiewicz 00b]. Denoting by δW_j^{s} the contribution of the jth structural element to the virtual work and by δW_i^{c} the contribution of the ith contact element to the total virtual work, the equilibrium condition is expressed by:

$$\sum_{j=1}^{Ne} \delta W_j^{\mathrm{s}} + \sum_{i=1}^{S} \delta W_i^{\mathrm{c}} = 0 \quad \rightarrow \quad \begin{Bmatrix} \delta \mathbf{u} \\ \delta \boldsymbol{\lambda} \end{Bmatrix}^{\mathrm{T}} \begin{Bmatrix} \mathbf{F}_u^{\mathrm{s}} + \mathbf{F}_u^{\mathrm{c}} \\ \mathbf{F}_\lambda^{\mathrm{c}} \end{Bmatrix} = 0, \tag{5.1}$$

where Ne is the total number of structural elements and S is the number of contact elements, which is equal to the number of slave nodes. Therefore, the number of algebraic equations expressed by (5.1) is equal to the number of free degrees of freedom (nodal displacements) plus the number of contact elements, both multiplied by the dimension of the problem (2D or 3D). The second expression in (5.1) provides a connection with the classical notation of the finite element method, which is composed by the full residual vector and the vector of virtual unknowns. This matricial form of the total virtual work presented in the right-hand side, involves the residual vector related to structural elements $\mathbf{F}_u^{\mathrm{s}}$ and the residual vector related to the contact elements $\mathbf{F}_i^{\mathrm{c}}$ $(i = u, \lambda)$, which are defined by the variation of the augmented Lagrangian functional (2.90). Since virtual nodal displacements and virtual Lagrange multipliers are arbitrary, the equilibrium condition (5.1) is satisfied only if:

$$\begin{Bmatrix} \mathbf{F}_u^{\mathrm{s}} + \mathbf{F}_u^{\mathrm{c}} \\ \mathbf{F}_\lambda^{\mathrm{c}} \end{Bmatrix} = \mathbf{0}, \tag{5.2}$$

which represents the set of algebraic nonlinear equations arising from the application of the augmented Lagrangian method to solve contact problems. Note that even if the residual vector of the structural part is defined by a set of linear algebraic equations, the nonlinearity related with the contact problem is maintained.

Based on the augmented Lagrangian approach, proposed by Alart and Curnier [Alart 91], the equilibrium of the discretized deformable body in frictional contact with an obstacle is governed by the following system of non-differentiable equations:

$$\begin{cases} \mathbf{F}^{\text{int}}(\mathbf{u}) - \mathbf{F}^{\text{ext}} + \mathbf{F}^{\text{c}}(\mathbf{u},\boldsymbol{\lambda}) = \mathbf{0} \\ -\dfrac{1}{\varepsilon}(\boldsymbol{\lambda} - \mathbf{F}^{\text{c}}(\mathbf{u},\boldsymbol{\lambda})) = \mathbf{0} \end{cases}, \tag{5.3}$$

where $\mathbf{F}^{\text{int}}(\mathbf{u})$ and $\mathbf{F}^{\text{ext}}$ denote the internal and the external force vectors, respectively. Besides, the external forces are assumed to be independent of the displacement field. The discretized frictional contact operator is denoted by $\mathbf{F}^{\text{c}}(\mathbf{u},\boldsymbol{\lambda})$ for a single contact element, which is defined generically for all contact status (gap, stick and slip) as follows:

$$\mathbf{F}^{\text{c}}(\mathbf{u},\boldsymbol{\lambda}) = \text{proj}_{\Re^-}(\hat{\lambda}_{\text{n}})\mathbf{n} + \text{proj}_{C^{\text{aug}}(\hat{\lambda}_{\text{n}})}(\hat{\boldsymbol{\lambda}}_{\text{t}}), \tag{5.4}$$

where $\hat{\lambda}_{\text{n}}$ and $\hat{\boldsymbol{\lambda}}_{\text{t}}$ are the augmented Lagrange multipliers defined in (2.82) and (2.84), respectively, while $\text{proj}_{\Re^-}(x)$ denotes the projection of x on $\Re^-$, i.e. the closest point to x belonging to $\Re^-$. In the second term of (5.4) the standard convex $C(\lambda_{\text{n}})$ is replaced by the augmented one, defined by:

$$C^{\text{aug}}(\hat{\lambda}_{\text{n}}) = C(\text{proj}_{\Re^-}(\hat{\lambda}_{\text{n}})), \tag{5.5}$$

which consists in the prolongation the multivalued convex set (friction cone) by the positive half-line, as illustrated in Figure 2.12 (b). For positive values of the normal component of the augmented Lagrange multiplier, the radius of the convex disk is equal to zero and the Coulomb's cone shrinks to the vertical half-line. The augmented projection is crucial to ensure the continuity of the frictional contact operator [Alart 91].

The solution of the system of nonlinear equations (5.3) is equivalent to the saddle point of the augmented Lagrangian (2.80), which is calculated as a variational problem through expression (2.90). Indeed, the system of equations (5.3) involves primal and dual variables, yielding a mixed formulation in which both the kinematic variables (nodal displacements) and the static variables (frictional contact forces) are the final unknowns of the problem. In order to highlight the contribution of the frictional contact to the global system of equations (5.3), it is useful separate the contact part from the structural part [Alart 91]. Thus, the mixed contact operator is defined from (5.3) by assembling the two sub-operators as follows:

$$\mathbf{F}^{*}(\mathbf{u},\boldsymbol{\lambda}) = \begin{Bmatrix} \mathbf{F}^{\text{c}}(\mathbf{u},\boldsymbol{\lambda}) \\ -\dfrac{1}{\varepsilon}(\boldsymbol{\lambda} - \mathbf{F}^{\text{c}}(\mathbf{u},\boldsymbol{\lambda})) \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda}) \\ \mathbf{F}^{\text{suppl}}(\mathbf{u},\boldsymbol{\lambda}) \end{Bmatrix}, \tag{5.6}$$

where $\mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda})$ denotes the constraints imposed by the obstacle in the equilibrium equation and $\mathbf{F}^{\text{suppl}}(\mathbf{u},\boldsymbol{\lambda})$ expresses the supplementary equations required to evaluate the frictional contact forces [Heege 96]. From a mathematical point of view, expression (5.3) defines a system of equations with six unknowns (3D) for each contact element, three

components of nodal displacements and three components of contact force defined in the artificial node (see Figure 5.1).

Depending on the contact status (gap, stick and slip), the contribution of the contact element to the virtual work of the system is given by different integrals, as illustrated in (2.90). Thus, the derivation of the mixed contact operator $\mathbf{F}^{*}(\mathbf{u},\boldsymbol{\lambda})$ for each contact status can be obtained from (5.6), using the frictional contact operator presented in (5.4) derived for each status. In the framework of the augmented Lagrangian method, the contact status is determined by the linear combination of kinematic and static variables (primal and dual variables) [Alart 91].

The non-contact between the slave node and the rigid obstacle (gap status) is defined by the normal component of the augmented Lagrange multiplier $\hat{\lambda}_{\mathrm{n}} > 0$. Thus, the mixed contact operator for the gap status is defined through (5.4) derived for the same contact status, given by:

$$\hat{\lambda}_{\mathrm{n}} > 0 \quad \Rightarrow \quad \mathbf{F}^{*}_{\mathrm{gap}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{Bmatrix} \mathbf{0} \\ -\dfrac{1}{\varepsilon}\boldsymbol{\lambda} \end{Bmatrix}, \tag{5.7}$$

which imposes that the contact force (Lagrange multiplier) connected with the slave node is zero, as expected since there is no contact. Besides, the residual vector related with the structural part is not affected by the contact element, consequently the equilibrium of the deformable body is reached when the internal forces are equal to the external forces (see (5.3)).

The stick contact status is defined through the condition $\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| \leq -\mu\hat{\lambda}_{\mathrm{n}}$. Therefore, from the definition of (5.4) for the stick status, the corresponding mixed contact operator for such contact status is given by:

$$\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| \leq -\mu\hat{\lambda}_{\mathrm{n}} \quad \Rightarrow \quad \mathbf{F}^{*}_{\mathrm{stick}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{Bmatrix} \hat{\lambda}_{\mathrm{n}}\mathbf{n} + \hat{\boldsymbol{\lambda}}_{\mathrm{t}} \\ g_{\mathrm{n}}\mathbf{n} + \mathbf{g}_{\mathrm{t}} \end{Bmatrix}, \tag{5.8}$$

which imposes that the relative displacement between the slave node and the rigid surface is zero, i.e. when the node is already in contact, its displacement is equal to the displacement of the obstacle. Analogously, the slip status of the slave node is defined by the condition $\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| > -\mu\hat{\lambda}_{\mathrm{n}}$ and the corresponding mixed contact operator derived from (5.4) is defined by:

$$\left\|\hat{\boldsymbol{\lambda}}_{\mathrm{t}}\right\| > -\mu\hat{p}_{\mathrm{n}} \quad \Rightarrow \quad \mathbf{F}^{*}_{\mathrm{slip}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{Bmatrix} \hat{\lambda}_{\mathrm{n}}(\mathbf{n} - \mu\mathbf{t}) \\ g_{\mathrm{n}}\mathbf{n} - \dfrac{1}{\varepsilon}(\boldsymbol{\lambda}_{\mathrm{t}} + \mu\hat{\lambda}_{\mathrm{n}}\mathbf{t}) \end{Bmatrix}, \tag{5.9}$$

where the tangential slip direction unit vector is defined by:

$$\mathbf{t} = \hat{\boldsymbol{\lambda}}_{\mathrm{t}} \Big/ \left\| \hat{\boldsymbol{\lambda}}_{\mathrm{t}} \right\|. \tag{5.10}$$

In fact, the expressions for the mixed contact operators according with the contact status of the slave node, presented in (5.7)–(5.9), are analogous to the contribution of the frictional contact interface to the virtual work of the system presented in (2.90). The difference between them arises in the augmented contact pressure $\hat{p}_{\mathrm{n}} = p_{\mathrm{n}} + \varepsilon_{\mathrm{n}} g_{\mathrm{n}}$, previously assumed as known, which is replaced by the augmented Lagrange multiplier $\hat{\lambda}_{\mathrm{n}}$ (unknown variable). Moreover, the difference between frictionless (only one supplementary degree of freedom) and frictional contact elements is highlighted in the augmented Lagrangian virtual work principle (2.90) through the split of normal and tangential components. Indeed, both kinematic and static contact variables are expressed with respect to a local frame defined on the surface of the rigid obstacle.

5.1.1. Contact linearization

The solution of the mixed system of nonlinear equations (5.3) can be achieved by two different solution schemes: a simultaneous updated of static and kinematic variables [Alart 91], [Heegaard 93] and a nested Uzawa-type iteration scheme [Simo 92]. In fact, the algorithms based in the Uzawa's scheme are stable but very slow, because they are based on an alternate treatment of the primal and dual variables [Yastrebov 13]. Therefore, in the present study, the resulting system of nonlinear and partially non-differentiable equations defined in (5.3) is solved simultaneously for both variables, using the generalized Newton method, previously presented in Section 3.2. The convergence of the generalized Newton method used to solve the resulting frictional contact augmented Lagrangian saddle point problem is discussed by Alart [Alart 97] and more recently by Renard [Renard 13] considering elasticity.

Firstly, the system of nonlinear equations, presented (5.3), arising from the application of the augmented Lagrangian method is splitted in two parts: (i) a differentiable structural part (virtual work of the deformable body) and (ii) a non-differentiable contact part (virtual work due to contact):

$$\mathbf{F}^{\mathrm{s}}(\mathbf{u}) + \mathbf{F}^{*}(\mathbf{u}, \boldsymbol{\lambda}) = \mathbf{0}. \tag{5.11}$$

Note that the contribution of the frictional contact to the global system is given by the mixed contact operator (5.6), which contains all non-differentiable part of the equations. Also, although the Lagrange multiplier (dual variable) contained in (5.6) is differentiable, it is also included in the non-differentiable part. On the other hand, the differentiable structural part is expressed by:

$$\mathbf{F}^{\mathrm{s}}(\mathbf{u}) = \begin{Bmatrix} \mathbf{F}^{\mathrm{int}}(\mathbf{u}) - \mathbf{F}^{\mathrm{ext}} \\ \mathbf{0} \end{Bmatrix}, \tag{5.12}$$

which contains the internal force vector defined in terms of primal variables (nodal displacements) and the external discretized forces (dead loads), which are independent of the solution. Since the system of equations (5.11) comprises both primal and dual variables, the pair $\mathbf{x} = (\mathbf{u}, \boldsymbol{\lambda})^{\mathrm{T}}$ is introduced. Therefore, the application of the generalized Newton method to the system of nonlinear and partially non-differentiable equations (5.11) presents the form:

$$\begin{cases} (\mathbf{K}^{\mathrm{s}}(\mathbf{x}_i) + \mathbf{J}^{*}(\mathbf{x}_i))\Delta\mathbf{x}_i = -(\mathbf{F}^{\mathrm{s}}(\mathbf{x}_i) + \mathbf{F}^{*}(\mathbf{x}_i)) \\ \mathbf{x}_{i+1} = \mathbf{x}_i + \Delta\mathbf{x}_i \end{cases}, \tag{5.13}$$

where $\mathbf{K}^{\mathrm{s}}(\mathbf{x}_i)$ denotes the structural tangent matrix and $\mathbf{J}^{*}(\mathbf{x}_i)$ denotes the generalized Jacobian of the mixed contact operator (5.6) at $\mathbf{x}_i$. Since the mixed contact operator is not differentiable everywhere, the Jacobian matrix is evaluated independently for each slave node as function of its contact status (gap, stick and slip), which are the regions where the operator is differentiable. Thus, for a given contact status, the generalized Jacobian is reduced to the classical Jacobian matrix.

The Newton method presented in (5.13) is used to solve the nonlinear contact problems incrementally. In order to attain quadratic convergence in the iterative solution scheme, a consistent linearization of the system of equations (5.11) is required. The closed form expressions for the tangent matrix derived from the structural finite elements is not given in the present study (see e.g. [Simo 85a] for rate-independent elastoplasticity). In fact, this dissertation is particularly focused in the tangent matrix resulting from the contact elements, which can be called tangent contact matrix. Therefore, in analogy to the tangent matrix of a structural finite element, the elemental contact Jacobian matrix is defined through the first-order partial derivatives of the mixed contact operator vector (5.6), expressed by:

$$\mathbf{J}^{*}(\mathbf{u}, \boldsymbol{\lambda}) = \begin{bmatrix} \nabla_{\mathbf{u}}\mathbf{F}^{\mathrm{equi}}(\mathbf{u}, \boldsymbol{\lambda}) & \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\mathrm{equi}}(\mathbf{u}, \boldsymbol{\lambda}) \\ \nabla_{\mathbf{u}}\mathbf{F}^{\mathrm{suppl}}(\mathbf{u}, \boldsymbol{\lambda}) & \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\mathrm{suppl}}(\mathbf{u}, \boldsymbol{\lambda}) \end{bmatrix}. \tag{5.14}$$

The definition of this matrix changes according with the contact status of the slave node belonging to the contact element. Thus, it is useful to split the tangent matrix into derivatives with respect to primal and dual variables, as shown in (5.14).

Taking into account (5.7), the elemental Jacobian matrix for the slave node with gap contact status is given by:

$$\mathbf{J}^{*}_{\text{gap}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -(1/\varepsilon)\mathbf{I} \end{bmatrix}, \tag{5.15}$$

where $\mathbf{I}$ is the second order identity tensor. The elemental Jacobian matrix proposed by Alart [Alart 91] for a slave node with stick status is defined as follows:

$$\mathbf{J}^{*}_{\text{stick}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{bmatrix} \varepsilon\mathbf{I} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}, \tag{5.16}$$

which is obtained from (5.8) assuming that the local frame defined on the master surface (see Figure 2.6) is fixed in all Newton iterations within an increment. Nevertheless, this assumption is not valid for strongly curved obstacles, particularly for large slip increments, because it is independent of the surface curvature. Therefore, Heege and Alart [Heege 96] developed an elemental Jacobian matrix which takes into account the contributions due to the curvature of the rigid obstacle, using the gradient of the unit normal vector at the contact point. However, according with what was previously mentioned about the contact operator (5.8), for stick contact status the nodal displacement of the slave node is known *a priori* (displacement of obstacle). Thus, it is possible to consider that the solution is path-independent and, consequently, adopt the Jacobian matrix defined in (5.16) [Alart 95a]. This was the assumption adopted in this work since the Jacobian matrix (5.16) is significantly simpler and the convergence rate of the Newton method is not affected [Oliveira 08]. Finally, the elemental Jacobian matrix for the node with slip status, which is more complex than the last two ones, is defined by Alart [Alart 91] from (5.9), assuming a fixed local frame attached to the obstacle:

$$\mathbf{J}^{*}_{\text{slip}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{bmatrix} \varepsilon\mathbf{M} & \mathbf{M} \\ \mathbf{M} & \dfrac{1}{\varepsilon}(\mathbf{M}-\mathbf{I}) \end{bmatrix}, \tag{5.17}$$

where the second-order tensor $\mathbf{M}$ is defined as:

$$\mathbf{M} = (\mathbf{n} - \mu\mathbf{t}) \otimes \mathbf{n} + \rho(\mathbf{I} - \mathbf{n} \otimes \mathbf{n} - \mathbf{t} \otimes \mathbf{t}), \tag{5.18}$$

which is independent of the rigid obstacle curvature. Its definition includes the unit slip vector defined in (5.10), as well as the unit normal vector of the master surface at the closest point. Besides, ρ is a projection scaling factor defined by:

$$\rho = -\mu\hat{\lambda}_{\text{n}} \Big/ \left\|\hat{\boldsymbol{\lambda}}_{\text{t}}\right\|, \tag{5.19}$$

which is always in the range between zero and one. Although Jacobian matrix (5.17) considers a fixed local frame, it is updated in each iteration of the Newton's method with the new projection contact point, providing a new slip direction. It should be noted that, in

opposite to the other contact status, the elemental Jacobian matrix defined for the slip status by (5.17) is not symmetric because $\mathbf{M}$ is not symmetric, excluding for the frictionless contact case $(\mu = 0)$. The Jacobian matrix developed by [Heege 96] for slip contact status takes into account the contributions of geometric terms arising from the gradient of the normal vector. Since the local frame defined by the surface normal vector varies within the iterative solution of the equilibrium equations (particularly for strongly curved rigid obstacles), the Jacobian matrix obtained from the contact operator (5.9) is expressed by:

$$\mathbf{J}^{*}_{\text{slip}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{bmatrix} \varepsilon\mathbf{E} + \mathbf{N}\nabla_{\mathbf{u}}\mathbf{n} & \mathbf{M} \\ \mathbf{E} + \dfrac{1}{\varepsilon}\mathbf{N}\nabla_{\mathbf{u}}\mathbf{n} & \dfrac{1}{\varepsilon}(\mathbf{M}-\mathbf{I}) \end{bmatrix}, \tag{5.20}$$

where the second-order tensor $\mathbf{E}$ is defined as:

$$\mathbf{E} = (\mathbf{n} - \mu\mathbf{t}) \otimes \nabla_{\mathbf{u}} g_{\text{n}} + \rho(\mathbf{I} - \mathbf{n}\otimes\mathbf{n} - \mathbf{t}\otimes\mathbf{t}), \tag{5.21}$$

which is independent of the obstacle curvature. The supplementary curvature terms can be easily identified in (5.20), since they are coupled to the gradient of the normal vector:

$$\mathbf{N} = (\mathbf{n} - \mu\mathbf{t}) \otimes \boldsymbol{\lambda} + \hat{\lambda}_{\text{n}}\mathbf{I} - \rho\left\{\mathbf{n}\otimes(\boldsymbol{\lambda} + \varepsilon\mathbf{g}) - ((\boldsymbol{\lambda} + \varepsilon\mathbf{g})\cdot\mathbf{n})(\mathbf{t}\otimes\mathbf{t} - \mathbf{I})\right\}. \tag{5.22}$$

The curvature terms included in (5.20) vanish for flat obstacles because the normal vector to the obstacle is the same everywhere (the second terms in the first column disappears). Besides, the gradient of the normal gap function involved in the definition of (5.21) results into the unit normal vector for flat surfaces, yielding the second-order tensor shown in (5.18), as proposed by [Alart 91]. Note that the derivative of the mixed contact operator (5.9) with respect to the Lagrange multiplier is not modified by the contact surface curvature (compare (5.17) with (5.20)). The numerical examples presented in Chapter 6 are performed with the Jacobian matrix (5.20) for the slip contact status.

In order to take into account large slips on curved contact surfaces, the gradients involved in the formulation of the Jacobian matrix (5.20) need to be evaluated with respect to the global coordinate system [Heege 96]. The gradient of the normal vector with respect to the local coordinates was previously expressed in (3.31). Thus, the subsequent transformation to the global Cartesian frame can be written as follows:

$$\nabla_{\mathbf{u}}\mathbf{n} \equiv \frac{\partial\mathbf{n}(\xi^1,\xi^2)}{\partial\mathbf{u}} = \frac{\partial\mathbf{n}}{\partial\xi^{\alpha}}\frac{\partial\xi^{\alpha}}{\partial\mathbf{u}} = -b_{\alpha\beta}(\xi^1,\xi^2)m^{\beta\gamma}(\xi^1,\xi^2)\boldsymbol{\tau}^{\text{m}}_{\gamma}(\xi^1,\xi^2)\frac{\partial\xi^{\alpha}}{\partial\mathbf{u}}, \quad \alpha,\beta,\gamma = 1,2, \tag{5.23}$$

where the gradient relating the curvilinear and the Cartesian system is evaluated introducing a third convective coordinate aligned with the surface normal vector, defining the local surface coordinate system [Heege 96], [Konyukhov 05]. This transformation can be defined for every point belonging to the parametric surface, which is obtained from the

normal projection method previously presented in (3.28). In fact, at the convergence of the numerical scheme (3.29), the first two lines of the inverse of the Jacobian matrix presented in (3.30) define the transformation matrix from the curvilinear system to the Cartesian system. Moreover, the third line defines the gradient of the normal gap function required for (5.21).

5.1.2. Mixed system of equations

The implicit formulation based on the coupled augmented Lagrangian method for *quasi*-static frictional contact problems leads to a nonlinear system of equations, which is solved iteratively by means of the generalized Newton method. Since the large deformation frictional contact problem involves three sources of nonlinearities (geometrical, material and contact), the external load/prescribed displacement is firstly divided into a set of increments. In fact, the rate-independent plasticity requires the numerical integration of the rate constitutive equations over a discrete sequence of time steps (increments). Besides, the friction force is path-dependent, thus requiring an incremental update procedure. Thus, the Newton method is used to determine each new equilibrium configuration of the body from the previous one. In each increment, the initial solution for the iterative scheme is calculated with an explicit algorithm (predictor step), which is successively corrected using an implicit method (corrector step), until achieving the static equilibrium of the deformable body (within a prescribed tolerance). This procedure is currently implemented in DD3IMP finite element code and it was schematically shown in Figure 3.1.

The adoption of the augmented Lagrangian method with simultaneous update of kinematic and static variables (contact forces) yields a mixed formulation where the primal and dual variables are the final unknowns of the problem, as shown in (5.13). The explicit form of the system of linear equations according to the Newton's method in each iteration is given by:

$$\left(\begin{bmatrix}\mathbf{K}^{\mathrm{s}}(\mathbf{u}_i) & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}+\begin{bmatrix}\nabla_{\mathbf{u}}\mathbf{F}^{\mathrm{equi}}(\mathbf{u}_i,\boldsymbol{\lambda}_i) & \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\mathrm{equi}}(\mathbf{u}_i,\boldsymbol{\lambda}_i)\\ \nabla_{\mathbf{u}}\mathbf{F}^{\mathrm{suppl}}(\mathbf{u}_i,\boldsymbol{\lambda}_i) & \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\mathrm{suppl}}(\mathbf{u}_i,\boldsymbol{\lambda}_i)\end{bmatrix}\right)\begin{Bmatrix}d\mathbf{u}_i\\ d\boldsymbol{\lambda}_i\end{Bmatrix}=\begin{Bmatrix}d\mathbf{f}^{\mathrm{s}}(\mathbf{u}_i)-\mathbf{F}^{\mathrm{equi}}(\mathbf{u}_i,\boldsymbol{\lambda}_i)\\ -\mathbf{F}^{\mathrm{suppl}}(\mathbf{u}_i,\boldsymbol{\lambda}_i)\end{Bmatrix}, \quad (5.24)$$

where the global tangent matrix is the assembly of the structural tangent matrix with the tangent contact matrix (5.14). Moreover, the global tangent matrix and the residual nodal force vector is given by the assembly of the elemental tangent matrices and the residual nodal force vectors. However, the system of linear equations (5.24) is slightly different in the explicit and the implicit steps of the time integration algorithm (see Figure 3.1). In case of the explicit approach (predictor step), $\mathbf{K}^{\mathrm{s}}$ denotes the global tangent stiffness matrix and $d\mathbf{f}^{\mathrm{s}}$ correspond to the incremental external force vector, while the resulting solution vector $d\mathbf{u}/d\boldsymbol{\lambda}$ expresses the displacement/incremental contact force for the current

increment. On the other hand, for the implicit stage (corrector step), $\mathbf{K}^{s}$ denotes the global consistent matrix and $d\mathbf{f}^{s}$ denotes the non-equilibrated force vector. In this case, the solution of the linear system of equations (5.24) expresses the corrections to the displacements/incremental contact forces evaluated in the predictor step. Concerning the contribution of the contact elements, the Jacobian matrix associated with the contact elements depends directly on the contact status of the slave node (5.15)–(5.17) and (5.20), but is the same for both predictor and corrector steps. Besides, the right-hand side vector of the mixed system includes the mixed contact operator that depends on the contact status of the node (5.7)–(5.9). Each one of the mixed linear system of equations presented in (5.24), resulting from the Newton method application is solved with the direct solver from the Intel MKL, described in Section 3.2.1.1.

The nonlinear system of equations presented in (5.3) involves both nodal displacements and nodal contact forces for the slave nodes, which are connected through the contact elements. Considering a generic slave node of the deformable body in contact with a rigid obstacle, the shortened form of (5.24) for a single Newton iteration (the indices related with the iteration number will be omitted) can be written as:

$$\begin{cases} \mathbf{K}^{s}\{d\mathbf{u}\} + \nabla_{\mathbf{u}}\mathbf{F}^{equi}\{d\mathbf{u}^{c}\} + \nabla_{\lambda}\mathbf{F}^{equi}\{d\boldsymbol{\lambda}^{c}\} = \{d\mathbf{f}^{s}\} - \mathbf{F}^{equi} \\ \nabla_{\mathbf{u}}\mathbf{F}^{suppl}\{d\mathbf{u}^{c}\} + \nabla_{\lambda}\mathbf{F}^{suppl}\{d\boldsymbol{\lambda}^{c}\} = -\mathbf{F}^{suppl} \end{cases}, \tag{5.25}$$

where $\{d\mathbf{u}^{c}\}$ denotes the nodal displacement of the slave node and $\{d\boldsymbol{\lambda}^{c}\}$ denotes the increment of nodal contact force vector in the artificial node (contact element).

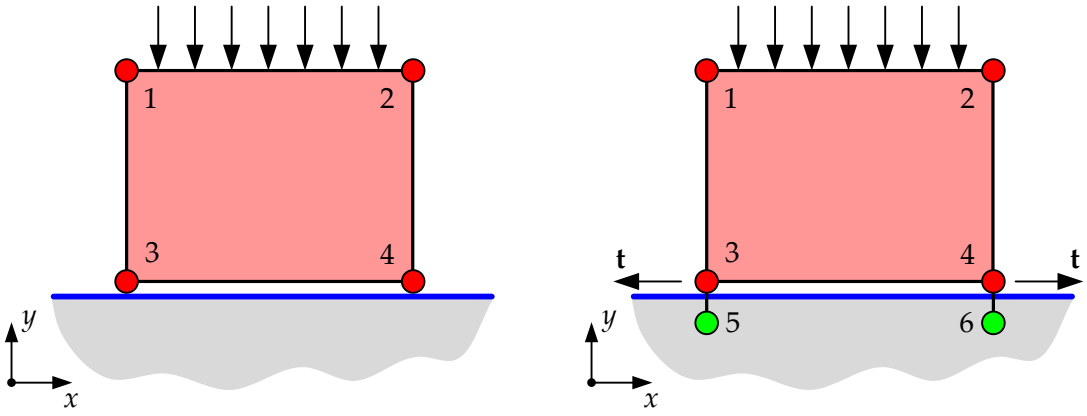


Figure 5.2. Example of a discretized deformable body in contact with a rigid flat obstacle: (a) absence of contact elements; (b) with contact elements.

In order to illustrate the pattern of the global tangent matrix due to the incorporation of the contact elements into the global system of equations, a simple 2D example is presented [Alart 95b]. The mechanical system involves a deformable body discretized by a

single structural finite element in frictional contact with a rigid flat obstacle, as shown in Figure 5.2. A constant pressure is imposed on the upper segment between nodes 1 and 2. In the absence of contact between the deformable body and the rigid obstacle, the global tangent matrix and the residual vector for such configuration have the pattern presented in Figure 5.3 (a). Indeed, only the deformable body contributed for the global system of equations through the assembly of the elemental matrices and nodal force vectors. The frictional contact with the rigid surface is treated by introducing two contact elements, as illustrated in Figure 5.2 (b). They are represented in a symbolic manner by a line connecting the slave node belonging to the deformable body with the artificial node (arbitrary position) that contains the contact force components as its degrees of freedom. The contact is incorporated into the global system of equations via the change of the tangent matrix and the residual vector, as shown in Figure 5.3 (b). In fact, the adoption of the coupled augmented Lagrangian method to deal with the contact constraints introduces additional degrees of freedom (see node 5 and 6 in Figure 5.2 (b)), leading to a mixed system of equations. The white squares in the representation of the global tangent matrix denotes the position of the zero entries, while the coloured squares denotes (possible) nonzero matrix entries. The red colour identifies the structural elements and the green colour is used for the contact elements, while the squares with both colours contain the contribution of both components. Each contact element is represented in the global tangent matrix by three squares and one triangle (Figure 5.3 (b)), which denote the four elements of elemental Jacobian matrix (5.14) included in the system of linear equations (5.25). The connection between the slave nodes (3 and 4) with the artificial nodes (5 and 6) is highlighted in the pattern of the tangent matrix. The global residual vector includes the mixed contact operator (5.6) in the nodes that define the contact element, as shown in Figure 5.3 (b).

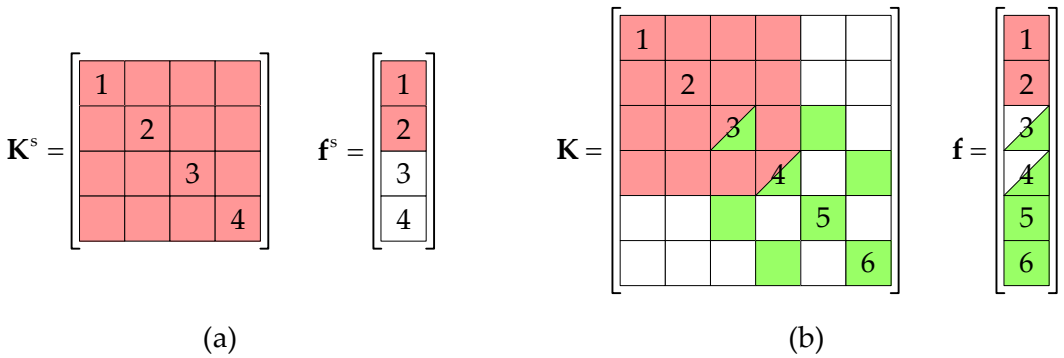


Figure 5.3. Pattern of the global tangent matrix and residual vector for a single structural finite element: (a) absence of contact elements; (b) contribution of two contact elements.

The global tangent matrix arising from the structural finite elements is always symmetrically structured, i.e. the pattern of nonzero entries is symmetric (see Figure 3.5 (b)). Indeed, the introduction of the contact elements into the global tangent matrix does

not changes the nonzero symmetric structure of the matrix, as shown in Figure 5.3. Therefore, the mixed system of linear equations (5.25) is solved using the Intel MKL library, considering the symmetrically structured square system matrix.

Since the example under analysis is 2D, each node of the deformable body (including the artificial ones) is represented by two degrees of freedom (horizontal and vertical components), leading to a global tangent matrix with 12 rows/columns. As aforementioned, the Jacobian matrix of the mixed contact operator (5.6) is evaluated according with the contact status of the slave node. Therefore, in order to describe explicitly the global tangent matrix of the present example, it is assumed backward slip for the node 3 and forward slip for the node 4 (see Figure 5.2 (b)). The Jacobian matrix for the slip status is given in (5.17) for flat obstacles, which should be evaluated for each slave node and, posteriorly, incorporated in (5.25). The slip direction unit vector (5.10) required to define the second-order tensor $\mathbf{M}$ is given in Figure 5.2 (b) for each slave node. Since the terms $\rho(\mathbf{I}-\mathbf{n}\otimes\mathbf{n}-\mathbf{t}\otimes\mathbf{t})$ arising in the second-order tensor (5.18) vanish for 2D examples [Alart 91], the tensors needed to determine the Jacobian matrices are expressed by:

$$\mathbf{M}=\begin{bmatrix}0 & \mu\\ 0 & 1\end{bmatrix} \quad \text{and} \quad \frac{1}{\varepsilon}(\mathbf{M}-\mathbf{I})=\begin{bmatrix}-1/\varepsilon & \mu/\varepsilon\\ 0 & 0\end{bmatrix}, \tag{5.26}$$

$$\mathbf{M}=\begin{bmatrix}0 & -\mu\\ 0 & 1\end{bmatrix} \quad \text{and} \quad \frac{1}{\varepsilon}(\mathbf{M}-\mathbf{I})=\begin{bmatrix}-1/\varepsilon & -\mu/\varepsilon\\ 0 & 0\end{bmatrix}, \tag{5.27}$$

for the node 3 and 4, respectively. The difference between the nodes 3 and 4 arises from the slip direction defined for each node, shown in Figure 5.2 (b). The global tangent matrix is defined by the assembly of the structural tangent matrix with the Jacobian matrix of each contact element. Thus, for the example comprising a discretized deformable body coming in contact with a rigid surface, presented in Figure 5.2, the global tangent matrix is given by:

$$\mathbf{K}=\begin{bmatrix}
K_{11}^{xx} & K_{11}^{xy} & K_{12}^{xx} & K_{12}^{xy} & K_{13}^{xx} & K_{13}^{xy} & K_{14}^{xx} & K_{14}^{xy} & 0 & 0 & 0 & 0 \\
K_{11}^{yx} & K_{11}^{yy} & K_{12}^{yx} & K_{12}^{yy} & K_{13}^{yx} & K_{13}^{yy} & K_{14}^{yx} & K_{14}^{yy} & 0 & 0 & 0 & 0 \\
K_{21}^{xx} & K_{21}^{xy} & K_{22}^{xx} & K_{22}^{xy} & K_{23}^{xx} & K_{23}^{xy} & K_{24}^{xx} & K_{24}^{xy} & 0 & 0 & 0 & 0 \\
K_{21}^{yx} & K_{21}^{yy} & K_{22}^{yx} & K_{22}^{yy} & K_{23}^{yx} & K_{23}^{yy} & K_{24}^{yx} & K_{24}^{yy} & 0 & 0 & 0 & 0 \\
K_{31}^{xx} & K_{31}^{xy} & K_{32}^{xx} & K_{32}^{xy} & K_{33}^{xx} & K_{33}^{xy}+\varepsilon\mu & K_{34}^{xx} & K_{34}^{xy} & 0 & \mu & 0 & 0 \\
K_{31}^{yx} & K_{31}^{yy} & K_{32}^{yx} & K_{32}^{yy} & K_{33}^{yx} & K_{33}^{yy}+\varepsilon & K_{34}^{yx} & K_{34}^{yy} & 0 & 1 & 0 & 0 \\
K_{41}^{xx} & K_{41}^{xy} & K_{42}^{xx} & K_{42}^{xy} & K_{43}^{xx} & K_{43}^{xy} & K_{44}^{xx} & K_{44}^{xy}-\varepsilon\mu & 0 & 0 & 0 & -\mu \\
K_{41}^{yx} & K_{41}^{yy} & K_{42}^{yx} & K_{42}^{yy} & K_{43}^{yx} & K_{43}^{yy} & K_{44}^{yx} & K_{44}^{yy}+\varepsilon & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & \mu & 0 & 0 & -1/\varepsilon & \mu/\varepsilon & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\mu & 0 & 0 & -1/\varepsilon & -\mu/\varepsilon \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\end{bmatrix}, \quad (5.28)$$

which presents the pattern illustrated in Figure 5.3 (b). The non-symmetry of the global tangent matrix (5.28) results from the non-associated character of the Coulomb friction law. However, for the specific case of frictionless contact, the tangent matrix is symmetric in terms of contact contribution, as can be observed through (5.28). Although some entries of the Jacobian contact matrices are zero, all coloured entries of the global tangent matrix are defined as nonzero in order to define a nonzero symmetric structure for the global tangent matrix.

5.1.3. Reduced system of equations

The solution of frictional contact problems with the augmented Lagrangian method leads to a nonlinear system of equations, involving both primal and dual unknowns, which are simultaneously updated using the generalized Newton method, as shown in (5.24). Besides, the Jacobian matrix associated with each contact element is dependent of the contact status (gap, stick and slip) of the slave node. However, according with the relationship between the two sub-operators defined in (5.6), expressed by:

$$\nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}}=\frac{1}{\varepsilon}\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}, \quad (5.29)$$

the linear system of equations (5.25) can be written in a simplified general form, valid for all contact status [Menezes 94]:

$$\begin{cases}\mathbf{K}^{s}\{d\mathbf{u}\}+\mathbf{A}\{d\mathbf{u}^{c}\}+\mathbf{C}\{d\boldsymbol{\lambda}^{c}\}=\{\boldsymbol{\mathcal{F}}^{1}\} \\ \mathbf{A}\{d\mathbf{u}^{c}\}+\mathbf{B}\{d\boldsymbol{\lambda}^{c}\}=\{\boldsymbol{\mathcal{F}}^{2}\}\end{cases}, \quad (5.30)$$

where the new tensors need to be evaluated for each contact status.

If the slave node presents the gap status, from the Jacobian matrix defined in (5.15) and the mixed contact operator (5.7), the system of equations (5.30) is defined by:

$$\mathbf{A} = \mathbf{C} = \mathbf{0} \quad \text{and} \quad \mathbf{B} = -\mathbf{I}, \tag{5.31}$$

$$\{\mathcal{F}^1\} = \{d\mathbf{f}^s\} \quad \text{and} \quad \{\mathcal{F}^2\} = \{\boldsymbol{\lambda}\}. \tag{5.32}$$

For stick contact status, the Jacobian matrix (5.16) and the contact operator (5.8) are used to express the system of linear equations (5.30), defined with:

$$\mathbf{A} = \varepsilon\mathbf{I}, \quad \mathbf{B} = \mathbf{0} \quad \text{and} \quad \mathbf{C} = \mathbf{I}, \tag{5.33}$$

$$\{\mathcal{F}^1\} = \{d\mathbf{f}^s\} - \{\hat{\boldsymbol{\lambda}}\} \quad \text{and} \quad \{\mathcal{F}^2\} = -\{\varepsilon\mathbf{g}\}. \tag{5.34}$$

Finally, for the slip contact status, considering the Jacobian matrix presented in (5.17) and the frictional contact operator defined in (5.9), the system of linear equations is defined by:

$$\mathbf{A} = \varepsilon\mathbf{M}, \quad \mathbf{B} = \mathbf{M} - \mathbf{I} \quad \text{and} \quad \mathbf{C} = \mathbf{M}, \tag{5.35}$$

$$\{\mathcal{F}^1\} = \{d\mathbf{f}^s\} - \{\hat{\lambda}_n(\mathbf{n} - \mu\mathbf{t})\} \quad \text{and} \quad \{\mathcal{F}^2\} = \{\boldsymbol{\lambda} - \hat{\lambda}_n(\mathbf{n} - \mu\mathbf{t})\}, \tag{5.36}$$

where the second-order tensor $\mathbf{M}$ is defined in (5.18). Note that the same procedure can be performed for the Jacobian matrix (5.20), which takes into account the obstacle curvature.

From the analysis of the mixed system of equations (5.30), it is possible to observe that the contact force variable is only dependent on the displacement of the slave node (contact element), i.e. the structural tangent matrix in not included in the second line of the system. Therefore, the unknown contact forces can be eliminated from the system, yielding a reduced system of equation with only nodal displacements as unknowns. Afterwards, the contact forces in the slave nodes can be evaluated from the nodal displacements. In fact, according with the definition of the matrices (5.31), (5.33) and (5.35) for the contact status gap, stick and slip, respectively, the relationship $\mathbf{B} = \mathbf{C} - \mathbf{I}$ is valid for all contact statuses. This means that the subtraction of the second line of (5.30) to the first one yields an expression to evaluate directly the nodal contact forces (Lagrange multipliers in the artificial nodes), given by:

$$\{d\boldsymbol{\lambda}^c\} = \{\mathcal{F}^1\} - \{\mathcal{F}^2\} - \mathbf{K}^s\{d\mathbf{u}\}. \tag{5.37}$$

The incorporation of (5.37) in the second line of (5.30) allows defining a linear system of equations with only nodal displacements as unknowns, defined by:

$$\mathbf{A}\{d\mathbf{u}^c\} + \mathbf{B}(\{\mathcal{F}^1\} - \{\mathcal{F}^2\}) - \mathbf{B}\mathbf{K}^s\{d\mathbf{u}\} = \{\mathcal{F}^2\}. \tag{5.38}$$

Therefore, the reduced system of equations is defined by (5.37) and (5.38) and can be written as:

$$\begin{cases} -\mathbf{B}\mathbf{K}^{s}\{d\mathbf{u}\} + \mathbf{A}\{d\mathbf{u}^{c}\} = \{\mathcal{F}^{*}\} \\ \{d\boldsymbol{\lambda}^{c}\} = \{\mathcal{F}^{**}\} - \mathbf{K}^{s}\{d\mathbf{u}\} \end{cases}, \tag{5.39}$$

where the vectors required for the right hand side of the global system are given by:

$$\{\mathcal{F}^{*}\} = \{\mathcal{F}^{2}\} - \mathbf{B}(\{\mathcal{F}^{1}\} - \{\mathcal{F}^{2}\}) \quad \text{and} \quad \{\mathcal{F}^{**}\} = \{\mathcal{F}^{1}\} - \{\mathcal{F}^{2}\}. \tag{5.40}$$

The first line of (5.39) denotes the reduced system of equations, which allows to calculate the nodal displacements using a global tangent matrix involving only primal variables. The contact forces in the artificial nodes (see Figure 5.1) are evaluated using the second line of (5.39), which make use of the nodal displacements previously calculated. The pattern of the tangent matrices obtained for the reduced system of equations, after the incorporation of the contact elements contribution, is highlighted using the example previously presented in Figure 5.2. In fact, the nonzero pattern of the global tangent matrix defining the reduced system of equations (5.39) is not altered with the introduction of contact elements. However, the equations (rows) associated with the displacement of nodes 3 and 4 (slave nodes) are modified by the contact elements, as illustrated in Figure 5.4 (a). In fact, all columns of these rows are changed due to the multiplication of the structural matrix by the second-order tensor $\mathbf{B}$ (see first line of (5.39)). The right hand side of the reduced system of equations (5.39) is modified according to (5.40), as shown in Figure 5.4 (a). Afterwards, the calculation of the Lagrange multipliers in the artificial nodes (number 5 and 6) uses the rows of the structural tangent matrix (associated with the slave nodes) multiplied by the solution vector containing the nodal displacements, as shown in Figure 5.3 (b). The grey colour is adopted to denote the solution vectors (displacements and contact forces). Note that the white/green squares composing the vectors presented in Figure 5.4 (a) and (b) do not contain the same information (see definition in (5.40)).

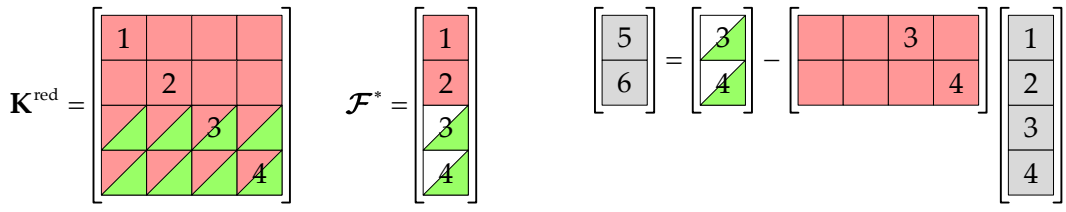


Figure 5.4. Pattern of the matrices required for the reduced system of equations: (a) global tangent matrix and residual vector to evaluate the nodal displacements; (b) structural tangent matrix and vector of nodal displacements to evaluate the nodal contact forces (grey colour denotes the solution vectors).

The solution of contact problems comprising a deformable body with a rigid obstacle can be achieved solving the reduced system of equations (5.39). The size of the global tangent matrix is defined only by the degrees of freedom related with the nodal displacements (primal variables). Thus, the size of the linear system of equations to be solved in each iteration of the Newton method is lower than the one required in the mixed system (5.25). Moreover, the tangent matrix of the reduced system does not contains any zero entry on diagonal, contrasting with the tangent matrix resulting from the mixed system of equations, as can be observed in the example (5.28). However, the contribution of the contact elements to the global tangent matrix using the standard assembly procedure disappears, in particular the idea of elemental contact Jacobian matrix (compare Figure 5.3 (b) with Figure 5.4). Moreover, the memory storage required for the tangent matrix of the reduced system of equations (Figure 5.4 (a)) together with the matrix necessary to calculate the contact forces (Figure 5.4 (b)) can be higher than the one necessary for the mixed system of equations.

5.2. Frictional contact between deformable bodies

In the previous section a relatively simple case of contact between a deformable body and a rigid surface was considered. In this section the formulation is generalized for the case of contact between two deformable bodies. The discrete formulation of the frictional contact problem between two deformable bodies using the augmented Lagrangian method is also implemented adopting the Node-to-Surface contact discretization technique, previously presented in Section 3.3.1. When two deformable bodies come in contact, the contact stresses arising in the contact interface are transferred from one body to the other according to the impenetrability and friction conditions. Due to the finite element discretization adopted, the slave surface is represented by a set of nodes while the master surface is defined by a collection of faceted finite elements. The required connection between the potential contact surfaces of the master and slave bodies is achieved with NTS contact elements. The classic NTS contact discretization does not pass the contact patch test [Taylor 91] when non-conforming meshes are used at the contact interface. However, the two-pass NTS approach is able to transfer correctly the uniform pressure through the contact interface (passes the patch test) for some particular interface discretizations (see Section 3.3.1.2).

The contribution of the frictional contact interface for the residual force vector involves the decomposition of the contact variables (kinematic and static) into normal and tangential components, as shown in (2.90). Thus, the classical piecewise bilinear representation of the master surface leads abrupt changes in the residual force vector for large sliding contact problems, causing serious convergence problems in the solution of the nonlinear system of

equations with the Newton method. In the present study the master surface is described by Nagata patches, yielding a smooth description of the contact surface (see the surface smoothing method in Chapter 4). The discontinuities introduced by the finite element discretization in the contact interface (see Figure 3.2) are reduced through the smoothing method, leading to a continuous distribution of the surface normal vector across patches. Since the master body is deformable, the interpolation coefficients of the Nagata patches, defined in Section 4.2.1, need to be re-evaluated at each iteration of each increment, using the new position and normal vectors of the nodes composing the master surface. Then, the normal gap function and tangential relative sliding (kinematic contact variables) are evaluated in the slave nodes taking into account the smooth description of the master surface.

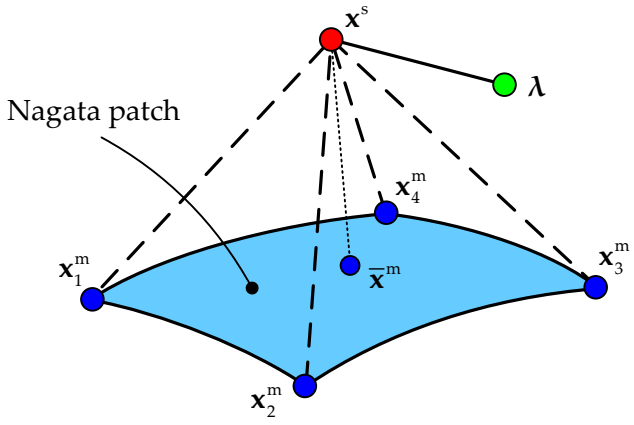


Figure 5.5. General form of a contact element of the type Node-to-Nagata patch using four master nodes (the artificial node for Lagrange multipliers is marked in green).

Each contact element used to connect the potential contact surfaces of two deformable bodies comprises one slave node and several master nodes attached to one structural finite element, as illustrated schematically in Figure 3.8 (b). The geometry of a contact element has a pyramid like shape, where the slave node is the top vertex and the base is composed by master nodes, as shown in Figure 5.5. As in the case of contact between a deformable body and a rigid obstacle, each contact element is complemented with an artificial node to store the Lagrange multipliers (contact force) resulting from the augmented Lagrangian method [Pietrzak 99]. The set of master nodes composing the contact element is defined by a single master facet pointing outside the body, i.e. it is defined by the intersection of a solid element with the master contact surface. For 3D meshes composed by 8-node hexahedral finite elements, the master contact boundaries are defined by quadrilateral surface elements (see e.g. Figure 3.14 (a)). On the other hand, the discretization of the body with 4-node tetrahedral elements leads to contact interfaces defined by triangular surface elements (see e.g. Figure 3.14 (b)). As previously mentioned, each bilinear facet that describes the master

surface is smoothed with a triangular or quadrilateral Nagata patch. Thus, each contact element involves a slave node and the associated Nagata patch on the master surface, which is selected by the normal projection of the slave node on the master surface (closest master segment). This is the key feature of the local contact search algorithm described in Section 3.4.2, which evaluates all the required quantities for each slave node. Besides, the contact element is complemented by an artificial node containing the contact force (Lagrange multipliers) as degrees of freedom (its location is arbitrary as shown in Figure 5.5).

The augmented Lagrangian frictional contact virtual work expressed in (2.90) involves the virtual work principle for the slave and master bodies, as well as the augmented virtual work developed by contact forces on the contact interface, which is splitted into three integrals over the non-contact, stick and slip zones. The virtual work of the discretized frictional contact problem is represented by a set of nonlinear equations, in which the nodal displacements of both bodies (slave and master) as well as the frictional contact force in the slave nodes are the final unknowns [Pietrzak 97]. For convenience, the kinematic variables (nodal displacements) of the slave and master bodies are gathered as $\mathbf{u}=(\mathbf{u}^1,\mathbf{u}^2)^T$ (see Figure 2.4 for notation). In fact, the contribution of the frictional contact to the global system of equations is dependent on the contact status of the slave node with respect to the master Nagata patch. Consequently, its definition comprises the configuration (nodal displacements) of both bodies in contact.

Following the augmented Lagrangian approach proposed by Alart and Curnier [Alart 91], the static equilibrium for the frictional contact problem between two discretized deformable bodies is governed by the system of nonlinear equations:

$$\begin{cases} (\mathbf{F}^{\text{int}}(\mathbf{u}^1)-\mathbf{F}^{\text{ext}})^1+\mathbf{F}^{\text{c}}{}_{\text{s}}(\mathbf{u},\boldsymbol{\lambda})=\mathbf{0} \\ (\mathbf{F}^{\text{int}}(\mathbf{u}^2)-\mathbf{F}^{\text{ext}})^2+\mathbf{F}^{\text{c}}{}_{\text{m}}(\mathbf{u},\boldsymbol{\lambda})=\mathbf{0} \\ -\frac{1}{\varepsilon}(\boldsymbol{\lambda}-\mathbf{F}^{\text{c}}{}_{\text{s}}(\mathbf{u},\boldsymbol{\lambda}))=\mathbf{0} \end{cases}, \tag{5.41}$$

where $\mathbf{F}^{\text{int}}$ and $\mathbf{F}^{\text{ext}}$ denote the internal and the external force vectors of each body, respectively. The first line of (5.41) is related with the equilibrium of the slave body while the second one expresses the equilibrium of the master body. The connection between both deformable bodies is expressed by the frictional contact operator, which is involved in the first two lines. The last line denotes the supplementary equations necessary to evaluate the contact force in the slave node. In the absence of contact between the bodies, the static equilibrium can be expressed independently for each body (two systems of equations), as illustrated in (5.41), where the last line vanishes as well as the frictional contact operator.

The difference between the system of equations (5.3) and the system (5.41) appears in the transmission of the frictional contact force from the slave node to the master body, which is preformed through the frictional contact operator. Since a single contact element involves one slave node, three or four master nodes and an artificial node for the Lagrange

multipliers, the mixed contact operator defined in (5.6) is extended to account the connection between the slave and master nodes, as shown in Figure 5.5 through dashed lines. The nodal displacements and contact forces in the slave nodes are directly included in the system of equations (5.41). However, both the displacement and the contact force associated with the contact point $\bar{\mathbf{x}}^{\mathrm{m}}$ on the master Nagata patch are not defined as independent variables. They are related to the same variables defined at the nodes of the Nagata patches, through the local coordinates of the contact point and local shape interpolation.

The contact interaction is governed by the principle of action and reaction, which is assured by the equilibrium between the discretized frictional contact operator defined for the slave node and the frictional contact operator in the contact point on the master surface. When the bodies are in contact, the slave node presents the same position than the contact point $\bar{\mathbf{x}}^{\mathrm{m}}$. The contact force in the slave node is defined as the force exerted by the slave node on the master body at the contact point. Since in general the contact point does not coincide with any master node (non-conforming meshes at contact interface or large sliding), this force should be distributed on the nodes of the master patch, according to the moment equilibrium. Considering that the contact element is composed by a quadrilateral Nagata patch (Figure 5.5), i.e. four master nodes, the frictional contact operator associated with the contact point is defined by:

$$\mathbf{F}^{\mathrm{c}}{}_{\mathrm{m}}(\mathbf{u},\boldsymbol{\lambda}) = \sum_{i=1}^{4} \mathbf{F}^{\mathrm{c}}{}_{\mathrm{m}i}(\mathbf{u},\boldsymbol{\lambda}) = -\mathbf{F}^{\mathrm{c}}{}_{\mathrm{s}}(\mathbf{u},\boldsymbol{\lambda}), \tag{5.42}$$

which is obtained from the summation over all master nodes composing the Nagata patch. Besides, according to the contact action–reaction principle, the frictional contact operator vector defined at the contact point (master) presents the same magnitude and opposite direction to the same operator defined in the slave node (first and last members of (5.42)). Therefore, according to (5.42), the frictional contact operator associated with each master node that composes the contact element can be expressed by a weighting factor of the frictional contact operator defined for the slave node:

$$\mathbf{F}^{\mathrm{c}}{}_{\mathrm{m}i}(\mathbf{u},\boldsymbol{\lambda}) = -w_i \mathbf{F}^{\mathrm{c}}{}_{\mathrm{s}}(\mathbf{u},\boldsymbol{\lambda}), \tag{5.43}$$

where the weight for each master node is evaluated according to the local coordinates of the contact point $(\bar{\eta},\bar{\zeta})$ in the Nagata patch domain. Since equation (5.43) represents the distribution of the contact force on the master nodes, the summation of all weights always gives one, i.e. constitute a partition of unity.

In the present study, the weight associated to each master node is given by the relative area of a portion of the patch, which is defined by the contact point coordinates. The area is measured in the parametric domain of the Nagata patch, as shown in Figure 5.6. Note that

the weight is defined using the area of the triangle/rectangle opposite to the node, as highlighted in the figure. The expressions for the weight associated to each master node in case of triangular Nagata patches are given by:

$$w_1 = 1-\bar{\eta}, \quad w_2 = \bar{\eta}-\bar{\zeta} \quad \text{and} \quad w_3 = \bar{\zeta}, \tag{5.44}$$

which involves the local coordinates of the contact point, calculated through the normal projection algorithm described in Section 3.4.2. The weight for each master node of the quadrilateral Nagata patch is obtained in a similar way, which can be written as:

$$w_1 = (1-\bar{\eta})(1-\bar{\zeta}), \quad w_2 = \bar{\eta}(1-\bar{\zeta}), \quad w_3 = \bar{\eta}\bar{\zeta} \quad \text{and} \quad w_4 = \bar{\zeta}(1-\bar{\eta}). \tag{5.45}$$

When the contact point is located exactly in the same position than one of the master nodes, the weight given by (5.44) or (5.45) for such master node is unitary, leading to the particular case of Node-to-Node contact element, illustrated in Figure 3.8 (a) for a 2D example.

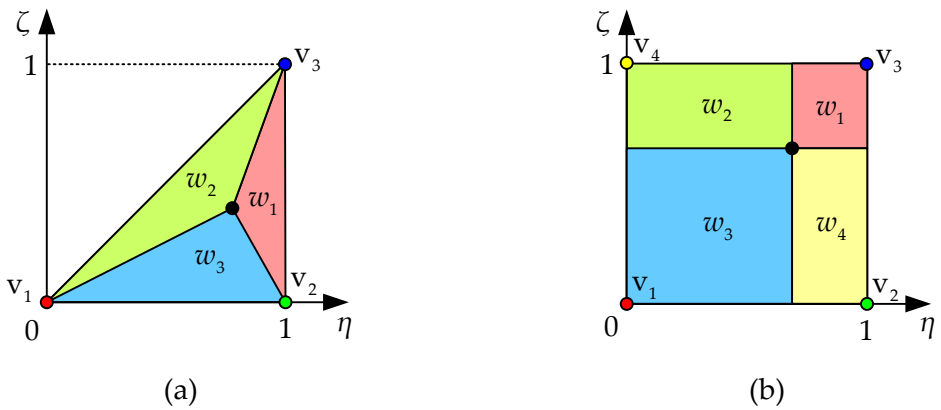


Figure 5.6. Definition of the weight for each master node based in the relative area: (a) triangular patch; (b) quadrilateral patch.

5.2.1. Residual vectors and Jacobian matrices

Since the contact contribution to the global system of equations (5.41) introduces non-differentiable equations [Alart 91], [Pietrzak 97], the contact part is separated from the structural part, as performed previously in case of contact with a rigid obstacle (5.11). Thus, the incorporation of the impenetrability and friction constraints into the static equilibrium problem is expressed by the mixed contact operator defined as follows:

$$\mathbf{F}^{*}(\mathbf{u},\boldsymbol{\lambda})=\left\{\begin{array}{c}\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda})\\ \mathbf{F}^{c}{}_{m1}(\mathbf{u},\boldsymbol{\lambda})\\ \mathbf{F}^{c}{}_{m2}(\mathbf{u},\boldsymbol{\lambda})\\ \mathbf{F}^{c}{}_{m3}(\mathbf{u},\boldsymbol{\lambda})\\ \mathbf{F}^{c}{}_{m4}(\mathbf{u},\boldsymbol{\lambda})\\ -\frac{1}{\varepsilon}(\boldsymbol{\lambda}-\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda}))\end{array}\right\}=\left\{\begin{array}{c}\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda})\\ -w_1\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda})\\ -w_2\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda})\\ -w_3\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda})\\ -w_4\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda})\\ -\frac{1}{\varepsilon}(\boldsymbol{\lambda}-\mathbf{F}^{c}{}_{s}(\mathbf{u},\boldsymbol{\lambda}))\end{array}\right\}=\left\{\begin{array}{c}\mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda})\\ -w_1\mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda})\\ -w_2\mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda})\\ -w_3\mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda})\\ -w_4\mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda})\\ \mathbf{F}^{\text{suppl}}(\mathbf{u},\boldsymbol{\lambda})\end{array}\right\},\tag{5.46}$$

which is similar to the one presented in (5.6) for the contact of a deformable body with a rigid obstacle. Nevertheless, the contact operator defined for the slave node is extended to the nodes of the master (deformable) body, according to the relationship expressed in (5.43). The mixed contact operator (5.46) is expressed for the specific case of a quadrilateral Nagata patch (four master nodes). For the general case of frictional contact between two deformable bodies using the Node-to-Segment contact discretization, the mixed contact operator (5.46) is defined by two distinct sub-operators. The sub-operator $\mathbf{F}^{\text{equi}}(\mathbf{u},\boldsymbol{\lambda})$ imposes the constraints due to contact with friction, which is function of the nodal displacements associated to the slave and master nodes. On the other hand, the sub-operator $\mathbf{F}^{\text{suppl}}(\mathbf{u},\boldsymbol{\lambda})$ introduces supplementary equations to calculate the frictional contact forces in the slave nodes. The above mentioned sub-operators are defined by the frictional contact operator (5.4), evaluated for each slave node according to its contact status.

The contribution of the contact elements to the equilibrium equations of the finite element model is given by the mixed contact operator (5.46), defined for each slave node, which is dependent of the contact status (gap, stick and slip) by means of (5.4). In the absence of contact between the bodies (gap status), the mixed contact operator is written as:

$$\mathbf{F}^{*}_{\text{gap}}(\mathbf{u},\boldsymbol{\lambda})=\left\{\begin{array}{c}\mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\\ -\frac{1}{\varepsilon}\boldsymbol{\lambda}\end{array}\right\},\tag{5.47}$$

which does not present any connection between the slave node and the master nodes composing the contact element. It only removes (if it exists) the contact force in the slave node. The mixed contact operator for stick status is given by:

$$\mathbf{F}^{*}_{\text{stick}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{Bmatrix} \hat{\lambda}_{\text{n}}\mathbf{n} + \hat{\boldsymbol{\lambda}}_{\text{t}} \\ -w_1(\hat{\lambda}_{\text{n}}\mathbf{n} + \hat{\boldsymbol{\lambda}}_{\text{t}}) \\ -w_2(\hat{\lambda}_{\text{n}}\mathbf{n} + \hat{\boldsymbol{\lambda}}_{\text{t}}) \\ -w_3(\hat{\lambda}_{\text{n}}\mathbf{n} + \hat{\boldsymbol{\lambda}}_{\text{t}}) \\ -w_4(\hat{\lambda}_{\text{n}}\mathbf{n} + \hat{\boldsymbol{\lambda}}_{\text{t}}) \\ g_{\text{n}}\mathbf{n} + \mathbf{g}_{\text{t}} \end{Bmatrix}, \tag{5.48}$$

where is highlighted the connection between the slave node (first line) and the four master nodes (next four lines), due to the contact action–reaction principle. Finally, for the slip status, the mixed contact operator is given by:

$$\mathbf{F}^{*}_{\text{slip}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{Bmatrix} \hat{\lambda}_{\text{n}}(\mathbf{n} - \mu\mathbf{t}) \\ -w_1(\hat{\lambda}_{\text{n}}(\mathbf{n} - \mu\mathbf{t})) \\ -w_2(\hat{\lambda}_{\text{n}}(\mathbf{n} - \mu\mathbf{t})) \\ -w_3(\hat{\lambda}_{\text{n}}(\mathbf{n} - \mu\mathbf{t})) \\ -w_4(\hat{\lambda}_{\text{n}}(\mathbf{n} - \mu\mathbf{t})) \\ g_{\text{n}}\mathbf{n} - \dfrac{1}{\varepsilon}(\boldsymbol{\lambda}_{\text{t}} + \mu\hat{\lambda}_{\text{n}}\mathbf{t}) \end{Bmatrix}, \tag{5.49}$$

where the tangential slip direction $\mathbf{t}$ is defined in (5.10), which is evaluated in the contact point using the Nagata patch definition (tangential directions). Note that in case of contact elements defined by a triangular Nagata patch, the mixed contact operators (5.47)–(5.49) only comprise three lines for the master nodes while the weight associated with each master node is given in (5.44).

The global system of nonlinear and partially non-differentiable equations (5.41) is solved using the generalized Newton method presented in Section 3.2. Since the coupled augmented Lagrangian method is adopted in the present study [Pietrzak 99], the nodal displacements of both deformable bodies and the contact forces (Lagrange multipliers) in the slave nodes are the unknowns of the linear system of equations shown in (5.13). The iterative solution scheme, based on the Newton method, requires a consistent linearization of the system of equations (5.41) to attain quadratic convergence. Therefore, in analogy to the tangent matrix of a structural finite element, the elemental contact Jacobian matrix is defined from the mixed contact operator vector (5.46), and it is expressed by:

$$\mathbf{J}^{*}=\begin{bmatrix}
\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}} \\
-w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_1w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_1w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_1w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_1w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_1\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}} \\
-w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_2w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_2w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_2w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_2w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_2\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}} \\
-w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_3w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_3w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_3w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_3w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_3\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}} \\
-w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_4w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_4w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_4w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & w_4w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}} & -w_4\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}} \\
\nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}} & -w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}} & -w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}} & -w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}} & -w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}} & \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{suppl}}
\end{bmatrix}, \tag{5.50}$$

which is evaluated according with the contact status of the slave node belonging to the contact element. Note that the matrix only contains the gradients of the two sub-operators involved in (5.46). Moreover, when the contact element is created using the triangular Nagata patch (three master nodes), the penultimate row and column are removed from the contact Jacobian matrix (5.50) and the new weights are given by (5.44).

Taking into account (5.50) and the mixed contact operators (5.47)–(5.49), the contact Jacobian matrices can be explicitly evaluated for each contact status (gap, stick and slip). The elemental Jacobian matrix for the gap status is obtained from (5.47) and it is written as:

$$\mathbf{J}^{*}_{\text{gap}}(\mathbf{u},\boldsymbol{\lambda})=\begin{bmatrix}
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -(1/\varepsilon)\mathbf{I}
\end{bmatrix}, \tag{5.51}$$

where $\mathbf{I}$ is the second order identity tensor. The elemental Jacobian matrix for the stick status is derived from (5.48) and it is expressed as follows:

$$\mathbf{J}^{*}_{\text{stick}}(\mathbf{u},\boldsymbol{\lambda})=\begin{bmatrix}
\varepsilon\mathbf{I} & -w_1\varepsilon\mathbf{I} & -w_2\varepsilon\mathbf{I} & -w_3\varepsilon\mathbf{I} & -w_4\varepsilon\mathbf{I} & \mathbf{I} \\
-w_1\varepsilon\mathbf{I} & w_1w_1\varepsilon\mathbf{I} & w_1w_2\varepsilon\mathbf{I} & w_1w_3\varepsilon\mathbf{I} & w_1w_4\varepsilon\mathbf{I} & -w_1\mathbf{I} \\
-w_2\varepsilon\mathbf{I} & w_2w_1\varepsilon\mathbf{I} & w_2w_2\varepsilon\mathbf{I} & w_2w_3\varepsilon\mathbf{I} & w_2w_4\varepsilon\mathbf{I} & -w_2\mathbf{I} \\
-w_3\varepsilon\mathbf{I} & w_3w_1\varepsilon\mathbf{I} & w_3w_2\varepsilon\mathbf{I} & w_3w_3\varepsilon\mathbf{I} & w_3w_4\varepsilon\mathbf{I} & -w_3\mathbf{I} \\
-w_4\varepsilon\mathbf{I} & w_4w_1\varepsilon\mathbf{I} & w_4w_2\varepsilon\mathbf{I} & w_4w_3\varepsilon\mathbf{I} & w_4w_4\varepsilon\mathbf{I} & -w_4\mathbf{I} \\
\mathbf{I} & -w_1\mathbf{I} & -w_2\mathbf{I} & -w_3\mathbf{I} & -w_4\mathbf{I} & \mathbf{0}
\end{bmatrix}, \tag{5.52}$$

which considers that the local frame defined on the master surface is fixed in all Newton iterations within an increment. This assumption is also effective for the contact between deformable bodies, because the solution obtained in the stick status is path-independent [Alart 95a]. Indeed, this contact status imposes null relative displacement between the slave node and the Nagata patch (master surface). Finally, the elemental Jacobian matrix for the slip status assuming a fixed local frame attached to the Nagata patch is defined by:

$$\mathbf{J}^{*}_{\text{slip}}(\mathbf{u},\boldsymbol{\lambda}) = \begin{bmatrix} \varepsilon\mathbf{M} & -w_1\varepsilon\mathbf{M} & -w_2\varepsilon\mathbf{M} & -w_3\varepsilon\mathbf{M} & -w_4\varepsilon\mathbf{M} & \mathbf{M} \\ -w_1\varepsilon\mathbf{M} & w_1w_1\varepsilon\mathbf{M} & w_1w_2\varepsilon\mathbf{M} & w_1w_3\varepsilon\mathbf{M} & w_1w_4\varepsilon\mathbf{M} & -w_1\mathbf{M} \\ -w_2\varepsilon\mathbf{M} & w_2w_1\varepsilon\mathbf{M} & w_2w_2\varepsilon\mathbf{M} & w_2w_3\varepsilon\mathbf{M} & w_2w_4\varepsilon\mathbf{M} & -w_2\mathbf{M} \\ -w_3\varepsilon\mathbf{M} & w_3w_1\varepsilon\mathbf{M} & w_3w_2\varepsilon\mathbf{M} & w_3w_3\varepsilon\mathbf{M} & w_3w_4\varepsilon\mathbf{M} & -w_3\mathbf{M} \\ -w_4\varepsilon\mathbf{M} & w_4w_1\varepsilon\mathbf{M} & w_4w_2\varepsilon\mathbf{M} & w_4w_3\varepsilon\mathbf{M} & w_4w_4\varepsilon\mathbf{M} & -w_4\mathbf{M} \\ \mathbf{M} & -w_1\mathbf{M} & -w_2\mathbf{M} & -w_3\mathbf{M} & -w_4\mathbf{M} & \dfrac{1}{\varepsilon}(\mathbf{M}-\mathbf{I}) \end{bmatrix}, \tag{5.53}$$

where the second-order tensor $\mathbf{M}$ is defined in (5.18). The Jacobian matrix for the slip status that takes into account the contributions of geometric terms arising from the gradient of the surface normal vector can be obtained in a similar way from (5.20). Indeed, the Jacobian matrix for the slip status adopted in the numerical examples presented in Chapter 6 is derived from (5.20), which takes into account the master surface curvature.

Considering a general contact element composed by four master nodes (quadrilateral Nagata patch) connecting two deformable bodies, the mixed linear system of equations for a single Newton iteration (the indices related with the iteration number will be omitted) can be written as:

$$\begin{cases} \mathbf{K}^1\{d\mathbf{u}^1\} + \nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{s}}\} + \sum_{i=1}^{4}(-w_i\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{m}i}\}) + \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}}\{d\boldsymbol{\lambda}^{\text{s}}\} = \{d\mathbf{f}^{\text{s}}\} - \mathbf{F}^{\text{equi}} \\ \mathbf{K}^2\{d\mathbf{u}^2\} - w_1\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{s}}\} + \sum_{i=1}^{4}(w_1w_i\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{m}i}\}) - w_1\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}}\{d\boldsymbol{\lambda}^{\text{s}}\} = \{d\mathbf{f}^{\text{m1}}\} + w_1\mathbf{F}^{\text{equi}} \\ \mathbf{K}^2\{d\mathbf{u}^2\} - w_2\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{s}}\} + \sum_{i=1}^{4}(w_2w_i\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{m}i}\}) - w_2\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}}\{d\boldsymbol{\lambda}^{\text{s}}\} = \{d\mathbf{f}^{\text{m2}}\} + w_2\mathbf{F}^{\text{equi}} \\ \mathbf{K}^2\{d\mathbf{u}^2\} - w_3\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{s}}\} + \sum_{i=1}^{4}(w_3w_i\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{m}i}\}) - w_3\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}}\{d\boldsymbol{\lambda}^{\text{s}}\} = \{d\mathbf{f}^{\text{m3}}\} + w_3\mathbf{F}^{\text{equi}} \\ \mathbf{K}^2\{d\mathbf{u}^2\} - w_4\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{s}}\} + \sum_{i=1}^{4}(w_4w_i\nabla_{\mathbf{u}}\mathbf{F}^{\text{equi}}\{d\mathbf{u}^{\text{m}i}\}) - w_4\nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{equi}}\{d\boldsymbol{\lambda}^{\text{s}}\} = \{d\mathbf{f}^{\text{m4}}\} + w_4\mathbf{F}^{\text{equi}} \\ \nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}}\{d\mathbf{u}^{\text{s}}\} + \sum_{i=1}^{4}(-w_i\nabla_{\mathbf{u}}\mathbf{F}^{\text{suppl}}\{d\mathbf{u}^{\text{m}i}\}) + \nabla_{\boldsymbol{\lambda}}\mathbf{F}^{\text{suppl}}\{d\boldsymbol{\lambda}^{\text{s}}\} = -\mathbf{F}^{\text{suppl}} \end{cases}, \tag{5.54}$$

where $\mathbf{K}^1$ and $\mathbf{K}^2$ denotes the global tangent matrix of the slave and master bodies, respectively. The nodal displacement of the slave node is denoted by $\{d\mathbf{u}^{\text{s}}\}$ and the nodal displacement of each master node is represented by $\{d\mathbf{u}^{\text{m}i}\}$, while the increment of nodal contact force vector in the slave node is designated by $\{d\boldsymbol{\lambda}^{\text{s}}\}$, which is stored in the artificial node. Moreover, the non-equilibrated force vector is denoted by $\{d\mathbf{f}^{\text{s}}\}$ for the slave node and by $\{d\mathbf{f}^{\text{m}i}\}$ in case of the master nodes. The definition of the linear system of equations (5.54) comprises the structural tangent matrix of each deformable body and the contact Jacobian matrix (5.50). The right hand side of the global system contains the mixed contact operator (5.46) associated to all nodes constituting the contact element.

In order to exemplify the structure of the global tangent matrix after the incorporation of the contact elements, a simple 2D example is presented [Aragón 13]. The mechanical system involves two discretized deformable bodies coming in frictional contact, as shown in Figure 5.7. The upper body is assigned as slave body and the lower one as master, each one discretized by a single structural element. A constant pressure is imposed on the upper segment between nodes 1 and 3. Besides, nodes 6 and 7 are fixed, thus they are not included in the global system of equations. In absence of contact, the global tangent matrix and the residual vector for such configuration have the pattern presented in Figure 5.8 (a). The contribution of each body to the tangent matrix is illustrated using two different colours, highlighting the geometrical independence of the bodies.

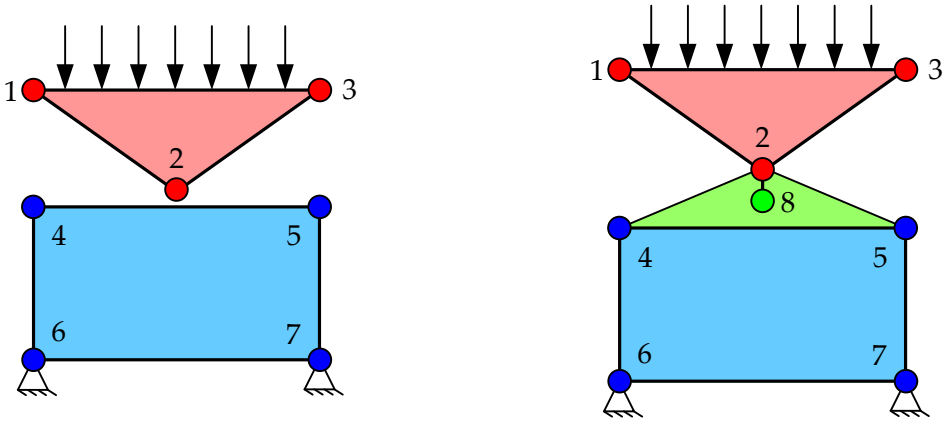


Figure 5.7. Example of two discretized deformable bodies coming in contact for the cases: (a) absence of contact elements; (b) with a contact element.

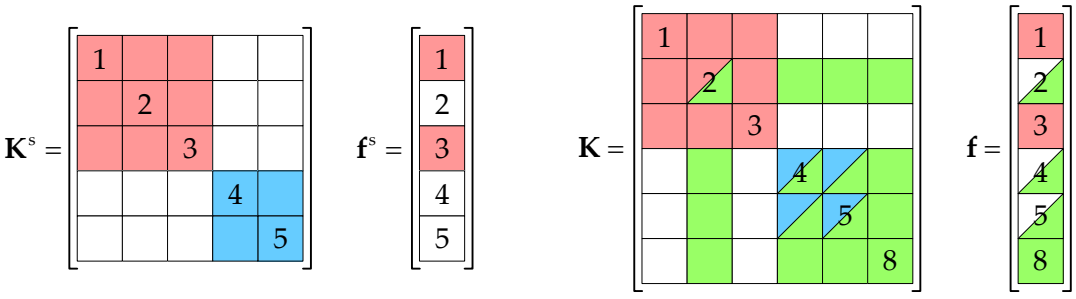


Figure 5.8. Pattern of the global tangent matrix and residual vector of two discretized bodies coming in contact: (a) absence of contact elements; (b) with a contact element.

In order to account for the frictional contact between the bodies, a NTS contact element is introduced at the contact interface, as shown in Figure 5.7 (b). Therefore, the connection between the bodies is created by means of such contact element (green), which is defined

by the slave node 2, the master segment joining nodes 4 and 5, and a complementary node 8 for the Lagrange multipliers. The selection of the master segment (nodes 4 and 5) that defines the contact element together with the salve node is performed by the contact search algorithm, previously presented in Section 3.4. Indeed, from the geometrical point of view, the slave node 2 may come in contact with the master segment located between nodes 4 and 5 (existence of the normal projection), as presented in Figure 5.7 (b). The contribution of the contact element to the global tangent matrix and residual vector is shown in Figure 5.8 (b) by means of the green colour. Since the considered example is 2D, the contact element is composed by four nodes (including the artificial one), leading to a contact Jacobian matrix (5.50) with four rows/columns. The sixteen elements of the Jacobian matrix are added to the global tangent matrix taking into account the four nodes that define the contact element, as shown in Figure 5.8 (b). The mixed contact operator (5.46) is introduced in the global residual vector through the nodes involved in the definition of the contact element, as shown in Figure 5.8 (b). As in the particular case of contact between a deformable body and a rigid obstacle, the pattern of nonzero entries in the arising global tangent matrix is symmetric (Figure 5.8 (b)). This important feature allows to solve the linear system of equations (5.54) considering the global tangent matrix as symmetrically structured (Intel MKL library).

5.2.2. Large sliding contact element

The main difficults in the treatment of large sliding frictional contact problems are the faceted description of the master surface, which is introduced by the finite element discretization. The usual piecewise bilinear representation of the master surface leads to abrupt changes in the surface normal vector across boundaries of adjacent facets. Consequently, this surface representation leads to jumps in the contact forces when a slave node slides from one master facet to a neighbouring one, which typically results in severe convergence problems [Heegaard 96], [Pietrzak 97]. These numerical problems are solved in the present study by means of the surface smoothing method presented in Chapter 4, which creates a new description of the master surface using Nagata patches.

As mentioned in the preceding section, a standard contact element consists of one slave node, one artificial nodes and one Nagata patch, where the latter can be composed by three or four master nodes (Figure 5.5). The purpose of the contact search algorithm, described in Section 3.4, is to select the correct Nagata patch for each slave node, which is based in the normal projection of the slave node on the master Nagata patches. However, due to the large sliding between the bodies, the patch selected in the current iteration/increment for a specific slave node should be updated in the subsequent steps, increasing the computational cost associated with contact detection. Moreover, the procedure necessary to create new contact elements (update master nodes) is relatively time consuming, since it

requires an update of the global tangent matrix pattern [Yastrebov 13]. Thus, this includes the invoking once each Intel MKL DSS interface routine, as shown in Figure 3.6, i.e. delete old and create new all data structures.

These difficulties can be easily illustrated through a simple example comprising two discretized deformable bodies undergoing large sliding, as presented in Figure 5.9. In order to account the sliding of the slave node 2 over several master segments, the master body is discretized by three finite elements. A prescribed displacement is imposed to the slave body, applying a movement from the left to the right side. Besides, nodes 8, 9, 10 and 11 are not included in the global system of equations because they are fixed (Dirichlet boundary conditions). The selection of the master segment composing the contact element is based in geometrical considerations, by means of the normal projection of the slave node on the master surface. Therefore, at the beginning of the sliding (Figure 5.9 (a)) the contact element is defined by the master nodes 4 and 5, while at the end of the sliding (Figure 5.9 (b)) the contact element is composed by the master nodes 6 and 7.

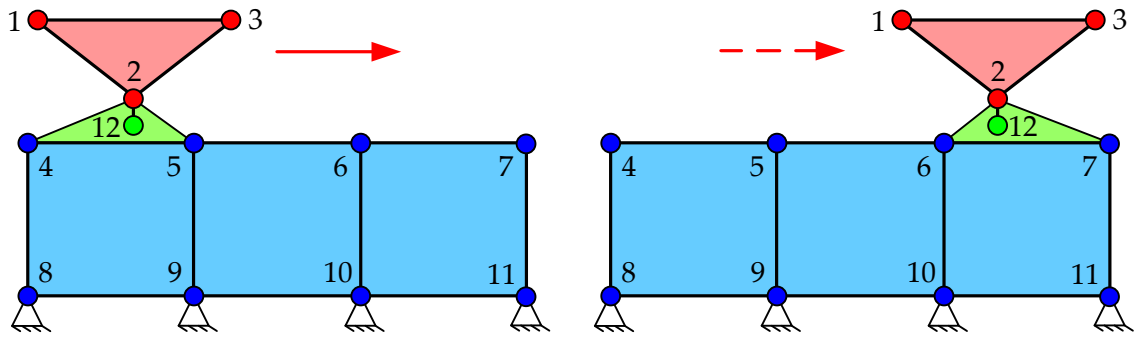


Figure 5.9. Example of two discretized deformable bodies undergoing large sliding: (a) configuration at the beginning; (b) configuration at the end.

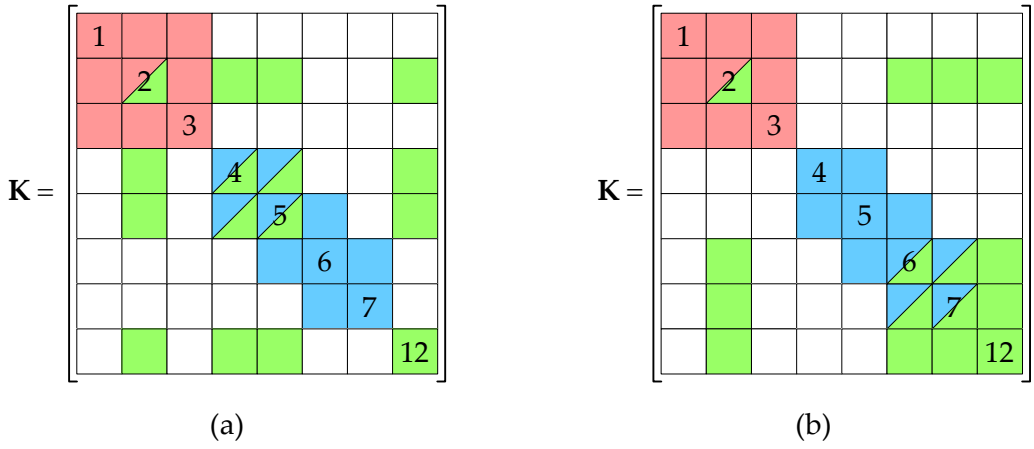


Figure 5.10. Pattern of the global tangent matrix of two discretized bodies undergoing large sliding: (a) at the beginning of sliding; (b) at the end of sliding.

The pattern of the global tangent matrix at the beginning of the sliding, obtained after the incorporation of contact element contribution, is presented in Figure 5.10 (a). The connection between the slave node 2, the master nodes 4 and 5, and the artificial node 12 is generated by the contact Jacobian matrix (5.50), whose elements are represented by green squares in the global matrix. However, the pattern of the global tangent matrix is modified when the slave node slides to another master facet, due to the change of the master nodes that constitute the contact element. The pattern of the global tangent matrix at the end of the sliding over three facets is presented in Figure 5.10 (b), highlighting the modification of the matrix pattern caused by the new contact Jacobian matrix, which comprises the master nodes 6 and 7. Nevertheless, the artificial node 12 for the Lagrange multipliers is always connected with the slave node (see Figure 5.9). Therefore, the large sliding frictional contact between deformable bodies requires the update of the global tangent matrix structure (pattern of nonzero entries). Note that in the case of large sliding contact between a deformable body and a rigid obstacle, the global tangent matrix presents always a fixed pattern. This important property arises from the fact that the contact force evaluated in the slave node is not transferred to the master body.

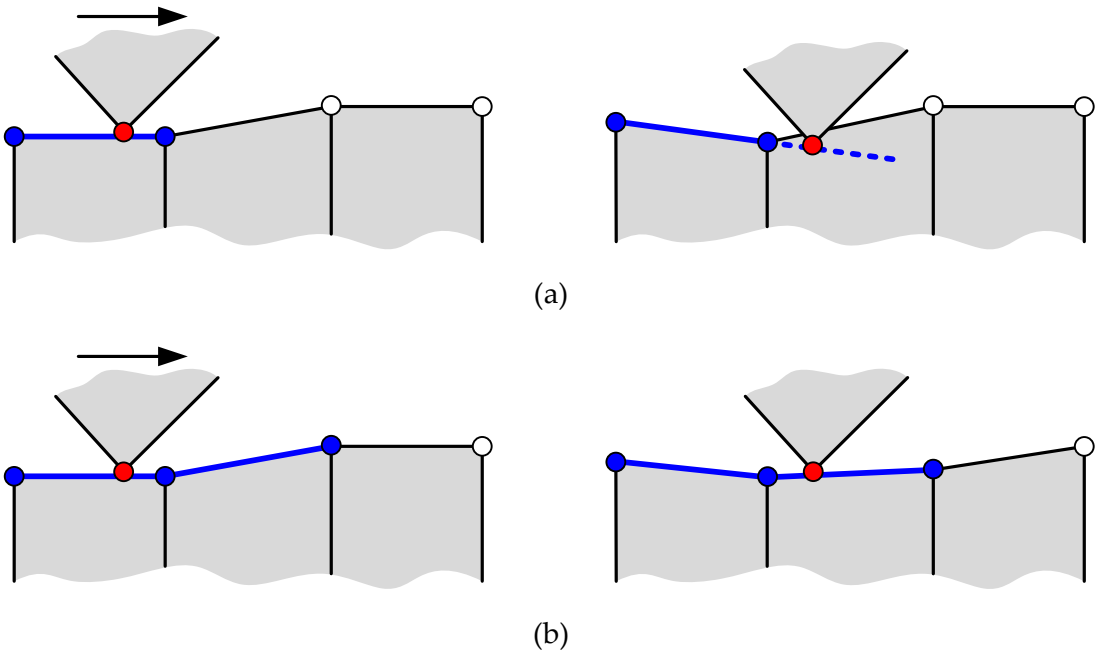


Figure 5.11. Example of large sliding contact comprising two distinct approaches: (a) extension of the master segment domain; (b) multi-face contact element (adapted from [Yastrebov 13]).

The standard contact element is defined by a single master Nagata patch, which is determined by the closest point projection of the slave node onto the master surface. When during the iterative process the slave node slides out of the selected Nagata patch, it is no longer possible to project the node on this patch. However, the assignment of the gap status

to this slave node would result in a geometrical penetration on the master surface (violation of the impenetrability condition). Therefore, the contact element should be recreated using the new patch, which is a procedure computationally expensive due to the modification of the nonzero pattern of the global tangent matrix (Figure 5.10). In order to avoid such drawbacks, two distinct approaches can be applied in large sliding frictional contact problems:

- Extension of the parametric (local) domain of the Nagata patch (see Figure 5.11 (a));
- Creation of multi-face contact elements in advance, i.e. incorporation of the adjacent Nagata patches into the contact element definition (see Figure 5.11 (b)).

The first approach yields a continuous sliding of the slave node along the extension of the actual patch, as shown in Figure 5.11 (a). Adopting such approach, the equilibrium is reached for a location of the slave node outside the patch domain, which results either in penetration, for a locally concave master surface, or in opening of the gap, in case of convex surface. The violation of the impenetrability conditions can be acceptable for small values of the sliding increment outside the patch domain in comparison to the size of the patch, otherwise the obtained solution becomes strongly inaccurate [Yastrebov 13]. The second approach uses the multi-face contact elements proposed by [Heegaard 93], which consists in creating contact elements composed by one slave node and several master patches, as shown in Figure 5.11 (b). This strategy avoids frequent updating of the contact elements and allows dealing with large sliding in a single increment, reducing the overall computation time and improving the contact accuracy in comparison with the first approach.

The structure of the residual vector and the tangent matrix are slightly different in the multi-face contact elements. Among all master patches composing the contact element, only one is active on the current iteration of the Newton's method. Thus, the master nodes of the contact element are divided into active and inactive nodes. The active master nodes, which are attached to the active Nagata patch, are connected with the slave node to create a contact element. Since the structure of the global tangent matrix is created taking into account all master nodes (active and inactive), when the slave node slides from one patch to another, the values of the tangent matrix are updated, but not its structure. Note that only active components introduce nonzero values to the residual vector and tangent matrix.

Another drawback in the numerical simulation of contact problems using the finite element method is the constant switching between two adjacent master segments (*flip–flop* effect), as illustrated in Figure 5.12. This effect occurs from one iteration to another when adopting the multi-face contact elements, due to the switch of active master segment, usually leading to the cycling of the numerical algorithms. Indeed, several master segments are available for a single slave node within an increment. However only one is active in

each iteration, which can alternate between two neighbouring master segments. Such oscillations can also occurs using the first approach, nevertheless in this case the oscillations take place from one increment to another, assuring the local convergence (see Figure 5.12). In fact, this drawback is more common when the master surface is represented by a piecewise bilinear finite element mesh, where such undesirable effect is sometimes avoided adopting the Node-to-Edge contact element, which connects the slave node with an edge avoiding switching.

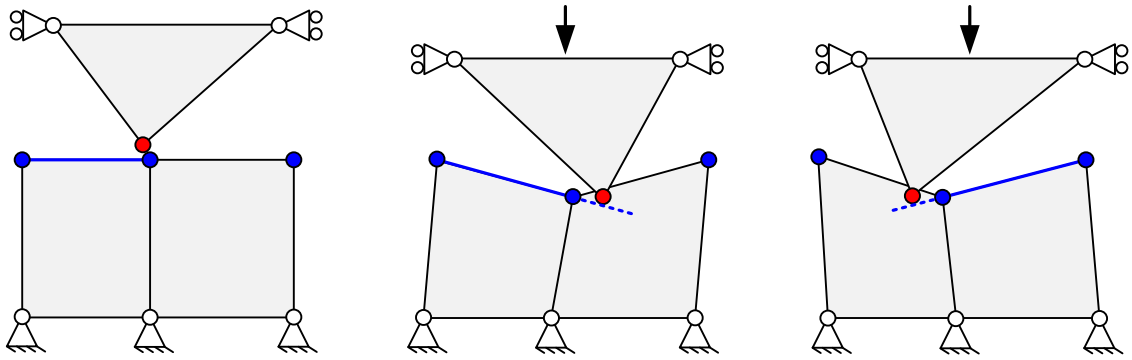


Figure 5.12. Example of constant switching between two adjacent master segments (*flip–flop* effect) (adapted from [Yastrebov 13]).

In the present study, the solution adopted to deal with large sliding frictional contact problems is based in both approaches previously described. The multi-face contact elements are composed by several Nagata patches, yielding a smooth description of the contact master surface. Moreover, the parametric domain of the active Nagata patch included in the definition of the contact element is extended 1% in each direction, allowing to eliminate the small blind spots observed in the normal projection of the slave nodes (see Figure 3.23 (a)). The amount of Nagata patches used in each multi-face contact element is defined by the neighbouring patches located around the patch selected in the contact detection, i.e. all Nagata patches sharing the master nodes of the patch on which the slave node is projected. Considering the master surface discretized by regular quadrilateral patches, each multi-face contact element is composed by 9 Nagata patches (16 master nodes). Nevertheless, for the general case of arbitrary surface mesh, the quantity of master patches is directly defined through the mesh topology, which can be different for each contact element (slave node).

The key point in the adoption of multi-face contact elements is the definition of the pattern to the global tangent matrix before starting the iterative solution procedure. The nonzero structure of the matrix is preserved in all iterations of the Newton's method, within an increment, and posteriorly updated at the beginning of the next increment, using the sequence presented in Figure 3.6. The example previously presented in Figure 5.9 is used again to highlight the advantages of the multi-face contact elements in case of frictional

contact problems involving large sliding. The multi-face contact element is represented in a symbolic manner by lines connecting the slave node with all master nodes assigned to this element, as presented in Figure 5.13 (a). The master segment joining nodes 5 and 6 is selected by the normal projection of the slave node onto the master surface, while the neighbouring are include to deal with large sliding. The pattern of the global tangent matrix is shown in Figure 5.13 (b) for a single multi-face contact element. Due to the existence of several master segments, the number of green squares in the tangent matrix increases, which represent elements of the Jacobian matrix. Nevertheless, only the active master nodes introduce nonzero values to the tangent matrix. In fact, the pattern of the matrix presented in Figure 5.13 (b) includes the two different patterns shown in Figure 5.10. Thus, the sliding of the slave node over the three master segments can be accurately evaluated without modify the nonzero pattern of the tangent matrix.

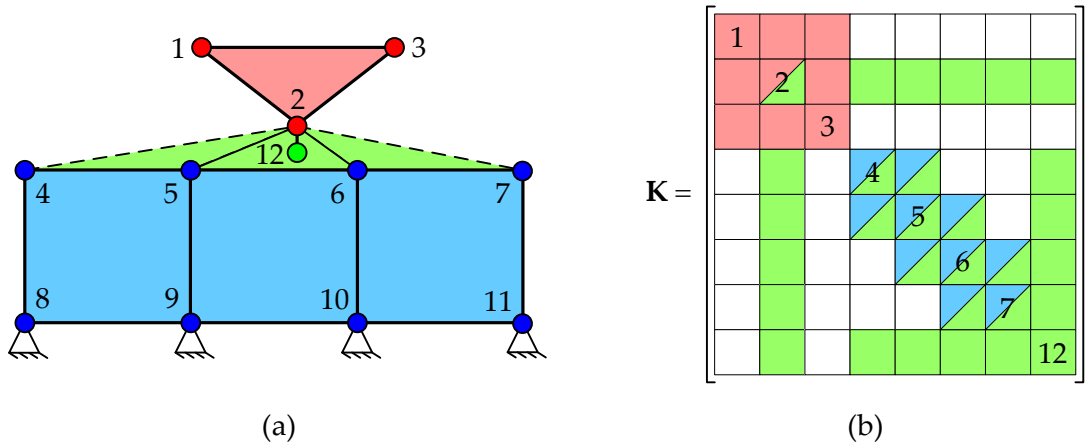


Figure 5.13. Example of large sliding contact: (a) two discretized bodies and representation of the multi-face contact element; (b) pattern of the global tangent matrix.

Chapter 6

Numerical Examples

This chapter presents the numerical results of some frictional contact problems, obtained using the developed algorithms and methods, currently implemented in the finite element code DD3IMP. Ten different numerical examples are presented, ranging from simple examples with known analytical solution to complex problems with significant industrial interest, namely sheet metal forming processes. The selected contact problems are divided into two groups: (i) contact between deformable and rigid bodies and (ii) contact between deformable bodies, all of them modelled in 3D framework. Besides, the examples are presented in such a way that the degree of nonlinearity increases problem by problem, i.e. large deformations, nonlinear material models, friction and large sliding. The accuracy, robustness and performance of the proposed 3D contact surface smoothing method is demonstrated by means of its comparison with the traditional piecewise bilinear finite element mesh representation. All numerical simulations are carried out on a computer machine equipped with an Intel Core™ i7–2600K Quad-Core processor (3.4 GHz), 8.0 GB RAM and the Windows 7 Professional (64-bits platform) operating system.

6.1. Contact between deformable and rigid bodies

This section contains four numerical examples involving the frictional contact between a deformable body and rigid obstacles. The first example considers the frictional sliding of a cube on a rigid flat surface. The Coulomb friction law is analysed, comparing the difference between local and global coefficients of friction. The second example presented is the ironing problem, which comprises large sliding between a curved rigid surface and the deformable elastic body. The influence of the surface smoothing method on the contact

force evolution is highlighted. The last two examples comprise the industrial application of the developed algorithms for sheet metal forming processes. The frictional contact between the sheet and the forming tools is studied for the reverse deep drawing and the automotive underbody cross member panel. The numerical results obtained with the Node-to-Nagata contact element are compared with the traditional representation of the tool surfaces using bilinear facets.

6.1.1. Frictional sliding of a cube on a plane

The first example studied involves the frictional sliding of a deformable cube on a rigid plane, which was firstly proposed by Yastrebov [Yastrebov 13]. The geometry of the bodies including the discretization of the cube are presented in Figure 6.1. The finite element mesh of the cube is composed by 512 solid finite elements (8-node hexahedral) and 729 nodes. The rigid flat surface (a single Nagata patch) is firstly moved towards the cube (0.05 mm in vertical direction) and afterwards it is moved along the plane (0.25 mm in horizontal direction). The elastic material properties of the unitary cube are indicated in Figure 6.1, as well as the boundary conditions. The purpose of this problem is to evaluate the difference between local and global (measured) coefficients of friction between the cube and the rigid surface, i.e. the fulfilment of the friction law constraints. Thus, three different coefficients of friction are considered in this study: $\mu = 0.1, 0.3, 0.6$, for the Coulomb's law.

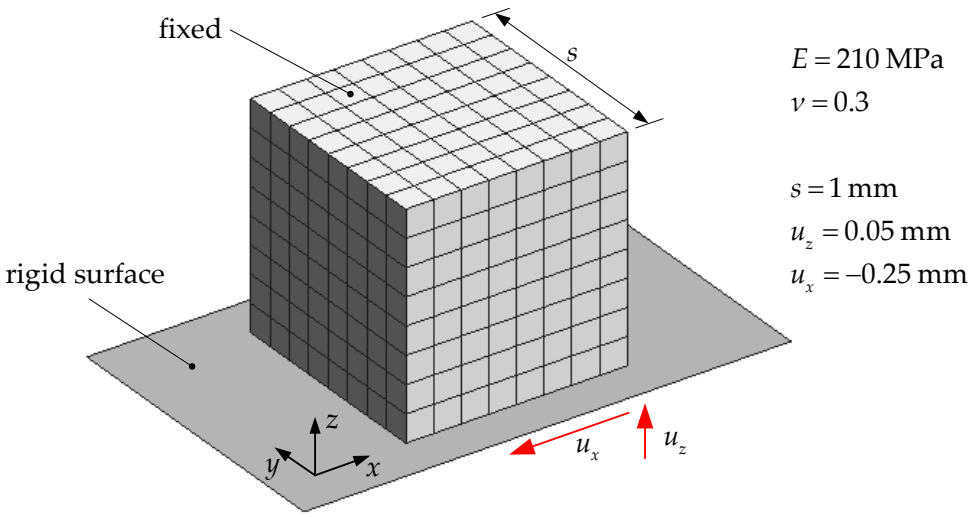


Figure 6.1. Geometrical setting of the sliding cube problem with finite element mesh.

The vertical displacement imposed to the rigid surface is applied within 10 equal increments, while the subsequent horizontal displacement is applied within 50 equal increments. Note that the horizontal displacement is 5 times higher than the vertical one. In order to represent the force evolution of both stages as a function of a single variable, the

first stage (vertical displacement) is represented in the pseudo-time from 0 to 1, while the second stage is represented in the pseudo-time from 1 to 3. The total vertical reaction (normal reaction force) scaled by the friction coefficient $\mu|F_n|$ and the tangential reaction $|T_x|$ are presented in Figure 6.2 as a function of the pseudo-time, for the three different values of friction coefficient. During the vertical movement of the rigid flat surface towards the cube, the tangential reaction force is zero for all values of friction coefficient, while the normal reaction increases linearly with the vertical displacement (pseudo-time), as shown in Figure 6.2. This does not means that the frictional force is zero in all nodes of the contact interface. On the contrary the sum of all tangential components of the force is zero due to geometrical symmetry conditions. The horizontal sliding of the rigid surface in frictional contact with the cube starts when the pseudo-time is equal to 1. Nevertheless, the transition of the slave nodes from the stick to the slip status occurs progressively according to the relative sliding between the bodies. In fact, the total tangential reaction force increases from zero until attaining the limit value defined by the product between the normal reaction force and the friction coefficient, as shown in Figure 6.2. The instant where this threshold value is attained increases with the global friction coefficient, since the transition of the slave nodes to slip status occurs for higher values of tangential force and consequently more deformation. However, the absolute value of the tangential reaction (x-direction) is always lower than the absolute value of the normal contact force multiplied by the friction coefficient, because the nodal contact force is not completely aligned with the x-direction. The relative difference between the global coefficient of friction (measured) and the predefined local one (input value) is always inferior to 0.7% (see Figure 6.2), when steady state conditions are considered.

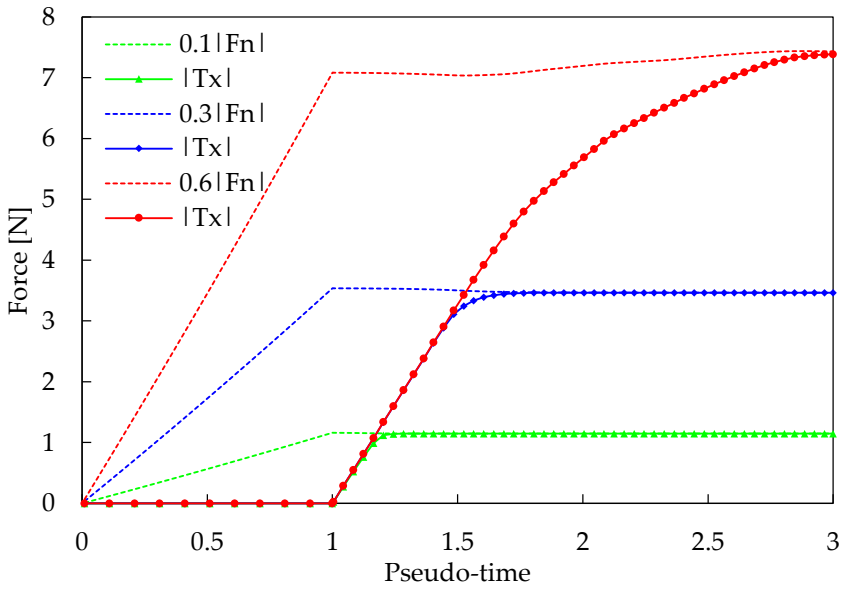


Figure 6.2. Evolution of the scaled normal force (dashed line) and the tangential force (solid line with marker) for three different values of friction coefficient.

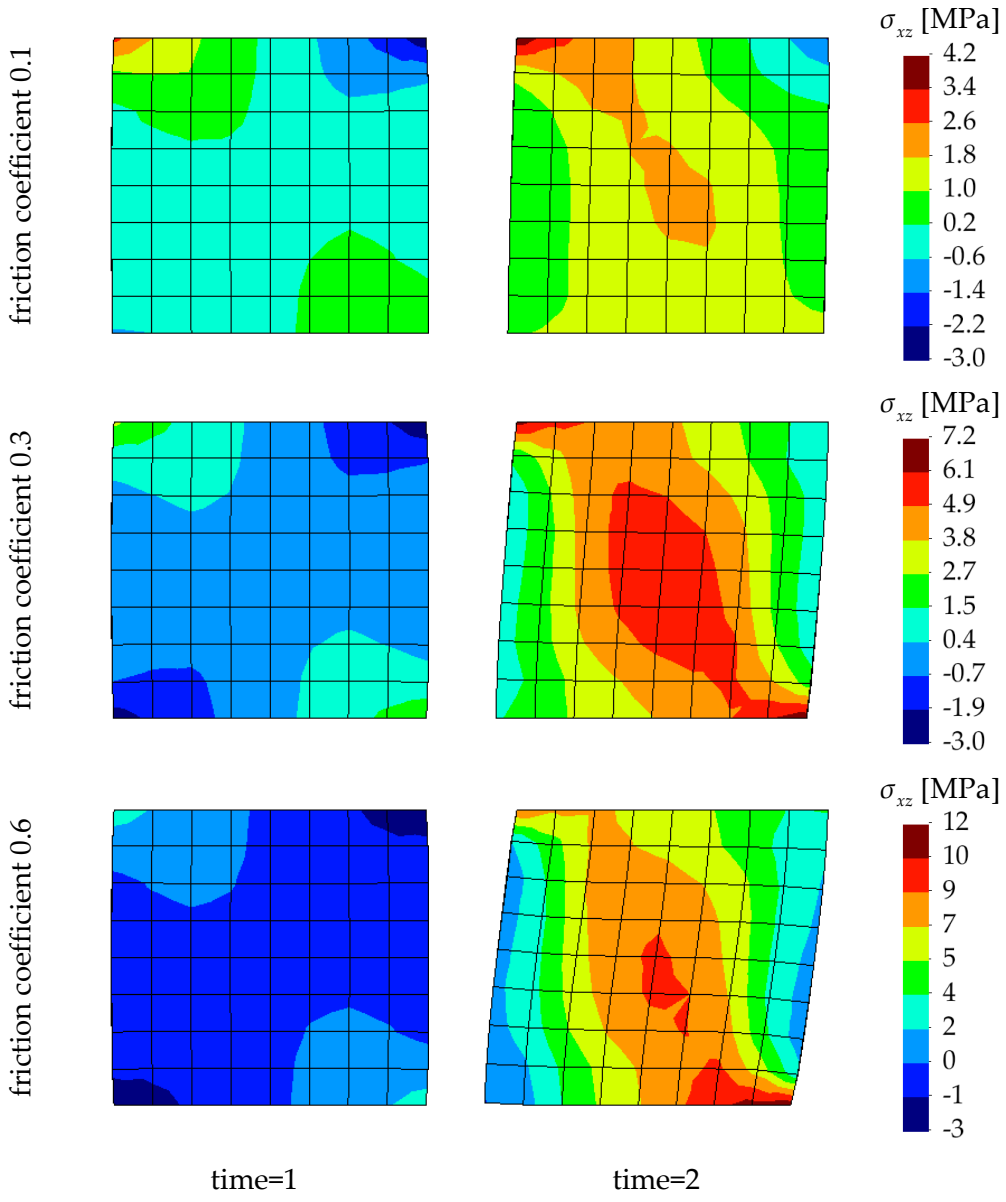


Figure 6.3. Contour plots of shear stress for different friction coefficients and time instants (side view).

The deformed configuration of the cube and the corresponding shear stress σ_{xz} distributions in the xOz lateral view are shown in Figure 6.3, for the three considered friction coefficients and two time instants. The first instant presented (time=1) corresponds to the end of the prescribed vertical displacement and, consequently, beginning of the horizontal sliding. During the first stage, the curvature of the vertical walls increases with the increase of the friction coefficient value, as shown in Figure 6.3. The second instant (time=2) was selected to show the deformed configuration of the cube in the middle of the prescribed horizontal displacement, i.e. 0.125 mm of displacement. At this time instant, the sliding of the cube on the rigid surface already reach the steady state for the two lower

values of friction coefficient, as shown through the constant tangential force evolution presented in Figure 6.2. Nevertheless, the case with higher value of friction coefficient requires a higher sliding distance to attain the maximum value of tangential force, because the detachment of a part of cube occurs (see the contact interface in Figure 6.3). In fact, at the instant presented in Figure 6.3, the last three layers of slave nodes located in the left-hand side of the cube do not present contact with the rigid surface. Since detachment occurs in the contact zone, the sliding velocity is altered, as can be observed through the slop of the tangential force curve illustrated in Figure 6.2. The location of the maximum value of shear stress changes from the top surface in the case with lower value of friction coefficient to the contact interface for the higher value of friction, as shown in Figure 6.3. The obtained results are in accordance with the ones presented by Yastrebov [Yastrebov 13].

6.1.2. Ironing problem

The second example comprises the sliding a rigid cylindrical die along a deformable slab, as illustrated in Figure 6.4. This example is inspired in the one proposed by Puso and Laursen [Puso 04b], which deals with the contact between two deformable bodies using the mortar based method. The slab dimensions are 9×4×3 mm and the die has 5.2 mm wide with a radius of 3 mm. The finite element mesh of the deformable body is presented in Figure 6.4, where the slab is discretized with 1,800 solid finite elements (8-node hexahedral). It is assumed that the slab is elastic $(E = 1\,\text{MPa}, \nu = 0.3)$, in order to focus the analysis on the interface behaviour. The friction between the die and the slab is modelled using Coulomb's law considering $\mu = 0.1$.

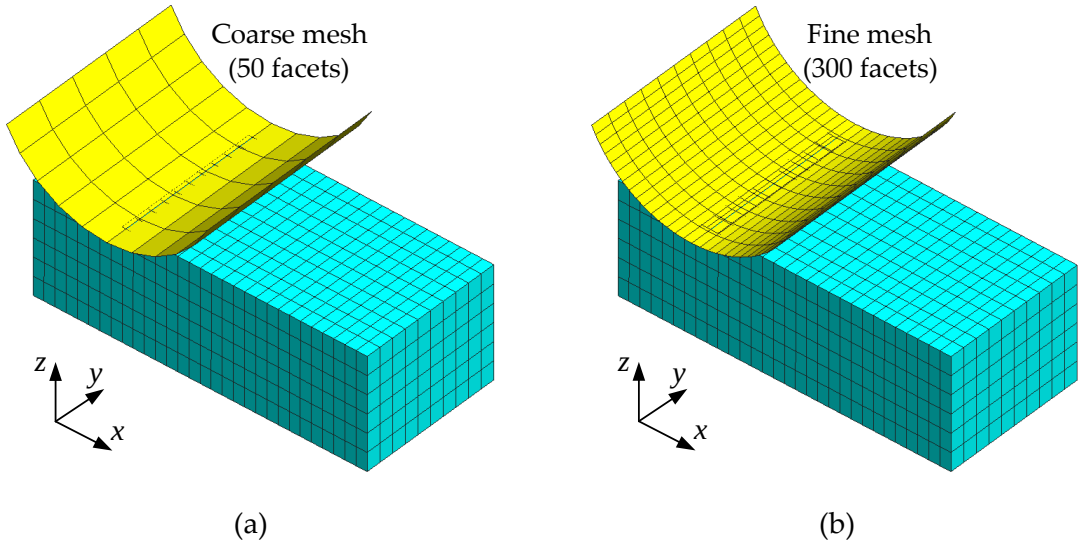


Figure 6.4. Definition of the ironing problem with the rigid cylindrical die described by: (a) coarse mesh; (b) fine mesh.

The cylindrical die, with the centre initially located 2.5 mm from the left end of the slab, travels –1.0 mm in the vertical z-direction and then 4 mm in the horizontal x-direction, while the slab is fixed on the bottom surface. The vertical displacement of the die in the z-direction is performed in 10 equal increments, while the horizontal displacement in the x-direction is applied in 100 increments. Concerning the rigid cylindrical die, two distinct structured surface meshes composed by quadrilateral finite elements are adopted, a coarse mesh and a fine mesh, as shown in Figure 6.4 (a) and (b), respectively. The coarse mesh of the die surface is generated using 10 finite elements in the circumferential direction, while the fine mesh uses 30 elements to describe such direction. Although the discretization in the axial direction does not influence the geometrical description accuracy, it was selected in order to avoid much distorted finite elements. Since the contact surface of the cylindrical die is a cylinder of radius 3 mm, the accuracy in the surface description can be evaluated from the detailed analysis performed in Section 4.3.2. The normalized arc length of the coarse and fine meshes are 0.224 and 0.075, respectively. Therefore, the maximum radial error is given in Figure 4.13 (a) for both faceted and smoothed descriptions of the rigid contact surface. Concerning the piecewise bilinear representation of the die surface, the maximum radial error of the coarse and fine meshes are 0.63% and 0.07%, respectively. On the other hand, the smoothing of contact surface with Nagata patches leads to a maximum value of radial error inferior to 0.002%, in both meshes. The maximum normal vector error is given in Figure 4.13 (b) for both surface description methods. Concerning the faceted representation of the die surface, the maximum value of error is 6.4° for the coarse mesh and 2.1° in the fine mesh. The application of the surface smoothing method leads to a maximum error value always inferior to 0.02°.

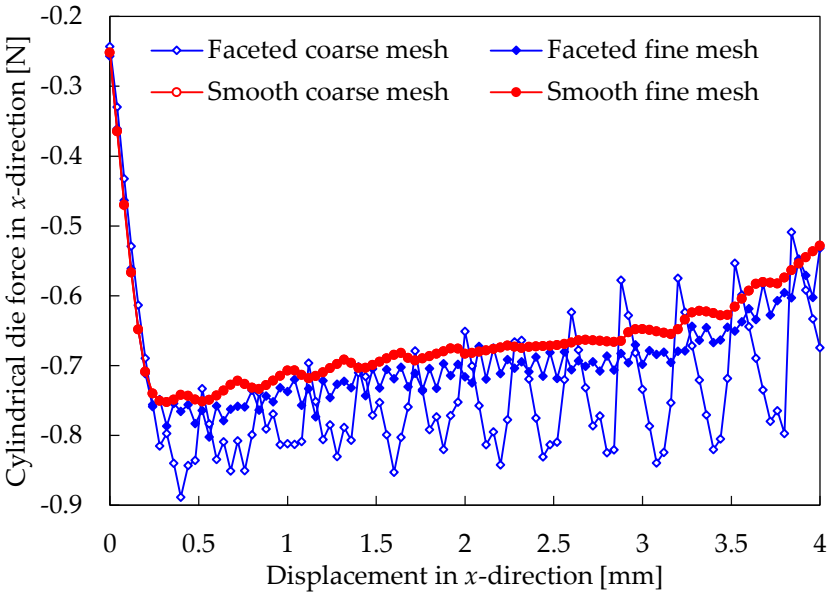
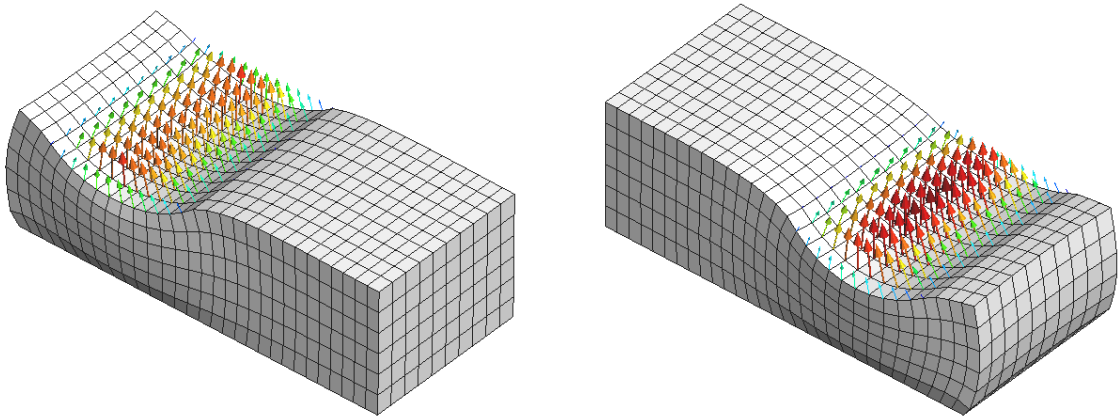


Figure 6.5. Influence of the surface smoothing method in the cylindrical die force as function of its horizontal displacement.

The cylindrical die force evolution with its displacement in the x-direction is presented in Figure 6.5, for both meshes of the rigid cylindrical die shown in Figure 6.4. The effect of the die surface description method (faceted or smoothed) is analysed by means of the force evolution. Concerning the description of the master rigid surface using the traditional piecewise bilinear representation, the mesh refinement reduces the chatter effect in the die force, because both the radial and the normal vector errors are substantially reduced (Figure 6.5). However, the chatter effect is not completely eliminated since the normal vector error is still relevant in the fine mesh (2.1°), leading to the overestimation of the predicted contact force amplitude. On the other hand, the application of the surface smoothing method to the meshes (coarse and fine) yields important improvements in the die force evolution, particularly the elimination of the chatter effect. Indeed, the application of the smoothing operation in both meshes lead exactly to the same die force evolution, as shown in Figure 6.5. The slight periodic oscillations observed in the die force evolution obtained with the smoothing method are related with the deformable body discretization.



(a) (b)

Figure 6.6. Nodal contact forces distribution in the deformed configuration of the slab (magnitude denoted by arrow size and colour) using the die described by Nagata patches: (a) end of vertical displacement; (b) end of horizontal displacement.

The deformed configuration of the slab is shown in Figure 6.6 at the end of each stage (vertical displacement followed by horizontal movement of the die), predicted by the model using Nagata patches in the die description. The deformable slab expands transversely with the sliding progresses such that slab edges slide in the radial direction of the cylindrical die. The nodal contact forces are depicted in Figure 6.6 for the same time instants. The total vertical force is approximately the same in both instants, which is the principal component of the nodal forces. Note that the value of the nodal contact forces is globally higher at the end of the ironing process because the number of slave nodes effectively in contact is smaller.

The computational performance of the ironing problem performed with both surface description methods (faceted and the smoothed) is presented in Table 6.1. The number of iterations necessary to perform the vertical displacement of the die is exactly the same for all models, because the sliding of the slave nodes over the rigid die surface is small. Although the number of increments is the same in all numerical simulations, the movement of the die in the x-direction involves more iterations when the faceted surface description is adopted. This is associated with the poorer geometric description of the rigid die, both in terms of shape and normal vector distribution. Indeed, for the faceted description, the required number of iterations is higher in the fine mesh than in the coarse one, despite the inferior value of the maximum radial error attained by the fine mesh, as well as the normal vector error. This can be explained by the fact that the fine mesh combines a substantial discontinuity (2.1°) of the normal direction between facets with a higher number of facets, which the slave nodes need to cross (higher number of discontinuities). On the other hand, the smoothing of the rigid die surface with Nagata patches leads to a numerical simulation independent of the die discretization, both in terms of computational time and number of iterations, as shown in Table 6.1. Moreover, the computational time is lower using the smoothed description of the rigid contact surface because the convergence problems related with the discontinuity of the surface normal vector field are eliminated.

Table 6.1. Computational performance of the ironing problem for two distinct finite element meshes of the die surface.

	Coarse mesh		Fine mesh	
	Faceted	Smoothed	Faceted	Smoothed
Nº iterations (vertical disp.)	56	56	56	56
Nº iterations (horizontal disp.)	629	510	712	510
Computational time [s]	46	38	52	38

6.1.3. Reverse deep drawing of a cylindrical cup

The third example considered is the reverse deep drawing process of a cylindrical cup, proposed as benchmark at the Numisheet 1999 conference [Gelin 99], which is schematically presented in Figure 6.7. The circular blank has 170 mm diameter and 0.98 mm of initial thickness. The main dimensions of the forming tools for both drawing stages are given in Table 6.2. The total punch displacement is 55 mm in the first stage and 85 mm for the second stage. The sheet material considered in this study is the mild steel DDQ, characterized by the elastic and plastic properties listed in Table 6.3. The material follows Hooke's law in the elastic region, while in the plastic domain the hardening is described by

Swift law. The plastic behaviour of the steel blank is described by Hill'48 quadratic yield criterion using the parameters presented in Table 6.3. The friction coefficient between the sheet and the rigid tools is taken from the benchmark specifications as $\mu = 0.15$.

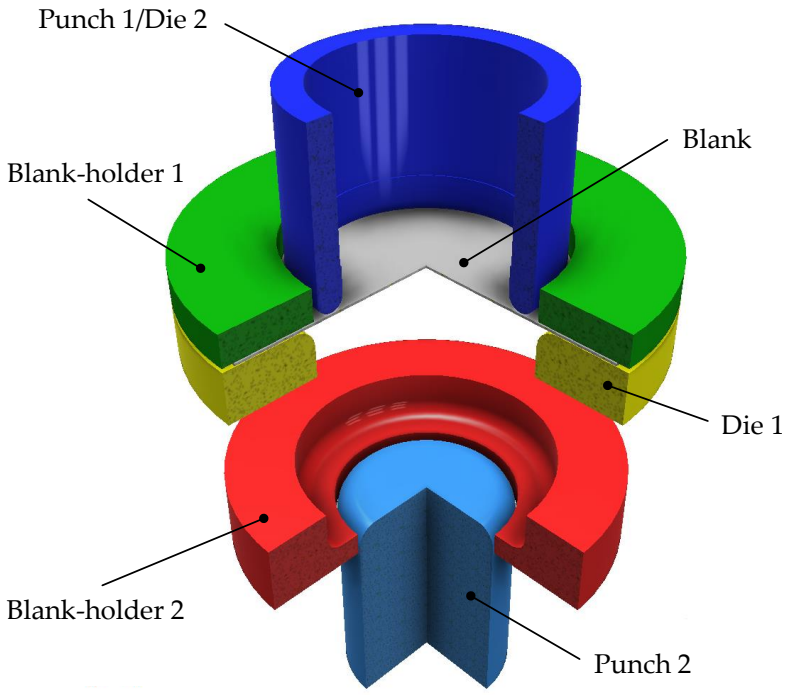


Figure 6.7. Scheme of the forming tools used in the reverse deep drawing of a cylindrical cup, including the blank properly positioned.

Table 6.2. Main dimensions of the tools for both forming stages (mm).

Forming tool		1st stage	2nd stage
Die	diameter	104.5	78.0
	radius	8.0	5.5
	height	21.0	16.0
Punch	diameter	100.0	73.4
	radius	5.5	8.5
Blank-holder	opening diameter	104.5	75.0
	radius	-	7.0
	height	-	20.0

Table 6.3. Elastic and plastic material properties of the DDQ steel used in the reverse deep drawing problem.

Elastic properties		Swift hardening law			Hill'48 yield criterion			
E [GPa]	ν	σ_0 [MPa]	K [MPa]	n	F	G	H	N
210	0.3	172.0	568.3	0.233	0.314	0.366	0.634	1.176

Due to geometric and material symmetry conditions, only one quarter of the model is simulated, as shown in Figure 6.8. The blank is discretized with 8-node hexahedral solid finite elements, using 2 layers through the thickness, making a total of 15,408 elements. The gap between the die and the blank-holder is held fixed in both stages, which is set equal to 1.13 mm in the first stage and 1.4 mm in the second stage in order to draw a cylindrical cup without wrinkles [Thuillier 02].

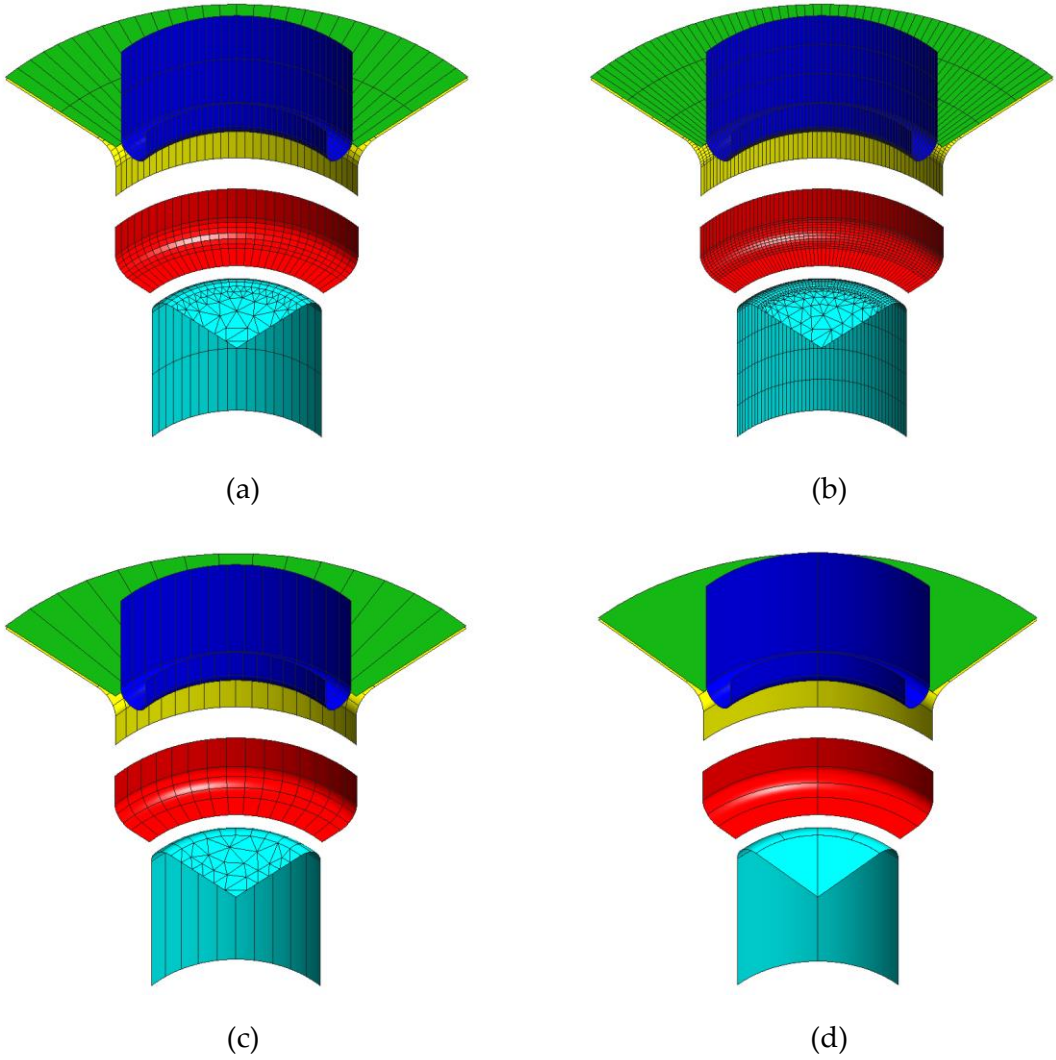


Figure 6.8. Description of the forming tools for the reverse deep drawing process using: (a) faceted coarse mesh; (b) faceted fine mesh; (c) Nagata patches; (d) Bézier patches.

In the present example, the surface of the forming tools is described using three different methods, as shown in Figure 6.8. Two models for the faceted surface description are created, using 6 and 12 bilinear facets to define each circular arc in the radial direction for the coarse and fine meshes (see Figure 6.8 (a) and (b)), respectively. The model composed by Nagata patches uses only 3 patches to describe each circular arc in the radial direction, as illustrated in Figure 6.8 (c). The last approach uses Bézier patches in the surface description, which are obtained with the aid of a specific CAD package [Oliveira 08]. Note that the required finite element meshes of the tool surfaces are generated from this CAD model, which is composed by 42 Bézier patches. Concerning the piecewise bilinear representation of the tool surface, the coarse mesh is defined by 1,418 bilinear facets, while the fine mesh comprises 4,672 facets. The surface discretization of the tools smoothed with Nagata patches is composed by 386 Nagata patches, i.e. about 10 times more than the total number of Bézier patches.

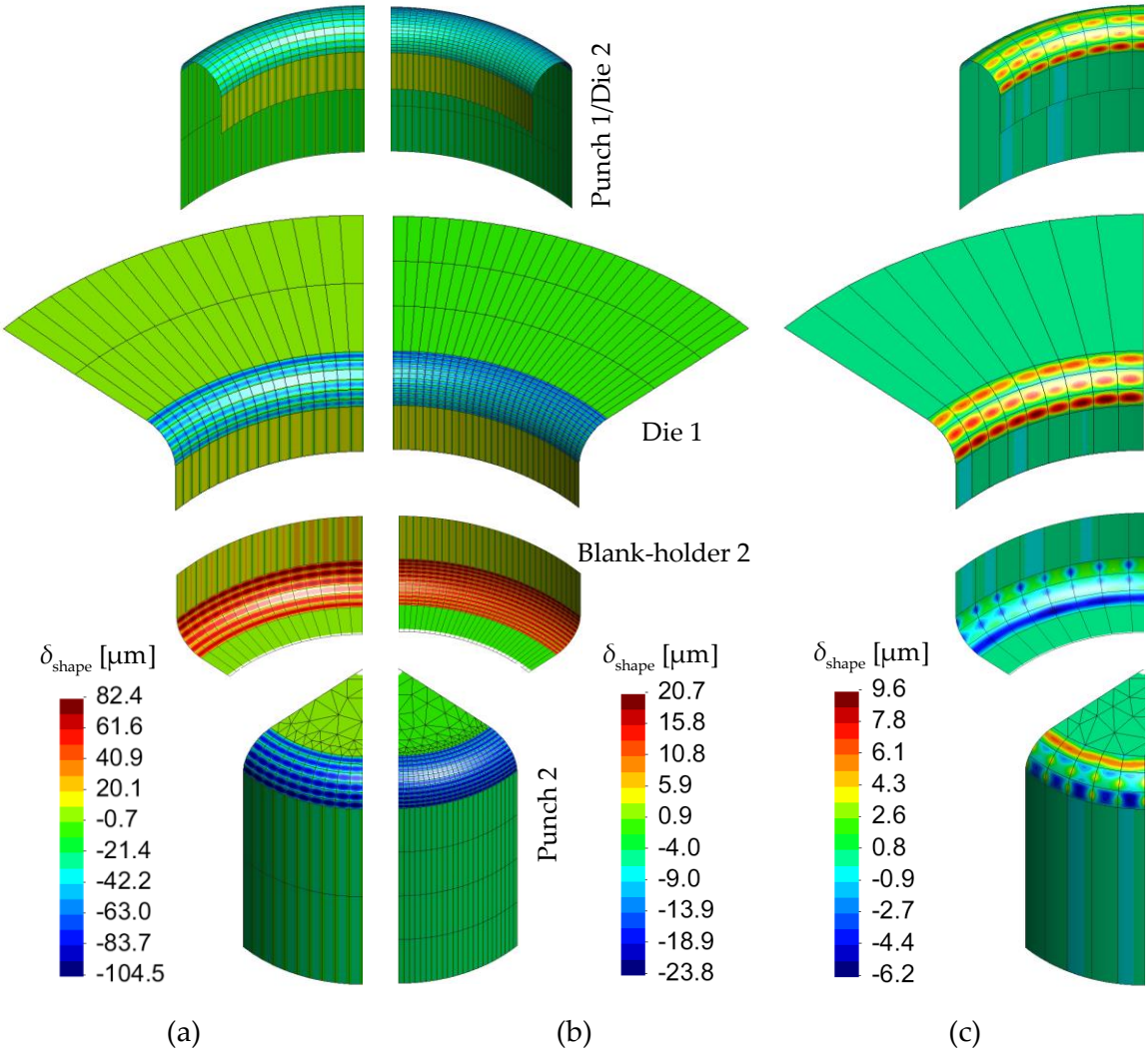


Figure 6.9. Shape error distribution on the tool surfaces described by: (a) coarse mesh of facets; (b) fine mesh of facets; (c) Nagata patches.

The shape error, previously defined in (4.26), is selected to assess the geometrical accuracy of the tool surfaces models, comparing all surface description methods, based on the discretization of the tool surfaces. The shape error distribution on the tool surfaces described by Nagata patches is presented in Figure 6.9 (c). The maximum positive error appears in the hyperbolic surface of die 1, while the maximum negative error occurs in elliptic surface of the punch 2. Nevertheless, the shape error ranges from −6.2 to 9.6 μm, thus the maximum value attained is always inferior to 10 μm. The comparison of the shape error distribution obtained with the faceted tool surface description for the coarse and fine mesh is presented in Figure 6.9 (a) and (b), respectively. Although the finer tool model composed by facets uses more than 10 times the number of Nagata patches model to describe the contact surfaces, the maximum error value reached is approximately twice the one attained using Nagata patches. Moreover, the sudden variation of the shape error within each facet, which can be observed in Figure 6.9, combined with the small size of the facets yields a higher error in the surface normal vector field.

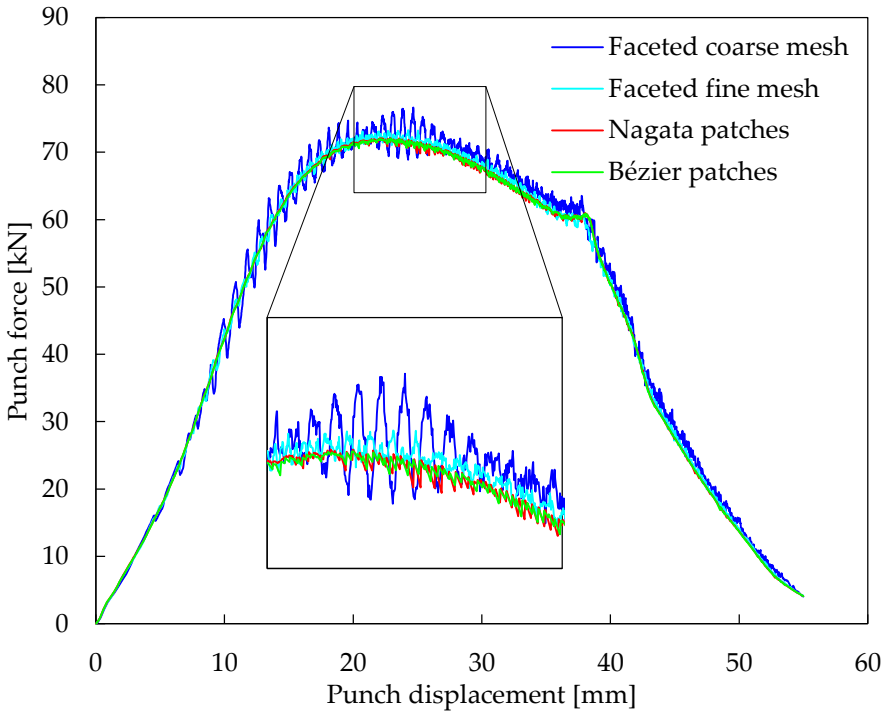


Figure 6.10. Punch force evolution during the 1st stage for different tool surface description methods and zoom of the chatter effect in the force.

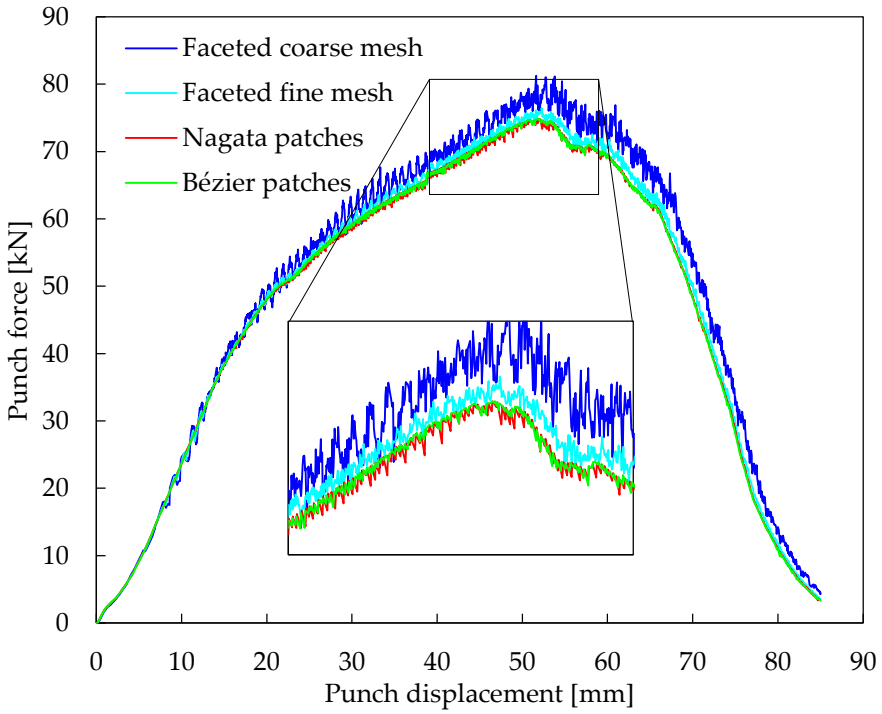
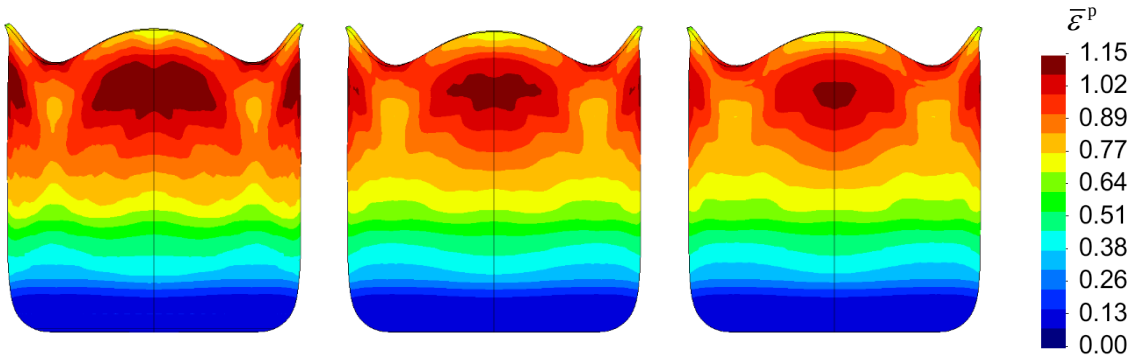


Figure 6.11. Punch force evolution during the 2nd stage for different tool surface description methods and zoom of the chatter effect in the force.

The comparison of the predicted punch force evolution using the different tool surface description methods is presented in Figure 6.10 and Figure 6.11, for the first and the second forming stage, respectively. The effect of the tool surface accuracy can be observed in both forming stages. Nevertheless, it is more emphasised during the second stage due to the lower value of the die radius. Regarding the faceted tool surface description, the increase of the number of facets reduces the oscillations in the punch force evolution. However, only the surface smoothing method with Nagata patches and the tools defined by Bézier patches lead to a force evolution with reduced oscillations, as can be observed in the zoom view included in Figure 6.10 and Figure 6.11.

The equivalent plastic strain distribution at the end of the second stage is presented in Figure 6.12, for each tool surface description method. Since the adoption of Bézier and Nagata patches in the tool surface description leads to identical results, only the distribution obtained with Nagata patches is presented in Figure 6.12 (c). However, the equivalent plastic strain predicted using the tool surfaces defined by facets is considerable different from the one obtained with smooth patches, particularly when adopting the coarser mesh. The maximum value of equivalent plastic strain is higher in the faceted coarse mesh, as well as the height of the cylindrical cup. This is related with the artificial roughness introduced in the contact surface by the faceted description of the tool surfaces, i.e. sudden changes in the surface normal between adjacent master facets. The refinement of the finite element mesh used in the description of the tool surfaces improves the accuracy of the numerical

results, converging to the ones obtained with Nagata patches or Bézier patches (see Figure 6.12). However, the fine mesh selected for this example is composed by 4,672 bilinear facets, while the tool model defined by Nagata patches presents only 386 finite elements, replaced by Nagata patches in the surface smoothing procedure. Therefore, the model composed by Nagata patches has a number of elements approximately 12 times lower, while the provided numerical results are more accurate, as shown in Figure 6.12.



(a) (b) (c)

Figure 6.12. Equivalent plastic strain distribution at the end of 2nd stage using tool surfaces described by: (a) faceted coarse mesh; (b) faceted fine mesh; (c) Nagata patches.

The evolution of the number of slave nodes in contact with the die surface for each tool surface description method is presented in Figure 6.13 and Figure 6.14, for the first and second forming stages, respectively. The higher number of nodes in contact predicted by the tool surfaces modelled by Bézier patches is associated with the excellent geometrical accuracy in the tool surfaces definition (at least C^1 continuity). On the other hand, the piecewise bilinear representation of the tool surfaces provides a lower number of nodes in contact, particularly in the first stage (see Figure 6.13), due to the artificial roughness induced in the contact surface. In fact, the surface mesh refinement is more effective in the second stage than in the first one, as can be seen in Figure 6.14. This is a consequence of the slight mesh distortion (ears generation) motivated by the material plastic anisotropy. In the second stage the mesh distortion is more pronounced, which allows to improve the contact conditions definition through a smooth transition of the nodes in contact.

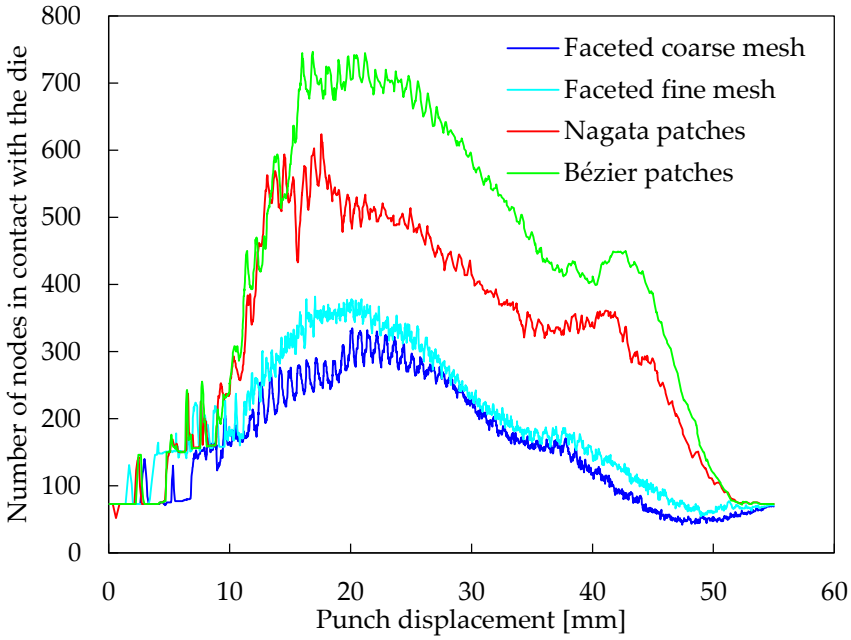


Figure 6.13. Evolution of the number of slave nodes in contact with the die during the 1st stage for different tool surface description methods.

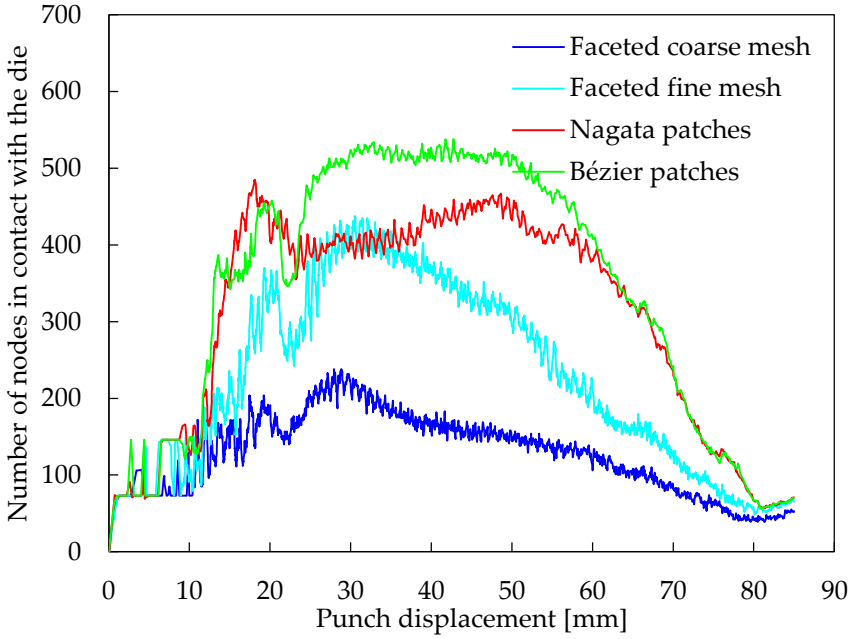


Figure 6.14. Evolution of the number of slave nodes in contact with the die during the 2nd stage for different tool surface description methods.

Table 6.4. Computational performance of the reverse deep drawing problem using different tool surface description methods.

	Faceted			
	Coarse mesh	Fine mesh	Nagata	Bézier
Nº increments	4,248	4,103	1,788	1,917
Nº iterations	25,939	24,103	12,460	12,925
Computational time [hours]	6.41	6.01	3.28	3.62

The computational performance of the numerical simulations is evaluated in this study through the number of increments, total number of equilibrium iterations and the computational time, which are compared in Table 6.4. The required number of increments decreases slightly with the mesh refinement of the faceted tool surface description. This is directly related with the error in the normal vector of the discretized tool surfaces, which decreases only linearly with the mesh refinement, as shown Figure 4.13 (b). On the other hand, the smooth tool surface descriptions (Bézier and Nagata patches) require approximately half the number of increments to complete the simulation, as shown in Table 6.4. Identical behaviour is observed for the total number of equilibrium iterations. The adoption of different tool surface description methods requires a dissimilar number of increments because the increment size is automatically reduced by means of the r_{min} strategy [Oliveira 04], used in order to improve the convergence during the corrector step. The reduction is more emphasised for the faceted tool surface description method (coarse and fine mesh), due to the sudden changes in the surface normal vector direction. Concerning the representation of the contact surfaces using Bézier or Nagata patches, the evolution of the increment size is identical.

The computational time is mainly dictated by the number of increments required to perform the simulation. Therefore, the computational time required to carry out the numerical simulation using the faceted tool surface description is approximately twice the one required when a smooth surface description is adopted, as shown in Table 6.4. Although the total number of Nagata patches used in tool surface description is higher than the number of Bézier patches (see Figure 6.8), the computational time needed to complete the simulation is slightly lower when applying Nagata patches. The same trend is observed for the total number iterations and increments (average 7 iterations per increment). This behaviour can be associated with the high interpolation degree used in some Bézier patches, which can create instabilities in the local contact search algorithm (normal projection).

6.1.4. Automotive underbody cross member

The industrial example selected consists in the sheet metal forming process of an automotive underbody cross member panel, proposed as benchmark at the Numisheet 2005

conference [Wu 05], which is schematically presented in Figure 6.15. The numerical simulation of the forming process consists of three operations: (i) forming; (ii) trimming and (iii) springback. Nevertheless, only the first operation is performed in this study since the forming stage dictates the success of the subsequent operations and it is largely affected by the contact surface description. The forming process involves three rigid tools and can be decomposed into two phases: blank-holder clamping until 1068 kN and subsequent punch displacement until 100 mm, while maintaining the clamping force constant. The geometry of the tools includes physical drawbeads to control material flow, which strongly increase its geometric complexity, as highlighted in the detail included in Figure 6.15. The selected material for the sheet is a DP600 dual-phase steel with an initial thickness of 1.62 mm, as proposed by the benchmark committee [Shi 05]. Regarding the material mechanical behaviour, the isotropic work hardening is assumed to be described by the Swift law, while the orthotropic behaviour is described by Hill's 1948 yield criterion. The material parameters were obtained from the experimental data provided by the benchmark committee [Shi 05] being presented in Table 6.5. The friction coefficient between the sheet and the tools is taken from the benchmark instructions as $\mu = 0.12$.

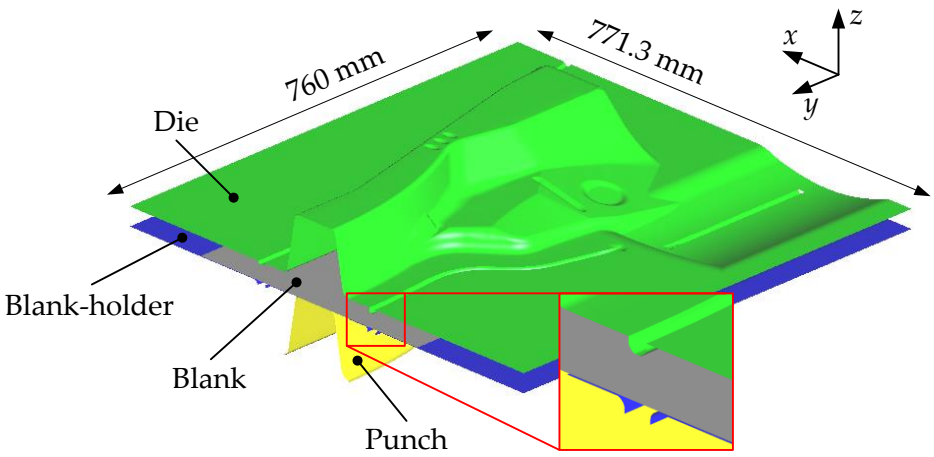


Figure 6.15. Scheme of the forming tools used in the automotive underbody cross member panel and zoom of drawbeads geometry.

Table 6.5. Elastic and plastic material properties of the dual-phase steel DP600 used in the automotive underbody cross member.

Elastic properties		Swift hardening law			Hill'48 yield criterion			
E [GPa]	ν	σ_0 [MPa]	K [MPa]	n	F	G	H	N
206	0.3	393.0	1038.8	0.162	0.460	0.576	0.424	1.446

The large dimensions of the component and the complex tool geometry dictate the use of a very large number of finite elements in the description of the sheet, in order to describe accurately the process and particularly the frictional contact conditions. The two main dimensions of the tools are indicated in Figure 6.15, while the radius of the drawbeads is approximately 4 mm. Due to geometric and material symmetry conditions at $y=0$, only half model is considered in the simulation, as shown in Figure 6.15. The blank geometry (half model), is shown in Figure 6.16 and can be defined by its width of 620 mm in x-direction and length of 607 mm in y-direction. It is discretized with 54,000 solid finite elements (8-node hexahedral), using 2 layers of elements through the thickness. The average finite element size in the x-direction is approximately 3 mm, while in y-direction is about 4 mm.

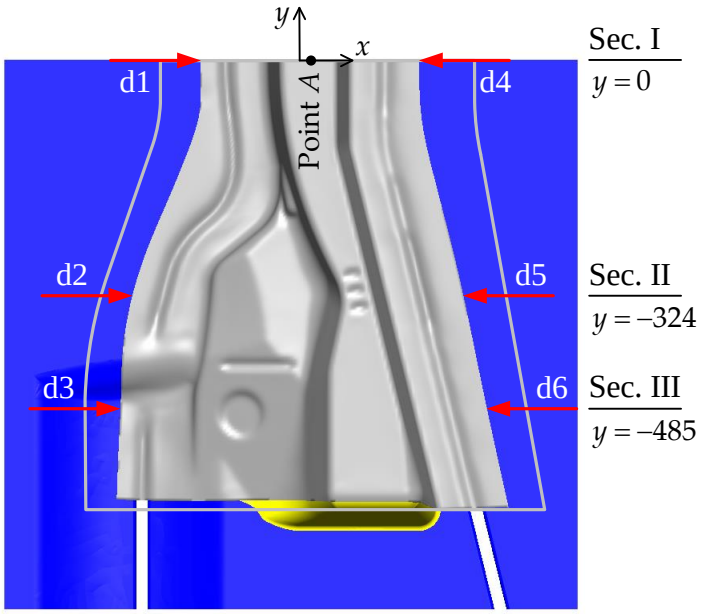
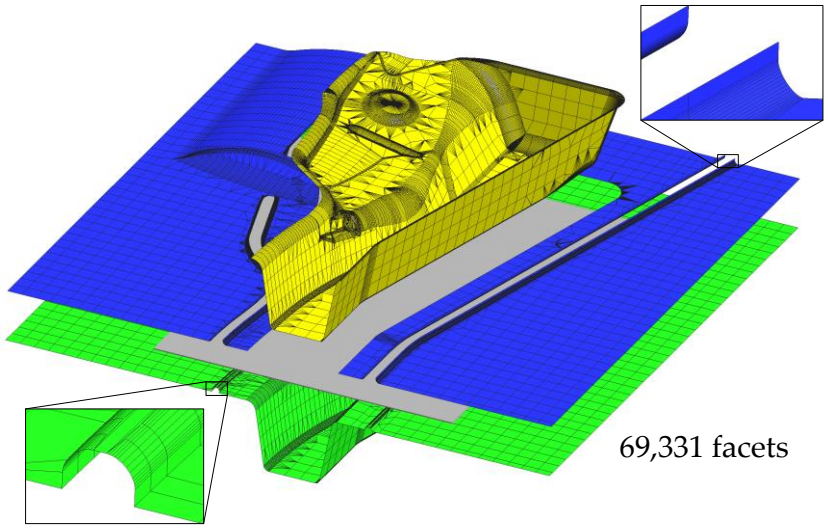


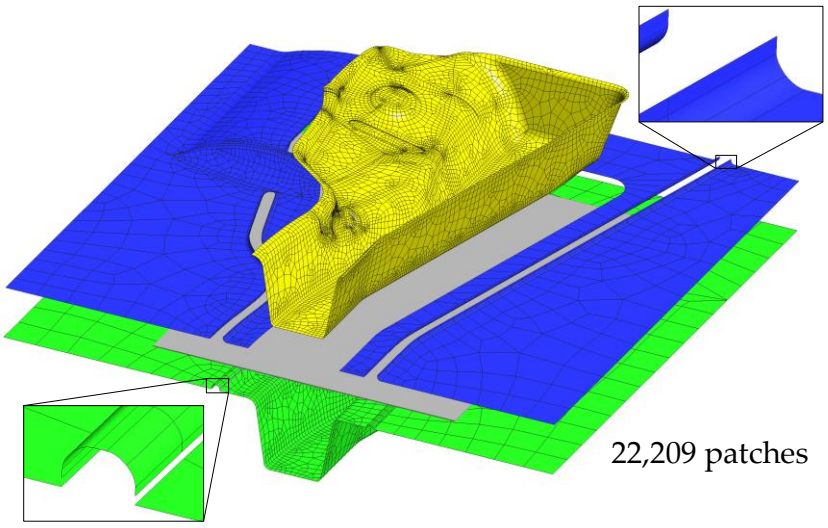
Figure 6.16. Sections for the blank draw-in and thickness measurement including the identification of point A.

Concerning the tool surface description, the information about the tools geometry was provided by the conference committee both in a Nastran mesh file and in an IGES format file. The available Nastran mesh is composed by a mixture of triangular and quadrilateral finite elements, as shown in Figure 6.17 (a). This discretization was generated to be used directly in the numerical simulation, without any type of additional surface smoothing procedure. On the other hand, the IGES file of each tool is used in this study to create the discretized tool surfaces, which are posteriorly smoothed with Nagata patches and then applied in the numerical analysis. Figure 6.17 (b) presents the model of the forming tools described by Nagata patches. Note that the total number of finite elements is approximately

three times higher in the Nastran mesh than the number of Nagata patches applied in the description of all tool surfaces. Besides, the faceted tool surface description is characterized by a mesh refinement in all curved surfaces, while flat surfaces are described by large finite elements (Figure 6.17 (a)). Since the smoothing method improves significantly the geometric accuracy of the discretized tool surfaces, the comparison is performed between a refined faceted tool mesh and a coarse smoothed tool mesh, with a similar level of geometric accuracy (shape error inferior to 0.05 mm).



(a)



(b)

Figure 6.17. Description of the forming tools for the automotive underbody cross member with rigid surfaces described by: (a) bilinear facets; (b) Nagata patches.

Although the proposed surface smoothing method allows to apply simultaneously both triangular and quadrilateral patches, the latter type of patches was selected to describe all tool surfaces. Indeed, the unstructured mesh topology was adopted to discretize the tool surfaces due to its easy and fast generation. This discretization was generated automatically using the pre-processor GiD [CIMNE 13], employing the chordal error option to assign the finite element size. The required parameter corresponds to the maximum relative chordal error (chordal error of an element divided by its characteristic size) in the whole model, which was taken as 7% to create at least two elements in the discretization of each quarter of circle arc (cf. Figure 4.13 (a) with a normalized arc length of 0.78). Additionally, the drawbeads geometry was described using a structured mesh in order to improve the accuracy attained in the subsequent smoothing procedure. The drawbeads definition is a very important issue in the treatment of frictional contact problems, due to the high contact forces involved and the large slip between the sheet and the tools. Typically the drawbeads geometry is approximately cylindrical, which allows the easy assessment of the geometric error introduced by the surfaces discretization procedure. In the present example, each quarter of the cylinder defining the drawbeads is described by 2 Nagata patches in the circumferential direction (see detail of Figure 6.17 (b)), leading to a maximum radial error (positive) inferior to 0.32%. On the other hand, the tool model defined using faceted finite elements contains 12 and 16 finite elements in the circumferential direction of each quarter of cylinder that composes the drawbeads of the die and blank-holder, respectively (see detail of Figure 6.17 (a)). This surfaces discretization leads to a maximum radial error (negative) of approximately 0.2% and 0.12% in the description of the drawbeads geometry inserted in the die and blank-holder, respectively (cf. Figure 4.13 (a), linear interpolation with a normalize arc length of 0.13 and 0.098). Therefore, the drawbeads model composed by faceted finite elements is slightly more accurate in terms of shape than the geometry defined by Nagata patches. However, the linear finite element discretization leads to sudden changes in the surface normal field, with a maximum normal vector error in the range 5.5 and 7.5°, whereas a *quasi*-continuous surface normal field is attained using Nagata patches in the surface description.

The first phase of the forming process comprises the blank-holder clamping, until attaining a force of 1068 kN. In order to correctly position the punch to the next phase, the punch and the blank-holder move together during this phase. The blank-holder force evolution with its displacement for both the smooth and the faceted tool surfaces descriptions is presented in Figure 6.18. Since the bending effects play a dominant role during the blank-holder closure, the blank-holder force evolution is nearly independent of the strategy adopted to describe the forming tools, as shown in the figure. Indeed, the sliding between the sheet and the tools is very small during this phase, which yields similar tool force evolutions. Until approximately 23 mm of blank-holder displacement the contact

between the sheet and the tool surfaces is established locally, being the effective clamping accomplished in the last 5 mm of displacement.

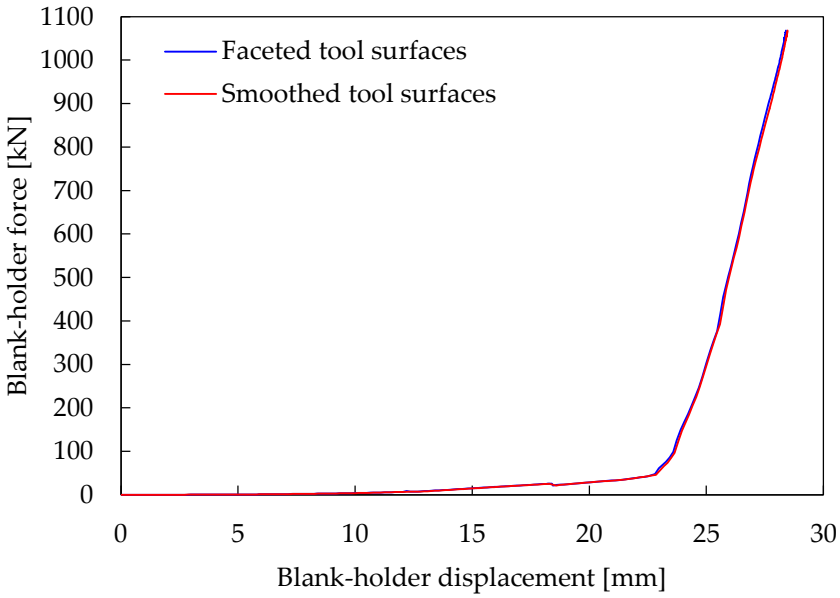


Figure 6.18. Blank-holder force evolution obtained with faceted and smoothed tool surface description methods.

The second phase of the forming process involves the punch displacement, keeping the holding force of the blank-holder constant. The punch force evolution obtained for both strategies, adopted to describe the tool surfaces, is presented in Figure 6.19. The occurrence of sliding between the sheet and the drawbeads during the punch movement becomes important in this second phase. Nevertheless, negligible differences are observed in the punch force evolution, since the tool geometry accuracy is similar for both models. The exponential growth of the punch force observed at the end of the forming process is related with the closing of the sheet between the punch and the die. The oscillations observed in the force evolution for both tool description strategies are due to numerical instabilities related to the ratio between the blank mesh element size and the curvature radius of the drawbeads, precluding the accurate definition of the frictional contact. Since the minimum radius in the tool geometry is about 4 mm, to avoid such problems the blank mesh size should be lower than 1 mm [Taylor 95]. However, this small element size leads to impracticable computational times. Thus, the blank mesh size used in this study about 3 mm to overcome this limitation. The maximum experimental force value attained by the die is 6,226 kN (sum of the punch and blank-holder forces) [Buranathiti 05a], which is close to the predicted numerical value. Nevertheless, this parameter is inadequate to verify the accuracy of the numerical results, due to the high force gradient that occurs at the end of the forming process (see Figure 6.19).

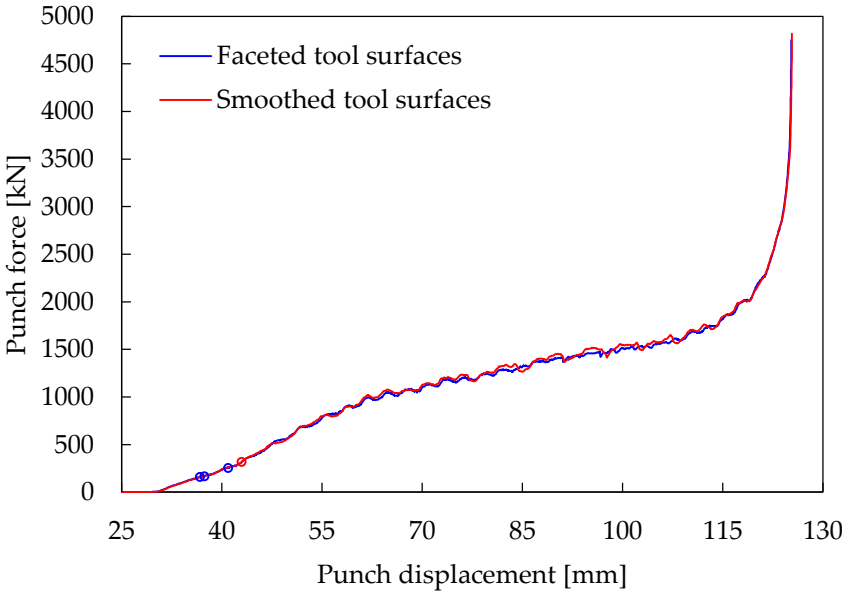


Figure 6.19. Punch force evolution obtained with faceted and smoothed tool surface description methods.

The flow stress distribution on the cross member panel at the end of the forming operation is presented in Figure 6.20, for both tool surface description methods (faceted and smoothed). As a consequence of the material flow between the drawbeads, the higher values of flow stress arise in the regions around them, highlighting the bending–unbending effect induced by the drawbeads in the sheet. Both tool description methods lead to a similar flow stress distribution, since the tool geometry accuracy is approximately the same in both models.

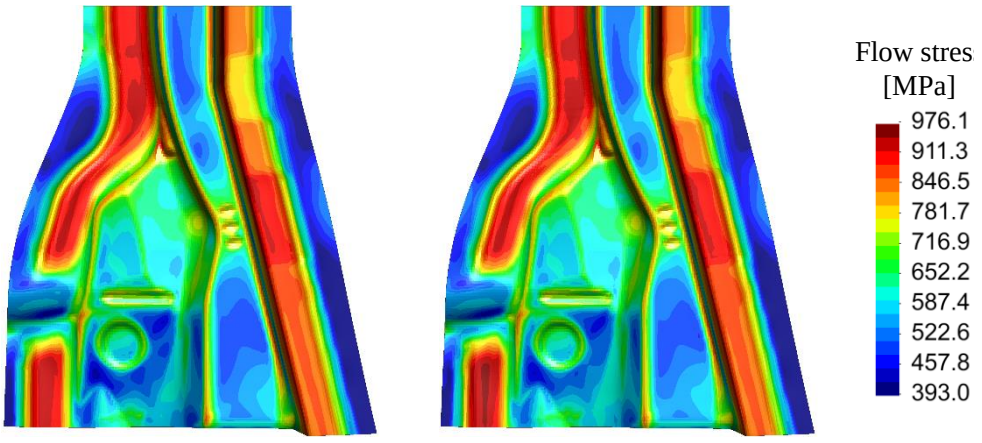


Figure 6.20. Flow stress distribution plotted in the fully deformed configuration using: (a) faceted tool surfaces; (b) smoothed tool surfaces.

The two parameters selected to evaluate the accuracy of the numerical results are the ones proposed in the benchmark specifications [Wu 05]. Figure 6.21 presents the comparison between experimental and numerical draw-in amounts, at the six positions predefined in Figure 6.16. Globally, the numerical prediction of the draw-in amount underestimates the experimental values, mainly in the left side of component (positions d1, d2 and d3). The maximum difference between experimental and numerical results (approximately 20%) occurs at position d2. In fact, the lower draw-in amount predicted is related with the artificial roughness introduced by the coarse mesh adopted for the sheet, which hampers the material flow through the drawbeads. Since the accuracy of the tool geometry is identical in both surface tool models, the draw-in amount predicted in the numerical simulation is also similar (Figure 6.21), which is slightly higher for the faceted tool description due to the negative radial error in the drawbeads regions. In fact, the positive radial error observed in these regions for the smoothed surfaces corresponds to a slight decrease of the gap between the die and blank-holder and, hence, inhibits the material flow.

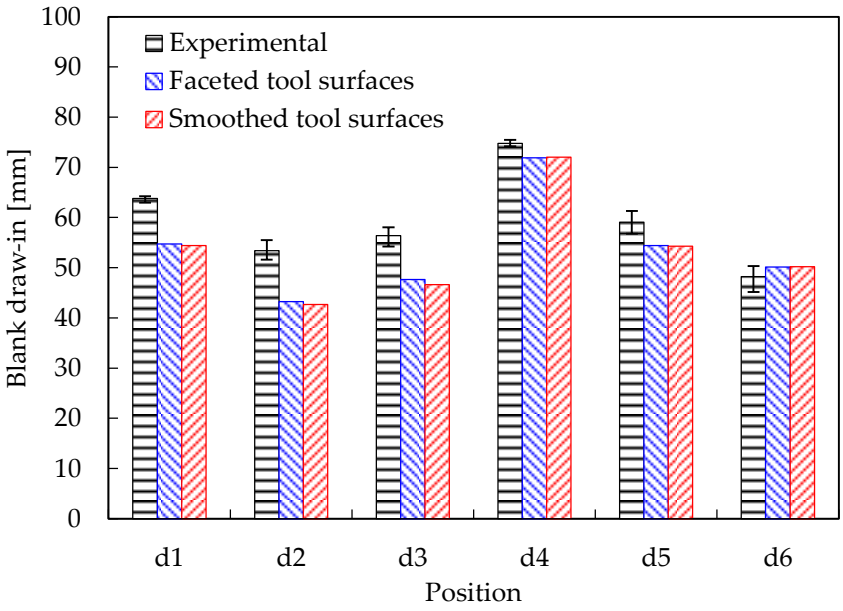


Figure 6.21. Comparison between experimental and numerical blank draw-in amount at specific localizations after forming (identified in Figure 6.16).

The sheet thickness distribution at the symmetry plane (Sec. I in Figure 6.16) after forming is presented in Figure 6.22, where the numerical results are compared with the experimental ones. The experimental tendency of the thickness distribution is well reproduced numerically, despite being underestimated in all points of the section. The positive correlation between the draw-in and the predicted thickness can be verified through the comparison of Figure 6.21 and Figure 6.22. Usually, the accurate prediction of

the draw-in amounts leads to a good prediction in the thickness strains, as referenced by Buranathiti and Cao [Buranathiti 05b]. The difference between the numerical and the experimental thickness distribution is globally inferior to 5%, except in the fillet edge region located at an arc-length distance from point A ($x = 7$ mm) around 33 mm (see Figure 6.22), where the difference is roughly 7%. The curvature radius of the fillet edge is approximately 5 mm in the punch and 6.5 mm in the die, making a strong bending effect in the sheet. Both numerical distributions are identical because of the reasons previously mentioned.

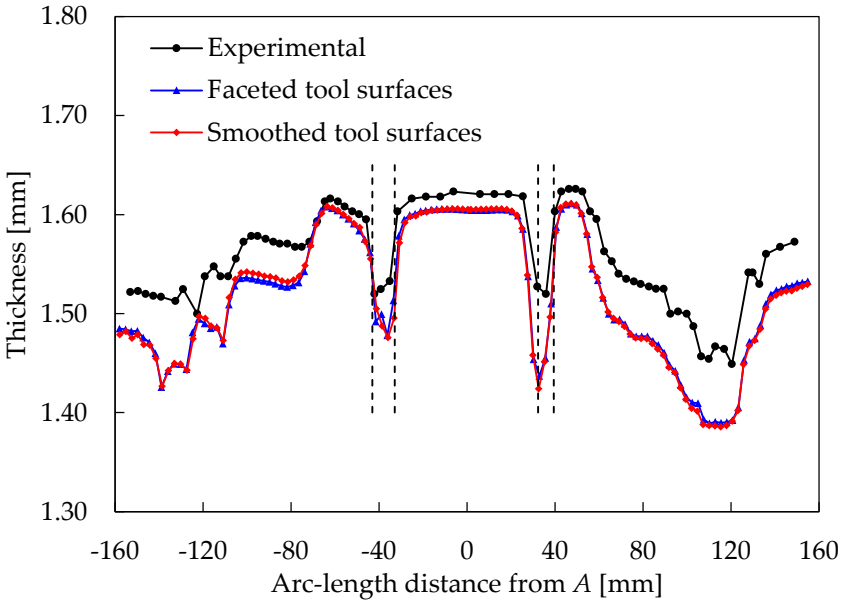


Figure 6.22. Comparison between experimental and numerical thickness distribution at the symmetry plane (Sec. I in Figure 6.16) after forming.

Table 6.6. Computational performance of the automotive underbody cross member using different tool surface description methods.

	Blank-holder clamping		Punch travel	
	Faceted	Smoothed	Faceted	Smoothed
Nº increments	368	352	1,919	1,106
Nº iterations	2,548	2,431	11,266	7,090
NTrial strategy	32	19	1,120	120
Nº divergent increments	0	0	3	1
Computational time [hours]	12.6	7.1	134.2	16.9

The computational performance of the numerical simulation of the automotive underbody cross member is shown in Table 6.6, comparing the faceted and the smoothed tool surface descriptions methods. The results show that both the number of increments

and total number of equilibrium iterations required to complete the blank-holder clamping phase is slightly lower when the smooth surface description is adopted. Nevertheless, the computational time is reduced to almost half using smooth tool surfaces. Therefore, the time consumed in each increment is lower when using the Nagata patches in the tool surface representation, which can be related with the smaller number of patches used to describe the tool surfaces (see Figure 6.17), since the global search procedure becomes computationally less expensive. In fact, the surface mesh provided in Nastran format presents a high ratio between the maximum and minimum finite element side length, leading to the maximum values allowed for the number of master nodes and number of grid divisions, expressed in (3.25) and (3.26), respectively. The *NTrial* strategy, currently implemented in DD3IMP finite element code, allows overcoming several convergence problems by automatically reducing the increment size when convergence is not attained within the maximum allowed number of iterations. During the blank-holder clamping phase all arising convergence problems are solved with the *NTrial* strategy, which automatically reduces the size for 32 increments in the faceted tool surfaces simulation and for 19 increments when using smoothed tool surfaces (see Table 6.6). Since the sliding between the sheet and the rigid tools is very small during this phase, the convergence is attained in all increments without user intervention.

Concerning the drawing phase (100 mm of punch travel), the computation time necessary to carry out this second phase is considerably lower when the smoothed surface description is applied, reducing from about 134 to only 17 hours, as shown in Table 6.6. This is a consequence of the necessary number of increments, which are 1,919 when using the piecewise bilinear surface representation and 1,106 when the proposed surface smoothing method is adopted. The convergence problems arising with the faceted tool model are motived by the surface normal changes, which are drastically reduced with the smoothing method. Hence, the number of increments with its size reduced by the *NTrial* strategy is 1,120 when the faceted tool surface description method is applied, and only 120 if the surfaces are smoothed with Nagata patches, as presented in Table 6.6. Therefore, the reduction of the computational time obtained with the adoption of the proposed surface smoothing method is related with two factors: lower number of increments and global contact search procedure computationally less expensive. Moreover, during the punch travel phase some increments do not attain the convergence within the maximum allowed number of iterations, which are indicated by the hollow circles in Figure 6.19. The divergence of the iterative procedure occurs only in one increment using the presented surface smoothing method, while the faceted tool description leads to three divergent increments. The occurrence of divergence requires the user intervention in order to change the penalty parameter and resume the simulation.

6.2. Contact between deformable bodies

This section contains six numerical examples involving the frictional contact between deformable bodies, including potential self-contact phenomena. The accuracy of the Node-to-Nagata contact elements developed to handle the contact between deformable bodies is firstly evaluated through the Hertz contact problem, involving frictionless contact of two elastic cylinders, for which the analytical solution is known. The second example takes into account the frictional effects at the contact interface and consists in the contact of a cylinder embedded in a bored infinite plane. In this case, the numerical results are compared with the corresponding semi-analytical solution. The third example analysis the contact problem between curved contact surfaces of deformable bodies. The developed surface smoothing procedure is compared with the traditional description of the contact surfaces using bilinear facets, for both frictionless and frictional large sliding. Indeed, the comparison between the developed surface smoothing procedure and the bilinear facets description will be performed for all the following examples, in order to highlight the benefits of its application to different problems. The extrusion of an aluminium billet into a deformable conical die is the fourth example studied, which includes nonlinear material behaviour. The ability of the global contact search algorithm, specifically developed for self-contact problems, is validated using the post-buckling of a thin walled tube. In this case, the results are compared with numerical solutions obtained by other authors. The last example shows the deep drawing of a cylindrical cup considering deformable tools. The numerical and experimental results are compared, showing the importance of the developed algorithms in terms of industrial applications.

6.2.1. Contact between two elastic cylinders

This example deals with the frictionless $(\mu = 0)$ contact of two elastic cylinders, which are pressed together by means of a load applied on the end of one of the bodies, as illustrated in Figure 6.23. The load is applied along their common symmetry plane, while the end of the other cylinder is fixed. The bodies are assumed to be elastic with identical material properties $(E = 1000\ \text{MPa}, \nu = 0.3)$. The radius r of both cylinders is 10 mm and a unitary length is assumed. The purpose of this problem is to evaluate the accuracy of the Node-to-Nagata contact element for small slip frictionless contact between two deformable bodies, i.e. fully nonlinear geometry. Therefore, two distinct values of applied force are analysed $f = 50\ \text{N}$ and $f = 200\ \text{N}$, in order to evaluate the impact of the assumption of a small contact area in the comparison between numerical and analytical results. The influence on the numerical results of using conforming and non-conforming meshes at the contact interface and of the master-slave choice is also studied.

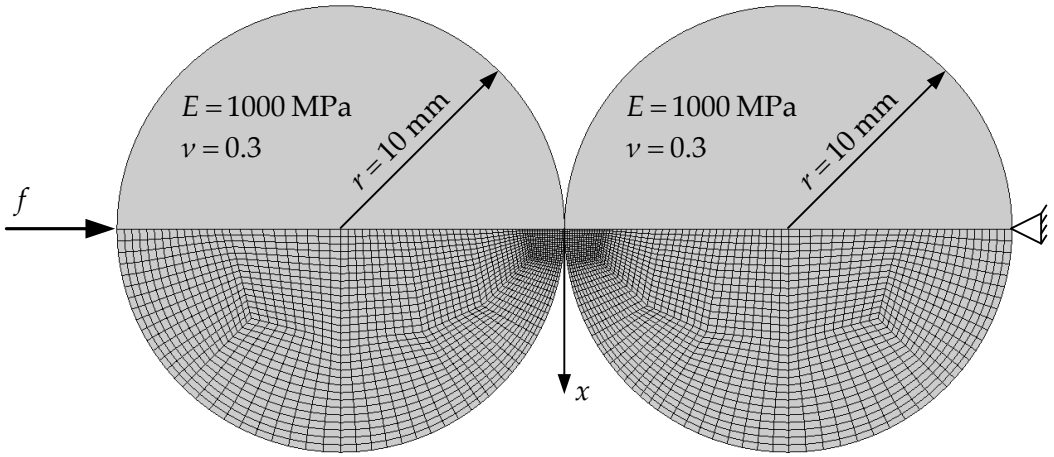


Figure 6.23. Contact between two elastic cylinders, problem definition (top) and finite element mesh (bottom).

The problem is modelled as 3D since this formulation is the main goal of this dissertation, though it is basically a 2D problem. Therefore, the out-of-plane displacements are fixed on both faces, i.e. plane strain condition is imposed. Due to the symmetry conditions of the problem, only one half of each cylinder is modelled. Each cylinder is discretized with 1,680 solid finite elements (8-node hexahedral) using a single layer through the cylinder length, yielding a total of 3,504 nodes. The finite element mesh is refined around the contact region, as depicted in Figure 6.23, allowing to achieve sufficient accuracy on the estimate of the contact area. Moreover, the mesh is constructed symmetrically for each cylinder, such that each slave node coincides with the master node on the opposite contact surface (conforming meshes). The cylinder on which the force is applied is assumed to be the slave body while the other one is assigned as master. The interface between the cylinders is modelled by a set of Node-to-Nagata contact elements. In the present study the external force applied on the top of the cylinder is replaced by an equivalent prescribed displacement (Dirichlet boundary conditions), which is applied in a single increment. Note that the displacement values are strongly dependent of the finite element mesh adopted near the zone where the load is applied.

The analytical solution of this problem can be obtained by the Hertz theory [Hertz 82]. The analytical contact pressure distribution along the perpendicular direction to the cylinder axis is given by:

$$p_{\mathrm{n}}(x) = \frac{2f}{\pi b^2}\sqrt{(b^2 - x^2)} \quad \text{for} \quad x < b, \tag{6.1}$$

with:

$$b = 2\sqrt{\frac{fr(1-\nu^2)}{\pi E}}, \tag{6.2}$$

where b is the half-width of the rectangular contact surface and x represents the coordinate of the current point on the contact zone, as depicted in Figure 6.23. The contact pressure is constant along the cylinder axis, while the maximal pressure value is always located on the symmetry axis. Note that the solution expressed by (6.1) is valid only in the case where the contact width is small compared to the cylinder radius.

The comparison between the numerical and the analytical solutions for the normal contact pressure distribution on the contact surface is presented in Figure 6.24. The frictionless analytical pressure (solid curve) has the shape of a half-ellipse with respect to the distance from the symmetry plane [Kikuchi 88]. For the two selected applied external forces, the numerical results are in very good agreement with the analytical solution, as shown in Figure 6.24. The contact pressure is slightly overestimated near the symmetry plane and underestimate in the borders of the contact region. Note that the agreement between analytical and numerical results is better for the lower value of applied external force due to the assumption of the Hertz theory, i.e. small contact width in comparison with the cylinder radius (10%). In order to analyse the influence of the contact interface discretization, the finite element mesh of the master body is slightly modified in the contact region, reducing the number of master facets from 16 to 15, yielding non-conforming meshes at contact interface. The predicted normal contact pressure distribution obtained with the new discretization is also shown in Figure 6.24, which is globally slightly higher than the one obtained with conforming meshes. However, the predicted contact pressure for the node located in the symmetry plane presents a slight negative deviation in comparison with the result obtained using conforming meshes. This is a consequence of the Node-to-Segment contact discretization using the single-pass algorithm, which underestimates the contact pressure of the slave nodes located in boundary, as previously presented in Figure 3.11 (a).

Note that the solution for the frictionless contact between two elastic cylinders (Figure 6.23) is similar to the one obtained for the problem involving an elastic cylinder in contact with a rigid plane. In that case, the second deformable cylinder (master) is replaced by a fixed rigid flat surface and the frictionless contact occurs between a deformable body and a rigid obstacle, reducing the required computation time (half the number of elements). The numerical solution obtained with such simplification (deformable–rigid) is exactly the same shown in Figure 6.24 for conforming meshes at the contact interface.

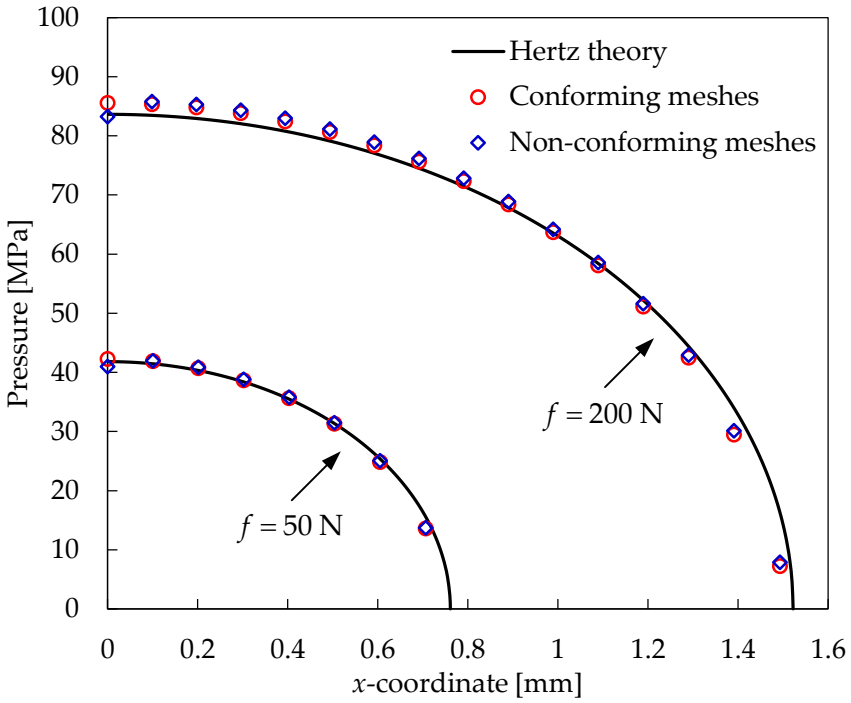
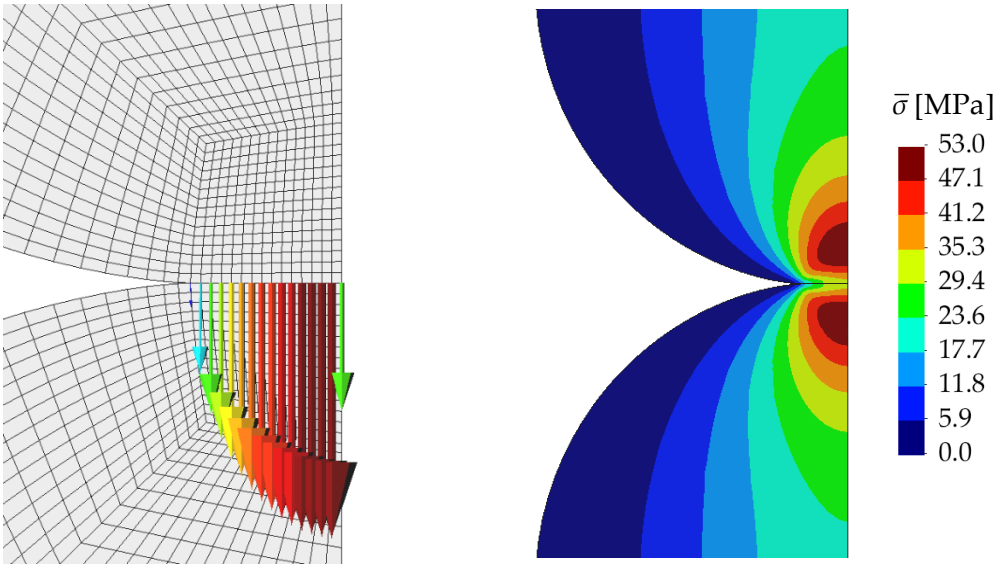


Figure 6.24. Comparison between numerical and analytical solution for the normal contact pressure distribution on the contact surface.

The distribution of the contact forces at the slave nodes is presented in Figure 6.25 (a) for the higher value of applied force ($f = 200$ N), using the discretization of the cylinders with conforming meshes at the contact interface. Since the impenetrability conditions are enforced at the nodes of the slave surface (single-pass Node-to-Segment), the resulting contact forces arise in these same nodes. The contact pressure field, as presented in Figure 6.24, is typically evaluated *a posteriori* from the nodal forces and tributary area methods. The shape of the nodal contact forces distribution shown in Figure 6.25 (a) is similar to the contact pressure distribution (half-ellipse), because the area of the master segments is approximately the same for all. Note that the contact force on the node located in the symmetry plane is half of its value due to the symmetry conditions, i.e. this node is shared by the missing half cylinder.

The von Mises stress distribution on the deformed contacting cylinders is plotted in Figure 6.25 (b), for the higher value of applied external force. The stress contours are perfectly identical for both bodies. Indeed, it was observed that they are not influenced by the arbitrary choice of the master and slave, when adopting conforming meshes at the contact interface. On the other hand, the adoption of non-conforming meshes at the contact interface yields a slight asymmetry in the stress distribution, which results from the incapacity to transfer a uniform pressure through the contact interface for an arbitrary discretization using the Node-to-Segment contact discretization. Nevertheless, the deviation is relatively small due to the very fine discretization of the interface region (16

slave nodes and 15 master quadrilateral Nagata patches). The maximum equivalent stress according to the von Mises criterion lies in the interior, as shown in Figure 6.25 (b), because the maximum shear stress value does not occurs in the interface region [Hughes 76].

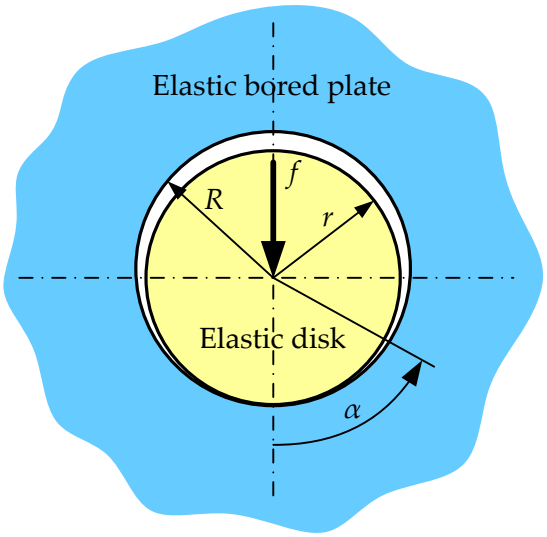


(a) (b)

Figure 6.25. Deformed configuration of the cylinders using conforming meshes at the contact interface for the higher value of applied force: (a) distribution of the nodal contact forces at the slave nodes; (b) von Mises stress distribution.

6.2.2. Disk embedded in a bored plate

This example comprises an elastic disk and an elastic infinite plate with a hole of almost the same radius of the disk. Both the disk and the plate have unitary thickness. The disk is pressed by a concentrated load, situated in its centre, against the cylindrical bore drilled in the infinite plate, as illustrated in Figure 6.26. The main dimensions of the two bodies are listed in the same figure and the Coulomb's friction law is assumed in the contact interface ($\mu = 0.4$). The material properties of the disk and the plate are identical ($E = 210$ GPa, $\nu = 0.3$). The value of the concentrated force is chosen such that the contact occurs at one third of the interface between the hole and the disc (half contact angle $\alpha = 60°$). This problem allows to deal with the frictional contact between two deformable bodies, where the contact surface is curved. The purpose of this example is to evaluate the friction shear stress distribution at the interface, as well to determine the angle for which the transition between the stick and the slip zones occurs.



$R = 60$ mm
$r = 59.99$ mm
$l = 1$ mm (plane strain)
$f = 1875$ N
$E = 210$ GPa
$\nu = 0.3$
$\mu = 0.4$
$\alpha = 60°$

Figure 6.26. Thin elastic disk embedded in a thin elastic infinite plane with a circular hole.

The problem is modelled as 3D using plane strain conditions, i.e. the out-of-plane displacements are fixed on both faces, as previously performed by other authors [Alart 91], [Yastrebov 13]. Due to the symmetry conditions, only one half of the disk and the plate are modelled. The finite element model of the problem is shown in Figure 6.27. The inner disk is discretized by 1896 solid finite elements (8-node hexahedral) using a single layer through the thickness, while the plate is defined by 3,228 finite elements. The external radius of the bored plate used in the finite element model is 10 times higher than its internal radius, allowing taking into account the half space influence. Besides, the nodal displacements are fixed in all nodes of the periphery (Figure 6.27). The finite element mesh is refined in the contact area, as shown in zoom of the contact region presented in Figure 6.27, to attain an accurate friction stress distribution. Moreover, the mesh is generated using the same number of finite elements in the circumferential direction of the disk and the hole of the plate, i.e. each slave node is located close a master node. The interface between the contacting bodies is modelled by a set of Node-to-Nagata contact elements, where the disk and the bored plate are chosen as the slave and the master, respectively. The concentrated force is applied through a rigid surface that is glue with the nodes of a small circular hole created in the centre of the disk (Figure 6.27), retaining the circular form of the small hole.

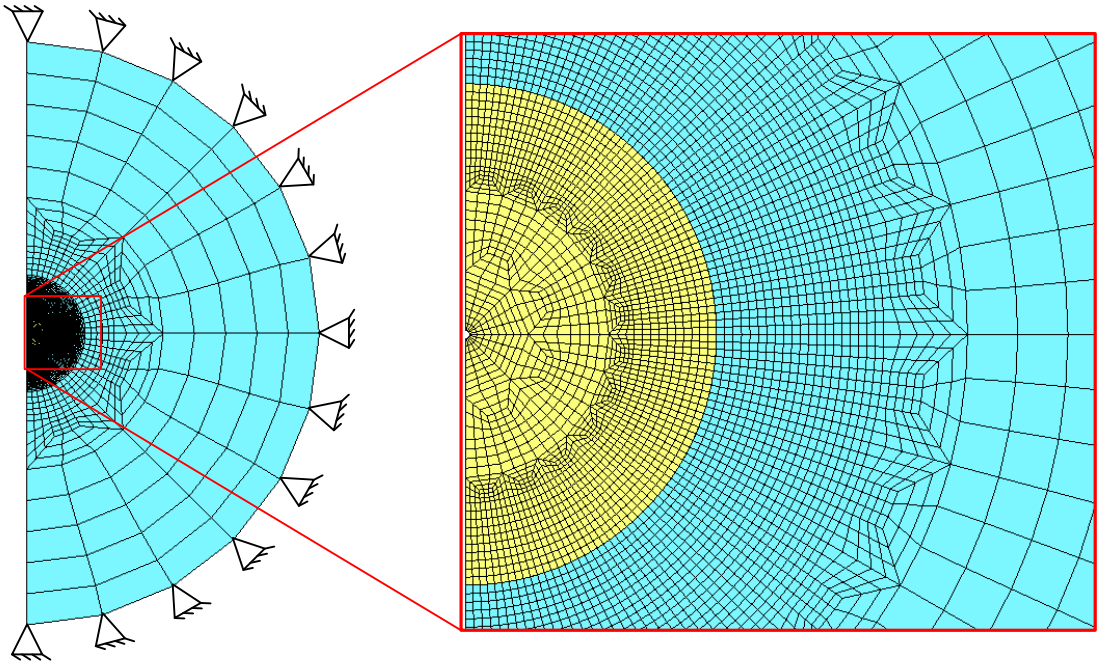


Figure 6.27. Finite element mesh (10,626 nodes, 74 active slave nodes in the interface) and zoom of the contact region.

The semi-analytical solution of this problem, found by Klang [Klang 79], involves integrals that cannot be expressed in closed form and have to be evaluated numerically. The semi-analytical shear stress distribution for the selected material and geometrical data is presented in Figure 6.28 (replicated from [Alart 91]). In the presence of friction, the contact surface can be divided into an inner stick region and an outer slip region, easily identified in the shear contact stress distribution. Since the friction is path-dependent but rate-independent, the semi-analytical solution is independent of the loading rate. Nevertheless, the predicted friction stress distribution is very sensitive to the contact point location, provided by the local contact search algorithm. Therefore, previous studies shown that the numerical solution is strongly dependent on the number of increments adopted to apply the load, particularly the shear stress distribution in the stick zone. A detailed study about the distribution of the shear stress in the contact zone for different number of increments can be found in [Yastrebov 13]. In order to use small increments at the beginning and larger at the end, the applied force value changes quadratically with the increment number in the models proposed by [Alart 91] and [Yastrebov 13]. The presented numerical solution (Figure 6.28) reported by [Yastrebov 13] was obtained using 100 increments, using the NTS contact discretization. On the other hand, the numerical results obtained by [Alart 91] make use of the NTN contact discretization.

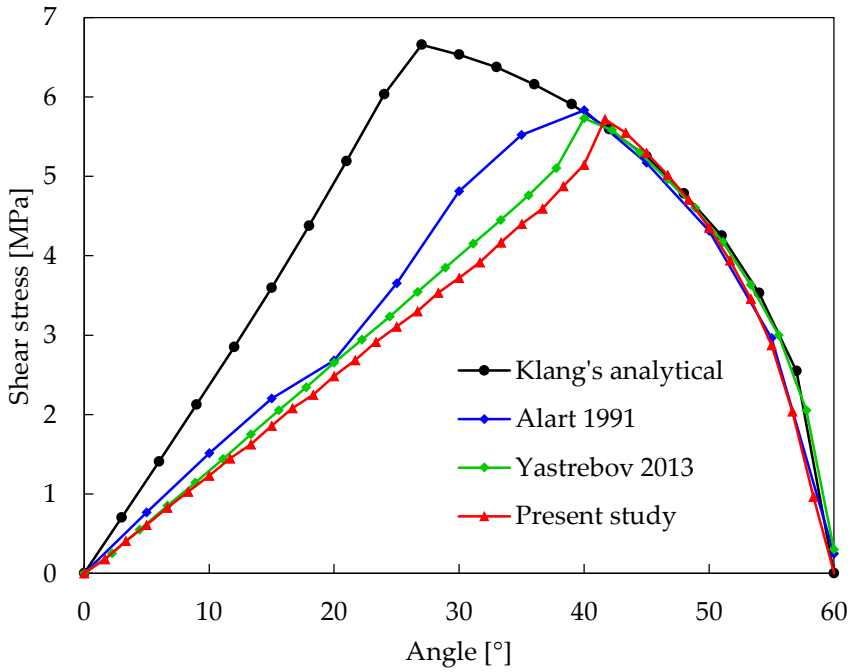
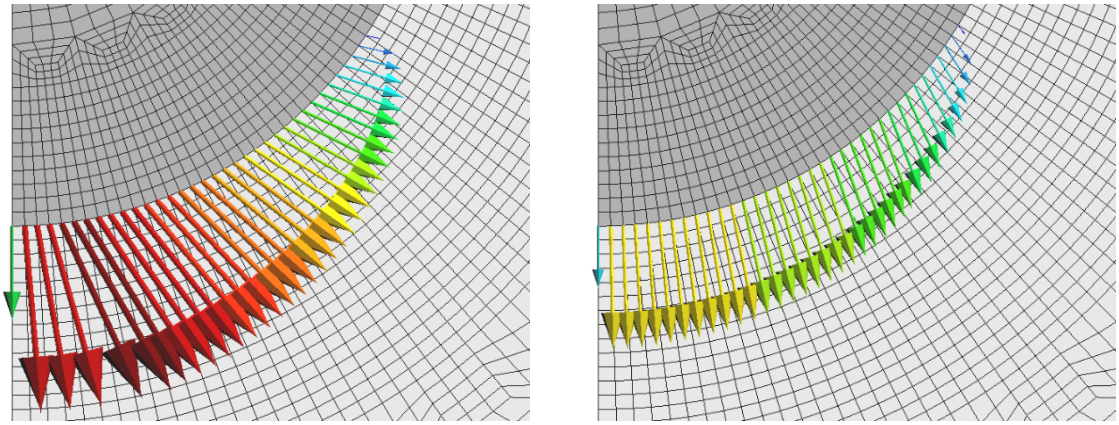


Figure 6.28. Comparison between semi-analytical and numerical solutions for the shear stress distribution in the contact surface.

In the present model, the external load is replaced by an equivalent prescribed displacement, which is applied in 100 increments, allowing an accurate evaluation of the contact point for each active slave node. However, all increments present the same size, i.e. the applied displacement increases linearly with the increment number. The comparison between semi-analytical and numerical solutions for the shear stress distribution in the contact surface is presented in Figure 6.28. The three presented numerical solutions are in good agreement, particularly for the estimate of the angle for which the division between stick and slip regions occurs (corresponding to the maximum shear stress value). Nevertheless, the value for this angle is substantially lower in the semi-analytical solution (approximately 27°), while the finite element solutions predict a value of about 41° (see Figure 6.28). The shear stress distribution in the slip region is in good agreement with the semi-analytical solution, which indicates the accurate prediction of the normal contact pressure distribution. On the other hand, the shear stress estimate provided by the present finite element model in the stick region is slightly lower than the others numerical solutions. However, its linear distribution in the stick region is accurately predicted, as shown in Figure 6.28.



(a) (b)

Figure 6.29. Fully deformed configuration and nodal contact forces distribution in the slave nodes of the disk, applying the external load using: (a) 1 increment; (b) 100 increments.

As previously mentioned, due to the presence of friction at the interface, the numerical solution of this problem is strongly dependent on the number of increments. In fact, the contact status (stick or slip) is defined through the comparison between the predicted tangential component (friction force) and the normal component of the contact force. Further, the direction and magnitude of the friction force is dictated by the location of the contact point (on the master surface), which defines the sliding path. The contact point is calculated by means of the local search algorithm (Section 3.4.2) considering the equilibrium configuration of the previous increment. Since for this example, the sliding path of each slave node over the master surface is inferior to 0.0001 mm (very small sliding), the success of the simulation is defined by the precision achieved in the calculation of the contact point. In order to emphasise this aspect, two loading cases are analysed: (i) 1 increment loading and (ii) 100 equal increments. The nodal contact forces distribution in the slave nodes is presented in Figure 6.29 for the two cases. The nodal contact forces distribution is very sensitive to the number of increments, both in magnitude and direction. In fact, the orientation of the friction force (direction of the nodal contact force) is incorrectly inverted when a single increment is used, as shown in Figure 6.29. In order to obtain an accurate solution, the increase of the external load should be performed such that only one new contact element is created in each load increment [Pietrzak 97]. This strategy allows to accurately define the contact point through the normal projection of the slave node on the Nagata patch, since the node coming in contact is very close to the master surface. Moreover, the asymmetry of the closest point definition (see Figure 2.9 (a)) is reduced by means of the proximity between slave node and master surface.

6.2.3. Rotating concentric hollow spheres

The third example involving deformable bodies is a simplification of the pressurized spheres contact problem proposed by Puso and Laursen [Puso 04a], which was originally solved with the mortar based method under frictionless contact conditions. The presented example is a 3D problem comprising two concentric hollow spheres undergoing frictional contact with large sliding, as illustrated in Figure 6.30, where the outer and the inner hollow spheres are depicted in dark grey and light grey, respectively. The principal dimensions of the two spheres are listed in Figure 6.30, as well as the linear elastic material properties (identical for both bodies: $E = 1000\,\text{MPa}$ and $\nu = 0.3$). Since the exterior radius of the inner sphere ($r_{\text{ext}} = 100.5\,\text{mm}$) is larger than the interior radius of the outer sphere ($R_{\text{int}} = 99.5\,\text{mm}$), an initial overlap is considered between the spheres. This initial overlap generates a contact pressure such that the exterior radius of the inner hollow sphere equals the interior radius of the outer hollow sphere. Moreover, the outer hollow sphere is fixed at its exterior radius, while the interior radius of the inner hollow sphere is rotated by 90° counter-clockwise (see Figure 6.30).

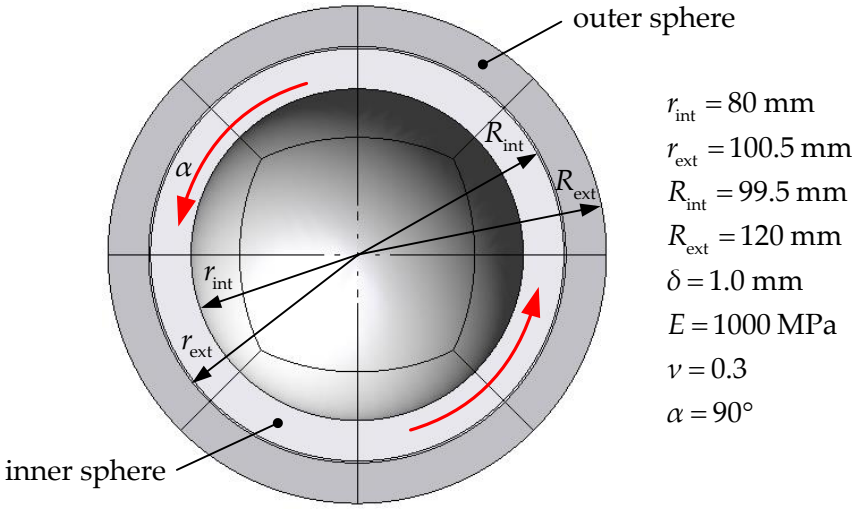
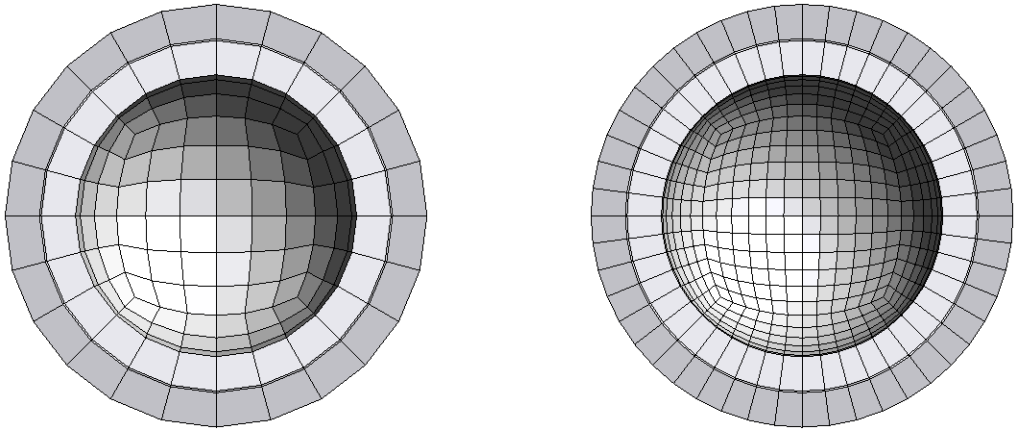


Figure 6.30. Two concentric hollow spheres undergoing frictional contact with large sliding, including geometrical and material properties.

The main purpose of this example is to demonstrate the ability of the developed Node-to-Nagata contact element to deal with large sliding between curved surfaces, highlighting the improvement in the numerical results when the master surface is smoothed. Due to the symmetry conditions, only one half of the hollow spheres are modelled and the finite element model is depicted in Figure 6.31. Two different structured meshes are considered in the numerical simulations: (i) a coarse mesh where each sphere is discretized by 108 solid finite elements (8-node hexahedral) using a single layer through the thickness (Figure 6.31

(a)) and (ii) a fine mesh where each sphere is discretized by 432 solid finite elements (Figure 6.31 (b)). Both finite element meshes are created in such a way that the slave nodes are coincident with the master nodes (conforming mesh) at the beginning of the simulation. The displacements of the nodes located in the exterior radius of the outer hollow sphere are fixed by means of a rigid surface glue to the nodes. The inner hollow sphere is rotated 90° counter-clockwise using a rigid surface glue to the nodes located in the interior radius of the inner hollow sphere. Two variants of the problem are considered for the present example: frictionless and frictional cases. For the frictionless case, the influence of the master-slave choice on the numerical results is also studied.



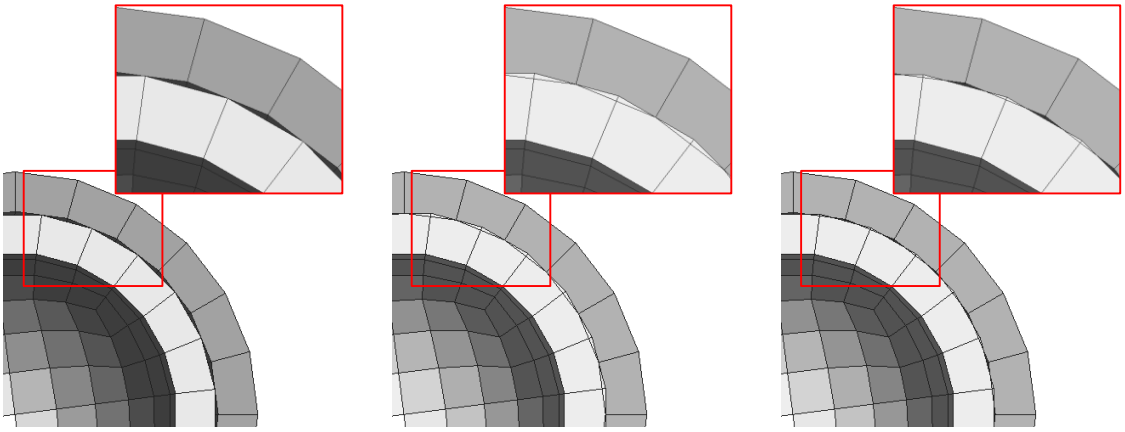
(a) (b)

Figure 6.31. Finite element mesh of the concentric hollow spheres: (a) coarse mesh with 484 nodes (121 active slave nodes); (b) fine mesh with 1,828 nodes (457 active slave nodes).

6.2.3.1. Frictionless case

Firstly, the frictionless case is analysed, which involves only the contact pressure between the inner and outer hollow spheres, resulting from the given initial overlap (see Figure 6.30), i.e. the direction of the nodal contact forces is exactly radial. The torque around the rotation axis is measured in the interior radius of the inner hollow sphere, as a function of the rotation angle. For the frictionless case its values should be zero. The numerical simulations are performed considering both faceted and smoothed definitions of the master surfaces. Concerning the surface smoothing method using Nagata patches, the required nodal normal vectors are approximated through the weighting factor MWSELR expressed in (4.43), which provides the exact normal vectors for spherical contact surfaces (see Figure 4.43). Indeed, the interface between the two hollow spheres is a spherical surface, as shown in Figure 6.31. The accuracy achieved in the description of the spherical master surface using both bilinear facets and Nagata patches is presented in Figure 4.19 (a) for the maximum radial error and in Figure 4.19 (b) for the maximum normal vector error. The

accuracy of the surface normal vector is an indirect measure of smoothness, which is very important in contact problems with large sliding.



(a) (b) (c)

Figure 6.32. Configuration of the concentric hollow spheres for a rotation angle of 7.5° considering: (a) outer hollow sphere defined as faceted master; (b) inner hollow sphere defined as faceted master; (c) outer hollow sphere defined as smoothed master.

The frictionless case of the present example is used to highlight the influence of the master-slave selection on the numerical results. Two alternatives are possible to define the contact interface: (i) the master surface is defined on the interior radius of the outer hollow sphere or (ii) the master surface is defined on the exterior radius of the inner hollow sphere. Since the impenetrability condition is enforced only at the slave nodes (Node-to-Nagata contact element), this approach precludes the penetration of the slave nodes into the master surface, where the latter is defined both by bilinear facets and Nagata patches. Therefore, the adoption of bilinear facets in the description of the curved master surface yields numerical results strongly dependent on the choice of the master surface, as shown in Figure 6.32 for a rotation angle of 7.5°, when adopting the coarse mesh. When the interior radius of the outer hollow sphere is assigned as master surface (Figure 6.32 (a)), the enforcement of the penetration condition for the slave nodes leads to a gap for the master nodes, producing very high nodal contact forces. On the other hand, the selection of the exterior radius of the inner hollow sphere as master surface leads to the penetration of the nodes of these surface into the slave body, as shown in Figure 6.32 (b). The configuration of the concentric hollow spheres using the master surface smoothed with Nagata patches is presented in Figure 6.32 (c) for the same rotation angle, using the interior radius of the outer hollow sphere as master. The adopted surface smoothing approach allows the slave nodes to apparently "penetrate" the opposing master surface, as highlighted in the detail of Figure 6.32 (c). An identical configuration is obtained by switching the master and slave surfaces.

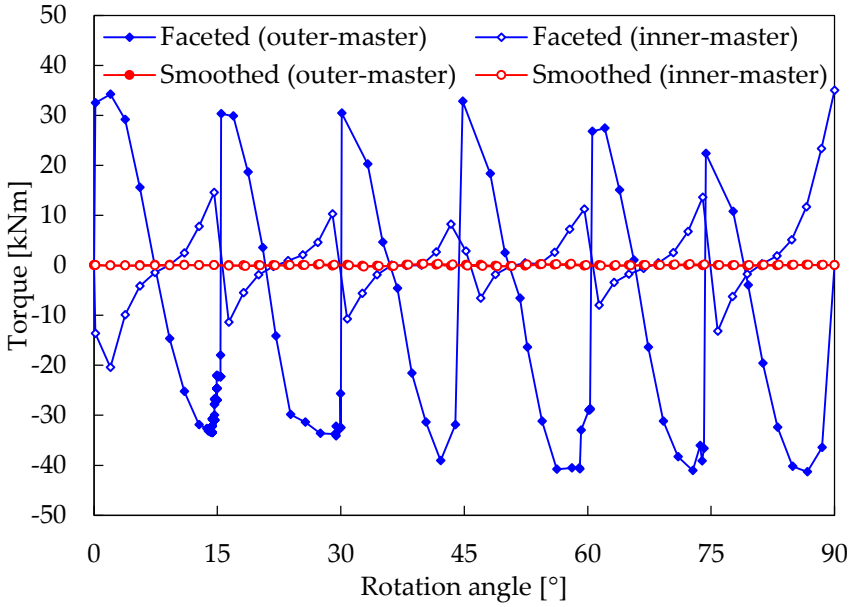


Figure 6.33. Influence of the master and slave surfaces selection in the torque evolution with the rotation angle, for both faceted and smoothed master surface descriptions (coarse mesh).

The evolution of the torque measured for the interior radius of the inner hollow sphere is presented in Figure 6.33, for the two alternative definitions of the master and slave surfaces, adopting the coarse mesh. Moreover, the faceted and smoothed descriptions of the master surface are compared, highlighting the higher accuracy obtained with the application of Nagata patches in the contact surface smoothing. Since the friction is not taken into account at the contact interface the torque should be zero for all rotation angles. Nevertheless, the faceted description of the master surface leads to strong oscillations in the resulting torque, particularly when the outer hollow sphere is assigned as master body, as shown in Figure 6.33. In fact, the existing gap between the hollow spheres resulting from the selection of the outer hollow sphere as master (see Figure 6.32 (a)) produces a higher value of contact pressure, which is reproduced in the large amplitude of the torque oscillations. Since the structured coarse mesh presents six finite elements to describe 90° of the circumferential direction (see Figure 6.31 (a)), the numerical simulation using the faceted master surface also generates six waves in the torque evolution.

The application of the Nagata patches in the smoothing of the master surface leads to a constant torque evolution being its value approximately zero, as shown in Figure 6.33. Moreover, the numerical solution is somewhat independent of the selected master surface (outer or inner hollow sphere). Therefore, the developed Node-to-Nagata contact element is much more accurate to solve contact problems involving large relative sliding between the contact surfaces, in comparison with the typical faceted description of the master surface.

The computational performance is affected by the choice of the master and slave surfaces, when the faceted surface description is adopted for the master surface. Indeed, the computational time required to carry out the simulation using the interior radius of the outer hollow sphere defined as master surface is substantially higher than the one obtained when using the exterior radius of the inner hollow sphere as master surface. This increase in the computational cost is directly related with the severe convergence problems resulting from the discontinuity of the master surface normal vector field. When the outer hollow sphere is defined as master body, the master surface is concave (Figure 3.19 (b)), while the definition of the inner hollow sphere as master body leads to a convex master surface (Figure 3.19 (a)). Consequently, the discretization of these curved surfaces with bilinear facets leads to the existence of internal and external blind spots, respectively. For the case of a concave master surface, the sliding of a slave node from one facet to a neighbouring facet leads to a constant switching between two adjacent master facets (*flip–flop* effect illustrated in Figure 5.12). In fact, the number of increments necessary to complete the simulation is higher for this case (see amount of points in Figure 6.33), because the increment size is automatically reduced when convergence is not attained within the allowed number of iterations (*NTrial* strategy). Moreover, it is necessary to increase the admissible tolerance to achieve the convergence. Thus, in the following simulations the master surface is always defined on the exterior radius of the inner hollow sphere. On the other hand, when the Nagata patches are used in the description of the master surface, the computational performance is not affected by its selection, as shown in Figure 6.33.

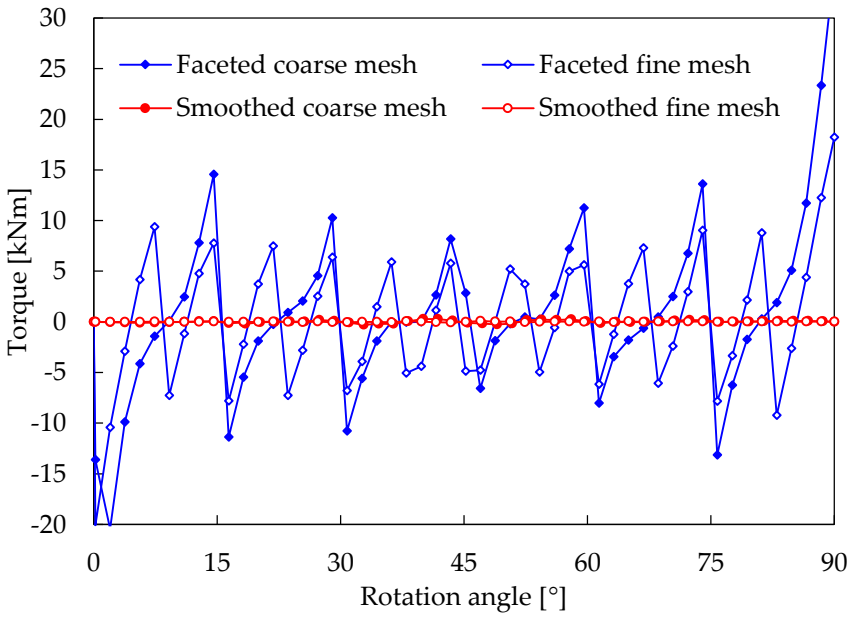


Figure 6.34. Influence of the finite element mesh (coarse and fine) in the torque evolution with the rotation angle for the frictionless case, for both faceted and smoothed master surface descriptions.

The evolution of the torque with the rotation angle obtained with the coarse and fine meshes (Figure 6.31) is presented in Figure 6.34, for both faceted and smoothed master surface descriptions. Note that the master surface is defined on the exterior radius of the inner hollow sphere. Considering the master surface described by bilinear facets, the amplitude of the oscillations in the resulting torque is only slightly reduced with the mesh refinement, as shown in Figure 6.34. Besides, the number of waves observed in the torque evolution increases from six to twelve (double) due to the mesh refinement in the circumferential direction of the hollow spheres (see Figure 6.31 (b)). On the other hand, the smoothed description of the master surface with Nagata patches produces a constant torque evolution of approximately zero, which is the analytical solution for the frictionless case. Thus, the chatter effect observed in the faceted description of the master surface is eliminated through the adoption of Nagata patches to smooth the master surface, as consequence of the higher accuracy in the surface normal vector. In fact, both finite element meshes adopted for this example lead to an accurate evolution of the torque when the smoothing procedure is applied, as shown in Figure 6.34.

Concerning the computational performance, the number of increments is the same for both faceted and smoothed master surface descriptions, i.e. the inner hollow sphere is rotated 90° counter-clockwise in 50 increments (1.8° per increment). Besides, the average number of iterations required to achieve convergence in each increment is the same for all models under analysis, about 5 iterations, indicating the good stability of the contact algorithm. The computational time required to carry out the simulations using the coarse mesh is approximately 3.3 seconds, while the models with fine mesh need about 15.4 seconds. The computational cost is only slightly reduced (about 2%) by the application of the surface smoothing method on the master surface.

6.2.3.2. Frictional case

The second variant of this example takes into account the frictional effects at the contact interface, considering a Coulomb friction coefficient of $\mu = 0.1$. This contact problem involves the evaluation of both the contact pressure and the tangential shear stress between the inner and outer hollow spheres, i.e. the direction of the nodal contact forces is not radial. The torque around the rotation axis is evaluated as a function of the rotation angle, which is different from zero due to the friction and initial overlap between the hollow spheres. Following the results shown in the previous section, the exterior radius of the inner hollow sphere is always considered to be the master surface.

The evolution of the resulting torque with the rotation angle is presented in Figure 6.35 for both spheres discretizations (coarse and fine meshes depicted in Figure 6.31), in order to compare the solution obtained with the faceted and smoothed master surface descriptions. Since the initial overlap and the friction coefficient are constant, the torque should also be constant during the rotation. Nevertheless, the faceted description of the

master surface leads to severe oscillations in the predicted torque, particularly for the coarse mesh. On the other hand, the smoothing of the master surface with Nagata patches yields a torque that increases from zero, during the stick contact phase, to a value that remains constant during the sliding of the contact interface, as shown in Figure 6.35. The constant value of the torque is approximately 33 kNm when the surface smoothing method is adopted in the numerical simulation. The amplitude of the oscillations in the torque evolution predicted by the model with the master surfaces modelled by bilinear facets are reduced with the mesh refinement, while its frequency increases to double (coincident with the master surface discretization), similarly to the frictionless case shown in Figure 6.34. Nevertheless, the constant value of torque is always underestimated by the faceted description of the master surface, due to the admissible overlap of the master nodes into the slave body (see Figure 6.32 (b)), which reduces the contact pressure and, consequently, the shear stress. The adoption of Nagata patches in the representation of the master surface produces a constant value of the torque, which increases approximately 2% with the mesh refinement. A similar result was obtained by Pietrzak and Curnier [Pietrzak 99] for a 2D problem comprising an elastic shaft rotating inside a bore, using Bézier and Spline interpolation for the smoothing method. The computational performance of the numerical simulations involving friction in the contact interface is identical to the one obtained for the frictionless case. The number of increments used to discretize the rotation angle is the same and the convergence is attained without problems in an average of 5 iterations per increment.

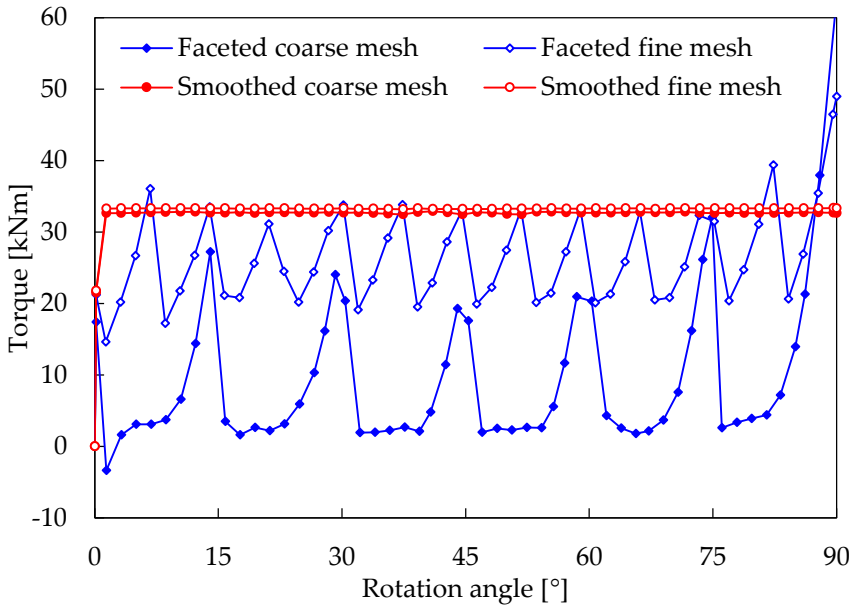
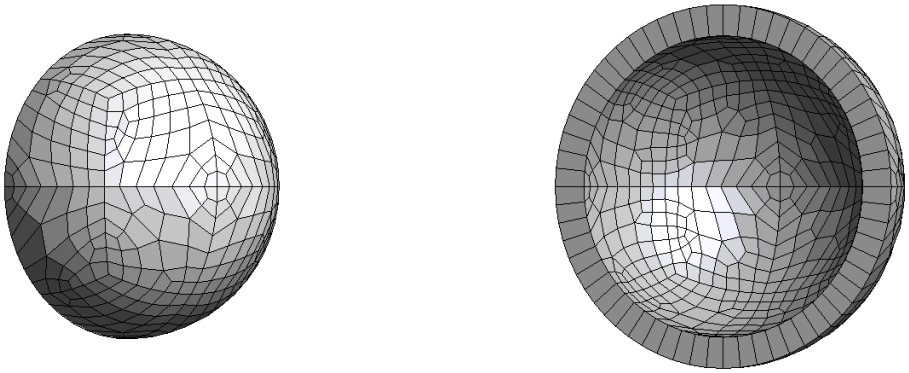


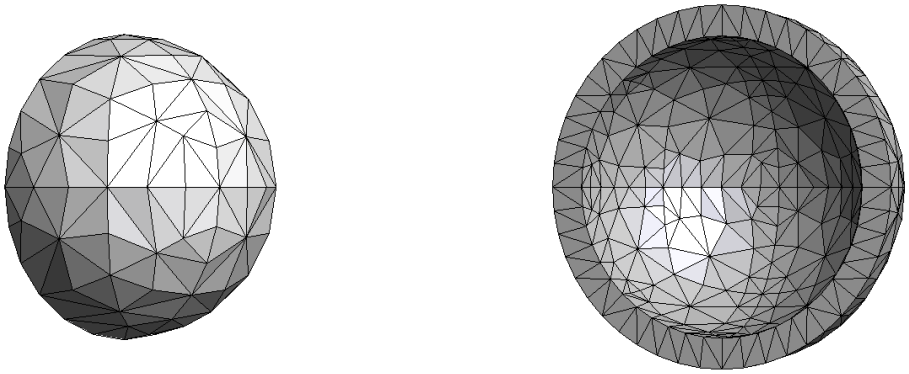
Figure 6.35. Influence of the finite element mesh (coarse and fine) in the torque evolution with the rotation angle for the frictional case, for both faceted and smoothed master surface descriptions.

In order to highlight the ability of the proposed surface smoothing method to deal with completely unstructured meshes at the frictional contact interface, the hollow spheres were also discretized using the finite element mesh shown in Figure 6.36. The inner hollow sphere is discretized with 402 solid finite elements (8-node hexahedral) and the outer hollow sphere with 538 finite elements, both using a single layer through the thickness. A similar problem (irregular mesh) was presented by Puso and Laursen [Puso 02], but involving a deformable hollow sphere surrounded by two rigid spherical shells. Moreover, the deformable bodies were also discretized with linear tetrahedral finite elements, as shown in Figure 6.37, in order to test the ability of the contact elements composed by triangular Nagata patches. The inner hollow sphere is discretized with 420 solid finite elements (4-node tetrahedral) while the outer hollow sphere is defined by 1182 finite elements. Although the number of finite elements used to describe the inner hollow sphere is approximately the same in both models, the number of master segments is significantly smaller in the tetrahedral finite element mesh (compare Figure 6.36 (a) with Figure 6.37 (a)).



(a) (b)

Figure 6.36. Unstructured discretization composed by hexahedral finite elements: (a) inner hollow sphere with 862 nodes; (b) outer hollow sphere with 1,142 nodes.



(a) (b)

Figure 6.37. Unstructured discretization composed by tetrahedral finite elements: (a) inner hollow sphere with 166 nodes; (b) outer hollow sphere with 444 nodes.

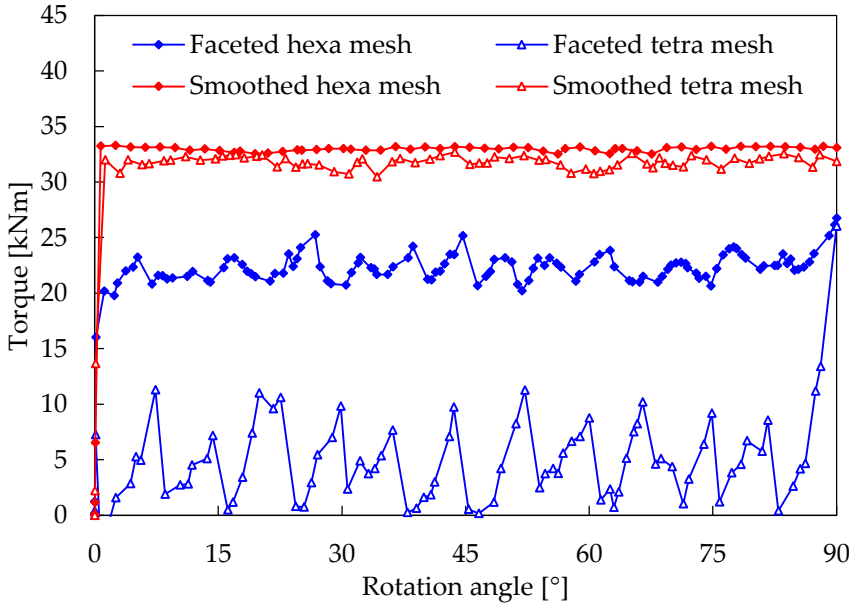


Figure 6.38. Torque evolution with the rotation angle using the unstructured finite element meshes, for both faceted and smoothed master surface descriptions.

The evolution of the predicted torque using the unstructured meshes, shown in Figure 6.36 and Figure 6.37, is presented in Figure 6.38 for both descriptions of the master surface (faceted and smoothed). The application of the surface smoothing method using Nagata patches to describe the master surfaces leads to a torque evolution approximately constant, with a value (33 kNm) which is close to the one obtained with the structured meshes. On the other hand, the faceted description of the master surface produces oscillations in the torque evolution, as shown in Figure 6.38, which are more severe in the tetrahedral finite element mesh than in the hexahedral mesh. Moreover, the torque provided by the tetrahedral finite element mesh is substantially lower because the amount of bilinear faceted used in the description of the master surface is also inferior (compare Figure 6.36 (a) and Figure 6.37 (a)). Although the number of finite elements used in the unstructured mesh composed by hexahedral finite elements (Figure 6.36) is only slightly higher than in the structured fine mesh (Figure 6.31 (b)), adopting the faceted description of the master surface, the oscillations in the torque evolution are smaller in the unstructured mesh (compare Figure 6.38 with Figure 6.35). This is related with the construction of the finite element mesh at the contact interface, which is always non-conforming in the unstructured mesh (inner and outer hollow spheres are discretized independently).

Concerning the computational performance for the hexahedral mesh, the number of increments required for the faceted surface description is approximately twice the one required when using the smoothed master surface, as shown in Figure 6.38 through the amount of dots. Indeed, the increment size is automatically reduced by means of the r_{min} strategy in order to improve the convergence during the corrector step. Consequently the

computational time increases from about 25 seconds, when the smooth representation of the master surface is adopted, to approximately 42 seconds, when using the faceted surface description. On the other hand, the application of the surface smoothing method in the discretization of the hollow spheres with tetrahedral finite elements leads to a reduction of 15% in the required number of increments. However, the computational time is approximately the same for both geometrical descriptions of the master surface, about 6 seconds.

6.2.4. Extrusion of an aluminium billet

The following example involves the extrusion of an aluminium billet into a deformable conical die. It has been previously studied by Padmanabhan and Laursen [Padmanabhan 01] in a 2D setting (axisymmetric model), to demonstrate the advantages of using a smoothing scheme on the discretized master contact surface. However, in the present study the problem is modelled in a 3D setting. Thus, the contact interface is treated as a surface instead of a curve. The cylindrical billet is drawn through a conical die with friction at their interface, as shown in Figure 6.39. This example illustrates the ability of the presented formulation for treating large sliding frictional contact problems between deformable bodies with elastoplastic material behaviour.

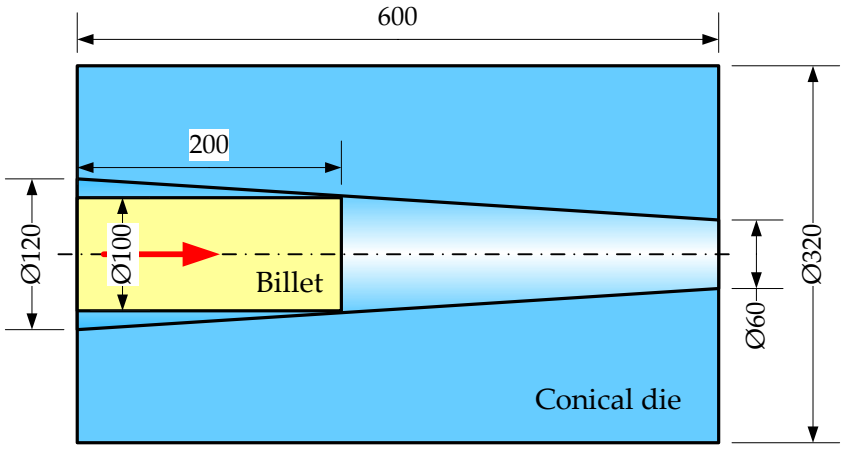


Figure 6.39. Extrusion of an aluminium billet in a conical die including geometrical and material properties (dimensions in mm).

The backside of the billet is subjected to an axial displacement into the conical die (total displacement of 150 mm), while the outer boundary (exterior radius) of the die is considered fixed in all directions. The Coulomb friction coefficient between the aluminium billet and the conical surface of the die is assumed to be $\mu = 0.15$. The dimensions of each body are indicated in Figure 6.39. The cylindrical billet is modelled assuming an

elastoplastic material behaviour, with the linear isotropic hardening law expressed by $\sigma = 31 + 261.2\bar{\varepsilon}^{p}$ [MPa]. The same type of mechanical behaviour is assumed for the conical die, with material with the hardening law given by $\sigma = 31 + 2.612\bar{\varepsilon}^{p}$ [MPa]. The material elastic properties are shown in Figure 6.39. These material properties were selected in order to ensure considerable deformation on both bodies [Padmanabhan 01].

Due to symmetry conditions, only one quarter of the problem is modelled. The finite element mesh (8-node hexahedral) of both deformable bodies is presented in Figure 6.40, where the billet is discretized by 540 finite elements and the conical die is defined by 840 finite elements. The element size in the axial direction of the billet is half the size used to discretize the die. The billet is defined as slave body while the conical die is assigned as master. Since the master surface (conical surface of the die) is convex, some convergence problems can arise in the numerical solution when using the faceted description of the master surface, namely when the slave and master nodes are very close. Therefore, the finite element mesh of the contact surfaces is created using six elements in the circumferential direction of the billet and seven elements for the die, as shown in the detail of Figure 6.40. This configuration allows avoiding the convergence problems related with the *flip–flop* effect in the circumferential direction. The displacements of the nodes located in the exterior radius of the die are fixed by means of a cylindrical rigid surface glue to the nodes. The axial displacement is applied incrementally in the backside of the billet, allowing its radial displacement.

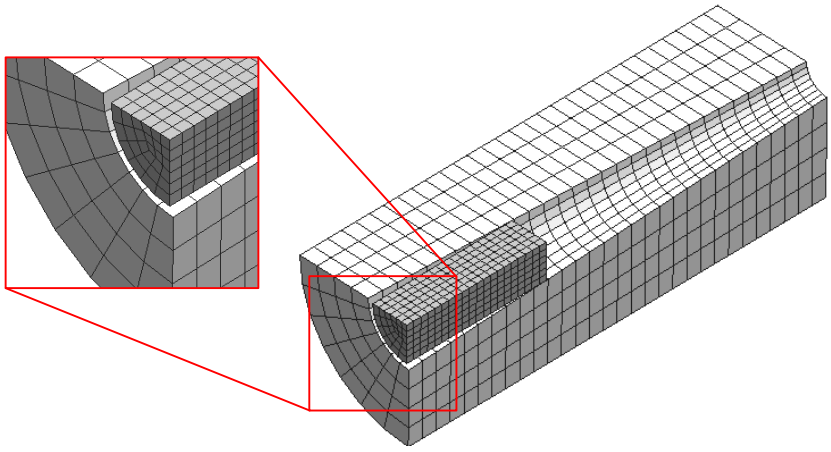


Figure 6.40. Finite element mesh of the billet (777 nodes) and the conical die (1,240 nodes) with detail of the discretization in the circumferential direction.

The evolution of the axial force acting on the backside of the billet as function of its displacement is presented in Figure 6.41, for both descriptions of the master surface (faceted and smoothed). The oscillations in the axial force for the frictional extrusion using the master surface described by bilinear facets are reduced through the smoothing procedure

with Nagata patches. The chatter effect is produced by the sudden changes in the surface normal between adjacent master facets, when the slave nodes slide on the master surface along the axial direction. Note that the sudden changes of the normal vector in the circumferential direction do not affect the frictional contact stability, because the sliding only occurs in the axial direction. Moreover, the value of axial force predicted by the model with faceted master surfaces is slightly higher than using the smoothed description of the surface, as shown in Figure 6.41. This is motived by the selection of the convex surface as master, which dictates a higher value of contact pressure due to the artificial gap between the billet and the die surface (see Figure 6.32 (a)).

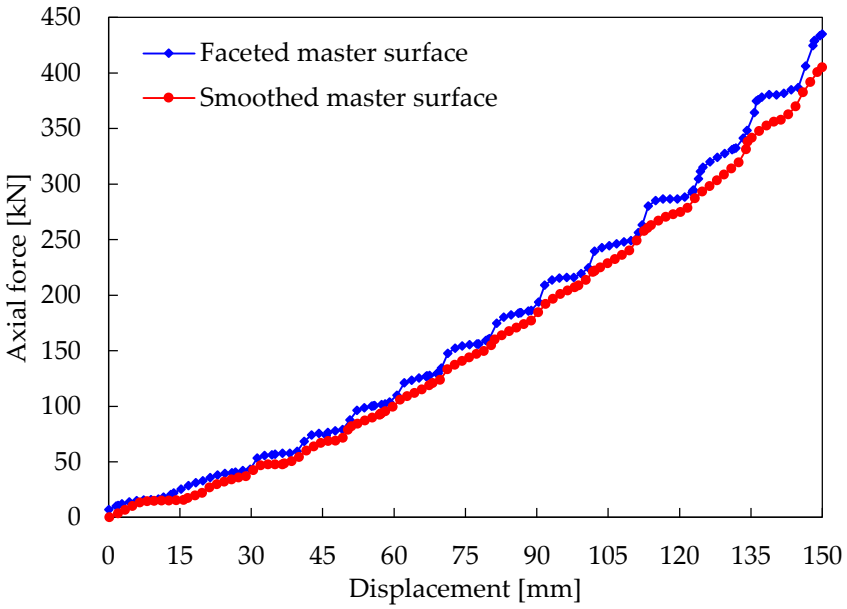
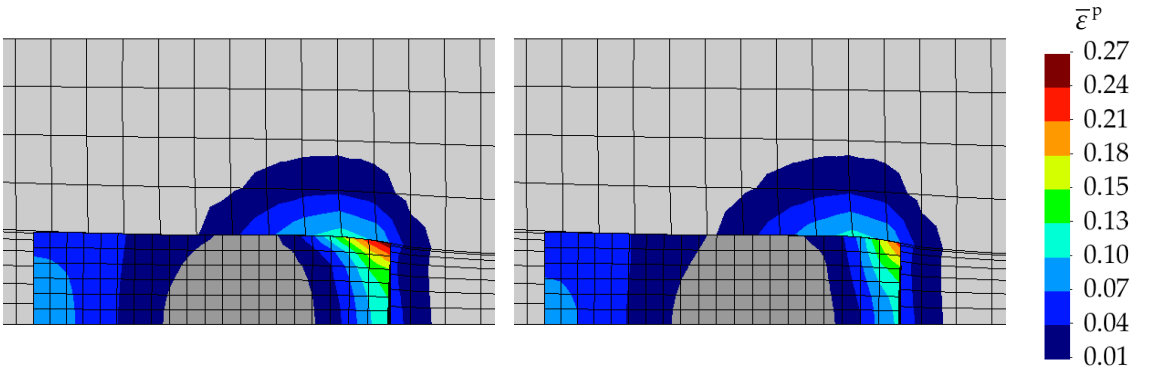


Figure 6.41. Axial force evolution as function of the displacement of the billet for the frictional extrusion problem, for both faceted and smoothed master surface descriptions.

The equivalent plastic strain distribution plotted in the deformed configuration (150 mm of displacement) for the frictional extrusion problem is presented in Figure 6.42, comparing faceted and smoothed master surface descriptions. The maximum value of equivalent plastic strain occurs in the periphery of the billet (forward side). Its value is approximately 27% for the faceted description of the master surface and 23% considering the surface smoothed with Nagata patches. An important amount of plastic strain arises in the backside of the billet, which is motived by the friction force. On the other hand, the plastic strain in the conical die is only located near the contact interface (Figure 6.42), where the nodal contact forces attain its maximum value. The die deformation at the end of the billet axial displacement is shown in Figure 6.42.



(a) (b)

Figure 6.42. Equivalent plastic strain distribution for the frictional extrusion problem plotted in the fully deformed configuration: (a) faceted master surface; (b) smoothed master surface.

The frictionless extrusion problem (highly lubricated process) is also analysed in the present study. The finite element model is exactly the same as for the frictional case, but the friction is not taken into account at the contact interface. The axial force evolution as function of the billet backside displacement is shown in Figure 6.43, for both faceted and smoothed master surface descriptions. Considering the piecewise bilinear representation of the master surface, the force evolution provided by the frictionless simulation contains oscillations with larger amplitude than the frictional case (compare Figure 6.43 and Figure 6.41). Indeed, for the frictionless case, the axial force decreases after an abrupt increase, while in the frictional case the predicted axial force is always increasing. Thus, the proposed surface smoothing method seems to be very efficient also in the elimination of the force oscillations in the frictionless simulation, as shown in Figure 6.43. In fact, the surface smoothing method is more effective in the frictionless extrusion problem, as previously reported by Padmanabhan and Laursen [Padmanabhan 01]. The oscillations in the axial contact force are not completely eliminated with the application of Nagata patches because the proposed interpolation method does not provide a continuous evolution of the normal vector field. In case of contact surfaces with inflection points, such as occurs in the present example for the conical surface of the die (see Figure 6.42), the continuity of the normal vectors across patches is not guaranteed.

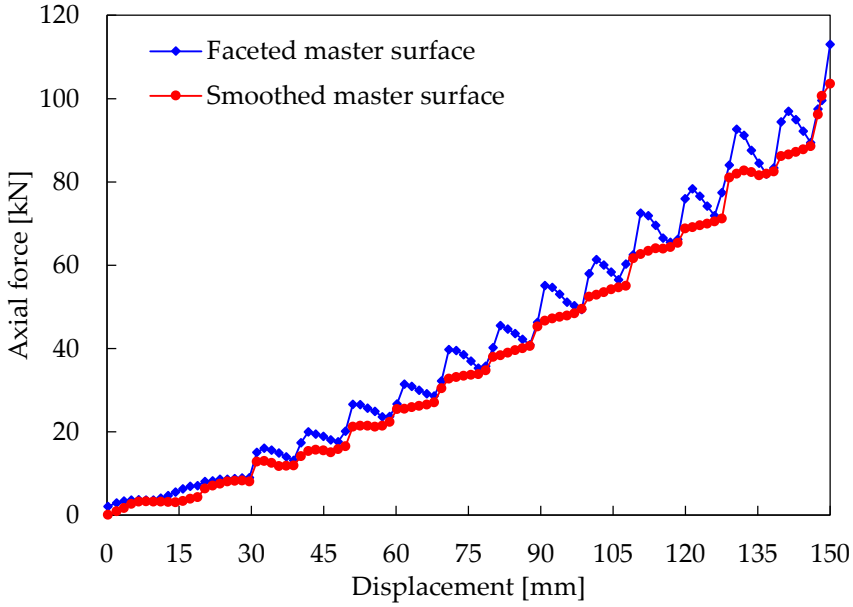


Figure 6.43. Axial force evolution as function of the displacement of the billet for the frictionless extrusion problem, for both faceted and smoothed master surface descriptions.

The equivalent plastic strain distribution for the frictionless extrusion problem is presented in Figure 6.44 for both master surface description methods (faceted and smoothed). The maximum value predicted in the frictionless case is approximately the same previously obtained in the frictional extrusion problem, as well as its location (see Figure 6.42). However, the backside of the billet does not present any plastic strain because the extrusion force involved in this process is significantly lower, as well as the effective contact area. Concerning the conical die, the volume submitted to plastic strain is slightly lower for the frictionless extrusion problem.

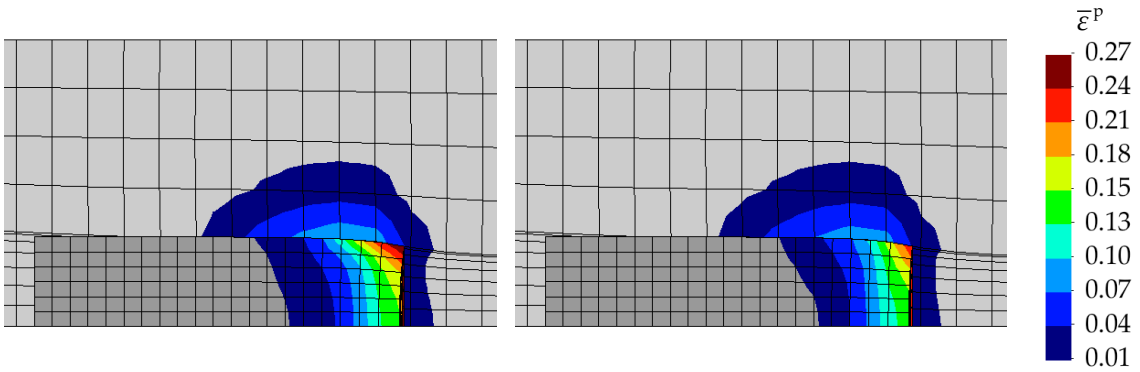


Figure 6.44. Equivalent plastic strain distribution for the frictionless extrusion problem plotted in the fully deformed configuration: (a) faceted master surface; (b) smoothed master surface.

Table 6.7. Computational performance of the extrusion problem using different master surface description methods.

	Frictional case		Frictionless case	
	Faceted	Smoothed	Faceted	Smoothed
Nº increments	128	110	100	100
Nº iterations	789	698	612	625
Computational time [s]	104	95	66	67

The comparison of the computational performance between the faceted and the smoothed surface description methods is presented in Table 6.7, for the frictional and the frictionless extrusion problems. Regarding the frictional extrusion problem, the proposed surface smoothing method improves the computational performance of the numerical simulation. The required number of increments used to impose the axial displacement is reduced from 128 to 110, while the computational time is reduced in approximately 10 seconds (10%). Indeed, the sudden changes in the surface normal vector are strongly reduced by means of the Nagata patches in the definition of the contact elements, avoiding some convergence problems. The average number of iterations required is only slightly higher in the smoothed description of the master surface because the increment size is larger. The computational performance of the frictionless extrusion problem is better than the homologue frictional problem, since the frictional effects play an essential role in the problem considered. Moreover, the necessary number of increments and the computational time is not affected by the surface smoothing method. In fact, the total number of iterations increases slightly with the description of the master surface by Nagata patches. This is related with the slow convergence occurred in some increments, which can be motived by *flip–flop* effect in the axial direction between two adjacent patches. Note that the geometry of each Nagata patch describing the master surface is updated in each iteration within each increment, which can introduce some perturbation in the numerical solution (switching between Nagata patch interpolation and bilinear interpolation)

6.2.5. Post-buckling of a thin walled tube

In this example, the *quasi*-static post-buckling of a thin walled tube is studied. This classical axisymmetric example involves large plastic deformations, post-buckling and self-contact phenomena. This problem was firstly introduced by Laursen and Simo [Laursen 93] in 2D setting (axisymmetric finite element model). More recently, a simplified version has been studied by Yang and Laursen [Yang 08b] to demonstrate the effectiveness of the mortar based method to deal with self-contact phenomena, undergoing large deformations and sliding. This type of problems is frequently encountered in crashworthiness research, where structural elements may buckle and subsequently wrinkle. All geometrical and

material properties of the example proposed by [Yang 08b] are adopted in the present study.

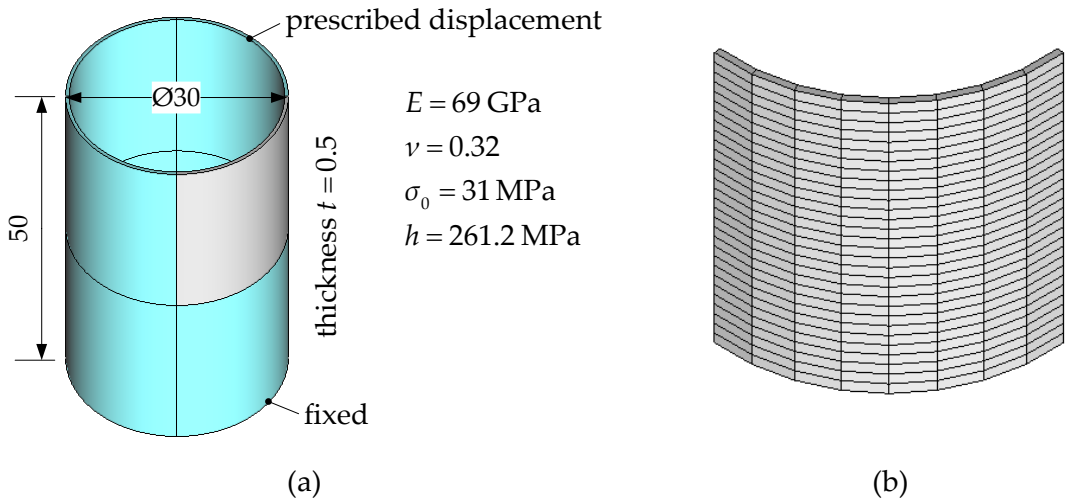


Figure 6.45. Post-buckling of a thin walled tube: (a) geometrical and material properties (dimensions in mm); (b) finite element mesh of one eighth of the tube (558 nodes).

The initial configuration and the finite element mesh of one eighth of the tube are shown in Figure 6.45. The bottom surface of the thin walled tube is fixed in all directions, while a vertical displacement is applied to the top surface of the tube. The prescribed displacement in the axial direction is 40 mm and the radial displacement is not allowed. The tube dimensions are given in Figure 6.45, as well as the elastoplastic material properties (linear isotropic hardening). Indeed, the material properties of the tube are exactly the same used in the aluminium billet of the previous example. Frictionless response is assumed in the self-contacting buckle regions of this problem. Since the application of the prescribed axial displacement (40 mm) in the top surface of the tube is equivalent to apply 20 mm of displacement in each face of the tube (top and bottom), it allows to model only half length of the tube. Thus, due to the symmetry conditions, only one eighth of the tube is simulated, as shown in Figure 6.45. The finite element mesh is presented in Figure 6.45 (b), where the tube is discretized with 240 solid finite elements (8-node hexahedral), using a single layer through the thickness. This finite element mesh is identical to the one adopted by [Yang 08b], which is denoted by coarse mesh due to the reduced number of elements.

In post-buckling problems it is difficult to predict in advance which portions of a surface will come into contact with each other. Therefore, the contact pairs required for the master-slave discretization are unknown *a priori* (a series of buckles occur). The straightforward solution for this class of contact problems is the application of the two-pass Node-to-Segment approach, discussed in Section 3.3.1.2, which is employed in this example. Thus, each node of the contact surface is defined as slave node, while all facets composing the contact surface are assigned as master segments. For the problem under

analysis, two independent self-contact surfaces are distinguished, i.e. the inner and outer surface of the tube. This means that the nodes belonging to the interior surface of the tube are not allowed to penetrate their own surface, while the nodes in the exterior surface of the tube are not allowed to penetrate the exterior surface. This example demonstrates the performance of the global contact search algorithm presented in Section 3.4.1.2, which was specifically developed to deal with self-contact problems. The adopted contact search algorithm is combined with the two-pass NTS contact formulation to solve contact problems involving self-contact, avoiding the specification of contact pairs.

The post-buckling geometry and the corresponding equivalent plastic strain distribution (four time steps) are shown in Figure 6.46, for half height of the tube. Note that the buckling occurs without introducing any initial geometric imperfection. The presented results are obtained with the smoothed master surface description. The numerical treatment of this problem using symmetric conditions in the middle of the tube height (see Figure 6.45) requires the introduction of additional contact boundary conditions to prevent the penetration of the tube wall into the plane of symmetry. Therefore, a rigid contact surface is introduced in the plane of symmetry, which imposes the necessary impenetrability conditions for the last buckle, as illustrated in the bottom of Figure 6.46 (d). The predicted number of buckles after the progressive folding of the tube (Figure 6.46) is exactly the same reported by [Yang 08b] and the shape is very similar. The maximum value of plastic strain is reached at the final state, in the interior of the buckles, which is approximately 57% for the last buckle. In fact, the entire tube comprises plastic deformation, with particularly the buckles with large strains.

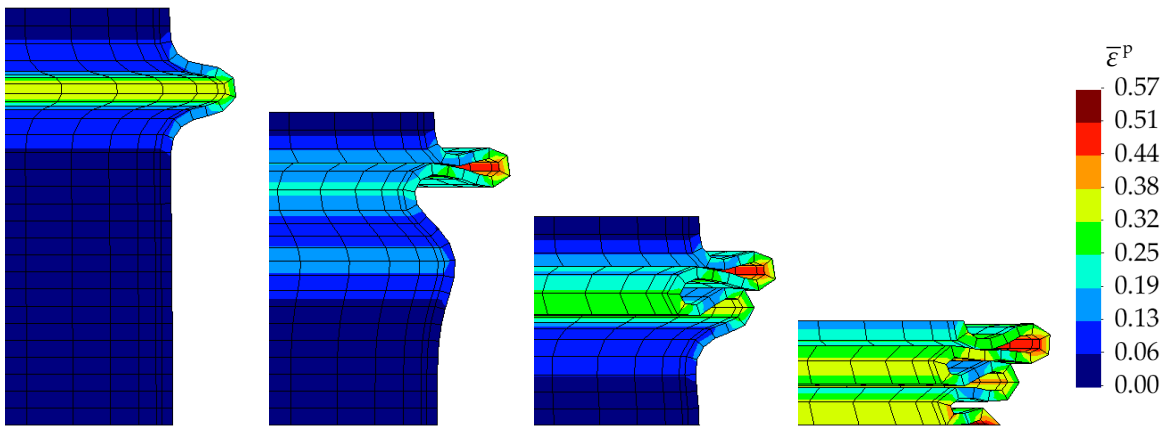


Figure 6.46. Post-buckling geometry and equivalent plastic strain distribution (smoothed coarse mesh) for different values of total displacement: (a) 10 mm; (b) 20 mm; (c) 30 mm; (d) 40 mm.

The total axial force induced in the tube as function of the total (top and bottom) axial displacement is presented in Figure 6.47. Moreover, the numerical results are compared with the ones obtained by [Yang 08b], which uses the coarse mesh to discretize the tube. The sequence of buckle cycles is clearly apparent in the curve, as well as the contribution of the self-contact. The buckling (indicated by the drop in force) occurs when a critical axial load is reached. The self-contact phenomena are illustrated as a reversal of force at the bottom of each buckle cycle, which provides the stiffening mechanism needed to activate the next buckle. The reaction force evolution predicted by the present model is identical to the one predicted by [Yang 08b] for the first buckle cycle, as shown in Figure 6.47. Nevertheless, the results reported by [Yang 08b] present a different buckling sequence, i.e. the first buckle cycle occurs simultaneously in both extremities of the tube, while the next buckle cycles arise alternating between the two extremities of the tube. Indeed, this behaviour cannot be reproduced using only half length of the tube (present model). Therefore, in order to validate the presented results, the simulation of the entire tube length was performed (one quarter model), which provides exactly the same results obtained with the presented model (one eighth of the tube). In fact, there is no reason for the asymmetry observed in the results presented by [Yang 08b]. However, since the post-buckling problems are very sensitive to the mesh configuration, some instabilities related with the contact treatment can be in the origin of such asymmetry. Somehow, the global trend of the reaction force evolution is similar in both models, as presented in Figure 6.47.

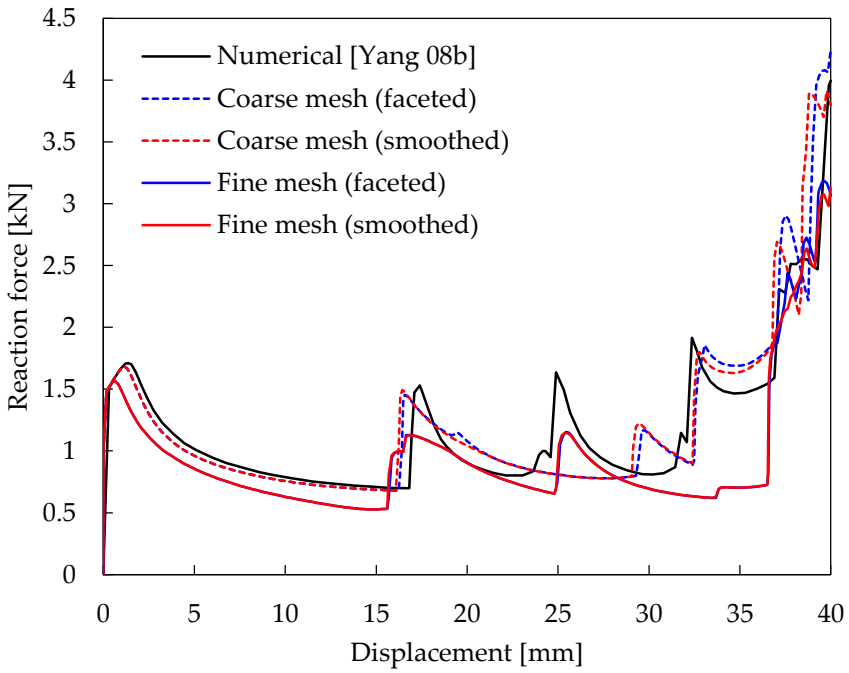


Figure 6.47. Axial force evolution as a function of the displacement in the post-buckling problem, for both faceted and smoothed master surface descriptions.

The smoothing of the master surface with Nagata patches induces the buckling slightly sooner than does the piecewise bilinear representation of the master surface. This offset in the curves due to the contact surface representation is more evident in the final state. Note that the deformation induces large plastic strains in the tube (see Figure 6.46), leading to a strongly nonlinear problem. In order to evaluate the influence of the tube discretization in the numerical results, a fine mesh is also created, where the one eighth of the tube is discretized by 4,800 solid finite elements (8-node hexahedral). The finite element mesh is created using 40 elements in the circumferential direction, 60 in the axial direction and two layers of elements through the thickness. The axial force evolution as a function of the total (top and bottom) axial displacement, obtained with the fine mesh, is shown in Figure 6.47. Globally the force is lower and the instants at which self-contact occurs are different. The post-buckling geometry and the corresponding equivalent plastic strain distribution (four time steps) are shown in Figure 6.48, for the fine mesh with smoothed master surface. The sequence of post-buckling shapes obtained in the simulation with the fine mesh is substantially different, as shown through the comparison with Figure 6.46. Typically, the finite element mesh refinement reduces the critical axial force necessary to produce buckling, as shown in Figure 6.47. Moreover, the application of the surface smoothing method to describe the master surface presents a small influence in the numerical solution (force evolution curves are almost coincident) because the finite element mesh at contact interface is relatively fine (see Figure 6.48).

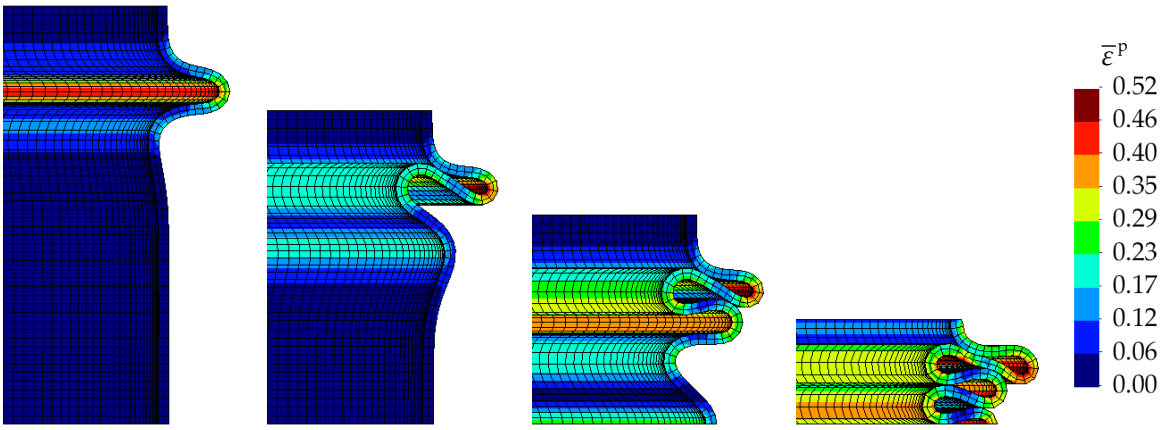
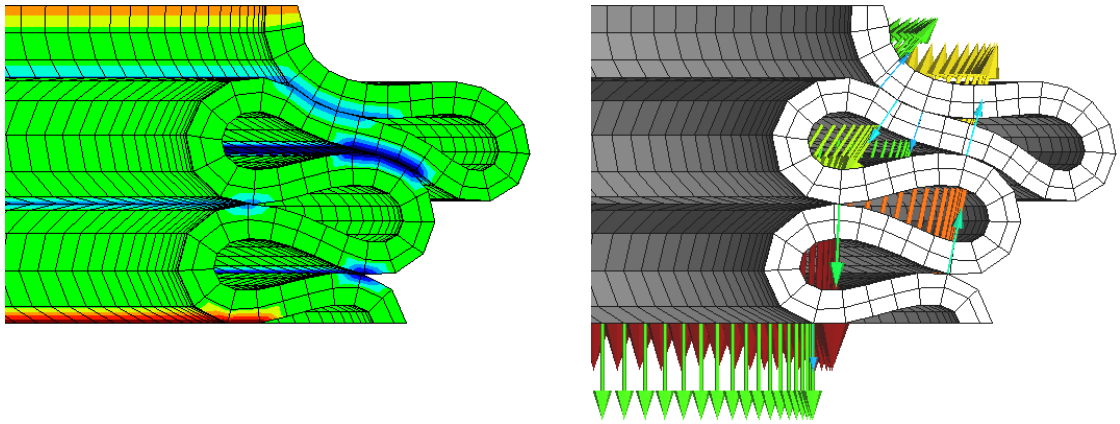


Figure 6.48. Post-buckling geometry and equivalent plastic strain distribution (smoothed fine mesh) for different values of total displacement: (a) 10 mm; (b) 20 mm; (c) 30 mm; (d) 40 mm.

The maximum value of accumulated equivalent plastic strain predicted by the fine mesh model occurs in the interior of the first buckle (see Figure 6.48 (d)), which is approximately 52%. Due to the better geometrical description of the folding of the tube

(buckles), globally the equivalent plastic strain provided by the fine mesh is lower than the one obtained when adopting the coarse mesh. Although the number of buckles at the final state is the same for both finite element meshes (see Figure 6.46 (d) and Figure 6.48 (d)), the instants at which self-contact occurs are different, as well as the post-buckling shape. At 30 mm of axial displacement there is no vertical tube wall portion when using the fine mesh (Figure 6.48 (c)), while the coarse mesh still has a small vertical wall, as shown in Figure 6.46 (c).



(a) (b)

Figure 6.49. Deformed configuration of the tube at the final state using the smoothed fine mesh: (a) potential contact surfaces denoted in blue; (b) nodal contact forces distribution.

The potential contact surfaces in which self-contact can occur, at the final state, are denoted in blue in Figure 6.49 (a), as determined by the global contact search algorithm. This algorithm is based in the nodal normal vectors and the maximal detection distance (see Section 3.4.1.2). Since both the inner (light blue) and the outer (dark blue) surfaces of the tube are predefined as self-contact surfaces, four potential contact interfaces are identified by the contact search algorithm, two in each self-contact surface. Moreover, the potential contact of the last buckle with the horizontal plane of symmetry is indicated in red in Figure 6.49 (a). The orange in the top of the tube indicates the nodes with prescribed axial displacement, which is performed in this example through a rigid surface glued to the nodes. Since the two-pass NTS contact algorithm is adopted in self-contact problems, the Node-to-Nagata contact elements are created in all nodes belonging to the potential contact surfaces.

The nodal contact forces arising at the final state of the post-buckling problem are shown in Figure 6.49 (b). The value of the contact force indicated in the nodes located in the symmetry plane is half of its real value due to the symmetry conditions. Since friction is not taken into account, the direction of the nodal contact force is given by the master surface normal vector at the contact point. Note that the contact is detected only within the potential

contact surfaces. Nevertheless, only some nodes present contact force in order to satisfy the impenetrability conditions, as shown in Figure 6.49. The last buckle is in contact with the rigid surface introduced in the horizontal plane of symmetry, producing the nodal contact forces shown in the bottom of Figure 6.49 (b).

Table 6.8. Computational performance of the post-buckling problem for two distinct finite element meshes.

	Coarse mesh		Fine mesh	
	Faceted	Smoothed	Faceted	Smoothed
Nº increments	269	270	260	258
Nº iterations	1,712	1,601	1,777	1,748
Computational time [min]	0.72	0.63	17.2	17.4

The computational performance of the post-buckling problem is presented in Table 6.8. The number of increments used to apply the axial displacement is approximately 270 for the coarse mesh and 260 for the fine mesh, which is a moderate number for this strongly nonlinear problem. Note that the smoothing of the master surfaces with Nagata patches has no effect on the increment size. Concerning the computational time, the numerical analysis is carried out in less than 1 minute when adopting the coarse mesh, while the fine mesh requires about 17 minutes to complete the simulation. The effectiveness of the surface smoothing method is more apparent in the coarse mesh than in the fine mesh. In fact, the computational time is approximately the same for both surface description methods (faceted and smoothed) when the fine mesh is adopted, as shown in Table 6.8. The same conclusion is extracted for the total number of iterations. On the other hand, the application of the surface smoothing in the coarse mesh yields a reduction of about 10%, both in terms of the total number of iterations as well as on the computational time. The average number of iterations required to achieve convergence in each increment is approximately the same for both models studied, which ranges between 6 and 7 iterations (see Table 6.8). Indeed, this numerical example presents poor convergence only when the onset of self-contact occurs, which requires the increment size reduction to overcome convergence problems.

6.2.6. Deep drawing using deformable tools

The last example presented in this dissertation involves the deep drawing of a cylindrical cup, taking into account the elastic deformation of the forming tools in the numerical model. The process parameters and tool dimensions adopted in this example are exactly the same employed during the first stage of the reverse deep drawing process considering rigid tools, presented in Section 6.1.3. The circular blank has 170 mm of

diameter and 0.98 mm of initial thickness, while the main dimensions of the forming tools are given in Table 6.2. The elastoplastic material properties of the blank (mild steel DDQ) are listed in Table 6.3. On the other hand, the elastic properties of the tools are identical to the ones adopted for the blank (see Figure 6.50). Due to geometric and material symmetry conditions, only one quarter of the model is simulated and all bodies (blank and tools) are discretized with 8-node hexahedral solid finite elements. The number of finite elements used in each body is shown in Figure 6.50. The discretization of the blank used in this example is the same previously adopted considering rigid tools (Section 6.1.3).

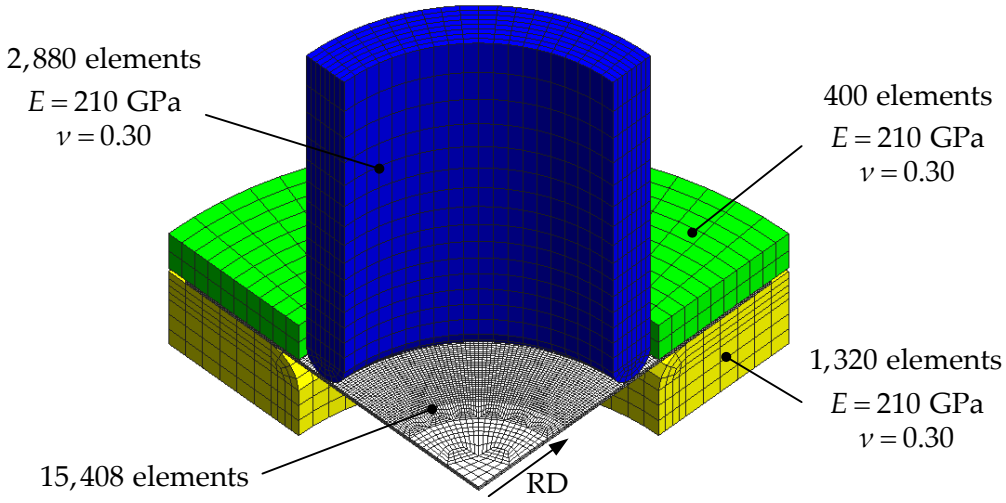


Figure 6.50. Finite element mesh of the forming tools (punch, blank-holder and die) and blank used in the deep drawing of a cylindrical cup.

The dimension of the tools used in this example, namely the thickness of the blank-holder (10 mm) and the die were defined from the experimental device, presented in [Thuillier 02]. The gap between the die and the blank-holder is keept constant by using 8 screws regularly space with adjustable keys. The gap value of 1.0 mm is determined as the largest possible to draw a cup without wrinkles [Thuillier 02]. The external diameter of the die and the blank-holder considered in the model (see Figure 6.50) is defined as the one where the screws are applied (190 mm). Thus, the nodal displacements of all nodes located on this diameter are fixed. The vertical displacement (50 mm) imposed to the punch is applied to its top surface.

The comparison between experimental and numerical punch force evolutions is presented in Figure 6.51, showing numerical results considering both rigid and deformable tools. In case of deformable tools, the gap between the die and the blank-holder is set equal to 1.00 mm (experimental value), which is prescribed in the external diameter of the die and the blank-holder. On the other hand, considering rigid tools, two distinct values of fixed gap are analysed and compared (1.13 mm and 1.10 mm). The experimental punch force

evolution is underestimated by the numerical simulation when rigid tools are considered, as shown in Figure 6.51. Moreover, the decrease of the gap (towards the experimental value) leads to an abrupt increase of the punch force at approximately 35 mm of displacement, which is a consequence of the ironing of the flange. On the other hand, the experimental punch force evolution is overestimated when the elastic deformation of the tools is taken into account in the numerical model. However, the adoption of the isotropic hardening law for the mild steel can contribute to this overestimation. In fact, typically the punch force decreases when taking also into account kinematic hardening (i.e. the Bauschinger effect) [Oliveira 07].

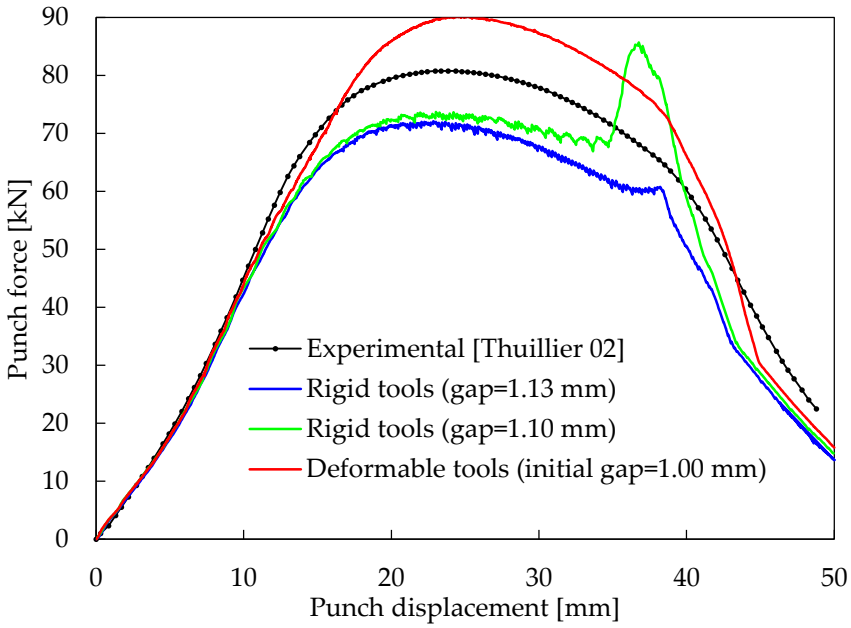
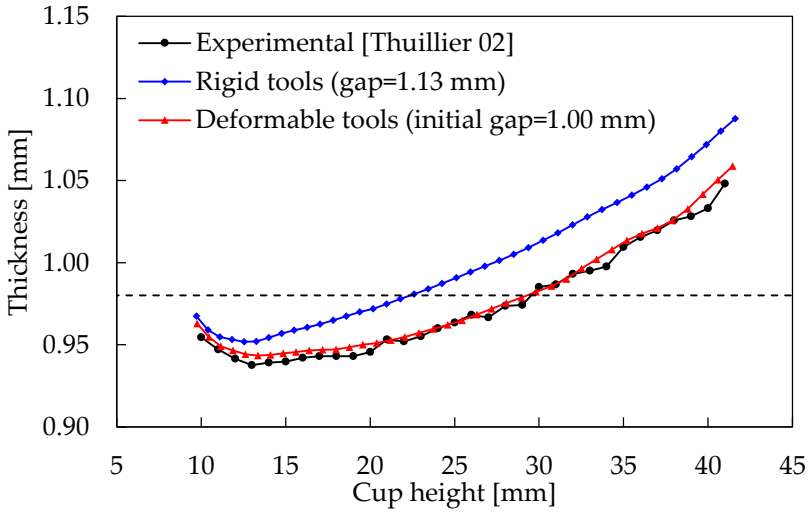
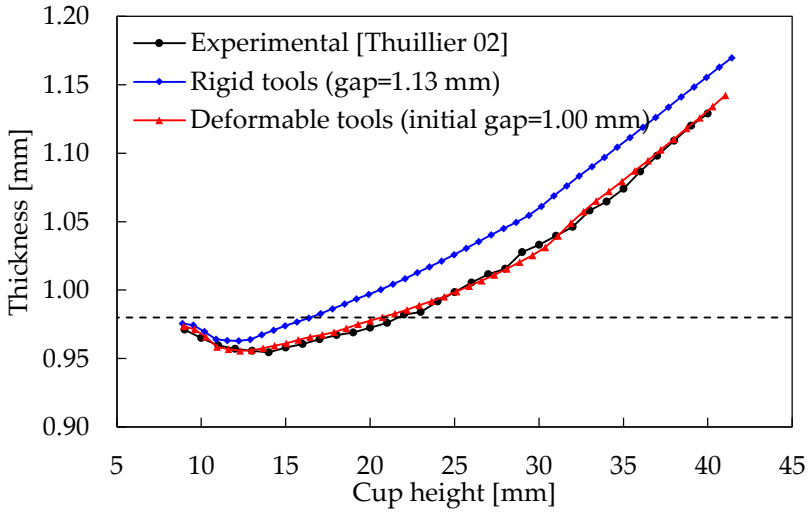


Figure 6.51. Comparison between experimental and numerical punch force evolutions in the deep drawing of a cylindrical cup.

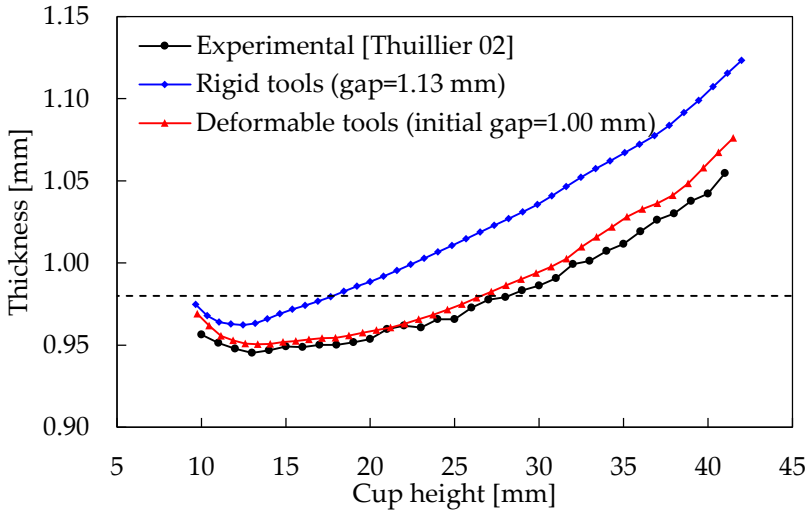
The comparison between experimental and numerical cup wall thickness distributions along the cup height at 0°, 45° and 90° to the rolling direction is presented in see Figure 6.52. Both the experimental and the numerical thickness distributions are calculated in the radial direction, always perpendicular to the cylindrical cup axis. The thickness distribution is overestimated in the three directions when considering rigid tools in the numerical simulation. The thickness is only slightly affected by the selected value for the gap (1.13 mm and 1.10 mm) because the decrease of the gap leads to the necking of the sheet in the punch radius. On the other hand, the incorporation of the elastic deformation of the tools in the numerical model leads to a numerical thickness distribution in very good agreement with the experimental results for all directions, as shown in Figure 6.52.



(a)



(b)



(c)

Figure 6.52. Experimental and numerical thickness distributions along the cup height measured in the: (a) rolling direction; (b) diagonal direction; (c) transverse direction.

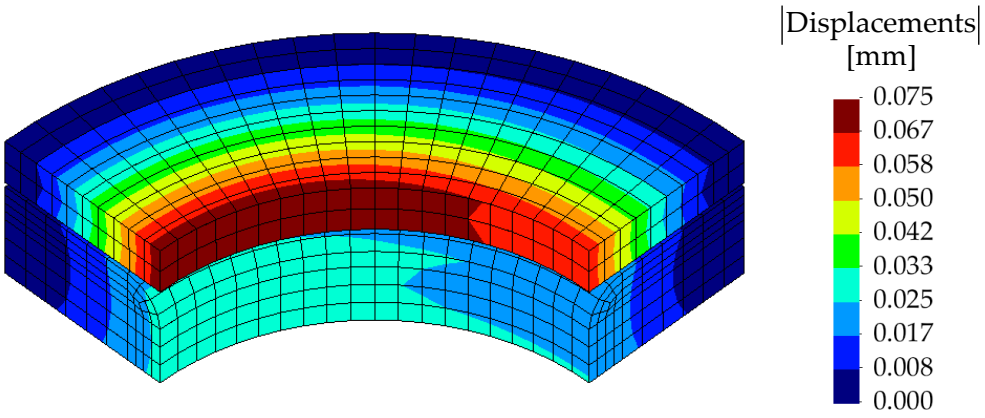


Figure 6.53. Contour plot of the nodal displacements magnitude in the die and blank-holder for 30 mm of punch displacement.

In order to quantify the deformation of the forming tools, predicted by the numerical simulation, the contour plot of nodal displacements in the die and blank-holder is presented in Figure 6.53, for 30 mm of punch displacement. The higher value of displacement occurs in the blank-holder opening diameter, which is approximately 0.075 mm. The nodal displacements of the die are around half the blank-holder ones for the same radial coordinate, due to its larger thickness (see Table 6.2). The deformation of the tools (mainly in the vertical direction) is caused by the increasing thickness of the flange resulting from the circumferential compression stress state (uniaxial compression strain path). Moreover, the anisotropy of the sheet is reproduced in the tools deformation distribution, as shown in Figure 6.53. Since the predicted thickness distributions is higher in the transverse direction (see Figure 6.52), the displacement of the die and the blank-holder is also higher along this direction.

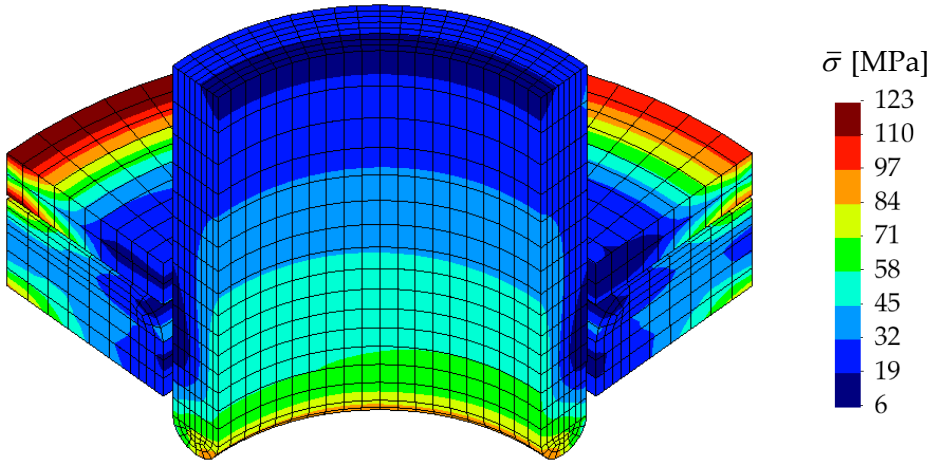
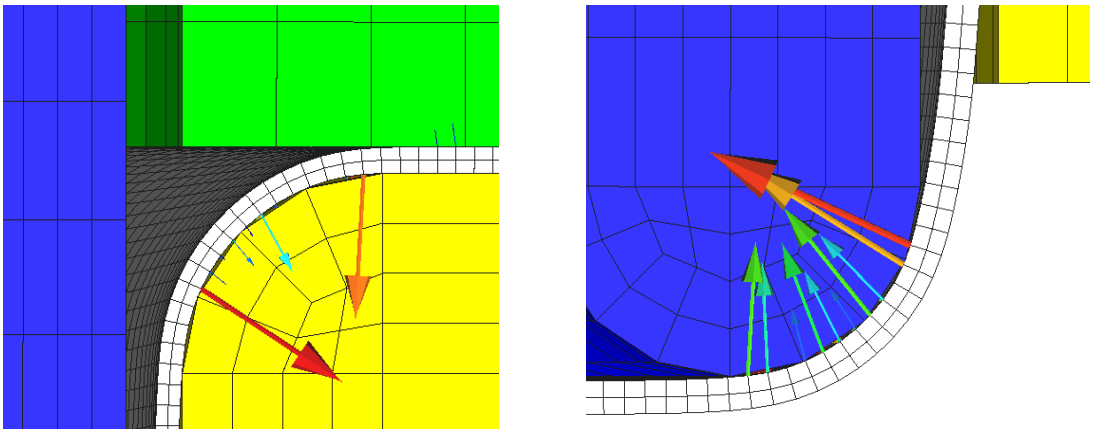


Figure 6.54. von Mises stress distribution in the forming tools for 30 mm of punch displacement.

The von Mises stress distribution in the forming tools (punch, blank-holder and die) is presented in Figure 6.54, for 30 mm of punch displacement. Due to the boundary conditions applied in the external diameter of the die and the blank-holder, the maximum value of stress arises in that area. Nevertheless, the low value of the equivalent stress indicates that only elastic deformation occurs in the tools, assuming that they are built with a convectional High-Speed Steel (HSS). As observed in the contour plot of the nodal displacements (Figure 6.53), the anisotropy of the sheet dictates a stress distribution on the tools that is not axisymmetric, as shown in Figure 6.54, predominantly in the blank-holder.



(a) (b)

Figure 6.55. Nodal contact forces in the slave nodes for 30 mm of punch displacement (magnitude denoted by arrow size and colour): (a) contact between sheet and die as well as between sheet and blank-holder; (b) contact between sheet and punch.

For the model that takes into account the deformation of the tools, the sheet is defined as slave body while the forming tools are assigned as master bodies (following the guidelines presented in Section 3.3.1.1). Moreover, the master surfaces are smoothed with Nagata patches, allowing an accurate description of the curved contact surfaces describing the tools. The nodal contact forces arising in the slave nodes, aligned with the rolling direction, are presented in Figure 6.55, for 30 mm of punch displacement. The contact occurs mainly in the curved zones of the tools, where the surface smoothing method is more effective. In fact, despite the coarse mesh adopted in the discretization of the tools, the contact forces are properly distributed on the smoothed surface, dictating a smooth evolution of the punch force, as shown Figure 6.51. In fact, although there is an apparent gap between the die and the sheet (Figure 6.55 (a)) and between the punch and the sheet (Figure 6.55 (b)) in the curved contact zones, the use of Nagata patches in the tool surface, allows recovering their curvature.

Concerning the computational performance of the sheet metal forming using deformable tools, the computational time is considerably higher in comparison with the

model that uses rigid tools. Indeed, the computational cost increases from 1.5 hours considering rigid tools to approximately 22 hours, when taking into account the elastic deformation of the forming tools in the numerical model. This increase is motived by several factors: (i) total number of degrees of freedom increases due to the discretization of the tools with solid elements; (ii) the nonzero pattern of the global tangent matrix is updated in each increment due to large sliding and (iii) the increment size has to be smaller, since the contact problem between deformable bodies is more non-linear. In fact, although the deformation of the forming tools is only linear, they are modelled with an elastoplastic material law, where the initial yield stress is $\sigma_0 = 1500\,\text{MPa}$ and a linear hardening rate is adopted.

Chapter 7

Conclusions and Future Perspectives

The accurate analysis of many physical systems requires the description of the mechanical interaction across contact interfaces. Some important industrial examples in the world range from the mechanical and civil engineering to the design of prosthetics in biomedical engineering. The principal objective of this thesis is the development of computational methods for the 3D treatment of frictional contact problems between deformable bodies undergoing large sliding, adopting the finite element method. Three main areas of the computational contact mechanics have been explored: (i) description of the contact surfaces; (ii) contact detection including self-contact and (iii) resolution procedure with the coupled augmented Lagrangian method. The developed original algorithms and the improved standard techniques have been implemented in the in-house finite element code DD3IMP. This chapter contains the principal conclusions of the present research work and provides some perspectives for future work.

7.1. Conclusions

The departure point for the work presented in this dissertation was the in-house finite element code DD3IMP, which has been specifically developed to simulate sheet metal forming processes. Since the forming tools are assumed as rigid, the contact was established between a deformable body (sheet) and various rigid obstacles (forming tools), which are modelled by Bézier patches. The extension of the finite element code capability to deal with frictional contact problems between deformable bodies is achieved in this work. Therefore, the range of application covered by the numerical simulations performed with DD3IMP increases significantly, due to the developments carried out in the present study. Moreover,

the limitation related with the definition of the Bézier patches from a standard STEP file generated with a specific CAD software is eliminated, by means of the new surface description method using Nagata patches. This is particularly important in industrial applications, where the complexity of the tools geometry is considerably high, typically involving hundreds of patches.

The Node-to-Segment discretization of the contact interface associated with the master-slave approach allows to deal with non-conforming meshes at the contact interface, large deformations and large sliding. Since the contact constraints are enforced only at the slave nodes, this discretization technique does not pass the contact patch test for non-conforming meshes using the traditional single-pass approach. However, the two-pass Node-to-Segment contact algorithm passes the contact patch test for some structured mesh configurations using low order finite elements, while the classification of the master and slave surfaces is arbitrary. Besides, the adopted discretization technique requires an efficient contact detection procedure to establish for each slave node the closest point on the master surface. The master segment defined by this point and the slave node constitutes the contact element. In order to obtain a smooth and accurate description of the master surface, a 3D contact surface smoothing method is developed based in the Nagata patch interpolation. It can be easily applied to both structured and unstructured surface finite element meshes, because the Nagata interpolation is completely local and requires only the surface normal vectors at the nodes. In fact, the principal advantage of the presented surface smoothing method over the existing ones is its ability to deal with arbitrary surface meshes. Moreover, the local support of the interpolation allows to combine triangular and quadrilateral Nagata patches in the description of a single contact surface.

The accuracy of the surface smoothing method is evaluated using simple geometries defined by analytical functions. Two types of error are analysed, the shape of the interpolated surface and the deviation in the surface normal vector, which are directly connected with the accuracy of the normal gap and the direction of the nodal contact forces, respectively. Considering a discretized circular arc, the maximum value of the radial error decreases quadratically with the normalized arc length using linear interpolation, while applying Nagata interpolation the convergence rate is quartic. Moreover, the maximum value of the normal vector error decreases linearly with the normalized arc length when the linear interpolation is adopted, while the Nagata interpolation method provides a cubic convergence rate. Identical convergence rates are obtained in 3D geometries adopting structured meshes in the surface discretization, such as cylinders, spheres and tori. In the cylindrical geometry, the error distribution is independent of the number of finite elements in the axial direction being dictated by the discretization in the circumferential direction. On the other hand, in the torus geometry, the equilibrium between number of patches and smoothed surface accuracy is given for approximately square finite elements aligned with the toroidal and poloidal directions. The adoption of unstructured meshes in the surface

discretization deteriorates slightly the surface accuracy. All these results confirm the weakness of the usually piecewise bilinear finite element mesh representation, both in terms of shape and normal vector distribution. Furthermore, the advantages of the contact surface smoothing with Nagata patches are highlighted. Indeed, this surface smoothing method allows to recover the original curvature with a relatively coarse surface discretization, ensuring *quasi*-G^1 continuity in patch boundaries, which is a vital feature in contact problems involving large sliding.

In the particular case of contact between a deformable body and a rigid obstacle, the master rigid surface is described by Nagata patches associated to each facet of the finite element mesh. Since the finite element mesh of the rigid surface is typically generated from a CAD model, the required nodal normal vectors are evaluated using the information contained in the IGES file, where the surface geometry is represented by trimmed NURBS surfaces. The finite element nodes are projected onto the corresponding trimmed NURBS surfaces, calculating the normal vector based on the cross product of the first-order partial derivatives, evaluated in the projection point. This strategy allows to accurately evaluate the nodal normal vectors, yielding an accurate Nagata interpolation. In fact, the interpolation accuracy is directly influenced by the adopted nodal normal vectors, which define the boundary conditions of the patches at the nodes.

In the general case of contact between deformable bodies, the master contacting surface is represented by a piecewise bilinear finite element mesh formed using the exterior of linear solid finite elements. Therefore, the nodal normal vectors required for the surface interpolation are approximated using the information provided by the surface mesh. The approach adopted in the present study consists in estimating the nodal normal vector using the weighted average of the normal vectors of facets adjacent to the node. Six different weighting factors were implemented and their accuracy is evaluated and compared by means of simple and complex geometries. Globally the MWSELR weighting factor provides the best approximation, particularly for spherical surfaces. Since the achieved level of accuracy is lower in some specific locations, mainly in the intersection between surfaces and in the boundaries, the nodal normal vector are adjusted in these locations to reduce the approximation error. The nodal normal vectors in the intersection between flat and curved geometries are adjusted for the normal vector of the flat surface, eliminating the error in the approximation. An identical approach is applied in the nodes located in symmetry planes, frequently used in the finite element method to reduce the computational time. In fact, the application of the proposed correction methodology improves significantly the accuracy of the nodal normal vector approximation, leading to clearly better results than any of the weighting algorithms analysed, particularly in coarse meshes.

The normal gap and the tangential relative sliding are the kinematic contact variables relevant for measuring relative positions and motions of deformable bodies, which are evaluated only in the slave nodes, when adopting the Node-to-Segment discretization.

Their definition is based in the normal projection of the slave nodes on the discretized master surface, described by Nagata patches. The main drawback of the Node-to-Segment is the asymmetric treatment of the bodies. The blind spots in the projection domain or multiple solutions arising in the piecewise bilinear representation of the master surface are strongly reduced with the adopted surface smoothing method. The developed contact search algorithm, which defines the contacting pairs on the discretized surfaces (slave node and master segment), is decomposed in two distinct phases: global and local search. The purpose of the global contact search is to minimize the number of operations performed by the local search algorithm, which is usually computationally more expensive. Concerning the contact of a deformable body against a rigid obstacle, the developed global contact search algorithm is inspired in the HITA algorithm. It is divided in three steps: (i) selection of an amount of closest master nodes and identification of their neighbouring patches; (ii) creation of a uniform grid of points on the patches and (iii) selection of the closest Nagata patches based on the grid of points. Since both the amount of master nodes and the grid dimension are calculated taking into account the surface mesh topology (irregular distribution of finite elements), the proposed algorithm is efficient and robust. Concerning the contact between deformable bodies, since typically their discretization does not involve distorted finite elements, the global search algorithm developed for this type of problems is based on the search for the closest master node and subsequent selection of the Nagata patches sharing such node. The global search algorithm for problems involving self-contact is implemented considering that the contact surface pairs (master-slave) are unknown *a priori*. It is developed taking into account the orientation of the nodal normal vectors to determine potential contacting surfaces, avoiding the contact with the reverse side in case of thin walled structures. The detection is performed only within a cylinder with its axis coincident with the normal vector of the contact node and its centre on the node. The key point of the local contact search procedure is the closest point projection solved numerically through the Newton–Raphson method. The quadratic definition of the Nagata patches yields a stable iterative solution, converging in few iterations.

The regularization of the non-differentiable impenetrability and friction conditions is provided by the augmented Lagrange method. This technique leads to the saddle point problem, where the objective functional is minimized by primal variables (displacements) and maximized by dual variables (Lagrange multipliers that represent contact forces). Its main advantage is the smooth objective functional and the transformation of the optimization problem with inequality constraints into a fully unconstrained problem. Moreover, the frictional contact constraints are exactly fulfilment with a finite value of the penalty parameter. The nonlinear system of equations resulting from the finite element method application is solved using the generalized Newton method, with a simultaneous update of both primal and dual variables. The solution of a linear system of equations in

each Newton iteration is obtained with the Direct Sparse Solver (DSS) from the Intel® Math Kernel Library.

The developed Node-to-Nagata contact element allows to handle large sliding over curved master surfaces. The geometrical part of the contact element is defined by a slave node and the associated Nagata patch (triangular or quadrilateral). Since the augmented Lagrangian method is adopted to deal with the frictional contact, each contact element is complemented by an artificial node to store the contact force (Lagrange multipliers). The elemental contribution of the new contact elements to the virtual work of the system and its linearization are derived, coherently with the surface smoothing approach. The residual vectors and tangent matrices are presented. In the particular case of frictional contact with a rigid surface, the contact forces arising in the slave nodes are not transferred to the master surface. Thus, the nonzero pattern of the global tangent matrix is not modified by the introduction of the contact elements, even when sliding of the slave nodes over the master Nagata patches occurs. This means that nonzero structure of the matrix is kept constant during all the increments, improving the computational performance of the numerical simulation. The reduced system of equations, which involves only the degrees of freedom related with the nodal displacements, is derived from the mixed system and presented. Although the size of global tangent is lower than in the mixed system of equations, the assembly of the contact contributions to the global tangent matrix is much more complex. On the other hand, the contact between deformable bodies (including self-contact phenomena) comprises the transfer of the contact forces arising in the slave nodes to the master body, according with the impenetrability and friction conditions. This dictates that nonzero pattern of the global tangent matrix should be always adjusted when the Nagata patch composing a contact element is updated due to sliding. Since this procedure is quite computationally expensive, multi-face contact elements are created when large sliding is present in the contact problem. The idea behind the multi-face contact element is adding the adjacent master patches during the creation of the nonzero pattern of the global tangent matrix. This allows updating the nonzero structure of the global tangent matrix only at the beginning of the each new increment, as adopted in the present study.

The accuracy, robustness and performance of the proposed contact surface smoothing method is validated through several numerical examples. The selected examples are separated into two categories: contact of a deformable body against rigid surfaces and contact between deformable bodies. The accuracy of the developed Node-to-Nagata contact element is shown by means of problems with available analytical solutions, where the obtained numerical results are in very good agreement with the exact results. In frictional contact problems involving large sliding, the comparison between the faceted and the smoothed surface description methods highlight the superior accuracy of the numerical results obtained with the master surface described by Nagata patches. Indeed, the non-physical oscillations in the contact force evolution arising from the piecewise bilinear

surface representation are eliminated or strongly reduced with the surface smoothing method. Moreover, the convergence problems are considerably reduced, improving the computational performance of the numerical simulation. The importance of the developed algorithms for industrial applications is shown by means of two real metal forming problems, namely the reverse deep drawing of a cylindrical cup and the automotive underbody cross member panel. In order to achieve approximately the same level of accuracy in the representation of the forming tools, the traditional piecewise bilinear surface requires 10 times more facets than Nagata patches in the first problem, and about 3 times more, for the second problem. Moreover, the computational time of the numerical simulation using smoothed tool surfaces is approximately half in comparison with the faceted description for the first problem and about 6 time lower for the second industrial problem. The good correlation between numerical and experimental results is presented for the automotive underbody cross member. The validation of the global contact search algorithm specifically developed for self-contact problems is performed using the post-buckling of a thin walled tube. The accuracy and efficiency of the global contact search combined with the two-pass Node-to-Nagata contact formulation is highlighted in this simple example. The last example presented in this dissertation, which deals with the sheet metal forming of a cylindrical cup taking into account the elastic deformation of the tools, shows that an accurate modelling of the forming tools behaviour is mandatory for attaining reliable numerical results. Since the elastic deformation of the blank-holder is important in the presented forming process, only the numerical results obtained considering deformable tools are in good agreement with the experimental ones.

7.2. Future Perspectives

The developed algorithms and its implementation in the finite element code DD3IMP allows to envisage some new challenges found in the field of the computational contact mechanics. Indeed, the extension of the existing finite element code to handle contact between deformable bodies (including self-contact) is the main contribution of the present work to enlarge the range of application currently covered by DD3IMP. Moreover, the accuracy of the proposed 3D contact surface smoothing method and its stability in contact problems involving large sliding are vital for the performance of the numerical simulations. Based in the knowledge acquired throughout this thesis, some perspectives concerning the continuity of the present work and future research are described.

The connection of non-conforming finite element meshes in nonlinear solid mechanics using the developed Node-to-Nagata contact element can be useful to reduce the computational time of the numerical simulation, as well as the time necessary to build the finite element model. This strategy allows to perform mesh refinement easily because the

subdomains to be coupled can be meshed independently. Besides, different types of finite elements can be connected, namely hexahedral and tetrahedral finite elements, in order to combine the advantages of each one. Due to the accuracy achieved by the surface smoothing method, the interface between the non-conforming meshes can be either flat or curve. The basic idea behind this strategy is the creation of contact elements at the interface of the subdomains, imposing the stick contact status during all simulation.

Concerning the sheet metal forming processes, the optimum tool design can be currently achieved using the numerical simulation considering rigid tools. Since the number of patches required to obtain an accurate surface description is considerably inferior to the number of facets used by the traditional piecewise bilinear surface representation, the number of variables involved in the optimization procedure is smaller. Moreover, the robustness provided by the smooth tool surface description allows to perform the optimization of complex geometries in a smaller amount of time without user intervention. The compensation of the tools shape is typically dictated by the springback predicted by the numerical simulation. However, the accurate simulation of springback in high strength steels requires the incorporation of the elastic properties of the tools into the finite element model. Thus, the implementation of contact elements to deal with contact between deformable bodies involving large sliding is essential for the accurate simulation of such forming processes.

The modelling of the friction at the contact interface is very important for several technical applications. The friction is assumed to be constant in most of the simulations, including in the present study. However, in general, the friction coefficient depends upon the normal contact pressure, the relative tangential velocity, the surface roughness, the temperature, among other parameters. Therefore, an accurate analysis of the frictional contact demands for more advanced friction laws in order to take into account these effects. The development of multi-scale contact mechanic models to describe the contact interaction should be take into account in future investigations to improve the knowledge concerning the contact condition. Regarding the roughness of contact surfaces, the contact elements developed allows to perform the 3D analysis of the frictional contact between two deformable bodies with rough surfaces. Different asperity shapes and geometrical characteristics can be studied using representative rough surfaces, as well as the interaction between asperities during the relative sliding.

Appendix A

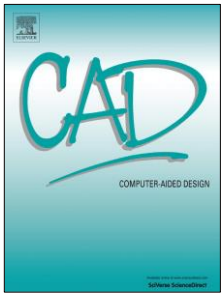
List of publications

Some parts of this thesis were included in scientific contributions that have been published in international scientific journals as well as in conference proceedings, including the oral presentation of the main results. The list of publications performed in the scope of this thesis, which summarises the results of the scientific research performed, is given in this appendix. Note that all the scientific work related with the frictional contact between deformable bodies, including the master surface smoothing with Nagata patches, is not yet published.

A.1. Scientific journals

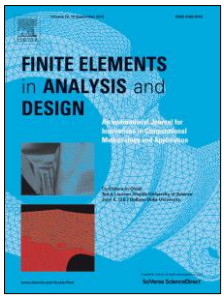
The dissemination of the scientific results has been done in international journals with referees. Besides, the four selected journals are indexed in ISI Web of Science and SciVerse Scopus, contributing for the diffusion of the obtained results.

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F. AND ALVES, J.L., 2013. Improving Nagata patch interpolation applied for tool surface description in sheet metal forming simulation. *Computer-Aided Design*, 45(3), pp. 639–656. http://dx.doi.org/10.1016/j.cad.2012.10.046



5-Year Impact Factor: 1.929 (© Thomson Reuters)

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F. AND ALVES, J.L., 2013. Nagata patch interpolation using surface normal vectors evaluated from the IGES file. *Finite Elements in Analysis and Design*, 72, pp. 35–46. http://dx.doi.org/10.1016/j.finel.2013.03.004



5-Year Impact Factor: 1.679 (© Thomson Reuters)

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F. AND ALVES, J.L., 2014. Applying Nagata patches to smooth discretized surfaces used in 3D frictional contact problems. *Computer Methods in Applied Mechanics and Engineering*, 271, pp. 296–320. http://dx.doi.org/10.1016/j.cma.2013.12.008

5-Year Impact Factor: 3.049 (© Thomson Reuters)

NETO, D.M., OLIVEIRA, M.C., ALVES, J.L. AND MENEZES, L.F., 2014. Comparing faceted and smoothed tool surface descriptions in sheet metal forming simulation. *International Journal of Material Forming*, pp. 1–17. In press. http://dx.doi.org/10.1007/s12289-014-1177-8

2013 Impact Factor: 1.418 (© Thomson Reuters)

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F. AND ALVES, J.L. A contact smoothing method for arbitrary surface meshes using Nagata patches. *Computer Methods in Applied Mechanics and Engineering*. In Preparation.

5-Year Impact Factor: 3.049 (© Thomson Reuters)

A.2. Conference proceedings

All papers published in conference proceedings are full-length papers, which were accompanied by the oral presentation conducted by the first author.

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F. AND ALVES, J.L., 2011. Formulation of frictional contact problems using tools described by Nagata patches. In A. Tadeu et al., eds. *CMNE 2011, Congresso de Métodos Numéricos e Computacionais em Engenharia*. Coimbra, pp. 1–20.

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F. AND ALVES, J.L., 2012. Evaluating the vertex normal vector in polyhedral mesh using the IGES file format. In A. Andrade-Campos, N. Lopes, R.A.F. Valente, & H. Varum, eds. *YIC 2012, First ECCOMAS Young Investigators Conference on Computational Methods in Applied Sciences*. Aveiro, pp. 1–12.

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F. AND ALVES, J.L., 2012. FEA of frictional contact problems using Nagata patches for surfaces description. In P.M. Pimenta & E.M. Campello, eds. *WCCM 2012, 10th World Congress on Computational Mechanics*. São Paulo, pp. 1–20.

NETO, D.M., OLIVEIRA, M.C., ALVES, J.L., MENEZES, L.F. AND MANACH, P.Y., 2013. Applying Nagata Patches in the Description of Smooth Tool Surfaces Used in Sheet Metal Forming Simulations. *Key Engineering Materials*, 554-557, pp. 2277–2284

NETO, D.M., OLIVEIRA, M.C., MENEZES, L.F., ALVES, J.L. AND MANACH, P.Y., 2013. Influence of the Blank-holder Gap in the Numerical Simulation of the Reverse Deep Drawing Process. In J.M. Blanco, I. Arias, A. Peña, J.M. Guedes, N. Silvestre, & M. Silva, eds. *CMN 2013, Congreso de Métodos Numéricos en Ingeniería*. Bilbao, pp. 1–17.

Appendix B

IGES file format

This appendix gives a summary of the Initial Graphics Exchange Specification (IGES) file format, which is widely used to transfer information between CAD and CAE software packages. This standard exchange format contains the information required for the mathematical definition of the surface model geometry. Moreover, it is organized in a structured manner following the standard specification [IGES 96]. The file organisation is illustrated in this appendix by means of a simple example, highlighting the fundamental data unit within the file, called *entity*. Two main entity categories are distinguished in an IGES file: geometric and non-geometric entities. The parameters involved in the geometric entities necessary for the definition of trimmed NURBS surfaces are described in the following. This is particularly important for the development of routines to read IGES files, which require the knowledge of the data storage format.

B.1. IGES file of a simple geometry

The geometry selected to analyse the IGES file format is composed by one quarter of a cylinder and an eighth of a sphere, as illustrated in Figure B.1. Therefore, the surface model is defined by two trimmed NURBS surfaces (blue and green in Figure B.1). Indeed, both trimmed surfaces and theoretically untrimmed ones are represented in the IGES as trimmed entities. The IGES file of such surface model is presented in Figure B.2, which contains a total of 150 lines divided by five sections. Start section (1 line), global section (4 lines), directory entry section (60 lines), parameter data section (84 lines) and terminate section (1 line). This highlights that the IGES files of complex mechanical parts can be extremely large.

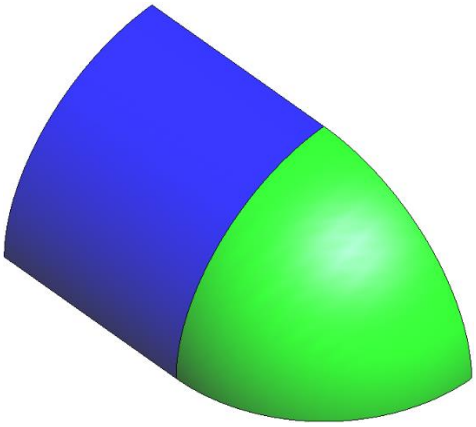


Figure B.1. Simple geometry composed by two trimmed NURBS surfaces.

```
21HTranslator GiD - IGES                                                S      1
1H,,1H;,3HGiD,33HC:/Users/user/Desktop/example.igs,3HGiD,6H10.0.9,32,38,G      1
6,308,15,,1.,2,2HMM,10,10.,13H140628.205450,1.00000000E-008,11.6302491,,G      2
65HCIMNE - International Center for Numerical Methods in Engineering,11,G      3
0,;                                                                     G      4
     314       1       0       0       0       0       0       000000200D      1
     314       0       1       1       0                   COLOR       1D      2
     406       2       0       0       1       0       0       000000300D      3
     406       0      -1       1       3                LEVELDEF       1D      4
     314       3       0       0       0       0       0       000000200D      5
     314       0       2       1       0                   COLOR       2D      6
     406       4       0       0       2       0       0       000000300D      7
     406       0      -5       1       3                LEVELDEF       2D      8
     128       5       0       0       1       0       0       000010000D      9
     128       0       0       6       0                NURBSURF       0D     10
     126      11       0       0       2       0       0       000010000D     11
     126       0       0       4       0                NURBLINE       7D     12
     126      15       0       0       1       0       0       000010000D     13
     126       0       0       3       0                NURBLINE      11D     14
     126      18       0       0       1       0       0       000010000D     15
     126       0       0       2       0                NURBLINE      10D     16
     126      20       0       0       1       0       0       000010000D     17
     126       0       0       2       0                NURBLINE       9D     18
     102      22       0       0       0       0       0       000010000D     19
     102       0       0       1       0                POLYLINE       0D     20
     126      23       0       0       0       0       0       000010500D     21
     126       0       0       2       1                NURBLINE       0D     22
     126      25       0       0       0       0       0       000010500D     23
     126       0       0       2       1                NURBLINE       0D     24
     126      27       0       0       0       0       0       000010500D     25
     126       0       0       2       1                NURBLINE       0D     26
     126      29       0       0       0       0       0       000010500D     27
     126       0       0      25       0                NURBLINE       0D     28
     102      54       0       0       1       0       0       000010500D     29
     102       0       0       1       0                  POLY_UV      0D     30
     142      55       0       0       1       0       0       000010500D     31
     142       0       0       1       0                CURVTRIM       0D     32
     144      56       0       0       1       0       0       000000000D     33
     144       0       0       1       0                SURFTRIM       6D     34
     128      57       0       0       2       0       0       000010000D     35
     128       0       0       4       0                NURBSURF       0D     36
     126      61       0       0       2       0       0       000010000D     37
     126       0       0       2       1                NURBLINE      20D     38
     126      63       0       0       2       0       0       000010000D     39
     126       0       0       3       0                NURBLINE      24D     40
     126      66       0       0       2       0       0       000010000D     41
```

```
     126       0       0       3       1                NURBLINE      12D     42
     126      69       0       0       2       0       0       000010000D     43
     126       0       0       4       0                NURBLINE       7D     44
     102      73       0       0       0       0       0       000010000D     45
     102       0       0       1       0                POLYLINE       0D     46
     126      74       0       0       0       0       0       000010500D     47
     126       0       0       2       1                NURBLINE       0D     48
     126      76       0       0       0       0       0       000010500D     49
     126       0       0       2       1                NURBLINE       0D     50
     126      78       0       0       0       0       0       000010500D     51
     126       0       0       2       1                NURBLINE       0D     52
     126      80       0       0       0       0       0       000010500D     53
     126       0       0       2       1                NURBLINE       0D     54
     102      82       0       0       2       0       0       000010500D     55
     102       0       0       1       0                  POLY_UV      0D     56
     142      83       0       0       2       0       0       000010500D     57
     142       0       0       1       0                CURVTRIM       0D     58
     144      84       0       0       2       0       0       000000000D     59
     144       0       0       1       0                SURFTRIM      10D     60
314,0.,100.,0.,12HRGB(0,255,0);                                              1P      1
406,2,1,6HLayer0;                                                            3P      2
314,0.,0.,100.,12HRGB(0,0,255);                                              5P      3
406,2,2,6HLayer1;                                                            7P      4
128,2,2,2,2,0,0,0,0,0,0.,0.,0.,1.,1.,1.,0.,0.,0.,1.,1.,1.,1.,               9P      5
0.707106781,1.,0.707106781,0.5,0.707106781,1.,0.707106781,1.,0.,             9P      6
10.,0.,0.,10.,10.,0.,-6.12323400E-016,10.,1.08881791E-015,10.,               9P      7
9.71014142E-016,10.,10.,10.,10.,-6.12323400E-016,10.,0.,10.,0.,              9P      8
10.,10.,6.12323400E-016,10.,-6.12323400E-016,6.12323400E-016,0.,             9P      9
1.,0.,1.;                                                                    9P     10
126,2,2,1,0,0,0,0.,0.,0.,1.,1.,1.,1.,0.707106781,1.,0.,                     11P     11
-6.12323400E-016,10.,10.,-6.12323400E-016,10.,10.,                          11P     12
-6.12323400E-016,6.12323400E-016,0.,1.,3.73383893E-032,1.,                  11P     13
3.76249879E-032;                                                            11P     14
126,2,2,1,0,0,0,0.,0.,0.,1.,1.,1.,1.,0.707106781,1.,10.,                    13P     15
-6.12323400E-016,6.12323400E-016,10.,10.,6.12323400E-016,0.,10.,            13P     16
0.,0.,1.,0.,0.,1.;                                                          13P     17
126,2,2,1,1,0,0,0.,0.,0.,1.,1.,1.,1.,0.707106781,1.,0.,10.,0.,              15P     18
1.22464680E-015,10.,1.22464680E-015,0.,10.,0.,0.,1.,0.,0.,1.;               15P     19
126,2,2,1,0,0,0,0.,0.,0.,1.,1.,1.,1.,0.707106781,1.,0.,10.,0.,              17P     20
0.,10.,10.,0.,-6.12323400E-016,10.,0.,1.,1.,0.,0.;                          17P     21
102,4,11,13,15,17;                                                          19P     22
126,1,1,1,0,1,0,0.,0.,1.,1.,1.,1.,1.,7.91461014E-012,0.,1.,1.,              21P     23
0.,0.,1.,0.,0.,1.;                                                          21P     24
126,1,1,1,0,1,0,0.,0.,1.,1.,1.,1.,1.,1.,0.,0.,1.,0.,0.,1.,0.,0.,            23P     25
1.;                                                                         23P     26
126,3,3,1,0,1,0,0.,0.,0.,0.,1.,1.,1.,1.,1.,1.,1.,1.,0.,1.,0.,0.,            25P     27
1.,0.,0.,1.,0.,0.,1.,0.,0.,1.,0.,0.,1.;                                     25P     28
126,36,3,1,0,1,0,0.,0.,0.,0.,0.507570747,0.522492845,                       27P     29
0.537414944,0.552337043,0.567259141,0.58218124,0.597103338,                 27P     30
0.612025437,0.626947536,0.641869634,0.656791733,0.671713831,                27P     31
0.68663593,0.701558028,0.716480127,0.731402226,0.746324324,                 27P     32
0.761246423,0.776168521,0.79109062,0.806012718,0.820934817,                 27P     33
0.835856916,0.850779014,0.865701113,0.880623211,0.89554531,                 27P     34
0.910467409,0.925389507,0.940311606,0.955233704,0.970155803,                27P     35
0.985077901,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,             27P     36
1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,1.,             27P     37
1.,1.,1.,0.,1.,0.,0.,0.63693868,0.,0.,0.,0.,0.0589286333,                   27P     38
0.000142665314,0.,0.0882071316,0.,0.,0.117654605,                           27P     39
1.02429087E-005,0.,0.147056802,0.,0.,0.17647113,7.35407758E-007,            27P     40
0.,0.205882208,0.,0.,0.235294157,5.27999015E-008,0.,0.264705872,            27P     41
0.,0.,0.29411765,3.79086237E-009,0.,0.323529411,0.,0.,                      27P     42
0.352941177,2.72171672E-010,0.,0.382352941,0.,0.,0.411764706,               27P     43
1.95410469E-011,0.,0.441176471,0.,0.,0.470588235,                           27P     44
1.40298404E-012,0.,0.5,0.,0.,0.529411765,1.00729726E-013,0.,                27P     45
0.558823529,0.,0.,0.588235294,7.23212260E-015,0.,0.617647059,0.,            27P     46
0.,0.647058823,5.19990238E-016,0.,0.67647059,0.,0.,0.705882346,             27P     47
4.77407305E-017,0.,0.735294142,0.,0.,0.764705792,                           27P     48
1.48379989E-016,0.,0.794117984,0.,0.,0.823528155,                           27P     49
```

```
2.02957911E-015,0.,0.852945867,0.,0.,0.882335435,                    27P     50
2.82657275E-014,0.,0.911830041,0.,0.,0.940932636,                    27P     51
3.93690607E-013,0.,0.971498238,0.,0.,0.988235294,                    27P     52
3.16584406E-012,0.,1.,7.91461014E-012,0.,0.,1.,0.,0.,1.;             27P     53
102,4,21,23,25,27;                                                   29P     54
142,0,9,29,19,2;                                                     31P     55
144,9,1,0,31;                                                        33P     56
128,1,2,1,2,0,0,0,0,0,0.,0.,1.,1.,0.,0.,0.,1.,1.,1.,1.,1.,           35P     57
0.707106781,0.707106781,1.,1.,-6.12323400E-016,0.,10.,               35P     58
-6.12323400E-016,-11.6302491,10.,10.,0.,10.,10.,-11.6302491,10.,     35P     59
10.,0.,0.,10.,-11.6302491,0.,0.,1.,0.,1.;                            35P     60
126,1,1,0,0,1,0,0.,0.,1.,1.,1.,1.,-6.12323400E-016,0.,10.,           37P     61
-6.12323400E-016,-11.6302491,10.,0.,1.,0.,0.,0.;                     37P     62
126,2,2,1,0,0,0,0.,0.,0.,1.,1.,1.,1.,0.707106781,1.,                 39P     63
-6.12323400E-016,-11.6302491,10.,10.,-11.6302491,10.,10.,            39P     64
-11.6302491,0.,0.,1.,5.59633784E-017,1.,1.13954377E-015;             39P     65
126,1,1,0,0,1,0,0.,0.,1.,1.,1.,1.,10.,-11.6302491,                   41P     66
1.22464680E-015,10.,-6.12323400E-016,6.12323400E-016,0.,1.,0.,       41P     67
0.,0.;                                                               41P     68
126,2,2,1,0,0,0,0.,0.,0.,1.,1.,1.,1.,0.707106781,1.,10.,             43P     69
-6.12323400E-016,6.12323400E-016,10.,-6.12323400E-016,10.,0.,        43P     70
-6.12323400E-016,10.,0.,1.,-3.76249879E-032,-1.,                     43P     71
-3.73383893E-032;                                                    43P     72
102,4,37,39,41,43;                                                   45P     73
126,1,1,1,0,1,0,0.,0.,1.,1.,1.,1.,0.,1.90049403E-018,0.,1.,0.,       47P     74
0.,0.,1.,0.,0.,1.;                                                   47P     75
126,1,1,1,0,1,0,0.,0.,1.,1.,1.,1.,1.,0.,0.,1.,1.,0.,0.,1.,0.,0.,     49P     76
1.;                                                                  49P     77
126,1,1,1,0,1,0,0.,0.,1.,1.,1.,1.,1.,1.,0.,4.44089210E-016,1.,       51P     78
0.,0.,1.,0.,0.,1.;                                                   51P     79
126,1,1,1,0,1,0,0.,0.,1.,1.,1.,1.,4.44089210E-016,1.,0.,0.,          53P     80
1.90049403E-018,0.,0.,1.,0.,0.,1.;                                   53P     81
102,4,47,49,51,53;                                                   55P     82
142,0,35,55,45,2;                                                    57P     83
144,35,1,0,57;                                                       59P     84
S      1G      4D     60P     84                                      T      1
```

Figure B.2. Sample of an IGES file.

B.2. Description of the parameters involved in the geometry entities

This section presents the a brief description of the parameters involved in each one of the five geometric entities presented in Table 4.3, which are necessary to define the trimmed NURBS surfaces composing the CAD model. In order to better understand the information in the parameter data section about the selected entities, it is recommended that the reader compare the provided description with the IGES file example shown in Figure B.2. The first entity analysed is the rational B-Spline surface (No. 128), which presents the following specification in the parameter data section:

$$128,n,m,p,q,u^{\mathrm{c}},v^{\mathrm{c}},pr,u^{\mathrm{p}},v^{\mathrm{p}},\overbrace{u_0,u_1,\ldots,u_r}^{\mathbf{U}},\overbrace{v_0,v_1,\ldots,v_s}^{\mathbf{V}},\overbrace{w_{0,0},w_{1,0},\ldots,w_{n,m}}^{w_{i,j}},$$
$$\underbrace{P^{\mathrm{x}}_{0,0},P^{\mathrm{y}}_{0,0},P^{\mathrm{z}}_{0,0},P^{\mathrm{x}}_{1,0},P^{\mathrm{y}}_{1,0},P^{\mathrm{z}}_{1,0},\ldots,P^{\mathrm{x}}_{n,m},P^{\mathrm{y}}_{n,m},P^{\mathrm{z}}_{n,m}}_{\mathbf{P}_{i,j}},u^{\mathrm{s}},u^{\mathrm{e}},v^{\mathrm{s}},v^{\mathrm{e}};$$

where the description of each parameter is presented in Table B.1. This entity starts with its identification number (128) and terminates with the semi-comma symbol. Typically, several lines of the IGES file are require to represent all information of this entity (see Figure B.2). Since the geometry of the example shown in Figure B.2 is defined by two trimmed NURBS surfaces, the IGES file is composed by two rational B-Spline surface entities.

Table B.1. Description of the parameters involved in the definition of the rational B-Spline surface entity (No. 128) [IGES 96].

Parameter	Description
n/m	Number of control points less one in u/v parametric direction
p/q	Surface degree in u/v parametric direction
$u^{\mathrm{c}}/v^{\mathrm{c}}$	Closed (=1) or open (=0) surface in u/v parametric direction
pr	Polynomial (=1) or rational (=0) surface representation
$u^{\mathrm{p}}/v^{\mathrm{p}}$	Periodic (=1) or non-periodic (=0) surface in u/v parametric direction
$\mathbf{U}=\{u_0,\ldots,u_r\}$	Surface's knot vector in u parametric direction (4.34)$_1$
$\mathbf{V}=\{v_0,\ldots,v_s\}$	Surface's knot vector in v parametric direction (4.34)$_2$
$w_{i,j}=\{w_{0,0},\ldots,w_{n,m}\}$	Weight associated to each control point
$\mathbf{P}_{i,j}=\{\mathbf{P}_{0,0},\ldots,\mathbf{P}_{n,m}\}$	Position vector of each control point
$u^{\mathrm{s}}/v^{\mathrm{s}}$	Starting value of u/v in the surface definition (4.33)
$u^{\mathrm{e}}/v^{\mathrm{e}}$	Ending value of u/v in the surface definition (4.33)

The second entity considered is the rational B-Spline curve (No. 126), which presents the following specification in the parameter data section:

$$126,h,l,c^{\mathrm{p}},t^{\mathrm{c}},pr,t^{\mathrm{p}},\overbrace{t_0,t_1,\ldots,t_g}^{\mathbf{T}},\overbrace{w_0,w_1,\ldots,w_h}^{w_i},$$
$$\underbrace{P_0^{\mathrm{x}},P_0^{\mathrm{y}},P_0^{\mathrm{z}},P_1^{\mathrm{x}},P_1^{\mathrm{y}},P_1^{\mathrm{z}},\ldots,P_h^{\mathrm{x}},P_h^{\mathrm{y}},P_h^{\mathrm{z}}}_{\mathbf{P}_i},t^{\mathrm{s}},t^{\mathrm{e}},n^{\mathrm{x}},n^{\mathrm{y}},n^{\mathrm{z}};$$

where the description of each parameter is presented in Table B.2. The control points are defined in the Euclidean space $(P_i^{\mathrm{x}},P_i^{\mathrm{y}},P_i^{\mathrm{z}})$, as well as in the parametric domain of the basis NURBS surface $(P_i^{\mathrm{u}},P_i^{\mathrm{v}},P_i^{\mathrm{w}}=0)$ when the curve describes the boundary of a trimmed NURBS surface. Considering the example of Figure B.2, each surface is bounded by four curves and all of them are trimming curves, then the IGES file is composed by 16 rational B-Spline curve entities.

Table B.2. Description of the parameters involved in the definition of the rational B-Spline curve entity (No. 126) [IGES 96].

Parameter	Description
h	Number of control points less one
l	Curve degree
c^{P}	Planar $(=1)$ or not planar $(=0)$ curve
t^{c}	Closed $(=1)$ or open $(=0)$ curve
pr	Polynomial $(=1)$ or rational $(=0)$ curve representation
t^{P}	Periodic $(=1)$ or non-periodic $(=0)$ curve
$\mathbf{T}=\left\{t_0,\ldots,t_g\right\}$	Curve's knot vector (4.32)
$w_i=\left\{w_0,\ldots,w_h\right\}$	Weight associated to each control point
$\mathbf{P}_i=\left\{\mathbf{P}_0,\ldots,\mathbf{P}_h\right\}$	Position vector of each control point defined in the Euclidean space or in the surface parametric domain
t^{s}	Starting value of t in the curve definition (4.30)
t^{e}	Ending value of t in the curve definition (4.30)
$\mathbf{n}=\left\{n^{\mathrm{x}},n^{\mathrm{y}},n^{\mathrm{z}}\right\}^{\mathrm{T}}$	Unit normal vector (if curve is planar)

The third entity described in this appendix is the composite curve (No. 102), which presents the following specification in the parameter data section:

$$102,n_{\mathrm{sc}},sc_1,sc_2,\ldots,sc_{n_{\mathrm{sc}}};$$

where the description of each parameter is presented in Table B.3. This entity defines a continuous curve composed by an ordered list of simple rational B-Spline curves (No. 126). For the example considered in Section B.1, all composite curves defined in the IGES file are composed by 4 simple curves $(n_{\mathrm{sc}}=4)$. Since the simple curves are defined in both the Euclidean space or in the surface parametric domain, the presented example comprises 4 composite curve entities.

Table B.3. Description of the parameters involved in the definition of the composite curve entity (No. 102) [IGES 96].

Parameter	Description
n_{sc}	Quantity of simple curves forming the composite curve
sc_i	Pointer to the i simple curve $(i=1,\ldots,n_{\mathrm{sc}})$

The forth entity described in this appendix is the composite curve (No. 142), which presents the following specification in the parameter data section:

$$142, wc, se, cc_{\text{uv}}, cc_{\text{xyz}}, rep;$$

where the description of each parameter is presented in Table B.4. This entity associates a given composite curve (No. 102) with a rational B-Spline surface (No. 128), classifying the curve as lying on the surface. The IGES file example shown in Figure B.2 involves two of these entities, one for each NURBS surface illustrated in Figure B.1. Note that this entity contains the information of the composite curve defined both in the parametric domain and in the Euclidean space.

Table B.4. Description of the parameters involved in the definition of the curve on a parametric surface entity (No. 142) [IGES 96].

Parameter	Description
wc	The way the curve on the surface was created
se	Pointer to the surface on which the curve lies
cc_{uv}	Pointer to the composite curve defined in the parametric domain
cc_{xyz}	Pointer to the composite curve defined in the Euclidean space
pre	Preferred representation in the sending system

The last entity considered in this appendix is the trimmed parametric surface (No. 144), which presents the following specification in the parameter data section:

$$144, se, tr, n_{\text{cci}}, cps^{\text{ob}}, cps_1^{\text{ib}}, cps_2^{\text{ib}}, \ldots, cps_{n_{\text{cci}}}^{\text{ib}};$$

where the description of each parameter is presented in Table B.5. This entity defines a trimmed NURBS surface as a junction of an untrimmed surface and a set of trimming closed curves. The number of closed curves that composes the inner boundary is designated by n_{cci}. The pointer cps^{ob} identifies the curve on the parametric surface entity (No. 142) that constitutes the outer boundary of the trimmed surface, while cps_i^{ib} designates the i closed inner boundary curve entity (No. 142).

Table B.5. Description of the parameters involved in the definition of the trimmed parametric surface entity (No. 144) [IGES 96].

Parameter	Description
se	Pointer to the surface that is to be trimmed
tr	Trimmed $(=1)$ or untrimmed $(=0)$ surface definition
n_{cci}	Quantity of closed curves defining the inner boundary of the trimmed surface
cps^{ob}	Pointer to the curve on parametric surface that define the outer boundary of the trimmed surface
cps_i^{ib}	Pointer to the i curve on parametric surface that define the inner boundary of the trimmed surface $(i=1,\ldots,n_{\text{cci}})$

Appendix C

Projection of a point on a NURBS surface

This appendix provides the basic equations necessary to evaluate the projection of a generic point on a NURBS surface. The evaluation of the nodal normal vectors from the original CAD geometry involves the normal projection of the nodes onto the CAD surface, which is performed in the local contact search procedure. The final purpose of the local search is finding the surface local coordinates of each node (called point inversion problem), in order to evaluate the surface normal vector from the first-order partial derivatives. The presented formulation is based in the textbook of Piegl and Tiller [Piegl 97].

C.1. Normal projection

Consider a generic point with position vector $\mathbf{x}_{\mathrm{P}}$ in the Euclidean space, which is orthogonally projected on the NURBS surface $\mathbf{S}(u,v)$, as shown in Figure C.1. According to Stadler et al. [Stadler 03], the distance vector connecting the generic point P to an arbitrary point of the surface is defined as:

$$\mathbf{r}(u,v) = \mathbf{S}(u,v) - \mathbf{x}_{\mathrm{P}}, \tag{C.1}$$

which is indirectly evaluated using the local coordinates of the NURBS surface, used to calculate the position vector of the point belonging the surface.

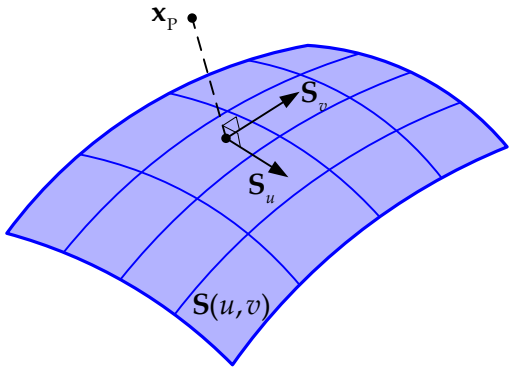


Figure C.1. Projection of a point on a NURBS surface.

The evaluation of the projection point (point of surface) requires the following orthogonal conditions:

$$\begin{cases} f(u,v) = \mathbf{S}_u(u,v) \cdot \mathbf{r}(u,v) = 0 \\ g(u,v) = \mathbf{S}_v(u,v) \cdot \mathbf{r}(u,v) = 0 \end{cases}, \tag{C.2}$$

where the $\mathbf{S}_u(u,v) = \partial \mathbf{S}(u,v)/\partial u$ and $\mathbf{S}_v(u,v) = \partial \mathbf{S}(u,v)/\partial v$ denote the first-order partial derivatives of the NURBS surface (see Figure C.1). In order to solve the nonlinear system of equations (C.2), the Newton–Raphson method can be applied, leading to the iterative solution procedure for iteration i as:

$$\mathbf{J}^i \Delta \mathbf{s}^i = \mathbf{k}^i, \tag{C.3}$$

where the Jacobian matrix of the system of equations (C.2) is expressed by:

$$\mathbf{J}^i = \begin{bmatrix} \dfrac{\partial f(u^i,v^i)}{\partial u} & \dfrac{\partial f(u^i,v^i)}{\partial v} \\ \dfrac{\partial g(u^i,v^i)}{\partial u} & \dfrac{\partial g(u^i,v^i)}{\partial v} \end{bmatrix} = \begin{bmatrix} |\mathbf{S}_u|^2 + \mathbf{r} \cdot \mathbf{S}_{uu} & \mathbf{S}_u \cdot \mathbf{S}_v + \mathbf{r} \cdot \mathbf{S}_{uv} \\ \mathbf{S}_u \cdot \mathbf{S}_v + \mathbf{r} \cdot \mathbf{S}_{vu} & |\mathbf{S}_v|^2 + \mathbf{r} \cdot \mathbf{S}_{vv} \end{bmatrix}^i, \tag{C.4}$$

where the second-order partial derivatives of the NURBS surface are given by $\mathbf{S}_{uu}(u,v) = \partial^2 \mathbf{S}(u,v)/\partial u^2$, $\mathbf{S}_{vv}(u,v) = \partial^2 \mathbf{S}(u,v)/\partial v^2$ and $\mathbf{S}_{uv}(u,v) = \partial^2 \mathbf{S}(u,v)/\partial u \partial v$. The incremental solution vector is given by:

$$\Delta \mathbf{s}^i = \begin{Bmatrix} \Delta u^i \\ \Delta v^i \end{Bmatrix} = \begin{Bmatrix} u^{i+1} - u^i \\ v^{i+1} - v^i \end{Bmatrix}, \tag{C.5}$$

which is the difference between two consecutive iterations in terms of surface local coordinates. The update procedure for the solution is directly obtained from (C.5). The residual vector is expressed by:

$$\mathbf{k}^i = -\begin{Bmatrix} f(u^i, v^i) \\ g(u^i, v^i) \end{Bmatrix}. \tag{C.6}$$

The initial values u^0 and v^0 should be given *a priori*, influencing the convergence rate. In the adopted approach the initial values are estimated based on the minimum distance between the generic point P and a set of points generated on the surface, equally spaced in the parametric domain. The convergence criterion defined for the solution of the nonlinear system of equations (C.2) is based on the simultaneously fulfilment of the following conditions. The first condition dictates that the distance between the generic point P and the calculated projection point is very small:

$$\left\| \mathbf{S}(u^i, v^i) - \mathbf{x}_\mathrm{P} \right\| \leq \varepsilon_1. \tag{C.7}$$

This condition is allowed because all surface mesh nodes are points generated on the NURBS surface. The essential orthogonal conditions expressed in (C.2) are satisfied through:

$$\frac{\left\| \mathbf{S}_u(u^i, v^i) \cdot (\mathbf{S}(u^i, v^i) - \mathbf{x}_\mathrm{P}) \right\|}{\left\| \mathbf{S}_u(u^i, v^i) \right\| \left\| \mathbf{S}(u^i, v^i) - \mathbf{x}_\mathrm{P} \right\|} \leq \varepsilon_2 \quad \text{and} \quad \frac{\left\| \mathbf{S}_v(u^i, v^i) \cdot (\mathbf{S}(u^i, v^i) - \mathbf{x}_\mathrm{P}) \right\|}{\left\| \mathbf{S}_v(u^i, v^i) \right\| \left\| \mathbf{S}(u^i, v^i) - \mathbf{x}_\mathrm{P} \right\|} \leq \varepsilon_2, \tag{C.8}$$

where the predefined threshold value ε_2 is very small to guaranty the zero cosine. Moreover, when the solution point slides out of the parametric domain of the NURBS surface during the iterative scheme, it is repositioned on the surface boundary as follows:

$$\begin{aligned} &\text{if } (u^{i+1} < u^\mathrm{s}) \quad u^{i+1} = u^\mathrm{s}; \quad \text{if } (u^{i+1} > u^\mathrm{e}) \quad u^{i+1} = u^\mathrm{e}, \\ &\text{if } (v^{i+1} < v^\mathrm{s}) \quad v^{i+1} = v^\mathrm{s}; \quad \text{if } (v^{i+1} > v^\mathrm{e}) \quad v^{i+1} = v^\mathrm{e}. \end{aligned} \tag{C.9}$$

where u^s, v^s, u^e and v^e are defined as presented in (4.33).

C.2. Partial derivatives of a NURBS surface

The closed form expressions necessary to evaluate the first- and second-order partial derivatives in a generic point of a NURBS surface are described in this section. The presented formulation is based on the studies of Wang et al. [Wang 04] and Tsai et al. [Tsai 03]. The NURBS surface of deegre p in the u direction and degree q in the v direction is given by:

$$\mathbf{S}(u,v)=\frac{\sum_{i=0}^{n}\sum_{j=0}^{m}N_{i,p}(u)N_{j,q}(v)w_{i,j}\mathbf{P}_{i,j}}{\sum_{i=0}^{n}\sum_{j=0}^{m}N_{i,p}(u)N_{j,q}(v)w_{i,j}}, \tag{C.10}$$

where all variable involved were previously defined in the presentation of (4.33). The numerator of (C.10) is expressed by:

$$\mathbf{A}(u,v)=\sum_{i=0}^{n}\sum_{j=0}^{m}N_{i,p}(u)N_{j,q}(v)w_{i,j}\mathbf{P}_{i,j}, \tag{C.11}$$

and the denominator of (C.10) is given by:

$$w(u,v)=\sum_{i=0}^{n}\sum_{j=0}^{m}N_{i,p}(u)N_{j,q}(v)w_{i,j}. \tag{C.12}$$

The formulas for computing the first-order partial derivatives of the NURBS surface defined in (C.10), can be obtained from the appropriate mathematical manipulations of the numerator and denominator expressions. Then, the first-order partial derivatives are defined by:

$$\frac{\partial\mathbf{S}(u,v)}{\partial\alpha}\equiv\mathbf{S}_{\alpha}(u,v)=\frac{\mathbf{A}_{\alpha}(u,v)-w_{\alpha}(u,v)\mathbf{S}(u,v)}{w(u,v)}, \tag{C.13}$$

where α denotes either u or v. The second-order partial derivatives of the NURBS surface are defined in a similar way, given by:

$$\begin{aligned}\frac{\partial^2\mathbf{S}(u,v)}{\partial\beta\partial\alpha}&\equiv\mathbf{S}_{\alpha\beta}(u,v)=\\&=\frac{\mathbf{A}_{\alpha\beta}(u,v)-w_{\alpha\beta}(u,v)\mathbf{S}(u,v)-w_{\alpha}(u,v)\mathbf{S}_{\beta}(u,v)-w_{\beta}(u,v)\mathbf{S}_{\alpha}(u,v)}{w(u,v)},\end{aligned} \tag{C.14}$$

where α denotes either u or v. On the other hand, the partial derivatives of the expressions (C.11) and (C.12), which are involved in the first- and second-order partial derivatives of the NURBS surface, are defined by:

$$
\begin{aligned}
\mathbf{A}_u(u,v) &= \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}^{(1)}(u)N_{j,q}(v)w_{i,j}\mathbf{P}_{i,j}\,, \quad w_u(u,v) = \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}^{(1)}(u)N_{j,q}(v)w_{i,j}\,,\\
\mathbf{A}_v(u,v) &= \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}(u)N_{j,q}^{(1)}(v)w_{i,j}\mathbf{P}_{i,j}\,, \quad w_v(u,v) = \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}(u)N_{j,q}^{(1)}(v)w_{i,j}\,,\\
\mathbf{A}_{uu}(u,v) &= \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}^{(2)}(u)N_{j,q}(v)w_{i,j}\mathbf{P}_{i,j}\,, \quad w_{uu}(u,v) = \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}^{(2)}(u)N_{j,q}(v)w_{i,j}\,,\\
\mathbf{A}_{vv}(u,v) &= \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}(u)N_{j,q}^{(2)}(v)w_{i,j}\mathbf{P}_{i,j}\,, \quad w_{vv}(u,v) = \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}(u)N_{j,q}^{(2)}(v)w_{i,j}\,,\\
\mathbf{A}_{uv}(u,v) &= \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}^{(1)}(u)N_{j,q}^{(1)}(v)w_{i,j}\mathbf{P}_{i,j}\,, \quad w_{uv}(u,v) = \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}^{(1)}(u)N_{j,q}^{(1)}(v)w_{i,j}\,,
\end{aligned}
\tag{C.15}
$$

where the derivatives of the B-spline basis functions (4.31) are involved. The *k*th derivative of $N_{i,p}(u)$ is denoted by $N_{i,p}^{(k)}(u)$, which is defined by its general recurrence formula given by [Piegl 97]:

$$
N_{i,p}^{(k)}(u) = p\left(\frac{N_{i,p-1}^{(k-1)}(u)}{u_{i+p}-u_i} - \frac{N_{i+1,p-1}^{(k-1)}(u)}{u_{i+p+1}-u_{i+1}}\right), \tag{C.16}
$$

where the derivatives of B-spline basis functions are efficiently represented in terms of B-spline lower order bases.

References

[Alart 91] ALART, P. AND CURNIER, A., 1991. A mixed formulation for frictional contact problems prone to Newton like solution methods. *Computer Methods in Applied Mechanics and Engineering*, 92(3), pp. 353–375.

[Alart 95a] ALART, P. AND HEEGE, A., 1995. Consistent tangent matrices of curved contact operators involving anisotropic friction. *Revue Européenne des Éléments*, 4(2), pp. 183–207.

[Alart 95b] ALART, P. AND LEBON, F., 1995. Solution of frictional contact problems using ILU and coarse/fine preconditioners. *Computational Mechanics*, 16(2), pp. 98–105.

[Alart 97] ALART, P., 1997. Méthode de Newton généralisée en mécanique du contact. *Journal de Mathématiques Pures et Appliquées*, 76(1), pp. 83–108.

[Al-Dojayli 02] AL-DOJAYLI, M. AND MEGUID, S.A., 2002. Accurate modeling of contact using cubic splines. *Finite Elements in Analysis and Design*, 38(4), pp. 337–352.

[Alves 03] ALVES, J.L., 2003. *Simulação numérica do processo de estampagem de chapas metálicas: modelação mecânica e métodos numéricos*. University of Minho, Portugal.

[Aragón 13] ARAGÓN, A.M., YASTREBOV, V.A. AND MOLINARI, J.-F., 2013. A constrained-optimization methodology for the detection phase in contact mechanics simulations. *International Journal for Numerical Methods in Engineering*, 96(5), pp. 323–338.

[Areias 04] AREIAS, P.M.A., CÉSAR DE SÁ, J.M.A. AND CONCEIÇÃO ANTÓNIO, C.A., 2004. Algorithms for the analysis of 3D finite strain contact problems. *International Journal for Numerical Methods in Engineering*, 61(7), pp. 1107–1151.

[Barboteu 03] BARBOTEU, M., HOARAU-MANTEL, T.-V. AND SOFONEA, M., 2003. On the frictionless unilateral contact of two viscoelastic bodies. *Journal of Applied Mathematics*, 2003(11), pp. 575–603.

[Barlat 91] BARLAT, F., LEGE, D.J. AND BREM, J.C., 1991. A six-component yield function for anisotropic materials. *International Journal of Plasticity*, 7(7), pp. 693–712.

[Basu 95] BASU, D. AND KUMAR, S.S., 1995. Importing mesh entities through IGES/PDES. *Advances in Engineering Software*, 23(3), pp. 151–161.

[Batailly 13] BATAILLY, A., MAGNAIN, B. AND CHEVAUGEON, N., 2013. A comparative study between two smoothing strategies for the simulation of contact with large sliding. *Computational Mechanics*, 51(5), pp. 581–601.

[Belgacem 98] BELGACEM, F.B., HILD, P. AND LABORDE, P., 1998. The mortar finite element method for contact problems. *Mathematical and Computer Modelling*, 28(4-8), pp. 263–271.

[Belytschko 00] BELYTSCHKO, T., LIU, W.K. AND MORAN, B., 2000. *Nonlinear finite elements for continua and structures*, John Wiley & Sons.

[Belytschko 02] BELYTSCHKO, T., DANIEL, W.J.T. AND VENTURA, G., 2002. A monolithic smoothing-gap algorithm for contact-impact based on the signed distance function. *International Journal for Numerical Methods in Engineering*, 55(1), pp. 101–125.

[Belytschko 91] BELYTSCHKO, T. AND NEAL, M.O., 1991. Contact-impact by the pinball algorithm with penalty and Lagrangian methods. *International Journal for Numerical Methods in Engineering*, 31(3), pp. 547–572.

[Benson 90] BENSON, D.J. AND HALLQUIST, J.O., 1990. A single surface contact algorithm for the post-buckling analysis of shell structures. *Computer Methods in Applied Mechanics and Engineering*, 78(2), pp. 141–163.

[Benzley 95] BENZLEY, S.E., PERRY, E., MERKLEY, K., CLARK, B. AND SJAARDEMA, G., 1995. A comparison of all hexagonal and all tetrahedral finite element meshes for elastic and elasto-plastic analysis. In *Proceedings of the 4th International Meshing Roundtable*. Sandia National Laboratories, pp. 179–191.

[Bernardi 90] BERNARDI, C., DEBIT, N. AND MADAY, Y., 1990. Coupling finite element and spectral methods: first results. *Mathematics of Computation*, 54(189), pp. 21–39.

[Bertram 08] BERTRAM, A., 2008. *Elasticity and Plasticity of Large Deformations*, Berlin: Springer-Verlag.

[Bhandarkar 00] BHANDARKAR, M.P., DOWNIE, B., HARDWICK, M. AND NAGI, R., 2000. Migrating from IGES to STEP: one to one translation of IGES drawing to STEP drafting data. *Computers in Industry*, 41(3), pp. 261–277.

[Bonet 97] BONET, J. AND WOOD, R.D., 1997. *Nonlinear Continuum Mechanics for Finite Element Analysis*, Cambridge University Press.

[Boschiroli 11] BOSCHIROLI, M., FÜNFZIG, C., ROMANI, L. AND ALBRECHT, G., 2011. A comparison of local parametric C0 Bézier interpolants for triangular meshes. *Computers & Graphics*, 35(1), pp. 20–34.

[Brunssen 07] BRUNSSEN, S., SCHMID, F., SCHÄFER, M. AND WOHLMUTH, B., 2007. A fast and robust iterative solver for nonlinear contact problems using a primal-dual active set strategy and algebraic multigrid. *International Journal for Numerical Methods in Engineering*, 69(3), pp. 524–543.

[Buranathiti 05a] BURANATHITI, T., 2005a. Benchmark Simulation Results: Automotive Underbody Cross Member (Benchmark 2). In *AIP Conference Proceedings*. pp. 1004–1112.

[Buranathiti 05b] BURANATHITI, T., 2005b. Numisheet2005 Benchmark Analysis on Forming of an Automotive Underbody Cross Member: Benchmark 2. In *AIP Conference Proceedings*. pp. 1113–1120.

[Cazacu 01] CAZACU, O. AND BARLAT, F., 2001. Generalization of Drucker's Yield Criterion to Orthotropy. *Mathematics and Mechanics of Solids*, 6(6), pp. 613–630.

[Cazacu 06] CAZACU, O., PLUNKETT, B. AND BARLAT, F., 2006. Orthotropic yield criterion for hexagonal closed packed metals. *International Journal of Plasticity*, 22(7), pp. 1171–1194.

[Chamoret 04] CHAMORET, D., SAILLARD, P., RASSINEUX, A. AND BERGHEAU, J.-M., 2004. New smoothing procedures in contact mechanics. *Journal of Computational and Applied Mathematics*, 168(1-2), pp. 107–116.

[Chow 97] CHOW, E. AND SAAD, Y., 1997. Experimental study of ILU preconditioners for indefinite matrices. *Journal of Computational and Applied Mathematics*, 86(2), pp. 387–414.

[CIMNE 13] CIMNE, 2013. GiD: The personal pre and postprocessor.

[Courtney-Pratt 57] COURTNEY-PRATT, J.S. AND EISNER, E., 1957. The Effect of a Tangential Force on the Contact of Metallic Bodies. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 238(1215), pp. 529–550.

[Cox 72] COX, M.G., 1972. The Numerical Evaluation of B-Splines. *IMA Journal of Applied Mathematics*, 10(2), pp. 134–149.

[Crisfield 00] CRISFIELD, M.A., 2000. Re-visiting the contact patch test. *International Journal for Numerical Methods in Engineering*, 48(3), pp. 435–449.

[Curnier 84] CURNIER, A., 1984. A theory of friction. *International Journal of Solids and Structures*, 20(7), pp. 637–647.

[de Boor 72] DE BOOR, C., 1972. On calculating with B-splines. *Journal of Approximation Theory*, 6(1), pp. 50–62.

[Dunne 05] DUNNE, F. AND PETRINIC, N., 2005. *Introduction to Computational Plasticity*, Oxford University Press.

[Duvaut 76] DUVAUT, G. AND LIONS, J.L., 1976. *Inequalities in Mechanics and Physics*, Berlin: Springer-Verlag.

[El-Abbasi 01] EL-ABBASI, N. AND BATHE, K.-J., 2001. Stability and patch test performance of contact discretizations and a new solution algorithm. *Computers & Structures*, 79(16), pp. 1473–1486.

[El-Abbasi 01] EL-ABBASI, N., MEGUID, S.A. AND CZEKANSKI, A., 2001. On the modelling of smooth contact surfaces using cubic splines. *International Journal for Numerical Methods in Engineering*, 50(4), pp. 953–967.

[Farouki 99] FAROUKI, R.T., 1999. Closing the gap between cad model and downstream application. *SIAM News*, 32(5), pp. 303–319.

[Fischer 06] FISCHER, K.A. AND WRIGGERS, P., 2006. Mortar based frictional contact formulation for higher order interpolations using the moving friction cone. *Computer Methods in Applied Mechanics and Engineering*, 195(37-40), pp. 5020–5036.

[Francavilla 75] FRANCAVILLA, A. AND ZIENKIEWICZ, O.C., 1975. A note on numerical computation of elastic contact problems. *International Journal for Numerical Methods in Engineering*, 9(4), pp. 913–924.

[Fujun 00] FUJUN, W., JIANGANG, C. AND ZHENHAN, Y., 2000. A contact searching algorithm for contact-impact problems. *Acta Mechanica Sinica*, 16(4), pp. 374–382.

[Gelin 99] GELIN, J.C., AND PICART, P., 1999. Benchmark C: reverse deep drawing of a cylindrical cup. In J.C. Gelin and P. Picart, ed. *Proceedings of Numisheet'99*. France, pp. 871–932.

[Gould 07] GOULD, N.I.M., SCOTT, J.A. AND HU, Y., 2007. A numerical evaluation of sparse direct solvers for the solution of large sparse symmetric linear systems of equations. *ACM Transactions on Mathematical Software*, 33(2), pp. 1–32.

[Gouraud 71] GOURAUD, H., 1971. Continuous Shading of Curved Surfaces. *IEEE Transactions on Computers*, C-20(6), pp. 623–629.

[Gurtin 81] GURTIN, M.E., 1981. *An Introduction to Continuum Mechanics*, Academic Press.

[Hachani 13] HACHANI, M. AND FOURMENT, L., 2013. A smoothing procedure based on quasi-C1 interpolation for 3D contact mechanics with applications to metal forming. *Computers & Structures*, 128, pp. 1–13.

[Hallquist 85] HALLQUIST, J.O., GOUDREAU, G.L. AND BENSON, D.J., 1985. Sliding interfaces with contact-impact in large-scale Lagrangian computations. *Computer Methods in Applied Mechanics and Engineering*, 51(1), pp. 107–137.

[Hama 08] HAMA, T., NAGATA, T., TEODOSIU, C., MAKINOUCHI, A. AND TAKUDA, H., 2008. Finite-element simulation of springback in sheet metal forming using local interpolation for tool surfaces. *International Journal of Mechanical Sciences*, 50(2), pp. 175–192.

[Hansson 90] HANSSON, E. AND KLARBRING, A., 1990. Rigid contact modelled by CAD surface. *Engineering Computations*, 7(4), pp. 344–348.

[Hartmann 08] HARTMANN, S. AND RAMM, E., 2008. A mortar based contact formulation for non-linear dynamics using dual Lagrange multipliers. *Finite Elements in Analysis and Design*, 44(5), pp. 245–258.

[Hartmann 09] HARTMANN, S., OLIVER, J., WEYLER, R., CANTE, J.C. AND HERNÁNDEZ, J.A., 2009. A contact domain method for large deformation frictional contact problems. Part 2: Numerical aspects. *Computer Methods in Applied Mechanics and Engineering*, 198(33), pp. 2607–2631.

[Heegaard 93] HEEGAARD, J.-H. AND CURNIER, A., 1993. An augmented Lagrangian method for discrete large-slip contact problems. *International Journal for Numerical Methods in Engineering*, 36(4), pp. 569–593.

[Heegaard 96] HEEGAARD, J.-H. AND CURNIER, A., 1996. Geometric properties of 2D and 3D unilateral large slip contact operators. *Computer Methods in Applied Mechanics and Engineering*, 131(3-4), pp. 263–286.

[Heege 96] HEEGE, A. AND ALART, P., 1996. A frictional contact element for strongly curved contact problems. *International Journal for Numerical Methods in Engineering*, 39(1), pp. 165–184.

[Hertz 82] HERTZ, H., 1882. Ueber die Berührung fester elastische Körper. *Journal für die reine und angewandte Mathematik*, 92, pp. 156–171.

[Hestenes 69] HESTENES, M.R., 1969. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 4(5), pp. 303–320.

[Hill 48] HILL, R., 1948. A Theory of the Yielding and Plastic Flow of Anisotropic Metals. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 193(1033), pp. 281–297.

[Holzapfel 00] HOLZAPFEL, G.A., 2000. *Nonlinear Solid Mechanics: A Continuum Approach for Engineering*, John Wiley & Sons.

[Hüeber 08] HÜEBER, S., STADLER, G. AND WOHLMUTH, B.I., 2008. A Primal-Dual Active Set Algorithm for Three-Dimensional Contact Problems with Coulomb Friction. *SIAM Journal on Scientific Computing*, 30(2), pp. 572–596.

[Hughes 05] HUGHES, T.J.R., COTTRELL, J.A. AND BAZILEVS, Y., 2005. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering*, 194(39-41), pp. 4135–4195.

[Hughes 76] HUGHES, T.J.R., TAYLOR, R.L., SACKMAN, J.L., CURNIER, A. AND KANOKNUKULCHAI, W., 1976. A finite element method for a class of contact-impact problems. *Computer Methods in Applied Mechanics and Engineering*, 8(3), pp. 249–276.

[Hughes 80] HUGHES, T.J.R., 1980. Generalization of selective integration procedures to anisotropic and nonlinear media. *International Journal for Numerical Methods in Engineering*, 15(9), pp. 1413–1418.

[IGES 96] IGES, 1996. *Initial Graphics Exchange Specification, IGES 5.3*, IGES/PDES Organization, U.S. Product Data Association.

[Intel 14] INTEL, 2014. *Intel Math Kernel Library Reference Manual*, Document Number: 630813-064US.

[Jin 05] JIN, S., LEWIS, R.R. AND WEST, D., 2005. A comparison of algorithms for vertex normal computation. *The Visual Computer*, 21(1-2), pp. 71–82.

[Kantorovich 48] KANTOROVICH, L.V., 1948. On Newton's method for functional equations. *Doklady Akademii nauk SSSR*, 59, pp. 1237–1240.

[Karafillis 93] KARAFILLIS, A.P. AND BOYCE, M.C., 1993. A general anisotropic yield criterion using bounds and a transformation weighting tensor. *Journal of the Mechanics and Physics of Solids*, 41(12), pp. 1859–1886.

[Khan 95] KHAN, A.S. AND HUANG, S., 1995. *Continuum Theory of Plasticity*, John Wiley & Sons.

[Kikuchi 88] KIKUCHI, N. AND ODEN, J.T., 1988. *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*, Philadelphia: SIAM.

[Klang 79] KLANG, M., 1979. *On interior contact under friction between cylindrical elastic bodies*. Linköping University, Sweden.

[Klarbring 90] KLARBRING, A., 1990. Examples of non-uniqueness and non-existence of solutions to quasistatic contact problems with friction. *Ingenieur-Archiv*, 60(8), pp. 529–541.

[Klarbring 92] KLARBRING, A. AND BJÖRKMAN, G., 1992. Solution of large displacement contact problems with friction using Newton's method for generalized equations. *International Journal for Numerical Methods in Engineering*, 34(1), pp. 249–269.

[Klarbring 95] KLARBRING, A., 1995. Large displacement frictional contact: a continuum framework for finite element discretization. *European Journal of Mechanics A/Solids*, 14(2), pp. 237–253.

[Konyukhov 05] KONYUKHOV, A. AND SCHWEIZERHOF, K., 2005. Covariant description for frictional contact problems. *Computational Mechanics*, 35(3), pp. 190–213.

[Konyukhov 06] KONYUKHOV, A. AND SCHWEIZERHOF, K., 2006. A special focus on 2D formulations for contact problems using a covariant description. *International Journal for Numerical Methods in Engineering*, 66(9), pp. 1432–1465.

[Konyukhov 08] KONYUKHOV, A. AND SCHWEIZERHOF, K., 2008. 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*, 197(33-40), pp. 3045–3056.

[Krstulović-Opara 02] KRSTULOVIC-OPARA, L., WRIGGERS, P. AND KORELC, J., 2002. A C1-continuous formulation for 3D finite deformation frictional contact. *Computational Mechanics*, 29(1), pp. 27–42.

[Laursen 02] LAURSEN, T.A., 2002. *Computational Contact and Impact Mechanics: Fundamentals of Modeling Interfacial Phenomena in Nonlinear Finite Element Analysis*, Berlin: Springer-Verlag.

[Laursen 12] LAURSEN, T.A., PUSO, M.A. AND SANDERS, J., 2012. Mortar contact formulations for deformable–deformable contact: Past contributions and new extensions for enriched and embedded interface formulations. *Computer Methods in Applied Mechanics and Engineering*, 205-208, pp. 3–15.

[Laursen 93] LAURSEN, T.A. AND SIMO, J.C., 1993. A continuum-based finite element formulation for the implicit solution of multibody, large deformation-frictional contact problems. *International Journal for Numerical Methods in Engineering*, 36(20), pp. 3451–3485.

[Laursen 94] LAURSEN, T.A., 1994. The convected description in large deformation frictional contact problems. *International Journal of Solids and Structures*, 31(5), pp. 669–681.

[Laursen 95] LAURSEN, T.A. AND MAKER, B.N., 1995. An augmented Lagrangian quasi-Newton solver for constrained nonlinear finite element applications. *International Journal for Numerical Methods in Engineering*, 38(21), pp. 3571–3590.

[Lee 69] LEE, E.H., 1969. Elastic-Plastic Deformation at Finite Strains. *Journal of Applied Mechanics*, 36(1), pp. 1–6.

[Lemaitre 90] LEMAITRE, J. AND CHABOCHE, J.-L., 1990. *Mechanics of Solid Materials*, Cambridge University Press.

[Lin 01] LIN, J., BALL, A.A. AND ZHENG, J.J., 2001. Approximating circular arcs by Bézier curves and its application to modelling tooling for FE forming simulations. *International Journal of Machine Tools and Manufacture*, 41(5), pp. 703–717.

[Lin 04] LIN, B.-T. AND KUO, C.-C., 2006. Application of an integrated CAD/CAE/CAM system for stamping dies for automobiles. *The*

International Journal of Advanced Manufacturing Technology, 35(9-10), pp. 1000–1013.

[Ling 97] LING, W. AND STOLARSKI, H.K., 1997. A contact algorithm for problems involving quadrilateral approximation of surfaces. *Computers & Structures*, 63(5), pp. 963–975.

[Litke 01] LITKE, N., LEVIN, A. AND SCHRÖDER, P., 2001. Trimming for subdivision surfaces. *Computer Aided Geometric Design*, 18(5), pp. 463–481.

[Luenberger 08] LUENBERGER, D.G. AND YE, Y., 2008. *Linear and Nonlinear Programming* Third edit., Boston: Springer-Verlag.

[Ma 03] MA, Y.L. AND HEWITT, W.T., 2003. Point inversion and projection for NURBS curve and surface: Control polygon approach. *Computer Aided Geometric Design*, 20(2), pp. 79–99.

[Malvern 69] MALVERN, L.E., 1969. *Introduction to the mechanics of a continuous medium*, Prentice- Hall.

[Martins 87] MARTINS, J.A.C. AND ODEN, J.T., 1987. Existence and uniqueness results for dynamic contact problems with nonlinear normal and friction interface laws. *Nonlinear Analysis: Theory, Methods & Applications*, 11(3), pp. 407–428.

[Martins 94] MARTINS, J.A.C., MONTEIRO MARQUES, M.D.P. AND GASTALDI, F., 1994. On an example of nonexistence of solutions to a quasi-static frictional contact problem. *European Journal of Mechanics A/Solids*, 13, pp. 113–133.

[Max 99] MAX, N., 1999. Weights for Computing Vertex Normals from Facet Normals. *Journal of Graphics Tools*, 4(2), pp. 1–6.

[McDevitt 00] MCDEVITT, T.W. AND LAURSEN, T.A., 2000. A mortar-finite element formulation for frictional contact problems. *International Journal for Numerical Methods in Engineering*, 48(10), pp. 1525–1547.

[Meek 00] MEEK, D.S. AND WALTON, D.J., 2000. On surface normal and Gaussian curvature approximations given data sampled from a smooth surface. *Computer Aided Geometric Design*, 17(6), pp. 521–543.

[Menezes 00] MENEZES, L.F. AND TEODOSIU, C., 2000. Three-dimensional numerical simulation of the deep-drawing process using solid finite elements. *Journal of Materials Processing Technology*, 97(1), pp. 100–106.

[Menezes 11] MENEZES, L.F., NETO, D.M., OLIVEIRA, M.C. AND ALVES, J.L., 2011. Improving Computational Performance through HPC Techniques: case study using DD3IMP in-house code. In *AIP Conference Proceedings*. pp. 1220–1225.

[Menezes 94] MENEZES, L.F., 1994. *Modelação tridimensional e simulação numérica dos processos de enformação por deformação plástica: aplicação à estampagem de chapas metálicas*, University of Coimbra, Portugal.

[Michalowski 78] MICHALOWSKI, R. AND MROZ, Z., 1978. Associated and non-associated sliding rules in contact friction problems. *Archives of Mechanics*, 30(3), pp. 259–276.

[Mijar 00] MIJAR, A.R. AND ARORA, J.S., 2000. Review of formulations for elastostatic frictional contact problems. *Structural and Multidisciplinary Optimization*, 20(3), pp. 167–189.

[Mijar 04] MIJAR, A.R. AND ARORA, J.S., 2004. An augmented Lagrangian optimization method for contact analysis problems, 1: formulation and algorithm. *Structural and Multidisciplinary Optimization*, 28(2-3), pp. 99–112.

[Mishin 10] MISHIN, Y., ASTA, M. AND LI, J., 2010. Atomistic modeling of interfaces and their impact on microstructure and properties. *Acta Materialia*, 58(4), pp. 1117–1151.

[Mooney 40] MOONEY, M., 1940. A Theory of Large Elastic Deformation. *Journal of Applied Physics*, 11(9), pp. 582–592.

[Morra 08] MORRA, E.A., ROSCA, M., GREENWALD, J.F.I. AND GREENWALD, A.S., 2008. The influence of contemporary knee design on high flexion: a kinematic comparison with the normal knee. *The Journal of bone and joint surgery. American volume*, 90(4), pp. 195–201.

[Nagata 05] NAGATA, T., 2005. Simple local interpolation of surfaces using normal vectors. *Computer Aided Geometric Design*, 22(4), pp. 327–347.

[Nagtegaal 74] NAGTEGAAL, J.C., PARKS, D.M. AND RICE, J.R., 1974. On numerically accurate finite element solutions in the fully plastic range. *Computer Methods in Applied Mechanics and Engineering*, 4(2), pp. 153–177.

[Ogden 84] OGDEN, R.W., 1984. *Non-Linear Elastic Deformations*, John Wiley & Sons.

[Oldenburg 94] OLDENBURG, M. AND NILSSON, L., 1994. The position code algorithm for contact searching. *International Journal for Numerical Methods in Engineering*, 37(3), pp. 359–386.

[Oliveira 03] OLIVEIRA, M.C., ALVES, J.L. AND MENEZES, L.F., 2003. Improvement of a frictional contact algorithm for strongly curved contact problems. *International Journal for Numerical Methods in Engineering*, 58(14), pp. 2083–2101.

[Oliveira 04] OLIVEIRA, M.C. AND MENEZES, L.F., 2004. Automatic correction of the time step in implicit simulations of the stamping process. *Finite Elements in Analysis and Design*, 40(13-14), pp. 1995–2010.

[Oliveira 07] OLIVEIRA, M.C., ALVES, J.L., CHAPARRO, B. AND MENEZES, L.F., 2007. Study on the influence of work-hardening modeling in springback prediction. *International Journal of Plasticity*, 23(3), pp. 516–543.

[Oliveira 08] OLIVEIRA, M.C., ALVES, J.L. AND MENEZES, L.F., 2008. Algorithms and Strategies for Treatment of Large Deformation Frictional Contact in the Numerical Simulation of Deep Drawing Process. *Archives of Computational Methods in Engineering*, 15(2), pp. 113–162.

[OuYang 05] OUYANG, D. AND FENG, H.-Y., 2005. On the normal vector estimation for point cloud data from smooth surfaces. *Computer-Aided Design*, 37(10), pp. 1071–1079.

[Padmanabhan 01] PADMANABHAN, V. AND LAURSEN, T.A., 2001. A framework for development of surface smoothing procedures in large deformation frictional contact analysis. *Finite Elements in Analysis and Design*, 37(3), pp. 173–198.

[Page 02] PAGE, D.L., SUN, Y., KOSCHAN, A.F., PAIK, J. AND ABIDI, M.A., 2002. Normal Vector Voting: Crease Detection and Curvature Estimation on Large, Noisy Meshes. *Graphical Models*, 64(3-4), pp. 199–229.

[Parisch 89] PARISCH, H., 1989. A consistent tangent stiffness matrix for three-dimensional non-linear contact analysis. *International Journal for Numerical Methods in Engineering*, 28(8), pp. 1803–1812.

[Piegl 91] PIEGL, L., 1991. On NURBS: a survey. *IEEE Computer Graphics and Applications*, 11(1), pp. 55–71.

[Piegl 95] PIEGL, L.A. AND RICHARD, A.M., 1995. Tessellating trimmed nurbs surfaces. *Computer-Aided Design*, 27(1), pp. 16–26.

[Piegl 97] PIEGL, L. AND TILLER, W., 1997. *The NURBS Book*, Berlin: Springer-Verlag.

[Pietrzak 97] PIETRZAK, G., 1997. *Continuum mechanics modelling and augmented Lagrangian formulation of large deformation frictional contact problems*. École Polytechnique Fédérale de Lausanne, Switzerland.

[Pietrzak 99] PIETRZAK, G. AND CURNIER, A., 1999. Large deformation frictional contact mechanics: continuum formulation and augmented Lagrangian treatment. *Computer Methods in Applied Mechanics and Engineering*, 177(3), pp. 351–381.

[Popp 09] POPP, A., GEE, M.W. AND WALL, W.A., 2009. A finite deformation mortar contact formulation using a primal-dual active set strategy. *International Journal for Numerical Methods in Engineering*, 79(11), pp. 1354–1391.

[Popp 12] POPP, A., 2012. *Mortar methods for computational contact mechanics and general interface problems*, Technische Universität München, Germany.

[Powell 69] POWELL, M.J.D., 1969. A method for nonlinear constraints in minimization problems. In R. Fletcher, ed. *Optimization*. New York, pp. 283–298.

[Puso 02] PUSO, M.A. AND LAURSEN, T.A., 2002. A 3D contact smoothing method using Gregory patches. *International Journal for Numerical Methods in Engineering*, 54(8), pp. 1161–1194.

[Puso 04a] PUSO, M.A. AND LAURSEN, T.A., 2004. A mortar segment-to-segment contact method for large deformation solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 193(6-8), pp. 601–629.

[Puso 04b] PUSO, M.A. AND LAURSEN, T.A., 2004. A mortar segment-to-segment frictional contact method for large deformations. *Computer Methods in Applied Mechanics and Engineering*, 193(45-47), pp. 4891–4913.

[Puso 08] PUSO, M.A., LAURSEN, T.A. AND SOLBERG, J., 2008. A segment-to-segment mortar contact method for quadratic elements and large deformations. *Computer Methods in Applied Mechanics and Engineering*, 197(6-8), pp. 555–566.

[Qian 14] QIAN, X., YUAN, H., ZHOU, M. AND ZHANG, B., 2014. A general 3D contact smoothing method based on radial point interpolation. *Journal of Computational and Applied Mathematics*, 257, pp. 1–13.

[Renard 13] RENARD, Y., 2013. Generalized Newton's methods for the approximation and resolution of frictional contact problems in elasticity. *Computer Methods in Applied Mechanics and Engineering*, 256, pp. 38–55.

[Rivlin 48] RIVLIN, R.S., 1948. Large Elastic Deformations of Isotropic Materials. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 241(835), pp. 379–397.

[Saad 03] SAAD, Y., 2003. *Iterative Methods for Sparse Linear Systems*, Society for Industrial and Applied Mathematics.

[Santos 95] SANTOS, A. AND MAKINOUCHI, A., 1995. Contact strategies to deal with different tool descriptions in static explicit FEM for 3-D sheet-metal forming simulation. *Journal of Materials Processing Technology*, 50(1-4), pp. 277–291.

[Sauer 07] SAUER, R.A. AND LI, S., 2007. A contact mechanics model for quasi-continua. *International Journal for Numerical Methods in Engineering*, 71(8), pp. 931–962.

[Schenk 04] SCHENK, O. AND GÄRTNER, K., 2004. Solving unsymmetric sparse systems of linear equations with PARDISO. *Future Generation Computer Systems*, 20(3), pp. 475–487.

[Sederberg 04] SEDERBERG, T.W., CARDON, D.L., FINNIGAN, G.T., NORTH, N.S., ZHENG, J. AND LYCHE, T., 2004. T-spline simplification and local refinement. *ACM Transactions on Graphics*, 23(3), pp. 276–283.

[Sekine 10] SEKINE, T. AND OBIKAWA, T., 2010. Normal-Unit-Vector-Based Tool Path Generation Using a Modified Local Interpolation for Ball-End Milling. *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 4(7), pp. 1246–1260.

[Shi 05] SHI, M.F., 2005. Specification For Benchmark Materials. In *AIP Conference Proceedings*. pp. 1173–1178.

[Shim 00] SHIM, H. AND SUH, E., 2000. Contact treatment algorithm for the trimmed NURBS surface. *Journal of Materials Processing Technology*, 104(3), pp. 200–206.

[Simo 85a] SIMO, J.C. AND TAYLOR, R.L., 1985. Consistent tangent operators for rate-independent elastoplasticity. *Computer Methods in Applied Mechanics and Engineering*, 48(1), pp.101–118.

[Simo 85b] SIMO, J.C., WRIGGERS, P. AND TAYLOR, R.L., 1985. A perturbed Lagrangian formulation for the finite element solution of contact problems. *Computer Methods in Applied Mechanics and Engineering*, 50(2), pp. 163–180.

[Simo 88] SIMO, J.C., 1988. A framework for finite strain elastoplasticity based on maximum plastic dissipation and the multiplicative decomposition: Part I. Continuum formulation. *Computer Methods in Applied Mechanics and Engineering*, 66(2), pp. 199–219.

[Simo 92] SIMO, J.C. AND LAURSEN, T.A., 1992. An augmented lagrangian treatment of contact problems involving friction. *Computers & Structures*, 42(1), pp. 97–116.

[Simo 98] SIMO, J.C. AND HUGHES, T.J.R., 1998. *Computational Inelasticity*, New York: Springer-Verlag.

[Stadler 03] STADLER, M., HOLZAPFEL, G.A. AND KORELC, J., 2003. Cn continuous modelling of smooth contact surfaces using NURBS and application to 2D problems. *International Journal for Numerical Methods in Engineering*, 57(15), pp. 2177–2203.

[Stadler 04] STADLER, M. AND HOLZAPFEL, G.A., 2004. Subdivision schemes for smooth contact surfaces of arbitrary mesh topology in 3D. *International Journal for Numerical Methods in Engineering*, 60(7), pp. 1161–1195.

[Taylor 91] TAYLOR, R. AND PAPADOPOULOS, P., 1991. On a patch test for contact problems in two dimensions. In P. Wriggers & W. Wagner, eds. *Computational Methods in Nonlinear Mechanics*. Springer-Verlag, pp. 690–702.

[Taylor 95] TAYLOR, L., CAO, J., KARAFILLIS, A.P. AND BOYCE, M.C., 1995. Numerical simulations of sheet-metal forming. *Journal of Materials Processing Technology*, 50(1-4), pp. 168–179.

[Tekkaya 00] TEKKAYA, A.E., 2000. State-of-the-art of simulation of sheet metal forming. *Journal of Materials Processing Technology*, 103(1), pp. 14–22.

[Tekkaya 09] TEKKAYA, A.E. AND MARTINS, P.A.F., 2009. Accuracy, reliability and validity of finite element analysis in metal forming: a user's perspective. *Engineering Computations*, 26(8), pp. 1026–1055.

[Thuillier 02] THUILLIER, S., MANACH, P.Y., MENEZES, L.F. AND OLIVEIRA, M.C., 2002. Experimental and numerical study of reverse re-drawing of anisotropic sheet metals. *Journal of Materials Processing Technology*, 125-126, pp. 764–771.

[Thürmer 98] THÜRMER, G. AND WÜTHRICH, C.A., 1998. Computing Vertex Normals from Polygonal Facets. *Journal of Graphics Tools*, 3(1), pp. 43–46.

[Todd 86] TODD, P.H. AND MCLEOD, R.J.Y., 1986. Numerical estimation of the curvature of surfaces. *Computer-Aided Design*, 18(1), pp. 33–37.

[Tsai 03] TSAI, M.-C., CHENG, C.-W. AND CHENG, M.-Y., 2003. A real-time NURBS surface interpolator for precision three-axis CNC

machining. *International Journal of Machine Tools and Manufacture*, 43(12), pp. 1217–1227.

[Ubach 13] UBACH, P.-A., ESTRUCH, C. AND GARCIA-ESPINOSA, J., 2013. On the interpolation of normal vectors for triangle meshes. *International Journal for Numerical Methods in Engineering*, 96, pp. 247–268.

[Wang 04] WANG, C.C.L., WANG, Y. AND YUEN, M.M.F., 2004. On increasing the developability of a trimmed NURBS surface. *Engineering with Computers*, 20(1), pp. 54–64.

[Wang 97] WANG, S.P. AND NAKAMACHI, E., 1997. The inside-outside contact search algorithm for finite element analysis. *International Journal for Numerical Methods in Engineering*, 40(19), pp. 3665–3685.

[Wohlmuth 00] WOHLMUTH, B.I., 2000. A Mortar Finite Element Method Using Dual Spaces for the Lagrange Multiplier. *SIAM Journal on Numerical Analysis*, 38(3), pp. 989–1012.

[Wohlmuth 01] WOHLMUTH, B.I., 2001. *Discretization Methods and Iterative Solvers Based on Domain Decomposition*, Berlin: Springer-Verlag.

[Wriggers 01] WRIGGERS, P., KRSTULOVIC-OPARA, L. AND KORELC, J., 2001. Smooth C1-interpolations for two-dimensional frictional contact problems. *International Journal for Numerical Methods in Engineering*, 51(12), pp. 1469–1495.

[Wriggers 06] WRIGGERS, P., 2006. *Computational Contact Mechanics*, Berlin: Springer-Verlag.

[Wriggers 08] WRIGGERS, P., 2008. *Nonlinear Finite Element Methods*, Berlin: Springer-Verlag.

[Wriggers 90] WRIGGERS, P., VU VAN, T. AND STEIN, E., 1990. Finite element formulation of large deformation impact-contact problems with friction. *Computers & Structures*, 37(3), pp. 319–331.

[Wriggers 95] WRIGGERS, P., 1995. Finite element algorithms for contact problems. *Archives of Computational Methods in Engineering*, 2(4), pp. 1–49.

[Wu 05] WU, J., 2005. Specifications For BM2: Underbody Cross Member Panel. In *AIP Conference Proceedings*. pp. 1150–1156.

[Yamada 68] YAMADA, Y., YOSHIMURA, N. AND SAKURAI, T., 1968. Plastic stress-strain matrix and its application for the solution of elastic-plastic problems by the finite element method. *International Journal of Mechanical Sciences*, 10(5), pp. 343–354.

[Yang 08] YANG, B. AND LAURSEN, T.A., 2008. A contact searching algorithm including bounding volume trees applied to finite sliding mortar formulations. *Computational Mechanics*, 41(2), pp. 189–205.

[Yang 08b] YANG, B. AND LAURSEN, T.A., 2008. A large deformation mortar formulation of self contact with finite sliding. *Computer Methods in Applied Mechanics and Engineering*, 197(6-8), pp. 756–772.

[Yastrebov 12] YASTREBOV, V.A., ANCIAUX, G. AND MOLINARI, J.-F., 2012. Contact between representative rough surfaces. *Physical Review E*, 86(3), p. 035601.

[Yastrebov 13] YASTREBOV, V.A., 2013. *Numerical Methods in Contact Mechanics*, John Wiley & Sons.

[Zavarise 09a] ZAVARISE, G. AND DE LORENZIS, L., 2009. A modified node-to-segment algorithm passing the contact patch test. *International Journal for Numerical Methods in Engineering*, 79(4), pp. 379–416.

[Zavarise 09b] ZAVARISE, G. AND DE LORENZIS, L., 2009. The node-to-segment algorithm for 2D frictionless contact: Classical formulation and special cases. *Computer Methods in Applied Mechanics and Engineering*, 198(41-44), pp. 3428–3451.

[Zavarise 98] ZAVARISE, G. AND WRIGGERS, P., 1998. A segment-to-segment contact strategy. *Mathematical and Computer Modelling*, 28(4-8), pp. 497–515.

[Zhong 89] ZHONG, Z.-H. AND NILSSON, L., 1989. A contact searching algorithm for general contact problems. *Computers & Structures*, 33(1), pp. 197–209.

[Zhong 93] ZHONG, Z.-H., 1993. *Finite Element Procedures for Contact-Impact Problems*, Oxford University Press.

[Zhong 96] ZHONG, Z.-H. AND NILSSON, L., 1996. A unified contact algorithm based on the territory concept. *Computer Methods in Applied Mechanics and Engineering*, 130(1-2), pp. 1–16.

[Zhu 13] ZHU, X.-F. ET AL., 2013. A new surface parameterization method based on one-step inverse forming for isogeometric analysis-suited geometry. *The International Journal of Advanced Manufacturing Technology*, 65(9-12), pp. 1215–1227.

[Zienkiewicz 00a] ZIENKIEWICZ, O.C. AND TAYLOR, R.L., 2000. *The Finite Element Method: Its Basis and Fundamentals, Volume 1*, Butterworth-Heinemann.

[Zienkiewicz 00b] ZIENKIEWICZ, O.C. AND TAYLOR, R.L., 2000. *The Finite Element Method: Solid mechanics, Volume 2*, Butterworth-Heinemann.