U.PORTO

FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Multi-scale Modelling of Contact Interfaces

Rodrigo Pinto Carvalho

Advisor:
Prof. Francisco Manuel Andrade Pires
Co-advisor:
Prof. Thiago Doca

In partial fulfilment of the requirements for the degree of
Doctor of Philosophy in the subject of
Mechanical Engineering

Porto, July 2021

CM2S | Computational Multi-Scale
Modeling of Solids and Structures

Institutional support

inegi driving science
& innovation

Funding

FCT Fundação
para a Ciência
e a Tecnologia

REPÚBLICA
PORTUGUESA | CIÊNCIA, TECNOLOGIA
E ENSINO SUPERIOR

NORTE2020
PROGRAMA OPERACIONAL REGIONAL DO NORTE

PORTUGAL
2020

UNIÃO EUROPEIA
Fundo Social Europeu

Supported by the Portuguese Science and Technology Foundation (FCT) under scholarship SFRH/BD/125281/2016

Dedicado aos meus pais,
Leonor e Ilídio

Page intentionally left blank

Abstract

Multi-Scale Modelling of Contact Interfaces

Keywords: Multi-Scale; Homogenization; Signorini Contact; Real Contact Area.

In virtually every structural, mechanical and biological system, it is by way of contact that bodies interact, coming in to touch and interchanging energy. Despite being something mundane, the physics of contact interaction is particularly complex and of multi-scale and multi-physics character. A great example is a fundamental observation that all engineering surfaces are rough, at a sufficiently small scale of inspection. It is intimately connected with the contact interaction properties, determining the morphology and physics at the interface. Therefore, contact problems are inherently non-linear, making them difficult to solve without resorting to approximate techniques. Taking into account that experimental research on this subject can be very expensive/time-consuming, or sometimes even non-feasible, the field of computational contact mechanics is an ever-growing field of research, demanding powerful numerical techniques that are able to accurately predict contact phenomena.

In this work, the main objective is to establish a computational framework for the analysis of tribological systems at multiple scales using the finite element method. The dual-mortar contact algorithm is adopted as the foundation for the contact interaction modelling of deformable bodies under a fully non-linear general hypothesis. Its formulation is presented in detail, providing a comprehensive review on what can be regarded, nowadays, as the state-of-the-art technique for computational contact mechanics. The dual-mortar method provides an accurate and robust alternative for contact modelling, owing to a well-established mathematical foundation. The trade-off, however, often regards its considerable computational complexity. In order to attenuate this difficulty, the algorithm is thus further optimised for the particular case of the Signorini problem (contact with a rigid obstacle). This configuration is commonly found in contact homogenisation and originates from the well-known result from contact mechanics that, under certain conditions, the contact between two rough surfaces can be mapped to the contact between an equivalent single rough surface and a rigid flat. The variational formulation using a Petrov-Galerkin scheme is investigated, also proposing a new definition for the nodal orthonormal moving frame attached to each contact node.

Then, the contact algorithm is embedded within a multi-scale framework for the modelling of rough contact. It is based on the establishment of a representative contact element, which is then separated into multiple smaller scales by splitting the power spectral density function of the rough topography. The statistics of the contact pressure field are integrated into a new multiplicative homogenization scheme, also investigating the definition of the splitting frequencies, the generation of individual scales and the violation of the principle of scale separation. From the numerical results, it is verified that the multi-scale algorithm can reproduce the original problem, with a considerable reduction in computational complexity. This is achieved by the information-passing scheme of the multi-scale algorithm, without compromising the flexibility of the finite element method.

Page intentionally left blank

Resumo

Modelação Multi-Escala de Interfaces de Contacto

Palavras-chave: Multi-Escala; Homogeneização; Contacto Signorini; Área Real de Contacto.

Em praticamente todos os sistemas estruturais, mecânicos e biológicos, é através de contacto que os corpos interagem, trocando energia entre si. A física do contacto é particularmente complexa e com uma forte componente multi-escala e multidisciplinar. Um exemplo pertinente é a observação de que todas as superfícies de engenharia são rugosas, quando visualizadas uma escala suficientemente pequena. Isto está intimamente relacionado com várias propriedades de contacto, determinando a morfologia e física na interface. Adicionalmente, os problemas são inerentemente não-lineares, tornando-os difíceis de resolver sem recorrer a técnicas de aproximação. Tendo em consideração que a investigação experimental neste tópico é geralmente dispendiosa, ou até por vezes impossível, a mecânica do contacto computacional tornou-se uma área de investigação em crescimento contínuo, que procura desenvolver novas técnicas numéricas, capazes de prever com precisão os fenómenos de contacto.

Neste trabalho, o objetivo principal é o estabelecimento de um algoritmo para a análise de sistemas tribológicos a várias escalas através do método dos elementos finitos. O método tem como base o algoritmo dual-mortar para modelar contacto não-linear entre corpos deformáveis. A sua formulação é apresentada em detalhe, sendo feita uma revisão do algoritmo que é considerado o estado-da-arte em modelação de contacto. Este método caracteriza-se pela sua precisão e robustez, ambos garantidos pela sua cuidada formulação matemática. No entanto, a principal desvantagem está relacionada com a considerável complexidade computacional. De forma a atenuar esta dificuldade, o algoritmo é otimizado para o caso particular do problema de Signorini (contacto com um obstáculo rígido). Esta configuração é encontrada com frequência em homogeneização de contacto, tendo como origem o resultado clássico da mecânica do contacto em que, sob determinadas hipóteses, o contacto entre duas superfícies rugosas pode ser mapeado para o contacto entre uma superfície com rugosidade equivalente e um liso rígido. A formulação variacional de Petrov-Galerkin é investigada, sendo também proposta uma nova definição para o sistema de eixos ortonormal associado com cada nó de contacto.

Na segunda parte do trabalho, o algoritmo de contacto é integrado num esquema de modelação de contacto rugoso a várias escalas. Na base do método está o estabelecimento de um elemento de contacto representativo, que é depois separado em várias escalas através da divisão da função de densidade espectral da topografia. É proposto um esquema de homogeneização multiplicativo, que incorpora a natureza estatística do campo de pressões de contacto. É também analisada a definição de frequências de separação, geração de escalas individuais e a violação do princípio de separação de escalas. Os resultados numéricos demonstram que o algoritmo multi-escala é capaz de reproduzir o problema original, beneficiando no entanto de uma redução significativa no tempo de cálculo. Isto é atingido pelo esquema de passagem de informação do algoritmo, sendo garantida a preservação da flexibilidade do método dos elementos finitos.

Page intentionally left blank

Agradecimentos

A realização deste trabalho não seria possível sem a ajuda, motivação e inspiração de uma longa lista de pessoas, que seguramente vai além dos nomes aqui mencionados.

Em primeiro lugar, agradeço de forma especial ao meu orientador Prof. Francisco Pires. A experiência, rigor científico, disponibilidade, visão estratégica e simpatia são algumas palavras que poderei destacar da sua orientação. Estarei sempre grato por acreditar em mim e por todas as condições que me disponibilizou. Em especial, agradeço por sentir que, por muito preenchida que esteja a sua agenda, a porta do gabinete está sempre aberta.

Ao meu co-orientador, Prof. Thiago Doca, agradeço todo o apoio, principalmente na fase inicial deste trabalho. A mecânica do contacto sendo um assunto novo para mim, todas as conversas foram preciosas e fundamentais para conseguir ambientar-me no tema.

Não poderia deixar de destacar também o contributo do Eng. António Carneiro. Sem a sua colaboração, não seria possível concretizar uma parte significativa deste trabalho. Agradeço profundamente toda a sua ajuda, paciência e a constante boa disposição. É um privilégio e uma inspiração poder trabalhar ao lado de alguém tão talentoso, prestável e extremamente humilde.

Deixo também uma palavra de apreço ao Prof. Alexander Popp pela disponibilidade. Foi um orgulho poder colaborar com alguém cujo trabalho na mecânica do contacto sempre foi uma grande referência.

A todos os membros do CM2S, agradeço a boa disposição e paciência para me aturar. Foi um privilégio ter feito parte do nascimento deste grupo, e sem a colaboração de todos seria impossível realizar este trabalho. Aos mais velhos, Igor, Fernando, Miguel e Bernardo, e aos mais novos, Francisca e Rui, o meu profundo obrigado.

À Inês, agradeço todo o carinho, que seguramente tornou este percurso mais fácil. Por todas as conversas aborrecidas sobre trabalho, desculpa e obrigado. É uma felicidade por ter-te a meu lado.

Por último, quero agradecer e dedicar todo este trabalho à minha família. À Juliana e ao Rafa, à Renata, Bruno e Maria, agradeço todo o carinho, paciência e apoio. Aos meus pais, Ilídio e Leonor, não há obrigado suficiente para exprimir a minha gratidão por tudo o que me ensinaram e proporcionaram.

A todos, o meu sincero obrigado.
Rodrigo Pinto Carvalho
Porto, Outono de 2021

Page intentionally left blank

Contents

List of Figures

Chapter 1
Introduction

Chapter 2
From Non-Linear Continuum Mechanics to Dual Mortar Contact

Chapter 3
Discrete Dual Mortar Contact Formulation and Global Solution Strategy

Chapter 4
Computational Implementation Approach and Numerical Validation

Chapter 5
An Efficient Dual Mortar Contact Technique for Rigid/Deformable Interaction

Chapter 6
An Efficient Multi-scale Strategy to Predict the Real Contact Area

Chapter A
Dual mortar contact consistent linearisation

Chapter C
RCE Boundary Conditions at Edges and Vertices

Page intentionally left blank

List of Tables

Chapter 3
Discrete Dual Mortar Contact Formulation and Global Solution Strategy

Chapter 4
Computational Implementation Approach and Numerical Validation

Chapter B
Mapping operations between quadratic finite elements and linear sub-elements

Page intentionally left blank

Nomenclature

General abbreviations

2D	**Two D**imensional
3D	**Three D**imensional
BEM	**B**oundary **E**element **M**ethod
BGT	**B**ush-**G**ibson-**T**homas
CPP	**C**losest **P**oint **P**rojection
DNS	**D**irect **N**umeric **S**imulation
DOP	**D**iscretised **O**rientation **P**olytopes
FE	**F**inite **E**lement
FEM	**F**inite **E**lement **M**ethod
FFT	**F**ast **F**ourier **T**transform
GW-SE	**G**reenwood-**W**illiamson **S**implified **E**lliptic
HDC	**H**eight **D**ifference **C**orrelation
HSM	**H**ertz-**S**ignorini-**M**oreau
IBVP	**I**nitial **B**oundary **V**alue **P**roblem
KKT	**K**arush-**K**uhn-**T**ucker
MS	**M**ulti-**S**cale
NCP	**N**on-linear **C**om**p**lementarity
NTS	**N**ote-**T**o-**S**egment
NURBS	**N**on-**u**niform **R**ational **B**-**s**plines
OO	**O**bject-**O**riented
PDASS	**P**rimal-**D**ual **A**ctive **S**et **S**trategy
PSD	**P**ower **S**pectral **D**ensity
PVW	**P**rinciple of **V**irtual **W**ork
RCE	**R**epresentative **C**ontact **E**lement
RMS	**R**oot **M**ean**S**quare
RVE	**R**epresentative **V**olume**E**lement
STS	**S**egment-**T**o-**S**egment
UML	**U**nified **M**odeling **L**anguage

Code abbreviations

`avg`	**Av**era**g**e
`bnd`	**B**ou**nd**ary
`cnd`	**C**a**nd**idate
`coord`	**Coord**inate
`dim`	**Dim**ension
`el`	**El**ement
`glb`	**Gl**o**b**al
`int`	**Int**erface/**int**egration
`iso`	**Iso**parametric
`lin`	**Lin**earisation
`mort`	**Mort**ar
`nod`	**Nod**e
`nonMort`	**Non-mort**ar
`nrm`	**N**o**rm**al
`seg`	**Seg**ment

Operators and symbols

δ_{ij}	Kronecker delta
$\delta(\bullet)$	Virtual quantity
$\Delta(\bullet)$	Increment of $(\bullet)$
$\mathscr{D}(\bullet)$	Directional derivative of $(\bullet)$
$\nabla_x(\bullet)$	Spatial gradient of $(\bullet)$
A	Finite element assembly operator
$\det(\bullet)$	Determinant of $(\bullet)$
H^1	Sobolev vector space
$\boldsymbol{I}, \mathbf{I}$	Second-order identity tensor, identity matrix
$\ln(\bullet)$	Natural logarithm of $(\bullet)$
$\max(\bullet)$	Maximum of a set of elements $(\bullet)$
$\operatorname{tr}(\bullet)$	Trace operator of $(\bullet)$
$\mathbf{0}$	Zero tensor, zero generic entity
$\emptyset$	Empty set
$(\dot{\bullet})$	Material time derivative of $(\bullet)$
$(\bullet)_{,a}$	Derivative of $(\bullet)$ with respect to a
$(\bullet)^{\mathrm{T}}$	Transpose of $(\bullet)$
$(\bullet)^{-1}$	Inverse of $(\bullet)$
$(\bullet)^{-\mathrm{T}}$	Transpose of the inverse of $(\bullet)$
$\lvert(\bullet)\rvert$	Absolute value of the scalar $(\bullet)$
$\lVert(\bullet)\rVert$	Euclidean norm of tensors or vectors

$\langle\cdot,\cdot\rangle_{\Gamma}$	Duality pairing on Γ
$(\bullet)\cdot(\bullet)$	Single contraction of vectors and tensors
$(\bullet):(\bullet)$	Double contraction of tensors
$(\bullet)\otimes(\bullet)$	Tensor product of vectors or tensors
$(\bullet)\times(\bullet)$	Vector product

Subscripts and superscripts

$(\bullet)^{\mathrm{a}}$	Adjacent element
$(\bullet)_{\mathcal{A}}$	Active set
$(\bullet)_{\mathsf{a}}$, $(\bullet)_{\mathsf{b}}$	Mortar segment end coordinate
$(\bullet)_{\mathrm{c}}$	Contact
$(\bullet)_{\mathrm{cell}}$	Integration cell
$(\bullet)^{\mathrm{den}}$	Denominator
$(\bullet)_{\mathrm{e}}$	Element
$(\bullet)_{\mathrm{ext}}$	External
$(\bullet)_{\mathrm{g}}$	Gauss point (integration segment)
$(\bullet)_{\mathrm{gp}}$	Gauss point (non-mortar element)
$(\bullet)^{\mathrm{h}}$	Discretised version of $(\bullet)$ with the FEM
$(\bullet)^{\mathsf{h}}$	Homogenised quantity
$(\bullet)_{\mathrm{int}}$	Internal
$(\bullet)_{\mathcal{I}}$	Inactive set
$(\bullet)^{i}$	Sub-domain
$(\bullet)^{\mathrm{k}}$	Newton iteration index
$(\bullet)_{l}$	Low cut-off
$(\bullet)_{\mathcal{L}}$	Slip set
$(\bullet)^{\mathrm{m}}$	Mortar
$(\bullet)^{\mathrm{num}}$	Numerator
$(\bullet)_{\mathcal{Q}}$	Stick set
$(\bullet)_{\mathrm{Ref}}$	Reference
$(\bullet)_{r}$	Roll-off
$(\bullet)^{\mathrm{r}}$	Rigid
$(\bullet)_{s}$	High cut-off
$(\bullet)_{\mathrm{split}}$	Split
$(\bullet)^{\mathrm{s}}$	Non-mortar
$(\bullet)^{\mathrm{slip}}$	Slip
$(\bullet)^{\mathrm{stick}}$	Stick
$(\bullet)^{\mathrm{sub}}$	Sub-element
$(\bullet)^{\eta}$	Normal direction
$(\bullet)^{\tau}$	Tangential direction

$(\bullet)^0$	Initial solution for iterative scheme
$(\bullet)_0$	Initial configuration
$(\bullet)^\Phi$	Dual shape functions construction
$(\bullet)_\odot$	Non-mortar element center
$(\bullet)_+$	Positive
$(\bullet)_-$	Negative

Accents

$(\tilde{\bullet})$	Weighted
$(\bar{\bullet})$	Prescribed
$(\check{\bullet})$	Non-unit
$(\hat{\bullet})$	Projected
$(\breve{\bullet})$	Abbreviation

Domains and boundaries

$\mathbb{R}$	Real set
$\mathbb{R}_0^+$	Semi-positive real half space
Ω_e	Finite element subdomain
Ω_0	Domain of a body in the reference configuration
Ω_t	Domain of a body in the current configuration
$\partial\Omega_0$	Boundary of a body in the reference configuration
$\partial\Omega_t$	Boundary of a body in the current configuration
$\partial\Omega_\mathrm{ext}$	Exterior boundary
$\partial\Omega_-$	Negative boundary
$\partial\Omega_+$	Positive boundary
Γ_c	Potential contact partition of boundary in the reference configuration
γ_c	Potential contact partition of boundary in the current configuration
Γ_a	Active contact partition of boundary in the reference configuration
Γ_i	Inactive contact partition of boundary in the reference configuration
Γ_σ	Neumann partition of boundary in the reference configuration
γ_σ	Neumann partition of boundary in the current configuration
Γ_u	Dirichlet partition of boundary in the reference configuration
γ_u	Dirichlet partition of boundary in the current configuration

Kinematics

A	Area
d	Number of spatial dimensions
$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$	Generic base vector of Cartesian coordinate system
$\boldsymbol{F}$	Deformation gradient

J	Determinant of the deformation gradient Jacobian
$\boldsymbol{n}$	Outward unit normal vector in the current configuration
$\boldsymbol{N}$	Outward unit normal vector in the reference configuration
$\boldsymbol{t}$	Surface traction vector in the current configuration
$\boldsymbol{u}$	Displacement vector field
V	Volume
$\boldsymbol{V}$	Material velocity field
$\boldsymbol{x}$	Position of a generic point in the current configuration
$\boldsymbol{X}$	Position of a generic point in the reference configuration
φ	Deformation map between the reference and current configuration

Strain, stresses and constitutive laws

$\boldsymbol{C}$	Right Cauchy-Green tensor
$\mathscr{C}$	Fourth-order constitutive tensor
E	Young's modulus
$\boldsymbol{E}$	Green-Lagrange strain tensor
$\boldsymbol{f}$	Force
$\boldsymbol{P}$	First Piola-Kirchhoff stress tensor
$\boldsymbol{t}$	Surface traction in the current configuration
$\boldsymbol{S}$	Second Piola-Kirchhoff stress tensor
ν	Poissoin's ratio
$\boldsymbol{\sigma}$	Cauchy stress tensor
Ψ	Strain energy function

Governing equations

$\boldsymbol{b}$	Body force in the current configuration
m	Mass of a body
t	Time
T	Total simulation time
$\bar{\boldsymbol{t}}$	Prescribed surface traction in the current configuration
$\bar{\boldsymbol{u}}$	Prescribed displacements in the current configuration
$\boldsymbol{\mathcal{U}}$	Vector space of kinematically admissible displacement fields
$\boldsymbol{\mathcal{V}}$	Vector space of virtual displacement fields
$\boldsymbol{\lambda}$	Lagrange multiplier vector
$\delta\Pi$	Virtual mechanical work
ρ	Material density

Contact mechanics

$\mathcal{A}$	Contact interface parameter space
$\mathcal{B}$	Friction bound $(d-1)$-dimensional sphere
g	Gap function in the normal direction
$\boldsymbol{g}$	Gap vector
$\boldsymbol{\mathcal{M}}$	Vector space of admissible Lagrange multiplier
p^{η}	Normal contact pressure in current configuration
$\boldsymbol{\mathcal{P}}$	Tangential plane projection operator
R	Contact mapping operator from the non-mortar to mortar boundary
$\boldsymbol{t}_{\mathrm{c}}$	Contact traction vector in the current configuration
$\boldsymbol{t}^{\tau}$	Tangential contact traction vector in the current configuration
$\mathbf{t}_{\alpha}$	Slip advected bases in the current configuration
$\mathbf{T}_{\alpha}$	Slip advected bases in the material configuration
β	Coulomb's friction law parameter
$\boldsymbol{\eta}$	Outward unit normal vector to the current non-mortar contact interface
μ	Coefficient of friction
$\boldsymbol{\varsigma}$	Position in contact interface parameter space
$\boldsymbol{v}$	Instantaneous relative velocity
$\boldsymbol{v}_{\mathrm{t}}$	Convective tangential relative velocity in the current configuration
$\boldsymbol{v}_{\mathrm{T}}$	Convective tangential relative velocity in the reference configuration
$\boldsymbol{v}^{\tau}$	Tangential relative velocity
$\boldsymbol{\mathcal{W}}$	Trace space of $\boldsymbol{\mathcal{U}}$ to the potential contact boundary $\gamma_{\mathrm{c}}^{\mathrm{s}}$
ψ	Coulomb's friction law slip function
Ψ_{t}	Mapping of contact interface parameter space to the current configuration
Ψ_{0}	Mapping of contact interface parameter space to the reference configuration

Finite element discretisation

$\mathbf{d}$	Global vector of nodal displacements
$\mathbf{d}_{k}$	Nodal displacements of node k
$\mathbf{f}$	Discrete force global vector
$\mathbf{K}$	Stiffness matrix
h	Finite element mesh size
$\mathbf{r}$	Discrete residual global vector
$\mathbf{x}$	Global vector of nodal coordinates in the current configuration
$\mathbf{x}_{k}$	Current coordinates of node k
n^{e}	Total number of elements
n_{nod}	Total number of nodes
N_{k}	Finite element shape function of node k

ε_r	Newton convergence tolerance
$\boldsymbol{\xi}$	Position in finite element parameter space

Mortar methods

a_{jk}	Linear coefficient relating Φ_j with N^{s}_k
$\mathbf{A}^{\Phi}_{\mathrm{e}}$	Coefficient matrix for dual-shape functions construction
$\mathcal{C}$	Clip polygon
$\mathbf{C}$	Discrete mortar contact operator
$\widetilde{\mathbf{C}}$	Contact stiffness matrix
d^{Φ}_{jk}	Component of the auxiliary coefficient matrix $\mathbf{D}^{\Phi}_{\mathrm{e}}$
D_{jk}	Individual contribution of a given integration element to **D**
$\mathbf{D}$	First mortar coupling matrix
$\mathbf{D}^{\Phi}_{\mathrm{e}}$	Auxiliary coefficient matrix for dual shape functions within the non-mortar element e
$\tilde{g}_j$	Weighted gap function at non-mortar node j
$\tilde{\mathbf{g}}_j$	Weighted gap vector at non-mortar node j
J	Generic Jacobian determinant
J^{cell}	Jacobian determinant of mortar integration cell
$J^{\mathrm{s}}_{\mathrm{e}}$	Jacobian determinant of non-mortar element e
J_{seg}	Jacobian determinant of mortar integration segment
$\mathbf{J}$	Discrete mortar projection operator
$\widetilde{\mathbf{K}}$	Effective stiffness matrix including contact
$\breve{\mathbf{K}}$	Abbreviation for the condensed stiffness matrix including contact
M_{jl}	Individual contribution of a given integration element to **M**
m^{Φ}_{jk}	Component of the auxiliary coefficient matrix $\mathbf{M}^{\Phi}_{\mathrm{e}}$
$\mathbf{M}$	Second mortar coupling matrix
$\mathbf{M}^{\Phi}_{\mathrm{e}}$	Auxiliary coefficient matrix for dual shape functions within the non-mortar element e
$\mathcal{M}$	Mortar set
n_{adj}	Number of non-mortar elements adjacent to a given non-mortar node
n_{cell}	Number of integration cells (three-dimensional problems)
n^{m}	Total number of mortar nodes
$n^{\mathrm{m}}_{\mathrm{e}}$	Number of nodes of a mortar element e
n^{s}	Total number of non-mortar nodes
$n^{\mathrm{s}}_{\mathrm{e}}$	Number of nodes of a non-mortar element e
$n_{\mathrm{g}}, n_{\mathrm{gp}}$	Number of Gauss integration points
n^{λ}	Number of nodes carrying Lagrange multiplier degrees of freedom
n^{λ}_{e}	Number of nodes carrying Lagrange multiplier degrees of freedom in a non-mortar element e

$\mathbf{n}_\mathrm{e}$	Vector of standard shape functions in non-mortar element e
N_l^m	Standard FE shape function of mortar node l
N_k^s	Standard FE shape function of non-mortar node k
$\mathcal{N}$	Set not related to $\mathcal{S}$ or $\mathcal{M}$
$\tilde{\mathbf{r}}$	Residual vector including contact
$\breve{\mathbf{r}}$	Residual vector including contact after condensation
$\mathcal{S}$	Non-mortar set
$\tilde{\boldsymbol{u}}_j^\tau$	Weighted tangential relative slip increment at non-mortar node j
$\tilde{\boldsymbol{v}}_j^\tau$	Weighted tangential relative velocity at non-mortar node j
w_g	Weighting factor for Gauss point g
$\mathbf{W}$	Tangent matrix of nodal projections
$\mathcal{X}$	Auxiliary plane for three-dimensional mortar integration
$\mathbf{z}$	Global vector of discrete nodal Lagrange multipliers
$\mathbf{z}_k$	Discrete Lagrange multiplier vector of node k
$\tilde{\beta}_j$	Coulomb's friction law parameter
$\boldsymbol{\zeta}$	Position in mortar integration segment parameter space
$\tilde{\boldsymbol{\eta}}_k$	Averaged unit normal vector at non-mortar node k
$\boldsymbol{\eta}_e^\mathrm{a}$	Unit normal vector to an adjacent non-mortar element e
Λ	Mapping from the element to the sub-element space
$\boldsymbol{\Xi}$	Jacobian of the mapping from the element to the sub-element space
$\tilde{\boldsymbol{\tau}}_k$	Averaged unit tangent vector at non-mortar node k
Φ_j	Dual Lagrange multiplier shape function of node j
$\boldsymbol{\Phi}_\mathrm{e}$	Vector of dual shape functions in a non-mortar element e

Contact constraints enforcement and semi-smooth Newton

$\mathcal{A}$	Set of active non-mortar nodes
$\mathbf{A}$	Contact normal constraint tangent matrices
$b^\eta, \boldsymbol{b}^\tau$	Auxiliary abbreviations for the NCP function in the tangential direction
c^η	Complementarity parameter in normal direction
c^τ	Complementarity parameter in tangential direction
C_j^η	Non-linear complementarity function in normal direction at non-mortar node j
$\boldsymbol{C}_j^\tau$	Non-linear complementarity function in tangential direction at non-mortar node j
$\tilde{\boldsymbol{C}}_j^\tau$	Auxiliary abbreviation for the complementarity function in tangential direction at non-mortar node j
$\tilde{\mathbf{g}}$	Global vector of weighted gaps
$\mathcal{I}$	Set of inactive non-mortar nodes
k	Semi-smooth Newton iteration step

$\mathcal{L}$	Set of slip non-mortar nodes
$\mathbf{L},\mathbf{H}$	Slip contact tangential constraint tangent matrices
$\breve{\mathbf{L}}$	Slip contact tangential constraint tangent matrices after condensation
n^{a}	Number of active contact nodes
n^{stick}	Number of stick contact nodes
n^{slip}	Number of slip contact nodes
$\mathcal{Q}$	Set of stick non-mortar nodes
$\mathbf{Q},\mathbf{T}$	Slip contact tangential constraint tangent matrices
$\breve{\mathbf{Q}}$	Slip contact tangential constraint tangent matrices after condensation
$\mathbf{r}^{\mathrm{stick}}$	Stick contact constraints residual
$\breve{\mathbf{r}}^{\mathrm{stick}}$	Stick contact constraints residual after condensation
$\mathbf{r}^{\mathrm{slip}}$	Slip contact constraints residual
$\breve{\mathbf{r}}^{\mathrm{slip}}$	Slip contact constraints residual after condensation
$\chi_{(\bullet)}$	Characteristic function for the set $(\bullet)$

Mortar-based contact linearisation

$\mathbf{b}_{\mathrm{e}}^{\mathrm{s}}$	Vector of standard shape functions derivatives in non-mortar element e
$\mathbf{B}$	Discrete FE gradient operator
C	Generic auxiliary constant
$\breve{\mathrm{d}}_{jk}^{\Phi}$	Component of the auxiliary matrix $\breve{\mathbf{D}}_{\mathrm{e}}^{\Phi}$
$\breve{\mathbf{D}}_{\mathrm{e}}^{\Phi}$	Auxiliary matrix used to construct $\mathbf{D}_{\mathrm{e}}^{\Phi}$
$\breve{\mathbf{I}}$	Auxiliary matrix used to multiply some quantity by nodal displacement increments
$\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}$	Tangent matrix of the non-mortar segment Jacobian determinant
$\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{s/m}}}$	Tangent matrix of the Gauss point global coordinates
$\mathbf{K}_{g^{\mathrm{h}}}$	Tangent vector of the discrete gap function
$\mathbf{K}_{\eta^{\mathrm{h}}}$	Tangent matrix of the interpolated unit normal vector
$\mathbf{K}_{\xi_{\mathrm{a/b}}^{\mathrm{s/m}}}$	Tangent vector of the integration element end coordinates
$\mathbf{K}_{\xi_{\mathrm{g}}^{\mathrm{s/m}}}$	Tangent vector of the integration Gauss point local coordinate
$\mathbf{K}_{\Phi_{\mathrm{e}}}$	Tangent matrix of dual shape functions within non-mortar element e
$l_{\breve{\eta}_{\mathrm{g}}^{\mathrm{h}}}$	Length of the non-unit interpolated normal vector at a Gauss point
$l_{\breve{\eta}_j}$	Length of the non-unit averaged normal at a non-mortar node j
$l_{\breve{\eta}_{\mathrm{e}}^{\mathrm{a}}}$	Length of the non-unit normal vector to adjacent non-mortar element e
$\breve{\mathrm{m}}_{jk}^{\Phi}$	Component of the auxiliary matrix $\breve{\mathbf{M}}_{\mathrm{e}}^{\Phi}$
$\mathbf{m}$	Vector connecting both mortar segment end coordinates
$\breve{\mathbf{M}}_{\mathrm{e}}^{\Phi}$	Auxiliary matrix used to construct $\mathbf{M}_{\mathrm{e}}^{\Phi}$
$\mathbf{N}$	Discrete FE interpolation matrix
$\mathbf{P}_{\mathrm{e}}$	Matrix gathering all tangent matrices of the averaged unit normals for a given non-mortar element e

$\mathbf{P}_j$	Tangent matrix of averaged unit normal vector at non-mortar node j
$\check{\mathbf{P}}_j$	Tangent matrix of averaged non-unit normal vector at non-mortar node j
$\check{\mathbf{P}}^{\mathrm{a}}_{\mathrm{e}}$	Tangent matrix of the non-unit normal vector to an adjacent element e
$\mathbf{R}$	Rotation matrix
$\mathbf{S}$	Matrix of shape function second derivatives
$\mathbf{W}_j$	Dyadic of the non-unit averaged normal at a non-mortar node j with itself
$\mathbf{W}^h_{\mathrm{g}}$	Dyadic of the non-unit interpolated normal at a Gauss point with itself
$\mathbf{W}^{\mathrm{a}}_{\mathrm{e}}$	Dyadic of the non-unit normal to an adjacent element e with itself
$\mathbf{x}_{\mathrm{e}}$	Vector of nodal coordinates in current configuration for element e
$\check{\boldsymbol{\eta}}_j$	Non-unit averaged normal at non-mortar node j
$\check{\boldsymbol{\eta}}^h_j$	Non-unit interpolated normal at non-mortar node j
$\tilde{\boldsymbol{\eta}}_{\mathrm{e}}$	Vector containing all averaged unit normals of a non-mortar element e
$\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}$	Non-unit normal vector to an adjacent non-mortar element e
$\check{\boldsymbol{\tau}}_j$	Non-unit averaged tangent vector at non-mortar node j
$\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e}}$	Non-unit tangent vector to an adjacent non-mortar element e

Rough contact

A	Nominal contact area
A_{c}	Real contact area
A_{r}	Relative contact area
C_0, C_0'	Power Spectral Density function
H	Hurst exponent
H_{Ref}	Height of the refined layer
H_{Sub}	Total height of the substrate
k	Frequency
n_{b}	Number of bins
n_{n}	Number of samples contained in a bin
n_{s}	Number of scales
L	Length of the RCE
$\bar{p}^{\eta\{1\}}$	Mean contact normal pressure
α	Nayak's parameter
ζ	Bandwidth
λ	Wavelength
σ_z	Root mean square height
Φ^{PSD}	Power Spectral Density function

Chapter 1

Introduction

Contact mechanics has always been a discipline of particular importance in the mechanics of solids. In virtually every structural, mechanical and even biological system, the event of deformable bodies coming in contact with each other is ubiquitous. To a great extent, it is indeed through contact that bodies interact, coming in to touch and exchanging energy (either it is, for example, in the form of mechanical loads, heat or electrical current). This has motivated the in-depth understanding of contact phenomena, especially when taking into account its technological and economic impact. Tzanakis *et al.* (2012) reported that frictional losses are often estimated to cost more than 1 percent of the gross national product in several nations. Another reported example is the power dissipation in tire-road interaction, which is responsible for 20-30% of the total fuel consumption (Nitsche, 2011).

However, the physics of contact interaction, is particularly rich and complex, turning it into a genuinely multi-scale and multi-disciplinary endeavour. Typical types of phenomena taking place at contact interfaces include mechanical (solid and fluid), thermal, electro-magnetic, metallurgical, quantum and others. Even from the simplest macroscopic continuum perspective, contact problems are inherently non-linear. Firstly, the actual contact surface on which bodies interact is unknown beforehand and must be determined as a part of the solution. Secondly, the boundary conditions at the contact interface are complex, typically involving unknown stresses, displacements and other relevant variables. This poses serious difficulties: conceptual, mathematical and computational.

Nonetheless, there are unique formulations of many classes of contact problems in which these difficulties are minimized, making it possible to employ ingenious analysis methods. Nowadays, processes such as deep-drawing and machine parts such as tires, bearings or gears, are only possible because science and engineering have exploited the principles of contact mechanics. Concrete applications that rely on the capacity to predict contact interactions include, for example: analysis of head/disk interfaces in computer magnetic storage devices; characterisation of adhesion and/or relative slip between concrete and reinforcing steel in structural engineering. Additionally, other representative examples, beyond the realm of engineering applications, that are largely dominated by contact and associated physical phenomena include: the analysis of plate tectonics process in continental drift; the study of the capillary flow of red blood cells through blood vessels in biological systems, just to name a few.

1.1 Motivation

As already mentioned, contact problems are inherently non-linear, making them extremely difficult to solve without resorting to approximate modelling techniques. This fact is even more pronounced when additional sources of non-linearity are taken into account, e.g., complex interface phenomena, damage evolution or thermo-mechanical effects. On top of that, if one wants to include the influence of surface roughness on the tribological behaviour, an additional layer of complexity is introduced.

The interest in conducting research on robust and efficient numerical algorithms to solve this class of problems is substantiated for many reasons. First of all, experimental research on this subject can be very expensive and/or time-consuming. Typical examples are the automotive industry simulations (e.g., crash analysis or rolling contact of car tires), which allow to reduce the expensive testing equipment and accelerate the development. Secondly, some experiments can be difficult or even non-feasible to be executed inside a laboratory, for example, the contact of a turbine blade and its frame case or *in situ* observation of micromechanical contact. Altogether, this makes the field of computational contact mechanics an ever-growing discipline among computational mechanics, demanding powerful numerical techniques that, ultimately, allow for a better understanding of complex systems.

1.2 Historical note

This section presents a short overview of the history of the knowledge about tribology; for a complete presentation, see Dowson (1991). In engineering, the study of bodies in relative motion undergoing contact is commonly designated *tribology*. The word, introduced in Jost (1966), is coined to David Tabor and Peter Jost and derives from the root *tribo* (from the Greek term for *rubbing*) and the suffix *logy* (which means *the study of*). Awareness of contact and friction can be traced back to the construction of the Egyptian pyramids (≈ 2600 BCE). The transportation of large blocks was a main technical difficulty, which was tackled by wetting the desert sand to reduce friction. The first recorded systematic study of friction has been credited to Leonardo da Vinci (1452-1519). Several drawings from Leonardo's notebooks, such as the one in Figure 1.1, became iconic among tribologists, showing that da Vinci was fully familiar with the basic tribological concepts of friction, lubrication and wear (see Hutchings [2016] for a historical overview). However, only two hundred years later (in 1699), the two fundamental friction laws were formally enunciated by Guillaume Amontons (1663-1705). These laws were later confirmed and extended to dynamic situations by Charles-Augustin de Coulomb (1736-1806), resulting in the widely known *Coulomb's friction law*. It states that the frictional force acting between two sliding surfaces is proportional to the magnitude of the load pressing the surfaces together, i.e., the forces have a constant ratio, commonly referred to as the coefficient of friction. Furthermore, the frictional force is assumed to be independent of the apparent area of contact and also the velocity (for ordinary sliding speeds). The first analysis of contact from a mathematical standpoint amounts to the Swiss mathematician Leonhard Euler (1707-1783), who assumed triangular section asperities in order to represent surface roughness (Euler, 1750b,a). It was Euler who introduced the symbol μ for the coefficient of friction.

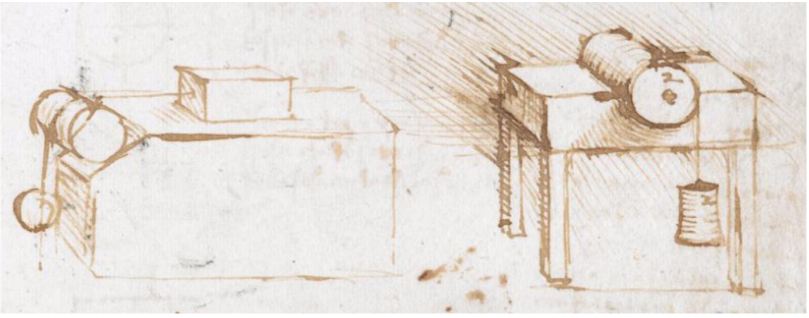


Figure 1.1: Sketch from Leonardo's notebooks showing a set of similar blocks in different orientations on a horizontal plane, from Codex Atlanticus, Biblioteca Ambrosiana, Milan (CA folio 532r c. 1506–8).

The application of the theory of elasticity to contact mechanics was carried out by Heinrich Hertz (1857-1894), who, at the age of 24, solved the problem of two non-conformal elastic bodies undergoing frictionless contact (Hertz, 1881). While studying Newton's optical interference fringes in the gap between two glass lenses, Hertz was concerned with the possible influence of the elastic deformation at the surfaces of the lenses due to the contact pressure between them. His solution is commonly considered to be the cornerstone of modern contact analysis and many of the analytical solutions available in the literature have been building on his work. The seminal models by Johnson *et al.* (1971) and Derjaguin *et al.* (1975), which describe adhesive contact between compliant or hard spheres, are well known examples. A comprehensive overview of the basic principles of contact mechanics, along with the essential analytical solution techniques, can be found in the textbooks by Johnson (1987) and Timoshenko and Goodier (1951).

1.3 Literature overview on computational contact mechanics

Although analytical models are undoubtedly essential to the understanding of the mechanics of contact interactions, their application is restricted to a small number of contact problems. In order to analyse real complex mechanical systems, the use of approximation techniques is usually prefered and, with the ongoing advent of modern computing power over the last decades, this has motivated both the industrial and scientific communities to develop robust algorithms able to treat contact problems. Nowadays, computational contact mechanics is one challenging but essential part of the discipline of computational mechanics, with methods like the Finite Element Method (FEM) and the Boundary Element Method (BEM) being widely used within continuum mechanics modelling. Throughout this work, all ideas and methods of computational contact mechanics are discussed exclusively in the context of the FEM, which proves to be a versatile method for solving boundary value problems in many fields of science and technology. The general FEM literature is abundant, referring, for example, to the monographs by T. J. Hughes (2000), Zienkiewicz *et al.* (2014), and Bathe (2006). A comprehensive overview of computational contact mechanics can be found in the textbooks by Laursen (2013) and Wriggers (2006).

1.3.1 Classical discretisation techniques

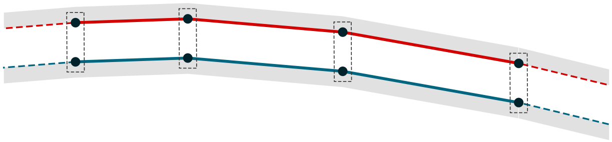
First applications of finite elements to contact problems between deformable bodies can be traced back to the late 1970s and early 1980s. In Francavilla and Zienkiewicz (1975) and T. J. Hughes, Taylor, Sackman, Curnier, *et al.* (1976), contact conditions are incorporated on a very simple, purely nodal basis. The method is restricted to node-matching finite element meshes at the contact interface under the hypothesis of small deformations, see Figure 1.2a. The extension for the general case of large deformations contact is not straightforward, as the actual active part of the contact interface is not known *a priori* and must be determined throughout de solution process. The most widely known technique for this class of problems is the so-called Node-To-Segment (NTS) approach, which is based on discrete, point-wise enforcement of the contact constraints at the finite element nodes, see Figure 1.2b. Early implementations can be found in T. J. Hughes, Taylor, Sackman, and Kanoknukulchai (1976) and J. O. Hallquist (1979), and have been extended to more general cases in J. Hallquist *et al.* (1985), Bathe and Chaudhary (1985), Wriggers, Vu Van, *et al.* (1990), Simo and Laursen (1992), and Laursen and Simo (1993), just to name a few.

Even though widely used—and currently available in most commercial finite element codes with contact formulations—there are well-known limitations in the robustness of NTS algorithms. It has been shown in Papadopoulos and Taylor (1992) that the so-called single-pass algorithms do not satisfy the contact patch test, being observed degradation of the spatial convergence rate. Moreover, due to the non-smoothness of the discretised surfaces, non-physical jumps in the contact forces can occur in problems undergoing large amounts of sliding within the contact interface. Several strategies have been proposed to overcome these issues, which are typically based on higher-order interpolations for the surfaces, see e.g., Wriggers, Krstulovic-Opara, *et al.* (2001), Puso and Laursen (2002), and M. Stadler and Holzapfel (2004). For a comprehensive overview on smooth contact discretisation the interested reader is referred to Wriggers (2006).

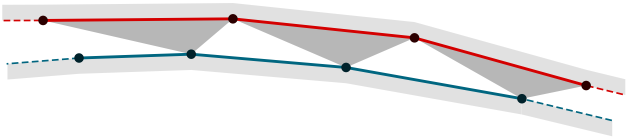
1.3.2 Mortar-based contact formulations

Over the last two decades, alternative formulations for improved robustness have gained considerable attention, particularly the application of the so-called mortar methods for contact problems. Before reviewing the literature on mortar methods, it is noteworthy to mention that the basic idea of using contact segments has been proposed by Simo, Wriggers, *et al.* (1985) and applied to problems involving large deformations, see Papadopoulos and Taylor (1992). The resulting algorithm, commonly referred to as Segment-To-Segment (STS), is based on a thorough sub-division of the contact surface into individual segments for numerical integration, see Figure 1.2c, together with an approximation for the contact pressure. This method can be interpreted as a precursor of mortar-based finite element methods for contact problems.

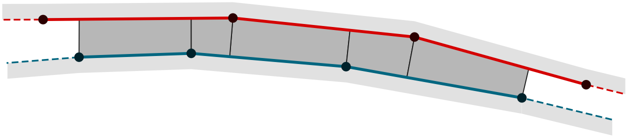
The mortar method (deriving from the old French word *mortier*) has been firstly introduced by Bernardi *et al.* (1993) and constituted a variational operator for the imposition of occurring interface constraints in an optimal weak sense. Although originally introduced as an abstract domain decomposition technique, mortar methods can be applied to partial differential equations for any type of domain, with few restrictions imposed on the discretization procedure (Ben Belgacem and Maday, 1994; Belhachmi and Bernardi, 1994;



(**a**) Node-to-node



(**b**) Node-to-segment



(**c**) Segment-to-segment

Figure 1.2: Graphical representation of different contact discretisations.

F. B. Belgacem, 1999). In the context of contact problems, this allows for a variationally consistent treatment of contact constraints, despite the inevitably non-matching interface discretisations in finite deformations. Early implementations of mortar methods for contact mechanics within the context of small deformations can be found, for example, in F. Belgacem *et al.* (1998), McDevitt and Laursen (2000), and Hild (2000). Subsequently, mortar-based contact formulations aiming at covering the realm of fully non-linear large deformation kinematics were gradually derived and, without claiming the following list to be exhaustive, the reader is referred to Puso (2004), Puso and Laursen (2004a,b), B. Yang, Laursen, and Meng (2005), Fischer and Wriggers (2005, 2006), Hesch and Betsch (2009), and Tur, Fuenmayor, *et al.* (2009) for successful implementations.

A fundamental aspect of mortar methods is that the solution is enforced by using Lagrange multipliers which, in order to preserve the accuracy of the solution, need to be judiciously chosen. Given its importance to the mathematical structure and robustness of the formulation, this topic has been extensively investigated and fundamental aspects such as inf-sup stability conditions and optimal *a priori* error estimates have been carefully examined. The reader is referred to B. I. Wohlmuth (2012) and the references therein for an overview in the context of domain decomposition techniques, and the contributions B. I. Wohlmuth (2011) and B. I. Wohlmuth, Popp, *et al.* (2012) in the context of frictional contact. Nowadays, a popular choice for the Lagrange multiplier space is the so-called dual Lagrange multipliers (B. I. Wohlmuth, 2000). In contrast with the standard choice of mortar methods, dual Lagrange multipliers are of special interest because they generate coupling conditions that are easier to realise, yet without compromising the optimal convergence of the discretisation error. Early applications of this technique to small deformation contact problems can be found in Hüeber and B. I. Wohlmuth (2005), Brunssen *et al.* (2007), Flemisch and B. I. Wohlmuth (2007), and Hüeber, G. Stadler, *et al.* (2008). In Hartmann, Brunssen, *et al.* (2007), the first steps towards the extension to finite deformations were derived in the context of shell finite elements, with the full extension for general contact problems being carried out by Popp, Gee, *et al.* (2009) and Popp, Gitterle, *et al.* (2010). The consistent treatment of particular cases such as partially integrated domains (associated with dropping edge contact problems) can be found in Cichosz and Bischoff (2011) and Popp, Seitz, *et al.* (2013) and second-order finite element interpolation is addressed in Popp, B. I. Wohlmuth, *et al.* (2012). The inclusion of frictional contact constraints amounts to Gitterle *et al.* (2010).

Another benefit of dual Lagrange multipliers is that the formulation is naturally suited for the application of the so-called Primal-Dual Active Set Strategies (PDASS) to enforce the contact constraints. The fundamental idea is to regularise the non-smoothness and multivalued character of contact inequality constraints using Non-linear Complementarity (NCP) functions. This makes it possible to apply a Newton-Raphson type algorithm, comprising in an integrated manner all sources of non-linearities of the problem (including the contact active set search itself). As well-known from the mathematical literature on constrained optimization and applications in contact mechanics, the resulting PDASS algorithms can be reinterpreted as semi-smooth Newton methods (Alart and Curnier, 1991; Strömberg *et al.*, 1996; Christensen *et al.*, 1998; Koziara and Bićanić, 2008).

For completeness, aside from mortar methods, a few alternative discretization methods have been proposed as well, see e.g., the contact domain method proposed in Oliver *et al.* (2009) and Hartmann, Oliver, *et al.* (2009). Nonetheless, mortar-based contact formulations have become well established and, in the meanwhile, can arguably be regarded as the state-of-the-art method for computational contact mechanics. Recent developments in the field of finite element mortar methods include to name only a few particularly active research fields:

- Isogeometric analysis using NURBS (De Lorenzis, Temizer, *et al.*, 2011; Temizer, Wriggers, and T. Hughes, 2011, 2012; Seitz, Farah, *et al.*, 2016), see De Lorenzis, Wriggers, and T. J. Hughes (2014) for a review;
- Improved integration schemes (Farah, Popp, *et al.*, 2015; Wilking and Bischoff, 2017);
- Smoothing techniques (Tur, Giner, *et al.*, 2012);
- Complex interface laws, such as wear (Cavalieri and Cardona, 2013; Farah, Gitterle, *et al.*, 2016; Farah, Wall, *et al.*, 2017; Doca and F. Andrade Pires, 2017);
- Mesh adaptivity (Kindo *et al.*, 2014);
- Multigrid methods (Wiesner *et al.*, 2018);
- Elasto-plastic contact (Seitz, Wall, *et al.*, 2018).
- Termo-mechanical contact (Seitz, Popp, *et al.*, 2015; Temizer, 2014; Dittmann *et al.*, 2014; Hüeber and B. Wohlmuth, 2009)

As a final remark, it is important to mention that mortar methods have caveats as any numerical method. The main drawbacks are the higher computational cost associated with the numerical evaluation of mortar integrals and the cumbersome consistent linearisation within Newton-Raphson iterative procedures.

1.4 Contact modelling across scales

The contact modelling techniques mentioned so far can be viewed as a particular case of solid and structural mechanics, in which a purely macroscopic viewpoint based on continuum assumptions is adopted; the actual interface is thought of as a surface with no thickness. However, the mathematical notion of a surface (i.e., two-dimensional manifold embedded in three-dimensional space) is an idealisation of a more complex reality, where multiple time and length scales phenomena coexist. It is well-known that, at a sufficiently small scale of observation, all engineering surfaces are, in fact, rough (Bowden and Tabor, 2001). For instance, Figure 1.3 shows experimental observations revealing that the real contact area is much smaller than the apparent area. Typically, this is due to inevitable abnormalities during manufacturing processes. Moreover, the surface can be covered by a series of surface films, such as oxide, contaminant and lubricant layers, which further increases the complexity of the physics occurring at the interface.

This has stimulated an active research field since the early 1970s, as surface roughness is deeply connected with the contact properties, determining the morphology and mechanical response at the interface. Roughness can be directly related to the performance of a mechanical component, since the irregularities on the surface may create nucleation points for the formation of cracks or corrosion. Examples include the frictional response of sliding contacts (Wriggers and Reinelt, 2009; Stupkiewicz *et al.*, 2014; Wagner, Wriggers, Veltmaat, *et al.*, 2017), wear (Kubiak *et al.*, 2011), electrical contact resistance (Greenwood, 1966), adhesion (Fuller and Tabor, 1975; Pastewka and Robbins, 2014) or thermal contact resistance (Mikic and Rohsenow, 1966; Temizer, 2011; Anciaux and Molinari, 2013), just to name a few.

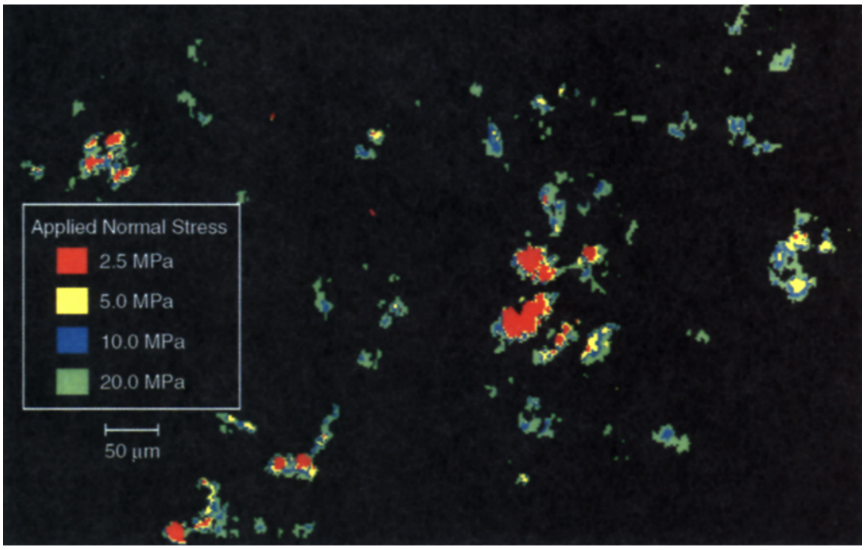


Figure 1.3: Photomicrographs by Dieterich and Kilgore (1994) showing the real contact area between rough acrylic plastic with increasing time of stationary contact at 10 MPa normal stress. In red is represented the area of real contact after 1 s, while the yellow and blue areas are the increase in the area of real contact after 100 s and 10000 s, respectively.

Throughout the decades, the two main approaches to understand the effect of roughness on contact have been:

- Analytical models based on experimental observations, phenomenological considerations and basic principles of mechanics;
- Numerical models, e.g., based on the BEM, the FEM or discrete methods based on molecular dynamics.

The multi-asperity analytical models have been, arguably, the most prominent methodology. Typically, in these techniques, the response function for the normal stresses and the true contact area is defined by using the solution of Hertz for elastic contact and/or other more complex mathematical concepts (such as self-affine surfaces) as the theoretical foundation. The pioneering work of Greenwood and Williamson (1966) is based on a distribution of heights and spheres, which has been the basis for several extensions, see, e.g., Greenwood (2006), Bush *et al.* (1975), and Afferrante *et al.* (2012). This type of analysis has also been carried out to study other contact phenomena, e.g. friction anisotropy Mróz and Stupkiewicz (1994) and Carbone, Lorenz, *et al.* (2009). Another popular analytical model amounts to Persson (2001b), which establishes a scaling approach that includes the presence of roughness on successive length scales; see Persson *et al.* (2005) for a review. For a comprehensive overview and extensive list of references on the modelling and simulation of tribology across scales, the interested reader is referred to the recent overview by Vakis *et al.* (2018).

1.4.1 Computational contact homogenisation

While effectively capturing the influence of various parameters on the frictional response, the aforementioned analytical models exhibit clear limitations when a quantitative prediction is sought. Generally, the set of assumptions is rather restrictive and might compromise the results for a wide range of conditions encountered in engineering applications, see Manners and Greenwood (2006), Carbone and Bottiglione (2008), Paggi and Ciavarella (2010), and V. A. Yastrebov, Anciaux, *et al.* (2015) for a discussion on the subject. This fact, allied to the growth of computational power and contact algorithms over the last decades, has fostered the numerical modelling of rough contact, which has become increasingly adopted. In comparison with analytical models, these techniques are, in general, more flexible, allowing for a more relaxed set of hypotheses and also for the direct analysis of topography realizations. Examples of existing formulations include the FEM (Hyun *et al.*, 2004; Pei *et al.*, 2005; V. A. Yastrebov, Durand, *et al.*, 2011; Stupkiewicz *et al.*, 2014; Couto Carneiro *et al.*, 2020), the BEM (Stanley and Kato, 1997; Polonsky and Keer, 1999; V. A. Yastrebov, Anciaux, *et al.*, 2015) and discrete methods based on molecular dynamics (Campañá and Müser, 2006; C. Yang *et al.*, 2006; Akarapu *et al.*, 2011).

The foundation of computational rough contact models can be traced back to the concept of homogenisation. By considering the interaction mechanisms at the microscale, the objective is to provide a refined description of the macroscopic properties by linking the two scales and, ultimately, an improved understanding of physical phenomena. The major advantage is that more complex physics can be elegantly incorporated in the formulation, with the potential to match experimental data both qualitatively and quantitatively well. The fundamental concept behind homogenisation is the *separation* *of scales*, which is fulfilled whenever two physical entities of the body have very different characteristic sizes—thus introducing the notion of macro-scale and micro-scale. The basic idea is to establish a so-called representative element (Hill, 1963) and then solve its equilibrium problem assuming a set of admissible boundary conditions, e.g., by means of the FEM. The macroscopic continuum quantities (often referred to as homogenised quantities) are typically linked to their micro-scale counterpart fields by means of some kind of averaging process. This type of analysis has been extensively applied to the analysis of heterogeneous bulk materials, which establishes the so-called Representative Volume Element (RVE), see e.g., Moulinec and Suquet (1998), Michel *et al.* (1999), Ghosh and Moorthy (1998), and Pinto Carvalho, Rodrigues Lopes, *et al.* (2018); see the overview by Matouš *et al.* (2017).

Despite being widely applied within the context of the volumetric modelling of the bulk of deformable solids, the concept of computational homogenisation can be applied to interfaces as well (Stupkiewicz, 2007). When thinking about contact roughness, an essential observation is that the characteristic dimension l of roughness features is much smaller than the dimension L of the contacting bodies, i.e., $L \ll l$; see Figure 1.4 for an illustration. This reassembles the principle of separation of scales and, therefore, naturally fits a homogenisation framework. The physics occurring at both scales, despite being intrinsically connected, is governed by different phenomena. At the micro-scale, contact interaction occurs at small spots, leading to a highly inhomogeneous deformation field governed by asperity interaction. At the macro-scale, it is the average contact traction within the contact interface that influences the overall deformation of the bodies. Contact homogenisation is based on the establishment of a Representative Contact Element (RCE). It explicitly describes the topography, constitutive properties and local interaction mechanisms of the contact surfaces, allowing to predict macroscopic contact properties of interest, including friction, contact compliance, real contact area, thermal or electrical contact conductance, wear, just to name a few. In the end, it allows replacing highly complex and inhomogeneous surface layers of bodies in contact with a homogenized interface, described by averaged properties obtained from the analysis of the micro-scale (Stupkiewicz, 2007).

The main challenge for numerical models is related to the topographies' discretization, containing geometrical features that span a wide range of scales, making the models computationally demanding. For this reason, the most popular numerical technique among the scientific community has been the BEM—precisely characterised by its reduced computational cost. However, its main disadvantage is that the bulk of the material is modelled with fundamental solutions, which are only fulfilled for simple situations, e.g., elastic half-spaces under small deformations. Thus, while being extremely powerful in understanding the compression process under certain loading conditions and judging the precision of analytical models, micromechanical contact is intrinsically a multi-physical problem with many phenomena occurring at the contacting interface. As a consequence, the application of the BEM to more complex scenarios becomes compromised (Vakis *et al.*, 2018). One of the most general and robust numerical methods is the FEM, which can deal with several non-linearities simultaneously. However, this is achieved with considerable computational effort. For this reason, the adoption of the finite element method within the frictionless elastic contact analysis is more limited.

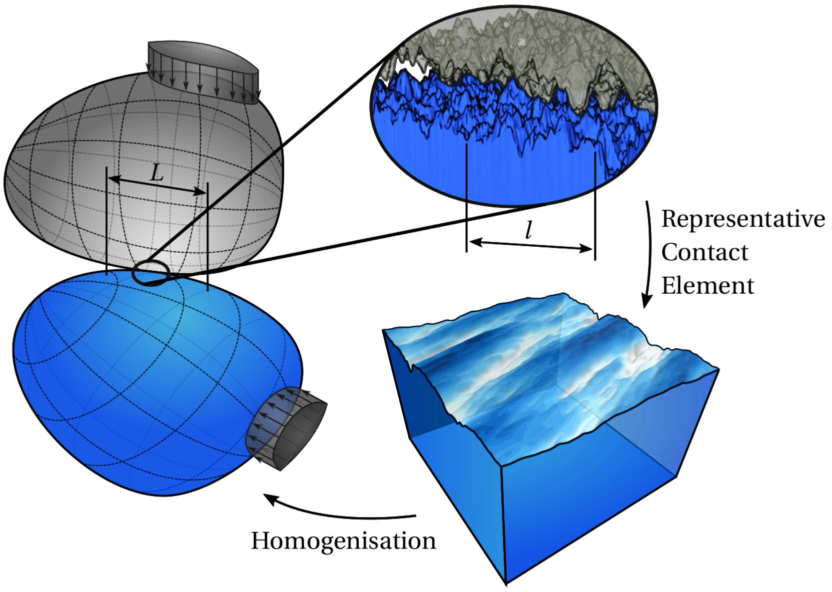


Figure 1.4: Scale separation in the contact of rough bodies.

Early applications of contact homogenisation combined with the FEM can be found in Tworzydlo *et al.* (1998), where contact homogenisation is applied in the context of asperity-based small deformation contact. The macroscopic frictional and normal contact behaviour of randomly rough surfaces is analysed through a combination of finite element analysis of the surface asperities and statistical homogenisation techniques. In Haraldsson and Wriggers (2000), the load and slip distance dependency of frictional behaviour between concrete and soil is investigated, considering elasto-plastic asperities under the hypothesis of small strains. Already accounting for finite deformations, Bandeira, Wriggers, *et al.* (2004) and Bandeira, Pimenta, *et al.* (2008) derived a micro-mechanically motivated law to describe the variation of normal pressure during contact, by carrying out three-dimensional finite element analyses of statistical surface models. Some notable contributions using the finite element method include the works by Hyun *et al.* (2004), Pei *et al.* (2005), and V. A. Yastrebov, Durand, *et al.* (2011). For completeness, other applications of computational contact homogenisation found in the literature using the FEM include:

- Prediction of the macroscopic thermal response of contact interfaces between rough surfaces (Temizer and Wriggers, 2010b; Temizer, 2014; Temizer, 2016);
- Friction anisotropy (Stupkiewicz *et al.*, 2014);
- Eletrical contact (V. A. Yastrebov, Cailletaud, *et al.*, 2015);
- Fretting wear (Pereira *et al.*, 2017);
- Lubrication (Zhu and Y.-Z. Hu, 2001);

It is noteworthy to mention that another concept closely related to contact homogenisation is that of concurrent and hierarchical multi-scale modelling—in which two (or more) different models associated with different scales exchange data continuously in a fully coupled way. The work by Temizer and Wriggers (2008) uses this strategy to model the contact interaction between a flat rubber block and a rigid base containing moving particles, also studying several aspects related to the RCE definition. Additional phenomenological phenomena like visco-elastic constitutive behaviour, non-circular particles and homogenisation studies can be found in Temizer and Wriggers (2010a) and Temizer (2013). One of the main challenges of these techniques regards the identification of the entities and mechanisms relevant to the phenomena to be modelled.

1.4.2 Multi-scale rough contact modelling

While the contact homogenisation techniques can upscale the influence of inhomogeneities at the micro-scale to the macroscopic contact response, the problem of modelling roughness at the micro-scale is still a major difficulty in terms of discretisation requirements. This has originated what can be classified as another multi-scale contact modelling class, in which the concepts of multi-scale modelling are applied to the micro-scale itself. On its foundation is the consideration that the geometrical features of rough profiles extend along with different length scales. In the work of Wriggers and Reinelt (2009), the frictional contact between a flat rubber block and a rigid rough surface is approached with a formulation based on the Height Difference Correlation (HDC) function. The rough surface is approximated by a sinusoidal function at each scale, whose amplitude is selected from the spectrum of the HDC function. The main result of each scale is a micromechanical friction law that can be inserted at larger scales. The algorithm has a recursive character, deriving an effective macroscopic coefficient of friction function dependent on the sliding velocity and normal pressure. Despite showing partially fair agreement with experimental results, the downside of this technique is that the number and frequency of the sinusoidal contributions in the contact description are not well-established in the literature. In De Lorenzis and Wriggers (2013), this strategy is further extended, and the numerical model improved by employing isogeometric analysis using a mortar contact formulation. However, it also reports the quality and convergence of the results to be difficult in some cases.

In the recent work by Wagner, Wriggers, Klapproth, *et al.* (2015), a simple information-passing multi-scale scheme is proposed, which, as opposed to the previous approach, directly models roughness with its asperities by its Power Spectrum Density (PSD). The homogenisation procedure is based on the spectral decomposition of the topography. Different scales are obtained by splitting the PSD at different frequencies and filtering the topography's unwanted components. The algorithm's main result is a constitutive frictional law obtained from the results obtained at each scale, following an information-passing scheme based on the evaluation of both the contact pressure and velocity distributions.

One key aspect of the techniques mentioned earlier is that the principle of scale separation is not fulfilled, since roughness covers a continuous spectrum of length scales, i.e., the scales are not naturally separated. The smallest wavelength at a given scale is the largest of the following smaller scale. This situation is pointed out by Wagner, Wriggers, Veltmaat, *et al.* (2017), which is part of the motivation for the simplified homogenization scheme proposed in that contribution (together with the inclusion of adhesion). It consists of applying a smoothing operation to the pressure downscaling procedure. The average contact pressure and the tangential velocity at a given scale are enforced on the following smaller scale. Then, the full-scale results are obtained through multiplicative homogenization by upscaling the friction coefficient and real contact area.

In conclusion, the continuous effort from the scientific and industrial communities to address contact interface problems from many different perspectives resulted in a collection of modelling techniques for a range of different types of problems. However, while nowadays elastic problems of great complexity can be solved successfully at various scales, significant effort is still required to incorporate effects like plasticity, adhesion, friction, wear, lubrication and surface chemistry in tribological surfaces. Although still challenging to develop (mainly due to the multi-disciplinary character), multi-scale and multi-physics models are entirely suitable to address this class of problems.

1.5 Objectives

The main goal of the present work is to develop a computational framework for the analysis of tribological systems at multiple scales using the FEM. The starting point is to review and implement an efficient and robust algorithm for general, fully non-linear contact problems. The class of mortar finite element algorithms is selected, focusing in particular on the so-called dual mortar method. As mentioned, mortar methods feature a well-established mathematical foundation for their application to many different physical problems, including finite deformation contact. Their mathematical foundations comprise mixed variational formulations and easy-to-construct Lagrange multiplier interpolations, satisfying necessary conditions such as inf-sup and mathematical optimality. Dual Lagrange multipliers, in particular, are especially advantageous for mortar-based contact algorithms, as they inherently lead to contact constraints that are easier to realise, while being perfectly suited for the application of efficient primal-dual active set strategies.

Then, attention is shifted towards contact homogenisation. However, an additional stage is introduced, in which the dual mortar contact algorithm is further optimised for the particular case of Signorini contact, i.e., the contact of a deformable body against a rigid foundation. This assumption is a commonly found scenario in contact homogenisation, with origin in a well-known result from contact mechanics. If there is no friction or adhesion between two rough surfaces and the surface slope is small, the elastic contact between two rough surfaces can be mapped to the contact between an equivalent single rough surface and a rigid flat (Johnson, 1987). As mentioned, the main drawback of mortar methods is, arguably, its higher computational cost, especially in three-dimensional problems. Therefore, this motivates the development of strategies to reduce the computational complexity, mainly by exploiting the simplifications associated with the Signorini contact. More specifically, the variational formulation using a Petrov-Galerkin scheme is further investigated, together with a new definition for the nodal orthonormal moving frame attached to each contact node.

The last point addressed in the present work regards the multi-scale modelling of rough contact. The combination of the micromechanical contact problem being approximated with the FEM, with the discretization requirements related to rough surfaces (covering geometrical features that span multiple physical scales), prompts the development of alternative techniques. As a general guideline, this is to be accomplished without affecting the flexibility achieved with the FEM—one of its strongest points. To achieve this goal, a new homogenisation algorithm is proposed to predict in an efficient way the real contact area. It enhances the information-passing scheme of the original algorithm proposed by Wagner, Wriggers, Veltmaat, *et al.* (2017) by considering the contact pressure field's statistics. The cornerstone of this algorithm is the split of the PSD function of the rough topography, in combination with a new multiplicative homogenization scheme. Additionally, the definition of the splitting roughness frequencies, the methodology for the generation of topographies of the individual scales and the numerical investigation of the violation of the principle of scale separation is also investigated.

1.6 Outline

After this introductory chapter covering the motivation and literature overview on computational contact mechanics, at both macro-scale and micro-scale, the remainder of this document is organized as follows.

Chapter 2 – From Non-Linear Continuum Mechanics to Dual Mortar Contact

In this chapter, the relevant governing equations describing the physical laws of large deformation contact mechanics are outlined. The starting point is the general non-linear continuum mechanics framework, introducing the fundamental concepts of kinematics, strain and stress measures, constitutive laws and balance equations. Then, the initial boundary value problem of finite deformation frictional contact is enunciated, discussing the fundamental concepts of continuum contact kinematics, together with normal and frictional contact constraints. The continuum framework is completed with a thorough description of the mathematical variational framework needed to develop weak statements for finite deformational frictional contact. It includes the introduction of the appropriate solution spaces, along with the reformulation of contact constraints as variational inequalities.

Chapter 3 – Discrete Dual Mortar Contact Formulation and Global Solution Strategy

This chapter is dedicated to the application of the FEM to approximate the dual-mortar contact problem derived in Chapter 2. It features a description of critical aspects such as the numerical integration procedure for the mortar integrals and the choice for the discrete Lagrange multiplier spaces. The employed primal-dual active set strategy for the treatment of contact constraints is introduced, followed by the resulting semi-smooth Newton solution method—including a detailed derivation of the required linearisations. Along with the final algorithm, a suitable algebraic representation for the resulting linearised system of equations is derived as well. Closely related to this chapter is Appendix A, which comprises all details concerning the consistent linearisation of mortar contact formulations, including a detailed algebraic representation of all the terms needed for the actual implementation of the algorithm.

Chapter 4 – Computational Implementation Approach and Numerical Validation

This chapter starts with a brief description of the devised implementation of the algorithm, emphasizing the employed object-oriented programming approach. Then, several examples are analysed in order to evaluate and validate the mortar-based contact algorithm, investigating various important properties of the method, such as consistency, convergence rates and capability to deal with simultaneous non-linearities.

Chapter 5 – An Efficient Dual Mortar Contact Technique for Rigid/Deformable Interaction

An efficient version of the dual mortar contact algorithm is proposed in this chapter. In the first stage, the simplifications involved in going from unilateral to rigid contact are described and the formulation of the dual mortar algorithm is derived. Then, the first concept to be analysed in more detail is the Petrov-Galerkin technique, showing its particular importance within rigid contact. The extension of this formulation towards quadratic finite element interpolation is also carried out, by employing a piecewise linear interpolation scheme for the variation of the Lagrange multipliers. The second idea regards a new definition for the nodal orthonormal frame attached to each contact node, based on the projection of frames from the rigid side to the deformable boundary. Lastly, several examples are shown in order to investigate critical aspects of the algorithm, such as spatial convergence rates, consistency and improvement on computational complexity.

Chapter 6 – An Efficient Multi-scale Strategy to Predict the Real Contact Area

The multi-scale approach to contact homogenisation proposed in this work is explained in this chapter. In a first step, the micromechanical rough contact problem is introduced by defining the concept of RCE. It is followed by an explanation of the technique used for the generation of rough boundaries using the power spectrum model. Then, the proposed multi-scale strategy is described in detail, including the topography decomposition method, the process for generating rough topographies for each scale, and the new multiplicative homogenization scheme. The proposed multi-scale formulation is validated, analysing the impact of key aspects such as the discretization error, the number of scales, or bulk material constitutive behaviour. Particular attention is given to the violation of the principle of scale separation by comparing the results obtained with the original full-size problem and the multi-scale solution. The improvement in the multi-scale algorithm’s computational performance over the original problem is quantitatively measured, including a comparison between the proposed method with several solutions available in the literature.

Chapter 7 – Summary and Outlook

The final chapter summarizes the most important results and accomplishments. Important aspects of mortar finite element methods are pointed out, giving insight into future challenges on multi-scale contact modelling as well.

1.7 Contributions

Most of the work developed in this thesis has either been published, submitted, or is currently in preparation for submission. The main contributions are summarized in what follows:

- Overview and detailed derivation of fully non-linear dual mortar contact, including algebraic representations for all the terms involved and necessary for the actual computational implementation.

- Incorporation of the dual mortar contact algorithm within the micro-scale modelling of heterogeneous materials. This work was motivated by the previous experience of the author in volumetric homogenisation, analysing the constitutive behaviour of a two-phase composite with debonded inclusions, and has been presented in the following communication:

  R. Pinto Carvalho, A. M. Couto Carneiro, F. M. Andrade Pires, T. Doca
  Computational Homogenization of Heterogeneous Materials
  in the Presence of Contact Interactions
  13th World Congress on Computational Mechanics (WCCM)
  New York, NY, USA, July 22-27, 2018

- Implementation of an RCE model under fully non-linear assumptions, establishing a computational homogenisation framework. This was the ground foundation for a contribution on the representativeness of the RCE for micromechanical contact of self-affine topographies:

  A. M. Couto Carneiro, R. Pinto Carvalho, F. M. Andrade Pires
  Representative Contact Element Size Determination
  for Micromechanical Contact Analysis of Self-Affine Topographies
  International Journal of Solids and Structures 206, pp. 262–281, 2020
  Reference: Couto Carneiro *et al.* (2020)

- New dual mortar contact method for the efficient modelling of rigid/deformable interactions. This contribution is submitted in a scientific journal and currently under review:

  R. Pinto Carvalho, A. M. Couto Carneiro, F. M. Andrade Pires, A. Popp
  An Efficient Contact Algorithm for Rigid/Deformable Interaction
  based on the Dual Mortar Method
  Reference: Pinto Carvalho, Couto Carneiro, F. M. Andrade Pires, and Popp (2022)

- Assessment on equivalent modelling of rough contact. This contribution ascertains the assumption of mapping the contact between two rough surfaces to the contact between an equivalent single rough surface and a rigid flat, and has been presented in the following communication:

  R. Pinto Carvalho, A. M. Couto Carneiro, F. M. Andrade Pires, T. Doca
  A Multiscale Approach to Equivalent Modelling of Rough Contact
  15th U. S. National Congress on Computational Mechanics (USNCCM)
  Austin, TX, USA, July 28–Aug. 1, 2019

- Efficient multi-scale algorithm for the evaluation of the evolution of the real contact area fraction between rough surfaces:

  R. Pinto Carvalho, A. M. Couto Carneiro, F. M. Andrade Pires, T. Doca
  An Efficient Multiscale Strategy to Predict the Evolution
  of the Real Contact Area of Rough Surfaces
  Tribology International 165, pp. 107255, 2022
  Reference: Pinto Carvalho, Couto Carneiro, F. M. Andrade Pires, and Doca (2022)

Chapter 2

From Non-Linear Continuum Mechanics to Dual Mortar Contact

An introduction of the fundamental governing equations that describe large deformation contact is presented in this chapter. Starting with a brief overview on the essential concepts of continuum mechanics, the important concepts of continuum contact kinematics and the associated contact constraints are introduced. The initial boundary value problem of finite deformation frictional contact is enunciated. Then, the final step in continuum specification of the frictional contact problem in large deformations is accomplished in the second part of this chapter, by deriving weak statements of the problem. The mathematical structure is analysed in a variational context, thus setting the basis for the application of displacement-based finite element discretization methods. Because the equations describing the behaviour of each body do not change from typical non-linear finite deformation problems, special attention is given to additional terms that emerge from the contact interactions.

2.1 Continuum mechanics and governing equations

A brief overview of the fundamental concepts of non-linear continuum mechanics sets the starting point for all other concepts discussed in this thesis. The presentation is kept short, having as objective only the establishment of basic conceptual and notational foundations for the study of computational contact mechanics to follow. For those who are not familiar with and do not have a strong background in these topics, the classical textbooks by Bonet and Wood (2008), Holzapfel (2000), and Gurtin *et al.* (2010), for example, are recommended.

2.1.1 Kinematics

The general continuum mechanics framework departs from the understanding of the geometrical description of motion and deformation kinematics. The classical (Boltzmann) continuum model in a Euclidean space description is assumed. As indicated, the Lipschitz bounded open set $\Omega_0 \subset \mathbb{R}^d$ represents the bounded body in the reference configuration, and point $P \in \Omega_0$ is referred by the position vector $\boldsymbol{X}$. The symbol d stands for the number of spatial dimensions, i.e., $d = 2, 3$. During deformation, each point $\boldsymbol{X}$ is mapped from the reference configuration Ω_0 to the current configuration Ω_t by a bijective non-linear deformation map φ at each time instant t, which tracks its motion,

$$\boldsymbol{x} = \varphi(\boldsymbol{X}, t) \,. \tag{2.1}$$

The function φ is a one-to-one mapping such that superposition of material points is not allowed. The absolute displacement of the material point is described as

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

At this stage, one needs to choose the type of description for the problem. In a Lagrangean (material) description of motion, the independent variable is the position vector in the reference configuration, $\boldsymbol{X}$, which is known *a priori*. The current position $\boldsymbol{x}$ is treated as a dependent variable on the primary unknown to be solved: the deformation map $\varphi(\boldsymbol{X}, t)$, or equivalently the displacement vector $\boldsymbol{u}(\boldsymbol{X}, t)$. In an Eulerian (spatial) description, a specific point in space is monitored, i.e., $\boldsymbol{x}$ is set as an independent variable.

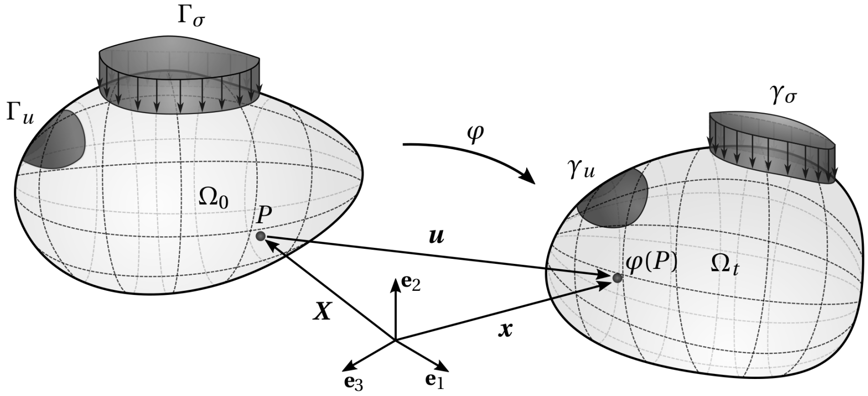


Figure 2.1: Illustration and nomenclature of the deformable body at the reference and current configurations.

The boundary of the deformable body in the reference configuration, denoted as $\partial\Omega_0$, is divided into two open disjoint open subsets: the Neumann partition Γ_σ, with prescribed stresses; the Dirichlet partition Γ_u, with prescribed displacements. These are the boundary conditions of the problem, and verify

$$\begin{aligned} \partial\Omega_0 &= \Gamma_u \cup \Gamma_\sigma \,, \\ \Gamma_u \cap \Gamma_\sigma &= \emptyset \,. \end{aligned} \tag{2.3}$$

The counterparts in the reference configuration read γ_σ and γ_u.

A fundamental measure of deformation and strain of the body within finite deformation is given by the deformation gradient. Here denoted by $\boldsymbol{F}$, it is a second-order tensor defined as the partial derivative of the current configuration $\boldsymbol{x}$ with respect to the reference configuration $\boldsymbol{X}$, i.e.,

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

Here, $\boldsymbol{I}$ is the second-order identity tensor. The deformation gradient is categorized as a two-point tensor, as it features one basis defined in the current configuration and the other in the reference configuration. Geometrically, it describes the mapping of an infinitesimal line element $\mathrm{d}\boldsymbol{X}$ in the reference configuration to a corresponding line element $\mathrm{d}\boldsymbol{x}$ in the current configuration. The Jacobian determinant of the deformation, J, corresponds to the determinant of the gradient tensor

$$J \equiv \det(\boldsymbol{F}) \tag{2.5}$$

and relates the volumes in the reference and current configuration, denoted $\mathrm{d}V_0$ and $\mathrm{d}V$, accordingly, i.e.,

$$\mathrm{d}V = J\mathrm{d}V_0 \,. \tag{2.6}$$

The physical condition of the material not being able to interpenetrate itself requires that

$$J > 0 \,, \tag{2.7}$$

which also ensures that the transformation between reference and current configurations is well-defined.

2.1.2 Strain measures

As a first step in quantifying finite strains, it may be noticed that the following decomposition can always be made:

$$\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} = \boldsymbol{V}\boldsymbol{R} \,. \tag{2.8}$$

Equation (2.8) is often referred to as the *polar decomposition theorem*, according to which any deformation can be split into two parts. A volume-preserving rigid body motion is represented by the orthogonal rotation tensor $\boldsymbol{R}$. A volume-changing part, represented by the positive definite tensors $\boldsymbol{U}$ and $\boldsymbol{V}$, also commonly called right and left stretches, respectively. This means that various strain measures can be derived from the deformation gradient, which may depend on either the reference of current configurations (recall that, being a two-point tensor, depends on both configurations).

When employing a Lagrangean description of motion, defining a strain measure described in the reference configuration makes sense. Therefore, one can define the right Cauchy-Green tensor,

$$\boldsymbol{C} = \boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} \,, \tag{2.9}$$

which proves to be a proper foundation for the definition of a strain measure—mainly because the $\boldsymbol{C}$ is an objective measure, i.e., it is not affected by rigid body motions contained in the deformation gradient. To ensure a consistent definition of zero strain state in the undeformed (reference) configuration, the so-called Green-Lagrange strain tensor $\boldsymbol{E}$ is defined as

$$\boldsymbol{E} = \frac{1}{2}\left(\boldsymbol{C} - \boldsymbol{I}\right) \,. \tag{2.10}$$

For the sake of completeness, it should be mentioned that strain measures are not unique and other alternatives can be derived, e.g., depending only on the current configuration, like the left Euler-Almansi tensor.

2.1.3 Stress measures

In parallel with motion and strain measures, the fundamental concept of internal stress is responsible for describing the internal forces that neighbouring particles of a continuous medium exert on each other. The Cauchy stress tensor, $\boldsymbol{\sigma}$, widely known from the theory of infinitesimal deformations, is a stress measure that states the true internal force in the body $\boldsymbol{f}$ to be

$$\mathrm{d}\boldsymbol{f} = \boldsymbol{\sigma} \cdot \mathrm{d}A\, \boldsymbol{n}\,, \tag{2.11}$$

with $\boldsymbol{n}$ denoting the outward normal vector to the surface element of area $\mathrm{d}A$ in the current configuration. The Cauchy stress tensor is symmetric due to equilibrium considerations, with diagonal and off-diagonal components being interpreted as normal and shear stresses.

Alternatively, one can map the surface element area in the reference configuration onto the resulting spatial force, by defining the first Piola-Kirchhoff stress tensor $\boldsymbol{P}$, i.e.,

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

Introducing the outward normal $\boldsymbol{N}$ to the surface element of area $\mathrm{d}A_0$ in the reference configuration, it follows that

$$\mathrm{d}\boldsymbol{f} = \boldsymbol{P} \cdot \mathrm{d}A_0\, \boldsymbol{N}\,. \tag{2.13}$$

By further mapping the resulting force vector to the reference configuration, it is possible to describe the internal forces in the body purely based on quantities in the reference configuration, leading to the definition of the second Piola-Kirchhoff tensor $\boldsymbol{S}$. Mathematically, it reads

$$\boldsymbol{S} = \boldsymbol{F}^{-1} \cdot \boldsymbol{P} = J\boldsymbol{F}^{-1} \cdot \boldsymbol{\sigma} \cdot \boldsymbol{F}^{-\mathrm{T}}\,. \tag{2.14}$$

Like the Cauchy stress, the second Piola-Kirchhoff tensor is a symmetric tensor. However, it does not have a clear physical interpretation, due to the pull-back operations involved.

As mentioned, a variety of strain and stress measures are available in non-linear continuum mechanics. Notwithstanding, these cannot be combined arbitrarily, but instead defining the strain-stress pairs based on the concept of energy conjugate pairs, which guarantee that the internal work is the same across configurations (Bonet and Wood, 2008).

2.1.4 Constitutive laws

Having introduced the descriptions for strain and stress, the missing element is the constitutive relation relating both quantities, i.e., linking kinematics and material response. Often termed as constitutive models, it can be expressed in terms of a strain energy function Ψ (also known as elastic potential). It only depends upon the current state of deformation, satisfying the physical requirements of objectivity and verification of the second law of thermodynamics. For example, a standard formulation relates the second Piola-Kirchhoff tensor $\boldsymbol{S}$ and the Green-Lagrange strain tensor $\boldsymbol{E}$ as

$$\boldsymbol{S} = \frac{\partial \Psi}{\partial \boldsymbol{E}}\,. \tag{2.15}$$

From this consideration, one can define the fourth-order constitutive tensor, $\mathscr{C}$, specifying the relation between increments in stress and strain to be

$$\mathscr{C} = \frac{\partial \boldsymbol{S}}{\partial \boldsymbol{E}} . \tag{2.16}$$

The choice for the definition of the strain energy function Ψ and constitutive tensor $\mathscr{C}$ goes along with the type of material being modelled and its properties which, for instance, can incorporate hyperelasticity or viscoelasticity. For further details concerning constitutive modelling, the interested reader is referred to the abundant literature, e.g. in Ogden (1997), Simo and T. J. R. Hughes (2006), Holzapfel (2000), and Neto *et al.* (2011).

2.1.5 Balance equations

Beyond the continuous descriptions of motion, strain, stresses and constitutive laws, mechanical systems must always verify the fundamental laws of physics—conservation of mass, equilibrium of linear and angular momentum and energy balance. Additionally, there exist also conservation laws which, altogether, are applicable to any material and must be satisfied at all times. Equilibrium of angular momentum reduces to the symmetry conditions of the Cauchy and second Piola-Kirchhoff stress tensors. As for energy balance, since purely mechanical systems are considered throughout this work (i.e., no other form of energy is taken into account), it becomes redundant with linear momentum equilibrium. The conservation of mass and equilibrium of linear momentum are introduced next. For the sake of brevity and scope of this work, no reference will be given to the thermodynamic principles.

2.1.5.1 Conservation of mass

Every continuum body has mass, here denoted by m, and is a fundamental physical property, directly related to the amount of material that the body contains. The conservation of mass ensures that this quantity remains invariant during motion. Denoting ρ_0 and ρ the body density at the reference and current configurations, respectively, it reads

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_0} \rho_0 \,\mathrm{d}V_0 = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_t} \rho \,\mathrm{d}V = 0 . \tag{2.17}$$

By applying Reynold's transport theorem and taking into account that the reference configuration does not depend on time, the conservation of mass reduces to the local verification of

$$\begin{aligned} \dot{\rho} + \rho \operatorname{div} \dot{\boldsymbol{u}} &= 0 , \\ \dot{\rho}_0 &= 0 . \end{aligned} \tag{2.18}$$

In the above, the notation $(\dot{\bullet})$ denotes the total time derivative and $\operatorname{div}(\bullet)$ the spatial divergence.

2.1.5.2 Equilibrium of linear momentum

The momentum balance principles can be interpreted as generalisations of Newton's first and second principles of motion to the context of continuum mechanics. In this work, one quasi-static problem is considered, which yields the linear momentum to be null. Under this condition, the balance of linear momentum reduces to the equilibrium between internal and external forces. By denoting body forces by $\boldsymbol{b}$ and the surface traction by $\boldsymbol{t}$, both in the current configuration, the equilibrium of forces writes

$$\int_{\Omega_t} \boldsymbol{b}\,\mathrm{d}V + \int_{\partial\Omega_t} \boldsymbol{t}\,\mathrm{d}A = 0\,. \tag{2.19}$$

After application of the Gauss divergence theorem, the linear momentum equilibrium can be conveniently rewritten as

$$\operatorname{div}\boldsymbol{\sigma} + \boldsymbol{b} = \mathbf{0}\,. \tag{2.20}$$

It should be noted that equivalent versions of the conditions above can be derived based on the reference configuration of the body. These equations are also widely known as the strong equilibrium equations.

2.1.6 The strong form of non-linear solid mechanics problems

At this point, all the fundamental concepts for the definition of the Initial Boundary Value Problem (IBVP) are introduced, leading to a set of coupled second-order partial differential equations satisfying a given set of boundary conditions:

Problem 2.1 (Strong form of IBVP of finite deformation non-linear solid mechanics) *For every solid Ω_t, the deformed configuration must verify the system of equations that encompasses both the momentum balance and the boundary conditions of the problem:*

$$\operatorname{div}\boldsymbol{\sigma} + \boldsymbol{b} = \mathbf{0}\,, \quad \text{in } \Omega_t\,, \tag{2.21a}$$

$$\boldsymbol{u} = \bar{\boldsymbol{u}}\,, \quad \text{on } \gamma_u\,, \tag{2.21b}$$

$$\boldsymbol{\sigma}\boldsymbol{n} = \bar{\boldsymbol{t}}\,, \quad \text{on } \gamma_\sigma\,, \tag{2.21c}$$

where $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{t}}$ denote the prescribed displacements and surface tractions at the current configuration Dirichlet and Neumann boundaries, accordingly.

2.2 Contact kinematics

In standard non-linear solid mechanics, both Dirichlet and Neumann boundaries are established *a priori*, meaning that either the prescribed displacements or the external forces are known for the entire boundary of the body throughout the deformation process. The extension of such a framework to account for potential contact interaction is based on the apparent observation that an additional contact boundary exists, which is *unknown* beforehand and changes continuously over time. This introduces an additional source of non-linearity into the continuum mechanics formulation.

From a mathematical problem formulation standpoint, contact methods can be classified into several different categories. The one-body contact problem against a rigid obstacle (commonly termed as *Signorini* contact), self contact and contact involving multiple bodies are examples of problem types within the field of contact mechanics. Nonetheless, attention is given to the general problem formulation of two deformable bodies undergoing large deformations and potentially contact interactions at this stage. It allows for the introduction of all the essential mathematical foundations of contact kinematics and contact constraints.

Therefore, without loss of generality, attention is directed to Figure 2.2, which illustrates contact interactions between two deformable bodies undergoing finite deformations. The superscript s stands for the *non-mortar* body and m for the *mortar* body.[*] As the two bodies experience motion and potentially come into contact, their boundaries $\partial\Omega_0^i$, $i \in \{\mathrm{s},\mathrm{m}\}$, can be divided into three open disjoint subsets, satisfying for each of the bodies i the relations

$$\begin{aligned} \partial\Omega_0^i = \Gamma_u^i \cup \Gamma_\sigma^i \cup \Gamma_\mathrm{c}^i\,, \\ \Gamma_u^i \cap \Gamma_\sigma^i = \Gamma_u^i \cap \Gamma_\mathrm{c}^i = \Gamma_\sigma^i \cap \Gamma_\mathrm{c}^i = \emptyset\,. \end{aligned} \tag{2.22}$$

Recall that Γ_u^i and Γ_σ^i refer to the well-known *Dirichlet* and *Neumann* boundaries, respectively, and Γ_c^i represents the *potential* contact surface. The counterparts in the current configuration Ω_t^s and $\Omega_t^\mathrm{m} \subset \mathbb{R}^d$, also satisfying Equation (2.22), are denoted as γ_u^i, γ_σ^i and γ_c^i, respectively. As already mentioned, one of the challenges of contact modelling is the fact that the so-called active boundary $\Gamma_\mathrm{a}^i \subseteq \Gamma_\mathrm{c}^i$ is unknown beforehand and continuously changes over time. For the sake of completeness, the currently inactive boundary $\Gamma_\mathrm{i}^i = \Gamma_\mathrm{c}^i \setminus \Gamma_\mathrm{a}^i$ should by interpreted as part of the Neumann boundary Γ_σ^i.

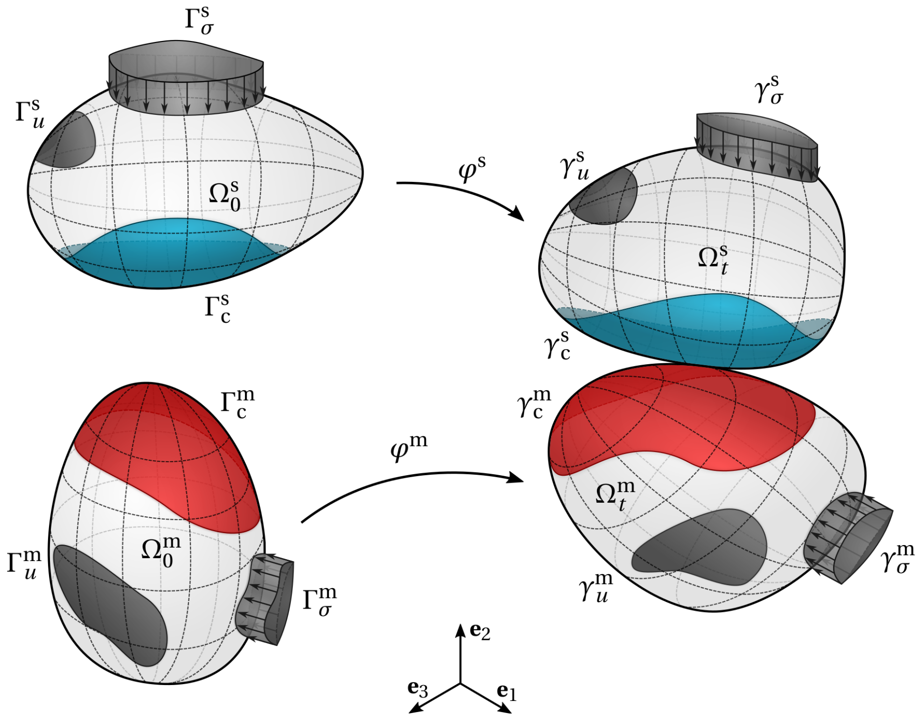


Figure 2.2: Two body unilateral contact problem illustration and nomenclature at the reference and current configuration.

---

[*] The classical nomenclature in contact mechanics of slave/master bodies can be easily recovered as, equivalently, *s* refers to the *slave* and *m* to the *master* domains.

2.2.1 Gap function

The first step towards the establishment of the normal contact constraints requires the definition of fundamental measures for the proximity, potential contact and penetration of two bodies. Therefore, and first of all, it proves convenient to choose one boundary—the non-mortar subset $\gamma_{\mathrm{c}}^{\mathrm{s}}$—as the one used to parametrise contact, such that the relative position of points belonging to this boundary is monitored with respect to the other boundary—the mortar subset $\gamma_{\mathrm{c}}^{\mathrm{m}}$. Consequently, the contact conditions are parametrised by $\boldsymbol{x}^{\mathrm{s}} \in \gamma_{\mathrm{c}}^{\mathrm{s}}$, with the opposing mortar subset $\gamma_{\mathrm{c}}^{\mathrm{m}}$ providing the additional geometric information necessary for the definitions.

Remark. *Since contact interaction takes place at bodies in their current configuration, the foregoing exposition follows an Eulerian description of continuum mechanics.*

For any point $\boldsymbol{x}^{\mathrm{s}} \in \gamma_{\mathrm{c}}^{\mathrm{s}}$, a contact point $\hat{\boldsymbol{x}}^{\mathrm{m}} \in \gamma_{\mathrm{c}}^{\mathrm{m}}$ on the mortar boundary is determined as the projection of $\boldsymbol{x}^{\mathrm{s}}$ onto $\gamma_{\mathrm{c}}^{\mathrm{m}}$ along the *current* outward unit normal vector $\boldsymbol{\eta}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)$ on the *non-mortar* boundary $\gamma_{\mathrm{c}}^{\mathrm{s}}$ (see Figure 2.3). This projection from the non-mortar onto the mortar contact boundaries is represented by the smooth interface mapping operator

$$R: \begin{cases} \gamma_{\mathrm{c}}^{\mathrm{s}} \rightarrow \gamma_{\mathrm{c}}^{\mathrm{m}} \\ \boldsymbol{x}^{\mathrm{s}} \mapsto \hat{\boldsymbol{x}}^{\mathrm{m}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) \end{cases} \tag{2.23}$$

which is assumed to be well defined. Note that the point $\hat{\boldsymbol{x}}^{\mathrm{m}}$ is the result of the smooth mapping φ^{m} of the corresponding point $\hat{\boldsymbol{X}}^{\mathrm{m}}$ in the reference configuration $\Gamma_{\mathrm{c}}^{\mathrm{m}}$, i.e.,

$$\hat{\boldsymbol{x}}^{\mathrm{m}} = \varphi^{\mathrm{m}}\left(\hat{\boldsymbol{X}}^{\mathrm{m}}, t\right) , \tag{2.24}$$

meaning that both $\hat{\boldsymbol{x}}^{\mathrm{m}}$ and $\hat{\boldsymbol{X}}^{\mathrm{m}}$ are associated with $\boldsymbol{x}^{\mathrm{s}}$ via the described projection. This also means that each contact pair possibly changes over time.

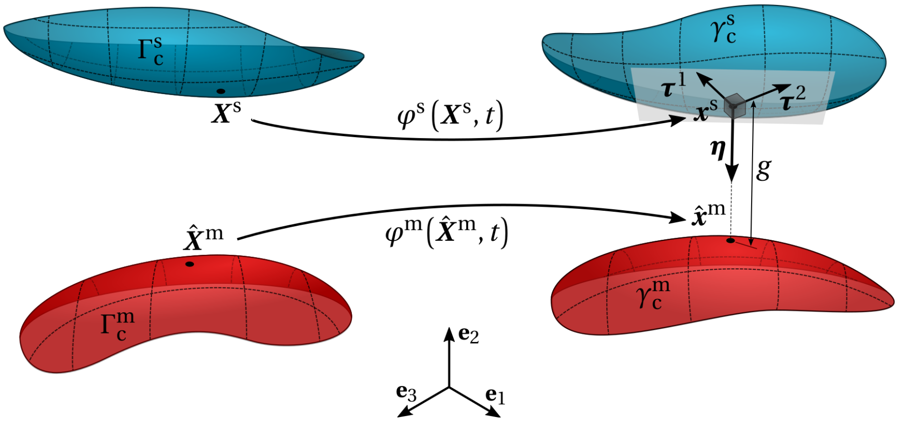


Figure 2.3: Schematic description of basis vectors and gap definition.

The so-called gap function $g\left(\boldsymbol{x}^{\mathrm{s}}, t\right)$ at a point $\boldsymbol{x}^{\mathrm{s}} \in \gamma_{\mathrm{c}}^{\mathrm{s}}$ in the current configuration may then be introduced as

$$g\left(\boldsymbol{x}^{\mathrm{s}}, t\right) \equiv -\boldsymbol{\eta}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) \cdot \left[\boldsymbol{x}^{\mathrm{s}} - \hat{\boldsymbol{x}}^{\mathrm{m}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)\right] . \tag{2.25}$$

This scalar quantity provides the fundamental *geometrical measure* employed to formulate the normal contact constraints. Figure 2.3 can be consulted for a schematic visualisation of its geometrical interpretation. Note that points currently in contact are characterised by the condition $g\left(\boldsymbol{x}^{\mathrm{s}}, t\right) = 0$. The gap vector $\boldsymbol{g}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)$ follows as

$$\boldsymbol{g}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) \equiv \boldsymbol{x}^{\mathrm{s}} - \hat{\boldsymbol{x}}^{\mathrm{m}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) . \tag{2.26}$$

It is noteworthy to mention that the approach followed here differs from the classical *Closest Point Projection* (CPP), which considers the projection based on the outward unit normal to the current *mortar* surface $\gamma_{\mathrm{c}}^{\mathrm{m}}$. Even though the CPP is commonly adopted in the context of note-to-segment discretisation techniques, for mortar finite element discretisation the approach followed here has some practical advantages, as discussed in more detail in Section 3.7. Nonetheless, a comprehensive overview of the closest point projection can be found in Konyukhov and Schweizerhof (2013), which includes an extensive mathematical treatment of potential pitfalls due to non-uniqueness.

2.2.2 Frictional kinematics on interfaces

Paramount to characterising the frictional constitutive behaviour in the tangential direction is the definition of the kinematic quantities on the contact surface. More precisely, the primary kinematic variable of interest is the *relative tangential velocity* between a given point $\boldsymbol{x}^{\mathrm{s}}$ of the non-mortar boundary $\gamma_{\mathrm{c}}^{\mathrm{s}}$ and its opposing point $\hat{\boldsymbol{x}}^{\mathrm{m}}$ on the mortar side $\gamma_{\mathrm{c}}^{\mathrm{m}}$. In a general framework entirely suitable for evolutionary frictional constitutive laws, this quantity serves as a slip rate change measure and is particularly characterised by depending on former states of motion. Even though several alternatives have been proposed in literature—which are usually motivated by the discretisation technique employed—the two most well-known approaches are introduced below: the formulation using *slip advected bases* and the formulation by the *difference of material velocities*.

2.2.2.1 Formulation using slip advected bases

The treatment of frictional contact based on slip advected bases has its foundations in the formulation of NTS algorithms, e.g., in Laursen and Simo (1993). This is mainly motivated by the spatial discretisation characteristics, which prove to be entirely suitable for mathematical treatment in light of differential geometry. When modelling tangential interactions in this framework, it is possible to capture tangential interactions in the covariant form and benefit from a clear physical interpretation of the quantities involved. A fully covariant description for a consistent formulation following this approach can be found in Schweizerhof and Konyukhov (2005).

The formulation of the slip advected bases requires the introduction of additional notation, namely the specification of the underlying parametric representation of the mortar boundary. Both subsets $\Gamma_{\mathrm{c}}^{\mathrm{m}}$ and $\gamma_{\mathrm{c}}^{\mathrm{m}}$ can be considered to be $(d-1)$-dimensional manifolds in $\mathbb{R}^{d}$, with coordinate systems given by the parametrizations

$$\begin{aligned} \Gamma_{\mathrm{c}}^{\mathrm{m}} &= \Psi_{0}^{\mathrm{m}}\left(\mathscr{A}^{\mathrm{m}}\right) , \\ \gamma_{\mathrm{c}}^{\mathrm{m}} &= \Psi_{t}^{\mathrm{m}}\left(\mathscr{A}^{\mathrm{m}}\right) , \end{aligned} \tag{2.27}$$

where the parameter space $\mathscr{A}^{\mathrm{m}} \subset \mathbb{R}^{d-1}$ and the continuous mappings Ψ_{0}^{m} and Ψ_{t}^{m} are assumed smooth. General points of $\mathscr{A}^{\mathrm{m}}$ are denoted as $\boldsymbol{\varsigma}$ and the particular point corresponding to $\hat{\boldsymbol{x}}^{\mathrm{m}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)$ is similarly marked with $\hat{\boldsymbol{\varsigma}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)$.

Since this formulation is established by writing frictional governing equations for a point $\boldsymbol{x}^{\mathrm{s}} \in \gamma_{\mathrm{c}}^{\mathrm{s}}$ opposing the boundary $\gamma_{\mathrm{c}}^{\mathrm{m}}$, it proves convenient to introduce a basis that is *advected* with each point $\boldsymbol{x}^{\mathrm{s}}$ throughout the deformation process—hence the designation *slip advected bases*. Therefore, the following definitions are introduced (see Figure 2.4)

$$\begin{aligned} \mathbf{T}_\alpha &\equiv \Psi^{\mathrm{m}}_{0,\alpha}\left(\hat{\boldsymbol{\varsigma}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)\right) , \\ \mathbf{t}_\alpha &\equiv \Psi^{\mathrm{m}}_{t,\alpha}\left(\hat{\boldsymbol{\varsigma}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)\right) , \quad \alpha = 1, \ldots, d-1 . \end{aligned} \tag{2.28}$$

Herein, the lowercase index α is reserved for quantities expressed on this basis. For notational convenience, the arguments of $\mathbf{T}_\alpha$ and $\mathbf{t}_\alpha$ are suppressed, with the consideration that they are always associated with a non-mortar point $\boldsymbol{x}^{\mathrm{s}} \in \Gamma_{\mathrm{c}}^{\mathrm{s}}$ and evaluated according to its projection to $\gamma_{\mathrm{c}}^{\mathrm{m}}$, which is parametrised by $\hat{\boldsymbol{\varsigma}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)$.

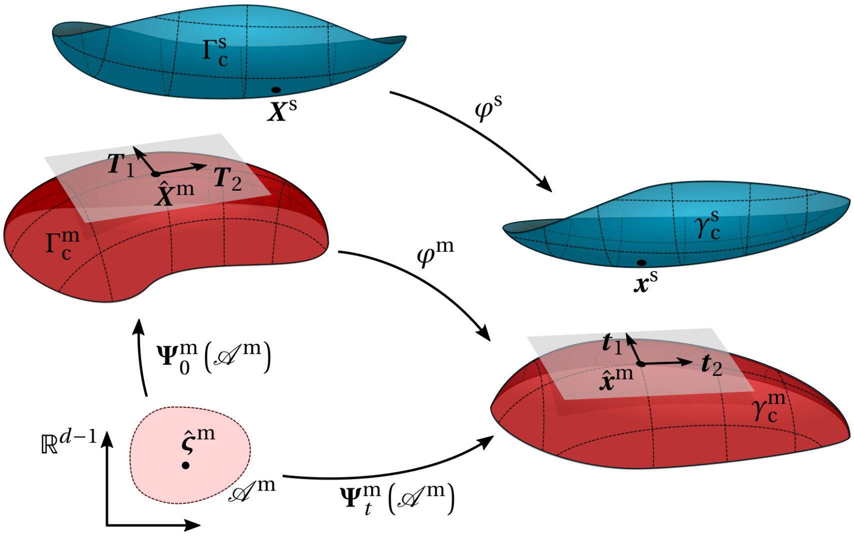


Figure 2.4: Illustration of the parametric representation of the mortar domain and slip advected bases.

Finally, to express the relative velocity by resorting to the slip advected bases, one may consider that both points $\boldsymbol{x}^{\mathrm{s}}$ and $\hat{\boldsymbol{x}}^{\mathrm{m}}$ remain *coincident in space* during perfect sliding. This implies that the gap vector $\boldsymbol{g}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)$ and its *total time derivative* remain zero, viz.

$$\dot{\boldsymbol{g}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) = \mathbf{0} . \tag{2.29}$$

Remark. *Because the fundamental kinematic field involved is the velocity field, a Lagrangian description is exceptionally preferred to facilitate the notation until the end of this section. Nonetheless, it should be kept in mind that the original Eulerian description can be easily recovered, understanding that both representations are related by*

$$\boldsymbol{X}^{\mathrm{s}} = \left[\varphi^{\mathrm{s}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)\right]^{-1} . \tag{2.30}$$

Recalling the definition of the gap vector given in Equation (2.26) (including the application of the chain rule), it follows that

$$
\begin{aligned}
\mathbf{0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\boldsymbol{x}^{\mathrm{s}} - \hat{\boldsymbol{x}}^{\mathrm{m}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left[\varphi^{\mathrm{s}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) - \varphi^{\mathrm{m}}\left(\hat{\boldsymbol{X}}^{\mathrm{m}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right), t\right)\right] \\
&= \mathbf{V}^{\mathrm{s}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) - \mathbf{V}^{\mathrm{m}}\left(\hat{\boldsymbol{X}}^{\mathrm{m}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right), t\right) - \boldsymbol{F}^{\mathrm{m}}\left(\Psi_0^{\mathrm{m}}\left(\hat{\boldsymbol{\varsigma}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\right)\right) \frac{\mathrm{d}}{\mathrm{d}t}\left[\hat{\boldsymbol{X}}^{\mathrm{m}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\right],
\end{aligned}
\tag{2.31}
$$

where $\boldsymbol{F}^{\mathrm{m}}$ is the deformation gradient of the mortar body. This equation can be conveniently rearranged as

$$
\mathbf{V}^{\mathrm{s}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) - \mathbf{V}^{\mathrm{m}}\left(\hat{\boldsymbol{X}}^{\mathrm{m}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right), t\right) = \boldsymbol{F}^{\mathrm{m}}\left(\Psi_0^{\mathrm{m}}\left(\hat{\boldsymbol{\varsigma}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\right)\right) \frac{\mathrm{d}}{\mathrm{d}t}\left[\hat{\boldsymbol{X}}^{\mathrm{m}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\right], \tag{2.32}
$$

and, upon examination, the quantity in the left-hand side is recognized as the difference of the material velocities associated with the points $\boldsymbol{X}^{\mathrm{s}}$ and $\hat{\boldsymbol{X}}^{\mathrm{m}}$, thus physically representing the *instantaneous relative velocity:*

$$
\boldsymbol{v}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) \equiv \mathbf{V}^{\mathrm{s}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) - \mathbf{V}^{\mathrm{m}}\left(\hat{\boldsymbol{X}}^{\mathrm{m}}, t\right). \tag{2.33}
$$

Moreover, it also follows that while sliding is occurring, the relative velocity can also be expressed by the right-hand side. As explained in more detail below, this geometric object comprises a change of contact projection, thus enabling to write the relative velocity in terms of vectors in the tangent spaces of $\Gamma_{\mathrm{c}}^{\mathrm{m}}$ and $\gamma_{\mathrm{c}}^{\mathrm{m}}$.

Remark. *It is noteworthy to mention that the slip rate measures derived from Equation (2.32) are only exact in the case of perfect sliding and persistent contact. In the case of violation of these assumptions (as would occur, for instance, if the point $\boldsymbol{x}^{\mathrm{s}}$ become out of contact), their precise physical interpretation becomes compromised. Nevertheless, typically this approach is still employed for quantifying relative tangential movement (Laursen, 2013).*

In view of Equation (2.32), the *spatial relative velocity* can be defined as

$$
\boldsymbol{v}_{\mathrm{t}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) \equiv \boldsymbol{F}^{\mathrm{m}}\left(\Psi_0^{\mathrm{m}}\left(\hat{\boldsymbol{\varsigma}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\right)\right) \frac{\mathrm{d}}{\mathrm{d}t}\left[\hat{\boldsymbol{X}}^{\mathrm{m}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\right], \tag{2.34}
$$

and, likewise, the *convected relative velocity* as

$$
\boldsymbol{v}_{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left[\hat{\boldsymbol{X}}^{\mathrm{m}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\right] = \dot{\hat{\varsigma}}^{\alpha}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\mathbf{T}_{\alpha}. \tag{2.35}
$$

Expressions for the components $\dot{\hat{\varsigma}}^{\alpha}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ can be found in Laursen (2013).

Examination of the definitions above reveal that $\boldsymbol{v}_{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ represents the pull back operation of $\boldsymbol{v}_{\mathrm{t}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ by the deformation map φ^{m}. Consequently, the spatial relative velocity can also be written as

$$
\boldsymbol{v}_{\mathrm{t}} = \dot{\hat{\varsigma}}^{\alpha}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)\mathbf{t}_{\alpha}, \tag{2.36}
$$

since, by virtue of the chain rule,

$$
\mathbf{t}_{\alpha} = \boldsymbol{F}^{\mathrm{m}}\mathbf{T}_{\alpha}. \tag{2.37}
$$

This means that the spatial relative velocity $\boldsymbol{v}_{\mathrm{t}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$, although associated with a material point $\boldsymbol{X}^{\mathrm{s}} \in \Gamma_{\mathrm{c}}^{\mathrm{s}}$, physically represents a vector expressed in spatial coordinates. Additionally, from Equation (2.35) and Equation (2.36), one can conclude that the components of $\boldsymbol{v}_{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ in the $\mathbf{T}_{\alpha}$ basis are equal to those of $\boldsymbol{v}_{\mathrm{t}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ in the $\mathbf{t}_{\alpha}$ basis. In other words, the components of $\boldsymbol{v}_{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ are the same as those obtained by expressing $\boldsymbol{v}_{\mathrm{t}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ in a basis attached—*convected*—to $\Gamma_{\mathrm{c}}^{\mathrm{m}}$ as it deforms. These two definitions for the relative velocity motivate the formulation of a frictional constitutive modelling framework either in the *spatial kinematic frame*—with the equations cast in terms of $\boldsymbol{v}_{\mathrm{t}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$—or in the *convected kinematic frame*—considering instead the quantity $\boldsymbol{v}_{\mathrm{T}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$. A comprehensive treatment of these formulations has been firstly presented in Laursen (1994).

2.2.2.2 Formulation by the difference of material velocities

The formulation of frictional contact based on the difference of material velocities has been carried out mainly in the context of mortar finite element discretisation. This is due to the fact that, even though not strictly required, mortar-based frictional sliding is typically described by means of *mortar projected kinematics*, which implies the governing frictional laws to be fulfilled point-wise at non-mortar nodes. Examples of this approach can be found in Puso and Laursen (2004b), B. Yang, Laursen, and Meng (2005), Puso, Laursen, and Solberg (2008), and Gitterle *et al.* (2010). Following an alternative technique, Fischer and Wriggers (2006) and Tur, Fuenmayor, *et al.* (2009) proposed a mortar-based framework in which frictional constitutive laws are evaluated at quadrature points. This method emerges as the result of the applied integration scheme, with the particular advantage within the frictional context of exploiting the quantities derived using slip advected bases. However, while benefiting from a strong continuum framework, the integration method possibly leads to discontinuities of the integrand. This has practical implications and numerical experience has shown very sensitive results—especially in the tangential direction—with regard to the employed number of integration points and finite element mesh size ratio between contacting domains (Fischer and Wriggers, 2006; Farah, Popp, *et al.*, 2015). Hence, aiming at a robust frictional contact modelling algorithm fully suitable for further extension, this approach is not followed here and attention is given to mortar projected kinematics instead.

Recalling Equation (2.32), the relative velocity can be represented in terms of material velocities by considering the quantity $\boldsymbol{v}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$ defined in Equation (2.33). For perfect sliding, which assumes the contact points to remain exactly coincident during active sliding, this quantity acts in the tangential plane. However, as already mentioned, when this assumption is not valid, this quantity may lose its precise interpretation as the relative tangential velocity. Therefore, in order to remove a potential normal component, the difference in material velocities is projected in the tangential plane as follows:

$$\boldsymbol{v}^{\tau}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) \equiv \left(\boldsymbol{I}_{d} - \boldsymbol{\eta} \otimes \boldsymbol{\eta}\right) \boldsymbol{v} \,. \tag{2.38}$$

Alternatively, one can expose the relative material velocity components, $v^{\tau_{\alpha}}\left(\boldsymbol{X}^{\mathrm{s}}, t\right)$, along each tangential direction $\boldsymbol{\tau}^{\alpha}$ by writing

$$\boldsymbol{v}^{\tau}\left(\boldsymbol{X}^{\mathrm{s}}, t\right) \equiv \sum_{\alpha=1}^{d-1} v^{\tau_{\alpha}} \boldsymbol{\tau}^{\alpha} \,. \tag{2.39}$$

Here, the vectors $\boldsymbol{\tau}^{\alpha}$ are unit tangent vectors, forming together with the normal vector $\boldsymbol{\eta}$ an orthonormal basis in $\mathbb{R}^{d}$, see Figure 2.3. Their orientation regarding the rotation around the normal vector $\boldsymbol{\eta}$ axis is therein arbitrary. These components are obtained as

$$v^{\tau_{\alpha}}\left(\boldsymbol{X}^{\mathrm{s}},t\right) \equiv \boldsymbol{\tau}^{\alpha}\cdot\boldsymbol{v}\,, \tag{2.40}$$

and can be conveniently arranged in a $(d-1)$-dimensional vector $\boldsymbol{v}^{\tau}$ as follows,

$$\boldsymbol{v}^{\tau}\left(\boldsymbol{X}^{\mathrm{s}},t\right) \equiv \left(v^{\tau_1},\ldots,v^{\tau_{d-1}}\right)\,. \tag{2.41}$$

It is noteworthy to mention that for a proper formulation of rate measures in the realm of finite deformations, such as the tangential relative velocity $\boldsymbol{v}^{\tau}$, frame indifference must be guaranteed. As explained in great detail in Laursen (2013), the definition of the relative tangential velocity according to Equation (2.38) is not frame indifferent. However, this aspect is kept in mind and carefully dealt with during the spatial discretisation of the formulation; further discussion about this topic follows in Section 3.4.2.

2.3 Contact constraints

With all the necessary notation and kinematic measures for finite deformation unilateral contact introduced, attention is now given to the fundamental contact conditions governing the interaction at the boundaries γ_{c}^{i} undergoing contact. These conditions encompass the normal constraints—enforcing the physical requirement of non-penetration and compressive interaction between bodies—and the frictional conditions—describing contact interactions along the tangential direction.

2.3.1 Contact traction

All the contact measures discussed in Section 2.2 are derived from purely kinematical considerations. Nevertheless, the establishment of contact constraints requires a compatible mathematical description of the forces that develop at the active contact region. Therefore, the contact traction $\boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}}\left(\boldsymbol{x}^{\mathrm{s}},t\right)$ acting on the *current* non-mortar contact region $\gamma_{\mathrm{c}}^{\mathrm{s}}$ is introduced and, similarly to the kinematic measures defined in Section 2.2, for three-dimensional problems, the decomposition into normal and tangential components yields

$$\boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}}\left(\boldsymbol{x}^{\mathrm{s}},t\right) = p^{\eta}\boldsymbol{\eta} + \boldsymbol{t}^{\tau}\,. \tag{2.42}$$

The term $p^{\eta}\left(\boldsymbol{\eta},\boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}}\right)$ represents the contact pressure,

$$p^{\eta}\left(\boldsymbol{\eta},\boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}}\right) \equiv \boldsymbol{\eta}\cdot\boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}}\,, \tag{2.43}$$

and the frictional traction vector $\boldsymbol{t}^{\tau}\left(\boldsymbol{x}^{\mathrm{s}},t\right)$ results from the projection on the tangential plane, i.e.,

$$\boldsymbol{t}^{\tau}\left(\boldsymbol{x}^{\mathrm{s}},t\right) \equiv \left(\boldsymbol{I}_{d}-\boldsymbol{\eta}\otimes\boldsymbol{\eta}\right)\boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}}\,. \tag{2.44}$$

Alternatively (and in a similar way as the relative tangential velocity), this last term can also be written as

$$\boldsymbol{t}^{\tau}\left(\boldsymbol{x}^{\mathrm{s}},t\right) \equiv \sum_{\alpha=1}^{d-1} t^{\tau_{\alpha}}\boldsymbol{\tau}^{\alpha}\,, \tag{2.45}$$

i.e., by exposing the tangential components t^{τ_α} of the contact traction $\boldsymbol{t}_\mathrm{c}^\mathrm{s}$ along each tangential direction $\boldsymbol{\tau}^\alpha$, which can also be arranged in the $(d-1)$-dimensional vector

$$\boldsymbol{t}^\tau\left(\boldsymbol{x}^\mathrm{s}, t\right) \equiv \left(t^{\tau_1}, \ldots, t^{\tau_{d-1}}\right) . \tag{2.46}$$

Additionally, as a result of the balance of linear momentum at the contact interface, the traction vector $\boldsymbol{t}_\mathrm{c}^\mathrm{m}\left(\hat{\boldsymbol{x}}^\mathrm{m}, t\right)$ acting on the current mortar contact region $\gamma_\mathrm{c}^\mathrm{m}$ becomes identical to $\boldsymbol{t}_\mathrm{c}^\mathrm{s}\left(\boldsymbol{x}^\mathrm{s}, t\right)$ except for the opposite sign, i.e.,

$$\boldsymbol{t}_\mathrm{c}^\mathrm{m}\left(\hat{\boldsymbol{x}}^\mathrm{m}, t\right) = -\boldsymbol{t}_\mathrm{c}^\mathrm{s}\left(\boldsymbol{x}^\mathrm{s}, t\right) . \tag{2.47}$$

2.3.2 Normal contact constraints

As mentioned, the contact constraints in the normal direction enforce the geometrical condition of non-penetration together with the physical requirement of compressive interaction between bodies. A closer inspection of Equation (2.25) reveals that a positive value $g > 0$ describes points *not* in contact, due to the sign convention. Therefore, the set of *Karush-Kuhn-Tucker* (KKT) optimality conditions† can be stated as

$$g\left(\boldsymbol{x}^\mathrm{s}, t\right) \geq 0 , \tag{2.48a}$$

$$p^\eta\left(\boldsymbol{\eta}, \boldsymbol{t}_\mathrm{c}^\mathrm{s}\right) \leq 0 , \tag{2.48b}$$

$$p^\eta\left(\boldsymbol{x}^\mathrm{s}, t\right) g\left(\boldsymbol{x}^\mathrm{s}, t\right) = 0 , \tag{2.48c}$$

which must hold for all $\boldsymbol{x}^\mathrm{s} \in \gamma_\mathrm{c}^\mathrm{s}$. The first KKT condition imposes the geometrical constraint of non-penetration, whereas the second KKT condition precludes adhesive stresses at the contact interface, thus only allowing for compressive interaction between bodies. Finally, the third constraint, the *complementarity condition*, ensures that the normal contact stress is only generated when contact is occurring, i.e., when $g = 0$. That is, $p^\eta\left(\boldsymbol{\eta}, \boldsymbol{t}_\mathrm{c}^\mathrm{s}\right)$ is forced to be zero when the out-of-contact condition $g > 0$ is verified (the gap is open).

Figure 2.5 gives a schematic representation of the admissible combinations of $p^\eta\left(\boldsymbol{\eta}, \boldsymbol{t}_\mathrm{c}^\mathrm{s}\right)$ and $g\left(\boldsymbol{x}^\mathrm{s}, t\right)$ that fulfil the set of KKT conditions given in Equation (2.48). As can be seen, it defines a contact law not only non-smooth and non-linear, but also of multi-valued character at $g = 0$.

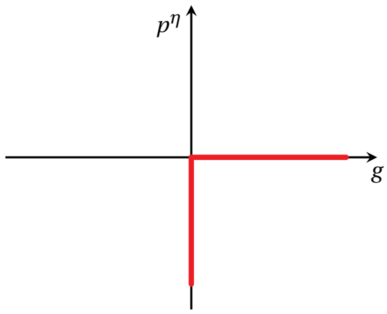


Figure 2.5: Schematic description of Karush-Kuhn-Tucker conditions.

---

† In the parlance commonly used in the contact mechanics community, also referred to as *Hertz-Signorini-Moreau* (HSM) conditions for frictionless unilateral contact.

2.3.3 Frictional contact constraints

Along with normal contact, the understanding and capacity to accurately model friction during relative motion is crucial for many technical systems. Due to the many different physical interface phenomena involved (such as the interaction between rough surfaces, adhesion and chemical interaction, to name a few), friction is recognised as a highly complex discipline among the scientific field of tribology. Despite the various modelling approaches available in the literature, the following overview is restricted to dry friction, which is commonly described in continuum mechanics by the classical Coulomb's friction law.

By introducing the coefficient of friction $\mu \geq 0$, the frictional contact conditions according to Coulomb's law can be stated as follows:

$$\psi\left(\boldsymbol{t}^{\tau}, p^{\eta}\right) \equiv \|\boldsymbol{t}^{\tau}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)\| - \mu |p^{\eta}\left(\boldsymbol{x}^{\mathrm{s}}, t\right)| \leq 0\,, \tag{2.49a}$$

$$\boldsymbol{v}^{\tau}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) + \beta \boldsymbol{t}^{\tau}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) = \boldsymbol{0}\,, \tag{2.49b}$$

$$\beta \geq 0\,, \tag{2.49c}$$

$$\psi\left(\boldsymbol{t}^{\tau}, p^{\eta}\right)\beta = 0\,. \tag{2.49d}$$

In the above, $\psi\left(\boldsymbol{t}^{\tau}, p^{\eta}\right)$ is termed the *slip function* and $\beta \geq 0$ is a scalar parameter. Due to its phenomenological character, Coulomb's friction law allows for an intuitive physical interpretation. The first condition, Equation (2.49a), also known as the *slip condition*, requires that the magnitude of the tangential stress $\boldsymbol{t}^{\tau}$ not exceeds a threshold defined by the coefficient of friction μ times the contact pressure p^{η}. Then, as a result of the *complementarity condition* given in Equation (2.49d), the frictional response at the contact interface can be characterised by two physically distinct regimes:

- **Stick condition:** $\psi < 0$
  The magnitude of the tangential stress $\boldsymbol{t}^{\tau}$ is less than the Coulomb limit and, in view of Equation (2.49b) (with $\beta = 0$), no relative tangential displacement in the contact zone is allowed, i.e., $\boldsymbol{v}^{\tau} = \boldsymbol{0}$.

- **Slip condition**: $\psi = 0$
  The scalar β can be greater than zero and Equation (2.49b), also known as *slip rule*, requires any tangential sliding that does occur to be collinear and in opposing direction with the frictional traction $\boldsymbol{t}^{\tau}$.

A graphical depiction of Coulomb's friction law for one-dimensional sliding (as would happen for two-dimensional problems) is shown in Figure 2.6. Just as KKT conditions in the normal direction, Coulomb friction describes a non-smooth and non-linear constitutive law with a multivalued structure, which has to be taken care of within the solution process. Moreover, a striking analogy between friction and elasto-plasticity formulation can be readily recognised. This has been commonly exploited in the development of numerical algorithms for friction, articulating and adapting well-known methodologies from the elasto-plasticity framework to model tangential sliding in the contact interface, see e.g., Laursen (2013).

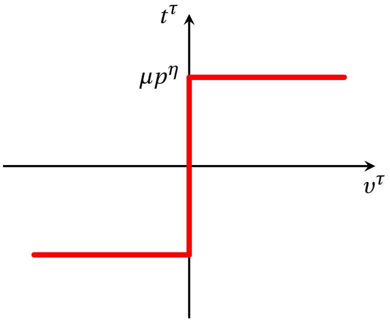


Figure 2.6: Schematic illustration of Coulomb's friction law for one-dimensional sliding.

2.4 Initial boundary value problem of finite deformation frictional contact

The main concepts of continuum contact kinematics have been introduced and, in order to facilitate future reference, the IBVP of finite deformation frictional contact is stated:

Problem 2.2 (The strong form of IBVP of finite deformation frictional contact)
On each sub-domain Ω_t^i, the momentum balance principle is defined by the system of equations

$$\operatorname{div}\boldsymbol{\sigma}^i + \boldsymbol{b}^i = \boldsymbol{0}\,, \quad \text{in } \Omega_t^i\,, \tag{2.50a}$$

$$\boldsymbol{u}^i = \bar{\boldsymbol{u}}^i\,, \quad \text{on } \gamma_u^i\,, \tag{2.50b}$$

$$\boldsymbol{\sigma}^i \boldsymbol{n}^i = \bar{\boldsymbol{t}}^i\,, \quad \text{on } \gamma_\sigma^i\,, \tag{2.50c}$$

and the contact constraints in the normal and tangential directions are given by

$$g \geq 0\,, \quad p^\eta \leq 0\,, \quad g p^\eta = 0 \quad \text{on } \gamma_\mathrm{c}^\mathrm{s}\,, \tag{2.51a}$$

$$\psi \leq 0\,, \quad \boldsymbol{v}^\tau + \beta \boldsymbol{t}^\tau = \boldsymbol{0}\,, \quad \beta \geq 0\,, \quad \beta\psi = 0 \quad \text{on } \gamma_\mathrm{c}^\mathrm{s}\,, \tag{2.51b}$$

accordingly.

2.5 Weak formulation

Building upon the continuous framework laid out so far, in the following the weak statements of the frictional contact problem in large deformations is derived. First of all, appropriate notions of solution space $\boldsymbol{\mathcal{U}}^i$ and weighting space $\boldsymbol{\mathcal{V}}^i$ are introduced as

$$\boldsymbol{\mathcal{U}}^i \equiv \left\{ \boldsymbol{u}^i \in [H^1(\Omega_t^i)]^d \mid \boldsymbol{u}^i = \bar{\boldsymbol{u}}^i \ \text{ on } \gamma_u^i \right\}, \tag{2.52}$$

$$\boldsymbol{\mathcal{V}}^i \equiv \left\{ \delta\boldsymbol{u}^i \in [H^1(\Omega_t^i)]^d \mid \delta\boldsymbol{u}^i = \boldsymbol{0} \ \text{ on } \gamma_u^i \right\}, \tag{2.53}$$

where $H^1(\Omega_t^i)$ represents the Sobolev space, i.e., all functions over Ω_t^i whose values and first derivatives are square integrable over the domain (e.g., refer to Marsden and T. J. R.

Hughes [2012] for additional details). For notational convenience, the collection of mappings defined by the product spaces $\boldsymbol{\mathcal{U}} \equiv \boldsymbol{\mathcal{U}}^{\mathrm{s}} \times \boldsymbol{\mathcal{U}}^{\mathrm{m}}$ and $\boldsymbol{\mathcal{V}} \equiv \boldsymbol{\mathcal{V}}^{\mathrm{s}} \times \boldsymbol{\mathcal{V}}^{\mathrm{m}}$ are introduced as well. With the solution and weighting spaces at hand, the so-called *Principle of Virtual Work* (PVW) can be obtained by multiplying the momentum balance (2.50) with an arbitrary displacement $\delta \boldsymbol{u}^i$ and integrating over Ω_t^i. It leads to

$$\delta\Pi_{\mathrm{int}}\left(\boldsymbol{u},\delta\boldsymbol{u}\right) - \delta\Pi_{\mathrm{ext}}\left(\delta\boldsymbol{u}\right) + \delta\Pi_{\mathrm{c}}\left(\boldsymbol{u},\delta\boldsymbol{u}\right) = 0\,, \quad \forall\, \delta\boldsymbol{u} \in \boldsymbol{\mathcal{V}}\,, \tag{2.54}$$

where $\delta\Pi_{\mathrm{int}}\left(\boldsymbol{u},\delta\boldsymbol{u}\right)$ represents the *internal* virtual work—the virtual work due to internal stresses—and $\delta\Pi_{\mathrm{ext}}\left(\delta\boldsymbol{u}\right)$ the *external* virtual work—the virtual work due to externally applied loadings. Each one of these contributions to the PVW equation is obtained by adding the individual weak forms on each body, meaning that it can be written in full length as

$$\delta\Pi_{\mathrm{int}}\left(\boldsymbol{u},\delta\boldsymbol{u}\right) \equiv -\sum_{i\in\{\mathrm{s},\mathrm{m}\}} \delta\Pi^i_{\mathrm{int}}\left(\boldsymbol{u}^i,\delta\boldsymbol{u}^i\right) \equiv -\sum_{i\in\{\mathrm{s},\mathrm{m}\}} \left[\int_{\Omega_t^i} \boldsymbol{\sigma}^i\left(\boldsymbol{u}^i\right) : \nabla_x\left(\delta\boldsymbol{u}^i\right)\,\mathrm{d}\Omega_t^i\right]\,, \tag{2.55a}$$

$$\delta\Pi_{\mathrm{ext}}\left(\delta\boldsymbol{u}\right) \equiv -\sum_{i\in\{\mathrm{s},\mathrm{m}\}} \delta\Pi^i_{\mathrm{ext}}\left(\delta\boldsymbol{u}^i\right) \equiv -\sum_{i\in\{\mathrm{s},\mathrm{m}\}} \left[\int_{\Omega_t^i} \boldsymbol{b}^i\cdot\delta\boldsymbol{u}^i\,\mathrm{d}\Omega_t^i + \int_{\gamma_\sigma^i} \bar{\boldsymbol{t}}^i\cdot\delta\boldsymbol{u}^i\,\mathrm{d}\gamma_\sigma^i\right]\,. \tag{2.55b}$$

The third term, $\delta\Pi_{\mathrm{c}}\left(\boldsymbol{u},\delta\boldsymbol{u}\right)$, represents the virtual work of contact forces,

$$\delta\Pi_{\mathrm{c}}\left(\boldsymbol{u},\delta\boldsymbol{u}\right) \equiv -\sum_{i\in\{\mathrm{s},\mathrm{m}\}} \delta\Pi^i_{\mathrm{c}}\left(\boldsymbol{u}^i,\delta\boldsymbol{u}^i\right) \equiv -\sum_{i\in\{\mathrm{s},\mathrm{m}\}} \left[\int_{\gamma_{\mathrm{c}}^i} \boldsymbol{t}^i_{\mathrm{c}}\left(\boldsymbol{u}^i\right)\cdot\delta\boldsymbol{u}^i\,\mathrm{d}\gamma_{\mathrm{c}}^i\right]\,, \tag{2.56}$$

which, by enforcing linear momentum across the contact interface, see Equation (2.47), can be conveniently rewritten as

$$\delta\Pi_{\mathrm{c}}\left(\boldsymbol{u},\delta\boldsymbol{u}\right) \equiv -\int_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \boldsymbol{t}^{\mathrm{s}}_{\mathrm{c}}\cdot\left(\delta\boldsymbol{u}^{\mathrm{s}} - \delta\hat{\boldsymbol{u}}^{\mathrm{m}}\right)\,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,. \tag{2.57}$$

In Equation (2.57), the contact traction $\boldsymbol{t}^{\mathrm{s}}_{\mathrm{c}}$ is subjected to contact constraints. Therefore, depending upon the solution method employed, different definitions can be derived; further discussion on this topic follows in the succeeding Section 2.6.

Remark. *Problems of constrained evolution, such as large deformation frictional contact, can be equally addressed in terms of the so-called variational inequalities. This alternative approach consists of ordinarily introducing all the kinematical restrictions of the problem on the solution space itself, thus making the solution and variational spaces equivalent. This leads to a variational inequality, instead of obtaining the weak solution from a minimization problem on a convex set. For a mathematical overview of this approach, the interested reader is referred to Kikuchi and Oden (1988).*

2.6 Treatment of the contact constraints

It is well-known that, with appropriate mathematical abstraction, a wide range of contact problems can be treated in the light of optimization theory as a problem of constrained minimization. Historically, this fact has been exploited to develop several contact algorithms and, among them, the two most widely implemented are the *Lagrange multipliers* and *penalty* methods (Wriggers, 2006; Laursen, 2013). In this work, the focus is on mortar-based finite element methods for contact mechanics which, as explained in more detail in the following, are based on Lagrange multipliers. Nonetheless, for a comprehensive overview of various existing techniques for the treatment of contact constraints, the interested reader is referred to Wriggers (2006) and references therein. Worth referring as well the landmark work of Kikuchi and Oden (1988) for the mathematical framework of contact problems in solid mechanics, especially the theory of variational inequalities.

2.6.1 Lagrange multiplier

As mentioned, mortar methods are based on the introduction of the *Lagrange multiplier vector* $\boldsymbol{\lambda}$ as an additional unknown. This sets the basis for a mixed variational approach, with the Lagrange multiplier vector physically modelling the contact pressure on the contact interface. Alternatively, Lagrange multipliers can be interpreted as the necessary forces to enforce the kinematic contact constraints—in a weak integral sense within the mortar method. In this work, the Lagrange multiplier vector $\boldsymbol{\lambda}$ is set to represent the *negative* contact traction on the non-mortar side, viz.

$$\boldsymbol{\lambda} = -\boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}} \,. \tag{2.58}$$

As the decomposition into normal and tangential components employed in the definition of the contact traction in Equation (2.42), the Lagrange multiplier vector can be understood as the sum of two components as well:

$$\boldsymbol{\lambda} = \lambda^{\eta}\boldsymbol{\eta} + \sum_{\alpha=1}^{d-1} \lambda^{\tau_\alpha}\boldsymbol{\tau}^{\tau_\alpha} \,. \tag{2.59}$$

Here, λ^{τ_α} denote the tangential components of the Lagrange multiplier vector $\boldsymbol{\lambda}$ along each tangential direction $\boldsymbol{\tau}^{\tau_\alpha}$, which as the tangential relative velocity $\boldsymbol{v}^{\tau}$ and tangential contact traction $\boldsymbol{t}^{\tau}$, can also be arranged in the $(d-1)$-dimensional vector

$$\boldsymbol{\lambda}^{\tau} \equiv \left(\lambda^{\tau_1}, \ldots, \lambda^{\tau_{d-1}}\right) . \tag{2.60}$$

By inserting Equation (2.58) into Equation (2.57), the contact virtual work comes as

$$\delta\Pi_{\mathrm{c}}(\boldsymbol{u}, \delta\boldsymbol{u}) \equiv \int_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \boldsymbol{\lambda} \cdot \left(\delta\boldsymbol{u}^{\mathrm{s}} - \delta\hat{\boldsymbol{u}}^{\mathrm{m}}\right) \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \,. \tag{2.61}$$

At this stage, the weak formulation of mortar-based frictional contact problems requires the selection of an appropriate solution space for the Lagrange multiplier. First, let the trace space $\boldsymbol{\mathcal{W}}^{\mathrm{s}}$ represent the restriction of the solution space $\boldsymbol{\mathcal{U}}^{\mathrm{s}}$ to the potential contact boundary $\gamma_{\mathrm{c}}^{\mathrm{s}}$. In terms of functional analysis, this means that $\mathcal{W}^{\mathrm{s}} = H^{1/2}(\gamma_{\mathrm{c}}^{\mathrm{s}})$, with $\mathcal{W}^{\mathrm{s}}$ denoting the single scalar components of the corresponding vector-valued space $\boldsymbol{\mathcal{W}}^{\mathrm{s}}$. This work employs the dual mortar method, which specifies the Lagrange multipliers to be in the space $\boldsymbol{\mathcal{M}}$: the dual space of the trace space $\boldsymbol{\mathcal{W}}^{\mathrm{s}}$. In terms of functional analysis characterization, the dual space components are now $\mathcal{M} = H^{-1/2}(\gamma_{\mathrm{c}}^{\mathrm{s}})$ (see, e.g., B. I. Wohlmuth [2012]).

Based on these considerations, according to Hüeber (2008), the Lagrange multiplier vector is chosen from the convex cone $\mathcal{M}(\boldsymbol{\lambda}) \subset \mathcal{M}$ given by

$$\mathcal{M}(\boldsymbol{\lambda}) \equiv \left\{ \delta\boldsymbol{\lambda} \in \mathcal{M} \mid \langle \delta\boldsymbol{\lambda}, \boldsymbol{\vartheta} \rangle_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \leq \langle \mu\lambda^{\eta}, \|\boldsymbol{\vartheta}^{\tau}\| \rangle_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \,, \; \boldsymbol{\vartheta} \in \mathcal{W}^{\mathrm{s}} \text{ with } \vartheta^{\eta} \leq 0 \right\}, \tag{2.62}$$

where $\langle \bullet, \bullet \rangle_{\gamma_{\mathrm{c}}^{\mathrm{s}}}$ stands for the $H^{1/2}$-duality pairing between $\mathcal{W}^{\mathrm{s}}$ and $\mathcal{M}$ on $\gamma_{\mathrm{c}}^{\mathrm{s}}$,

$$\langle \boldsymbol{\lambda}, \delta\boldsymbol{u} \rangle_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \equiv \int_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \boldsymbol{\lambda}\, \delta\boldsymbol{u} \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \,, \tag{2.63}$$

and $\delta\boldsymbol{\lambda}$ represents a trial contact traction. The solution cone for the Lagrange multipliers space in Equation (2.62) simultaneously satisfies, in a weak sense, the KKT conditions in the normal direction and the Coulomb's tangential law of friction.

2.6.2 Contact constraints as variational inequalities

In order to derive weak statements of the contact constraints (2.51), it proves convenient first to express them in the alternative form of variational inequalities. Only the principal ideas are emphasized in the following, referring the reader to Hüeber (2008) for a comprehensive description and proof of equivalence between the variational inequalities and their continuum counterparts introduced in Chapter 2.

The KKT optimality conditions (2.48) in the normal direction can be equivalently rewritten as

$$\lambda^{\eta} \in \mathbb{R}_0^+ : \quad g\left(\delta\lambda^{\eta} - \lambda^{\eta}\right) \geq 0 \quad \forall\, \delta\lambda^{\eta} \in \mathbb{R}_0^+ \,, \tag{2.64}$$

where $\mathbb{R}_0^+$ represents the semi-positive real half-space. Likewise, the frictional contact conditions along the tangential direction, given in Equation (2.49), follow as

$$\boldsymbol{\lambda}^{\tau} \in \mathcal{B}\left(\mu\lambda^{\eta}\right) : \quad \boldsymbol{v}^{\tau} \cdot \left(\delta\boldsymbol{\lambda}^{\tau} - \boldsymbol{\lambda}^{\tau}\right) \leq 0 \quad \forall\, \delta\boldsymbol{\lambda}^{\eta} \in \mathcal{B}\left(\mu\lambda^{\eta}\right) . \tag{2.65}$$

Here, $\mathcal{B}\left(\mu\lambda^{\eta}\right)$ represents a $(d-1)$-dimensional sphere centred at $\mathbf{0}$ and of radius $\mu\lambda^{\eta}$—also commonly referred to as friction bound. Note that the condition (2.65) can be interpreted as a representation of the principle of dissipation inequality representing Coulomb's friction law, e.g., see Strömberg *et al.* (1996) or Strömberg (1997).

2.7 Summarized weak form

With the fundamental aspects of the variational continuum framework discussed, the final weak saddle point type formulation for mortar-based finite deformation frictional contact can be obtained by introducing the Lagrange multiplier in the general weak form (2.54) and by integrating the variational inequalities (2.64) and (2.65) over the non-mortar contact interface $\gamma_{\mathrm{c}}^{\mathrm{s}}$. The problem can be stated as:

Problem 2.3 (Weak form of IBVP of finite deformation frictional contact) *Given the internal forces and boundary conditions, find the kinematically admissible displacement function,* $\boldsymbol{u} \in \boldsymbol{\mathcal{U}}$, *and Lagrange multiplier,* $\boldsymbol{\lambda} \in \boldsymbol{\mathcal{M}}(\boldsymbol{\lambda})$, *such that, for all* $t \in [t_0, T]$, *the virtual work equations*

$$\delta\Pi_{\text{int}}\big(\boldsymbol{u}, \delta\boldsymbol{u}\big) - \delta\Pi_{\text{ext}}\big(\delta\boldsymbol{u}\big) + \int_{\gamma_{\text{c}}^{\text{s}}} \boldsymbol{\lambda} \cdot \big(\delta\boldsymbol{u}^{\text{s}} - \delta\hat{\boldsymbol{u}}^{\text{m}}\big)\, \mathrm{d}\gamma_{\text{c}}^{\text{s}} = 0\,, \quad \forall\, \delta\boldsymbol{u} \in \boldsymbol{\mathcal{V}}\,, \tag{2.66}$$

$$\langle\, g\,, \delta\lambda^{\eta} - \lambda^{\eta}\,\rangle_{\gamma_{\text{c}}^{\text{s}}} \geq 0\,, \quad \forall\, \delta\boldsymbol{\lambda} \in \boldsymbol{\mathcal{M}}(\boldsymbol{\lambda})\,, \tag{2.67}$$

$$\langle\, \boldsymbol{v}^{\tau}, \delta\boldsymbol{\lambda}^{\tau} - \boldsymbol{\lambda}^{\tau}\,\rangle_{\gamma_{\text{c}}^{\text{s}}} \leq 0\,, \quad \forall\, \delta\boldsymbol{\lambda} \in \boldsymbol{\mathcal{M}}(\boldsymbol{\lambda})\,, \tag{2.68}$$

are satisfied for any admissible test functions $\delta\boldsymbol{u} \in \boldsymbol{\mathcal{V}}$ *and* $\delta\boldsymbol{\lambda} \in \boldsymbol{\mathcal{M}}(\boldsymbol{\lambda})$.

Note that, since the underlying potentials of the weak form introduced in Section 2.5 are convex (thus representing positive definite operators with unique minima), the solution of Problem 2.3 constitutes a saddle point of the associated potential (2.66), i.e., the solution corresponds to a minimum with respect to the displacements $\boldsymbol{u}$ and a maximum concerning the Lagrange multipliers $\boldsymbol{\lambda}$, see e.g., Kikuchi and Oden (1988) and Luenberger and Ye (2008).

Chapter 3

Discrete Dual Mortar Contact Formulation and Global Solution Strategy

Building upon the continuous framework laid out in Chapter 2, this chapter focuses on the development of the discrete form of mortar-based frictional contact using the FEM. A brief overview of discrete Lagrange multiplier spaces is given (with particular emphasis on dual Lagrange multipliers), followed by a detailed description of the evaluation and numerical integration of discrete mortar finite element contributions. Then, the primal-dual approach is presented, with the reformulation of nodal inequality constraints with non-linear complementarity functions. The Newton-type algorithm for the solution of the spatially discrete frictional contact problem is introduced, describing in detail the subsequent linearisation of all the terms related to contact. Finally, the algebraic representation for the final linearised system of equations is explained in detail, emphasizing the elimination of the Lagrange multiplier.

3.1 Finite element approximation

Within this work, the weak form introduced in Problem 2.3 is discretised using the FEM exclusively. Nonetheless, since a detailed description of all the FEM concepts is beyond the scope of this work, only the basic notions and notation are emphasized. For a complete treatment of finite element methods, the reader is referred to the abundant literature regarding this topic, see, e.g., T. J. Hughes (2000), Zienkiewicz *et al.* (2014), Bathe (2006), Reddy (2014), and Belytschko *et al.* (2013).

In a nutshell, the concept of finite element discretisation consists of finding an approximate solution for the primary variable of the problem (i.e., the displacement field $\boldsymbol{u}$) at discrete points of interest—commonly referred to as nodes. The nodes are conveniently connected to establish elements, which altogether (as a *mesh*) allows formulating the following geometrical approximation[*]:

$$\Omega \approx \Omega^{\mathrm{h}} \equiv \bigcup_{\mathrm{e}=1}^{n^{\mathrm{e}}} \Omega_{\mathrm{e}}\,. \tag{3.1}$$

Essentially, the approximation is obtained by partitioning the domain $\Omega = \Omega^{\mathrm{s}} \cup \Omega^{\mathrm{m}}$ into n^{e} element subdomains $\Omega_{\mathrm{e}} \subset \Omega^{\mathrm{h}}$. Mathematically speaking, this introduces the *finite-*

---

[*] Throughout this document the superscript $(\bullet)^{\mathrm{h}}$ denotes a spatially discretised quantity.

dimensional subsets $\mathcal{U}^{\mathrm{h}} \subset \mathcal{U}$ and $\mathcal{V}^{\mathrm{h}} \subset \mathcal{V}$ as approximations for their corresponding functional sets. These approximations are typically expressed as an expansion in terms of *shape* (or *interpolation*) functions, which are usually low-order polynomials (e.g., Lagrange polynomials, associated with specific nodal positions in the mesh) meeting the differentiability requirements of the weak form. Following the *isoparametric approach*, both geometry and field variables are approximated using the same shape functions and, furthermore, the element subdomain Ω_{e} is typically mapped to the parameter (or reference) space:

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d) \tag{3.2}$$

With focus placed on finite element discretisation of frictional contact, only the associated physical quantities are considered in the following. Moreover, it is important to refer that, throughout this work, the actual mortar-based contact modelling is completely independent of the finite element formulation employed for the remaining terms (internal forces, external loads, etc.).

3.1.1 Finite element interpolation at the contact interface

As already mentioned in Chapter 2, the contact boundaries γ_{c}^{i} can be understood as $(d-1)$-dimensional manifolds embedded in the corresponding d-dimensional space. This means that contact interactions are evaluated on the contact subdomains γ_{c}^{i} and, consequently, a direct inheritance between the employed finite element discretisation and resulting contact interface elements is typically exploited. The following general form for geometry approximation on the discrete non-mortar and mortar boundaries holds:

$$\boldsymbol{x}^{\mathrm{s}} \approx \left\{\boldsymbol{x}^{\mathrm{s}}\right\}^{\mathrm{h}}\Big|_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \equiv \sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\mathbf{x}_k^{\mathrm{s}}\,; \tag{3.3a}$$

$$\boldsymbol{x}^{\mathrm{m}} \approx \left\{\boldsymbol{x}^{\mathrm{m}}\right\}^{\mathrm{h}}\Big|_{\{\gamma_{\mathrm{c}}^{\mathrm{m}}\}^{\mathrm{h}}} \equiv \sum_{l=1}^{n^{\mathrm{m}}} N_l^{\mathrm{m}}\left(\boldsymbol{\xi}^{\mathrm{m}}\right)\mathbf{x}_l^{\mathrm{m}}\,. \tag{3.3b}$$

Herein, n^{s} and n^{m} denote the total number of nodes on discrete non-mortar subdomain $\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}$ and mortar counterpart $\{\gamma_{\mathrm{c}}^{\mathrm{m}}\}^{\mathrm{h}}$, respectively, and the corresponding discrete nodal coordinates (in current configuration) are represented by $\mathbf{x}_k^{\mathrm{s}}$ and $\mathbf{x}_l^{\mathrm{m}}$. The shape functions N_k^{s} and N_l^{m} are defined with respect to the associated finite element parameter space $\boldsymbol{\xi}^{i}$, $i \in \{\mathrm{s},\mathrm{m}\}$, which for contact interface elements is defined similarly to the reference space introduced in Equation (3.2), with the exception of being a $(d-1)$-dimensional space, i.e., for two-dimensional problems it reduces to ξ^{i} and in three dimensions is denoted as $\boldsymbol{\xi}^{i} = \left(\xi_1^{i}, \xi_2^{i}\right)$. The displacement field interpolation is obtained from the nodal displacements $\mathbf{d}_k^{\mathrm{s}}$ and $\mathbf{d}_l^{\mathrm{m}}$ in a similar way

$$\boldsymbol{u}^{\mathrm{s}} \approx \left\{\boldsymbol{u}^{\mathrm{s}}\right\}^{\mathrm{h}}\Big|_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \equiv \sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\mathbf{d}_k^{\mathrm{s}}\,; \tag{3.4a}$$

$$\boldsymbol{u}^{\mathrm{m}} \approx \left\{\boldsymbol{u}^{\mathrm{m}}\right\}^{\mathrm{h}}\Big|_{\{\gamma_{\mathrm{c}}^{\mathrm{m}}\}^{\mathrm{h}}} \equiv \sum_{l=1}^{n^{\mathrm{m}}} N_l^{\mathrm{m}}\left(\boldsymbol{\xi}^{\mathrm{m}}\right)\mathbf{d}_l^{\mathrm{m}}\,. \tag{3.4b}$$

In addition to geometry and displacement fields, a suitable discretisation strategy for the Lagrange multipliers $\boldsymbol{\lambda}$ is required too. Within this work, the Lagrange multipliers interpolation is employed on the non-mortar side and its approximation is based on the introduction of the discrete Lagrange multiplier space $\mathcal{M}^{\mathrm{h}} \subset \mathcal{M}$. All details concerning the choice of this discrete space, its fundamental properties and implementation strategy follow in Section 3.2. At this stage, though, a generic notation can be introduced as

$$\boldsymbol{\lambda} \approx \boldsymbol{\lambda}^{\mathrm{h}} \equiv \sum_{j=1}^{n^{\lambda}} \Phi_j \left(\boldsymbol{\xi}^{\mathrm{s}}\right) \mathbf{z}_j \,, \tag{3.5}$$

with Φ_j representing the Lagrange multiplier interpolation, n^{λ} the total number of non-mortar nodes carrying additional Lagrange multiplier degrees of freedom and $\mathbf{z}_j$ the discrete nodal Lagrange multipliers. In mortar methods it is a common practice to consider that every non-mortar node serves as coupling node, thus $n^{\lambda} = n^{\mathrm{s}}$. However, for generality, this is not considered at this stage.

3.2 Discrete Lagrange multiplier spaces

This section focuses on the choice of the discrete Lagrange multiplier space $\mathcal{M}^{\mathrm{h}}$. Given its importance to the mathematical structure, robustness and numerical efficiency of the resulting mortar mixed formulation, this has been a topic of extensive research over the last two decades. In particular, fundamental aspects, such as inf-sup stability conditions and optimal *a priori* error estimates, have been carefully examined. As a result, a well-established mathematical framework has been gradually built up, thus creating appropriate theoretical foundations for the application of mortar methods to many different physical problems. With the focus of the present work not being on the mathematical analysis of mortar methods though, the reader is instead referred to B. I. Wohlmuth (2012) and references therein for an overview in the context of domain decomposition techniques, or the review article B. I. Wohlmuth (2011) and the contribution B. I. Wohlmuth, Popp, *et al.* (2012) in the context of frictional contact. Historically, two classes of discrete Lagrange multipliers—*standard* and *dual* Lagrange multipliers—are distinguished and described in the following.

3.2.1 Standard Lagrange multipliers

Standard Lagrange multipliers are the classical choice of discrete Lagrange multipliers within mortar methods. This approach consists of taking the interpolation functions from the finite-dimensional subset $\mathcal{W}^{\mathrm{h}} \subset \mathcal{W}$ (recall that $\mathcal{W}$ stands for the trace space of the virtual displacements $\mathcal{V}^{\mathrm{s}}$ on the non-mortar side), which leads to *identical* shape functions for Lagrange multipliers and non-mortar displacement interpolation, i.e.,

$$\Phi_j = N_j^{\mathrm{s}} \,. \tag{3.6}$$

One important property of this methodology is that it leads to mortar coupling conditions with global support. Mathematically, and explained in more detail later on, this property is displayed in the first mortar coupling becoming densely populated. Further details and illustrations on how to construct typical shape functions for finite element interpolation can be found, for instance, in T. J. Hughes (2000).

3.2.2 Dual Lagrange multipliers

The dual Lagrange multipliers for mortar methods have firstly been presented in B. I. Wohlmuth (2000). Exploiting the local dual basis functions, originally used in Scott and Zhang (1990) to define global projection-like operators, the dual mortar approach uses the same dual basis functions, to define the discrete Lagrange multipliers space. As emphasized in B. I. Wohlmuth (2000), this technique is motivated by the observation that, in the continuous setting, the Lagrange multipliers physically represent fluxes (tractions) on the contact interface. This duality argument is reflected by the use of the so-called *bi-orthogonality condition* with the displacements in $\mathcal{W}^{\mathrm{h}}$ to construct dual Lagrange multiplier shape functions. A typical notation for the bi-orthogonality condition is

$$\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j N_k^{\mathrm{s}} \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} = \delta_{jk} \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} N_k^{\mathrm{s}} \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \, . \tag{3.7}$$

Here, δ_{jk} represents the Kronecker delta, and this relation holds for every non-mortar node $j = \{1, \ldots, n^{\lambda}\}$ and $k = \{1, \ldots, n^{\mathrm{s}}\}$. The main advantage of this approach is that it heavily facilitates the treatment of mortar problems, without impinging upon the optimal convergence of the discretisation error. Firstly, the coupling conditions are now *localized*. With this, the first coupling mortar matrix $[\mathbf{D}]$ reduces to a diagonal matrix, thus allowing at later stages of the algorithm the application of efficient condensation procedures of the discrete Lagrange multiplier degrees of freedom—ultimately removing the troublesome saddle point structure of the underlying mixed formulation. Furthermore, in the context of contact problems, the local support character of dual mortar coupling is particularly attractive, as it allows to decouple point-wise the contact constraints, thus setting the basis for the application of efficient solution algorithms (Hüeber, 2008).

Remark. *During the analysis of the dual shape functions, it is crucial to consider the unique requirements for mortar-based finite deformation contact. These are essentially associated with the inequality nature of contact constraints and, thus, positivity requirements for the Lagrange multiplier basis functions. For the sake of simplicity, however, it is only mentioned as a remark at this stage, referring to Section 5.3 of Chapter 5 for a more in-depth overview. In doing so, it is not considered second-order interpolation in three-dimensions, as proper techniques to establish dual shape functions for contact problems successfully are required. This part of the work is focused on the building blocks of the dual mortar method instead and, therefore, this particular case is only analysed later in Chapter 5. Notwithstanding, all the remaining concepts in this chapter remain thoroughly valid for every possible combination of interpolation degree and finite element geometry. In the following, a brief overview of the construction of the actual dual Lagrange multiplier shape functions is given, mentioning the issue of consistency of dual Lagrange multipliers in partially integrated non-mortar elements.*

3.2.2.1 Construction of dual Lagrange multiplier space

Within this work, dual shape functions are constructed by employing the strategy proposed by Flemisch and B. I. Wohlmuth (2007). The theoretical details are carefully described in the original publication, including a set of mathematical proofs and *a priori* results. Nonetheless, here only the main ideas needed for the actual implementation of the method are emphasized. First, note that the bi-orthogonality condition in Equation (3.7) can be equally evaluated *element-wise* on each non-mortar element $\{\gamma_\mathrm{c}^\mathrm{s}\}_\mathrm{e}^\mathrm{h}$, viz.

$$\int_{\{\gamma_\mathrm{c}^\mathrm{s}\}_\mathrm{e}^\mathrm{h}} \Phi_j N_k^\mathrm{s} \, \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} = \delta_{jk} \int_{\{\gamma_\mathrm{c}^\mathrm{s}\}_\mathrm{e}^\mathrm{h}} N_k^\mathrm{s} \, \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} \,, \tag{3.8}$$

which hold for every non-mortar element node $j = \{1, \ldots, n_\mathrm{e}^\lambda\}$ and $k = \{1, \ldots, n_\mathrm{e}^\mathrm{s}\}$ within the non-mortar element. Herein, n_e^s represents the number of nodes in a given non-mortar element e and n_e^λ the corresponding number of non-mortar element nodes carrying Lagrange multipliers. Then, the dual shape function Φ_j is defined by introducing unknown linear coefficients a_{jk}^Φ such that

$$\Phi_j = \sum_{k=1}^{n_\mathrm{e}^\mathrm{s}} \mathrm{a}_{jk}^\Phi N_k^\mathrm{s} \,, \tag{3.9}$$

that is, a linear combination of standard shape functions N_k^s. Therefore, by introducing the vector $\{\mathbf{\Phi}_\mathrm{e}\} \in \mathbb{R}^{n_\mathrm{e}^\lambda}$ conveniently gathering all dual shape functions defined at the non-mortar element e,

$$\{\mathbf{\Phi}_\mathrm{e}\} \equiv \begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_{n_\mathrm{e}^\lambda} \end{bmatrix} , \tag{3.10}$$

and the vector $\{\mathbf{n}_\mathrm{e}\} \in \mathbb{R}^{n_\mathrm{e}^\mathrm{s}}$ containing the associated standard shape functions,

$$\{\mathbf{n}_\mathrm{e}\} \equiv \begin{bmatrix} N_1 \\ \vdots \\ N_{n_\mathrm{e}^\mathrm{s}} \end{bmatrix} , \tag{3.11}$$

one can determine the dual shape functions from the transformation

$$\{\mathbf{\Phi}_\mathrm{e}\} = [\mathbf{A}_\mathrm{e}^\Phi]\{\mathbf{n}_\mathrm{e}\} \,. \tag{3.12}$$

The coefficients matrix $[\mathbf{A}_\mathrm{e}^\Phi] = [\mathrm{a}_{jk}^\Phi] \in \mathbb{R}^{n_\mathrm{e}^\lambda \times n_\mathrm{e}^\mathrm{s}}$ is computed by inserting Equation (3.9) into Equation (3.8), which yields the local mass matrix system to be solved on each element:

$$[\mathbf{A}_\mathrm{e}^\Phi] = [\mathbf{D}_\mathrm{e}^\Phi][\mathbf{M}_\mathrm{e}^\Phi]^{-1} \,. \tag{3.13}$$

Individual entries for the matrices $[\mathbf{D}_\mathrm{e}^\Phi] = [\mathrm{d}_{jk}^\Phi] \in \mathbb{R}^{n_\mathrm{e}^\lambda \times n_\mathrm{e}^\mathrm{s}}$ and $[\mathbf{M}_\mathrm{e}^\Phi] = [\mathrm{m}_{jk}^\Phi] \in \mathbb{R}^{n_\mathrm{e}^\mathrm{s} \times n_\mathrm{e}^\mathrm{s}}$ are defined as

$$\mathrm{d}_{jk}^\Phi \equiv \delta_{jk} \int_{\{\gamma_\mathrm{c}^\mathrm{s}\}_\mathrm{e}^\mathrm{h}} N_k^\mathrm{s}(\boldsymbol{\xi}^\mathrm{s}) \, \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} \approx \delta_{jk} \sum_{\mathrm{gp}=1}^{n_\mathrm{gp}} w_\mathrm{gp} N_k^\mathrm{s}(\boldsymbol{\xi}_\mathrm{gp}^\mathrm{s}) J_\mathrm{e}^\mathrm{s} \,, \tag{3.14}$$

$$\mathrm{m}_{jk}^\Phi \equiv \int_{\{\gamma_\mathrm{c}^\mathrm{s}\}_\mathrm{e}^\mathrm{h}} N_j^\mathrm{s}(\boldsymbol{\xi}^\mathrm{s}) N_k^\mathrm{s}(\boldsymbol{\xi}^\mathrm{s}) \, \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} \approx \sum_{\mathrm{gp}=1}^{n_\mathrm{gp}} w_\mathrm{gp} N_j^\mathrm{s}(\boldsymbol{\xi}_\mathrm{gp}^\mathrm{s}) N_k^\mathrm{s}(\boldsymbol{\xi}_\mathrm{gp}^\mathrm{s}) J_\mathrm{e}^\mathrm{s} \,, \tag{3.15}$$

where J_e^s stands for the non-mortar element Jacobian determinant and n_gp the number of Gauss points for numerical integration over the non-mortar element. Lastly, the dual basis functions are constructed by a node-wise gluing process, just as in standard finite element methods.

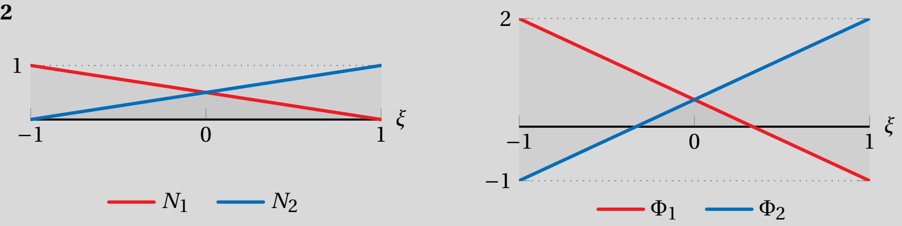
It is essential to point out that the bi-orthogonality condition, both in Equation (3.7) or in its element-wise version in Equation (3.8), must be satisfied in the *current* physical space. Consequently, dual shape functions not only depend on the actual distortion of the individual finite element—which means that cannot be defined *a priori* for non-constant non-mortar element Jacobian determinant—but also need to be constantly re-evaluated. Illustrative examples of dual shape functions for first-order *undistorted* finite elements are represented in Figure 3.1, along with their standard counterparts. For further details on the construction of the dual basis, the interested reader is referred to Flemisch and B. I. Wohlmuth (2007) and Popp, B. I. Wohlmuth, *et al.* (2012). Worth mentioning as well Lamichhane *et al.* (2005), for a comprehensive mathematical analysis of higher-order dual mortar methods, and also Hartmann, Brunssen, *et al.* (2007) for a clear illustration and numerical results.

Besides its intricate construction process, it should also be noticed that dual shape functions themselves are deformation-dependent too. Therefore, consistent linearisation is required for the application of Newton-Raphson methods, which turns out to represent one of the main challenges for dual mortar contact implementation. As suggested by Popp, Gee, *et al.* (2009) and Popp, Gitterle, *et al.* (2010), the numerical procedure can be based on linearising the bi-orthogonality condition, which is provided in Appendix A with considerable detail, together with a suitable algebraic representation of the associated directional derivatives.

3.2.2.2 Consistency of dual Lagrange multipliers at boundaries

As a final topic to be addressed in this section, attention is shifted towards the algorithmic aspect of a consistent definition of bi-orthogonality and dual Lagrange multiplier shape functions in partially integrated non-mortar elements. In the foregoing discussion, it is assumed that all non-mortar elements on the contact boundary $\gamma_{\mathrm{c}}^{\mathrm{s}}$ are *completely* integrated. However, in finite deformation contact, there most likely exist scenarios where non-mortar elements do not fully participate in the integration area, e.g., the so-called dropping edge problems (when one body slides off another at an edge). This may lead to consistency issues, as different integration domains are employed. While the bi-orthogonality condition in Equation (3.7) is still defined on each *entire* non-mortar element, element contributions to mortar integrals are calculated on the sub-domains where, as explained in more detail in Section 3.7, the projection onto the mortar boundary is still feasible. As a consequence, the dual shape functions obtained by the strategy described lose the sought-after localized character. As originally proposed by Cichosz and Bischoff (2011) for two-dimensional problems, and later on extended to three dimensions in Popp, Seitz, *et al.* (2013), a solution to make the bi-orthogonality condition consistent in any event is to replace the integration domain in Equation (3.8) by the actual overlapping region of the non-mortar element. It is important to mention that, by employing this strategy, exceptional attention to some aspects such as integral positivity and an ill-conditioned system of equations is needed. For further details on appropriate algorithmic treatment to these issues, refer to Popp, Seitz, *et al.* (2013).

(a) line2



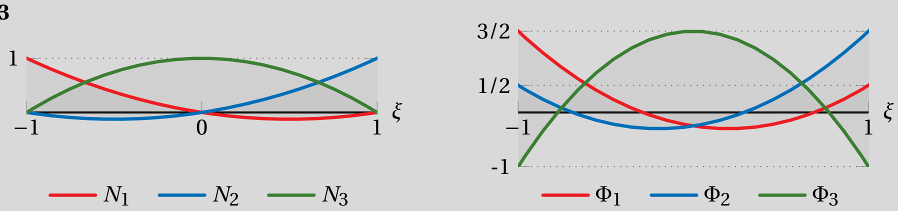
$N_1(\xi) = (1-\xi)/2$

$N_2(\xi) = (1+\xi)/2$

$\Phi_1(\xi) = (1-3\xi)/2$

$\Phi_2(\xi) = (1+3\xi)/2$

(b) line3



$N_1(\xi) = \xi(\xi-1)/2$

$N_2(\xi) = (1-\xi)(1+\xi)$

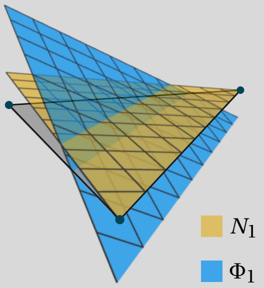
$N_3(\xi) = \xi(\xi+1)/2$

$\Phi_1(\xi) = N_1(\xi) - (3/4)N_3(\xi) + (1/2)$

$\Phi_2(\xi) = (5/2)N_3(\xi) - 1$

$\Phi_3(\xi) = N_2(\xi) - (3/4)N_3(\xi) + (1/2)$

(c) tri3



$N_1(\boldsymbol{\xi}) = 1-\xi_1-\xi_2$

$N_2(\boldsymbol{\xi}) = \xi_1$

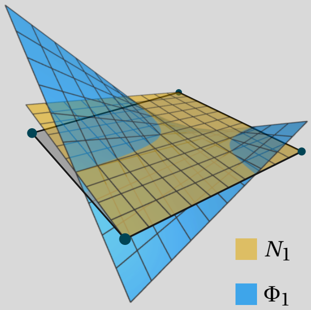
$N_3(\boldsymbol{\xi}) = \xi_2$

$\Phi_1(\boldsymbol{\xi}) = 3-4\xi_1-4\xi_2$

$\Phi_2(\boldsymbol{\xi}) = 4\xi_1-1$

$\Phi_3(\boldsymbol{\xi}) = 4\xi_2-1$

(d) quad4



$N_1(\boldsymbol{\xi}) = 1/4\,(1-\xi_1)(1-\xi_2)$

$N_2(\boldsymbol{\xi}) = 1/4\,(1+\xi_1)(1-\xi_2)$

$N_3(\boldsymbol{\xi}) = 1/4\,(1+\xi_1)(1+\xi_2)$

$N_4(\boldsymbol{\xi}) = 1/4\,(1-\xi_1)(1+\xi_2)$

$\Phi_1(\boldsymbol{\xi}) = 4N_1(\boldsymbol{\xi}) - 2N_2(\boldsymbol{\xi}) + N_3(\boldsymbol{\xi}) - 2N_4(\boldsymbol{\xi})$

$\Phi_2(\boldsymbol{\xi}) = -2N_1(\boldsymbol{\xi}) + 4N_2(\boldsymbol{\xi}) - 2N_3(\boldsymbol{\xi}) + N_4(\boldsymbol{\xi})$

$\Phi_3(\boldsymbol{\xi}) = N_1(\boldsymbol{\xi}) - 2N_2(\boldsymbol{\xi}) + 4N_3(\boldsymbol{\xi}) - 2N_4(\boldsymbol{\xi})$

$\Phi_4(\boldsymbol{\xi}) = -2N_1(\boldsymbol{\xi}) + N_2(\boldsymbol{\xi}) - 2N_3(\boldsymbol{\xi}) + 4N_4(\boldsymbol{\xi})$

Figure 3.1: Dual shape functions for undistorted first-order finite elements.

3.3 Discrete contact virtual work

Following the finite element scheme introduced in the previous section, the discretised version of the virtual contact work given in Equation (2.61) yields

$$\delta\Pi_{\mathrm{c}} \approx \delta\Pi_{\mathrm{c}}^{\mathrm{h}} \equiv \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \boldsymbol{\lambda}^{\mathrm{h}} \cdot \left[\{\delta \boldsymbol{u}^{\mathrm{s}}\}^{\mathrm{h}} - \{\delta \hat{\boldsymbol{u}}^{\mathrm{m}}\}^{\mathrm{h}} \right] \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \tag{3.16}$$

which, recalling the interpolations in Equations (3.4) and (3.5), can be rearranged as

$$\begin{aligned} \Pi_{\mathrm{c}}^{\mathrm{h}} = & \sum_{j=1}^{n^{\lambda}} \sum_{k=1}^{n^{\mathrm{s}}} \left\{ \mathbf{z}_j \left[\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j(\boldsymbol{\xi}^{\mathrm{s}}) N_k^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \right] \delta \mathbf{d}_k^{\mathrm{s}} \right\} \\ & - \sum_{j=1}^{n^{\lambda}} \sum_{l=1}^{n^{\mathrm{m}}} \left\{ \mathbf{z}_j \left[\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j(\boldsymbol{\xi}^{\mathrm{s}}) N_l^{\mathrm{m}}(\hat{\boldsymbol{\xi}}^{\mathrm{m}}) \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \right] \delta \mathbf{d}_l^{\mathrm{m}} \right\} . \end{aligned} \tag{3.17}$$

This expression motivates the introduction of two fundamental entities of mortar methods: the *first mortar coupling matrix*, herein denoted by $[\mathbf{D}] \in \mathbb{R}^{(d \cdot n^{\lambda}) \times (d \cdot n^{\mathrm{s}})}$, and the *second mortar coupling matrix*, represented by $[\mathbf{M}] \in \mathbb{R}^{(d \cdot n^{\lambda}) \times (d \cdot n^{\mathrm{m}})}$. Their nodal blocks $\mathbf{D}_{[j,k]}$ and $\mathbf{M}_{[j,l]}$ are readily identified in Equation (3.17) as

$$\mathbf{D}_{[j,k]} \equiv \mathrm{D}_{jk} \, \mathbf{I}_d \,, \quad j = 1, \ldots, n^{\lambda}, \; k = 1, \ldots, n^{\mathrm{s}} \,, \tag{3.18a}$$

$$\mathbf{M}_{[j,l]} \equiv \mathrm{M}_{jl} \, \mathbf{I}_d \,, \quad j = 1, \ldots, n^{\lambda}, \; l = 1, \ldots, n^{\mathrm{m}} \,, \tag{3.18b}$$

with $\mathbf{I}_d$ representing the d-dimensional identity matrix, together with the abbreviations:

$$\mathrm{D}_{jk} \equiv \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j(\boldsymbol{\xi}^{\mathrm{s}}) N_k^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \,; \tag{3.19a}$$

$$\mathrm{M}_{jl} \equiv \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j(\boldsymbol{\xi}^{\mathrm{s}}) N_l^{\mathrm{m}}(\hat{\boldsymbol{\xi}}^{\mathrm{m}}) \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \,. \tag{3.19b}$$

As mentioned, due to the bi-orthogonality condition (3.7), the first mortar coupling matrix $[\mathbf{D}]$ becomes diagonal and can be evaluated as

$$\mathrm{D}_{jk} = \delta_{jk} \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} N_j(\boldsymbol{\xi}^{\mathrm{s}}) \, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \,. \tag{3.20}$$

It is important to note that, while the first mortar coupling matrix $[\mathbf{D}]$ is exclusively concerned with terms defined on the non-mortar side (recall that the Lagrange multipliers are interpolated on the non-mortar boundary), the second mortar coupling matrix $[\mathbf{M}]$ is more intricate, since it involves terms related to both sides. This becomes clear by having in mind the projection operation underlying the definition of the point $\{\delta \hat{\boldsymbol{u}}^{\mathrm{m}}\}^{\mathrm{h}}$, typically represented by the notation

$$\{\delta \hat{\boldsymbol{u}}^{\mathrm{m}}\}^{\mathrm{h}} \equiv \{\delta \boldsymbol{u}^{\mathrm{m}}\}^{\mathrm{h}} \circ R^{\mathrm{h}} \,, \tag{3.21}$$

with R^{h} denoting the discrete version of the contact mapping operator introduced in Equation (2.23). This reveals to be one of the challenges of mortar-based contact algorithms, mainly because the discrete mapping changes due to relative movement between contact subdomains. Consequently, mortar coupling terms must be continuously re-evaluated and, at later stages, with regard to nonlinear solution schemes of Newton-Raphson type, this requires the consistent linearisation of all the above deformation-dependent quantities as well. Details on the actual numerical integration of the aforementioned mortar coupling terms follow in Section 3.7.

3.4 Discrete contact constraints

In addition to the discrete contact virtual work discussed in Section 3.3, the discrete version of contact constraints given in Equations (2.67) and (2.68) must be derived as well. Following the same line of thought as Chapter 2, the next paragraphs treat contact constraints separately in normal and tangential directions.

3.4.1 Discrete contact constraints in the normal direction

The discrete counterpart of the contact constraints in the normal direction, Equation (2.67), follows as

$$\int_{\gamma_{\mathrm{c}}^{\mathrm{s}}} g\left(\delta\lambda^{\eta} - \lambda^{\eta}\right) \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \approx \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} g^{\mathrm{h}}\left(\left\{\delta\lambda^{\eta}\right\}^{\mathrm{h}} - \left\{\lambda^{\eta}\right\}^{\mathrm{h}}\right) \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \geq 0\,, \tag{3.22}$$

with $g^{\mathrm{h}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)$ representing the discretised version of the gap function introduced in Equation (2.25), i.e.,

$$g^{\mathrm{h}}(\boldsymbol{\xi}^{\mathrm{s}}) \equiv -\boldsymbol{\eta}^{\mathrm{h}}(\boldsymbol{\xi}^{\mathrm{s}}) \cdot \left[\sum_{j=1}^{n^{\mathrm{s}}} N_j^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\mathbf{x}_j^{\mathrm{s}} - \sum_{l=1}^{n^{\mathrm{m}}} N_l^{\mathrm{m}}(\hat{\boldsymbol{\xi}}^{\mathrm{m}})\mathbf{x}_l^{\mathrm{m}}\right]. \tag{3.23}$$

At this stage, an abstract compact notation for the interpolated unit normal vector $\boldsymbol{\eta}^{\mathrm{h}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)$ is adopted, referring to Section 3.7 for additional details on the employed discretisation strategy. The normal component of the Lagrange multiplier $\left\{\lambda^{\eta}\right\}^{\mathrm{h}}$ is, as the contact traction vector, obtained by projecting the discrete Lagrange multiplier vector $\boldsymbol{\lambda}^{\mathrm{h}}$ along the normal direction $\boldsymbol{\eta}^{\mathrm{h}}$, i.e.,

$$\left\{\lambda^{\eta}\right\}^{\mathrm{h}} \equiv \boldsymbol{\lambda}^{\mathrm{h}} \cdot \boldsymbol{\eta}^{\mathrm{h}}\,; \tag{3.24}$$

the same applies to the discrete test functions $\left\{\delta\lambda^{\eta}\right\}^{\mathrm{h}}$. Then, by considering the finite element interpolation (3.5) for the Lagrange multiplier, it follows that Equation (3.22) can be rewritten as

$$\sum_{j=1}^{n^{\lambda}} \left(\delta \mathrm{z}_j^{\eta} - \mathrm{z}_j^{\eta}\right) \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j\left(\boldsymbol{\xi}^{\mathrm{s}}\right) g^{\mathrm{h}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right) \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,. \tag{3.25}$$

It is important to mention that in Equation (3.25), the tangential plane, and consequently the interpolated normal vector $\boldsymbol{\eta}^{\mathrm{h}}$, is assumed to be constant within the local support of node j. Finally, by choosing carefully in a targeted oriented way the discrete Lagrange multiplier component z_j^{η} and test function $\delta\mathrm{z}_j^{\eta}$, i.e., by assuming

$$\delta \mathrm{z}_j^{\eta} \geq 0 \tag{3.26}$$

for a given non-mortar node j and

$$\mathrm{z}_k^{\eta} \geq 0 \tag{3.27}$$

for the remaining non-mortar nodes k—thus defining a subset of the trace space $\mathcal{W}^{\mathrm{h}}$—the use of constrained dual basis enables to *decouple* the weak form of the contact constraints in the normal direction in Equation (2.67), leading to the following set of *point-wise* conditions:

$$\tilde{g}_j \geq 0\,, \tag{3.28a}$$
$$\mathrm{z}_j^{\eta} \geq 0\,, \tag{3.28b}$$
$$\tilde{g}_j\, \mathrm{z}_j^{\eta} = 0\,, \tag{3.28c}$$

holding for all non-mortar discrete nodes $j = 1, \dots, n^{n^{\lambda}}$. Herein, the discrete weighted gap $\tilde{g}_j$ at a given non-mortar node j is defined as

$$\tilde{g}_j \equiv \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j\left(\boldsymbol{\xi}^{\mathrm{s}}\right) g^{\mathrm{h}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,. \tag{3.29}$$

It should be pointed out that the transition from Equation (3.25) to the point-wise conditions (3.28) is only possible for dual Lagrange multiplier interpolation; see Hüeber (2008) for a thorough explanation, including a mathematical proof for equivalence between Equation (3.28) and Equation (2.67). Examining the contact constraints (3.28) more carefully, a striking analogy with the set of KKT optimality conditions given in Equation (2.48) becomes apparent. Essentially, the decoupled point-wise conditions (3.28) can be interpreted as a discrete formulation of the continuum KKT conditions. Furthermore, the former must be enforced independently at discrete non-mortar nodal points, as in classical NTS discretisation schemes. This similarity, however, proves only to be superficial, since the conditions (3.28) for mortar methods carry substantially more information.

3.4.2 Discrete contact constraints in the tangential direction

The discrete version of the tangential contact constraints introduced in Equation (2.68) follows as

$$\int_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \boldsymbol{v}^{\tau} \cdot \left(\delta\boldsymbol{\lambda}^{\tau} - \boldsymbol{\lambda}^{\tau}\right) \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \approx \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \{\boldsymbol{v}^{\tau}\}^{\mathrm{h}} \cdot \left(\{\delta\boldsymbol{\lambda}^{\tau}\}^{\mathrm{h}} - \{\boldsymbol{\lambda}^{\tau}\}^{\mathrm{h}}\right) \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \leq 0\,, \tag{3.30}$$

with the term $\{\boldsymbol{\lambda}^{\tau}\}^{\mathrm{h}}$ representing the vector containing the tangential components of the discrete Lagrange multiplier vector $\boldsymbol{\lambda}^{\mathrm{h}}$; analogously, $\{\delta\boldsymbol{\lambda}^{\tau}\}^{\mathrm{h}}$ denotes the tangential projection of the discrete test functions. The discrete relative tangential velocity $\{\boldsymbol{v}^{\tau}\}^{\mathrm{h}}$ is determined such that it satisfies the requirement of frame indifference and, given its importance, is discussed in more detail in the following paragraphs. Nonetheless, one can already further rearrange Equation (3.30) by considering the finite element interpolation of Lagrange multipliers in Equation (3.5), viz.

$$\sum_{j=1}^{n^{\lambda}} \left(\delta\mathbf{z}_j^{\tau} - \mathbf{z}_j^{\tau}\right) \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j\left(\boldsymbol{\xi}^{\mathrm{s}}\right) \cdot \{\boldsymbol{v}^{\tau}\}^{\mathrm{h}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,; \tag{3.31}$$

note that, once again, the tangent plane is assumed constant within the local support of node j. Then, by defining the weighted tangential relative velocity at a node j as

$$\tilde{\boldsymbol{v}}_j^{\tau} \equiv \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j\left(\boldsymbol{\xi}^{\mathrm{s}}\right) \cdot \{\boldsymbol{v}^{\tau}\}^{\mathrm{h}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\, \mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,, \tag{3.32}$$

and following a similar approach to the discrete normal constraints in the previous section, that is, by carefully choosing the discrete test function $\delta\mathbf{z}_j^\tau$ at node j to be

$$\|\delta\mathbf{z}_j^\tau\| \le \mu\,|\mathrm{z}_j^\eta| \tag{3.33}$$

and

$$\|\mathbf{z}_k^\tau\| \le \mu\,|\mathrm{z}_k^\eta| \tag{3.34}$$

for the remaining nodes k, the discrete test functions are in the admissible space $\delta\boldsymbol{\lambda}^\mathrm{h} \in \boldsymbol{\mathcal{M}}^\mathrm{h}(\boldsymbol{\lambda}^\mathrm{h})$. By means of the constrained dual basis, the tangential contact constraints yield the following set of *point-wise* conditions:

$$\begin{aligned} \psi_j \equiv \|\mathbf{z}_j^\tau\| - \mu\,|\mathrm{z}_j^\eta| &\le 0\,, && (3.35a)\\ \tilde{\boldsymbol{v}}_j^\tau + \beta_j\,\mathbf{z}_j^\tau &= \mathbf{0}\,, && (3.35b)\\ \beta_j &\ge 0\,, && (3.35c)\\ \psi_j\beta_j &= 0\,, && (3.35d) \end{aligned}$$

which must hold for all non-mortar nodes $j = 1,\ldots,n^\mathrm{s}$. Similar to the normal constraints, the frictional conditions (3.35) can be interpreted as a weak statement of the original Coulomb's friction law in Equation (2.49), with an additional weighting based on the Lagrange multiplier dual shape functions Φ_j. The integral character of the weak formulation is embedded in the mortar projected discrete geometrical measures, namely the weighted tangential relative velocity $\tilde{\boldsymbol{v}}_j^\tau$ (and weighted gap function $\tilde{g}_j$, in the normal constraints). Again, for a comprehensive mathematical treatment of the aforementioned deviations, the interested reader is referred to Hüeber (2008).

3.4.2.1 Frame indifference of relative tangential velocity

An aspect in the formulation of large deformation frictional contact widely discussed in the literature is that of *frame indifference* of the rate measures necessary for the expression of the evolution laws associated with friction. The concept of frame indifference, also commonly referred to as *objectivity*, evaluates if a given physical quantity is properly invariant, i.e., unaffected by any rigid body motion the material may undergo during the deformation process. Mathematically speaking, such evaluation can be performed by defining an alternative reference frame composed by a proper rotation and translation, with respect to the coordinate system in which the problem is formulated. A given quantity is frame indifferent if and only if it remains unaffected by such transformation.

Within the ongoing formulation, frame indifference of the frictional problem in Equation (3.35) must be guaranteed. Therefore, the definition of a frame indifferent mortar-projected tangential relative velocity $\{\boldsymbol{v}^\tau\}^\mathrm{h}$ is required. However, the relative tangential velocity definition introduced in Section 2.2.2 is not frame indifferent, even in the continuous setting, see e.g., Laursen (2013). Therefore, it is no surprise that its discrete version and, consequently, the weighted tangential relative velocity in Equation (3.32) are not frame indifferent either. In order to overcome this problem, and assure that the frictional constitutive relation (3.35) is objective, in Puso and Laursen (2004b), a solution based on adding in a dilatant portion to the tangential velocity measure is proposed. This idea has been broadly exploited and adopted, e.g., in B. Yang, Laursen, and Meng (2005), Puso, Laursen, and Solberg (2008), B. Yang and Laursen (2009), Popp, Gitterle, *et al.* (2010), and Gitterle *et al.* (2010), and is briefly described here within the context of mortar-based contact. The starting point is the definition of the weighted gap vector $\tilde{\boldsymbol{g}}_j$ at a node j as

$$\tilde{\boldsymbol{g}}_j \equiv \int_{\{\gamma_c^s\}^h} \Phi_j\left(\boldsymbol{\xi}^s\right) \boldsymbol{g}^h\left(\boldsymbol{\xi}^s\right) \, d\gamma_c^s \,, \tag{3.36}$$

which, taking into account the structure of the discrete version of the gap vector,

$$\boldsymbol{g}^h(\boldsymbol{\xi}^s) \equiv \sum_{k=1}^{n^s} N_k^s(\boldsymbol{\xi}^s)\mathbf{x}_k^s - \sum_{l=1}^{n^m} N_l^m(\hat{\boldsymbol{\xi}}^m)\mathbf{x}_l^m \,, \tag{3.37}$$

can be further rearranged as

$$\tilde{\boldsymbol{g}}_j = \sum_{k=1}^{n^s} \mathbf{D}_{[j,k]} \, \mathbf{x}_k^s - \sum_{l=1}^{n^m} \mathbf{M}_{[j,l]} \, \mathbf{x}_l^m \,. \tag{3.38}$$

Thus, similarly to the approach presented in Section 2.2.2, the contacting points are assumed to remain coincident in space. This implies that the total time derivative of Equation (3.38) remains zero, i.e.,

$$\dot{\tilde{\boldsymbol{g}}}_j = \left(\sum_{k=1}^{n^s} \mathbf{D}_{[j,k]} \, \dot{\mathbf{x}}_k^s - \sum_{l=1}^{n^m} \mathbf{M}_{[j,l]} \, \dot{\mathbf{x}}_l^m \right) + \left(\sum_{k=1}^{n^s} \dot{\mathbf{D}}_{[j,k]} \, \mathbf{x}_k^s - \sum_{l=1}^{n^m} \dot{\mathbf{M}}_{[j,l]} \, \mathbf{x}_l^m \right) \approx \mathbf{0} \,. \tag{3.39}$$

The first term in brackets leads precisely to the non-objective measure for the tangential relative velocity previously introduced in Equation (3.32). The second term in brackets, however, is frame indifferent under all conditions of contact, as the time derivatives of the mortar coupling matrices nodal blocks $\dot{\mathbf{D}}_{[j,k]}$ and $\dot{\mathbf{M}}_{[j,l]}$ remain zero during rigid body motions. This means that by changing its sign, it can be used as a measure for the tangential relative velocity. Although, it is important to note that the normal contact constraints only enforce the normal projection of the weighted gap $\tilde{\boldsymbol{g}}^h$ to be zero (not the vector itself). So, in order to remove any potential normal component, Equation (3.39) is projected in the nodal tangential plane. The components of the nodal tangential relative velocity follow as

$$\tilde{\boldsymbol{v}}_j^\tau \equiv -\boldsymbol{\mathcal{P}}_j \left(\sum_{k=1}^{n^s} \dot{\mathbf{D}}_{[j,k]} \, \mathbf{x}_k^s - \sum_{l=1}^{n^m} \dot{\mathbf{M}}_{[j,l]} \, \mathbf{x}_l^m \right) . \tag{3.40}$$

Here, the tangential projection matrix $\left[\boldsymbol{\mathcal{P}}_j\right] \in \mathbb{R}^{(d-1)\times d}$ has been introduced, and is defined as the assembly of interpolated unit tangent vectors $\{\boldsymbol{\tau}_j^\alpha\}^h$ at the non-mortar node j, i.e.,

$$\boldsymbol{\mathcal{P}}_j \equiv \begin{bmatrix} \{\boldsymbol{\tau}_j^1\}^h \\ \vdots \\ \{\boldsymbol{\tau}_j^{d-1}\}^h \end{bmatrix} . \tag{3.41}$$

While this expression is general for mortar-based contact methods, further simplification can still be employed for dual Lagrange multipliers. Owing to the bi-orthogonality condition (3.7), the first mortar coupling matrix $[\mathbf{D}]$ has local support, which allows rewriting

Equation (3.40) as

$$\tilde{\boldsymbol{v}}_j^{\tau} \equiv -\boldsymbol{\mathcal{P}}_j \left(\dot{\mathbf{D}}_{[j,j]} \mathbf{x}_j^{\mathrm{s}} - \sum_{l=1}^{n^{\mathrm{m}}} \dot{\mathbf{M}}_{[j,l]} \mathbf{x}_l^{\mathrm{m}} \right) , \tag{3.42}$$

At this stage, a time discretisation scheme for the mortar coupling terms in Equation (3.42) is required, which is described in the following for the adopted method.

3.4.2.2 Time discretisation

Although this work focuses on quasi-static problems, an appropriate time discretisation scheme for the relative tangential velocity discussed above is still required. Thus, by choosing a *backward Euler scheme*, the so-called weighted tangential relative slip increment $\tilde{\boldsymbol{u}}_j^{\tau}$ at a non-mortar node j is introduced as

$$\tilde{\boldsymbol{u}}_j^{\tau} \equiv \tilde{\boldsymbol{v}}_j^{\tau} \cdot \Delta t , \tag{3.43}$$

which results in

$$\tilde{\boldsymbol{u}}_j^{\tau} \equiv -\boldsymbol{\mathcal{P}}_j \cdot \left[\left(\mathbf{D}_{[j,j]}^{t_n} - \mathbf{D}_{[j,j]}^{t_{n-1}} \right) \mathbf{x}_j^{\mathrm{s}} - \sum_{l=1}^{n^{\mathrm{m}}} \left(\mathbf{M}_{[j,l]}^{t_n} - \mathbf{M}_{[j,l]}^{t_{n-1}} \right) \mathbf{x}_l^{\mathrm{m}} \right] . \tag{3.44}$$

Since the underlying integration method is implicit, it can be shown to be unconditionally stable, thus providing a robust algorithm even compatible with large time step sizes. By defining $\tilde{\beta}_j = \beta_j \cdot \Delta t$, the nodal tangential contact conditions at each non-mortar node j can be re-written as

$$\psi_j \equiv \|\mathbf{z}_j^{\tau}\| - \mu |\mathrm{z}_j^{\eta}| \leq 0 , \tag{3.45a}$$

$$\tilde{\boldsymbol{u}}_j^{\tau} + \tilde{\beta}_j \mathbf{z}_j^{\tau} = \mathbf{0} , \tag{3.45b}$$

$$\tilde{\beta}_j \geq 0 , \tag{3.45c}$$

$$\psi_j \tilde{\beta}_j = 0 . \tag{3.45d}$$

3.5 Averaged orthonormal moving frame

With the discrete version of the virtual contact work and contact constraints at hand, attention is now shifted towards their computational treatment. One fundamental ingredient of contact methods regards the local orthonormal frame attached to each contact node carrying a Lagrange multiplier. Its definition has a profound impact on the enforcement of contact constraints, as it establishes the basis for the split of surface contact tractions into normal and tangential components. Moreover, as explained in more detail in Section 3.7, within the numerical evaluation of mortar integrals, the discrete projection of Gauss points between sides of the contact interface also requires a continuous field of normals. The idea originally proposed by B. Yang, Laursen, and Meng (2005) defines, at a given non-mortar node j, the *averaged nodal unit normal* as[†]

$$\tilde{\boldsymbol{\eta}}_j\left(\xi_j^{\mathrm{s}}\right) \equiv \frac{\sum_{\mathrm{e}=1}^{n_{\mathrm{adj}}} \boldsymbol{\eta}_{\mathrm{e}}^{\mathrm{a}}\left(\xi_j^{\mathrm{s}}\right)}{\|\sum_{\mathrm{e}=1}^{n_{\mathrm{adj}}} \boldsymbol{\eta}_{\mathrm{e}}^{\mathrm{a}}\left(\xi_j^{\mathrm{s}}\right)\|} , \tag{3.46}$$

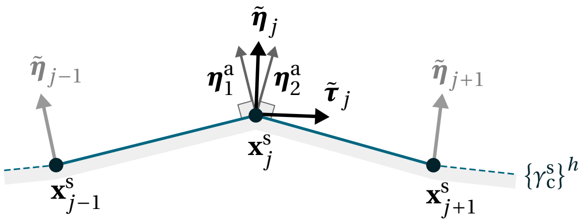
[†] As suggested by Popp, Gee, *et al.* (2009), no weighting whatsoever of the element normals $\boldsymbol{\eta}_{\mathrm{e}}^{\mathrm{a}}$ is considered.

Here, $\boldsymbol{\eta}_{\mathrm{e}}^{\mathrm{a}}$ stands for the *outward* unit normal of the adjacent non-mortar element e, evaluated at the slave node k, and n_{adj} is the total number of adjacent non-mortar facets (for further details, see Appendix A). Although this definition is equally applicable to both two- and three-dimensional problems, conceptually, the definition of the averaged unit tangent vectors requires additional consideration. While in two dimensions, the tangent vector is uniquely defined. In three dimensions, the pair of tangent vectors is not unique within the tangent plane to the non-mortar node. The strategy adopted in this work is based on considering the projection of the first edge connected with the non-mortar node j, see Figure 3.2 for a schematic representation.

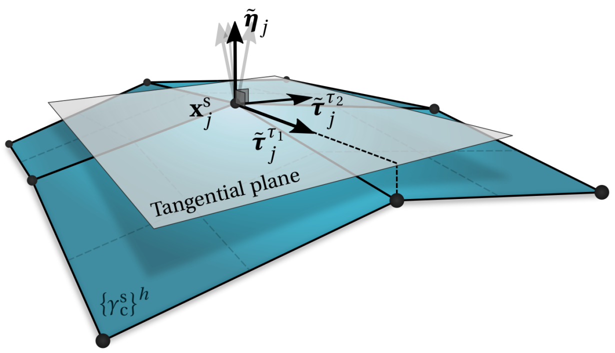
Based on this definition, a C^0-continuous field of normals is, then, defined through finite element interpolation as

$$\boldsymbol{\eta} \approx \boldsymbol{\eta}^{\mathrm{h}}\left(\xi^{\mathrm{s}}\right) \equiv \frac{\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} N_k^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right)\tilde{\boldsymbol{\eta}}_k}{\left\|\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} N_k^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right)\tilde{\boldsymbol{\eta}}_k\right\|}\,. \tag{3.47}$$

The normalization procedure is required, as the finite element interpolation may lead to a non-unit normal vector—which ultimately compromises the correct evaluation of the gap function. The same interpolation procedure is applied to the definition of the unit tangent vectors. As explained in B. Yang, Laursen, and Meng (2005), this procedure has clear advantages compared with traditional mortar segmentation techniques based on closest point projection. It smooths the discontinuities associated with the discretization of the contact interface, proving specially powerful at avoiding degenerate cases such as the non-uniqueness of the CPP.



(a) Averaged unit normal and tangent vector in two-dimensions.



(b) Averaged unit normal vector and tangent vectors in three-dimensions.

Figure 3.2: Averaged orthonormal moving frame for first-order interpolation.

3.6 Contact search

The next key step to be addressed is the contact search procedure. This is an essential algorithmic aspect of any contact formulation, in particular within the context of finite deformations, where large sliding motions can occur. The main goal is to identify bodies or finite elements that might possible come into contact, i.e., it defines the list of element pairs to be integrated. Historically, contact search has been an intense research topic since the origins of computational contact mechanics, with some popular approaches being found in Benson and J. O. Hallquist (1990), Zhi-Hua and Nilsson (1990), and Williams and O'Connor (1999). The reader is referred to Wriggers (2006) for a comprehensive overview of the subject.

In this work, the algorithm proposed by B. Yang and Laursen (2008) has been implemented. It consists of two components: a hierarchical global search structure (search tree), together with a simplified geometry representation using bounding volumes. It utilises discretised orientation polytopes with k edges (k-DOPs), which are typically 8-DOPs in two-dimensions and 18-DOPs in the three-dimensional case. During the problem initialisation, both the mortar and non-mortar boundaries are stored within a hierarchical binary tree structure in a top-down way. Starting from a root node, containing the entire associated boundary, each level is obtained by dividing into halves until reaching the individual finite element elements (leaf nodes). The search tree must be updated at each geometry update for finite deformations, i.e., at each iteration step. The contact search procedure itself consists of a recursive algorithm based on intersection tests between both boundaries. Starting at the root nodes, wherever an overlap of the bounding volumes is detected, the algorithm proceeds to the next lower level of the search tree. This procedure continues until reaching the leaf nodes, where the contact pair can be identified.

The computational complexity of the algorithm is expected to be $O(N \cdot \log(N^2))$.[‡] This is a considerably faster procedure than the naive brute-force approach of checking all the possible combinations of finite elements, which has a computational complexity of $O(N^2)$. Further development towards the parallelization of the algorithm can be found in Popp (2012).

3.7 Numerical integration of mortar integrals

As briefly mentioned in Section 3.3, the computation of mortar integrals is one of the main challenges of mortar contact algorithms. The reason for that is mainly related to the terms involving the relative description of between boundaries (e.g., the second mortar coupling matrix). They require integration over the non-mortar interface (i.e., curve and surface integrals for two-dimensional and three-dimensional contact problems, accordingly) with an integrand containing quantities defined on both mortar and non-mortar sides. Since several fundamental quantities, e.g., the weighted gap and relative tangential velocity, directly depend on the mortar coupling terms, this algorithmic aspect dramatically impacts on the global accuracy and computational efficiency of the overall numeri-

---

[‡] For the sake of simplicity, it is assumed that $n^{\mathrm{s}} \approx n^{\mathrm{m}} = N$.

cal procedure. Therefore, the selection of an adequate numerical integration scheme for the interface coupling terms is one fundamental aspect of the mortar-based frictional contact modelling.

In the literature, two approaches have been predominantly adopted: mortar integration by employing a *segmentation* strategy or utilizing an appropriate *element-wise* integration scheme. The first method, originally outlined by Simo, Wriggers, *et al.* (1985) and Zavarise and Wriggers (1998) for the classical segment-to-segment contact formulations, has been extensively applied in the context of mortar contact (Puso and Laursen, 2004a; B. Yang, Laursen, and Meng, 2005; Puso, Laursen, and Solberg, 2008; Popp, Gitterle, *et al.*, 2010) and is based on subdividing each integration domain into sub-domains—herein termed as *integration segments*. The idea is to prevent all possible discontinuities (typically in the form of kinks) that might deteriorate the achievable accuracy of the Gauss quadrature. The second approach, firstly introduced by Fischer and Wriggers (2005, 2006), ignores the possibility of discontinuities in the integrands and tries to minimize the resulting error by employing high-order integration rules. Both methodologies are explained in more detail in Section 3.7.1 and Section 3.7.2, accordingly.

A comprehensive comparison between these methods is presented in Farah, Popp, *et al.* (2015), where it has been concluded that both techniques demonstrate solid results, but with the element-based integration being significantly more efficient. However, the segment-based integration provided better results for particular scenarios, such as quadratic interpolation or sensitive frictional problems. In the same publication, an alternative integration strategy is also proposed in order to reach an ideal compromise between the characteristics of the both aforementioned techniques—designated as *boundary-segmentation*. The idea is to employ the segment-based strategy for problematic non-mortar elements having strong discontinuities, while using the element-based integration for the remaining non-critical elements. The identification of critical elements is performed by detecting if any Gauss point projection has failed. Based on the numerical experience of the author of the present work, this integration scheme provides solid results in general, while being much more computationally efficient than the segmentation strategy. However, as pointed out by Farah, Popp, *et al.* (2015), it should be applied carefully when dealing with quadratic interpolation or frictional problems, where the error in the mortar integrals may deteriorate the convergence of the global algorithm. Therefore, the logical recommendation is to use, whenever possible, the segment-based integration, while recurring to the boundary-segmentation for computationally demanding problems.

In the following paragraphs, both the element-wise and segment-based mortar integration algorithms are outlined for one integration segment, formed by one pair (s, m) of non-mortar and mortar elements. As mentioned in Section 3.6, this information about element pairs can be found using an efficient search strategy.

Remark. *It is noteworthy mentioning that the first mortar coupling matrix* $[\mathbf{D}]$ *could be evaluated directly by element-wise integration on* $\gamma_{\mathrm{c}}^{\mathrm{s}}$ *(theoretically without any compromise) since the integrand exclusively contains quantities defined on the non-mortar side. However, the segmentation integration scheme mentioned earlier is employed anyhow. As shown, for example, in Puso and Laursen (2004a) or Popp, Gitterle,* et al. *(2010), this reveals to be advantageous with regard to the fulfilment of the fundamental conservation laws in the discrete setting. A description of the integration strategy in two and three dimensions, including the set of projection rules and Gaussian quadrature, is described in the following paragraphs.*

3.7.1 Element-wise integration

Originally proposed in Fischer and Wriggers (2005, 2006), the element-based integration scheme defines the integration domain to be identical to the non-mortar boundary, i.e., the numerical integration is performed directly on the non-mortar elements without any segmentation, see Figure 3.3. This requires finding the projected coordinate on the mortar domain, $\hat{\boldsymbol{\xi}}_{\mathrm{g}}^{\mathrm{m}}$, which can be obtained by solving the local system of equations

$$\sum_{j=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left[N_j^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\mathbf{x}_j^{\mathrm{s}}\right] + \alpha\boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}) - \sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left[N_l^{\mathrm{m}}(\hat{\boldsymbol{\xi}}_{\mathrm{g}}^{\mathrm{m}})\mathbf{x}_l^{\mathrm{m}}\right] = \mathbf{0}\,. \tag{3.48}$$

Here, α denotes the normal distance between the points, and once again, the compact notation for the interpolated unit normal vector at the Gauss point $\boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}$ has been adopted. While for 2-noded line elements or 3-noded triangle surface elements, it leads to a linear system of equations (which can be solved directly), for the remaining cases, it requires the application of a local Newton-Raphson scheme. In three dimensions, each iteration reads

$$\begin{Bmatrix}\Delta\hat{\xi}_1^{\mathrm{m}}\\ \Delta\hat{\xi}_2^{\mathrm{m}}\\ \Delta\alpha\end{Bmatrix} = [\mathbf{W}]^{-1}\left\{\sum_{j=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left[N_j^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\mathbf{x}_j^{\mathrm{s}}\right] + \alpha\boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}) - \sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left[N_l^{\mathrm{m}}(\hat{\boldsymbol{\xi}}_{\mathrm{g}}^{\mathrm{m}})\mathbf{x}_l^{\mathrm{m}}\right]\right\}, \tag{3.49}$$

where the tangent matrix $[\mathbf{W}] \in \mathbb{R}^{d\times d}$ comes as

$$[\mathbf{W}] \equiv \left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left[N_{l,\hat{\xi}_1^{\mathrm{m}}}^{\mathrm{m}}(\hat{\boldsymbol{\xi}}_{\mathrm{g}}^{\mathrm{m}})\mathbf{x}_l^{\mathrm{m}}\right] \;\middle|\; \sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left[N_{l,\hat{\xi}_2^{\mathrm{m}}}^{\mathrm{m}}(\hat{\boldsymbol{\xi}}_{\mathrm{g}}^{\mathrm{m}})\mathbf{x}_l^{\mathrm{m}}\right] \;\middle|\; \boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\right]. \tag{3.50}$$

The two-dimensional version is rather straightforward by simply omitting the second column. As mentioned, this strategy will most likely lead to discontinuities in the integrands involving both mortar and non-mortar boundaries simultaneously (e.g., the discrete gap or second mortar matrix) and, therefore, should be used with caution.

3.7.2 Segment-based integration

As mentioned, the segment-based integration is based on the prevention of discontinuities by the establishment of smooth integrable segments. This is accomplished by the joint transmission of geometrical information between both boundaries, by means of projections based on the continuous field of normals. Even though the algorithmic approach to both two- and three-dimensional mortar contact shares the same concept, they require different strategies and, therefore, are discussed separately in the following.

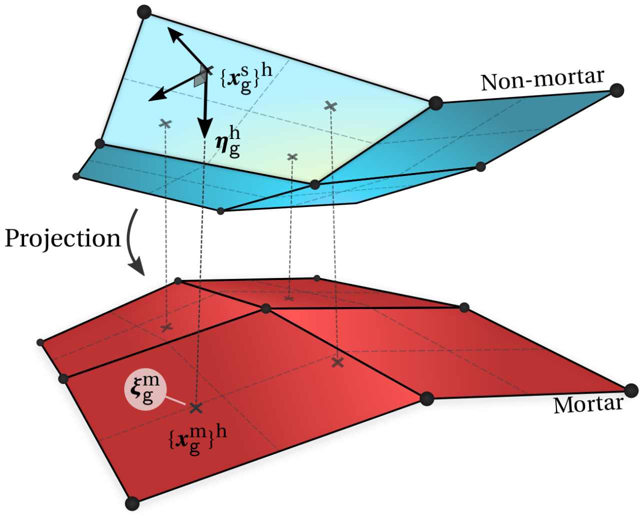


Figure 3.3: Illustration of the element-wise scheme. The integration Gauss points are projected directly from the discrete non-mortar boundary onto the mortar side, without caring for discontinuities. For example, in the particular case in the illustration each integration point falls onto different mortar elements.

3.7.2.1 Two-dimensional problems

As schematically represented in Figure 3.4, in two dimensions, the integration segments *endpoints* can be either nodes themselves or a projection of mortar nodes onto the non-mortar side—and vice versa. The projections are based on the *collinearity* condition with the outward normal vector, i.e., the projection of a non-mortar node $\mathbf{x}_a^s$ onto the related mortar element e with nodes $\mathbf{x}_l^m$ $(l = 1, \ldots, n_e^m)$ is obtained by solving

$$\left[\sum_{l=1}^{n_e^m}\left(N_l^m(\xi_a^m)\,\mathbf{x}_l^m\right)-\mathbf{x}_a^s\right]\times\tilde{\boldsymbol{\eta}}_a=\mathbf{0}\,, \tag{3.51}$$

where ξ_a^m is the sought-after projection coordinate; the projection of a mortar node $\mathbf{x}_b^m$ onto the related non-mortar element e, with nodes $\mathbf{x}_k^s$ $(k = 1, \ldots, n_e^s)$, can be determined by finding the coordinate ξ_b^s satisfying

$$\left[\mathbf{x}_b^m-\sum_{k=1}^{n_e^s}\left(N_k^s(\xi_b^s)\mathbf{x}_k^s\right)\right]\times\left[\sum_{k=1}^{n_e^s}\left(N_k^s(\xi_b^s)\tilde{\boldsymbol{\eta}}_k\right)\right]=\mathbf{0}\,. \tag{3.52}$$

In general, Equations (3.51) and (3.52) yield non-linear conditions for the projection coordinates, thus usually solved by a local Newton-Raphson scheme.

Finally, the local parameter space $\zeta \in [-1, 1]$ is introduced and the integrals are evaluated using Gauss quadrature. Considering the notation ξ_a^s and ξ_b^s for the integration segment end coordinates on the non-mortar side, the mapping from integration segment coordinates, ζ, to non-mortar coordinates, ξ^s, reads

$$\xi_g^s=\frac{1}{2}\left(1-\zeta_g\right)\xi_a^s+\frac{1}{2}\left(1+\zeta_g\right)\xi_b^s\,. \tag{3.53}$$

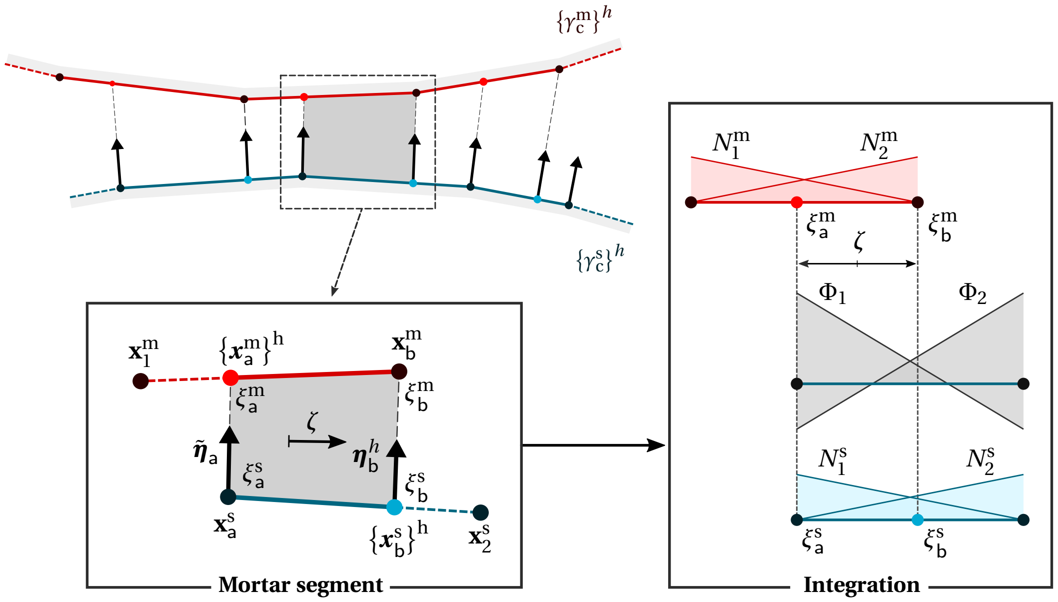


Figure 3.4: Schematic description of the segmentation method for two-dimensional first-order interpolation.

With regard to the mapping from the local integration coordinates ζ to the mortar coordinates, ξ^{m}, the technique proposed by Popp (2012) is adopted and, instead of establishing an interpolation rule for ξ^{m}, the individual Gauss points ζ_{g} (with global coordinates $\mathbf{x}_{\mathrm{g}}^{\mathrm{s}}$) are projected from the non-mortar the mortar side just as in Equation (3.51), i.e.,

$$\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\mathbf{x}_l^{\mathrm{m}}\right)-\mathbf{x}_{\mathrm{g}}^{\mathrm{s}}\right]\times\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\boldsymbol{\eta}}_k\right)\right]=\mathbf{0}\,. \tag{3.54}$$

The Gauss quadrature can finally be written as:

$$\mathrm{D}_{jk}^{(\mathrm{s,m})}\approx\sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}}\,\Phi_j\big(\xi_{\mathrm{g}}^{\mathrm{s}}(\zeta_{\mathrm{g}})\big)\,N_k^{\mathrm{s}}\big(\xi_{\mathrm{g}}^{\mathrm{s}}(\zeta_{\mathrm{g}})\big)J_{\mathrm{seg}}\,, \tag{3.55}$$

$$\mathrm{M}_{jl}^{(\mathrm{s,m})}\approx\sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}}\,\Phi_j\big(\xi_{\mathrm{g}}^{\mathrm{s}}(\zeta_{\mathrm{g}})\big)\,N_l^{\mathrm{m}}\big(\xi_{\mathrm{g}}^{\mathrm{m}}(\zeta_{\mathrm{g}})\big)J_{\mathrm{seg}}\,, \tag{3.56}$$

where J_{seg} comprises the set of mappings

$$J_{\mathrm{seg}}\equiv\left\|\frac{\partial\{\boldsymbol{x}^{\mathrm{s}}\}^{\mathrm{h}}}{\partial\xi^{\mathrm{s}}}\right\|\left|\frac{\partial\xi^{\mathrm{s}}}{\partial\zeta}\right|, \tag{3.57}$$

i.e., the mapping from physical space to non-mortar element parameter space and from non-mortar element space to local integration segment space.

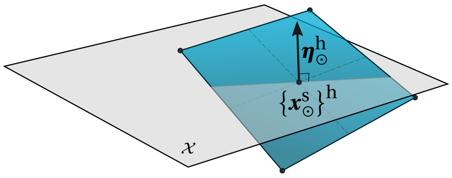
3.7.2.2 Three-dimensional problems

In the three-dimensional space, the geometrical treatment is altogether much more intricate. For instance, the definition of integration elements now involves arbitrarily shaped polygons. Beyond that, even the choice of an adequate integration surface itself is quite difficult. The approach followed in two dimensions of performing the integration directly on the non-mortar side is not trivial because of the possible warping of surface facets. Aiming to solve these difficulties, a simplified algorithm employing piecewise flat segments was originally proposed in Puso (2004) and is adopted here. For further details, the interested reader is also referred to Puso and Laursen (2004a,b). The main steps of the algorithm are outlined in Table 3.1.

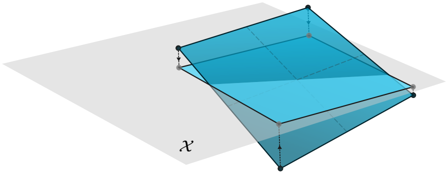
While capable of providing accurate results under any circumstances, the computational complexity of this algorithm is a significant drawback. The total number of projection operations grows considerably with problem size and, therefore, strategies to accelerate the algorithm are crucial. In Wilking and Bischoff (2017), this topic is tackled by combining adjacent triangular together to form quadrilateral integration domains. This significantly reduces the total number of Gauss points while maintaining integration accuracy. Despite the Jacobian determinant of quadrilateral integration cells being non-constant, its impact on the overall performance is negligible and, when combined with the suitable Gauss-Legendre quadrature, yields stable and accurate results. For example, in Figure 3.5 all the integration cells are represented for a simulation of a half torus, in which the vast majority of the cells can be merged to quadrilaterals. In Wilking and Bischoff (2017), the use of alternative quadratures with reduced integration points is also explored. However, it resulted in integration errors, mainly due to the rational polynomials emanating from the nodal projections and non-constant Jacobian determinants.

Table 3.1: Main steps of three-dimensional mortar integration algorithm.

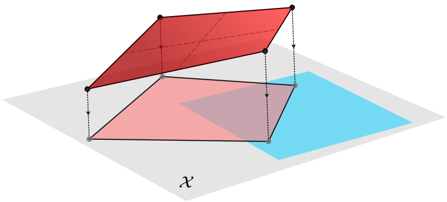
Step 1
Create an auxiliary plane $\mathcal{X}$ for numerical integration, based on the non-mortar element center $\{\boldsymbol{x}^{\mathrm{s}}_{\odot}\}^{\mathrm{h}}$ and the corresponding unit normal vector $\boldsymbol{\eta}^{\mathrm{h}}_{\odot}$.



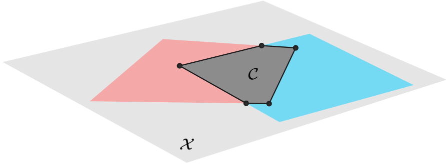
Step 2
Project all $n^{\mathrm{s}}_{\mathrm{e}}$ non-mortar element nodes $\mathbf{x}^{\mathrm{s}}_{k}$ $(k = 1,\dots,n^{\mathrm{s}}_{\mathrm{e}})$ onto the auxiliary plane along $\boldsymbol{\eta}^{\mathrm{h}}_{\odot}$. This can be interpreted as an approximation for the non-mortar surface obtained by removing element warping.



Step 3
Project all $n^{\mathrm{m}}_{\mathrm{e}}$ mortar element nodes $\mathbf{x}^{\mathrm{s}}_{l}$ $(l = 1,\dots,n^{\mathrm{m}}_{\mathrm{e}})$ onto the auxiliary plane along $\boldsymbol{\eta}^{\mathrm{h}}_{\odot}$.

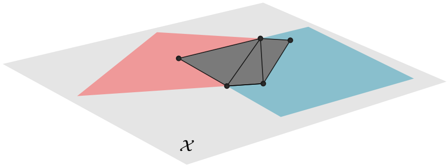


Step 4
Find the clip polygon $\mathcal{C}$ (see e.g., Foley *et al.* [1996]) to define the overlapping region of the projected non-mortar and mortar elements on the auxiliary plane $\mathcal{X}$.

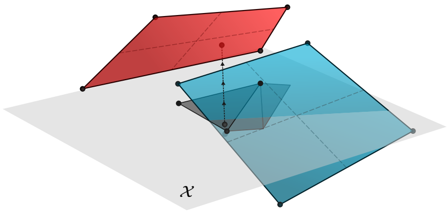


Step 5
Perform a decomposition (e.g., by applying Delauney triangulation) of the clip polygon $\mathcal{C}$ into n_{cell} easy-to-integrate triangular subdomains: herein termed *integration cells.* On each integration cell, a standard triangular finite element interpolation is employed, thus introducing the integration cell parameter space,

$$\boldsymbol{\zeta} = \left\{(\zeta_1,\zeta_2) \mid \zeta_1 \geq 0\,,\ \zeta_2 \geq 0\,,\ \zeta_1+\zeta_2 \leq 1\right\}. \tag{3.58}$$



Step 6
Define n_{g} Gauss integration points with coordinates $\boldsymbol{\zeta}_{\mathrm{g}}\,(\mathrm{g} = 1,\dots,n_{\mathrm{g}})$ on *each* cell (see, e.g., Cowper [1973]), and project back along $\boldsymbol{\eta}^{\mathrm{h}}_{\odot}$ to non-mortar and mortar elements to obtain $\boldsymbol{\xi}^{\mathrm{s}}(\boldsymbol{\zeta}_{\mathrm{g}})$ and $\boldsymbol{\xi}^{\mathrm{m}}(\boldsymbol{\zeta}_{\mathrm{g}})$. Then, perform Gauss integration of $\mathrm{D}^{(\mathrm{s,m})}_{jk}$ and $\mathrm{D}^{(\mathrm{s,m})}_{jk}$ (with $j,k = 1,\dots,n^{\mathrm{s}}_{\mathrm{e}}$ and $l = 1,\dots,n^{\mathrm{m}}_{\mathrm{e}}$) on *all* integration cells.



$$\mathrm{D}^{(\mathrm{s,m})}_{jk} \approx \sum_{c=1}^{n_{\mathrm{cell}}} \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}}\, \Phi_j\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}(\boldsymbol{\zeta}_{\mathrm{g}})\big)\, N^{\mathrm{s}}_{k}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}(\boldsymbol{\zeta}_{\mathrm{g}})\big) J^{\mathrm{cell}}_{c} \tag{3.59}$$

$$\mathrm{M}^{(\mathrm{s,m})}_{jl} \approx \sum_{c=1}^{n_{\mathrm{cell}}} \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}}\, \Phi_j\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}(\boldsymbol{\zeta}_{\mathrm{g}})\big)\, N^{\mathrm{m}}_{l}\big(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}}(\boldsymbol{\zeta}_{\mathrm{g}})\big) J^{\mathrm{cell}}_{c} \tag{3.60}$$

where J^{cell}_{c} is the integration cell c Jacobian determinant.

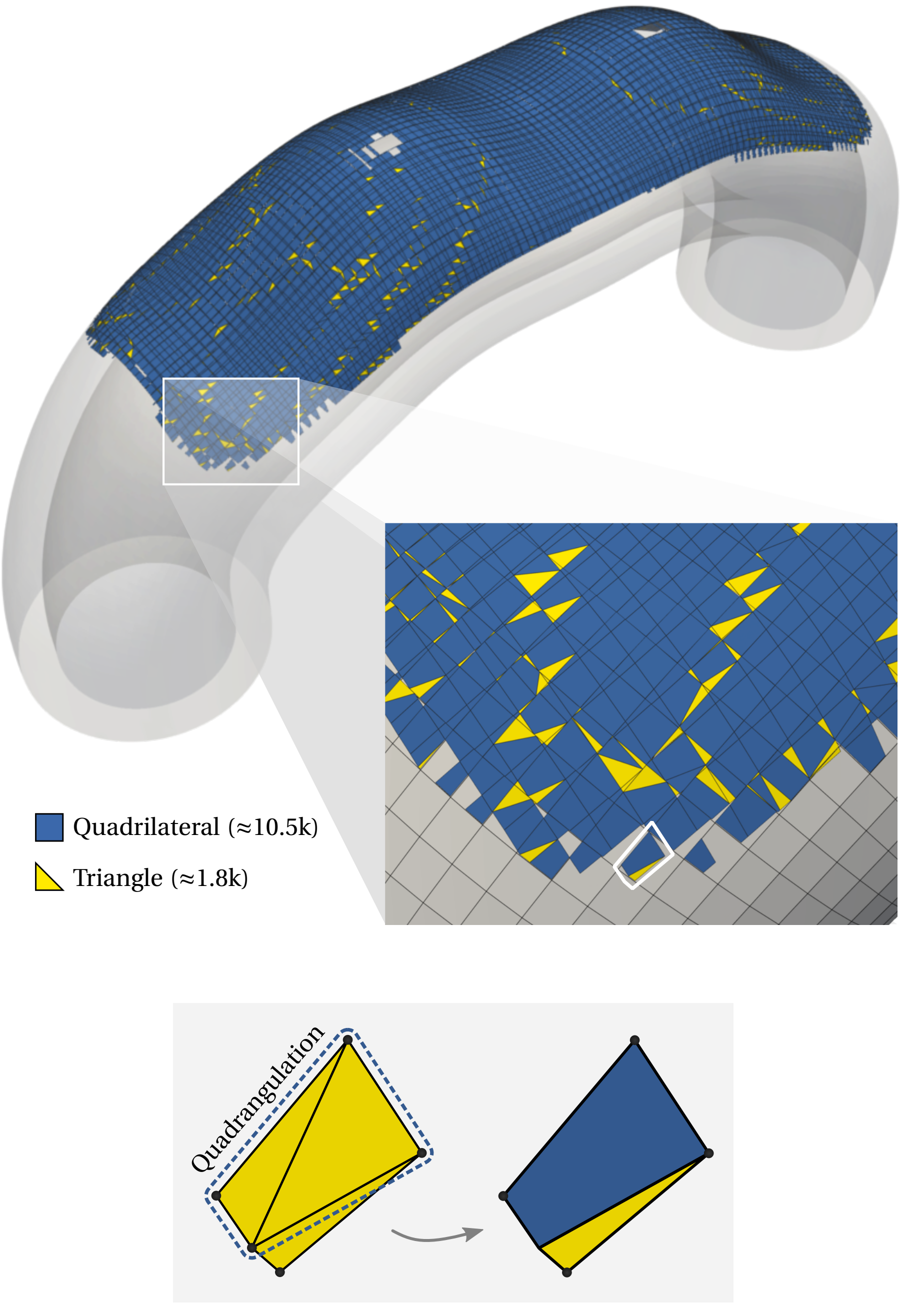


Figure 3.5: Integration cells in the simulation of a half torus with quadrangulation of integration cell triangles.

3.8 Summarized discrete form

With all the main aspects of the finite element approximation of mortar-based frictional contact discussed, the employed algebraic representation of the discrete entities involved is now introduced. This is done with the actual computational implementation of the algorithm in mind, and the resulting expressions are the starting point of the global solution method discussed in the rest of the chapter.

As a first step, all nodes (and corresponding degrees of freedom) of the two subdomains Ω^{s} and Ω^{m} are partitioned into three disjoint sets $\mathcal{S} \cup \mathcal{M} \cup \mathcal{N}$: a group $\mathcal{S}$ containing all non-mortar quantities, a group $\mathcal{M}$ of all mortar quantities and a group $\mathcal{N}$ associated with all remaining nodes or degrees of freedom. This allows for the definition of the displacement global vector $\{\mathbf{d}\} \in \mathbb{R}^{d \cdot n_{\mathrm{nod}}}$ as

$$\{\mathbf{d}\} \equiv \left\{ \begin{array}{c} \mathbf{d}_{\mathcal{N}} \\ \mathbf{d}_{\mathcal{M}} \\ \mathbf{d}_{\mathcal{S}} \end{array} \right\} . \tag{3.61}$$

The same structure applies to the virtual displacement global vector $\{\delta \mathbf{d}\}$. The discrete contact virtual work Equation (3.17) can thus be written as

$$\Pi_{\mathrm{c}}^{\mathrm{h}} = \{\delta \mathbf{d}\}^{\mathrm{T}} \{\mathbf{f}_{\mathrm{c}}\} , \tag{3.62}$$

where the discrete vector of contact forces $\{\mathbf{f}_{\mathrm{c}}(\mathbf{d},\mathbf{z})\} \in \mathbb{R}^{d \cdot n_{\mathrm{nod}}}$ is constructed as

$$\{\mathbf{f}_{\mathrm{c}}(\mathbf{d},\mathbf{z})\} \equiv [\mathbf{C}]\{\mathbf{z}\} . \tag{3.63}$$

Herein, the global vector $\{\mathbf{z}\} \in \mathbb{R}^{d \cdot n^{\lambda}}$ contains all the Lagrange multipliers and the discrete mortar contact operator $[\mathbf{C}(\mathbf{d})] \in \mathbb{R}^{(d \cdot n_{\mathrm{nod}}) \times (d \cdot n^{\lambda})}$ is defined as

$$[\mathbf{C}(\mathbf{d})] \equiv \begin{bmatrix} \mathbf{0} \\ -\mathbf{M}^{\mathrm{T}} \\ \mathbf{D}^{\mathrm{T}} \end{bmatrix} . \tag{3.64}$$

Finally, by introducing the global vectors of internal forces $\{\mathbf{f}_{\mathrm{int}}(\mathbf{d})\} \in \mathbb{R}^{d \cdot n_{\mathrm{nod}}}$ and external forces $\{\mathbf{f}_{\mathrm{ext}}\} \in \mathbb{R}^{d \cdot n_{\mathrm{nod}}}$, the discrete finite deformation contact problem can finally be stated as:

Problem 3.1 (The discrete form of finite deformation frictional contact)
Given the internal forces and boundary conditions, find the displacement vector, $\mathbf{d}$, *and Lagrange multiplier vector,* $\mathbf{z}$, *such that, for all* $t \in [t_0, T]$, *the set of discrete non-linear ordinary differential equations of motion including frictional contact are satisfied:*

$$\mathbf{f}_{\mathrm{int}}(\mathbf{d}) - \mathbf{f}_{\mathrm{ext}} + \mathbf{f}_{\mathrm{c}}(\mathbf{d},\mathbf{z}) = \mathbf{0} , \tag{3.65}$$

$$\tilde{g}_j \geq 0 , \quad z_j^{\eta} \geq 0 , \quad \tilde{g}_j z_j^{\eta} = 0 , \qquad \forall j \in \mathcal{S} , \tag{3.66}$$

$$\psi_j \leq 0 , \quad \tilde{\boldsymbol{u}}_j^{\tau} + \tilde{\beta}_j \mathbf{z}_j^{\tau} = \mathbf{0} , \quad \tilde{\beta}_j \geq 0 , \quad \psi_j \tilde{\beta}_j = 0 , \qquad \forall j \in \mathcal{S} . \tag{3.67}$$

3.9 Primal-dual active set strategy

When compared with the classical non-linear solid mechanics, the discrete contact problem comprises an additional source of non-linearity apart from the well-known geometrical and material non-linearities: the contact *inequality* constraints. In the discrete setting, the set of discrete normal constraints—enforced point-wise at each non-mortar node—is divided into two *a priori* unknown sets of active and inactive constraints. On top of that, the discrete tangential constraints for the active contact nodes must be fulfilled as well. Within the context of Coulomb's friction constitutive law, in particular, stick and sliding nodes need to be identified and the associated kinematical constraints enforced.

The idea of an active set strategy for frictional unilateral contact comes down to finding the correct subset of all non-mortar nodes which are effectively in contact and at the stick/slip condition at the end of each step of the global incremental algorithm. The simplest possible approach to this problem consists of defining an iterative scheme for the active set where, for each guess, the solution of the underlying convex problem is obtained. Then, the constraints are checked and updated, if necessary. However, this approach is undesirable from a computational efficiency standpoint, as it leads to two nested iterative solution schemes: the outer loop (fixed-point type) searching for the active set and the inner loop (Newton-Raphson type) solving the constrained non-linear finite element problem. To overcome this problem, an efficient and robust alternative is employed here: the so-called *Primal-Dual Active Set Strategy* (PDASS). The fundamental idea is to regularise the problem by reformulating the discrete nodal inequality constraints with the so-called *Non-linear Complementarity* (NCP) functions. This introduces a certain regularisation on the active set search and, as explained in more detail in Section 3.10, enables the application of a Newton-Raphson type algorithm, comprising not only geometrical and material non-linearities, but also the contact active set search itself as well. In the following paragraphs, the complementarity functions used to express the contact constraints in the normal and tangential directions are introduced.

3.9.1 Non-linear complementarity function for normal constraints

According to Hüeber and B. I. Wohlmuth (2005), the complementarity function in normal direction can be defined as

$$C_j^{\eta}\left(\mathbf{d},\mathbf{z}_j\right) \equiv \mathrm{z}_j^{\eta} - \max\left\{0\,,\, \mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\,\tilde{g}_j\right\}\,, \qquad \mathrm{c}^{\eta} > 0\,, \tag{3.68}$$

where c^{η} represents the so-called normal complementarity parameter (its concrete role is discussed later in this paragraph). With the NCP function (3.68) defined, it can be easily shown that the normal *inequality* constraints in Equation (3.66) can be equivalently rewritten as the following *equality* condition:

$$C_j^{\eta} = 0\,, \qquad \forall\, j \in \mathcal{S}\,. \tag{3.69}$$

Figure 3.6 schematically illustrates the nodal complementarity function, highlighting the equivalence with the KKT conditions. The distinction between the active and inactive sets is contained in the solution branch structure of the non-smooth max-operator. Thus, even though the NCP function (3.68) is continuous, it is non-smooth and has no uniquely defined derivative at $\mathrm{z}_j^\eta - \mathrm{c}^\eta\,\tilde{g}_j = 0$. Yet, the max-function can be classified as a semi-smooth function and directional derivatives can still be extracted, which allows for the application of a semi-smooth Newton method, see, e.g., Hintermüller *et al.* (2002). Herewith, all sources of non-linearity, i.e,. geometrical, material and contact itself, can be treated within a single iterative scheme.

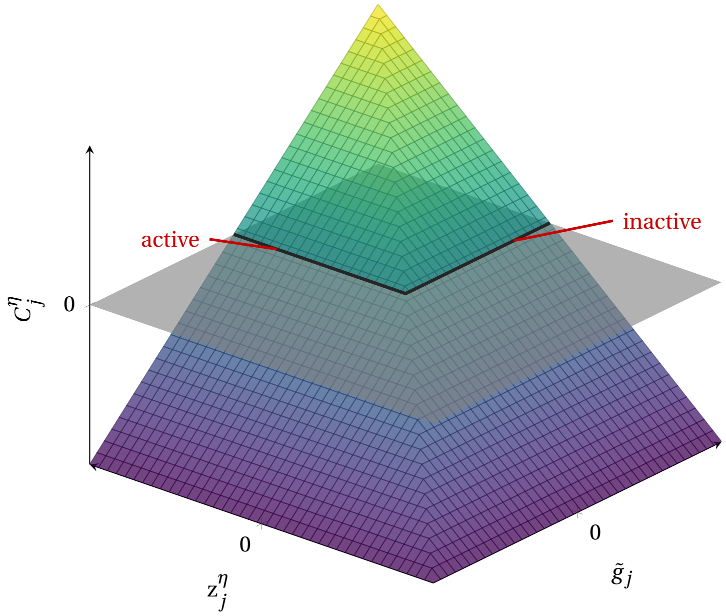


Figure 3.6: Nodal complementarity function C_j^η for the normal contact constraints (with $\mathrm{c}^\eta = 1$).

As a final remark, it is pointed out that the normal complementarity parameter c^η represents a purely algorithmic parameter. Despite its similarities, at first sight, with a penalty parameter, it only affects the convergence behaviour of the semi-smooth Newton algorithm—not the accuracy of the results. The value of c^η has been suggested in Hüeber and B. I. Wohlmuth (2005) to be of the order of the Young's modulus of the involved contacting bodies to obtain optimal convergence. However, several numerical investigations have shown very little influence on the semi-smooth Newton convergence rate (Popp, Gee, *et al.*, 2009; Popp, Gitterle, *et al.*, 2010).

3.9.2 Non-linear complementarity function for frictional constraints

Similar to the normal contact constraints, the PDASS strategy also reformulates the frictional inequality constraints (3.67) within a NCP function. There are various possibilities to define a suitable complementarity function, see, e.g., Alart and Curnier (1991), Christensen *et al.* (1998), and Hüeber (2008) for a comprehensive overview. In this work, the formulation introduced in Hüeber (2008) and successfully validated in Gitterle *et al.* (2010) is employed, which defined the tangential NCP function as

$$\begin{aligned}\boldsymbol{C}_j^\tau\left(\mathbf{d},\mathbf{z}_j\right) \equiv \mathbf{z}_j^\tau \max\left\{\mu\left(\mathrm{z}_j^\eta - \mathrm{c}^\eta\,\tilde{g}_j\right), \|\mathbf{z}_j^\tau + \mathrm{c}^\tau\,\tilde{\boldsymbol{u}}_j^\tau\|\right\}& \\ -\mu\left(\mathbf{z}_j^\tau + \mathrm{c}^\tau\,\tilde{\boldsymbol{u}}_j^\tau\right)\max\left\{0, \mathrm{z}_j^\eta - \mathrm{c}^\eta\,\tilde{g}_j\right\}, \quad \mathrm{c}^\eta, \mathrm{c}^\tau > 0\,.&\end{aligned} \tag{3.70}$$

As the normal complementarity parameter c^η, the tangential complementarity parameter c^τ is a purely algorithmic parameter affecting only the convergence behaviour (Gitterle *et al.*, 2010). Note that the NCP function (3.70) is formulated for the three-dimensional case and, thus, constitutes a vector with two components associated with each tangential direction at a contact point. For two-dimensional problems, Equation (3.70) reduces to a scalar NCP function. In analogy to the reformulated normal contact constraints in Equation (3.69), the solution of Coulomb's friction law in (3.67) can be equivalently expressed by the *equality* condition

$$\boldsymbol{C}_j^\tau = \boldsymbol{0}\,, \qquad \forall\, j \in \mathcal{S}\,. \tag{3.71}$$

Figure 3.7 gives a schematic description of the tangential NCP function (3.70) in the two-dimensional case and emphasizes the equivalence with the frictional contact constraints (3.67). The distinction between slip and stick nodes is again implicitly contained in the NCP function (3.70) due to the non-smoothness of the max operator.

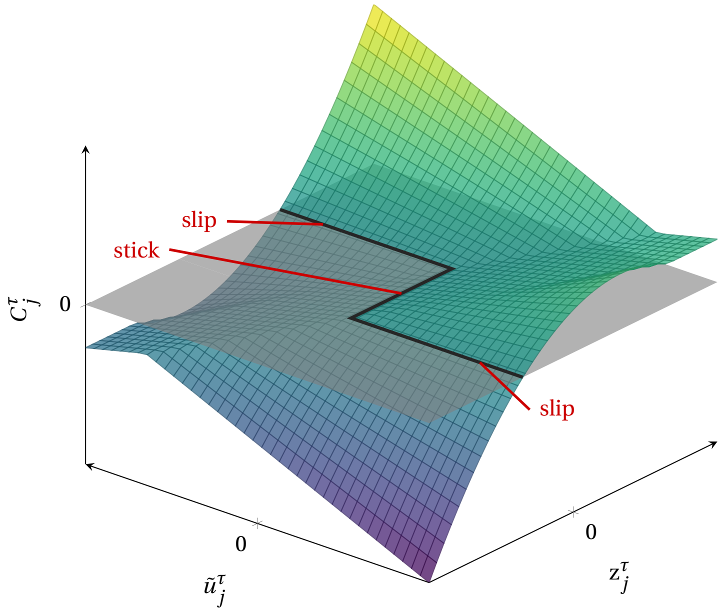


Figure 3.7: Nodal complementarity function C_j^τ for the tangential contact constraints in two-dimensions (with $c^\eta = 1$ and a constant value of $\mu z_j^\eta = 7$).

3.10 Semi-smooth Newton solution method

With the active set strategy thought out, attention is now shifted towards the actual numerical solution of the spatially discrete frictional contact problem stated in Problem 3.1. As mentioned in Section 3.9, the reformulation of inequality contact constraints as NCP functions can be interpreted as an algorithmic regularisation for the non-smoothness and multivalued character of contact conditions. The resulting PDASS algorithm is well-known from the general mathematical literature on constrained minimization and, as pointed out by Christensen *et al.* (1998) or Hintermüller *et al.* (2002), can be conveniently reinterpreted as a semi-smooth Newton method. For additional information on semi-smooth Newton methods, the interested reader is referred to Hintermüller *et al.* (2002) or Qi and Sun (1993). Regarding applications of semi-smooth methods for contact problems, early implementations within the context of classical NTS formulations can be found, e.g., in Alart and Curnier (1991), Strömberg *et al.* (1996), and Christensen *et al.* (1998) also referring to Koziara and Bićanić (2008) for frictional contact between pseudo-rigid bodies. In what concerns mortar-based contact formulations, in Hüeber and B. I. Wohlmuth (2005), the algorithm is applied for small deformation contact and, later on, extended by Popp, Gee, *et al.* (2009) and Popp, Gitterle, *et al.* (2010) for finite deformation dual mortar contact. The inclusion of frictional constraints is credited to Hüeber, G. Stadler, *et al.* (2008).

By introducing the NCP functions (3.69) and (3.71) in Problem 3.1, the discrete frictional contact problem can be rewritten as the set of *equality* conditions:

$$\mathbf{r}\left(\mathbf{d},\mathbf{z}\right) \equiv \mathbf{f}_{\text{int}}\left(\mathbf{d}\right) - \mathbf{f}_{\text{ext}} + \mathbf{f}_{\text{c}}\left(\mathbf{d},\mathbf{z}\right) = \mathbf{0}\,, \tag{3.72}$$

$$C_j^{\eta}\left(\mathbf{d},\mathbf{z}_j\right) = 0\,, \qquad \forall\, j \in \mathcal{S}\,, \tag{3.73}$$

$$\boldsymbol{C}_j^{\tau}\left(\mathbf{d},\mathbf{z}_j\right) = \mathbf{0}\,, \qquad \forall\, j \in \mathcal{S}\,. \tag{3.74}$$

At this point, Equations (3.72) to (3.74) establish the basis for the application of a Newton-type algorithm. Essentially, the discrete balance of linear momentum equation (here represented by the residual **r**) and discrete contact constraints (now expressed using the NCP functions) are *linearised*. Each non-linear solution step consists of solving the resulting linearised system of equations. An incremental update of the unknown displacements **d** and Lagrange multipliers **z** is applied repeatedly until a user-defined convergence criterion is met. Within the global incremental load algorithm, this solution method is applied for each load factor. With regard to the current formulation, the application of a Newton-type algorithm requires subsequent linearisation of the terms in Equations (3.72) to (3.74), as explained in the following paragraphs.

3.10.1 Consistent linearisation

The aforementioned semi-smooth Newton method is solved in each iteration k for the primal-dual pair $\left(\mathbf{d}^{\text{k}+1},\mathbf{z}^{\text{k}+1}\right)$ and is defined by:

$$\mathscr{D}\mathbf{r} = -\left\{\mathbf{r}\right\}^{\text{k}}\,, \tag{3.75}$$

$$\mathscr{D}C_j^{\eta} = -\left\{C_j^{\eta}\right\}^{\text{k}}\,, \qquad \forall\, j \in \mathcal{S}\,, \tag{3.76}$$

$$\mathscr{D}\boldsymbol{C}_j^{\tau} = -\left\{\boldsymbol{C}_j^{\tau}\right\}^{\text{k}}\,, \qquad \forall\, j \in \mathcal{S}\,, \tag{3.77}$$

followed by the update

$$\left(\mathbf{d}^{\text{k}+1},\mathbf{z}^{\text{k}+1}\right) = \left(\mathbf{d}^{\text{k}},\mathbf{z}^{\text{k}}\right) + \left(\Delta\mathbf{d},\Delta\mathbf{z}\right)\,. \tag{3.78}$$

Herein, the compact notation $\mathscr{D}\left(\bullet\right)$ has been introduced for the so-called *directional derivative*, i.e.

$$\mathscr{D}\left(\bullet\right) \equiv \left.\frac{\partial\left(\bullet\right)}{\partial\mathbf{d}}\right|^{\text{k}} \Delta\mathbf{d} + \left.\frac{\partial\left(\bullet\right)}{\partial\mathbf{z}}\right|^{\text{k}} \Delta\mathbf{z}\,. \tag{3.79}$$

The linearisation of the discrete balance equation residual $\mathbf{r}$, Equation (3.75), includes the linearisation of the internal forces vector $\mathbf{f}_{\mathrm{int}}$ and the contact forces vector $\mathbf{f}_{\mathrm{c}}$, viz.

$$\mathscr{D}\mathbf{r} = \mathscr{D}\mathbf{f}_{\mathrm{int}} + \mathscr{D}\mathbf{f}_{\mathrm{c}} \, . \tag{3.80}$$

The linearisation of the internal force vector $\mathbf{f}_{\mathrm{int}}$ with respect to the displacements leads to the well-known tangent stiffness matrix, herein represented by $\mathbf{K}$; since it only depends on the finite element technology employed—it is entirely independent of the contact enforcement—it is not addressed here. The remaining terms, related to contact, are treated individually in the following paragraphs.

Remark. *For notational convenience, the Newton iteration step superscript* k *is omitted from the linearisation derivations.*

3.10.1.1 Contact forces vector

By recalling the definition of the contact forces global vector in Equation (3.63), its linearisation yields

$$\mathscr{D}\mathbf{f}_{\mathrm{c}} = \begin{bmatrix} \mathbf{0} \\ -\mathscr{D}\mathbf{M}^{\mathrm{T}}\mathbf{z} - \mathbf{M}^{\mathrm{T}}\mathscr{D}\mathbf{z} \\ \mathscr{D}\mathbf{D}^{\mathrm{T}}\mathbf{z} + \mathbf{D}^{\mathrm{T}}\mathscr{D}\mathbf{z} \end{bmatrix} . \tag{3.81}$$

As can be observed, the directional derivative of the contact forces vector requires the consistent linearisation of the mortar coupling matrices $[\mathbf{D}]$ and $[\mathbf{M}]$, whose computation, as explained in great detail in Section 3.7, is performed by employing a segmentation-based strategy. Therefore, by taking into account the Gauss quadrature scheme given in Equations (3.55) and (3.56) for two-dimensional problems and Equations (3.59) and (3.60) for three-dimensional problems, the linearisation of one integration segment contribution to the entries D_{jk} of the first mortar coupling matrix $[\mathbf{D}]$ can be written as

$$\begin{aligned} \mathscr{D}\mathrm{D}_{jk} = & \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \, \mathscr{D}\Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, N_k^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \\ & + \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \, \Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, \mathscr{D}N_k^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \\ & + \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \, \Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, N_k^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, \mathscr{D}J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, , \end{aligned} \tag{3.82}$$

and the linearisation of the entries M_{jk} to the second mortar coupling matrix $[\mathbf{M}]$ reads

$$\begin{aligned} \mathscr{D}\mathrm{M}_{jl} = & \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \, \mathscr{D}\Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, N_l^{\mathrm{m}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}}\big) \, J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \\ & + \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \, \Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, \mathscr{D}N_l^{\mathrm{m}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}}\big) \, J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \\ & + \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \, \Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, N_l^{\mathrm{m}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}}\big) \, \mathscr{D}J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \, . \end{aligned} \tag{3.83}$$

Here, the generic notation J for the integration segment Jacobian determinant has been used, considering that it corresponds to the Jacobian determinant of the integration segment J_{seg} in two dimensions and the integration cell Jacobian determinant J^{cell} in three dimensions.

Examining the directional derivatives (3.82) and (3.83) in more detail, one can readily identify several element contributions: the linearisation of the dual shape functions Φ_j, the displacement shape functions N_k^{s} and N_l^{m} and the Jacobian determinant J. Applying the chain rule of differentiation, further elementary linearisations will naturally emerge, e.g., that of the averaged unit normal vector $\tilde{\boldsymbol{\eta}}_j$ and of the integration Gauss points $\xi_{\text{g}}^{\text{s}}$ or $\xi_{\text{g}}^{\text{m}}$. Since this stage of the algorithm accounts for the more significant part of the numerical effort of mortar-based finite deformation contact, in Appendix A, a detailed derivation and corresponding algebraic representation for all individual terms necessary to evaluate the directional derivatives of mortar coupling matrices is provided.

3.10.1.2 NCP function for normal constraints

While the contact forces global vector is differentiable in the classical sense, the NCP functions (3.68) and (3.70) contain the max-operator and, therefore, must be handled properly as a differentiable function (Hintermüller *et al.*, 2002). In terms of mathematical characterization, the resulting condition is classified as of Robin type since it involves both the displacement and the Lagrange multiplier.

As a preliminary step, recall that the generalised derivative of the max-function can be defined as

$$f(x) = \max\{a, x\} \longrightarrow \frac{\partial f(x)}{\partial x} = \begin{cases} 0\,, & \text{if } x \le a\,, \\ 1\,, & \text{if } x > a\,, \end{cases} \tag{3.84}$$

i.e., it reproduces the underlying branch structure of the max-function. Since active and inactive branches are separated, one can define the disjoint partitioning of the non-mortar set $\mathcal{S}$ into an inactive nodes set $\mathcal{I}$ and active nodes set $\mathcal{A}$ as

$$\mathcal{I} \equiv \left\{ j \in \mathcal{S} \mid \mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j \le 0 \right\}, \tag{3.85}$$

$$\mathcal{A} \equiv \left\{ j \in \mathcal{S} \mid \mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j > 0 \right\}, \tag{3.86}$$

which is tantamount to an update scheme to be applied after each iterative step. Defining $\chi_{(\bullet)}$ as a characteristic function for the set $(\bullet)$, i.e.

$$\chi_{(\bullet)} \equiv \begin{cases} 0\,, & \text{if } j \in (\bullet)\,, \\ 1\,, & \text{if } j \not\in (\bullet)\,, \end{cases} \tag{3.87}$$

the linearisation of the NCP function in the normal direction (3.68) can be written as

$$\mathscr{D} C_j^{\eta} \equiv \mathscr{D} \mathrm{z}_j^{\eta} - \chi_{\mathcal{A}} \left[\mathscr{D} \mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \mathscr{D} \tilde{g}_j \right], \tag{3.88}$$

with the abbreviation

$$\mathscr{D} \mathrm{z}_j^{\eta} \equiv \mathscr{D} \boldsymbol{\eta}_j^{\text{h}} \cdot \mathbf{z}_j + \boldsymbol{\eta}_j^{\text{h}} \cdot \mathscr{D} \mathbf{z}_j\,. \tag{3.89}$$

The numerical integration of the weighted gap function $\tilde{g}_j$, by considering its definition in Equation (3.29) and the employed segmentation scheme for numerical integration of mortar integrals, can be expressed by the generic notation

$$\tilde{g}_j \approx \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\, g^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}, \boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}}\big)\, J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\,, \tag{3.90}$$

whose consistent linearisation yields

$$\begin{aligned} \mathscr{D}\tilde{g}_j = & \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}} \mathscr{D}\Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\, g^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}, \boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}}\big)\, J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \\ & + \sum_{\mathrm{g}}^{n_{\mathrm{g}}} w_{\mathrm{g}} \Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\, \mathscr{D}g^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}, \boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}}\big)\, J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \\ & + \sum_{\mathrm{g}}^{n_{\mathrm{g}}} w_{\mathrm{g}} \Phi_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\, g^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}, \boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}}\big)\, \mathscr{D}J\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\,. \end{aligned} \tag{3.91}$$

Once again, the generic notation J for the integration segment Jacobian determinant has been used. For a detailed description of the linearisation of all the involved terms, refer to Appendix A.

Finally, consideration of the definitions (3.85) and (3.86) in Equation (3.88) results in the following compact notation:

- *Inactive nodes:* $\mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j \leq 0$

$$\{C_j^{\eta}\}_{\mathcal{I}} \equiv \mathrm{z}_j^{\eta} \tag{3.92}$$

$$\{\mathscr{D}C_j^{\eta}\}_{\mathcal{I}} = \mathscr{D}\mathrm{z}_j^{\eta} \tag{3.93}$$

- *Active nodes:* $\mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j > 0$

$$\{C_j^{\eta}\}_{\mathcal{A}} \equiv \mathrm{c}^{\eta} \tilde{g}_j \tag{3.94}$$

$$\{\mathscr{D}C_j^{\eta}\}_{\mathcal{A}} = \mathrm{c}^{\eta} \mathscr{D}\tilde{g}_j \tag{3.95}$$

3.10.1.3 NCP function for tangential constraints

First of all, for notational convenience, the abbreviations

$$b^{\eta}\left(\mathbf{d}, \mathbf{z}_j\right) = \mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j\,, \tag{3.96}$$

$$\boldsymbol{b}^{\tau}\left(\mathbf{d}, \mathbf{z}_j\right) = \mathbf{z}_j^{\tau} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}_j^{\tau}\,, \tag{3.97}$$

are introduced; this allows to rewrite Equation (3.70) using the compact notation

$$\boldsymbol{C}_j^{\tau} \equiv \mathbf{z}_j^{\tau} \max\{\mu\, b^{\eta}, \|\boldsymbol{b}^{\tau}\|\} - \mu\, \boldsymbol{b}^{\tau} \max\{0, b^{\eta}\}\,. \tag{3.98}$$

Like the NCP function in the normal direction, at this stage, it is convenient to introduce a characteristic function for the relevant sets. While in the normal direction, the set of non-mortar nodes is divided into two disjoint sets of active and inactive nodes; along the tangential direction, there are three sets: the set $\mathcal{I}$ of *inactive* nodes (not in contact); the set $\mathcal{Q}$ of nodes in contact and *stick* state; the set $\mathcal{L}$ of nodes in contact and *slip* condition. Mathematically, these disjoint sets follow as[§]:

$$\mathcal{I} \equiv \left\{ j \in \mathcal{S} \mid \mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j \le 0 \right\}, \tag{3.99}$$

$$\mathcal{Q} \equiv \left\{ j \in \mathcal{A} \mid \|\mathbf{z}_j^{\tau} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}_j^{\tau}\| - \mu \left[\mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j\right] < 0 \right\}, \tag{3.100}$$

$$\mathcal{L} \equiv \left\{ j \in \mathcal{A} \mid \|\mathbf{z}_j^{\tau} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}_j^{\tau}\| - \mu \left[\mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \tilde{g}_j\right] \ge 0 \right\}. \tag{3.101}$$

By using the characteristic function χ already introduced in Equation (3.87), the directional derivative of Equation (3.98) yields

$$\begin{aligned} \mathscr{D}\mathbf{C}_j^{\tau} = {} & \mathscr{D}\mathbf{z}_j^{\tau} \Big(\chi_{\mathcal{Q}}\, \mu\, b^{\eta} + \chi_{\mathcal{I}\cup\mathcal{L}}\, \|\boldsymbol{b}^{\tau}\| \Big) \\ & + \mathbf{z}_j^{\tau} \left(\chi_{\mathcal{Q}}\, \mu\, \mathscr{D} b^{\eta} + \chi_{\mathcal{I}\cup\mathcal{L}}\, \frac{\boldsymbol{b}^{\tau}}{\|\boldsymbol{b}^{\tau}\|} \cdot \mathscr{D}\boldsymbol{b}^{\tau} \right) \\ & - \chi_{\mathcal{Q}\cup\mathcal{L}} \Big(\mu\, b^{\eta}\, \mathscr{D}\boldsymbol{b}^{\tau} + \mu\, \boldsymbol{b}^{\tau}\, \mathscr{D} b^{\eta} \Big), \end{aligned} \tag{3.102}$$

where the individual linearisations $\mathscr{D} b^{\eta}$ and $\mathscr{D}\boldsymbol{b}^{\tau}$ are computed as

$$\mathscr{D} b^{\eta} = \mathscr{D}\mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\, \mathscr{D}\tilde{g}_j\,, \tag{3.103}$$

$$\mathscr{D}\boldsymbol{b}^{\tau} = \mathscr{D}\mathbf{z}_j^{\tau} + \mathrm{c}^{\tau}\, \mathscr{D}\tilde{\boldsymbol{u}}_j^{\tau}\,, \tag{3.104}$$

respectively. Here, the directional derivative of the nodal Lagrange multiplier projection in the tangential plane $\mathbf{z}_j^{\tau}$ reads

$$\mathscr{D}\mathbf{z}_j^{\tau} = \mathscr{D}\boldsymbol{\mathcal{P}}_j \mathbf{z}_j + \boldsymbol{\mathcal{P}}_j \mathscr{D}\mathbf{z}_j\,, \tag{3.105}$$

and, recalling the nodal weighted slip increment $\tilde{\boldsymbol{u}}_j^{\tau}$ definition in Equation (3.44), its linearisation results in

$$\begin{aligned} \mathscr{D}\tilde{\boldsymbol{u}}_j^{\tau} = {} & - \mathscr{D}\boldsymbol{\mathcal{P}}_j^{\mathrm{h}} \left[\left(\mathbf{D}_{[j,j]}^{t_n} - \mathbf{D}_{[j,j]}^{t_{n-1}} \right) \mathbf{x}_j^{\mathrm{s}} - \sum_{l=1}^{n^{\mathrm{m}}} \left(\mathbf{M}_{[j,l]}^{t_n} - \mathbf{M}_{[j,l]}^{t_{n-1}} \right) \mathbf{x}_l^{\mathrm{m}} \right] \\ & - \boldsymbol{\mathcal{P}}_j \left[\left(\mathscr{D}\mathbf{D}_{[j,j]}^{t_n} - \mathscr{D}\mathbf{D}_{[j,j]}^{t_{n-1}} \right) \mathbf{x}_j^{\mathrm{s}} - \sum_{l=1}^{n^{\mathrm{m}}} \left(\mathscr{D}\mathbf{M}_{[j,l]}^{t_n} - \mathscr{D}\mathbf{M}_{[j,l]}^{t_{n-1}} \right) \mathbf{x}_l^{\mathrm{m}} \right] \\ & - \boldsymbol{\mathcal{P}}_j \left[\left(\mathbf{D}_{[j,j]}^{t_n} - \mathbf{D}_{[j,j]}^{t_{n-1}} \right) \Delta\mathbf{x}_j^{\mathrm{s}} - \sum_{l=1}^{n^{\mathrm{m}}} \left(\mathbf{M}_{[j,l]}^{t_n} - \mathbf{M}_{[j,l]}^{t_{n-1}} \right) \Delta\mathbf{x}_l^{\mathrm{m}} \right]. \end{aligned} \tag{3.106}$$

The linearisation of the discrete tangential projection operator $\boldsymbol{\mathcal{P}}_j$ at non-mortar node j gathers the derivatives of individual unit tangent vectors, i.e.,

$$\mathscr{D}\boldsymbol{\mathcal{P}}_j = \begin{bmatrix} \{\mathscr{D}\boldsymbol{\tau}_j^1\}^{\mathrm{h}} \\ \vdots \\ \{\mathscr{D}\boldsymbol{\tau}_j^{d-1}\}^{\mathrm{h}} \end{bmatrix}. \tag{3.107}$$

[§] It should be noted that, although leading to identical sets (cf. Equation (3.85) and Equation (3.99), for instance), this division is *independent* of the partitioning resulting from the NCP function in the normal direction.

Again, all the individual entries in the expressions above can be found in Appendix A. Finally, by isolating each solution branch and expanding the notation, the following expressions are obtained:

- *Inactive nodes:* $\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \tilde{g}_j \leq 0$

$$\{\boldsymbol{C}_j^\tau\}_{\mathcal{I}} \equiv \mathbf{z}_j^\tau \|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\| \tag{3.108}$$

$$\{\mathscr{D}\boldsymbol{C}_j^\tau\}_{\mathcal{I}} = \mathscr{D}\mathbf{z}_j^\tau \|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\| + \mathbf{z}_j^\tau \left[\left(\frac{\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau}{\|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\|} \right) \cdot \left(\mathscr{D}\mathbf{z}_j^\tau + \mathrm{c}^\tau \mathscr{D}\tilde{\boldsymbol{u}}_j^\tau \right) \right] \tag{3.109}$$

- *Stick nodes:* $\|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\| - \mu \left[\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \tilde{g}_j \right] < 0$

$$\{\boldsymbol{C}_j^\tau\}_{\mathcal{Q}} \equiv -\mu \, \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau \left(\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \tilde{g}_j \right) \tag{3.110}$$

$$\{\mathscr{D}\boldsymbol{C}_j^\tau\}_{\mathcal{Q}} = -\mu \, \mathrm{c}^\tau \mathscr{D}\tilde{\boldsymbol{u}}_j^\tau \left(\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \tilde{g}_j \right) - \mu \, \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau \left(\mathscr{D}\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \mathscr{D}\tilde{g}_j \right) \tag{3.111}$$

- *Slip nodes:* $\|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\| - \mu \left[\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \tilde{g}_j \right] \geq 0$

$$\{\boldsymbol{C}_j^\tau\}_{\mathcal{L}} \equiv \mathbf{z}_j^\tau \|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\| - \mu \left(\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \tilde{g}_j \right) \left(\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau \right) \tag{3.112}$$

$$\begin{aligned} \{\mathscr{D}\boldsymbol{C}_j^\tau\}_{\mathcal{L}} = {} & \mathscr{D}\mathbf{z}_j^\tau \|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\| + \mathbf{z}_j^\tau \left[\left(\frac{\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau}{\|\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau\|} \right) \cdot \left(\mathscr{D}\mathbf{z}_j^\tau + \mathrm{c}^\tau \mathscr{D}\tilde{\boldsymbol{u}}_j^\tau \right) \right] \\ & - \mu \left(\mathscr{D}\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \mathscr{D}\tilde{g}_j \right) \left(\mathbf{z}_j^\tau + \mathrm{c}^\tau \tilde{\boldsymbol{u}}_j^\tau \right) - \mu \left(\mathrm{z}_j^\eta - \mathrm{c}^\eta \, \tilde{g}_j \right) \left(\mathscr{D}\mathbf{z}_j^\tau + \mathrm{c}^\tau \mathscr{D}\tilde{\boldsymbol{u}}_j^\tau \right) \end{aligned} \tag{3.113}$$

3.10.2 Semi-smooth Newton algorithm

With all the required derivatives obtained, the global semi-smooth Newton algorithm can finally be summarized. Before doing that, however, it should be noted that the Lagrange multipliers only enter the discrete frictional contact problem in a linear fashion. This means that one can write

$$\mathscr{D}\mathbf{z} = \mathbf{z}^{\mathrm{k}+1} - \mathbf{z}^{\mathrm{k}} \tag{3.114}$$

for all linearisations above, thus allowing to solve directly for the unknown Lagrange multipliers $\mathbf{z}^{\mathrm{k}+1}$ at each iteration, without any compromises (i.e., without employing an incremental formulation). Moreover, a closer inspection of the conditions in Equations (3.92) and (3.93) and Equations (3.108) and (3.109) for the inactive set $\mathcal{I}$ reveals that an additional simplification can still be adopted. Since for inactive nodes, the friction bound is zero, there is no need to enforce collinearity between $\mathbf{z}_j^\tau$ and $\tilde{\boldsymbol{u}}_j^\tau$ and, thus, the choice

$$\mathbf{z}_j = \mathbf{0}\,, \qquad \forall\, j \in \mathcal{I}\,, \tag{3.115}$$

is consistent within the iterative process.

The global algorithm is described in Algorithm 1, considering a semi-smooth Newton step at the current iterate k to be solved for the primal-dual pair of discrete variables

$\left(\mathbf{d}^{\mathrm{k}+1}, \mathbf{z}^{\mathrm{k}+1}\right)$. Herein, the variable ε_r represents an absolute Newton convergence tolerance for the L^2-norm of the total residual vector $\mathbf{r}_{\mathrm{tot}}$, which comprises not only the force residual, but the residual of the contact constraints as well.

The procedure described in Algorithm 1, besides being able to handle in an aggregate manner all sources of non-linearities within a single iterative loop—including the search for the correct sets for the contact constraints—has also shown to have a good convergence behaviour, see, e.g., Hüeber, G. Stadler, *et al.* (2008), Popp, Gee, *et al.* (2009), and Gitterle *et al.* (2010). Typically, while the contact sets are not found, locally superlinear convergence rates are expected due to the consistent linearisation (Hintermüller *et al.*, 2002). Once the sets are resolved and fixed, a standard (smooth) Newton-Raphson is recovered and locally quadratic convergence rates can be achieved.

Algorithm 1: Semi-smooth Newton algorithm

(1) Set $\mathrm{k} = 0$ and initialize the solution $\left(\mathbf{d}^0, \mathbf{z}^0\right)$

(2) Initialise $\mathcal{I}^0$, $\mathcal{A}^0$, $\mathcal{Q}^0$ and $\mathcal{L}^0$ such that:

$$\mathcal{I}^0 \cup \mathcal{A}^0 = \mathcal{S} \tag{3.116}$$

$$\mathcal{Q}^0 \cup \mathcal{L}^0 = \mathcal{A}^0 \tag{3.117}$$

$$\mathcal{I}^0 \cap \mathcal{Q}^0 = \mathcal{I}^0 \cap \mathcal{L}^0 = \mathcal{Q}^0 \cap \mathcal{L}^0 = \emptyset \tag{3.118}$$

(3) Find the primal-dual pair $\left(\Delta\mathbf{d}, \mathbf{z}^{\mathrm{k}+1}\right)$ by solving:

$$\mathscr{D}\mathbf{r} = -\mathbf{r}^{\mathrm{k}} \tag{3.119}$$

$$\mathbf{z}_j^{\mathrm{k}+1} = \mathbf{0}\,, \qquad \forall\, j \in \mathcal{I}^{\mathrm{k}} \tag{3.120}$$

$$\left\{\mathscr{D}C_j^\eta\right\}_{\mathcal{A}} = -\left\{C_j^\eta\right\}_{\mathcal{A}}^{\mathrm{k}}\,, \qquad \forall\, j \in \mathcal{A}^{\mathrm{k}} \tag{3.121}$$

$$\left\{\mathscr{D}\boldsymbol{C}_j^\tau\right\}_{\mathcal{Q}} = -\left\{\boldsymbol{C}_j^\tau\right\}_{\mathcal{Q}}^{\mathrm{k}}\,, \qquad \forall\, j \in \mathcal{Q}^{\mathrm{k}} \tag{3.122}$$

$$\left\{\mathscr{D}\boldsymbol{C}_j^\tau\right\}_{\mathcal{L}} = -\left\{\boldsymbol{C}_j^\tau\right\}_{\mathcal{L}}^{\mathrm{k}}\,, \qquad \forall\, j \in \mathcal{L}^{\mathrm{k}} \tag{3.123}$$

(4) Update displacements:

$$\mathbf{d}^{\mathrm{k}+1} = \mathbf{d}^{\mathrm{k}} + \Delta\mathbf{d} \tag{3.124}$$

(5) Set $\mathcal{I}^{\mathrm{k}+1}$, $\mathcal{Q}^{\mathrm{k}+1}$ and $\mathcal{L}^{\mathrm{k}+1}$ to:

$$\begin{aligned}
&\mathcal{I}^{\mathrm{k}+1} \equiv \left\{ j \in \mathcal{S} \mid \{\mathrm{z}_j^\eta\}^{\mathrm{k}+1} - \mathrm{c}^\eta \{\tilde{g}_j\}^{\mathrm{k}+1} \le 0 \right\} \\
&\mathcal{A}^{\mathrm{k}+1} \equiv \left\{ j \in \mathcal{S} \mid \{\mathrm{z}_j^\eta\}^{\mathrm{k}+1} - \mathrm{c}^\eta \{\tilde{g}_j\}^{\mathrm{k}+1} > 0 \right\} \\
&\mathcal{Q}^{\mathrm{k}+1} \equiv \left\{ j \in \mathcal{A}^{\mathrm{k}+1} \mid \|\{\mathbf{z}_j^\tau\}^{\mathrm{k}+1} + \mathrm{c}^\tau \{\tilde{\boldsymbol{u}}_j^\tau\}^{\mathrm{k}+1}\| - \mu \left[\{\mathrm{z}_j^\eta\}^{\mathrm{k}+1} - \mathrm{c}^\eta \{\tilde{g}_j\}^{\mathrm{k}+1} \right] < 0 \right\} \\
&\mathcal{L}^{\mathrm{k}+1} \equiv \left\{ j \in \mathcal{A}^{\mathrm{k}+1} \mid \|\{\mathbf{z}_j^\tau\}^{\mathrm{k}+1} + \mathrm{c}^\tau \{\tilde{\boldsymbol{u}}_j^\tau\}^{\mathrm{k}+1}\| - \mu \left[\{\mathrm{z}_j^\eta\}^{\mathrm{k}+1} - \mathrm{c}^\eta \{\tilde{g}_j\}^{\mathrm{k}+1} \right] \ge 0 \right\}
\end{aligned} \tag{3.125}$$

(6) `If` $\mathcal{I}^{\mathrm{k}+1} = \mathcal{I}^{\mathrm{k}}$, $\mathcal{A}^{\mathrm{k}+1} = \mathcal{A}^{\mathrm{k}}$, $\mathcal{Q}^{\mathrm{k}+1} = \mathcal{Q}^{\mathrm{k}}$, $\mathcal{L}^{\mathrm{k}+1} = \mathcal{L}^{\mathrm{k}}$ and $\|\mathbf{r}_{\mathrm{tot}}\| \le \varepsilon_r$ `then`

stop

`else`

set $\mathrm{k} := \mathrm{k} + 1$ and go to step **(3)**.

3.11 Algebraic representation for Dual mortar methods

As a final topic to be addressed in this chapter, an algebraic representation for the globally assembled matrix notations is provided. As before, special attention is given to the terms related to contact and, with the assembly procedure itself being relatively well-known from finite element technology, only abstract definitions of the individual blocks are given. Moreover, the condensation of the discrete Lagrange multipliers is explained as well. Due to the bi-orthogonality condition of dual Lagrange multipliers, it can be performed in an efficient and variationally consistent way for dual mortar methods.

3.11.1 Linearised system

As already mentioned, Algorithm 1 is constructed around the linearised system to be solved within each semi-smooth Newton iteration. In the following paragraphs, algebraic notations for the linearised versions of all involved terms are provided.

3.11.1.1 Discrete force equilibrium

The first term to be analysed is the discrete force equilibrium given in Equation (3.119). By isolating in the directional derivative of the contact forces vector derived in Equation (3.81), the terms associated with the derivative with respect to the discrete displacements vector $\{\mathbf{d}\}$ and Lagrange multipliers vector $\{\mathbf{z}\}$—thus exposing the vector $\{\Delta\mathbf{d}\}$ of unknown displacement increments—one can write that

$$\{\mathscr{D}\mathbf{f}_\mathrm{c}\} = \begin{bmatrix} \mathbf{0} \\ -\mathscr{D}\mathbf{M}^\mathrm{T} \\ \mathscr{D}\mathbf{D}^\mathrm{T} \end{bmatrix} \{\mathbf{z}\} + \begin{bmatrix} \mathbf{0} \\ -\mathbf{M}^\mathrm{T} \\ \mathbf{D}^\mathrm{T} \end{bmatrix} \{\mathscr{D}\mathbf{z}\} = \big[\widetilde{\mathbf{C}}\big]\{\Delta\mathbf{d}\} + \big[\mathbf{C}\big]\{\mathscr{D}\mathbf{z}\}\,. \tag{3.126}$$

Herein, the contact stiffness matrix $\big[\,\widetilde{\mathbf{C}}(\mathbf{d},\mathbf{z})\,\big] \in \mathbb{R}^{(d\cdot n_\mathrm{nod})\times(d\cdot n_\mathrm{nod})}$ contains the directional derivatives of both mortar coupling matrices $\big[\mathbf{D}\big]$ and $\big[\mathbf{M}\big]$ concerning the discrete displacements vector $\{\mathbf{d}\}$, together with the current Lagrange multipliers vector $\{\mathbf{z}\}$. Note that, by employing the nodal ordering for the incremental displacements vector $\{\Delta\mathbf{d}\}$ given in Equation (3.61), the matrix $\big[\widetilde{\mathbf{C}}\big]$ contains blocks associated with non-mortar and mortar degrees of freedom only, i.e.,

$$\big[\,\widetilde{\mathbf{C}}(\mathbf{d},\mathbf{z})\,\big] \equiv \begin{bmatrix} \mathbf{0} \\ \widetilde{\mathbf{C}}_\mathcal{M} \\ \widetilde{\mathbf{C}}_\mathcal{S} \end{bmatrix}\,. \tag{3.127}$$

Considering the directional derivative of the Lagrange multipliers vector in Equation (3.114), the following notation for the semi-smooth Newton step k for the contact forces vector $\{\mathbf{f}_\mathrm{c}\}$ can be obtained:

$$\begin{gathered} \{\mathscr{D}\mathbf{f}_\mathrm{c}\} = -\{\mathbf{f}_\mathrm{c}(\mathbf{d}^\mathrm{k},\mathbf{z}^\mathrm{k})\} \\ \big[\widetilde{\mathbf{C}}(\mathbf{d}^\mathrm{k},\mathbf{z}^\mathrm{k})\big]\{\Delta\mathbf{d}\} + \big[\mathbf{C}(\mathbf{d}^\mathrm{k})\big]\Big(\{\mathbf{z}^{\mathrm{k}+1}\} - \{\mathbf{z}^\mathrm{k}\}\Big) = -\big[\mathbf{C}(\mathbf{d}^\mathrm{k})\big]\{\mathbf{z}^\mathrm{k}\} \\ \big[\widetilde{\mathbf{C}}(\mathbf{d}^\mathrm{k},\mathbf{z}^\mathrm{k})\big]\{\Delta\mathbf{d}\} + \big[\mathbf{C}(\mathbf{d}^\mathrm{k})\big]\{\mathbf{z}^{\mathrm{k}+1}\} = \mathbf{0}\,. \end{gathered} \tag{3.128}$$

Thus, each iteration of the discrete force equilibrium in Equation (3.119) yields

$$\begin{gathered} \{\mathscr{D}\mathbf{r}\} = -\{\mathbf{r}(\mathbf{d}^\mathrm{k},\mathbf{z}^\mathrm{k})\} \\ \big[\widetilde{\mathbf{K}}(\mathbf{d}^\mathrm{k},\mathbf{z}^\mathrm{k})\big]\{\Delta\mathbf{d}\} + \big[\mathbf{C}(\mathbf{d}^\mathrm{k})\big]\{\mathbf{z}^{\mathrm{k}+1}\} = -\{\widetilde{\mathbf{r}}(\mathbf{d}^\mathrm{k})\} \end{gathered} \tag{3.129}$$

with the matrix $\left[\widetilde{\mathbf{K}}(\mathbf{d},\mathbf{z})\right] \in \mathbb{R}^{(d\cdot n_{\mathrm{nod}})\times(d\cdot n_{\mathrm{nod}})}$ representing the *effective stiffness matrix including contact*, i.e.

$$\left[\widetilde{\mathbf{K}}(\mathbf{d})\right] \equiv \left[\mathbf{K}(\mathbf{d})\right] + \left[\widetilde{\mathbf{C}}(\mathbf{d},\mathbf{z})\right]\,, \tag{3.130}$$

and the vector $\left\{\tilde{\mathbf{r}}(\mathbf{d})\right\} \in \mathbb{R}^{d\cdot n_{\mathrm{nod}}}$ an abbreviation for the residual,

$$\left\{\tilde{\mathbf{r}}(\mathbf{d})\right\} \equiv \left\{\mathbf{f}_{\mathrm{int}}\right\} - \left\{\mathbf{f}_{\mathrm{ext}}\right\}. \tag{3.131}$$

Note that, by solving directly for the current Lagrange multipliers vector $\left\{\mathbf{z}^{\mathrm{k}+1}\right\}$, the current contact forces vector $\left\{\mathbf{f}_{\mathrm{c}}\left(\mathbf{d}^{\mathrm{k}},\mathbf{z}^{\mathrm{k}}\right)\right\}$ vanishes, see Equation (3.128), and the final residual $\left\{\tilde{\mathbf{r}}\right\}$ simplifies to the difference between internal and external forces, as in standard finite element methods.

3.11.1.2 Normal contact constraints

As can be observed in Equations (3.94) and (3.95), the NCP function in the normal direction for the active set $\{C_j^{\eta}\}_{\mathcal{A}}$ and its linearisation $\{\mathscr{D}C_j^{\eta}\}_{\mathcal{A}}$ involve only the weighted gap function $\tilde{g}_j$ and its directional derivative, accordingly. Therefore, the semi-smooth Newton iteration step k can be simply written as

$$\mathscr{D}\tilde{g}_j = -\tilde{g}_j^{\mathrm{k}}\,, \qquad \forall\, j \in \mathcal{A}^{\mathrm{k}}\,, \tag{3.132}$$

or, by converting to an algebraic representation, equivalently as

$$\left[\mathbf{A}_{\mathcal{M}}\left(\mathbf{d}^{\mathrm{k}}\right)\right]\left\{\Delta\mathbf{d}_{\mathcal{M}}\right\} + \left[\mathbf{A}_{\mathcal{S}}\left(\mathbf{d}^{\mathrm{k}}\right)\right]\left\{\Delta\mathbf{d}_{\mathcal{S}}\right\} = -\left\{\tilde{\mathbf{g}}\left(\mathbf{d}^{\mathrm{k}}\right)\right\}. \tag{3.133}$$

Here, $\left\{\tilde{\mathbf{g}}(\mathbf{d})\right\} \in \mathbb{R}^{n^{\mathrm{a}}}$ represents a global vector gathering the weighted gaps $\tilde{g}_j$ for all the n^{a} active non-mortar nodes,

$$\underset{j\in\mathcal{A}}{\mathsf{A}}\, \tilde{g}_j = \left\{\tilde{\mathbf{g}}(\mathbf{d})\right\}, \tag{3.134}$$

and the matrices $\left[\mathbf{A}_{\mathcal{M}}(\mathbf{d})\right] \in \mathbb{R}^{n^{\mathrm{a}}\times(d\cdot n^{\mathrm{m}})}$ and $\left[\mathbf{A}_{\mathcal{S}}(\mathbf{d})\right] \in \mathbb{R}^{n^{\mathrm{a}}\times(d\cdot n^{\mathrm{s}})}$ stand for the assembly of all directional derivatives covered in $\mathscr{D}\tilde{g}_j$, i.e.

$$\underset{j\in\mathcal{A}}{\mathsf{A}}\, \mathscr{D}\tilde{g}_j = \left[\mathbf{A}_{\mathcal{M}}\right]\left\{\Delta\mathbf{d}_{\mathcal{M}}\right\} + \left[\mathbf{A}_{\mathcal{S}}\right]\left\{\Delta\mathbf{d}_{\mathcal{S}}\right\}. \tag{3.135}$$

Herein, A denotes the standard finite element assembly operator.

3.11.1.3 Tangential contact constraints

While in the normal direction, only matrix notations for the active set are needed, in the tangential direction, both stick and slip conditions need to be imposed throughout the iterative scheme. Starting by the stick regime, according to Equations (3.110) and (3.111), the semi-smooth Newton step k can be written in full extent as

$$\begin{aligned} -\mu\,\mathrm{c}^{\tau}\,\mathscr{D}\tilde{\boldsymbol{u}}_j^{\tau}\left(\left\{\mathrm{z}_j^{\eta}\right\}^{\mathrm{k}} - \mathrm{c}^{\eta}\left\{\tilde{g}_j\right\}^{\mathrm{k}}\right) - \mu\,\mathrm{c}^{\tau}\left\{\tilde{\boldsymbol{u}}_j^{\tau}\right\}^{\mathrm{k}}\left(\mathscr{D}\mathrm{z}_j^{\eta} - \mathrm{c}^{\eta}\,\mathscr{D}\tilde{g}_j\right) \\ = \mu\,\mathrm{c}^{\tau}\left\{\tilde{\boldsymbol{u}}_j^{\tau}\right\}^{\mathrm{k}}\left(\left\{\mathrm{z}_j^{\eta}\right\}^{\mathrm{k}} - \mathrm{c}^{\eta}\left\{\tilde{g}_j\right\}^{\mathrm{k}}\right), \qquad \forall\, j \in \mathcal{Q}^{\mathrm{k}}\,. \end{aligned} \tag{3.136}$$

Considering that insertion of Equation (3.114) into Equation (3.89) yields

$$\mathscr{D}\mathrm{z}_j^{\eta} = \mathscr{D}\boldsymbol{\eta}_j^{\mathrm{h}}\cdot\left\{\mathbf{z}_j\right\}^{\mathrm{k}} + \left\{\boldsymbol{\eta}_j^{\mathrm{h}}\right\}^{\mathrm{k}}\cdot\left(\left\{\mathbf{z}_j\right\}^{\mathrm{k}+1} - \left\{\mathbf{z}_j\right\}^{\mathrm{k}}\right), \tag{3.137}$$

it follows that Equation (3.136) can be further simplified to

$$\{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{Q}} = -\{\widetilde{\boldsymbol{C}}^{\tau}_{j}\}^{\mathrm{k}}_{\mathcal{Q}}\,, \qquad \forall\, j \in \mathcal{Q}^{\mathrm{k}}\,, \tag{3.138}$$

with the abbreviations:

$$\{\widetilde{\boldsymbol{C}}^{\tau}_{j}\}^{\mathrm{k}}_{\mathcal{Q}} \equiv -\,\mu\,\mathrm{c}^{\eta}\,\mathrm{c}^{\tau}\,\{\tilde{\mathrm{g}}_{j}\}^{\mathrm{k}}\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}\,; \tag{3.139}$$

$$\begin{aligned}\{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{Q}} \equiv &-\,\mu\,\mathrm{c}^{\tau}\,\mathscr{D}\tilde{\boldsymbol{u}}^{\tau}_{j}\left(\{\boldsymbol{\eta}^{\mathrm{h}}_{j}\}^{\mathrm{k}}\cdot\{\mathbf{z}_{j}\}^{\mathrm{k}} - \mathrm{c}^{\eta}\,\{\tilde{\mathrm{g}}_{j}\}^{\mathrm{k}}\right)\\ &-\,\mu\,\mathrm{c}^{\tau}\,\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}\left(\mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{j}\cdot\{\mathbf{z}_{j}\}^{\mathrm{k}} + \{\boldsymbol{\eta}^{\mathrm{h}}_{j}\}^{\mathrm{k}}\cdot\{\mathbf{z}_{j}\}^{\mathrm{k+1}} - \mathrm{c}^{\eta}\,\mathscr{D}\tilde{\mathrm{g}}_{j}\right)\,.\end{aligned} \tag{3.140}$$

The equations for all n^{stick} sticking nodes can be equally written in matrix form as

$$[\mathbf{Q}_{\mathcal{M}}(\mathbf{d}^{\mathrm{k}},\mathbf{z}^{\mathrm{k}})]\{\Delta\mathbf{d}_{\mathcal{M}}\} + [\mathbf{Q}_{\mathcal{S}}(\mathbf{d}^{\mathrm{k}},\mathbf{z}^{\mathrm{k}})]\{\Delta\mathbf{d}_{\mathcal{S}}\} + [\mathbf{T}(\mathbf{d}^{\mathrm{k}})]\{\mathbf{z}^{\mathrm{k+1}}_{\mathcal{Q}}\} = -\{\mathbf{r}^{\mathrm{stick}}(\mathbf{d}^{\mathrm{k}})\}\,, \tag{3.141}$$

where the vector $\{\mathbf{r}^{\mathrm{stick}}(\mathbf{z})\} \in \mathbb{R}^{n^{\mathrm{stick}}}$ represents the residual,

$$\mathop{\mathsf{A}}_{j\in\mathcal{Q}}\{\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{Q}} = \{\mathbf{r}^{\mathrm{stick}}\}\,, \tag{3.142}$$

and $[\mathbf{Q}_{\mathcal{M}}(\mathbf{d},\mathbf{z})] \in \mathbb{R}^{n^{\mathrm{stick}}\times(d\cdot n^{\mathrm{m}})}$, $[\mathbf{Q}_{\mathcal{S}}(\mathbf{d},\mathbf{z})] \in \mathbb{R}^{n^{\mathrm{stick}}\times(d\cdot n^{\mathrm{s}})}$ and $[\mathbf{T}(\mathbf{d})] \in \mathbb{R}^{n^{\mathrm{stick}}\times(d\cdot n^{\mathrm{stick}})}$ are the matrices containing the directional derivatives covered in $\{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{Q}}$,

$$\mathop{\mathsf{A}}_{j\in\mathcal{Q}}\{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{Q}} = [\mathbf{Q}_{\mathcal{M}}]\{\Delta\mathbf{d}_{\mathcal{M}}\} + [\mathbf{Q}_{\mathcal{S}}]\{\Delta\mathbf{d}_{\mathcal{S}}\} + [\mathbf{T}]\{\mathbf{z}^{\mathrm{k+1}}_{\mathcal{Q}}\}\,. \tag{3.143}$$

Focusing now on the slip state, by recalling Equations (3.112) and (3.113), the semi-smooth Newton step k is defined as

$$\begin{aligned}&\mathscr{D}\mathbf{z}^{\tau}_{j}\left\|\{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}} + \mathrm{c}^{\tau}\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}\right\| + \{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}}\frac{\{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}} + \mathrm{c}^{\tau}\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}}{\left\|\{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}} + \mathrm{c}^{\tau}\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}\right\|}\cdot\left(\mathscr{D}\mathbf{z}^{\tau}_{j} + \mathrm{c}^{\tau}\,\mathscr{D}\tilde{\boldsymbol{u}}^{\tau}_{j}\right)\\ &-\mu\left(\mathscr{D}\mathrm{z}^{\eta}_{j} - \mathrm{c}^{\eta}\,\mathscr{D}\tilde{\mathrm{g}}_{j}\right)\left(\{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}} + \mathrm{c}^{\tau}\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}\right) - \mu\left(\{\mathrm{z}^{\eta}_{j}\}^{\mathrm{k}} - \mathrm{c}^{\eta}\,\{\tilde{\mathrm{g}}_{j}\}^{\mathrm{k}}\right)\left(\mathscr{D}\mathbf{z}^{\tau}_{j} + \mathrm{c}^{\tau}\,\mathscr{D}\tilde{\boldsymbol{u}}^{\tau}_{j}\right)\\ &= -\{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}}\left\|\{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}} + \mathrm{c}^{\tau}\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}\right\| + \mu\left(\{\mathrm{z}^{\eta}_{j}\}^{\mathrm{k}} - \mathrm{c}^{\eta}\,\{\tilde{\mathrm{g}}_{j}\}^{\mathrm{k}}\right)\left(\{\mathbf{z}^{\tau}_{j}\}^{\mathrm{k}} + \mathrm{c}^{\tau}\{\tilde{\boldsymbol{u}}^{\tau}_{j}\}^{\mathrm{k}}\right), \quad \forall\, j \in \mathcal{L}^{\mathrm{k}}\,.\end{aligned} \tag{3.144}$$

Then, substitution of Equation (3.114) into Equation (3.105) yields

$$\mathscr{D}\mathbf{z}^{\tau}_{j} = \mathscr{D}\boldsymbol{\mathcal{P}}_{j}\{\mathbf{z}_{j}\}^{\mathrm{k}} + \boldsymbol{\mathcal{P}}_{j}\left(\{\mathbf{z}_{j}\}^{\mathrm{k+1}} - \{\mathbf{z}_{j}\}^{\mathrm{k}}\right)\,, \tag{3.145}$$

and Equation (3.144) can be rewritten as

$$\{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{L}} = -\{\widetilde{\boldsymbol{C}}^{\tau}_{j}\}^{\mathrm{k}}_{\mathcal{L}}\,, \tag{3.146}$$

with

$$\{\widetilde{\boldsymbol{C}}^{\tau}_{j}\}^{\mathrm{k}}_{\mathcal{L}} \equiv \mu\, \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}^{\tau}_{j} \mathrm{c}^{\eta}\, \tilde{\mathrm{g}}_{j} + \left[\mu \boldsymbol{\eta}^{\mathrm{h}}_{j} \cdot \mathbf{z}_{j} - \left(\frac{\boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}^{\tau}_{j}}{\| \boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}^{\tau}_{j} \|} \right) \cdot \left(\boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} \right) \right] \boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j}\,, \tag{3.147}$$

$$\begin{aligned} \{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{L}} \equiv{}& \| \boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}^{\tau}_{j} \| \left(\mathscr{D}\boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \boldsymbol{\mathcal{P}}_{j} \{\mathbf{z}_{j}\}^{\mathrm{k+1}} \right) \\ &+ \left[\left(\frac{\boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}^{\tau}_{j}}{\| \boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}^{\tau}_{j} \|} \right) \cdot \left(\mathscr{D}\boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \boldsymbol{\mathcal{P}}_{j} \{\mathbf{z}_{j}\}^{\mathrm{k+1}} + \mathrm{c}^{\tau}\, \mathscr{D}\tilde{\boldsymbol{u}}^{\tau}_{j} \right) \right] \boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} \\ &- \mu \left(\mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{j} \cdot \mathbf{z}_{j} + \boldsymbol{\eta}^{\mathrm{h}}_{j} \cdot \{\mathbf{z}_{j}\}^{\mathrm{k+1}} - \mathrm{c}^{\eta}\, \mathscr{D}\tilde{\mathrm{g}}_{j} \right) \left(\boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \mathrm{c}^{\tau}\, \tilde{\boldsymbol{u}}^{\tau}_{j} \right) \\ &- \mu \left(\boldsymbol{\eta}^{\mathrm{h}}_{j} \cdot \mathbf{z}_{j} - \mathrm{c}^{\eta}\, \tilde{\mathrm{g}}_{j} \right) \left(\mathscr{D}\boldsymbol{\mathcal{P}}_{j} \mathbf{z}_{j} + \boldsymbol{\mathcal{P}}_{j} \{\mathbf{z}_{j}\}^{\mathrm{k+1}} + \mathrm{c}^{\tau}\, \mathscr{D}\tilde{\boldsymbol{u}}^{\tau}_{j} \right). \end{aligned} \tag{3.148}$$

Here, for ease of notation, the iteration index k has been omitted, with the exception of the sought-after unknown Lagrange multiplier vector $\{\mathbf{z}_j\}^{\mathrm{k+1}}$; it should be kept in mind, though, that all the remaining terms are evaluated at the current iteration state k. Each semi-smooth Newton step comprising all n^{slip} slipping nodes can be alternatively be written using the algebraic notation

$$[\mathbf{L}_{\mathcal{M}}(\mathbf{d}^{\mathrm{k}},\mathbf{z}^{\mathrm{k}})]\{\Delta\mathbf{d}_{\mathcal{M}}\} + [\mathbf{L}_{\mathcal{S}}(\mathbf{d}^{\mathrm{k}},\mathbf{z}^{\mathrm{k}})]\{\Delta\mathbf{d}_{\mathcal{S}}\} + [\mathbf{H}(\mathbf{d}^{\mathrm{k}})]\{\mathbf{z}^{\mathrm{k+1}}_{\mathcal{Q}}\} = -\{\mathbf{r}^{\mathrm{slip}}(\mathbf{d}^{\mathrm{k}})\}\,, \tag{3.149}$$

with the residual vector $\{\mathbf{r}^{\mathrm{slip}}(\mathbf{z})\} \in \mathbb{R}^{n^{\mathrm{slip}}}$ assembled as

$$\mathop{\mathsf{A}}_{j\in\mathcal{L}} \{\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{L}} = \{\mathbf{r}^{\mathrm{slip}}\}\,, \tag{3.150}$$

and the matrices $[\mathbf{L}_{\mathcal{M}}(\mathbf{d},\mathbf{z})] \in \mathbb{R}^{n^{\mathrm{slip}} \times (d \cdot n^{\mathrm{m}})}$, $[\mathbf{L}_{\mathcal{S}}(\mathbf{d},\mathbf{z})] \in \mathbb{R}^{n^{\mathrm{slip}} \times (d \cdot n^{\mathrm{s}})}$ and $[\mathbf{H}(\mathbf{d})] \in \mathbb{R}^{n^{\mathrm{slip}} \times (d \cdot n^{\mathrm{slip}})}$ comprising the directional derivatives in $\{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{L}}$,

$$\mathop{\mathsf{A}}_{j\in\mathcal{L}} \{\mathscr{D}\widetilde{\boldsymbol{C}}^{\tau}_{j}\}_{\mathcal{L}} = [\mathbf{L}_{\mathcal{M}}]\{\Delta\mathbf{d}_{\mathcal{M}}\} + [\mathbf{L}_{\mathcal{S}}]\{\Delta\mathbf{d}_{\mathcal{S}}\} + [\mathbf{H}]\{\mathbf{z}^{\mathrm{k+1}}_{\mathcal{L}}\}\,. \tag{3.151}$$

3.11.1.4 Final equation system before Lagrange multipliers condensation

Having established all the algebraic representations, the final equation system to be solved at each semi-smooth Newton iteration can finally be assembled. However, in order to handle in a fully integrated fashion all equations (both in normal and tangential direction), one needs to further separate from the active set $\mathcal{A}$ the non-mortar nodes in the stick set $\mathcal{Q}$ and in the slip set $\mathcal{L}$ (keeping in mind that $\mathcal{A} = \mathcal{Q} \cup \mathcal{L}$). Thus, the global vector of displacement increments $\{\Delta\mathbf{d}\}$ is reorganized as

$$\{\Delta\mathbf{d}\} \equiv \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{N}} \\ \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{I}} \\ \Delta\mathbf{d}_{\mathcal{Q}} \\ \Delta\mathbf{d}_{\mathcal{L}} \end{Bmatrix} \tag{3.152}$$

and the vector of unknown Lagrange multipliers $\{\mathbf{z}^{\mathrm{k+1}}\}$ split as

$$\{\mathbf{z}^{\mathrm{k+1}}\} \equiv \begin{Bmatrix} \mathbf{z}_{\mathcal{I}}^{\mathrm{k+1}} \\ \mathbf{z}_{\mathcal{Q}}^{\mathrm{k+1}} \\ \mathbf{z}_{\mathcal{L}}^{\mathrm{k+1}} \end{Bmatrix} . \tag{3.153}$$

Moreover, taking into account that, for dual Lagrange multipliers, the first mortar coupling matrix $[\mathbf{D}]$ reduces to a diagonal matrix, it can be rearranged as follows

$$[\mathbf{D}] \equiv \begin{bmatrix} \mathbf{D}_{\mathcal{I}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{\mathcal{Q}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{\mathcal{L}} \end{bmatrix} . \tag{3.154}$$

The second mortar coupling matrix $[\mathbf{M}]$, on the other hand, yields

$$[\mathbf{M}] \equiv \begin{bmatrix} \mathbf{M}_{\mathcal{I}} \\ \mathbf{M}_{\mathcal{Q}} \\ \mathbf{M}_{\mathcal{L}} \end{bmatrix} . \tag{3.155}$$

Finally, the assembled system to be solved within each semi-smooth Newton step can be expressed as:

$$\begin{bmatrix}
\mathbf{K}_{\mathcal{N}\mathcal{N}} & \mathbf{K}_{\mathcal{N}\mathcal{M}} & \mathbf{K}_{\mathcal{N}\mathcal{I}} & \mathbf{K}_{\mathcal{N}\mathcal{Q}} & \mathbf{K}_{\mathcal{N}\mathcal{L}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{K}_{\mathcal{M}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{M}\mathcal{M}} & \widetilde{\mathbf{K}}_{\mathcal{M}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{M}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{M}\mathcal{L}} & -\mathbf{M}_{\mathcal{I}}^{\mathrm{T}} & -\mathbf{M}_{\mathcal{Q}}^{\mathrm{T}} & -\mathbf{M}_{\mathcal{L}}^{\mathrm{T}} \\
\mathbf{K}_{\mathcal{I}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{M}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{L}} & \mathbf{D}_{\mathcal{I}}^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\
\mathbf{K}_{\mathcal{Q}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{Q}\mathcal{M}} & \widetilde{\mathbf{K}}_{\mathcal{Q}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{Q}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{Q}\mathcal{L}} & \mathbf{0} & \mathbf{D}_{\mathcal{Q}}^{\mathrm{T}} & \mathbf{0} \\
\mathbf{K}_{\mathcal{L}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{L}\mathcal{M}} & \widetilde{\mathbf{K}}_{\mathcal{L}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{L}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{L}\mathcal{L}} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{\mathcal{L}}^{\mathrm{T}} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{\mathcal{I}} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{A}_{\mathcal{M}} & \mathbf{A}_{\mathcal{I}} & \mathbf{A}_{\mathcal{Q}} & \mathbf{A}_{\mathcal{L}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{Q}_{\mathcal{M}} & \mathbf{Q}_{\mathcal{I}} & \mathbf{Q}_{\mathcal{Q}} & \mathbf{Q}_{\mathcal{L}} & \mathbf{0} & \mathbf{T} & \mathbf{0} \\
\mathbf{0} & \mathbf{L}_{\mathcal{M}} & \mathbf{L}_{\mathcal{I}} & \mathbf{L}_{\mathcal{Q}} & \mathbf{L}_{\mathcal{L}} & \mathbf{0} & \mathbf{0} & \mathbf{H}
\end{bmatrix}
\begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{N}} \\ \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{I}} \\ \Delta\mathbf{d}_{\mathcal{Q}} \\ \Delta\mathbf{d}_{\mathcal{L}} \\ \mathbf{z}_{\mathcal{I}}^{\mathrm{k+1}} \\ \mathbf{z}_{\mathcal{Q}}^{\mathrm{k+1}} \\ \mathbf{z}_{\mathcal{L}}^{\mathrm{k+1}} \end{Bmatrix}
= -\begin{Bmatrix} \mathbf{r}_{\mathcal{N}} \\ \widetilde{\mathbf{r}}_{\mathcal{M}} \\ \widetilde{\mathbf{r}}_{\mathcal{I}} \\ \widetilde{\mathbf{r}}_{\mathcal{Q}} \\ \widetilde{\mathbf{r}}_{\mathcal{L}} \\ \mathbf{0} \\ \widetilde{\mathbf{g}} \\ \mathbf{r}^{\mathrm{stick}} \\ \mathbf{r}^{\mathrm{slip}} \end{Bmatrix} . \tag{3.156}$$

Herein, the current iteration index k has been dropped, once again for ease of notation. Examining the linear system (3.156), one can recognise the first five rows as the linearised form of the discrete force equilibrium in Equation (3.72). The sixth row represents the trivial contact constraints for non-mortar nodes of the inactive set $\mathcal{I}$, with $\mathbf{I}_{\mathcal{I}} \in \mathbb{R}^{d\cdot(n^{\mathrm{s}}-n^{\mathrm{a}})\times d\cdot(n^{\mathrm{s}}-n^{\mathrm{a}})}$ as an identity matrix. With regard to the active set $\mathcal{A}$, the contact constraints in the normal direction are enforced in the seventh row, while stick and slip tangential contact constraints are imposed in rows eight and nine, accordingly.

It should be noted that, although the algebraic representation in Equation (3.156) consists of nine rows and eight columns, the system block matrix still reassembles a square matrix structure. This is due to the fact that the normal and tangential contact constraints are enforced independently in the last three rows, while using a vector-valued Lagrange multiplier. When compared with standard FE problems without contact, the final global equation system displays an increased size (recall that the Lagrange multipliers are introduced as unknowns) with the typical saddle point structure easily recognised. Nonetheless, as explained in the following, the use of dual Lagrange multipliers allows to simplify the equation system further.

3.11.2 Elimination of the Lagrange multiplier

As already mentioned, the dual Lagrange multiplier shape functions are constructed such that the interface contact coupling problem is reduced to a *localized* form. Algebraically, this is characterized by the first mortar coupling matrix $[\mathbf{D}]$ becoming a diagonal matrix. This also means that its inversion becomes trivial, thus allowing the condensation of the discrete Lagrange multipliers in the linear system (3.156) to be performed in an efficient manner. Ultimately, the undesirable saddle point structure can be conveniently removed while maintaining the original saddle point formulation.

First of all, note that by considering the separation between the mortar set $\mathcal{M}$, the non-mortar set $\mathcal{S}$ and the set $\mathcal{N}$ comprising the remaining degrees of freedom, based on Equation (3.129), the discrete Lagrange multipliers vector $\{\mathbf{z}\}^{k+1}$ can be obtained from the following general expression:

$$\{\mathbf{z}\}^{k+1} = -[\mathbf{D}]^{-T}\Big(\{\tilde{\mathbf{r}}_{\mathcal{S}}\} + [\mathbf{K}_{\mathcal{S}\mathcal{N}}]\{\Delta\mathbf{d}_{\mathcal{N}}\} + [\tilde{\mathbf{K}}_{\mathcal{S}\mathcal{M}}]\{\Delta\mathbf{d}_{\mathcal{M}}\} + [\tilde{\mathbf{K}}_{\mathcal{S}\mathcal{S}}]\{\Delta\mathbf{d}_{\mathcal{S}}\}\Big)\,. \tag{3.157}$$

Alternatively, and following the nodal ordering scheme used in Equation (3.156), the Lagrange multipliers $\{\mathbf{z}_{\mathcal{I}}\}^{k+1}$ associated with inactive non-mortar nodes can easily be condensed by extracting the identity $\{\mathbf{z}_{\mathcal{I}}\}^{k+1} = \mathbf{0}$ from the sixth row in Equation (3.156); this means that the sixth row and column can be removed. Then, based on the fourth and fifth rows, one can write that

$$\begin{aligned}\{\mathbf{z}_{\mathcal{Q}}\}^{k+1} = -[\mathbf{D}_{\mathcal{Q}}]^{-T}\Big(&\{\tilde{\mathbf{r}}_{\mathcal{Q}}\} + [\mathbf{K}_{\mathcal{Q}\mathcal{N}}]\{\Delta\mathbf{d}_{\mathcal{N}}\} + [\tilde{\mathbf{K}}_{\mathcal{Q}\mathcal{M}}]\{\Delta\mathbf{d}_{\mathcal{M}}\}\\ &+ [\tilde{\mathbf{K}}_{\mathcal{Q}\mathcal{I}}]\{\Delta\mathbf{d}_{\mathcal{I}}\} + [\tilde{\mathbf{K}}_{\mathcal{Q}\mathcal{Q}}]\{\Delta\mathbf{d}_{\mathcal{Q}}\} + [\tilde{\mathbf{K}}_{\mathcal{Q}\mathcal{L}}]\{\Delta\mathbf{d}_{\mathcal{L}}\}\Big)\,,\end{aligned} \tag{3.158}$$

$$\begin{aligned}\{\mathbf{z}_{\mathcal{L}}\}^{k+1} = -[\mathbf{D}_{\mathcal{L}}]^{-T}\Big(&\{\tilde{\mathbf{r}}_{\mathcal{L}}\} + [\mathbf{K}_{\mathcal{L}\mathcal{N}}]\{\Delta\mathbf{d}_{\mathcal{N}}\} + [\tilde{\mathbf{K}}_{\mathcal{L}\mathcal{M}}]\{\Delta\mathbf{d}_{\mathcal{M}}\}\\ &+ [\tilde{\mathbf{K}}_{\mathcal{L}\mathcal{I}}]\{\Delta\mathbf{d}_{\mathcal{I}}\} + [\tilde{\mathbf{K}}_{\mathcal{L}\mathcal{Q}}]\{\Delta\mathbf{d}_{\mathcal{L}}\} + [\tilde{\mathbf{K}}_{\mathcal{L}\mathcal{L}}]\{\Delta\mathbf{d}_{\mathcal{L}}\}\Big)\,.\end{aligned} \tag{3.159}$$

Examining the definitions above, one can easily conclude that Equations (3.158) and (3.159) have the same structure as Equation (3.157), with the only difference being the involved nodal subsets. Finally, the substitution of the expressions above into Equation (3.156) leads to the final condensed system:

$$\begin{bmatrix} \mathbf{K}_{\mathcal{N}\mathcal{N}} & \mathbf{K}_{\mathcal{N}\mathcal{M}} & \mathbf{K}_{\mathcal{N}\mathcal{I}} & \mathbf{K}_{\mathcal{N}\mathcal{Q}} & \mathbf{K}_{\mathcal{N}\mathcal{L}} \\ \breve{\mathbf{K}}_{\mathcal{N}} & \breve{\mathbf{K}}_{\mathcal{M}} & \breve{\mathbf{K}}_{\mathcal{I}} & \breve{\mathbf{K}}_{\mathcal{Q}} & \breve{\mathbf{K}}_{\mathcal{L}} \\ \mathbf{K}_{\mathcal{I}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{M}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{L}} \\ \mathbf{0} & \mathbf{A}_{\mathcal{M}} & \mathbf{A}_{\mathcal{I}} & \mathbf{A}_{\mathcal{Q}} & \mathbf{A}_{\mathcal{L}} \\ \breve{\mathbf{Q}}_{\mathcal{N}} & \breve{\mathbf{Q}}_{\mathcal{M}} & \breve{\mathbf{Q}}_{\mathcal{I}} & \breve{\mathbf{Q}}_{\mathcal{Q}} & \breve{\mathbf{Q}}_{\mathcal{L}} \\ \breve{\mathbf{L}}_{\mathcal{N}} & \breve{\mathbf{L}}_{\mathcal{M}} & \breve{\mathbf{L}}_{\mathcal{I}} & \breve{\mathbf{L}}_{\mathcal{Q}} & \breve{\mathbf{L}}_{\mathcal{L}} \end{bmatrix} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{N}} \\ \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{I}} \\ \Delta\mathbf{d}_{\mathcal{Q}} \\ \Delta\mathbf{d}_{\mathcal{L}} \end{Bmatrix} = - \begin{Bmatrix} \mathbf{r}_{\mathcal{N}} \\ \breve{\mathbf{r}}_{\mathcal{M}} \\ \widetilde{\mathbf{r}}_{\mathcal{I}} \\ \widetilde{\mathbf{g}} \\ \breve{\mathbf{r}}^{\text{stick}} \\ \breve{\mathbf{r}}^{\text{slip}} \end{Bmatrix} . \tag{3.160}$$

Here, several algebraic functions and abbreviations have been introduced to facilitate the notation. First, the algebraic function $\left[\breve{\mathbf{K}}_{(\bullet)}\right]$ appearing in the second row is defined as

$$\left[\breve{\mathbf{K}}_{(\bullet)}\right] \equiv \begin{cases} \left[\mathbf{K}_{\mathcal{M}(\bullet)}\right] + \left[\mathbf{J}_{\mathcal{Q}}\right]^{\mathrm{T}}\left[\mathbf{K}_{\mathcal{Q}(\bullet)}\right] + \left[\mathbf{J}_{\mathcal{L}}\right]^{\mathrm{T}}\left[\mathbf{K}_{\mathcal{L}(\bullet)}\right], & \text{if } (\bullet) \in \mathcal{N}, \\ \left[\widetilde{\mathbf{K}}_{\mathcal{M}(\bullet)}\right] + \left[\mathbf{J}_{\mathcal{Q}}\right]^{\mathrm{T}}\left[\widetilde{\mathbf{K}}_{\mathcal{Q}(\bullet)}\right] + \left[\mathbf{J}_{\mathcal{L}}\right]^{\mathrm{T}}\left[\widetilde{\mathbf{K}}_{\mathcal{L}(\bullet)}\right], & \text{if } (\bullet) \notin \mathcal{N}, \end{cases} \tag{3.161}$$

where the stick and slip parts of the so-called *mortar projection operator*, denoted as $\left[\mathbf{J}_{\mathcal{Q}}\right] \in \mathbb{R}^{(d \cdot n^{\text{stick}}) \times (d \cdot n^{\text{m}})}$ and $\left[\mathbf{J}_{\mathcal{L}}\right] \in \mathbb{R}^{(d \cdot n^{\text{slip}}) \times (d \cdot n^{\text{m}})}$, have been introduced:

$$\left[\mathbf{J}_{\mathcal{Q}}\right] \equiv \left[\mathbf{D}_{\mathcal{Q}}\right]^{-1}\left[\mathbf{M}_{\mathcal{Q}}\right]; \qquad \left[\mathbf{J}_{\mathcal{L}}\right] \equiv \left[\mathbf{D}_{\mathcal{L}}\right]^{-1}\left[\mathbf{M}_{\mathcal{L}}\right]. \tag{3.162}$$

The residual $\left\{\breve{\mathbf{r}}_{\mathcal{M}}\right\} \in \mathbb{R}^{d \cdot n^{\text{m}}}$ denotes the abbreviation

$$\left\{\breve{\mathbf{r}}_{\mathcal{M}}\right\} \equiv \left\{\widetilde{\mathbf{r}}_{\mathcal{M}}\right\} + \left[\mathbf{J}_{\mathcal{Q}}\right]^{\mathrm{T}}\left\{\widetilde{\mathbf{r}}_{\mathcal{Q}}\right\} + \left[\mathbf{J}_{\mathcal{L}}\right]^{\mathrm{T}}\left\{\widetilde{\mathbf{r}}_{\mathcal{L}}\right\}. \tag{3.163}$$

Focusing now on the fifth row of Equation (3.160), the function algebraic function $\left[\breve{\mathbf{Q}}_{(\bullet)}\right]$ is defined as

$$\left[\breve{\mathbf{Q}}_{(\bullet)}\right] \equiv \begin{cases} -\left[\mathbf{T}\right]\left[\mathbf{D}_{\mathcal{Q}}\right]^{-\mathrm{T}}\left[\mathbf{K}_{\mathcal{Q}(\bullet)}\right], & \text{if } (\bullet) \in \mathcal{N}, \\ \left[\mathbf{Q}_{(\bullet)}\right] - \left[\mathbf{T}\right]\left[\mathbf{D}_{\mathcal{Q}}\right]^{-\mathrm{T}}\left[\widetilde{\mathbf{K}}_{\mathcal{Q}(\bullet)}\right], & \text{if } (\bullet) \notin \mathcal{N}, \end{cases} \tag{3.164}$$

and the residual $\left\{\breve{\mathbf{r}}^{\text{stick}}\right\} \in \mathbb{R}^{n^{\text{stick}}}$ yields

$$\left\{\breve{\mathbf{r}}^{\text{stick}}\right\} \equiv \left\{\mathbf{r}^{\text{stick}}\right\} - \left[\mathbf{T}\right]\left[\mathbf{D}_{\mathcal{Q}}\right]^{-\mathrm{T}}\left\{\widetilde{\mathbf{r}}_{\mathcal{Q}}\right\}. \tag{3.165}$$

At last, in the sixth row of the system (3.160), the function $\left[\breve{\mathbf{L}}_{(\bullet)}\right]$ is constructed as

$$\left[\breve{\mathbf{L}}_{(\bullet)}\right] \equiv \begin{cases} -\left[\mathbf{H}\right]\left[\mathbf{D}_{\mathcal{L}}\right]^{-\mathrm{T}}\left[\mathbf{K}_{\mathcal{L}(\bullet)}\right], & \text{if } (\bullet) \in \mathcal{N}, \\ \left[\mathbf{L}_{(\bullet)}\right] - \left[\mathbf{H}\right]\left[\mathbf{D}_{\mathcal{L}}\right]^{-\mathrm{T}}\left[\widetilde{\mathbf{K}}_{\mathcal{L}(\bullet)}\right], & \text{if } (\bullet) \notin \mathcal{N}, \end{cases} \tag{3.166}$$

and the residual $\left\{\breve{\mathbf{r}}^{\text{slip}}\right\} \in \mathbb{R}^{n^{\text{slip}}}$ as

$$\left\{\breve{\mathbf{r}}^{\text{slip}}\right\} \equiv \left\{\mathbf{r}^{\text{slip}}\right\} - \left[\mathbf{H}\right]\left[\mathbf{D}_{\mathcal{L}}\right]^{-\mathrm{T}}\left\{\widetilde{\mathbf{r}}_{\mathcal{L}}\right\}. \tag{3.167}$$

It should be noted that by using this approach, the discrete Lagrange multipliers are evaluated by using Equations (3.158) and (3.159) as a post-processing step at the end of each iterative Newton step—within mortar formulations, in a variationally consistent way. Given its physical interpretation as contact stresses, this stage of the algorithm is critical to several contact algorithms extensions, such as wear modelling or termo-mechanical problems; e.g., refer to Wilking, Bischoff, and Ramm (2018), where contact stress recovery for dual mortar formulations is analysed.

As a final remark, a recapitulation of some numerical properties of the condensed system in Equation (3.160) is given. As already mentioned, the mortar-typical generalised saddle point structure is algebraically removed and the global system size is reduced to the number of primary (i.e., the displacement field) degrees of freedom. While being beyond the scope of the current work, this allows the application of state-of-the-art iterative solvers with algebraic multigrid preconditioners for efficiently solving large contact problems. For further information on this topic, the interested reader is referred to B. I. Wohlmuth and Krause (2003), Brunssen *et al.* (2007), and Wiesner *et al.* (2018) and the references therein.

Chapter 4

Computational Implementation Approach and Numerical Validation

The contact algorithm has been implemented within an in-house Fortran program named LINKS (Large Strain Implicit Non-linear Analysis of Solids Linking Scales, see Figure 4.1). It is a multi-scale finite element code for implicit small and large strain analysis of hyperelastic and elasto-plastic solids, developed jointly by CM2S (Computational Multi-Scale Modelling of Solids and Structures) at the Faculty of Engineering of the University of Porto. Because the contact implemented has been entirely made from scratch, the general design approach has been carefully analysed.

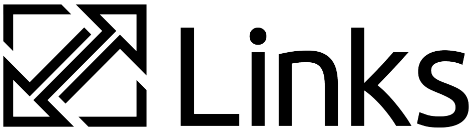


Figure 4.1: LINKS logo.

The first section of this chapter describes the devised implementation of the algorithm, emphasizing the adopted software design approach and concrete solutions to mortar-based finite element algorithms. Then, in order to evaluate and validate the mortar algorithm implementation for finite deformation contact, several numerical examples are presented and analysed.

4.1 Object-oriented approach

The main objective is to develop a robust framework suitable to various contact algorithms and flexible enough to facilitate the future development of multi-scale contact formulations. In order to achieve this goal, the code architecture has been throughout designed by means of *Object-Oriented* (OO) software techniques since they inherently promote code modularity and data-hiding—both critical concepts for flexible software design. Naturally, this choice is only possible because Fortran 2003 and subsequent standard levels provide OO language constructs, thus allowing to fully exploit the concepts of OO programming (Metcalf *et al.*, 2011). In the following paragraphs, a brief introduction to the general principles of OO programming is presented. Then, the most important classes and abstractions created to implement the present dual-mortar contact framework are explained.

4.1.1 General principles of Object-Oriented programming

The object-oriented paradigm is a software development strategy based on the concept that systems should be built from a collection of reusable components called *objects*. While in purely structured programming (the base concept of procedural programming, the classical approach in Fortran), data and functionality are treated separately, objects encompass both in a self-contained manner: the data is contained in the form of *attributes* and the procedures in the form of *methods*. This means that instead of having data structures passing around to all the program logic that acts on it, a OO program consists of a set of interacting objects, which combine data and logic altogether.

As mentioned, central to OO programming lies the goal of code reuse. Usually, it takes the form of *inheritance*. Authors frequently refer to inheritance as an "is a" relationship, in the sense that one object is a specific sub-category of another object. The lower-level objects *inherit* both the instance variables and methods of the object from which it is derived while still being able to add additional attributes or methods. This allows programmers to create objects that are built upon existing objects and, ultimately, reuse code.

Another fundamental concept is that of *polymorphism*, i.e., the ability to present the same interface for different objects. While each object provides its own implementation of the interface, different objects can still be treated as generic ones. Ultimately, this allows having code working with multiple objects without needing to know its specific type. In practice, a polymorphic variable is a variable whose data type is dynamic at runtime, thus eliminating the need to use conditional statements to select the proper methods.

The benefits of the OO paradigm can be further exploited if *encapsulation* is employed, i.e., if the implementation details of an object are hidden from the other objects in the program. Also commonly referred to as *data hiding*, it implies that objects can only interact with each other by invoking their methods. Encapsulation provides two primary benefits to software developers. First, it provides code modularity since an object can be written and maintained independently from the remaining objects in the system. Additionally, while the public interface is not changed, an object's variables and methods can be changed whenever needed without introducing side effects in the other objects that depend on it.

The last topic to be addressed in this short discussion about the general concepts of OO programming is *design patterns*. The effort put into thinking abstractly about software structure ultimately leads towards high-level designs that prove useful and independent of the application and implementation language. This motivated software engineers to create the so-called design patterns, reusable designs formalized as best practices used by experienced OO software developers to solve common occurring problems. It gained popularity with the book *Design Patterns: Elements of Reusable Object-Oriented Software*, which classifies design patterns into creational, structural and behavioural patterns (Gamma *et al.*, 1995). Whenever convenient, patterns are employed in the present implementation in order to take full advantage of the capabilities of OO programming.

4.2 Code structure

Another essential feature of OO programming is that it allows the use of *Unified Modeling Language* (UML), an abstract schematic toolset for describing object-oriented software independently from any chosen implementation language (Rumbaugh *et al.*, 1999). The following paragraphs describe the most important classes of the present contact algorithm implementation by using *class diagrams*. It is the most common UML diagram, in which class attributes and methods are represented together with interactions and relationships among different classes. For additional information on UML notation, the reader is referred to the vast literature available (Rumbaugh *et al.*, 1999; Rouson *et al.*, 2011).

4.2.1 Contact facade

The main class of contact implementation is the contact package. Following the facade design pattern guidelines, the contact main class provides a unified higher-level interface that makes the contact algorithm sub-system easier to use. All the interactions with the remaining code are established using this class, which contains all the geometrical information, main variables (e.g., mortar matrices or contact constraints) and operations needed within the algorithm (e.g., contact search). Figure 4.2 shows the class diagram.

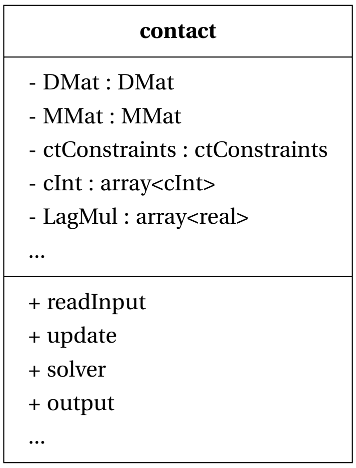


Figure 4.2: UML diagram for the main contact class.

This class contains the variables needed to describe the contact interactions with the mortar method, e.g., the mortar coupling matrices and the contact constraints. As described in the following paragraphs, the entire geometrical information is contained within a list of candidate interfaces. Moreover, all the global nodal variables are stored, e.g., the nodal Lagrange multipliers. In order to maintain encapsulation, all attributes are set as private, leaving only the main operations open to the rest of the code. This includes the information reading, output and the two main procedures: the update, which is needed after every update in the body configuration, and the solver procedure, which is responsible for evaluating all the derivatives and assembly of the global equation system.

4.2.2 Candidate interface

One of the main challenges is how to define the potential contact boundaries. The objective is to derive a class entirely capable of containing the geometrical information and operations needed within the algorithm, e.g., numerical integration of mortar integrals. Figure 4.3 represents the class diagram of the fundamental concepts and classes and is described as follows.

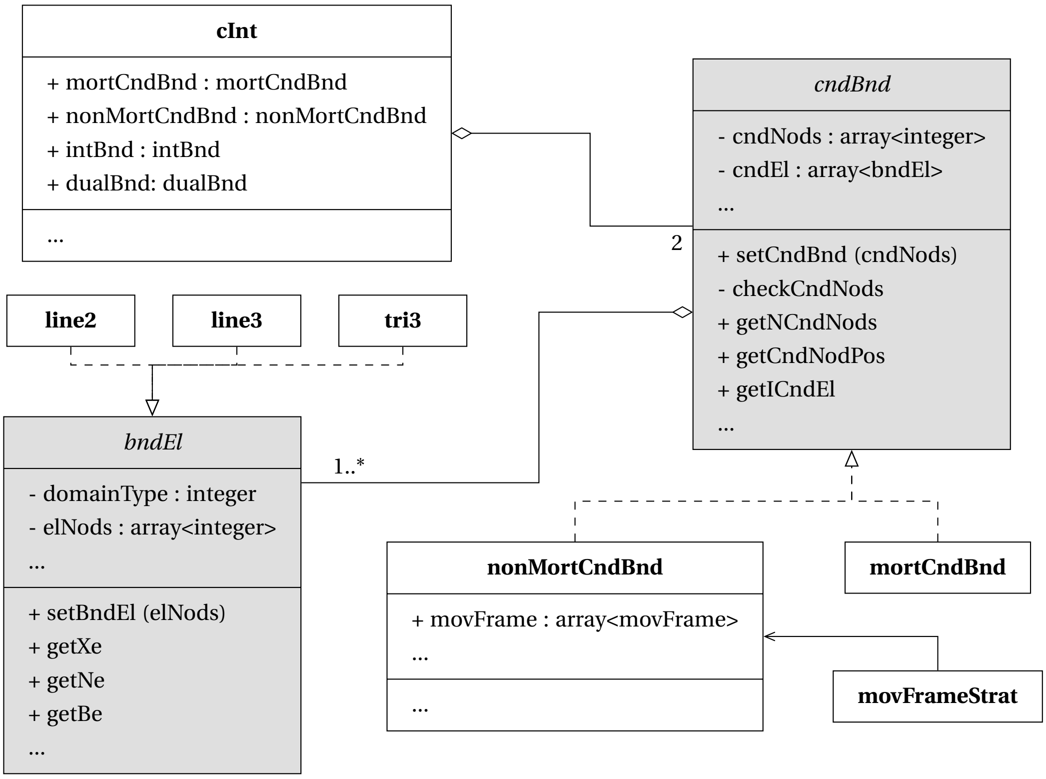


Figure 4.3: UML diagram for the contact interface class.

First, the class *contact interface* (`cInt`) is introduced, which is defined as the aggregation of the non-mortar and mortar candidate boundaries, i.e., a collection of nodes/facets that, altogether, form the potential contact zone. Considering the fact that either the non-mortar and mortar boundaries are formed by candidate elements, they are defined as extensions of the abstract class *candidate boundary* (`cndBnd`)—thus taking advantage of code re-usability in inheritance constructs. This abstract class is initialised based on a list of candidate nodes specified by the user, from which the associated collection of finite elements can be extracted from the finite element mesh. This requires the definition of another fundamental object: that of *boundary elements* (`bndEl`). Besides other auxiliary variables (such as type of domain, number of spatial dimensions, etc.), a boundary element is essentially defined as a list of nodal connectivities and, to exploit once more code re-use, it is defined as an abstract class, with concrete implementations being specific types of finite elements. It is important to note that this strategy makes it possible to set the boundary element independently from the global mesh.

Lastly, because the mortar integration scheme requires projections of Gauss points between the non-mortar and mortar boundaries, the *non-mortar candidate boundary* (`nonMortCndBnd`) further contains the list of the nodal moving frame. The strategy pattern is adopted and concrete algorithms are encapsulated within the interface *moving frame strategy* (`movFrameStrat`) to provide flexibility in implementing different methods of defining the averaged normal vector

4.2.3 Integration boundary

With the potential contact interface defined, attention is now focused on the abstractions needed to implement the mortar integration algorithm. As shown in Figure 4.4, the class *integration boundary* (`intBnd`) is defined as the collection of integration segments together with a quadrature rule. Its initialisation procedure takes as input the candidate boundaries and the list of contact element pairs, which is to be obtained by the contact search algorithm (for the sake of brevity, not discussed here).

Following an approach similar to the one proposed in Badia *et al.* (2018), the numerical quadrature is based on the abstract class *quadrature* (`quadrature`), which can be interpreted as a place-holder for quadrature points and weights. Concrete numerical quadratures, such as Gauss-Legendre or Gauss-Lobatto rules, must implement this interface. The main goal here is to exploit polymorphism and treat different domain types dynamically at runtime.

In the segmentation strategy described in Chapter 3 to evaluate mortar integrals, the fundamental entities are two- and three-dimensional integration segments. Therefore, they are both set as implementations of the abstract class *integration segment* (`intSeg`), where information like the type of domain is shared in both two and three dimensional problems. The interfaces for the deferred methods encompass operations such as the mapping from integration segment space to boundary element finite element space or the corresponding Jacobian determinant, together with the corresponding directional derivatives. Once again, this allows the use of polymorphism and reuse the same code for mortar integrals evaluation in two- and three-dimensional problems.

Lastly, note that either type of integration segment is composed of a list of nodes. This motivates the introduction of the abstract class *integration segment node* (`intSegNod`), which, besides some common properties (e.g., knowing if the node is projected or not), is defined depending on the type of problem. While in two-dimensions, an integration segment node is defined by its boundary element isoparametric coordinate, for three-dimensional integration cells, the nodes are defined by their global coordinates.

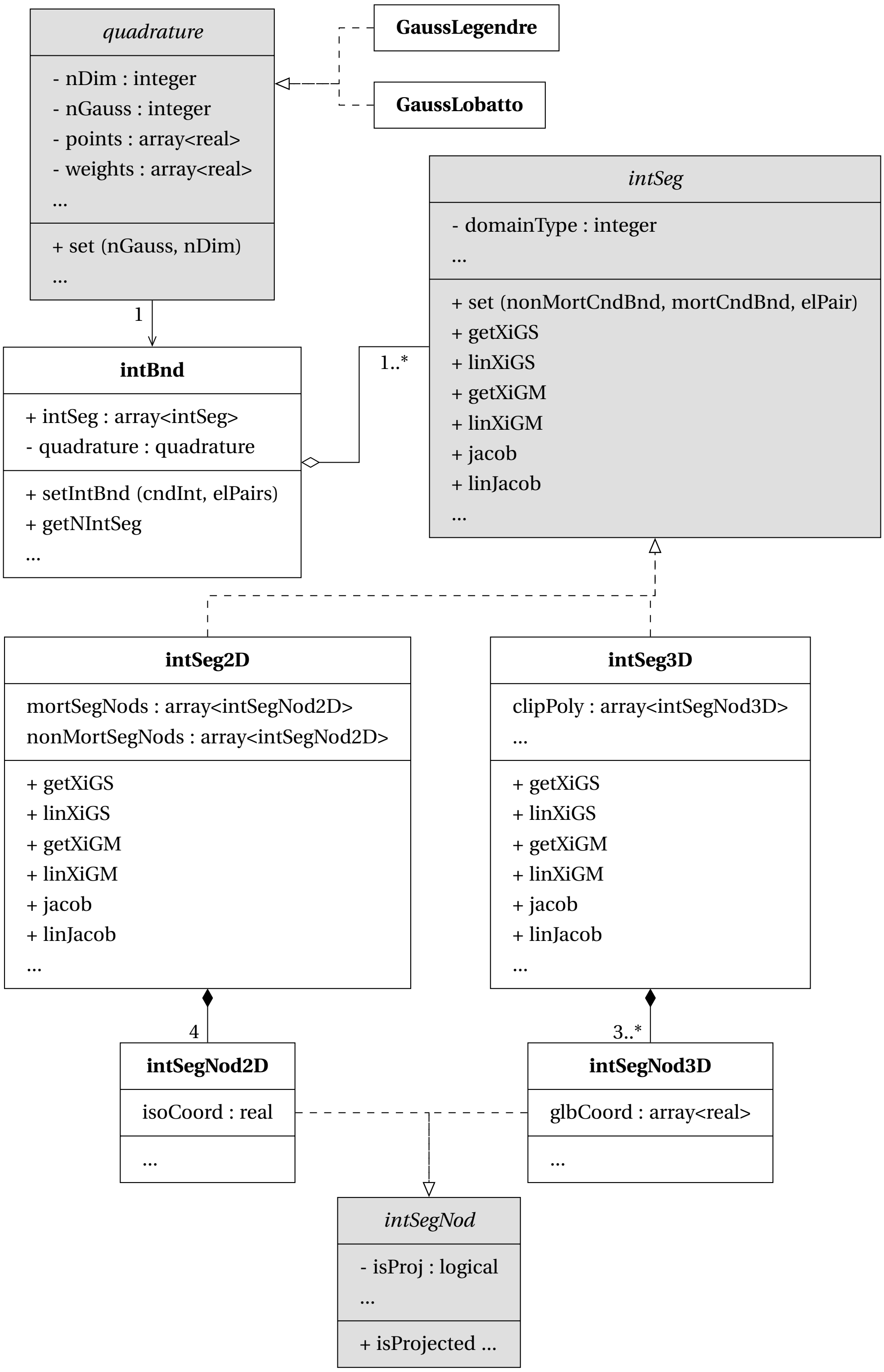


Figure 4.4: UML diagram for the integration boundary class.

4.2.4 Dual boundary

The last fundamental abstraction to be discussed regards the dual basis. As represented in Figure 4.5, the class *dual boundary* (`dualBnd`) is introduced as the collection of dual elements, used to interpolate Lagrange multipliers within the integration boundary, and a quadrature rule, necessary to evaluate dual shape functions numerically. Its initialisation takes the candidate boundaries, the list of contact element pairs and the integration boundary. It should be noted that the Gauss quadrature used to evaluate the dual basis does not need to be the same used to evaluate mortar integrals.

As described in Chapter 3, dual shape functions are defined as a linear combination of standard shape functions. Thus, the object *dual element* (`dualEl`) can be simply defined as a coefficient matrix attached to a non-mortar boundary element, to be numerically determined at each global iteration. However, in order to deal with partially integrated elements, additional information must be stored. Namely, a flag identifying partially integrated elements and the list of integration segments related to that boundary element. Lastly, besides being able to determine the dual shape functions, the dual element must also evaluate their directional derivatives.

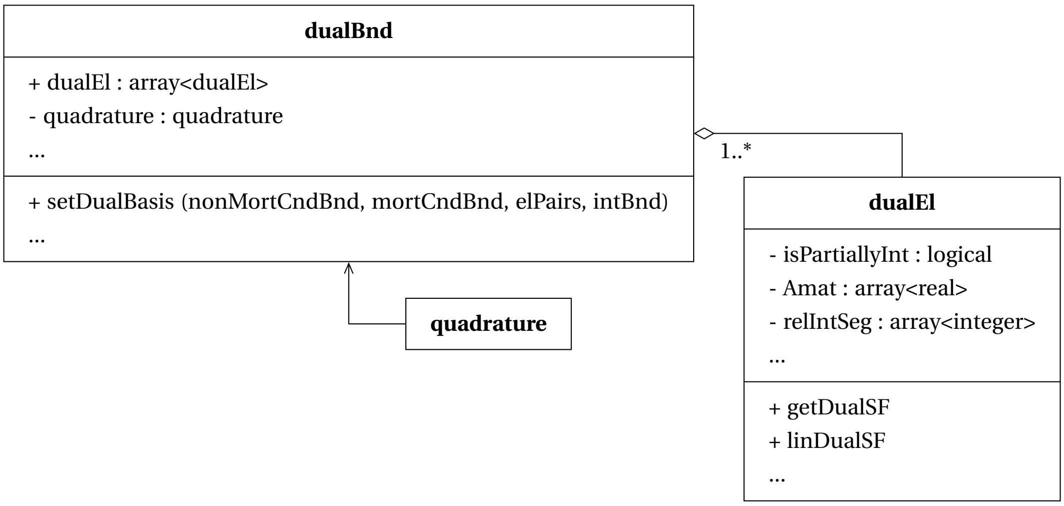


Figure 4.5: UML diagram for the dual boundary class.

4.3 Numerical examples

In what follows, various numerical examples are presented and analysed, to evaluate the mortar algorithm implementation. The first example is the contact patch test, whose main objective is to check for consistency for non-matching discretisations at the contact interface. Then, attention is given to the classical Hertzian contact problem of a cylinder being pressed against a block. The results are compared with both the analytical solution and other commercial finite element packages. Optimal spatial convergence rates are verified, as well as the numerical efficiency of the semi-smooth Newton algorithm. The following example is dedicated to the large deformation problems involving plasticity—an additional source of non-linearity apart from the enforcement of contact constraints. It consists of the contact problem of two deformable rings, assuming for both solids plasticity under large deformations. The last two examples focus on frictional contact problems. The shallow ironing problem is analysed, comparing results with other algorithms proposed in the literature. Lastly, the three-dimensional compression of a cylindrical plastic tube is simulated in a 5 stage non-monotonic loading configuration.

If not stated otherwise, first-order finite elements combined with dual Lagrange multiplier interpolation are used. The normal complementarity parameter is set to $\mathrm{c}^{\eta} = E$, and a total of $n_{\mathrm{g}} = 9$ Gauss points for numerical evaluation of mortal integrals is considered.

4.3.1 Patch test

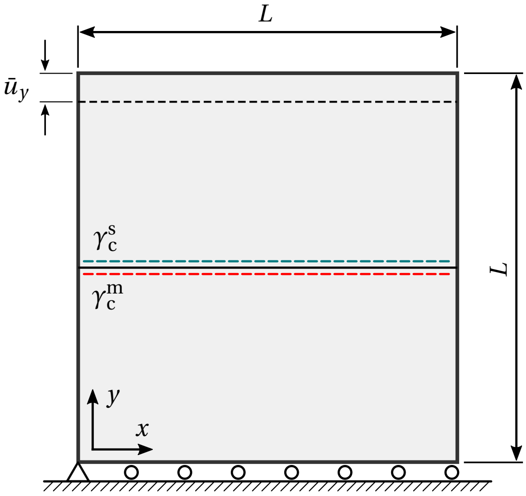
Originally introduced in Bazeley *et al.* (1966), patch tests are one of the most common validation techniques within finite element technology. It checks the robustness of a finite element form by ascertaining if an arbitrary *patch* of assembled elements subjected to boundary displacements, consistent with constant straining, can reproduce exactly the behaviour of an elastic solid material. Satisfaction of the patch test constitutes a necessary condition for algorithmic convergence while assessing the asymptotic convergence rate (Taylor, Simo, *et al.*, 1986). In the context of contact problems, in particular, patch tests can be used to assess the capability of a contact formulation to exactly transmit a *continuous* field of normal stresses between two contacting surfaces, regardless of their discretisation. If the algorithm fails the patch test, it may lead to solution errors at the contacting surfaces, which do not necessarily decrease with mesh refinement. As pointed out in the early work of Taylor and Papadopoulos (1991) and further analysed in Crisfield (2000), collocation methods such as the classical NTS discretisation schemes have difficulties passing the contact patch test.[*] In contrast, mortar-based finite element methods guarantee the exact satisfaction of patch tests by design due to their variationally consistent interpolation of the contact tractions via discrete Lagrange multipliers.

The considered contact patch test is depicted in Figure 4.6 and is inspired by the patch test proposed in Crisfield (2000). The problem consists of two blocks (with non-conforming meshes) making up a square of size $L = 1$ m, interacting with each other at a plane contact interface. The lower block is supported at its bottom boundary and a uniform displacement $\bar{u}_y = 0.01$ m is applied at the top surface of the upper block. The

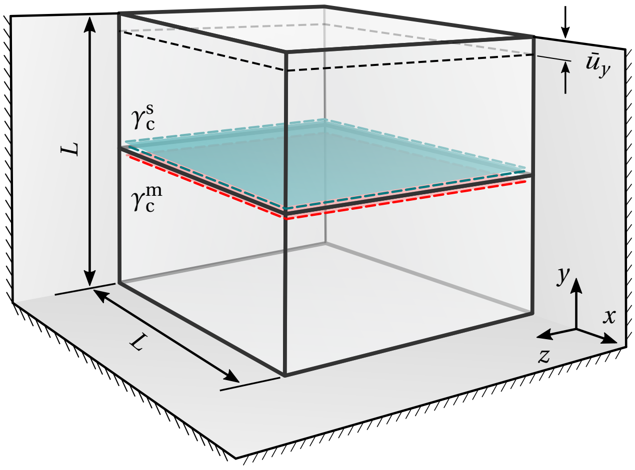
---

[*] Nonetheless, it should be mentioned that modified versions of the algorithm aiming at guaranteeing the patch test satisfaction have been proposed, referring to Zavarise and De Lorenzis (2009) for a successful implementation of such algorithms.

material is considered to be linear elastic with Young's modulus $E = 10\,\text{GPa}$ and Poisson's ratio $\nu = 0.35$. The top contact boundary is chosen as non-mortar and the bottom surface as mortar. The obtained results with both first- and second-order mortar finite element interpolation follow in Figures 4.7 and 4.8. As expected, the algorithm is capable of exactly (up to machine precision) transmitting a constant contact pressure along with the interface, thus leading to a linear displacement field, see Figure 4.7, and constant stress field, see Figure 4.8. Finally, it is worth mentioning that the exact patch test compliance is not influenced by choice of non-mortar and mortar sides and, ultimately, both lead to the same results.

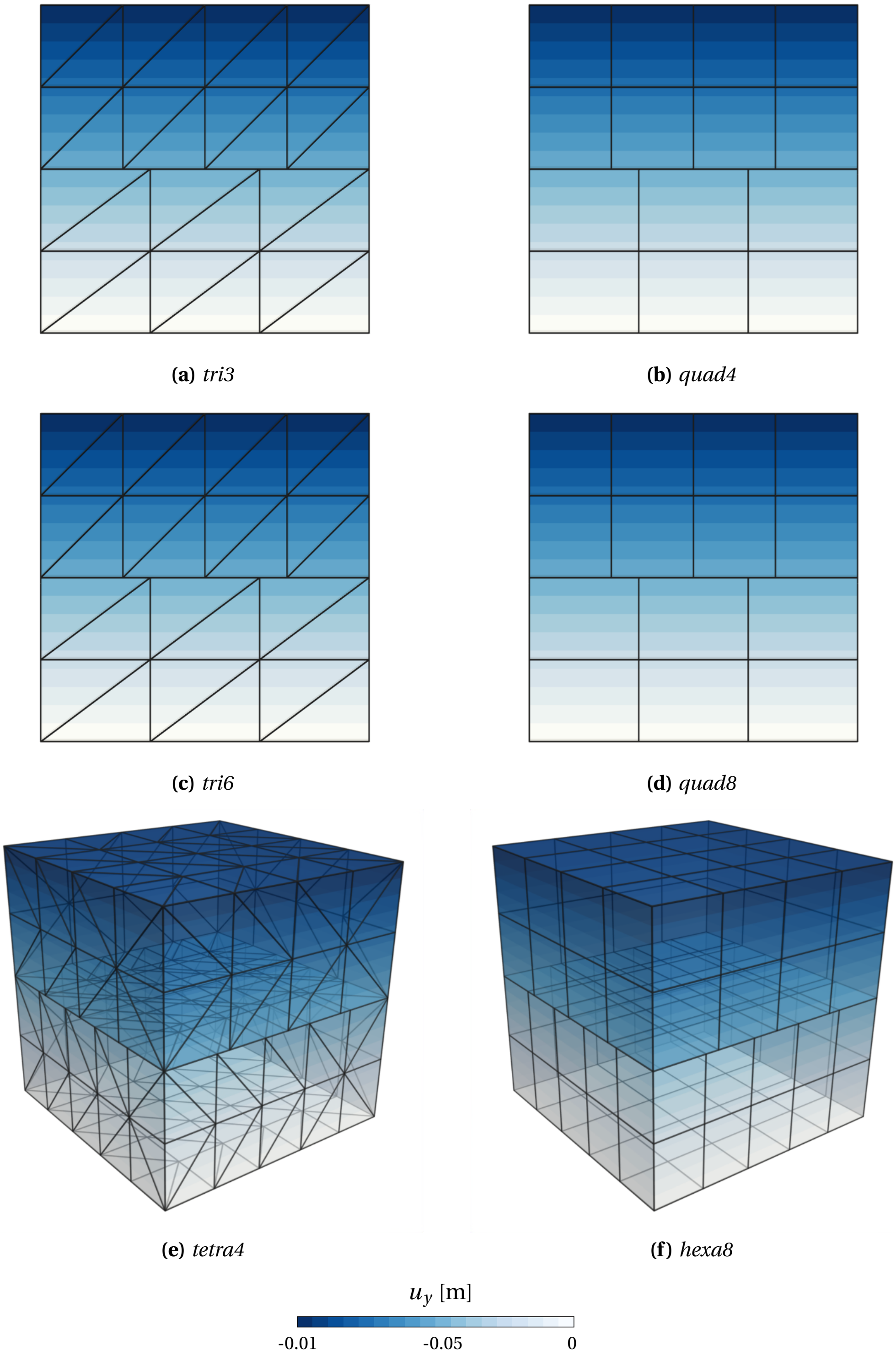


(a) Two-dimensional patch test.



(b) Three-dimensional patch test.

Figure 4.6: Contact patch test – schematic representation of the problem setting.

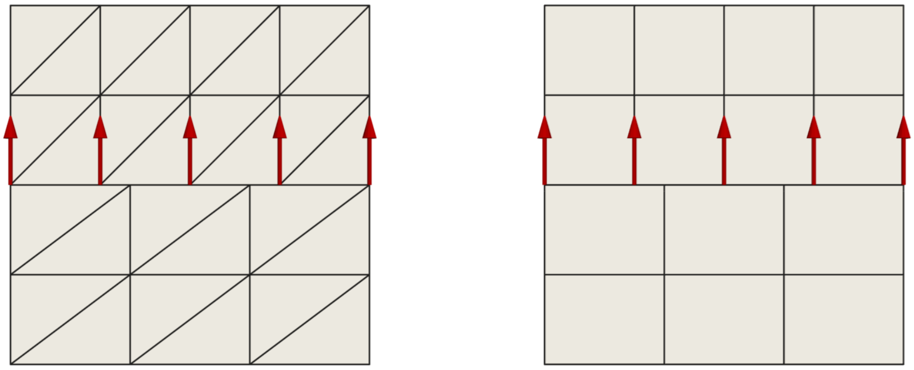


(a) *tri3* **(b)** *quad4*

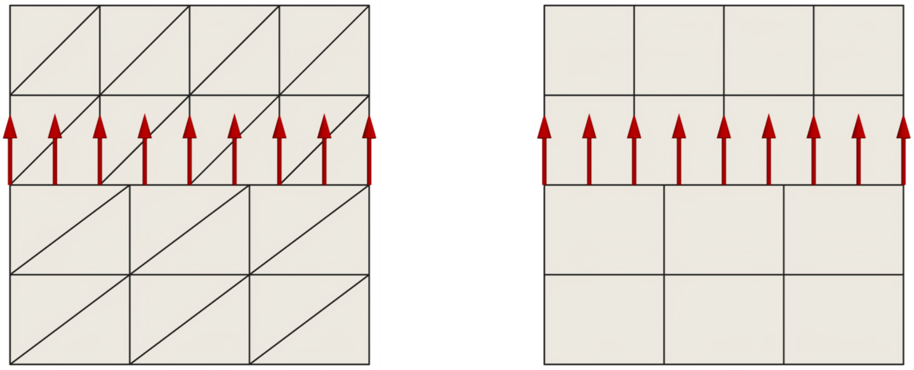
(c) *tri6* **(d)** *quad8*

(e) *tetra4* **(f)** *hexa8*

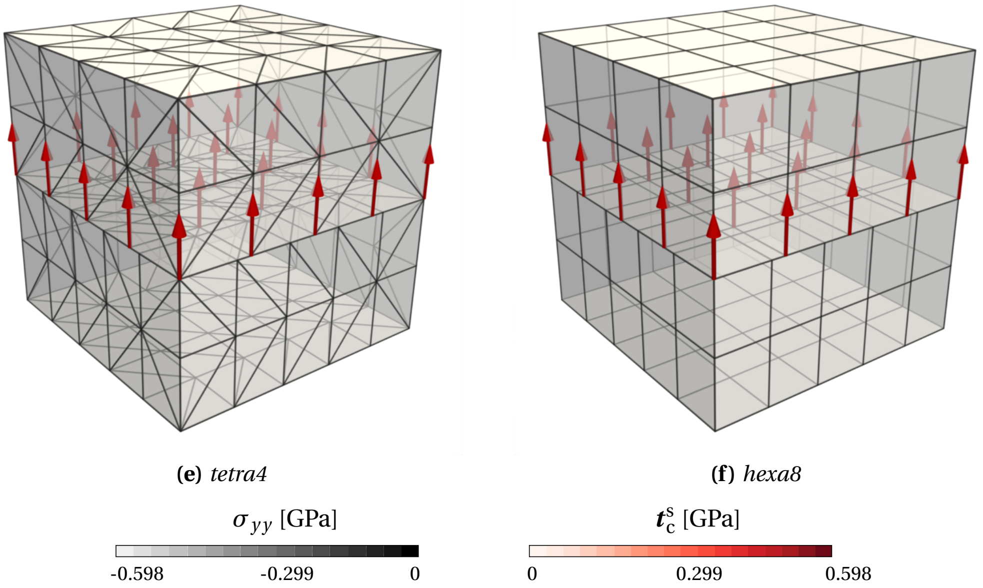
Figure 4.7: Contact patch test – displacement along y-direction for different types of first- and second-order mortar finite element interpolation.



(a) *tri3* **(b)** *quad4*



(c) *tri6* **(d)** *quad8*



(e) *tetra4* **(f)** *hexa8*

Figure 4.8: Contact patch test – Cauchy stress along y-direction and contact traction for different types of first- and second-order mortar finite element interpolation.

4.3.2 Hertzian contact

In this section, attention is given to the solution of the classical contact problem of an elastic cylinder being pressed against an elastic block. It belongs to a class of problems commonly referred to as *Hertzian contact* and is recognised as one of the most popular examples within contact mechanics. It features a highly non-linear geometry setting stemming from the curved and evolving contact interface, and, more importantly, the great asset of this benchmark relies on the existence of an analytical solution originally proposed by Hertz (1881). In his treatment of the problem, he considers the equilibrium of two elastic bodies in contact on surfaces whose projection in the plane were conic sections, under the assumption that the contact area is obtained within an ellipse. Moreover, for the purpose of calculating the local deformations, at least one body is considered as an elastic half-space loaded over a small elliptical region of its place surface. As a result, Hertz derived formulas for the contact pressure and indentation between the two bodies.

The geometry and boundary conditions of the problem are schematically represented in Figure 4.9. It features a hard steel cylinder being pressed against an aluminium alloy block, both assumed to be linear elastic, with the properties summarized in Table 4.1. Only half of the cylinder is modelled, under the hypothesis of plane strain and frictionless contact. The lower boundary of the block is fixed and a concentrated vertical load F is applied in the top cylinder (various load magnitudes will be analysed). The radius of the cylinder is $R_1 = 50\,\text{mm}$, while the lower block has $W_2 = 100\,\text{mm}$ and $H_2 = 50\,\text{mm}$ of width and height, respectively. The surface of the cylinder is set as non-mortar and the block top surface as mortar. By using this problem setting, various aspects of the mortar-based contact formulation will be addressed and individually discussed in the following.

Table 4.1: Hertzian contact - material properties.

	Young's modulus, E (GPa)	Poisson's ratio, ν
Cylinder	210	0.3
Block	70	0.3

4.3.2.1 Comparison with analytical solution and commercial FE software

The first aspect to be studied is a comparison of the contact pressure distribution along with the contact interface according to different solutions for a load of magnitude $F = 35$ kN. The analytical solution can be obtained from the Hertzian contact for two cylinders (line contact), which is characterized via the maximum contact pressure,

$$p^{\eta}_{\max} = \sqrt{\frac{FE^*}{2\pi R^*}}\,, \tag{4.1}$$

and contact width

$$a = \sqrt{\frac{8FR^*}{\pi E^*}}\,. \tag{4.2}$$

Here, F denotes the applied normal force and E^* the combined elasticity modulus, which can be determined from the modulus of elasticity and Poisson's ratio of the cylinder and

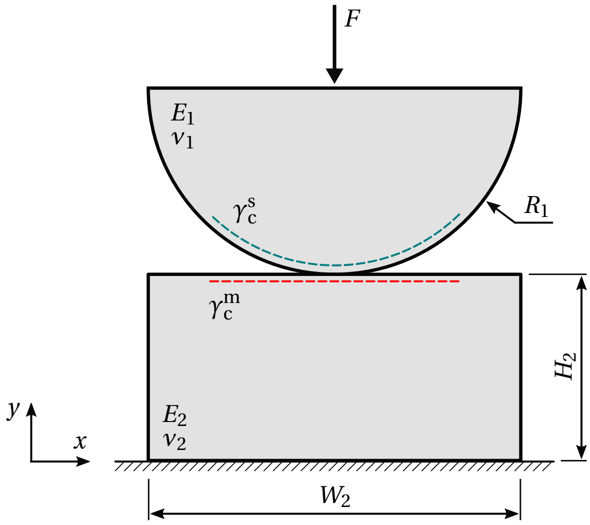


Figure 4.9: Hertzian contact – schematic representation of the problem setting.

block, E_1, ν_1 and E_2, ν_2, respectively, using the equation

$$E^* = \frac{2E_1E_2}{E_2\left(1-\nu_1^2\right)+E_1\left(1-\nu_2^2\right)} \,. \tag{4.3}$$

The combined radius of curvature, R^*, is evaluated from the radius of curvature of the cylinder, R_1, and block, R_2, as follows

$$R^* = \frac{R_1R_2}{R_1+R_2} \,. \tag{4.4}$$

For the particular type of problem at hand, it follows that $R_2 \to +\infty$, which yields

$$R^* = \lim_{R_2\to+\infty} \frac{R_1R_2}{R_1+R_2} = R_1 \,. \tag{4.5}$$

The pressure distribution along the contact width is determined by

$$p^\eta = p^\eta_{\max}\sqrt{1-\left(\frac{x}{a}\right)^2} \,. \tag{4.6}$$

Considering the numerical values for the problem at hand, the following results are obtained:

$$a \approx 6.21\,\text{mm}\,, \qquad p^\eta_{\max} \approx 3585.37\,\text{MPa}\,. \tag{4.7}$$

Note that half of the contact length is equal to 6.21 mm, which corresponds to ≈ 7.1° of the cylinder. This aspect should be kept in mind since, in order to simulate this problem appropriately, a very fine finite element mesh near the contact zone is required.

Regarding the numerical solution of the problem, besides the mortar algorithm presented in this work, results obtained with the commercial programs Abaqus/Standard® and MSC.MARC® are considered as well (Konter, 2006). Both linear and quadratic finite element interpolations are employed using the element types summarized in Table 4.2.

For the dual mortar algorithm, the 4-noded F-bar quadrilateral (QUAD4-FBAR) element and the 8-noded quadrilateral element with reduced integration (QUAD8-RI) are used for linear and quadratic interpolations, respectively. The two-dimensional bi-linear F-bar finite element is used because of its improved behaviour under nearly isochoric deformations. Essentially, the concept of this type of element is to substitute the volumetric component of the deformation gradient at the Gauss point that is being evaluated, calculating this term at the centroid of the element (de Souza Neto *et al.*, 1996). Along with the main advantages of such an approach, in the present context of contact interactions, the fact that this element is able to avoid volumetric locking and is suitable to capture strain localisation is of particular interest. In what concerns the commercial programs, linear elements with reduced integration and fully integrated quadratic elements are used for linear and quadratic interpolations, accordingly. The hard contact algorithm is used, which is based on direct coupling of the displacements using automatically generated constraint equations. The slave nodes (contacting nodes) correspond to the nodes on the cylinder, while the master nodes (contacted nodes) are set to the upper edge of the block. For all numerical solutions, the generated finite element mesh is such that the smallest element edges near the contact zone have approximately 0.59 mm for the cylinder and 1.5 mm for the block (see Figure 4.11 for an exemplary mesh).

Table 4.2: Hertzian contact - finite element types.

	LINKS	Abaqus/Standard®	MSC.MARC®
Linear	QUAD4-FBAR	CPE4R	Type 115
Quadratic	QUAD8-RI	CPE8	Type 27

The results for the contact pressure distribution along the contact interface are summarized in Figure 4.10, representing as well in Figure 4.11 an exemplary deformed configuration (obtained with the dual mortar algorithm and linear finite element interpolation). A close agreement of the dual mortar algorithm with the analytical reference contact pressure distribution is confirmed for both linear and quadratic interpolations. It should be noted, however, that an oscillating type of behaviour is obtained for the commercial FE programs, especially when using the quadratic elements.

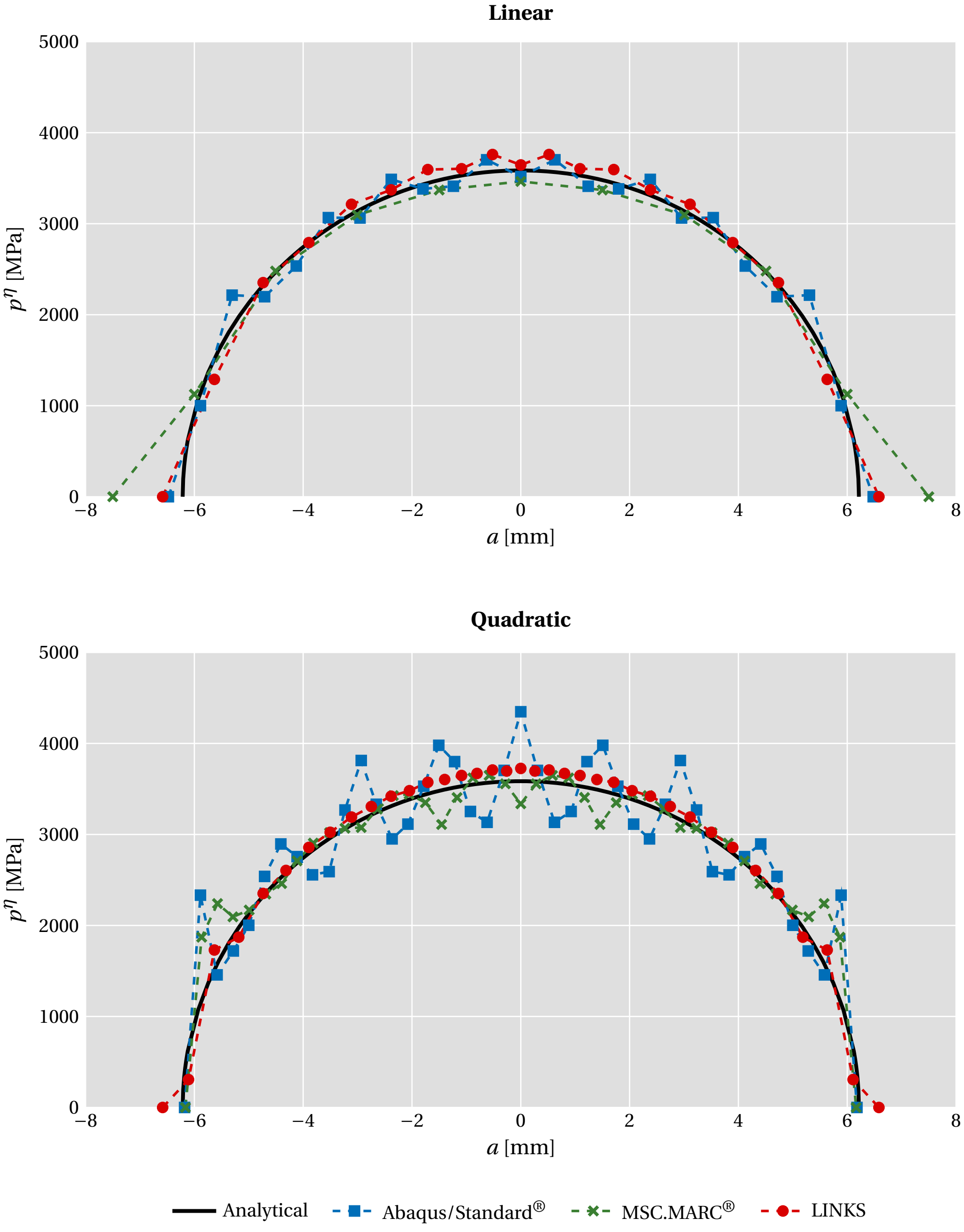


Figure 4.10: Hertzian contact – contact pressure distribution along with the contact interface for the analytical and numerical solutions.

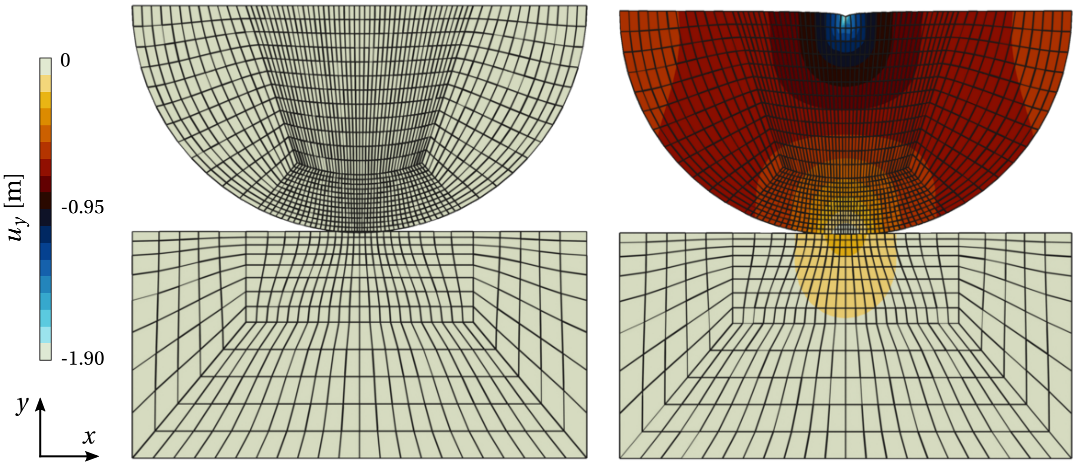
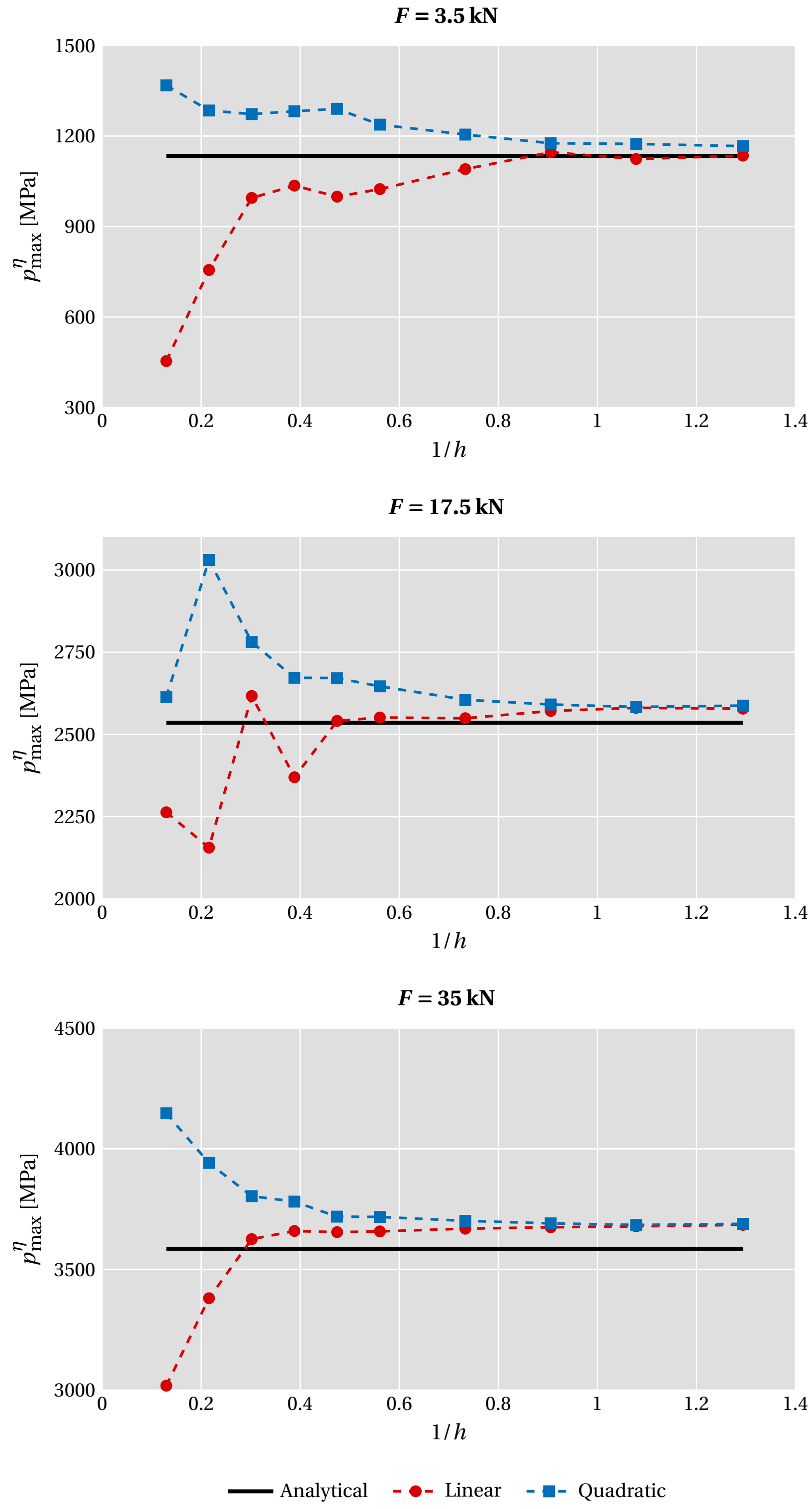


Figure 4.11: Hertzian contact – finite element mesh (on the left) and deformed geometry and displacement along the y-direction (on the right).

4.3.2.2 Spatial convergence rate

The next aspect to be addressed is the optimal convergence rates of the devised mortar algorithm. In order to investigate this property, progressive mesh refinement is applied to the entire model and the convergence of the numerical solution for the maximum contact normal pressure, p^{η}_{max}, is analysed. Furthermore, aiming at covering the convergence properties over a wide range of deformation magnitudes, three specific loads are considered: $F = 3.5$, 17.5 and 35 kN. The results are all gathered in Figure 4.12, from which expected convergence rates can be observed while reaching an asymptotic numerical solution. Another trend in the results is that by increasing the load magnitude, the relative error of the numerical results with respect to the analytical solution increases as well. This deviation is solely due to the fact that while the analytical solution is based on the small deformation assumption, the numerical algorithm fully considers all sources of non-linearity in the problem, including the geometrical finite deformation hypothesis. For higher loads, the deformation magnitude of the bodies is higher and, as a consequence, the impact of this fact in the results will be more pronounced and, ultimately, both solutions will diverge.

Lastly, in order to further analyse the capacity of the algorithm to reproduce the stress state along with the contact interface (and not only the maximum contact pressure p^{η}_{max}), in Figure 4.13, the contact normal pressure distribution is represented for each load magnitude, considering two different discretisations ($1/h \approx 0.56$ and $1/h \approx 1.29$). Again, a close agreement with the analytical solution is observed for both discretisations, even for the smallest load case ($F = 3.5\,\text{kN}$), where only a few elements are in contact.



$F = 3.5$ kN

$F = 17.5$ kN

$F = 35$ kN

Figure 4.12: Hertzian contact – maximum normal contact pressure, $p^{\eta}_{\max}$, with mesh refinement for three different values of load magnitude.

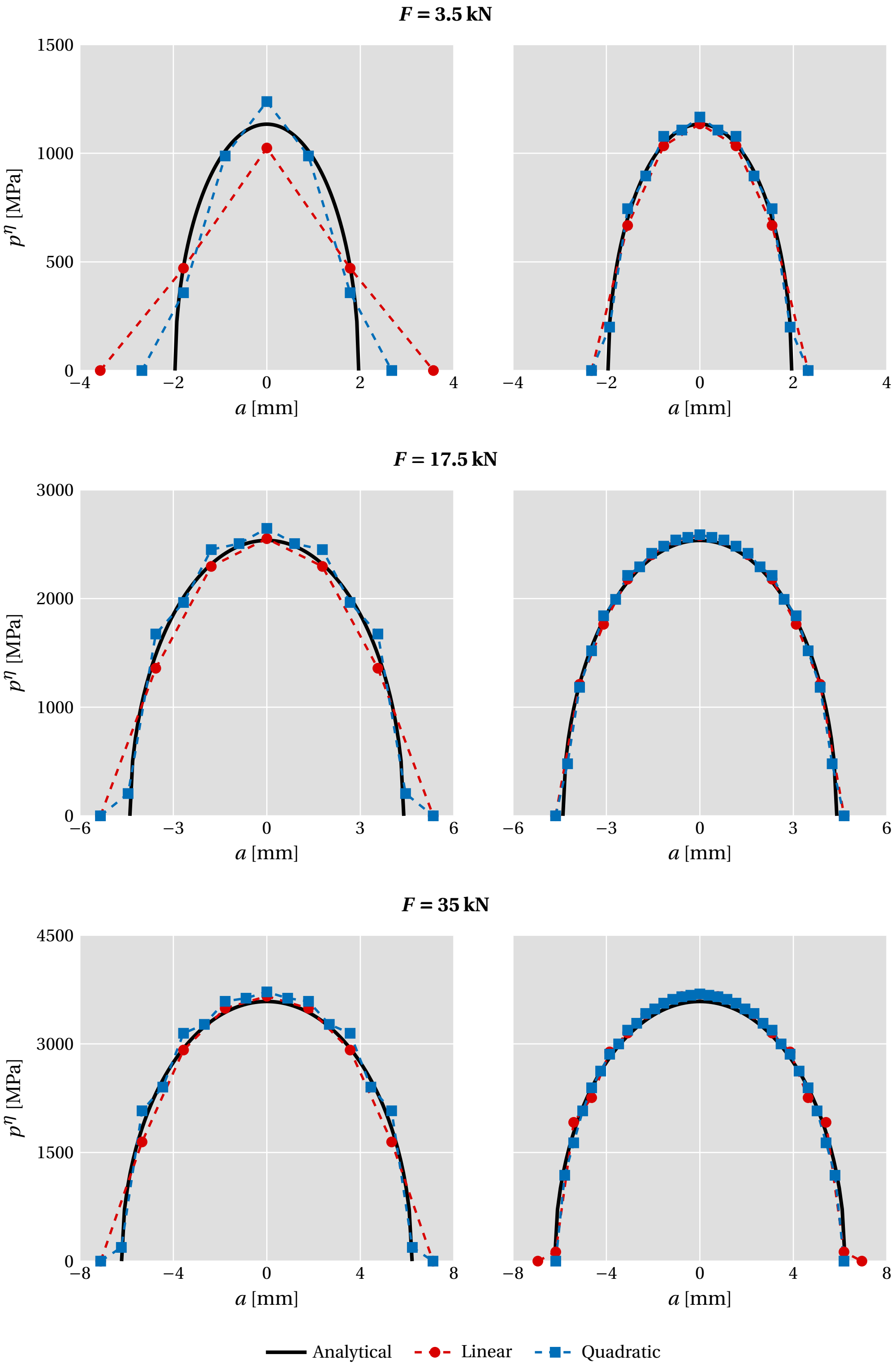


Figure 4.13: Hertzian contact – normal contact pressure distribution for three different values of load magnitude using two different discretizations ($1/h \approx 0.56$ on the left and $1/h \approx 1.29$ on the right).

4.3.2.3 Numerical efficiency of the semi-smooth Newton algorithm

The last property to be analysed is the numerical efficiency of the employed semi-smooth Newton algorithm. Characteristic results are given in Figure 4.14 for both linear and quadratic finite element interpolations. Regardless of the interpolation order, the active set strategy algorithm finds the correct active set within a few iterations steps. Once stabilised, the non-linear iterative scheme reduces to a standard (smooth) Newton-Raphson method and, ultimately, quadratic convergence is obtained, due to the underlying consistent linearisation.

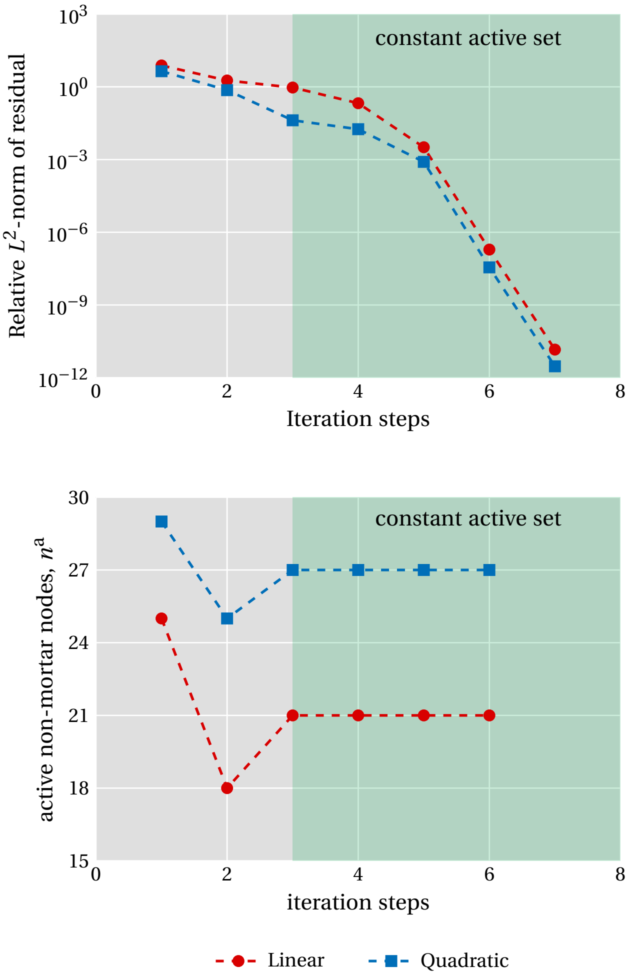


Figure 4.14: Hertzian contact – exemplary convergence behaviour of the semi-smooth Newton algorithm in terms of the L^2-norm of the residual (top) and active contact set (bottom). The shaded regions indicate the iterations steps with contact active set.

4.4 Large deformation rings

In this section, the large deformation contact between two rings is analysed. This problem has been analysed by several authors, see, e.g., Tur, Giner, *et al.* (2012), and its geometry and boundary conditions are schematically illustrated in Figure 4.15. Essentially, it consists of a ring of size $R_1^{\mathrm{int}} = 8\,\mathrm{mm}$ and $R_1^{\mathrm{ext}} = 10\,\mathrm{mm}$ contacting against a fixed ring of size $R_2^{\mathrm{int}} = 10\,\mathrm{mm}$ and $R_2^{\mathrm{ext}} = 12\,\mathrm{mm}$. The upper ring is subjected to a horizontal displacement of $\bar{u}_x = 35$ mm, applied at its top boundary in 70 increments. The two rings are assumed to be made of the same material and, with the aim of incorporating additional sources of non-linearity in the problem, an elasto-perfectly-plastic constitutive behaviour described by the von Mises model is considered. The Young's modulus is $E = 689.56\,\mathrm{MPa}$, the Poisson ratio $\nu = 0.32$ and the yield strength $\sigma_{\mathrm{y}} = 31\,\mathrm{MPa}$.

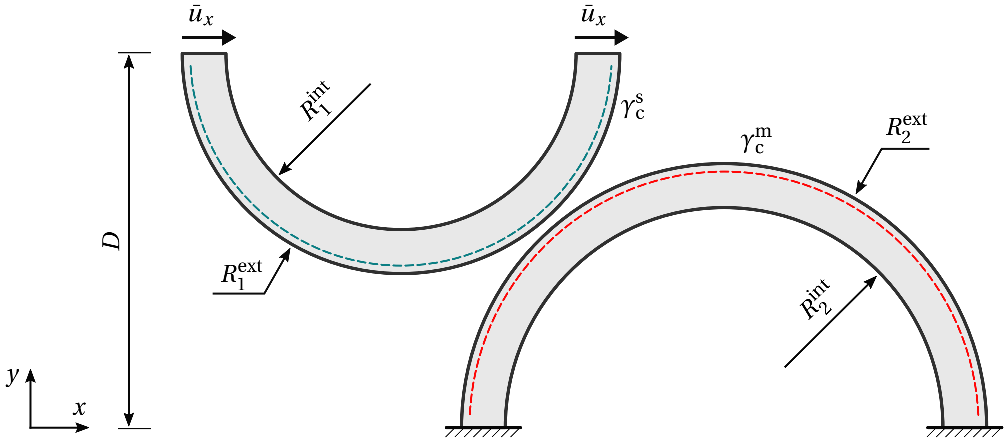
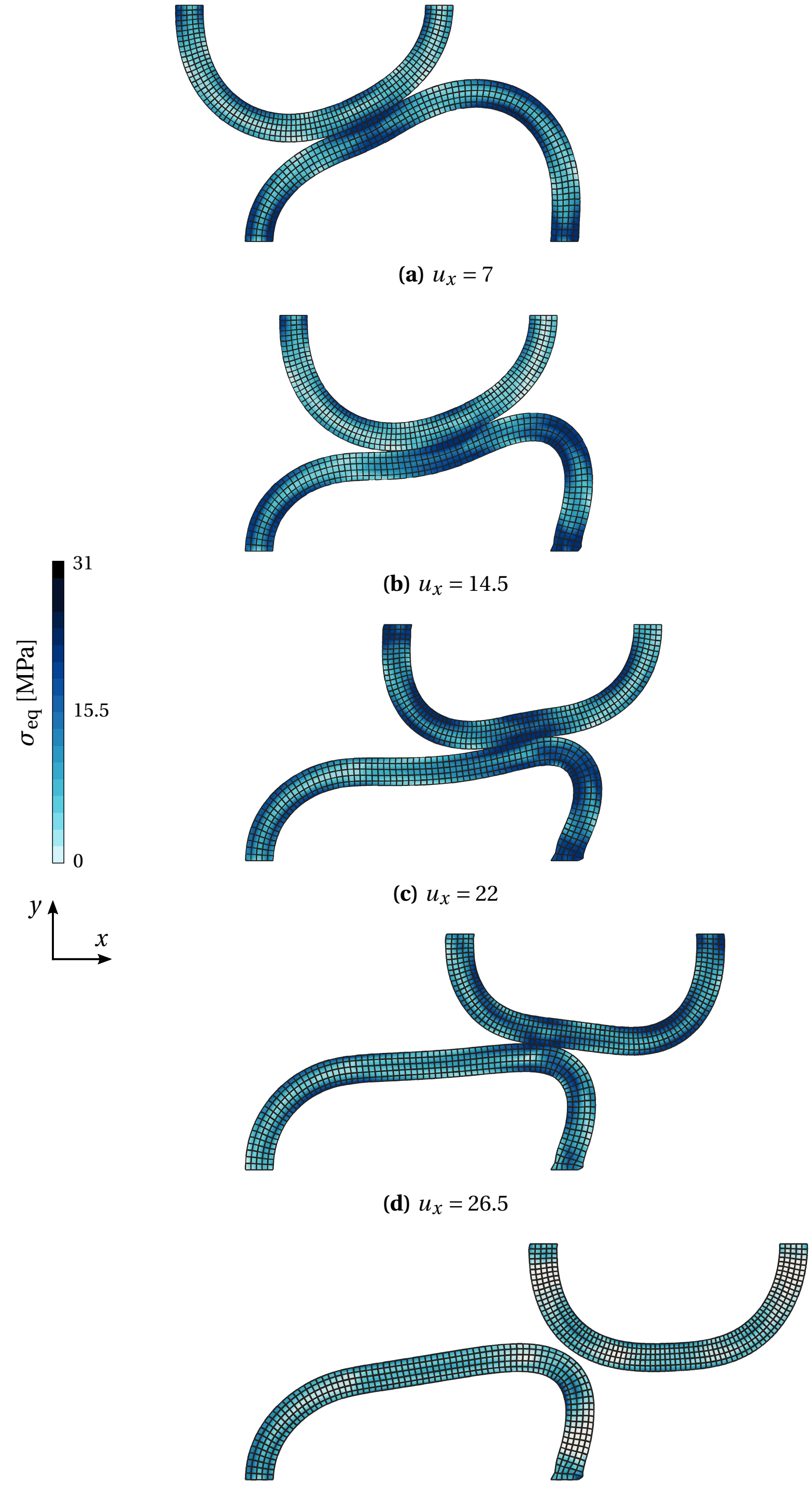


Figure 4.15: Large deformation rings – schematic representation of the problem.

Figure 4.16 shows the deformed configuration and the von Mises equivalent stress for various incremental steps. Both rings undergo a significant level of deformation and, due to the amount of plastic deformation accumulated, the initial geometry is not recovered and the final configuration changes significantly in relation to the initial one. Nonetheless, the optimal convergence properties of the algorithm are kept throughout the simulation and the imposed horizontal displacement to the top ring is successfully applied.

Remark. *In Tur, Giner,* et al. *(2012), two material behaviours are compared, considering the material to be elastic as well. However, such comparison cannot be carried out due to the presence of a snap-through phenomenon which, without appropriate algorithmic treatment, severely affects the convergence properties of the algorithm. This instability needs to be handled by using continuation techniques, such as the arc-length scheme.*



(a) $u_x = 7$

(b) $u_x = 14.5$

(c) $u_x = 22$

(d) $u_x = 26.5$

(e) $u_x = 32.5$

Figure 4.16: Large deformation rings – distribution of the von Mises equivalent stress at various stages of the deformation process.

4.5 Shallow ironing

The class of ironing contact problems is a popular choice for the analysis of finite deformation frictional contact between two deformable bodies undergoing a non-monotonic loading path. The geometry of the specific problem is schematically represented in Figure 4.17, which consists of the contact problem between a deformable indenter with a rounded side and a rectangular flat base. The foundation is fixed at the bottom and the indenter's top boundary is prescribed in a two-stage deformation process. Firstly, it is pressed down against the deformable base under the prescribed vertical displacement $\bar{u}_y = 1$. Secondly, the horizontal displacement $\bar{u}_x = 10$ makes the indenter slide along the deformable base. Each stage lasts 1 second and the Coulomb's law with the friction coefficient $\mu = 0.3$ is considered. The solution of the problem requires modelling friction during two distinct contact scenarios. During the compression stage, the main difficulty involves determining which nodes come into contact, with contact nodes at both stick and slip conditions. During the sliding stage, the indenter starts sliding and the correct evaluation of the tangential forces becomes critical. It is expected that, as the relative motion between the indenter and the base occurs, nodes at the stick state become sliding nodes.

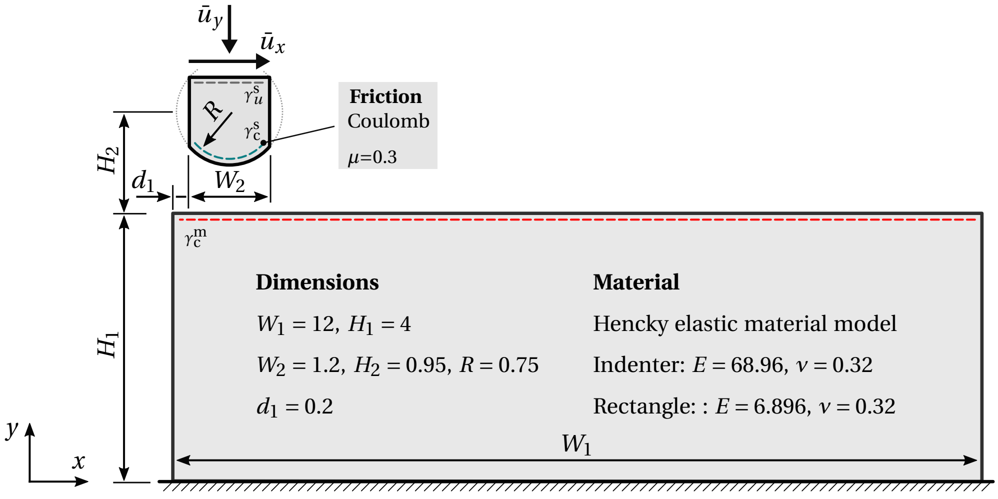
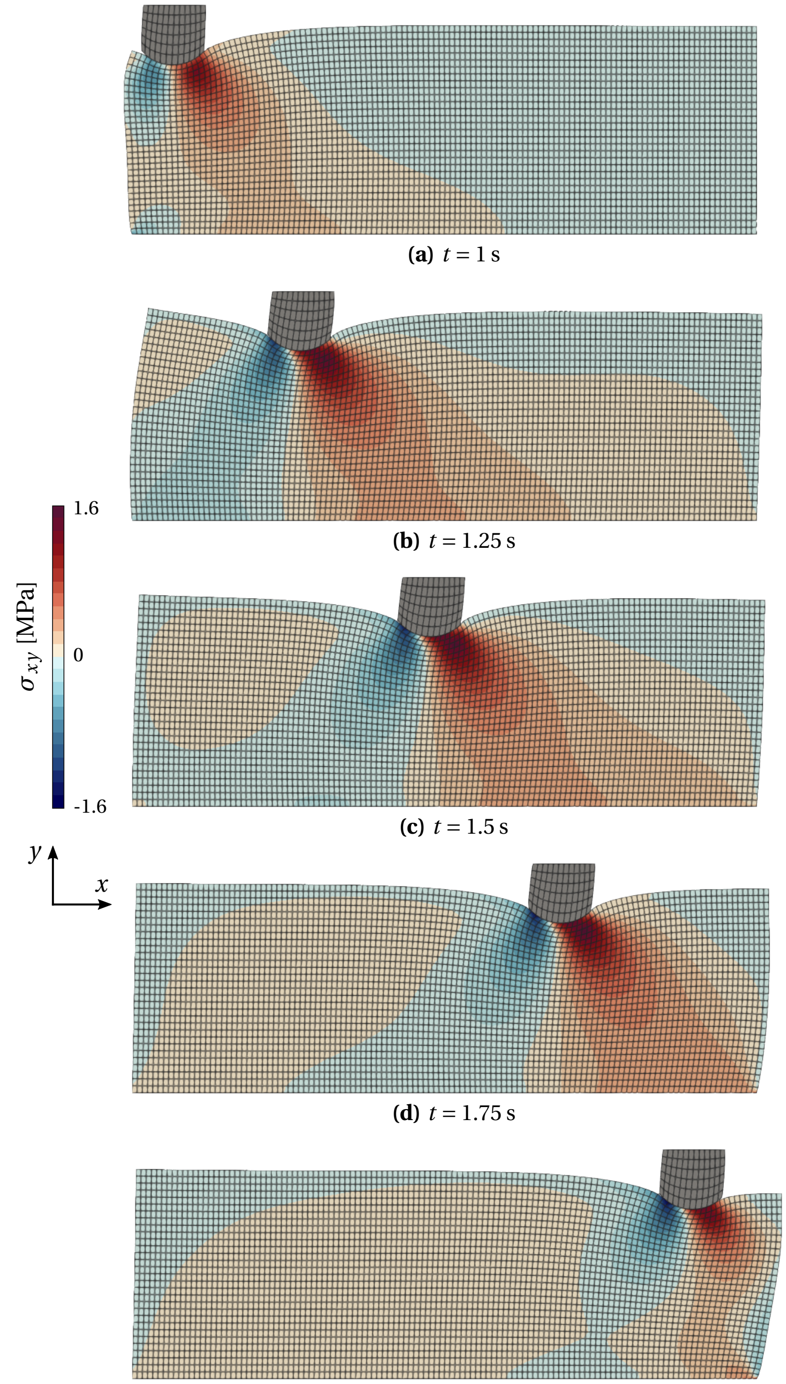


Figure 4.17: Shallow ironing – schematic representation of the problem setting.

The incrementation originally proposed in Fischer and Wriggers (2006) is adopted, according to which the compression stage is performed in 10 increments and the horizontal displacement in 500 increments. For the numerical model employed in this work, both bodies are modelled using the Ogden elastic model and discretised using 4-noded quadrilateral F-bar finite elements. The number of elements, $n^{\mathrm{e}} = 3672$, is such that the total number of degrees of freedom remains the same as in Fischer and Wriggers (2006).[†] Figure 4.18 shows the deformed configuration throughout the process, including a contour plot of the shear stresses at the deformable base. Two regions with opposite shear stresses can be identified, which move along the base as the indenter slides over it.

[†] The numerical model in Fischer and Wriggers (2006) uses 9-noded quadratic finite elements.



(a) $t = 1$ s

(b) $t = 1.25$ s

(c) $t = 1.5$ s

(d) $t = 1.75$ s

(e) $t = 2.0$ s

Figure 4.18: Shallow ironing – distribution of the Cauchy shear stress σ_{xy} at various stages of the deformation process.

In this example, the sum of the reaction forces at the top of the indenter is typically measured throughout the simulation, existing several results available in the literature. Figure 4.19 shows a comparison between the ongoing formulation and other solutions. The green curve corresponds to the formulation proposed by Fischer and Wriggers (2006), based on a mortar contact formulation using the concept of the moving friction cone and a penalty regularisation scheme. In blue, the results amount to Hartmann, Oliver, *et al.* (2009) using the contact domain method. The results in red, given in V. Yastrebov (2011), are obtained with the NTS discretisation and the augmented Lagrangian method.

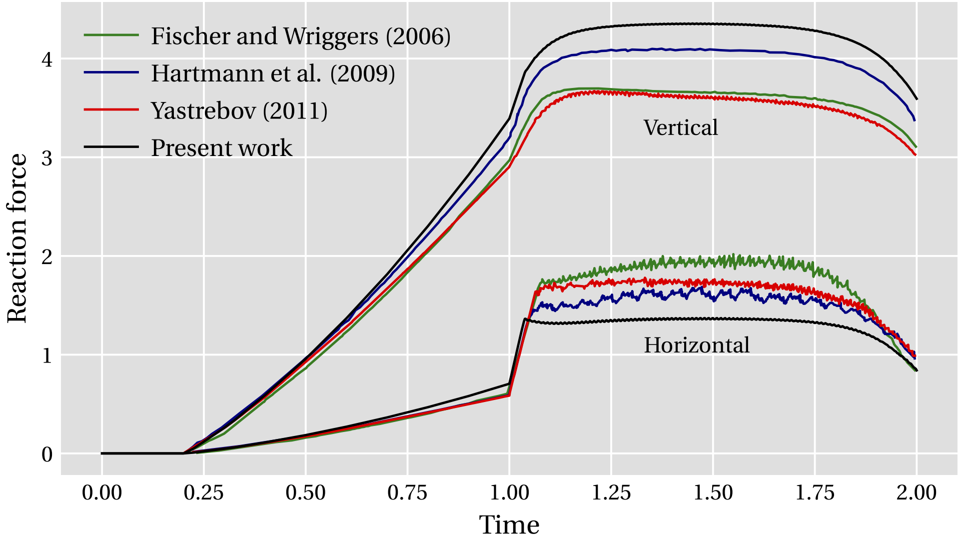


Figure 4.19: Shallow ironing – vertical and horizontal reactions on the indenter compared with the results obtained by Fischer and Wriggers (2006), Hartmann, Oliver, *et al.* (2009) and V. Yastrebov (2011).

As reported by all the authors mentioned in Figure 4.19, the two stages of the problem can be identified in the plot. During the compression process, both the vertical and horizontal components of the reactions at the indenter increase as it is pressed against the base. Once the indenter starts sliding, there is an increase in both vertical and horizontal reactions and subsequent stabilisation. This jump represents the resistance that the foundation offers to the sliding motion and is most likely originated by the change in contact conditions (since there are multiple contact nodes transitioning from stick to slide state). Throughout this stage, the ratio between the normal and horizontal reactions approximates the value of the coefficient of friction. Comparing the present results with the remaining authors, the ongoing formulation estimates the vertical reaction to be higher and the horizontal component to be smaller. Since the algorithms are significantly different, the answer for the origin of such differences is not obvious. Nonetheless, the oscillations observed in the remaining methods are more pronounced than in the present work. Since the total number of degrees of freedom is kept the same across all results, this demonstrates the quality of results achieved with the dual mortar contact formulation.

4.6 Squeezed plastic tube

The last numerical example presented in this chapter is a three-dimensional problem that demonstrates the capability of the dual mortar contact algorithm to deal with multiple sources of non-linearity simultaneously. It features frictional contact combined with plasticity at finite deformations. The problem setting is schematically represented in Figure 4.20 and has been originally presented in Seitz, Popp, *et al.* (2015)—which is also inspired by a similar example analysed in Hager and B. I. Wohlmuth (2009). It consists of a cylindrical tube, prescribed at both ends, that is squeezed by two rigid semi-cylindrical tools in a five-stage process. Starting from a stress-free configuration, the tools are prescribed along the vertical direction in the first stage, squeezing the tube for the first time (stroke 1). Then, the load is released by imposing the opposite displacement to the tools, finishing the first cycle. In the third intermediate stage, a 90 degrees rotation is imposed to the tube's ends and, then, the second compression cycle is applied (stroke 2 and release). In this work, the problem is solved using the 8-noded F-bar hexahedron finite element and the outer boundary of the cylindrical tube is set as mortar and the tools as non-mortar. The plastic behaviour of the cylindrical tube is described by the von-Mises constitutive model and the finite element mesh consists of ≈ 30k nodes. The geometry, dimensions and properties of the problem are all indicated in Figure 4.20.

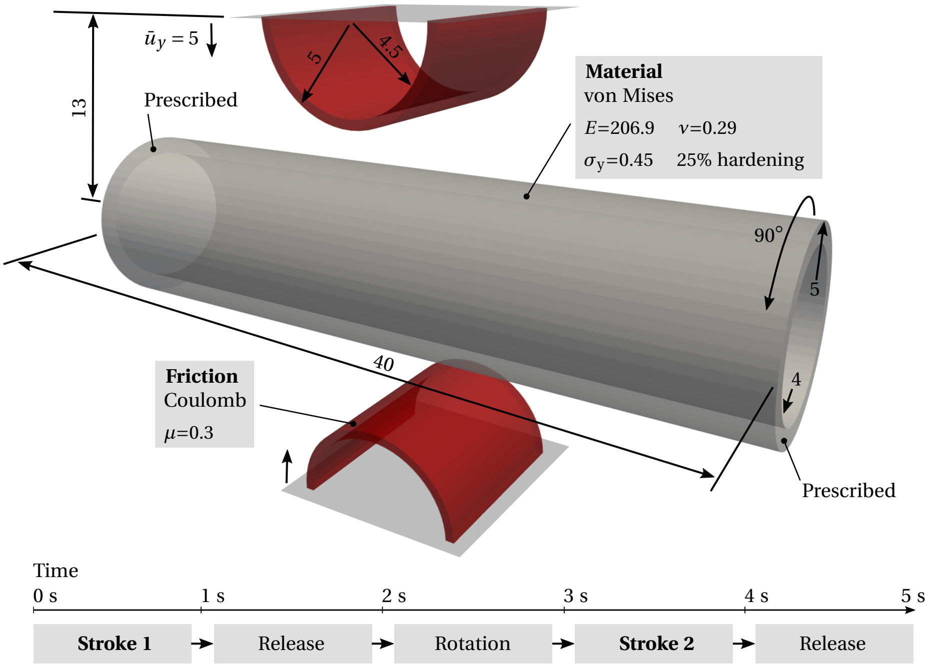
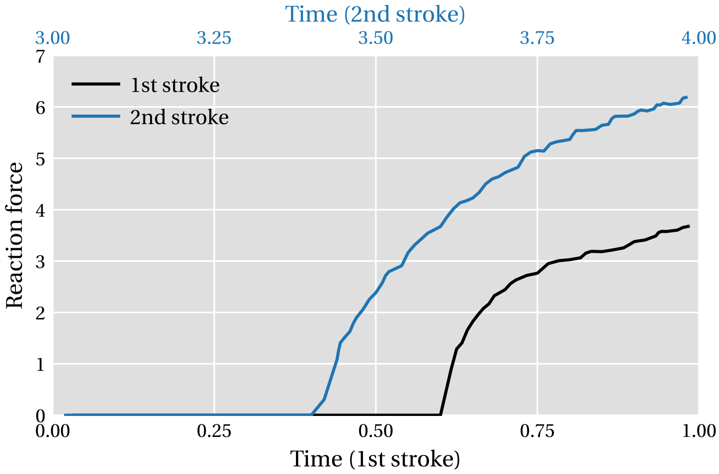


Figure 4.20: Squeezed plastic tube – schematic representation of the problem setting.

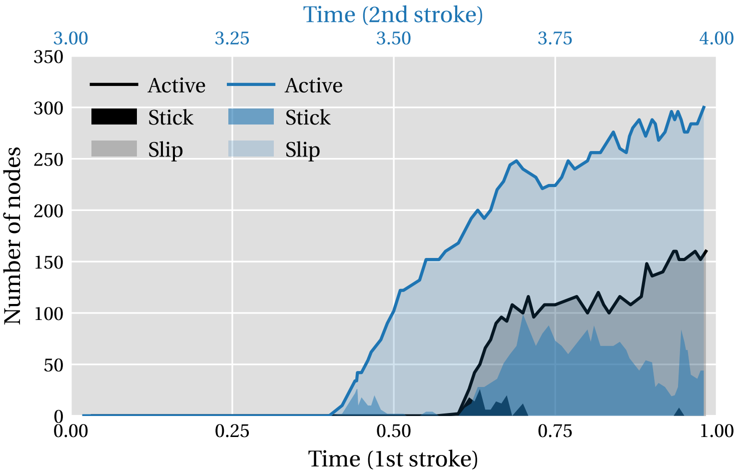
The deformed configuration at the end of each stage of the process is given in Figure 4.22. The contour plot represents the von-Mises equivalent stress, from which it is possible to identify the residual stress present in the cylindrical tube after each stroke. During the release of the tool and rotation of the tube, there is not any external solicitation to the cylinder and, therefore, the stresses are entirely due to the accumulated plastic strain, see Figure 4.23 for its distribution.

In order to have a better understanding of the problem, Figure 4.21a shows the evolution of the vertical reaction forces at the tool during each stroke. The main observation obtained from these results is that, at the second stroke, the tool touches the tube earlier than the first stroke. This is happening due to the permanent deformation present in the cylinder after the first compression. After the 90 degrees rotation, it decreases the initial gap distance between the cylinder and the tools. Moreover, the reaction forces are more significant during the second stroke, which is consistent with the earlier contact interaction and material hardening.

Lastly, in Figure 4.21b the total number of active contact nodes is also presented, also representing the partial sum of stick and slip nodes. It can be seen that, for both stroke cycles, the number of active contact nodes increases at a more significant rate during the initial contact interactions (finding several nodes at stick state), after which a slight stabilisation is observed (most of the contact nodes become sliding). This behaviour is more pronounced during the first stroke, while in the second, the existent plastic deformation changes the geometry of the contact zone and originates a more significant contact region (with nodes at both stick and slip states).

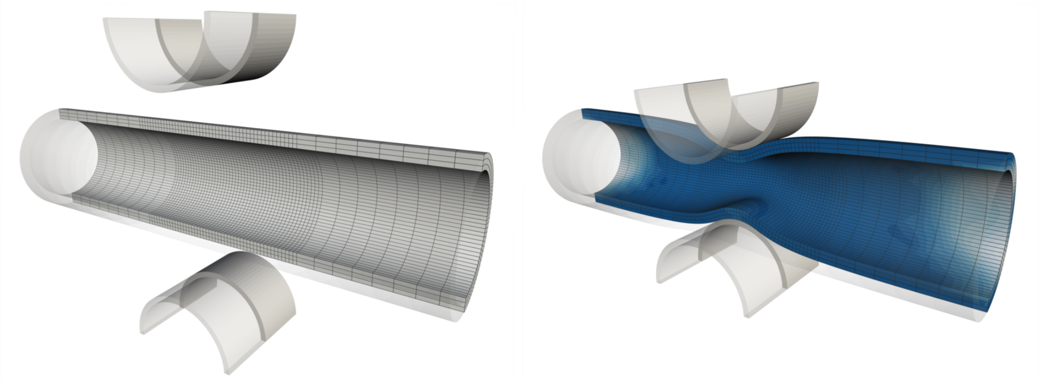


(a) Sum of the vertical reaction forces at the outer boundary of the upper tool.

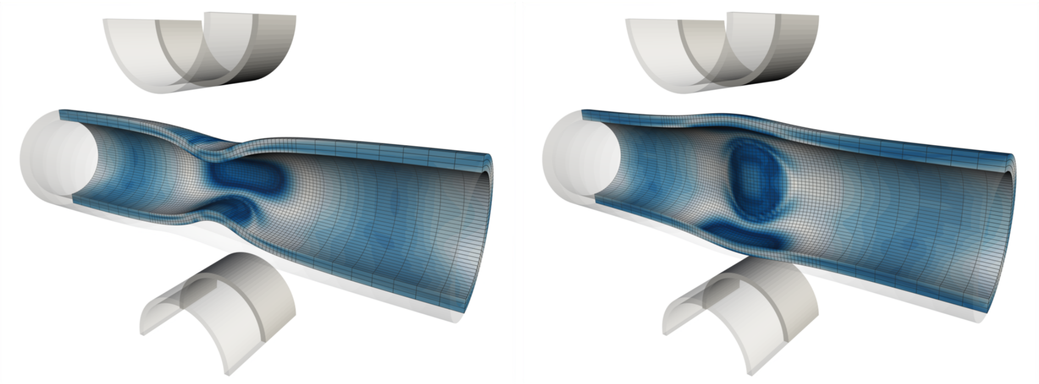


(b) Total number of active contact nodes and partial count of stick and slip nodes.

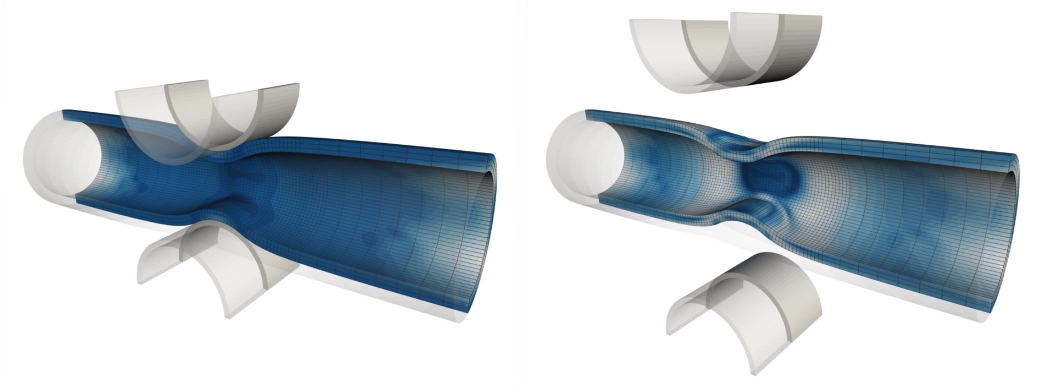
Figure 4.21: Squeezed plastic tube – reactions and number of active nodes at each stroke stage.



(a) Initial configuration ($t = 0.0$ s) **(b)** First stroke ($t = 1.0$ s)



(c) Release ($t = 2.0$ s) **(d)** Rotation ($t = 3.0$ s).



(e) Second stroke ($t = 4.0$ s) **(f)** Release ($t = 5.0$ s)

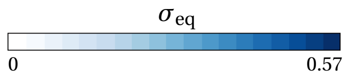
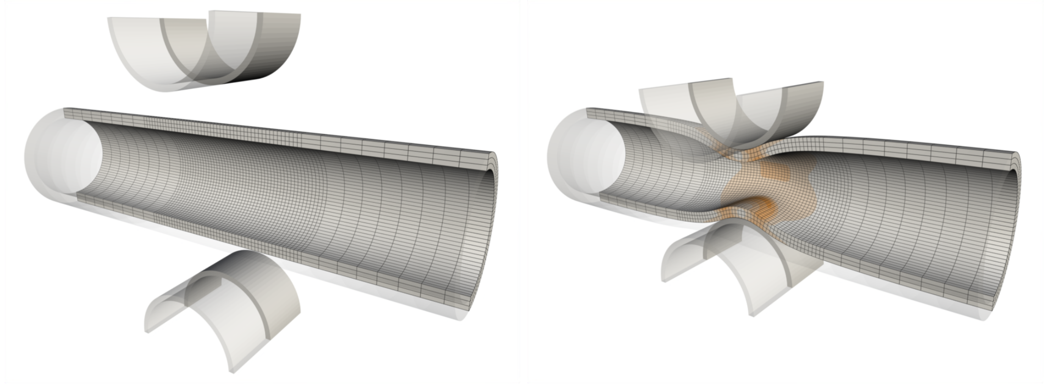
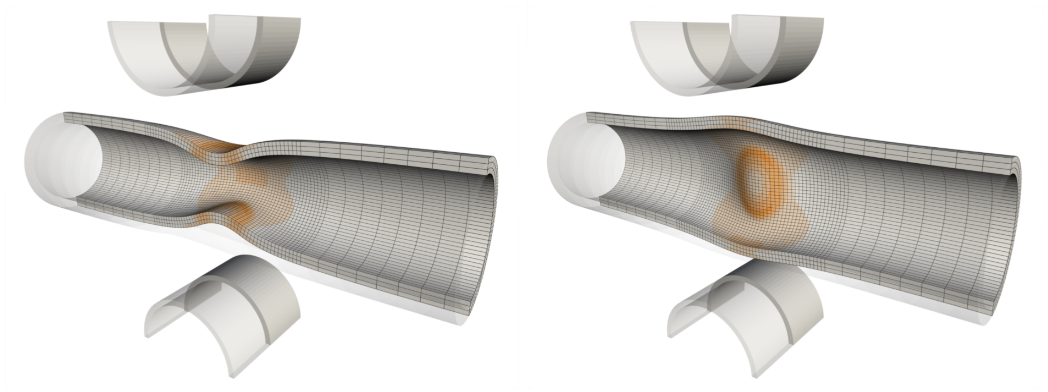


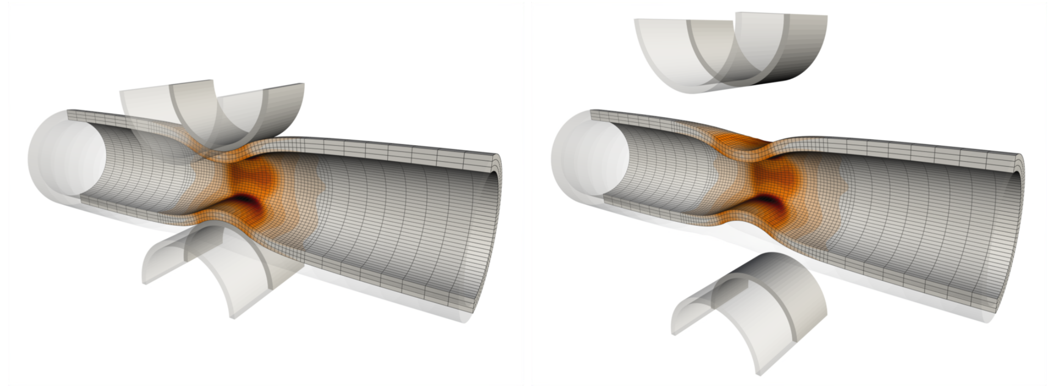
Figure 4.22: Squeezed plastic tube – distribution of the equivalent stress at the end of each stage of the simulation.



(a) Initial configuration ($t = 0.0$ s) **(b)** First stroke ($t = 1.0$ s)



(c) Release ($t = 2.0$ s) **(d)** Rotation ($t = 3.0$ s).



(e) Second stroke ($t = 4.0$ s) **(f)** Release ($t = 5.0$ s)

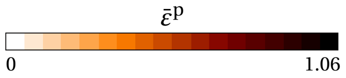


Figure 4.23: Squeezed plastic tube – distribution of the accumulated plastic strain at the end of each stage of the simulation.

Page intentionally left blank

Chapter 5

An Efficient Dual Mortar Contact Technique for Rigid/Deformable Interaction

The classification of contact and impact procedures is typically based upon the problem configuration, with factors like the total number of bodies contacting each other or their physical behaviour originating different classes of contact problems. While the contact between two deformable bodies is typically termed as unilateral contact, the setup consisting of one single deformable body going against a rigid obstacle is commonly referred to as Signorini contact. The former can be interpreted as the most general class, turning out essential to derive the fundamental mathematical and computational frameworks. However, the underlying assumptions in Signorini contact promotes opportunities for simplifications. This chapter is precisely dedicated to the development of such techniques, having in mind the goal of reducing the computational complexity of the dual mortar contact algorithm for Signorini contact.

The topic of rigid/deformable contact is encountered in a wide range of practically important problems, such as forming operations and impact tests. Under these circumstances, assuming that one of the contacting bodies is rigid is an excellent approximation to the physical phenomenon, which can significantly simplify the contact analysis. In addition, within the framework of multi-scale contact modelling, it is also convenient to adopt this simplification, which is supported by a classic result from contact mechanics. For micromechanical contact without friction or adhesion between two rough surfaces with a small slope, the elastic contact can be mapped to an equivalent problem precisely involving one deformable body (with a combined roughness profile) and a rigid flat (Johnson, 1987). Taking into consideration that the computational complexity of the FEM typically discourages its application to model rough contact, there is a natural motivation for the development of strategies that aim at attenuating the overall complexity of the finite element algorithm. This aspect becomes even more relevant when the dual mortar method is employed, which involves complex concepts and algorithms.

This chapter starts with the mathematical formulation of rigid/deformable contact under large deformations, both in strong and weak forms. Then, the finite element discretisation is described, emphasizing the two main concepts proposed as simplifications of the algorithm. Firstly, the variational formulation using a Petrov-Galerkin scheme is further investigated, as it unlocks a significant simplification of the algorithm by removing the need to evaluate the dual shape functions. In this regard, the extension for quadratic dual mortar interpolation is also carried out. Secondly, a new definition for the nodal orthonormal moving frame attached to each contact node is proposed. Then, several numerical examples are presented and discussed, in order to validate and inspect the computational performance of the algorithm.

5.1 Continuum mechanics of the Signorini contact problem

From the viewpoint of mathematical problem formulation, the Signorini contact is a popular alternative to introduce the fundamental concepts of contact mechanics within the linear regime. Nonetheless, the approach to rigid/deformable contact derived in this work assumes large deformations and generic constitutive behaviour. This leads to a fully non-linear problem, which can be obtained based upon the concepts of unilateral contact between two deformable bodies introduced in Chapter 2. The problem is derived in its strong and weak forms in the following.

5.1.1 Strong form

The nomenclature and illustration of the problem are schematically illustrated in Figure 5.1. The (only) deformable body is designated as non-mortar and identified with the superscript $(\bullet)^{\mathrm{s}}$. The open set $\Omega_0^{\mathrm{s}} \subset \mathbb{R}^d\,(d=2,3)$ represents its reference configuration and the boundary $\Gamma_{\mathrm{c}}^{\mathrm{r}}$ stands for the rigid body contour. Herein, the superscript $(\bullet)^{\mathrm{r}}$ relates to quantities associated with the rigid boundary (thus replacing the notion of the mortar side). The boundary $\partial\Omega_0^{\mathrm{s}}$ in the reference configuration can be divided into three open disjoint subsets: the Dirichlet boundary Γ_u^{s}, with prescribed displacements $\bar{\boldsymbol{u}}^{\mathrm{s}}$, the Neumann boundary $\Gamma_\sigma^{\mathrm{s}}$, satisfying a given surface traction $\bar{\boldsymbol{t}}^{\mathrm{s}}$, and the non-mortar potential contact interface $\Gamma_{\mathrm{c}}^{\mathrm{s}}$. As the body undergoes motion, denoted by the smooth mapping φ^{s}, the counterparts in the current configuration $\Omega_t^{\mathrm{s}} \subset \mathbb{R}^d$ are denoted as γ_u^{s}, $\gamma_\sigma^{\mathrm{s}}$ and $\gamma_{\mathrm{c}}^{\mathrm{s}}$, respectively. Note that the rigid boundary $\Gamma_{\mathrm{c}}^{\mathrm{r}}$ remains stationary throughout the entire process, thus having a null displacement field.

Remark. *The mathematical formulation and computer implementation for the case of a rigid boundary undergoing rigid-body motions poses no additional difficulties at all to the problem setting. Only for the sake of simplicity, the particular case of a non-moving rigid boundary is exclusively considered in the present work.*

The continuum mathematical framework presented in Chapter 2 is equally applicable to this problem, with the only exception being a slight modification in the gap vector definition. As a fundamental entity of contact kinematics, the gap vector is involved in the definition of both the gap function g and instantaneous relative velocity $\boldsymbol{v}^{\tau}$, and for rigid/deformable contact follows as

$$\boldsymbol{g}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) \equiv \boldsymbol{x}^{\mathrm{s}} - \hat{\boldsymbol{x}}^{\mathrm{r}}\left(\boldsymbol{x}^{\mathrm{s}}, t\right) . \tag{5.1}$$

The contact point $\hat{\boldsymbol{x}}^{\mathrm{r}}$ on the rigid boundary $\Gamma_{\mathrm{c}}^{\mathrm{r}}$ stands for the projection of the non-mortar point $\boldsymbol{x}^{\mathrm{s}} \in \gamma_{\mathrm{c}}^{\mathrm{s}}$ along its current outward unit normal vector $\boldsymbol{\eta}$. Compared with the general counterpart in Equation (2.26), the difference is that the coordinates on the opposing

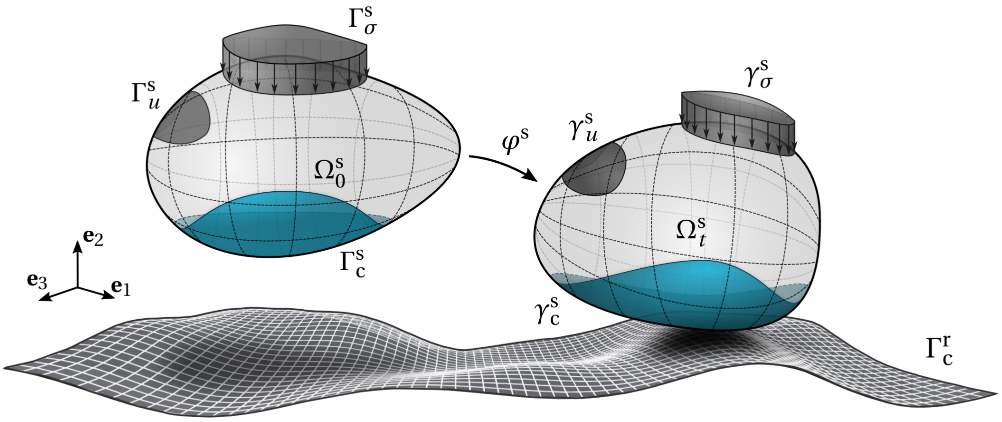


Figure 5.1: Illustration and nomenclature of the Signorini contact problem in the reference and current configurations.

side to the non-mortar boundary $\gamma_{\mathrm{c}}^{\mathrm{s}}$ remain fixed, i.e.,

$$\boldsymbol{x}^{\mathrm{r}} \equiv \boldsymbol{X}^{\mathrm{r}} . \tag{5.2}$$

Nonetheless, the projection itself, $\hat{\boldsymbol{x}}^{\mathrm{r}}$, still depends on the deformation of the opposite contact boundary $\gamma_{\mathrm{c}}^{\mathrm{s}}$ since the projection possibly changes over time. As explained in more detail in the following paragraphs describing the discrete version of the problem, the fact that one of its terms remains fixed simplifies the algorithm. Notwithstanding, the strong form of the IBVP of finite deformation frictional contact between a deformable body and a rigid obstacle is stated in the following.

Problem 5.1 (The strong form of IBVP of finite deformation frictional rigid/deformable contact)
In the domain Ω_t, the momentum balance principle and the set of Dirichlet and Neumann boundary conditions read

$$\operatorname{div}\boldsymbol{\sigma}^{\mathrm{s}} + \boldsymbol{b}^{\mathrm{s}} = \boldsymbol{0} , \quad \text{in } \Omega_t^{\mathrm{s}} , \tag{5.3a}$$

$$\boldsymbol{u}^{\mathrm{s}} = \bar{\boldsymbol{u}}^{\mathrm{s}} , \quad \text{on } \gamma_u^{\mathrm{s}} , \tag{5.3b}$$

$$\boldsymbol{\sigma}^{\mathrm{s}}\boldsymbol{n}^{\mathrm{s}} = \bar{\boldsymbol{t}}^{\mathrm{s}} , \quad \text{on } \gamma_\sigma^{\mathrm{s}} , \tag{5.3c}$$

and the contact constraints in the normal and tangential directions are given by

$$g \geq 0 , \quad p^{\eta} \leq 0 , \quad g p^{\eta} = 0 \quad \text{on } \gamma_{\mathrm{c}}^{\mathrm{s}} , \tag{5.4a}$$

$$\psi \leq 0 , \quad \boldsymbol{v}^{\tau} + \beta \boldsymbol{t}^{\tau} = \boldsymbol{0} , \quad \beta \geq 0 , \quad \beta\psi = 0 \quad \text{on } \gamma_{\mathrm{c}}^{\mathrm{s}} , \tag{5.4b}$$

accordingly.

5.1.2 Weak form

The derivation of the weak version of the problem begins with the introduction of the solution space $\mathcal{U}^{\mathrm{s}}$ and weighting space $\mathcal{V}^{\mathrm{s}}$:

$$\mathcal{U}^{\mathrm{s}} \equiv \left\{ \boldsymbol{u}^{\mathrm{s}} \in \left[H^1(\Omega_t^{\mathrm{s}})\right]^d \mid \boldsymbol{u}^{\mathrm{s}} = \bar{\boldsymbol{u}}^{\mathrm{s}} \text{ on } \gamma_u^{\mathrm{s}} \right\}, \tag{5.5}$$

$$\mathcal{V}^{\mathrm{s}} \equiv \left\{ \delta\boldsymbol{u}^{\mathrm{s}} \in \left[H^1(\Omega_t^{\mathrm{s}})\right]^d \mid \delta\boldsymbol{u}^{\mathrm{s}} = \boldsymbol{0} \text{ on } \gamma_u^{\mathrm{s}} \right\}. \tag{5.6}$$

These spaces are conceptually similar to the general case of unilateral contact, although involving only the mappings on the deformable non-mortar body. Consideration of the principle of virtual work allows rewriting the momentum balance, Equation (5.3a), as

$$\delta\Pi_{\mathrm{int}}^{\mathrm{s}} - \delta\Pi_{\mathrm{ext}}^{\mathrm{s}} - \delta\Pi_{\mathrm{c}}^{\mathrm{s}} = 0\,, \quad \forall\, \delta\boldsymbol{u}^{\mathrm{s}} \in \mathcal{V}^{\mathrm{s}}\,, \tag{5.7}$$

where $\delta\Pi_{\mathrm{int}}^{\mathrm{s}}(\boldsymbol{u}^{\mathrm{s}},\delta\boldsymbol{u}^{\mathrm{s}})$ represents the internal virtual work, $\delta\Pi_{\mathrm{ext}}^{\mathrm{s}}(\delta\boldsymbol{u}^{\mathrm{s}})$ the external virtual work and the third term, $\delta\Pi_{\mathrm{c}}^{\mathrm{s}}(\boldsymbol{u}^{\mathrm{s}},\delta\boldsymbol{u}^{\mathrm{s}})$, the virtual work of contact forces. These terms remain unchanged from the classical unilateral contact problem, with the exception being the virtual work of contact forces. For the particular case of Signorini contact, this contribution involves only the virtual displacements of the non-mortar side, i.e.,

$$\delta\Pi_{\mathrm{c}}^{\mathrm{s}}(\boldsymbol{u}^{\mathrm{s}},\delta\boldsymbol{u}^{\mathrm{s}}) \equiv -\int_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \boldsymbol{t}_{\mathrm{c}}^{\mathrm{s}} \cdot \delta\boldsymbol{u}^{\mathrm{s}} \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,. \tag{5.8}$$

As discussed in more detail in Section 5.2, this is the origin of one of the main simplifications of the algorithm. The so-called second mortar coupling matrix can be removed, thus evaluating the virtual work of contact forces using solely the first mortar coupling matrix.

The last step towards the formulation of the weak version of rigid/deformable large deformation contact comprises the treatment of the contact constraints. In this regard, the concepts introduced in Chapter 2 are thoroughly applicable. The Lagrange multiplier is introduced as the negative contact traction on the non-mortar side and chosen from the convex cone $\mathcal{M}(\boldsymbol{\lambda})$, see Section 2.6. Then, by adopting the concept of variational inequalities, the final weak form of the problem can be stated.

Problem 5.2 (The weak form of IBVP of finite deformation frictional rigid/deformable contact)

Given the internal forces and boundary conditions, find the kinematically admissible displacement function, $\boldsymbol{u} \in \mathcal{U}^{\mathrm{s}}$, and Lagrange multiplier, $\boldsymbol{\lambda} \in \mathcal{M}(\boldsymbol{\lambda})$, such that, for all $t \in [t_0, T]$, the virtual work equations

$$\delta\Pi_{\mathrm{int}}^{\mathrm{s}}(\boldsymbol{u},\delta\boldsymbol{u}) - \delta\Pi_{\mathrm{ext}}^{\mathrm{s}}(\delta\boldsymbol{u}) + \int_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \boldsymbol{\lambda} \cdot \delta\boldsymbol{u}^{\mathrm{s}} \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} = 0\,, \quad \forall\, \delta\boldsymbol{u} \in \mathcal{V}^{\mathrm{s}}\,, \tag{5.9}$$

$$\langle\, g\,, \delta\lambda^{\eta} - \lambda^{\eta}\, \rangle_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \geq 0\,, \quad \forall\, \delta\boldsymbol{\lambda} \in \mathcal{M}(\boldsymbol{\lambda})\,, \tag{5.10}$$

$$\langle\, \boldsymbol{v}^{\tau}, \delta\boldsymbol{\lambda}^{\tau} - \boldsymbol{\lambda}^{\tau}\, \rangle_{\gamma_{\mathrm{c}}^{\mathrm{s}}} \leq 0\,, \quad \forall\, \delta\boldsymbol{\lambda} \in \mathcal{M}(\boldsymbol{\lambda})\,, \tag{5.11}$$

are satisfied for any admissible test functions $\delta\boldsymbol{u} \in \mathcal{V}^{\mathrm{s}}$ and $\delta\boldsymbol{\lambda} \in \mathcal{M}(\boldsymbol{\lambda})$.

5.2 Mortar finite element discretisation

The weak form derived in Section 5.1.2 is approximated using the FEM, whose foundations are not affected by the presence of a rigid boundary. Therefore, most of the concepts described in Chapter 3 remain unchanged: the domain Ω^{s} is partitioned into n^{e} element sub-domains and the geometry and displacement field are approximated at discrete points of interest. The finite element interpolation at the contact interface is repeated here for convenience:

$$\boldsymbol{x}^{\mathrm{s}} \approx \left\{\boldsymbol{x}^{\mathrm{s}}\right\}^{\mathrm{h}}\Big|_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \equiv \sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\mathbf{x}_k^{\mathrm{s}}\,, \tag{5.12a}$$

$$\boldsymbol{u}^{\mathrm{s}} \approx \left\{\boldsymbol{u}^{\mathrm{s}}\right\}^{\mathrm{h}}\Big|_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \equiv \sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\mathbf{d}_k^{\mathrm{s}}\,. \tag{5.12b}$$

Here, n^{s} denotes the total number of nodes on discrete non-mortar subdomain $\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}$. The corresponding discrete nodal coordinates (in the current configuration) are represented by $\mathbf{x}_k^{\mathrm{s}}$ and the nodal displacement by $\mathbf{d}_k^{\mathrm{s}}$. The shape functions N_k^{s} are defined with respect to the associated finite element parameter space $\boldsymbol{\xi}^{\mathrm{s}}$.

One important notion to introduce at this stage is that the geometry of the rigid boundary $\Gamma_{\mathrm{c}}^{\mathrm{r}}$ is also discretised, i.e.

$$\boldsymbol{x}^{\mathrm{r}} \approx \left\{\boldsymbol{x}^{\mathrm{r}}\right\}^{\mathrm{h}} \equiv \sum_{l=1}^{n^{\mathrm{r}}} N_l^{\mathrm{r}}\left(\boldsymbol{\xi}^{\mathrm{r}}\right)\mathbf{x}_l^{\mathrm{r}}\,, \tag{5.13}$$

with n^{r} denoting the total number of nodes defining the rigid boundary. From a conceptual point of view, this step is not mandatory, as the rigid boundary could be equally defined by some analytical function. From a practical perspective, however, finite element interpolation is possibly more convenient, as it allows for the treatment of arbitrarily complex geometries found in engineering applications (which can be difficult to describe analytically). It also allows reusing several procedures and algorithms already established for unilateral contact, such as contact search and the numerical evaluation of mortar integrals. Notwithstanding, it should be kept in mind that all the techniques described in the following are still applicable for analytical representations of rigid boundaries (and, in some cases, even simplified).

5.3 Petrov-Galerkin approach to dual mortar contact

This work is entirely based on the so-called dual mortar method. It consists of defining dual shape functions to interpolate the Lagrange multipliers satisfying the bi-orthogonality condition, given by Equation (3.7). As already mentioned, this technique becomes particularly advantageous within contact problems, as it localizes the coupling conditions while preserving the optimal convergence of the discretisation error. The first mortar coupling matrix becomes diagonal and the contact constraints decouple to point-wise conditions, thus creating a perfect fit for the application of efficient active set strategies. However, when thinking about the requirements of the dual shape functions for contact mechanics, several aspects need to be carefully analysed. More specifically, the inequality nature of contact constraints requires positivity for the Lagrange multiplier basis functions.

First of all, it can be easily shown that dual Lagrange multiplier shape functions are guaranteed to satisfy *partition of unity* on each non-mortar element, i.e.,

$$\sum_{j=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \Phi_j = 1\,, \qquad j = 1,\ldots,n_{\mathrm{e}}^{\mathrm{s}}\,. \tag{5.14}$$

This property is assured by the bi-orthogonality condition given in Equation (3.8), see Flemisch and B. I. Wohlmuth (2007) for a proof. Moreover, the partition of unity yields another important feature of dual Lagrange multiplier basis:

$$\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}_{\mathrm{e}}^{\mathrm{h}}} \Phi_j \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} = \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}_{\mathrm{e}}^{\mathrm{h}}} N_j^{\mathrm{s}} \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,. \tag{5.15}$$

As mentioned below, this property plays a crucial role in ensuring the integral positivity of Lagrange multiplier interpolation.

While Equations (5.14) and (5.15) are proven properties of dual shape functions, when considering the transmission of contact stresses across the interface additional requirements are needed. First of all, in order to render the first mortar matrix non-singular, thus invertible, the condition of non-zero integrals arises, i.e.,

$$\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}_{\mathrm{e}}^{\mathrm{h}}} \Phi_j \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \neq 0\,. \tag{5.16}$$

While this requirement is sufficient for mesh tying applications, inequality constraints in contact problems require the Lagrange multiplier shape functions to satisfy at least *integral positivity*, i.e.,

$$\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}_{\mathrm{e}}^{\mathrm{h}}} \Phi_j \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} > 0\,. \tag{5.17}$$

This condition becomes necessary within contact modelling due to the physical interpretation of the KKT conditions. In their discrete form, the KKT conditions are written using the weighted nodal gap $\tilde{g}_j$, which is constructed based on the nodal shape functions used to interpolate the variation of the Lagrange multipliers, see Equation (3.29). Therefore, it turns out reasonable to require that arbitrary positive discrete gap functions, $g^{\mathrm{h}} > 0$, correspond to weighted gaps $\tilde{g}_j$ that are also positive—otherwise, it would lose its physical meaning. From a closer inspection of the definition of the weighted gap, Equation (3.29), it is possible to conclude that this is only satisfied if integral positivity of Lagrange multiplier shape functions is guaranteed, i.e., if Equation (5.17) holds. Moreover, the first mortar coupling matrix, is responsible for characterizing the Lagrange multiplier distribution as contact forces on the non-mortar boundary. Negative entries in this matrix can compromise the physical interpretation of the third KKT condition, and, in turn, positivity of D_{jk} is a critical assumption in the mathematical proof of optimal spatial convergence rates (B. I. Wohlmuth, Popp, *et al.*, 2012).

According to Equation (5.15), integral positivity is directly inherited from the corresponding displacement shape functions N_j^{s}. Unfortunately, while for first-order finite element interpolation, this property is fulfilled, for second-order approximation in three dimensions, this aspect reveals to be particularly troublesome. Integral positivity does not hold for specific elements there, e.g., corner nodes of eight-noded quadrilateral (quad8) or six-noded triangular (tri6) facets. For standard Lagrange multipliers, in Puso, Laursen, and Solberg (2008), this issue has been addressed either by choosing $n^{\lambda} < n^{\mathrm{s}}$ (i.e., only the corner nodes carry discrete Lagrange multipliers) or by employing *piecewise* linear polynomials on subsegments. In the dual Lagrange multiplier case, a solution based on special basis transformation procedures is proposed in Popp, B. I. Wohlmuth, *et al.* (2012) and B. I. Wohlmuth, Popp, *et al.* (2012). In this approach, a modified basis of interpolation functions based on a well-designed linear combination of the standard interpolation functions N_j^{s} replaces the latter in the bi-orthogonality condition to construct the corresponding dual shape functions.

5.3.1 Motivation for the Petrov-Galerkin technique

When dealing with contact problems under large deformations, integral positivity can be viewed as a *minimum* requirement for contact modelling. It assures convergence of the global active set algorithm at least under approximately constant gap conditions, but certainly not in every situation and most likely not for severe gradients found in coarse meshes with high curvatures. An illustrative two-dimensional example with first-order interpolation is represented in Figure 5.2, in which the sign of contributions to the weighted gap $\tilde{g}_j$ at a given non-mortar node j is highlighted. Even though there is no overlap between the two boundaries, node j will be erroneously identified as active. The reason for that has to do with the negative part of the dual shape function Φ_j, which yields for those regions weighted gap values $\tilde{g}_j$ with the opposite sign of the discrete gap g^{h}. Whenever severe geometrical curvatures are found, this effect can be amplified and, possibly, yield unphysical results that can compromise the convergence of the active set algorithm.

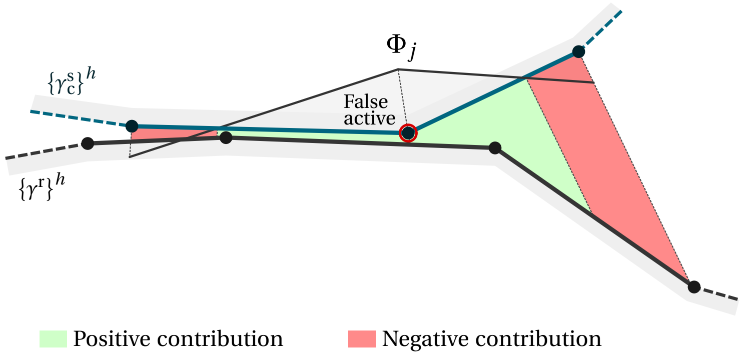


Figure 5.2: Illustration of a possible unphysical contact state in two-dimensions.

Since the severity of these artefacts is h-dependent, a possible solution to avoid these situations is the refinement of the finite element mesh. However, these options are not always available in practice, due to possible limitations in computational resources or even time to iterate on the numerical model. The ideal solution is a further restriction to *strictly positive* dual shape functions, i.e., by imposing

$$\Phi_j \geq 0\,, \tag{5.18}$$

since the possibility of obtaining unphysical contact states is entirely eliminated. However, in the case of strictly positive standard interpolation functions, as in two- and three-dimensional first-order elements, it is mathematically impossible to generate strictly positive dual shape functions. In the work by Popp, Seitz, *et al.* (2013), this problem is tackled by means of a Petrov-Galerkin technique. While in the standard Bubnov-Galerkin method, the dual shape functions Φ_j are used to interpolate *both* the Lagrange multipliers, $\boldsymbol{\lambda}^{\mathrm{h}}$, and their variation, $\delta\boldsymbol{\lambda}^{\mathrm{h}}$, the Petrov-Galerkin type uses different interpolations for each term. The Lagrange multipliers are discretised using the dual basis,

$$\boldsymbol{\lambda}^{\mathrm{h}} \equiv \sum_{j=1}^{n^{\mathrm{s}}} \Phi_j\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\mathbf{z}_j\,, \tag{5.19}$$

and their variation using the standard shape functions,

$$\delta\boldsymbol{\lambda}^{\mathrm{h}} \equiv \sum_{j=1}^{n^{\mathrm{s}}} N_j^{\mathrm{s}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\delta\mathbf{z}_j\,. \tag{5.20}$$

Under these circumstances, all the previously introduced discrete quantities still remain valid, except for the weighted gap, which now writes

$$\tilde{g}_j \equiv \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} N_j^{\mathrm{s}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right) g^{\mathrm{h}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}}\,. \tag{5.21}$$

Following this technique, the Lagrange multipliers can still be condensed from the global system of equations, while preserving the robustness of standard Lagrange multipliers for the treatment of the contact constraints. However, it should be pointed out that this approach also leads to a non-symmetric global system of equations (even though one could argue that from a physics-based perspective, symmetry is only achieved for frictionless contact anyway).

5.3.2 Additional advantages for rigid/deformable interaction

For the case of rigid/deformable contact, in particular, the application of the Petrov-Galerkin technique becomes extremely important for an additional reason: it allows us to completely eliminate the evaluation of the dual shape functions. As already mentioned, the second mortar coupling matrix vanishes due to the fact that one of the opposing contact surfaces remains fixed. Moreover, the bi-orthogonality condition leads to a diagonal first mortar coupling matrix, which can be evaluated using the standard shape functions only, see Equation (3.20). Therefore, the only terms left involving the dual shape functions are the contact constraints. As a consequence, and as explained in more detail in the following paragraphs, if the variation of the Lagrange multipliers is interpolated using standard shape functions, the dual shape functions are not explicitly used.

5.3.3 Extension to quadratic elements using piecewise linear interpolation

The establishment of quadratic dual shape functions in three dimensions has always been a challenging topic, especially for contact problems. Additional techniques are required for assuring integral positivity, which, as mentioned previously, is twofold: ensure optimal spatial convergence and physically meaningful weighted gaps. In this work, the locally quadratic technique proposed in Popp, B. I. Wohlmuth, *et al.* (2012) is employed to address the first point. Essentially, it consists of combining the bi-orthogonality condition, Equation (3.7), with a basis transformation procedure. As suggested in Flemisch and B. I. Wohlmuth (2007), feasible dual shape functions are constructed from the element-wise bi-orthogonality condition, although based on the introduction of modified shape functions $\tilde{N}_j^{\mathrm{s}}$, i.e.,

$$\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}_{\mathrm{e}}^{\mathrm{h}}} \Phi_j \tilde{N}_k^{\mathrm{s}} \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} = \delta_{jk} \int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}_{\mathrm{e}}^{\mathrm{h}}} \tilde{N}_k^{\mathrm{s}} \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \,. \tag{5.22}$$

It is noteworthy to mention that this technique leads to a non-diagonal first mortar coupling matrix that, notwithstanding, can still be trivially inverted due to the closed-form character of the transformation scheme. For additional details on this strategy, the reader is referred to the original publication in Popp, B. I. Wohlmuth, *et al.* (2012).

If strict positivity is to be further pursued, the situation becomes even more complicated. The original motivation for the Petrov-Galerkin technique relies on the preservation of the properties of standard shape functions in the contact constraints, while keeping the localization character of the dual basis in the coupling conditions. However, it is rather obvious that the requirement of non-negativity is only fully met by standard Lagrange multiplier basis functions for first-order finite element interpolation. For example, Figure 5.3 represents the integral value of the quadratic shape function associated with the first node of an 8-noded (serendipity) quadrilateral. The integration domain is considered rectangular, and, as can be observed in the contour plot, there is a region over which the integral value becomes negative. Within the context of the contact formulation, this compromises the physical interpretation of the weighed gap function and impairs the convergence of the active set search. Therefore, an additional modification of the interpolation scheme for the variation of the Lagrange multipliers is required for quadratic elements.

This topic has been firstly addressed in Sitzmann *et al.* (2016) within the context of dual mortar contact with regularisations. Similar to the concepts already introduced in Puso, Laursen, and Solberg (2008), an alternative is proposed to define Lagrange multipliers only at corner nodes, interpolating their variation using the associated first-order standard shape functions. Despite leading to a semi-smooth Newton method with a smaller active set to be iterated, this technique has the disadvantage of losing surface information for curved boundaries. The enforcement of contact constraints at edge nodes is ignored and, for coarse meshes, this can lead to unphysical results. Furthermore, as this approach inherently relies on $n^{\lambda} < n^{\mathrm{s}}$, the construction of the dual basis is slightly more involved and requires substantial algorithmic adaptations, see Popp, B. I. Wohlmuth, *et al.*, 2012.

In this work, an alternative technique for quadratic Petrov-Galerkin dual mortar contact is proposed. In the spirit of the concepts presented in Puso, Laursen, and Solberg (2008) for the evaluation of the mortar integrals for quadratic elements, it is based on the establishment of linear sub-elements and corresponding piecewise interpolation. The

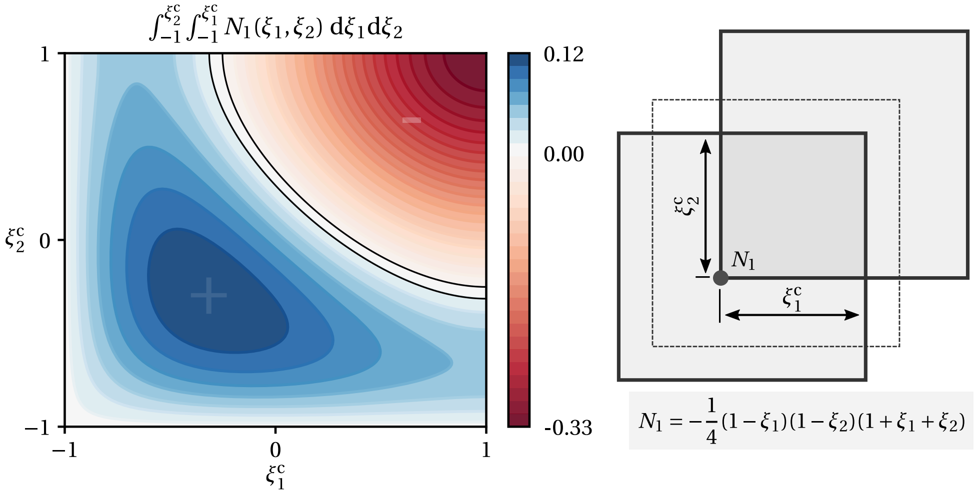


Figure 5.3: The integral value of the standard shape function node 1, N_1, of the 8-noded quadratic element over partial rectangular domains.

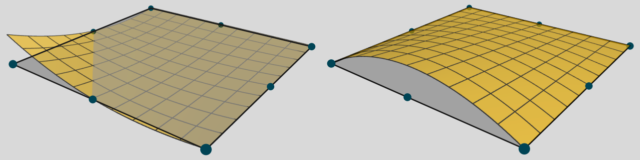
Lagrange multipliers are approximated using the quadratic dual shape functions, while the piecewise linear standard shape functions are used for its variation, i.e.,

$$\delta\boldsymbol{\lambda}^{\mathrm{h}} \equiv \sum_{j=1}^{n^{\mathrm{s}}} N_j^{\mathrm{sub}}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\delta\mathbf{z}_j \,. \tag{5.23}$$

Compared with the alternative proposed by Sitzmann *et al.* (2016), all non-mortar nodes are checked for contact, thus preserving more information about the interface geometry. As the condition $n^{\lambda} = n^{\mathrm{s}}$ is retained, the implementation of the proposed approach requires little effort and enables recycling of most of the pre-existing algorithmic components. In combination with the certainty of strict positivity of the piecewise linear shape functions, this allows for a robust algorithm that becomes less sensitive to discretization problems. However, as a trade-off, the computational complexity in the numerical evaluation of mortar integrals increases. Instead of approximating a single linear element (containing only the corner nodes), each quadratic facet is divided into multiple linear/bilinear sub-elements, which may require further sub-divisions within the clipping polygon algorithm. This downside, however, is not intrinsically related to the proposed approach but is already present in popular mortar segmentation procedures for quadratic elements in deformable/deformable contact (Puso, Laursen, and Solberg, 2008). Notwithstanding, for rigid/deformable contact, this aspect can be counterbalanced by employing the efficient projected orthonormal frame to be described in Section 5.6.

The application of the piecewise linear interpolation requires the establishment of proper mappings, referring to Appendix B for their explicit expressions and associated Jacobian matrices. Figure 5.4 gives an overview of the various finite element interpolation schemes used within the Petrov-Galerkin dual mortar contact algorithm, also representing the shape functions characteristic of both corner and edge nodes of a quadratic 8-noded quadrilateral (quad8).

(a) Standard shape functions



Geometry

$$\boldsymbol{x}^{\mathrm{r}} \approx \{\boldsymbol{x}^{\mathrm{r}}\}^{\mathrm{h}} \equiv \textstyle\sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\,\mathbf{x}_k^{\mathrm{s}}$$

$$\boldsymbol{x}^{\mathrm{s}} \approx \{\boldsymbol{x}^{\mathrm{s}}\}^{\mathrm{h}} \equiv \textstyle\sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\,\mathbf{x}_k^{\mathrm{s}}$$

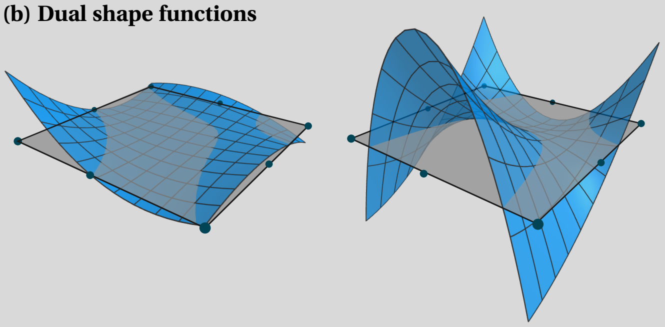
Displacements

$$\boldsymbol{u}^{\mathrm{s}} \approx \{\boldsymbol{u}^{\mathrm{s}}\}^{\mathrm{h}} \equiv \textstyle\sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\,\mathbf{d}_k^{\mathrm{s}}$$

Displacements variation

$$\delta\boldsymbol{u}^{\mathrm{s}} \approx \{\delta\boldsymbol{u}^{\mathrm{s}}\}^{\mathrm{h}} \equiv \textstyle\sum_{k=1}^{n^{\mathrm{s}}} N_k^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\,\delta\mathbf{d}_k^{\mathrm{s}}$$

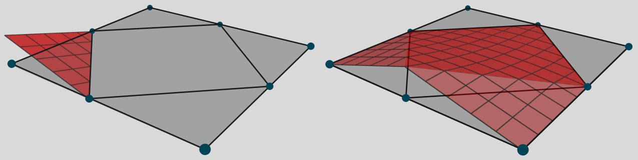
(b) Dual shape functions



Lagrange multipliers

$$\boldsymbol{\lambda} \approx \boldsymbol{\lambda}^{\mathrm{h}} \equiv \textstyle\sum_{j=1}^{n^{\mathrm{s}}} \Phi_j(\boldsymbol{\xi}^{\mathrm{s}})\,\mathbf{z}_j$$

(c) Piecewise linear shape functions



Lagrange multipliers variation

$$\delta\boldsymbol{\lambda} \approx \delta\boldsymbol{\lambda}^{\mathrm{h}} \equiv \textstyle\sum_{j=1}^{n^{\mathrm{s}}} N_j^{\mathrm{sub}}(\boldsymbol{\xi}^{\mathrm{s}})\,\delta\mathbf{z}_j$$

Figure 5.4: Finite element interpolations for rigid/deformable Petrov-Galerkin dual mortar contact and shape functions for a quadratic 8-noded quadrilateral.

Remark. *From a mathematical point of view, the Petrov-Galerkin technique introduces modifications into the variational structure of the problem. As demonstrated in Hüeber (2008), decoupling of the contact constraints into point-wise conditions is only possible due to the localized character of the dual Lagrange multipliers. Therefore, only for the Bubnov-Galerkin approach, the bi-orthogonality condition is fulfilled and, consequently, the contact conditions are localized. In the Petrov-Galerkin technique, however, the discrete dual space is switched to the trace space of the displacements for the variation of the Lagrange multipliers. In doing so, it should be expected that the localization property will be affected and additional coupling terms will originate. Nevertheless, the transformation of weak contact constraints to a set of point-wise conditions has been employed extensively in the literature, e.g., see Puso and Laursen (2004a,b), Puso, Laursen, and Solberg (2008), Popp, Seitz,* et al. *(2013), and Sitzmann* et al. *(2016). In analogy to mass lumping techniques, it can be interpreted as a form of lumping of the contact constraints (Sitzmann* et al.*, 2016). For an exemplary formulation considering fully coupled constraints, see Blum* et al. *(2016). The assessment of the impact of the Petrov-Galerkin approach on the mathematical structure of the method would fall beyond the scope of the present work and, therefore, this topic is addressed from a practical perspective in the numerical investigations presented in Section 5.9.*

5.4 Discrete contact virtual work

Based upon the finite element interpolation scheme introduced in the previous section, the discretised version of the contact virtual work $\delta\Pi_{\mathrm{c}}$, given in Equation (5.8), can be written as

$$\delta\Pi_{\mathrm{c}}^{\mathrm{s}} \approx \{\Pi_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}} \equiv \sum_{j=1}^{n^{\mathrm{s}}} \sum_{k=1}^{n^{\mathrm{s}}} \left\{ \mathbf{z}_j \left[\int_{\{\gamma_{\mathrm{c}}^{\mathrm{s}}\}^{\mathrm{h}}} \Phi_j(\boldsymbol{\xi}^{\mathrm{s}}) N_k^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) \,\mathrm{d}\gamma_{\mathrm{c}}^{\mathrm{s}} \right] \delta\mathbf{d}_k^{\mathrm{s}} \right\} . \tag{5.24}$$

It should be noticed that the contact virtual work is entirely based on the first mortar coupling matrix. In contrast to the mortar boundary in unilateral contact problems, the rigid side remains completely prescribed (fixed), it does not contribute to the virtual work by the contact forces—the second mortar matrix is not required.

5.5 Discrete contact constraints

The treatment of the contact constraints starts by the definition of the discrete version of the gap vector, $\boldsymbol{g}^{\mathrm{h}}(\boldsymbol{\xi}^{\mathrm{s}})$, which for rigid/deformable contact reads

$$\boldsymbol{g} \approx \boldsymbol{g}^{\mathrm{h}}(\boldsymbol{\xi}^{\mathrm{s}}) \equiv \sum_{j=1}^{n^{\mathrm{s}}} N_j^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\mathbf{x}_j^{\mathrm{s}} - \sum_{l=1}^{n^{\mathrm{r}}} N_l^{\mathrm{r}}(\hat{\boldsymbol{\xi}}^{\mathrm{r}})\mathbf{x}_l^{\mathrm{r}} . \tag{5.25}$$

When compared with the general case of unilateral contact introduced in Chapter 3, they both have the same structure. However, because the derivative of the rigid surface coordinates vanishes, its linearisation is simplified.

As a final remark, it is worth mentioning that if the rigid boundary geometry is discretised using finite elements, the projection operations can be performed by employing the techniques already described for unilateral contact. These typically include an efficient global search algorithm and a continuous field of normals associated with the non-mortar boundary, see Section 5.6.

5.5.1 Discrete contact constraints in the normal direction

The discrete point-wise conditions (3.66) described in Chapter 3 remain unchanged, except for the weighted gap $\tilde{g}_j$ definition, which by employing the Petrov-Galerkin approach is interpolated using the standard shape functions, viz.

$$\tilde{g}_j \equiv \int_{\{\gamma_\mathrm{c}^\mathrm{s}\}^\mathrm{h}} N_j^\mathrm{s}\left(\boldsymbol{\xi}^\mathrm{s}\right) g^\mathrm{h}\left(\boldsymbol{\xi}^\mathrm{s}\right) \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} \,. \tag{5.26}$$

The discrete gap function g^h is still evaluated as

$$g^\mathrm{h}\left(\boldsymbol{\xi}^\mathrm{s}\right) \equiv \boldsymbol{\eta}^\mathrm{h}\left(\boldsymbol{\xi}^\mathrm{s}\right) \cdot \mathbf{g}^\mathrm{h}\left(\boldsymbol{\xi}^\mathrm{s}\right) \,, \tag{5.27}$$

yet considering the discrete gap vector in Equation (5.25).

For quadratic finite elements, the piecewise linear interpolation is employed, i.e.,

$$\tilde{g}_j \equiv \int_{\{\gamma_\mathrm{c}^\mathrm{s}\}^\mathrm{h}} N_j^\mathrm{sub}(\boldsymbol{\xi}^\mathrm{sub}) g^\mathrm{h}(\boldsymbol{\xi}^\mathrm{s}(\boldsymbol{\xi}^\mathrm{sub})) \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} \,. \tag{5.28}$$

It is worth mentioning that the discrete gap is still defined based on the quadratic shape functions, with coordinates obtained using the mappings given in Appendix B. This allows to retain the quadratic interpolation of the geometry and, therefore, capture smooth curvatures in the evaluation of the discrete gap function.

5.5.2 Discrete contact constraints in the tangential direction

In the same way, both the point-wise tangential conditions (3.67) and all the concepts regarding the backward Euler scheme for time integration remain consistent. The only difference regards the evaluation of the incremental slip, which now must be derived from the modified weighted gap vector, $\tilde{\mathbf{g}}_j\left(\boldsymbol{\xi}^\mathrm{s}\right)$, which reads

$$\begin{aligned} \tilde{\mathbf{g}}_j\left(\boldsymbol{\xi}^\mathrm{s}\right) \equiv & \sum_{k=1}^{n^\mathrm{s}} \left[\int_{\{\gamma_\mathrm{c}^\mathrm{s}\}^\mathrm{h}} N_j^\mathrm{s}\left(\boldsymbol{\xi}^\mathrm{s}\right) N_k^\mathrm{s}(\boldsymbol{\xi}^\mathrm{s}) \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} \right] \mathbf{x}_k^\mathrm{s} \\ & - \sum_{l=1}^{n^\mathrm{r}} \left[\int_{\{\gamma_\mathrm{c}^\mathrm{s}\}^\mathrm{h}} N_j^\mathrm{s}\left(\boldsymbol{\xi}^\mathrm{s}\right) N_l^\mathrm{r}(\hat{\boldsymbol{\xi}}^\mathrm{r}) \mathrm{d}\gamma_\mathrm{c}^\mathrm{s} \right] \mathbf{x}_l^\mathrm{r} \,. \end{aligned} \tag{5.29}$$

5.6 Projected averaged moving frame

With the discrete version of the contact virtual work and contact constraints thought out, attention is now shifted towards their computational treatment. The importance of mortar integral evaluation within mortar methods is well-known and represents one of the main challenges. Their correct evaluation is essential to preserve the sough-after properties of the mortar-based variational formulation. It requires approximating surface integrals with complex geometrical operations involved and, therefore, ends up being one of the main contributors to the overall cost of the algorithm. Therefore, there is a clear motivation for new strategies that are able to facilitate the complexity of the algorithm without affecting its accuracy and robustness. For the particular case at hand of rigid/deformable contact under large deformations, the fact that one of the boundaries remains fixed can be exploited and, as explained in the following paragraphs, a new definition for the continuous field of orthonormal frames is proposed.

In all contact problems, it is necessary to define a local orthonormal moving frame attached to each contact node containing a Lagrange multiplier. It splits the surface contact tractions into normal and tangential components and establishes projection rules necessary for the evaluation of mortar coupling terms. The overall idea of the proposed method consists in defining an initial field of orthonormal frames on the rigid side, which is then continuously projected onto the non-mortar side throughout the deformation process. On the ground foundation of this idea is the fact that, while contact occurs, the boundaries of both sides tend to coincide and, for the regions in full contact, become practically identical. Therefore, the field of orthonormal frames on the rigid side can be defined using sophisticated and accurate methodologies, which are then transmitted to the deformable side by means of simple projections rules, similar to the ones already used within the integration algorithm. The individual steps are described in more detail in the following.

Firstly, it is necessary to define the field of orthonormal frames on the rigid side. This operation is only realised once during problem initialisation and the averaged normal approach described in Chapter 3 is employed here. Originally proposed by B. Yang, Laursen, and Meng (2005), and later on slightly simplified by Popp, Gee, *et al.* (2009), it is based on a continuous field of normal vectors defined on the non-mortar side, which smoothens the discontinuities associated with the discretization of the contact interface by averaging the nodal unit normals.

Next, the unit normal vector is projected to the non-mortar contact boundary and inverted in order to still point outwards, see Figure 5.5. For a given non-mortar node j with coordinates $\mathbf{x}_j^\mathrm{s}$, the projection consists of finding the isoparametric coordinate $\hat{\boldsymbol{\xi}}^\mathrm{r}$ on the rigid side such that the following condition holds:

$$\sum_{k=1}^{n_\mathrm{e}^\mathrm{r}} N_k^\mathrm{r}(\hat{\boldsymbol{\xi}}^\mathrm{r})\,[\mathbf{x}_k^\mathrm{r} + \alpha\tilde{\boldsymbol{\eta}}_k^\mathrm{r}] - \mathbf{x}_j^\mathrm{s} = \mathbf{0}\,. \tag{5.30}$$

Here, n_e^r denotes the total number of nodes of the rigid element and $\tilde{\boldsymbol{\eta}}_k^\mathrm{r}$ the averaged unit normal vectors. The parameter α relates to the normal distance between the points.* This system of equations can be solved with a local Newton-Raphson procedure, where each iteration reads

$$\begin{Bmatrix} \Delta\hat{\xi}_1^\mathrm{r} \\ \Delta\hat{\xi}_2^\mathrm{r} \\ \Delta\alpha \end{Bmatrix} = [\mathbf{W}]^{-1} \left\{ \sum_{l=1}^{n_\mathrm{e}^\mathrm{r}} N_l^\mathrm{r}(\hat{\boldsymbol{\xi}}^\mathrm{r}) \left(\mathbf{x}_l^\mathrm{r} + \alpha\tilde{\boldsymbol{\eta}}_k^\mathrm{r}\right) - \mathbf{x}_j^\mathrm{s} \right\}. \tag{5.31}$$

The matrix $[\mathbf{W}] \in \mathbb{R}^{d\times d}$ is obtained from the derivative of the projection condition with

---

*For the sake of simplicity, the interpolated normal vector is not normalized. This does not change the solution of the projection procedure, thus affecting the physical meaning of the parameter α only.

respect to the rigid side coordinate $\boldsymbol{\xi}^{\mathrm{r}}$ and the parameter α. In three dimensions, it yields

$$[\mathbf{W}] \equiv \left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N^{\mathrm{r}}_{l,\xi^{\mathrm{r}}_1} \left(\mathbf{x}^{\mathrm{r}}_l + \tilde{\boldsymbol{\eta}}_l \right) \;\middle|\; \sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N^{\mathrm{r}}_{l,\xi^{\mathrm{r}}_2} \left(\mathbf{x}^{\mathrm{r}}_l + \tilde{\boldsymbol{\eta}}_l \right) \;\middle|\; \sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N^{\mathrm{r}}_l \tilde{\boldsymbol{\eta}}_l \right] . \tag{5.32}$$

The two-dimensional version is relatively straightforward by simply omitting the second column. Lastly, with the projection coordinate $\hat{\boldsymbol{\xi}}^{\mathrm{r}}$ at hand, the associated frame can be inverted, which means that the unit normal vector $\boldsymbol{\eta}^{\mathrm{h}}_j$ at the non-mortar node j yields

$$\boldsymbol{\eta}^{\mathrm{h}}_j = - \frac{\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N^{\mathrm{r}}_l \left(\hat{\boldsymbol{\xi}}^{\mathrm{r}} \right) \tilde{\boldsymbol{\eta}}^{\mathrm{r}}_k}{\left\| \sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N^{\mathrm{r}}_l \left(\hat{\boldsymbol{\xi}}^{\mathrm{r}} \right) \tilde{\boldsymbol{\eta}}^{\mathrm{r}}_k \right\|} . \tag{5.33}$$

Having the unit normal vector at hand, the tangent vector can be freely chosen from the tangential plane. A recommended technique is, for example, considering the direction of the interpolated tangent vector on the rigid side.

In practice, the projection is performed considering various rigid interface elements until a valid solution is found. The application of an efficient global search algorithm (see Section 3.6) is, thus, highly recommended in order to perform this iterative procedure based on a reduced list of target elements. The averaged frame can still be employed in situations where no valid projection is found (e.g., in dropping edge problems), as it depends only on the deformable boundary itself.

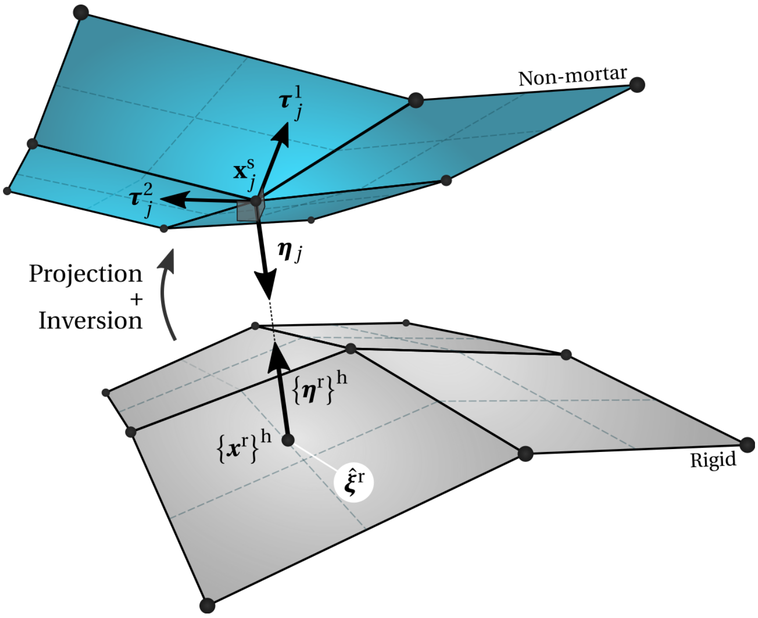


Figure 5.5: Projection of the averaged orthonormal frame from the rigid onto the non-mortar boundary.

The local frame is deformation-dependent for contact under large deformations, and, therefore, needs to be linearised within the Newton-Raphson algorithm. In fact, this now exposes the main advantage of this technique, which regards the self-contained character of the projection procedure. As demonstrated in more detail in Section 5.8.1, the projected frame derivative is guaranteed to have the minimum bandwidth, i.e., it contains only the degrees of freedom associated with the non-mortar node itself. This has a beneficial impact on the computational complexity of the algorithm by reducing the total number of individual operations. As the derivatives of the normal and tangent vectors appear (directly and indirectly) in every term of the formulation, the reduction in the derivatives' number of terms becomes amplified in the overall computational cost. Recall that, beyond the contact constraints, the integration scheme is based on projections that also use the continuous field of normals on the non-mortar side. The improvements in computational complexity are carefully investigated and quantitatively measured in the numerical examples shown in Section 5.9.2.

5.7 Numerical evaluation of mortar integrals

The prime cause for the difficulty in evaluating mortar integrals is related to the quantities with terms belonging to different sides of the contact interface—usually connected by means of projection rules that, within the ongoing formulation, are based on the projected frame. Generally speaking, these include the transmission of contact stresses across the interface, described by the second mortar coupling matrix, and the kinematics describing the relative motion between both boundaries, namely in the form of the weighted gap and the incremental slip. One of the main simplifications for the particular case of rigid/deformable contact is that the second mortar matrix vanishes. However, even in the simplest case of frictionless contact, the weighted gap still needs to be evaluated and, therefore, the challenge of correctly evaluating mortar integrals is still present.

The primary source of complexity involves the transmission of geometrical information between boundaries (to determine the overlap of both domains) and, in the three-dimensional case, the evaluation of surface integrals with complex geometries. As the rigid boundary is interpolated using finite elements, the techniques described in the overview work by Farah, Popp, *et al.* (2015) are equally applicable to the problem at hand. Both the segmentation and element-wise integration schemes remain practically unchanged, with the only exception being the displacements of the mortar side vanishing. Nonetheless, for quadratic interpolation using piecewise linear interpolation, the understanding of how the subdivision of the interface element into multiple sub-elements affects the integration scheme is noteworthy to mention. In the spirit of preserving all the techniques derived for linear elements, each individual sub-element of the parent quadratic facet is treated accordingly. The two strategies employed for the numerical approximation of mortar integrals are described in the following.

5.7.1 Evaluation of mortar integrals for piecewise linear interpolation

For the segmentation strategy in two dimensions, segments are formed using each pair of nodes of the involved 3-noded line elements, while in three dimensions, the clipping polygon is established using the sub-elements (3-noded triangles or 4-noded quadrilaterals). Because the clipping polygon technique for three-dimensional problems is only valid for linear facets (otherwise, it would be impractical to perform the clipping algorithm based on curved domains), this means that the quadratic facets on the rigid side also need to be divided into sub-elements. The basic steps of the algorithm remain the same: after defining the clipping polygon, it is divided into multiple cells to be numerically integrated (triangles and quadrilaterals) and the Gauss points are projected back to the sub-element. At this stage, however, an additional step needs to be performed by recovering the original isoparametric coordinate at the parent element (quadratic) using the mappings described in Appendix B. The contribution of a given pair of non-mortar and rigid elements, for instance, to the first mortar coupling matrix, becomes

$$\mathrm{D}_{jj}^{(\mathrm{s,r})} \approx \sum_{s=1}^{n_{\mathrm{sub}}} \sum_{c=1}^{n_{\mathrm{cell}}} \sum_{\mathrm{g}=1}^{n_{\mathrm{g}}} w_{\mathrm{g}}\, N_j\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{sub}}(\boldsymbol{\zeta}_{\mathrm{g}})\big)\big) J_c^{\mathrm{cell}}\,, \tag{5.34}$$

where n_{sub} stands for the total number of sub-elements and n_{cell} is the total number of integration cells. The Gaussian quadrature is defined over n_{g} integration points with w_{g} weights and coordinates $\boldsymbol{\zeta}_{\mathrm{g}}$. The Jacobian determinant, J_c, defines the transformation from the integration cell of the sub-element to the global spatial configuration, i.e.,

$$J_c^{\mathrm{cell}}(\boldsymbol{\xi}^{\mathrm{sub}}) = \left\| \frac{\partial \{\boldsymbol{x}^{\mathrm{s}}\}^{\mathrm{h}}}{\partial \boldsymbol{\xi}^{\mathrm{s}}} \right\| \cdot \left\| \frac{\partial \boldsymbol{\xi}^{\mathrm{s}}}{\partial \boldsymbol{\xi}^{\mathrm{sub}}} \right\| . \tag{5.35}$$

For the element-based integration, the Gauss points are defined on the sub-element and directly projected to the opposite side. Figure 5.6 schematically represents the boundary-segmentation method in three dimensions, which is based on the combination of both strategies. The algorithm is illustrated for a pair of non-mortar and rigid elements, which, after division into sub-elements, yields both types of integration cells. For the sub-element represented in red, all the Gauss points are successfully projected, whereas the remaining sub-elements require segmentation.

It is important to mention that, from a practical perspective, there exists a slight difference in both strategies that motivates a modified physical interpretation of the Jacobian determinant. For element-wise integration, all the nodes of the integration cell belong to the parent finite element and, therefore, the mapping to the parent element is explicitly employed and the Jacobian evaluated using the two terms in Equation (5.35). However, when applying the segmentation scheme, the sub-element is projected to the auxiliary plane, which leads to a relative loss of geometrical information (interpreted as a kind of faceting of the quadratic element).[†] The nodes of the integration cells become either projections of the sub-element nodes to the auxiliary plane or nodes generated by the clipping algorithm. The Jacobian determinant is evaluated directly by using the global coordinates of the integration cell nodes and, thus, overlooking the two-step operation in Equation (5.35).

As a final remark, it is typical to address conservation properties of the spatially discrete problem in mortar contact formulations, especially concerning linear momentum conservation. In deformable/deformable contact settings, this is trivially ensured by adopting the same quadrature rules for both mortar coupling terms. In contrast, in the current scenario dealing with rigid/deformable contact, linear momentum conservation is not affected by the occurrence of contact and, thus, this topic is not relevant here. For the sake of simplicity, the same integration scheme is employed for every mortar integral, with respect to the first mortar matrix and weighting gap.

[†] Strictly speaking, this becomes relevant only for 4-noded quadrilateral facets, in which the possibility of element warping exists.

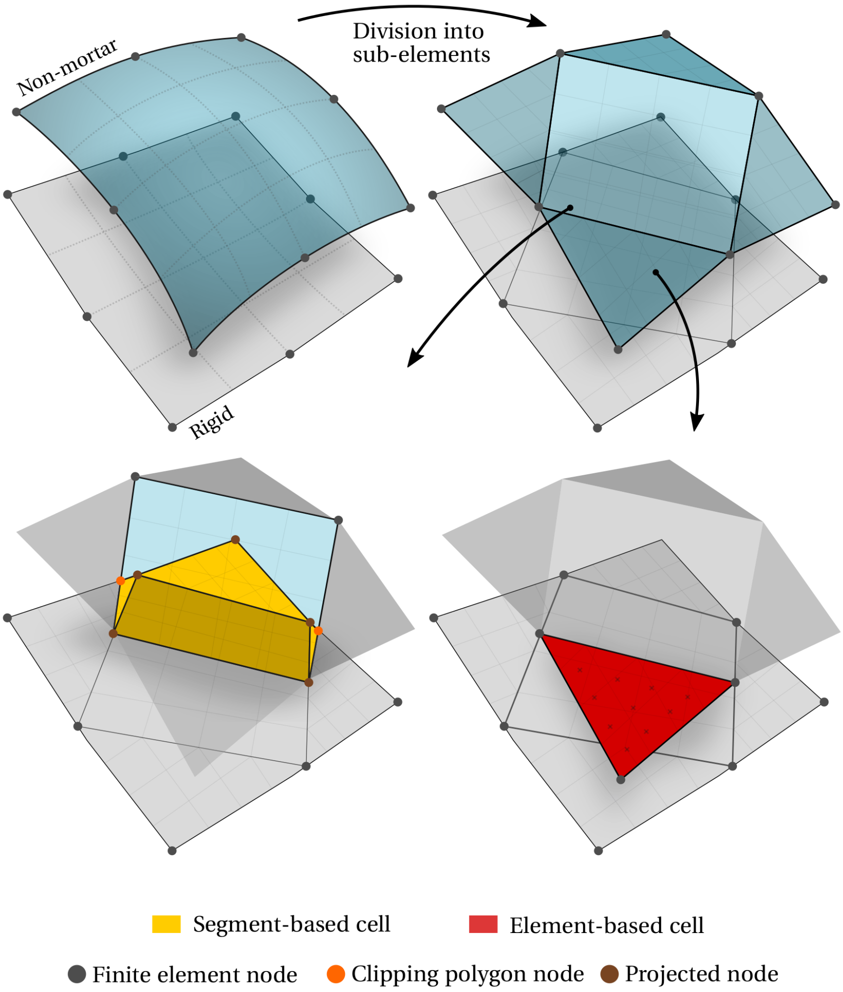


Figure 5.6: Schematic illustration of the boundary-segmentation integration method for the piecewise linear interpolation in three dimensions. The top two figures represent the division into sub-elements and the bottom two figures correspond to the integration cells originated from two of the five sub-elements.

5.8 Global solution algorithm

The final step to obtain the final discrete non-linear contact problem between a deformable body and a rigid obstacle regards the active set strategy. As for the general case of unilateral contact between two deformable bodies, the contact inequality constraints require identifying the subset of interface nodes effectively in contact, as well as the necessary tangential contact states (stick/slip). To address this problem, the PDASS described in Chapter 3 is thoroughly applied without any modifications. In a nutshell, it consists of reformulating the discrete nodal inequality constraints using NCP functions. This introduces a certain regularisation on the active set search and enables the application of a Newton-Raphson type algorithm, comprising not only geometrical and material non-linearities, but also the contact active set search itself. After all, the final discrete contact problem can be written as the entire set of equality conditions, which is repeated here for convenience:

$$\begin{aligned} \mathbf{r}(\mathbf{d},\mathbf{z}) \equiv \mathbf{f}_{\text{int}}(\mathbf{d}) - \mathbf{f}_{\text{ext}} + \mathbf{f}_{\text{c}}(\mathbf{d},\mathbf{z}) = \mathbf{0}\,, \\ C_j^{\eta}\left(\mathbf{d},\mathbf{z}_j\right) = 0\,, \qquad \forall\, j \in \mathcal{S}\,, \\ \boldsymbol{C}_j^{\tau}\left(\mathbf{d},\mathbf{z}_j\right) = \mathbf{0}\,, \qquad \forall\, j \in \mathcal{S}\,. \end{aligned}$$

Recall that $\mathbf{r}(\mathbf{d},\mathbf{z})$ represents the residual vector, $\mathbf{f}_{\text{int}}(\mathbf{d})$ and $\mathbf{f}_{\text{ext}}$ the internal and external forces vectors and $\mathbf{f}_{\text{c}}(\mathbf{d},\mathbf{z})$ for the discrete contact forces. At this stage, the foundations for the application of a Newton-type algorithm are complete. Thus, we describe in the following the consistent linearisation of the problem.

5.8.1 Consistent linearisation

The application of the semi-smooth Newton algorithm requires consistent linearisation of both the discrete balance equations and the NCP functions. In this section, the focus is exclusively placed on the terms introduced within the current formulation for non-linear rigid/deformable contact, namely the piecewise linear interpolation for quadratic dual mortar contact and the projected orthonormal frame. The remaining derivations remain unchanged from the unilateral contact case and, therefore, the reader is referred to Chapter 3 and Appendix A and the discussions in Puso and Laursen, 2004a; Popp, Gee, *et al.*, 2009; Popp, Gitterle, *et al.*, 2010, to name a few.

5.8.1.1 Piecewise linear interpolation

The piecewise linear interpolation affects the definition of the isoparametric coordinates at each Gauss point. As the interface element is divided into sub-elements, the numerical integration requires the application of mappings between domains. These need to be taken into account within the derivative chain rule and, for example, the derivative of the isoparametric non-mortar coordinate, $\mathscr{D}\boldsymbol{\xi}^{\text{s}}$, yields

$$\mathscr{D}\boldsymbol{\xi}^{\text{s}} = \frac{\partial \boldsymbol{\xi}^{\text{s}}}{\partial \boldsymbol{\xi}^{\text{sub}}}\, \mathscr{D}\boldsymbol{\xi}^{\text{sub}}\,. \tag{5.36}$$

The first term regards the Jacobian matrix of the mappings, referring to Appendix B for its definition. The derivative $\mathscr{D}\boldsymbol{\xi}^{\text{sub}}$ contains the directional derivative of the sub-element parameter space, which is computed using the typical first-order element procedures.

5.8.1.2 Projected orthonormal frame

Considering the definition of the projected unit normal given in Equation (5.33), its linearisation reads

$$\mathscr{D}\boldsymbol{\eta}_j^{\mathrm{h}} = \left(\frac{1}{l_{\check{\eta}^{\mathrm{r}}}} \boldsymbol{I} - \frac{1}{l_{\check{\eta}^{\mathrm{r}}}^3} \check{\boldsymbol{\eta}}^{\mathrm{r}} \otimes \check{\boldsymbol{\eta}}^{\mathrm{r}} \right) \mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{r}} \,. \tag{5.37}$$

Here, $\check{\boldsymbol{\eta}}_j^{\mathrm{r}}(\hat{\boldsymbol{\xi}}^{\mathrm{r}})$ has been introduced as an abbreviation for the non-unit normal vector on the rigid boundary,

$$\check{\boldsymbol{\eta}}_j^{\mathrm{r}}(\hat{\boldsymbol{\xi}}^{\mathrm{r}}) \equiv \sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N_l^{\mathrm{r}}(\hat{\boldsymbol{\xi}}^{\mathrm{r}}) \tilde{\boldsymbol{\eta}}_k^{\mathrm{r}} \,, \tag{5.38}$$

whose derivative yields

$$\mathscr{D}\check{\boldsymbol{\eta}}_j^{\mathrm{r}}(\hat{\boldsymbol{\xi}}^{\mathrm{r}}) = \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N_{k,\xi_1^{\mathrm{r}}}^{\mathrm{r}}(\hat{\boldsymbol{\xi}}^{\mathrm{r}}) \, \mathscr{D}\hat{\xi}_1^{\mathrm{r}} \, \tilde{\boldsymbol{\eta}}_k^{\mathrm{r}} + \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{r}}} N_{k,\xi_2^{\mathrm{r}}}^{\mathrm{r}}(\hat{\boldsymbol{\xi}}^{\mathrm{r}}) \, \mathscr{D}\hat{\xi}_2^{\mathrm{r}} \, \tilde{\boldsymbol{\eta}}_k^{\mathrm{r}} \,. \tag{5.39}$$

Note that the only derivative needed arises from the projection procedure, given that the averaged unit normal linearisation vanishes (the rigid boundary remains fixed). The linearisation of the projected rigid isoparametric coordinate $\hat{\boldsymbol{\xi}}^{\mathrm{r}}$ can be extracted from the projection condition in Equation (5.31) as follows

$$\begin{Bmatrix} \mathscr{D}\hat{\xi}_1^{\mathrm{r}} \\ \mathscr{D}\hat{\xi}_2^{\mathrm{r}} \\ \mathscr{D}\alpha \end{Bmatrix} = [\mathbf{W}]^{-1} \{\Delta \mathbf{x}_j^{\mathrm{s}}\} \,. \tag{5.40}$$

It is noteworthy to mention that the matrix [**W**] to be inverted is already computed during the projection, see Equation (5.32), which makes the computational evaluation of the orthonormal moving frame relatively straightforward and efficient.

5.8.2 Algebraic representation

This last section provides the algebraic representation of the discrete entities involved in the contact algorithm. As the approach adopted in Chapter 3, all nodes and corresponding degrees of freedom are partitioned into two (instead of three) disjoint sets $\mathcal{S} \cup \mathcal{N}$: a group $\mathcal{S}$ containing all non-mortar quantities and a group $\mathcal{N}$ associated with all remaining nodes or degrees of freedom. Then, the non-mortar set $\mathcal{S}$ is further partitioned into three disjoint sets: the inactive nodes set $\mathcal{I}$, the set $\mathcal{Q}$ of nodes in contact and stick state and the set $\mathcal{L}$ of nodes in contact and slip condition. The assembled system to be solved within each semi-smooth Newton step k in order to obtain the incremental displacements vector $\Delta\mathbf{d}$ and current Lagrange multipliers $\mathbf{z}^{k+1}$ can be expressed as:

$$\begin{bmatrix} \mathbf{K}_{\mathcal{N}\mathcal{N}} & \mathbf{K}_{\mathcal{N}\mathcal{I}} & \mathbf{K}_{\mathcal{N}\mathcal{Q}} & \mathbf{K}_{\mathcal{N}\mathcal{L}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_{\mathcal{I}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{I}\mathcal{L}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{I}\mathcal{I}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{I}\mathcal{Q}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{I}\mathcal{L}} \\ \mathbf{K}_{\mathcal{Q}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{Q}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{Q}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{Q}\mathcal{L}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{I}\mathcal{Q}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{Q}\mathcal{Q}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{Q}\mathcal{L}} \\ \mathbf{K}_{\mathcal{L}\mathcal{N}} & \widetilde{\mathbf{K}}_{\mathcal{L}\mathcal{I}} & \widetilde{\mathbf{K}}_{\mathcal{L}\mathcal{Q}} & \widetilde{\mathbf{K}}_{\mathcal{L}\mathcal{L}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{L}\mathcal{I}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{L}\mathcal{Q}} & \mathbf{D}^{\mathrm{T}}_{\mathcal{L}\mathcal{L}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{\mathcal{I}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{\mathcal{I}} & \mathbf{A}_{\mathcal{Q}} & \mathbf{A}_{\mathcal{L}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\mathcal{I}} & \mathbf{Q}_{\mathcal{Q}} & \mathbf{Q}_{\mathcal{L}} & \mathbf{0} & \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{\mathcal{I}} & \mathbf{L}_{\mathcal{Q}} & \mathbf{L}_{\mathcal{L}} & \mathbf{0} & \mathbf{0} & \mathbf{H} \end{bmatrix} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{N}} \\ \Delta\mathbf{d}_{\mathcal{I}} \\ \Delta\mathbf{d}_{\mathcal{Q}} \\ \Delta\mathbf{d}_{\mathcal{L}} \\ \mathbf{z}^{k+1}_{\mathcal{I}} \\ \mathbf{z}^{k+1}_{\mathcal{Q}} \\ \mathbf{z}^{k+1}_{\mathcal{L}} \end{Bmatrix} = - \begin{Bmatrix} \mathbf{r}_{\mathcal{N}} \\ \widetilde{\mathbf{r}}_{\mathcal{I}} \\ \widetilde{\mathbf{r}}_{\mathcal{Q}} \\ \widetilde{\mathbf{r}}_{\mathcal{L}} \\ \mathbf{0} \\ \widetilde{\mathbf{g}} \\ \mathbf{r}^{\mathrm{stick}} \\ \mathbf{r}^{\mathrm{slip}} \end{Bmatrix}. \tag{5.41}$$

For convenience, the nomenclature is recalled in the following. The blocks $\mathbf{K}$ denote the stiffness matrix resulting from the linearisation of the internal forces vector. The blocks $\widetilde{\mathbf{K}}$ represent the effective stiffness matrix, obtained from the summation of the respective stiffness blocks $\mathbf{K}$ with the linearization terms of the contact force vector $\mathbf{f}_c$ in order of the displacements, see Equation (3.130). Blocks $\mathbf{A}$ contain the derivatives in order to the displacements of the NCP function for the normal contact constraints, while $\mathbf{Q}$ and $\mathbf{L}$ regard the stick and slip equations derivatives, respectively. The matrices $\mathbf{T}$ and $\mathbf{H}$ aggregate the derivative of the tangential NCP function in order to the Lagrange multiplier. The residual blocks $\widetilde{\mathbf{r}}$ comes as a result of solving directly for the unknown Lagrange multipliers $\mathbf{z}^{k+1}$ at each iteration (i.e., without employing an incremental formulation), see Equation (3.131). The vector $\widetilde{\mathbf{g}}$ gathers all the nodal weighted gaps and $\mathbf{r}^{\mathrm{stick}}$ and $\mathbf{r}^{\mathrm{slip}}$ are the residuals for the stick and slip equations. Compared with the counterpart in Equation (3.156) for unilateral contact, the only difference is that there are no mortar degrees of freedom, thus leading to a smaller total number of stiffness matrix blocks.

Remark. *It should be mentioned that, in Equation (5.41), the case of a non-diagonal first mortar coupling matrix is assumed for the sake of generality. Nonetheless, for first-order finite elements, it retains the diagonal structure due to the bi-orthogonality condition given in Equation (3.7). For second-order interpolation, on the other hand, the application of the modified dual shape functions, see Equation (5.22), leads to a non-diagonal first mortar matrix (although, still easily inverted).*

5.8.2.1 Elimination of the Lagrange multipliers

As already mentioned, the use of dual Lagrange multipliers allows for a straightforward simplification of the system of equations by performing the condensation of the Lagrange multipliers (thus removing the unwanted saddle point structure). This is possible because the first mortar coupling matrix $\mathbf{D}$ can be trivially inverted. The evaluation of the Lagrange multipliers at a given configuration starts with the consideration of the following system of equations

$$\begin{Bmatrix} \mathbf{z}_{\mathcal{I}}^{k+1} \\ \mathbf{z}_{\mathcal{Q}}^{k+1} \\ \mathbf{z}_{\mathcal{L}}^{k+1} \end{Bmatrix} = - \begin{bmatrix} \mathbf{D}_{\mathcal{I}\mathcal{I}}^{-\mathrm{T}} & \mathbf{D}_{\mathcal{I}\mathcal{Q}}^{-\mathrm{T}} & \mathbf{D}_{\mathcal{I}\mathcal{L}}^{-\mathrm{T}} \\ \mathbf{D}_{\mathcal{Q}\mathcal{I}}^{-\mathrm{T}} & \mathbf{D}_{\mathcal{Q}\mathcal{Q}}^{-\mathrm{T}} & \mathbf{D}_{\mathcal{Q}\mathcal{L}}^{-\mathrm{T}} \\ \mathbf{D}_{\mathcal{L}\mathcal{I}}^{-\mathrm{T}} & \mathbf{D}_{\mathcal{L}\mathcal{Q}}^{-\mathrm{T}} & \mathbf{D}_{\mathcal{L}\mathcal{L}}^{-\mathrm{T}} \end{bmatrix} \begin{Bmatrix} \tilde{\mathbf{r}}_{\mathcal{I}} + \sum_{\mathcal{X}\in\{\mathcal{N},\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{K}_{\mathcal{I}\mathcal{X}} \Delta\mathbf{d}_{\mathcal{X}} \\ \tilde{\mathbf{r}}_{\mathcal{Q}} + \sum_{\mathcal{X}\in\{\mathcal{N},\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{K}_{\mathcal{Q}\mathcal{X}} \Delta\mathbf{d}_{\mathcal{X}} \\ \tilde{\mathbf{r}}_{\mathcal{L}} + \sum_{\mathcal{X}\in\{\mathcal{N},\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{K}_{\mathcal{L}\mathcal{X}} \Delta\mathbf{d}_{\mathcal{X}} \end{Bmatrix} . \tag{5.42}$$

Consideration of the fifth row of Equation (5.41) yields $\mathbf{z}_{\mathcal{I}}^{k+1} = \mathbf{0}$ and, therefore, one has

$$\{\mathbf{z}_{\mathcal{Q}}^{k+1}\} = - \sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{Q}\mathcal{Y}}^{-\mathrm{T}} \left[\tilde{\mathbf{r}}_{\mathcal{Y}} + \sum_{\mathcal{X}\in\{\mathcal{N},\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{K}_{\mathcal{Y}\mathcal{X}} \Delta\mathbf{d}_{\mathcal{X}} \right] , \tag{5.43}$$

$$\{\mathbf{z}_{\mathcal{L}}^{k+1}\} = - \sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{L}\mathcal{Y}}^{-\mathrm{T}} \left[\tilde{\mathbf{r}}_{\mathcal{Y}} + \sum_{\mathcal{X}\in\{\mathcal{N},\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{K}_{\mathcal{Y}\mathcal{X}} \Delta\mathbf{d}_{\mathcal{X}} \right] . \tag{5.44}$$

This also allows the third and fourth rows of Equation (5.41) to be solved. It should be mentioned that the presented condensation procedure becomes simplified for first-order finite elements, as the first mortar coupling matrix maintains a diagonal structure. Substitution of the expressions above into Equation (5.41) leads to the final condensed system:

$$\begin{bmatrix} \mathbf{K}_{\mathcal{N}\mathcal{N}} & \mathbf{K}_{\mathcal{N}\mathcal{I}} & \mathbf{K}_{\mathcal{N}\mathcal{Q}} & \mathbf{K}_{\mathcal{N}\mathcal{L}} \\ \breve{\mathbf{K}}_{\mathcal{N}} & \breve{\mathbf{K}}_{\mathcal{I}} & \breve{\mathbf{K}}_{\mathcal{Q}} & \breve{\mathbf{K}}_{\mathcal{L}} \\ \mathbf{0} & \mathbf{A}_{\mathcal{I}} & \mathbf{A}_{\mathcal{Q}} & \mathbf{A}_{\mathcal{L}} \\ \breve{\mathbf{Q}}_{\mathcal{N}} & \breve{\mathbf{Q}}_{\mathcal{I}} & \breve{\mathbf{Q}}_{\mathcal{Q}} & \breve{\mathbf{Q}}_{\mathcal{L}} \\ \breve{\mathbf{L}}_{\mathcal{N}} & \breve{\mathbf{L}}_{\mathcal{I}} & \breve{\mathbf{L}}_{\mathcal{Q}} & \breve{\mathbf{L}}_{\mathcal{L}} \end{bmatrix} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{N}} \\ \Delta\mathbf{d}_{\mathcal{I}} \\ \Delta\mathbf{d}_{\mathcal{Q}} \\ \Delta\mathbf{d}_{\mathcal{L}} \end{Bmatrix} = - \begin{Bmatrix} \mathbf{r}_{\mathcal{N}} \\ \breve{\mathbf{r}}_{\mathcal{I}} \\ \tilde{\mathbf{g}} \\ \breve{\mathbf{r}}^{\text{stick}} \\ \breve{\mathbf{r}}^{\text{slip}} \end{Bmatrix} . \tag{5.45}$$

Here, several algebraic functions and abbreviations have been introduced to facilitate the notation. The algebraic function $\breve{\mathbf{K}}_{(\bullet)}$ in the second row of Equation (5.45) is defined as

$$[\breve{\mathbf{K}}_{(\bullet)}] \equiv \begin{cases} \mathbf{K}_{\mathcal{I}(\bullet)} - \sum\limits_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{I}\mathcal{Q}}^{\mathrm{T}} \mathbf{D}_{\mathcal{Q}\mathcal{Y}}^{-\mathrm{T}} \mathbf{K}_{\mathcal{Y}\mathcal{N}} \Delta\mathbf{d}_{\mathcal{N}} - \sum\limits_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{I}\mathcal{L}}^{\mathrm{T}} \mathbf{D}_{\mathcal{L}\mathcal{Y}}^{-\mathrm{T}} \mathbf{K}_{\mathcal{Y}\mathcal{N}} \Delta\mathbf{d}_{\mathcal{N}}, & \text{if } (\bullet) \in \mathcal{N}, \\ \tilde{\mathbf{K}}_{\mathcal{I}(\bullet)} - \sum\limits_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{I}\mathcal{Q}}^{\mathrm{T}} \mathbf{D}_{\mathcal{Q}\mathcal{Y}}^{-\mathrm{T}} \mathbf{K}_{\mathcal{Y}(\bullet)} \Delta\mathbf{d}_{(\bullet)} - \sum\limits_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{I}\mathcal{L}}^{\mathrm{T}} \mathbf{D}_{\mathcal{L}\mathcal{Y}}^{-\mathrm{T}} \mathbf{K}_{\mathcal{Y}(\bullet)} \Delta\mathbf{d}_{(\bullet)}, & \text{if } (\bullet) \notin \mathcal{N}, \end{cases} \tag{5.46}$$

and the corresponding residual comes as

$$\{\breve{\mathbf{r}}_{\mathcal{I}}\} \equiv \tilde{\mathbf{r}}_{\mathcal{I}} - \sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{I}\mathcal{Q}}^{\mathrm{T}} \mathbf{D}_{\mathcal{Q}\mathcal{Y}}^{-\mathrm{T}} \tilde{\mathbf{r}}_{\mathcal{Y}} - \sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{D}_{\mathcal{I}\mathcal{L}}^{\mathrm{T}} \mathbf{D}_{\mathcal{L}\mathcal{Y}}^{-\mathrm{T}} \tilde{\mathbf{r}}_{\mathcal{Y}} . \tag{5.47}$$

The algebraic function $[\breve{\mathbf{Q}}_{(\bullet)}]$ on the fourth row of Equation (5.45) is defined as

$$[\breve{\mathbf{Q}}_{(\bullet)}] \equiv \begin{cases} - \sum\limits_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{T}\mathbf{D}_{\mathcal{Q}\mathcal{Y}}^{-\mathrm{T}} \mathbf{K}_{\mathcal{Y}\mathcal{N}} \Delta\mathbf{d}_{\mathcal{N}} , & \text{if } (\bullet) \in \mathcal{N} , \\ \mathbf{Q}_{(\bullet)} - \sum\limits_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{T}\mathbf{D}_{\mathcal{Q}\mathcal{Y}}^{-\mathrm{T}} \mathbf{K}_{\mathcal{Y}(\bullet)} \Delta\mathbf{d}_{(\bullet)} , & \text{if } (\bullet) \notin \mathcal{N} , \end{cases} \tag{5.48}$$

and the corresponding residual $\{\check{\mathbf{r}}^{\text{stick}}\} \in \mathbb{R}^{n^{\text{stick}}}$ yields

$$\{\check{\mathbf{r}}^{\text{stick}}\} \equiv \mathbf{r}^{\text{stick}} - \sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{T}\mathbf{D}^{-\mathrm{T}}_{\mathcal{Q}\mathcal{Y}}\tilde{\mathbf{r}}_{\mathcal{Y}} \,. \tag{5.49}$$

On the fifth row of Equation (5.45), the function $[\check{\mathbf{L}}_{(\bullet)}]$ is constructed as

$$[\check{\mathbf{L}}_{(\bullet)}] \equiv \begin{cases} -\displaystyle\sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{H}\mathbf{D}^{-\mathrm{T}}_{\mathcal{L}\mathcal{Y}}\mathbf{K}_{\mathcal{Y}\mathcal{N}}\Delta\mathbf{d}_{\mathcal{N}}\,, & \text{if } (\bullet) \in \mathcal{N}\,, \\ \mathbf{L}_{(\bullet)} - \displaystyle\sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{H}\mathbf{D}^{-\mathrm{T}}_{\mathcal{L}\mathcal{Y}}\mathbf{K}_{\mathcal{Y}(\bullet)}\Delta\mathbf{d}_{(\bullet)}\,, & \text{if } (\bullet) \notin \mathcal{N}\,, \end{cases} \tag{5.50}$$

and the residual $\{\check{\mathbf{r}}^{\text{slip}}\} \in \mathbb{R}^{n^{\text{slip}}}$ as

$$\{\check{\mathbf{r}}^{\text{slip}}\} \equiv \mathbf{r}^{\text{slip}} - \sum_{\mathcal{Y}\in\{\mathcal{I},\mathcal{Q},\mathcal{L}\}} \mathbf{H}\mathbf{D}^{-\mathrm{T}}_{\mathcal{L}\mathcal{Y}}\tilde{\mathbf{r}}_{\mathcal{Y}} \,. \tag{5.51}$$

5.9 Numerical results

In the following, several numerical examples are presented and analysed in order to validate the proposed formulation for rigid/deformable finite deformation contact. The set of numerical examples presented is exclusively focused on particular aspects of the proposed formulation. Firstly, the optimal convergence rate of the Petrov-Galerkin approach is discussed in Section 5.9.1, considering the classical Hertzian contact problem in two dimensions under uniform mesh refinement. The spatial convergence of the proposed piecewise linear interpolation scheme for quadratic finite elements is also analysed. The improvement in computational complexity is measured in Section 5.9.2 by analysing the three-dimensional contact problem of a half torus going against a complex rigid boundary. Lastly, the proposed piecewise linear interpolation scheme for quadratic finite elements in three dimensions is employed in Section 5.9.3 to solve the contact of a deformable base against a rigid punch.

5.9.1 Spatial convergence - Hertzian contact

The first numerical example to be analysed regards the Hertzian frictionless contact between two cylinders under plane strain conditions. The objective is to assess the spatial convergence rate of the Bubnov-Galerkin and the Petrov-Galerkin approaches, including also the proposed piecewise linear interpolation scheme. The geometry of the problem is schematically represented in Figure 5.7. It consists of one deformable cylinder with radius $R_1 = 8$, described by the St.-Venant-Kirchhoff hyperelastic material model, with Young's modulus $E = 200$ and Poisson's ratio $\nu = 0.3$, contacting against a rigid cylinder with the same radius $R_2 = 8$. For simplicity, only half of the deformable cylinder is modelled and the constant pressure of $p_0 = -0.8$ is applied to the top surface of the hemisphere. The problem is discretised using a structured mesh of standard 4- and 8-noded quadrilateral elements for first- and second-order interpolation. The external pressure is applied incrementally in 20 steps, considering a relative convergence tolerance for the nonlinear solver of $\varepsilon_r = 1 \times 10^{-10}$. Only the segment-based strategy is considered in order to not

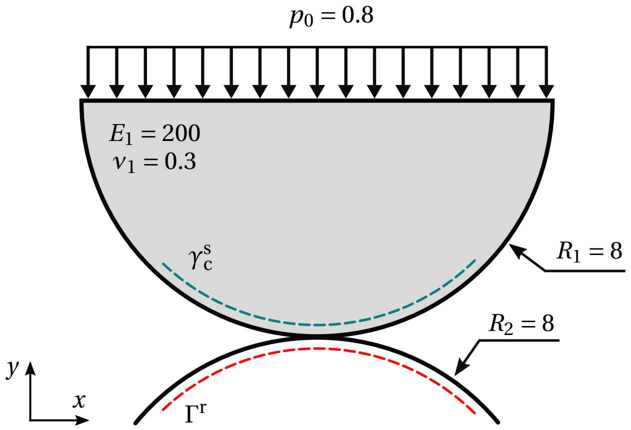


Figure 5.7: Hertzian contact – schematic representation of the problem setting.

compromise the spatial convergence properties of the underlying mixed finite element formulation.

Firstly, in order to inspect the convergence rate, successive uniform mesh refinement steps are employed in a structured manner. Figure 5.8a shows the entire coarse finite element mesh, including a snippet of the remaining refinement levels. The contact region starts with a discretisation based on 4 elements, which is then doubled four times, i.e., until reaching 64 elements. The idea is to compare the solution of each level with a reference solution, here obtained using the second-order finite element mesh with 256 elements (two refinement levels above the most refined mesh). The Bubnov-Galerkin and the Petrov-Galerkin formulations are considered, including the piecewise interpolation scheme for second-order interpolation as well.

Figure 5.8b shows the discretisation error based on the H^1-norm of the error in the displacement field, i.e., by evaluating directly $\|\boldsymbol{u} - \boldsymbol{u}^{\mathrm{h}}\|$ between solutions. For first-order interpolation, $O(h)$ convergence is observed, while for second-order interpolation, optimal results of $O(h^{3/2})$ are achieved. These are in accordance with the theoretical estimates and numerical investigations carried out within the context of unilateral contact due to a reduced regularity of contact solutions, e.g., see B. I. Wohlmuth, Popp, *et al.* (2012) and Popp, B. I. Wohlmuth, *et al.* (2012). Regarding the ongoing formulation for rigid/deformable contact, one significant result is that the Petrov-Galerkin convergence rates are practically identical with the classical Bubnov-Galerkin approach. This indicates that the mathematical structure of the dual mortar formulation remains unaffected by the mixed interpolation scheme used in the Petrov-Galerkin approach. The same is observed for the piecewise linear interpolation.

Numerical results for the contact pressure distribution are illustrated in Figure 5.9 for first-order interpolation using different discretization levels.[‡] As expected, no oscillations on the contact normal pressure are observed throughout the active contact region, and the solution with the second level of refinement, $n = 2$, is practically identical with the remaining finer meshes. The coarse mesh, $n = 1$, appears to be inappropriate to model the contact problem accurately, as the active region is only described by three

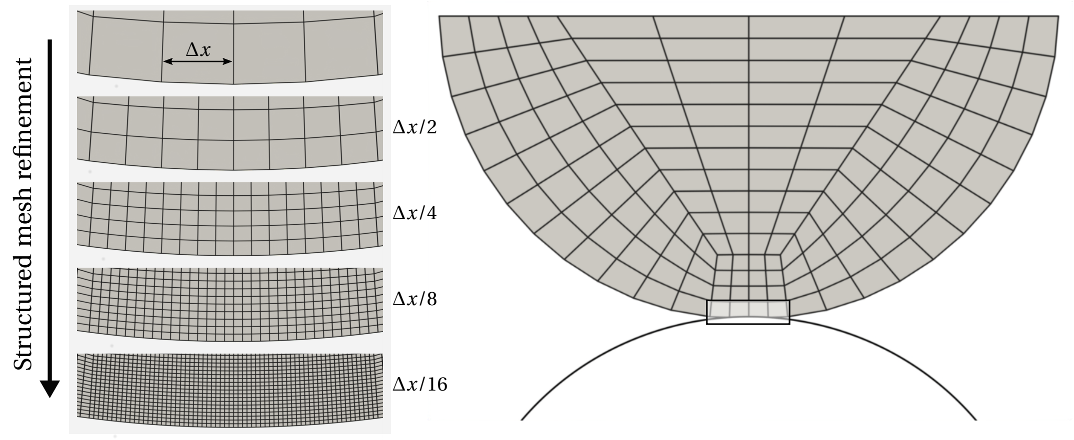
---

[‡] Even though not documented here for the sake of brevity, similar results are obtained with the second-order interpolation.

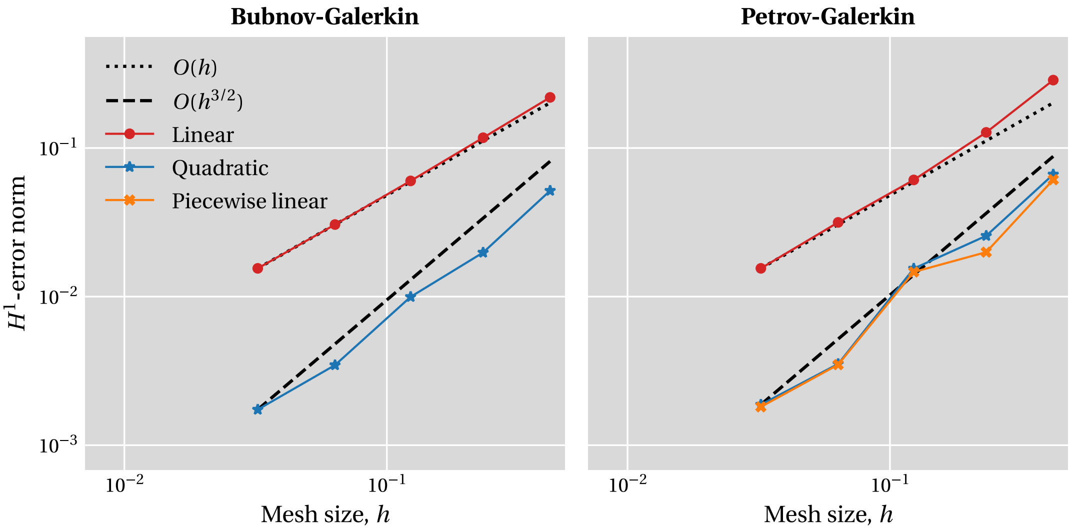
elements. Notwithstanding, this problem configuration has been intentionally considered beforehand to identify possible limitations of the formulations. Even in this case, no particular formulation becomes clearly more compelling in terms of accuracy and, therefore, reinforces the conclusion that the Petrov-Galerkin approach behaves similarly to the Bubnov-Galerkin scheme.

Figure 5.10 shows the maximum normal contact pressure, p^{η}_{max}, with mesh refinement for all the combinations considered in this section. All the results converge to the same value, with the only significant difference being the results obtained with the coarse mesh. While the first-order Bubnov-Galerkin tends to underestimate the maximum contact pressure, the first-order Petrov-Galerkin technique tends to follow the trend of second-order interpolation and overestimates the result. The results obtained with the piecewise linear interpolation have the slightest variation.

Lastly, in order to have a complete picture of the results, Figure 5.11 shows the deformed configuration of the cylinder, including a coloured representation of the vertical displacement. Only the case with the coarse mesh is represented, as the remaining meshes are similar. In fact, even for the coarse mesh, the difference in the displacement field is so slight that the only significant difference regards the results obtained with the Petrov-Galerkin approach with first-order elements. A small penetration of the nodes at the end of the active contact zone is observed, which reflects the difference in the way the weighted gap is evaluated. Nonetheless, this is consistent with similar studies based on the Petrov-Galerkin technique, e.g., in Popp, Seitz, *et al.* (2013) and Sitzmann *et al.* (2016), and is a behaviour that tends to vanish with mesh refinement. No overlap is observed in the results obtained with the next finer mesh, $n = 2$.



(a) Successive finite element mesh refinement.



(b) Convergence of the discretization error, $\|\boldsymbol{u} - \boldsymbol{u}^{\mathrm{h}}\|$.

Figure 5.8: Hertzian contact – convergence of the H^1-norm of discretization error $\|\boldsymbol{u} - \boldsymbol{u}^{\mathrm{h}}\|$, for both first- and second-order finite interpolation based on quadrilaterals.

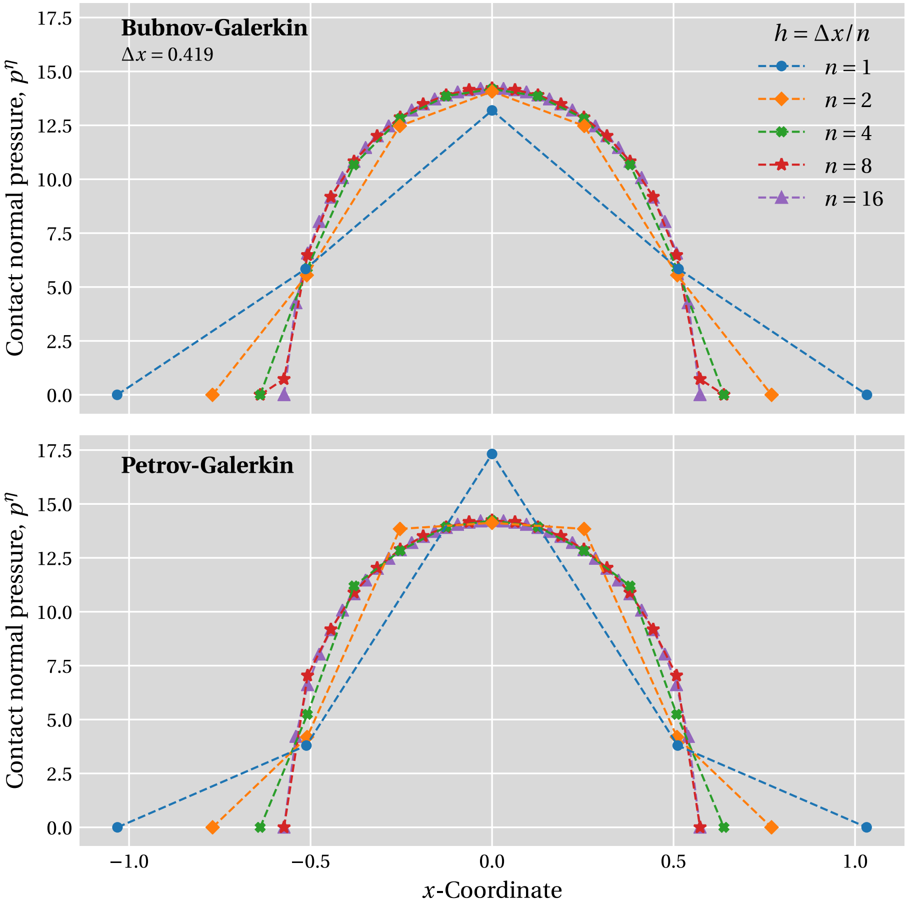


Figure 5.9: Hertzian contact – convergence of the contact pressure distribution p^{η} with mesh refinement for first-order quadrilateral meshes.

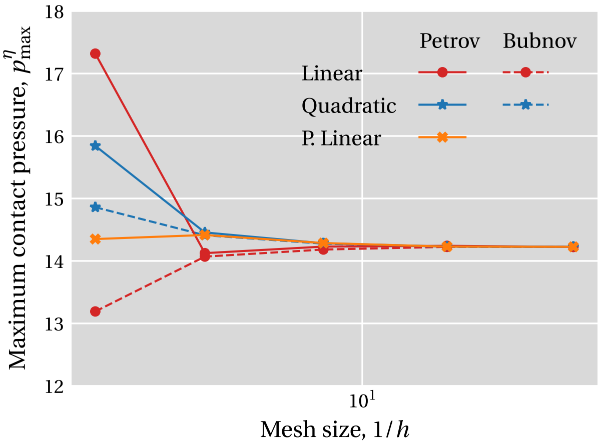


Figure 5.10: Hertzian contact – maximum normal contact pressure, $p^{\eta}_{\max}$, with mesh refinement for first-order and second-order quadrilateral meshes.

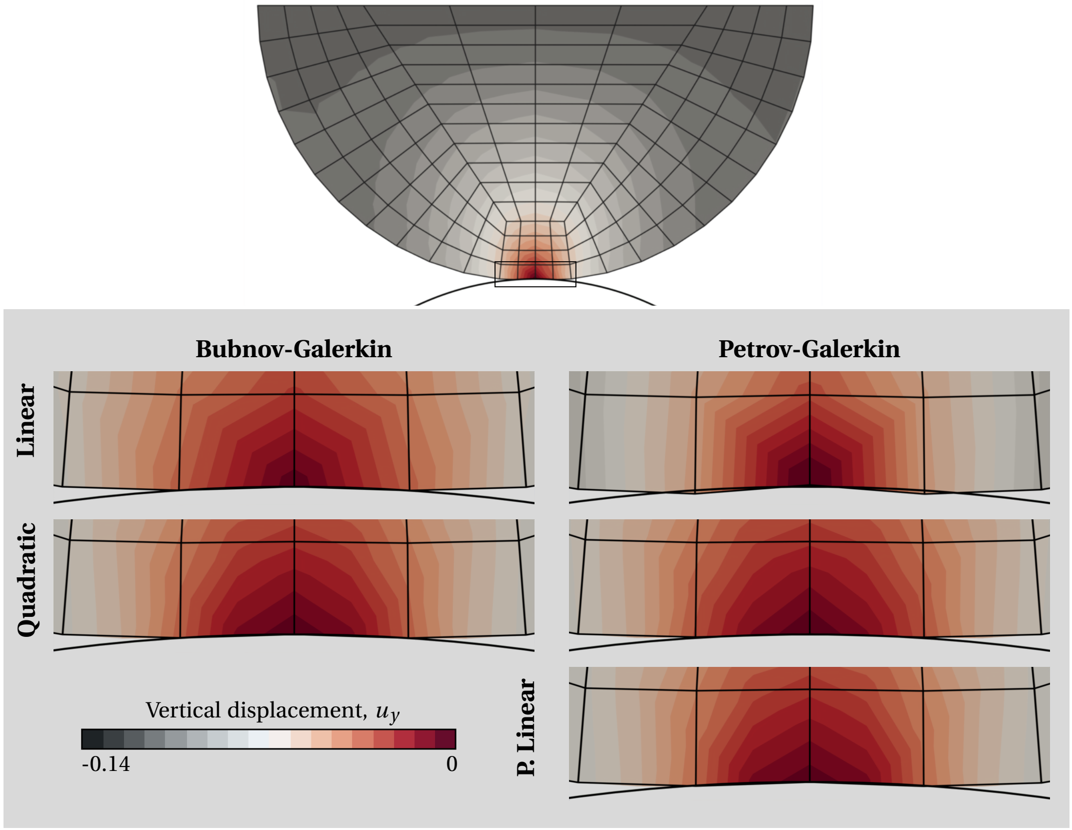


Figure 5.11: Hertzian contact – vertical displacement for different methods. The penetration obtained with the Petrov-Galerkin approach and linear finite elements is highlighted.

5.9.2 Computational complexity - Half-torus Signorini contact

In this section, the computational complexity of dual mortar formulations for rigid/deformable contact is quantitatively measured. Besides robustness, the motivation for the proposed techniques is mainly driven by improving the numerical efficiency of contact modelling. Therefore, a numerical example designed to unveil the computational performance of the contact algorithm has been carried out. The problem setup and dimensions are schematically illustrated in Figure 5.12, which consists of the contact of a deformable half-torus against a rigid surface with a relatively complex shape. This example renders a high ratio between the contact interface and total degrees of freedom, while involving a non-trivial rigid boundary to define the projected normal vectors. The Neo-Hookean material model is considered for the half-torus, which is subjected to an incremental vertical displacement and has its outer surface set as non-mortar. Frictionless contact is assumed based on four combinations of formulations:

- Bubnov/Petrov-Galerkin techniques;
- Averaged/projected orthonormal moving frame.

The idea is to measure the impact of the two main aspects discussed in this work (individually and combined), thus getting an estimate of the overall computational performance of the algorithms for rigid/deformable contact.

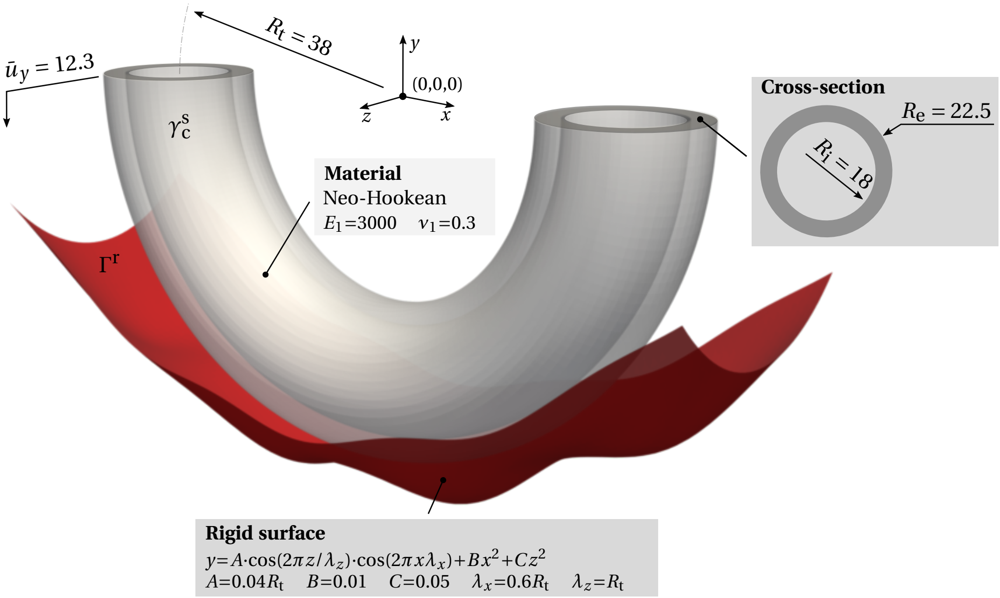


Figure 5.12: Half-torus – schematic representation of the problem setting.

The body is discretised using a structured 8-noded hexahedron mesh with F-bar elements (de Souza Neto *et al.*, 1996) and the rigid side is discretised using 4-noded bilinear quadrilaterals. The total dimension of the problem is around 85k nodes, from which 13k are rigid. The displacement is applied in 27 equally spaced increments and a relative convergence tolerance of $\varepsilon_r = 1 \times 10^{-6}$ is considered. In order to accentuate the complexity of the contact algorithm, the mortar integrals are evaluated using the segmentation method exclusively. Using a desktop workstation, the total simulation time has varied between 3 and 5 hours. An exemplary representation of the deformed configuration at the end of the simulation is given in Figure 5.13, including glyphs representing the contact stress field.

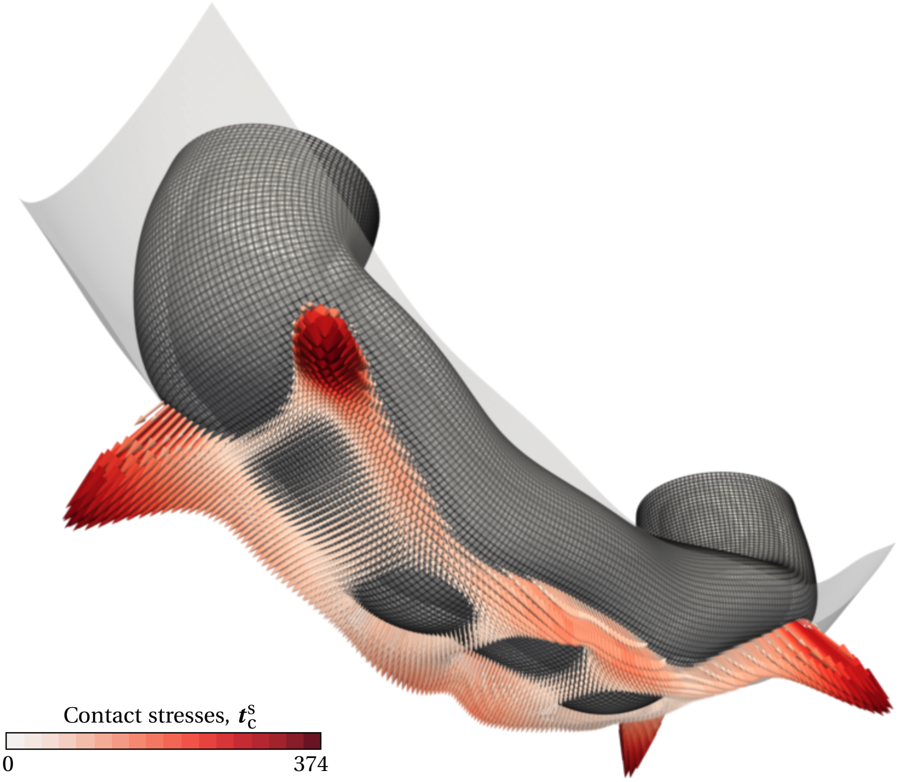


Figure 5.13: Half-torus – deformed configuration contact stress field.

The time needed to evaluate all the terms related to contact is represented in Figure 5.14. Here, the average time per iteration of a given increment is plotted against the ratio between the total number of active nodes and the total number of nodes of the finite element mesh. As the number of active nodes increases monotonically with the pseudo-time, the horizontal axis can be interpreted as the pseudo-time, or currently prescribed displacement, yet adjusted to a more relevant quantity for the current analysis. In turn, the ratio of active node numbers measures the impact of contact modelling within the global finite element problem. On the vertical axis, the contact time (mainly dominated by the linearisation update procedures) is summed per iteration and, after the global Newton algorithm converges, it is averaged over the total number of iterations needed to achieve equilibrium conditions. It is observed that, as expected, the computational time increases with the total number of active nodes. The classical formulation based on the Bubnov-Galerkin approach with the averaged orthonormal frame is the most computational demanding combination. In contrast, the newly proposed methodology based on the Petrov-Galerkin/projected frame is the fastest. In order to evaluate the differences more clearly, the bottom graph in Figure 5.14 shows the speed-up of each formulation in relation to the Bubnov-Galerkin/averaged frame method. The most significant improvement in efficiency is achieved by switching from Bubnov- to Petrov-Galerkin techniques, with an average reduction of $\approx 35\%$ in computation time. The projected frame achieves a reduction of $\approx 10\%$, which means that the two methodologies combined sum up to a $\approx 45\%$ reduction in the computation time for contact evaluation.

At this stage, it seems appropriate to discuss the impact of the contact algorithm within the global framework of finite element modelling. One can acknowledge that the impact of this reduction in the global simulation time is highly dependent on the problem size, the computational implementation and the computer hardware. All the examples presented in this work have been solved using a desktop workstation. The global linear system of equations have been solved by employing a direct solver and all the operations have been carried out in serial. Under these conditions, the problem is dominated by the contact algorithm and, therefore, the global computation times follow approximately the same tendency of the times plotted in Figure 5.14. By employing strategies such as parallelisation, one should expect the problem to become dominated more by the linear solver. Nonetheless, even for such an optimised scenario, the computational complexity associated with the contact algorithm (especially the geometrical operations and subsequent linearisation) is typically not negligible. It is expected to remain an essential part of the overall computational performance for more demanding problems with a high ratio of active contact nodes.

In order to understand more clearly the reduction in computational complexity, one can look at memory usage. Figure 5.15 shows the total number of non-zero entries in the derivatives of the unit normal vector at each active contact node for both the averaged and projected techniques. For the sake of simplicity, only the results obtained with the Petrov-Galerkin approach are shown, as the results of Bubnov-Galerkin are exactly the same. As expected, this number increases throughout the simulation, and the bandwidth of the derivative of the projected frame is smaller than for the averaged method. The derivative of the unit normal vector appears in every term of the mortar formulation. As already mentioned, it has an amplified effect by dictating the total number of individual operations within the sparse matrix procedures.

Besides the total number of non-zero entries, one can look at the sparsity pattern of the global stiffness matrix to understand the improvements in computational complexity. Figure 5.16 shows a visualisation of an exemplary square system matrix originating from the averaged and projected methodologies, in which a non-zero entry in the matrix is marked with a black pixel. For ease of interpretation, a magnified representation of the blocks associated with the contact constraints is also included, as any modification in the contact algorithm will be reflected there. As expected, the pattern is more compact in the projected frame variant because the projected frame derivative involves fewer terms than the averaged strategy. Nonetheless, both global system matrices exhibit the pronounced band structure obtained with mortar methods, see e.g., Popp (2018).

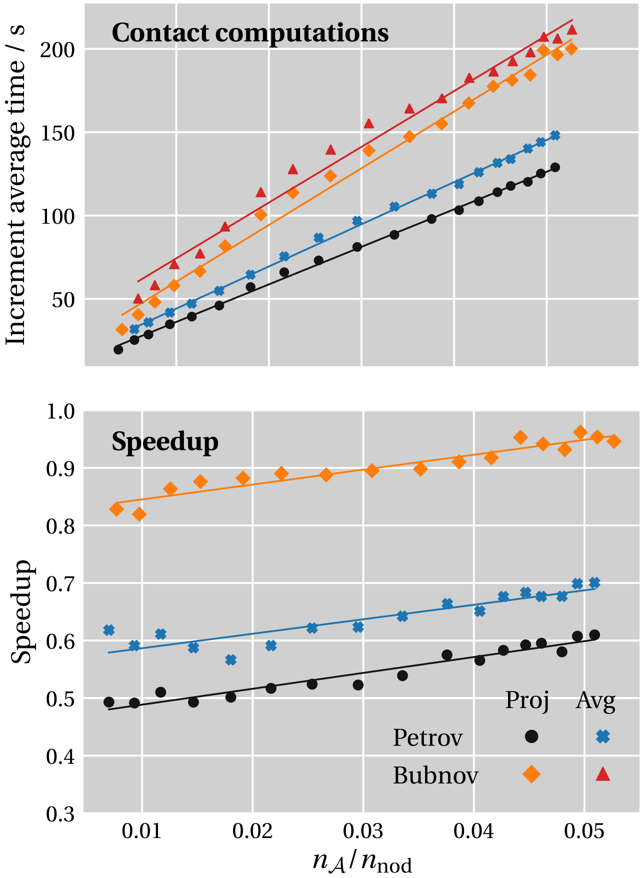


Figure 5.14: Time required to complete all operations related with contact (the solid lines are linear fittings).

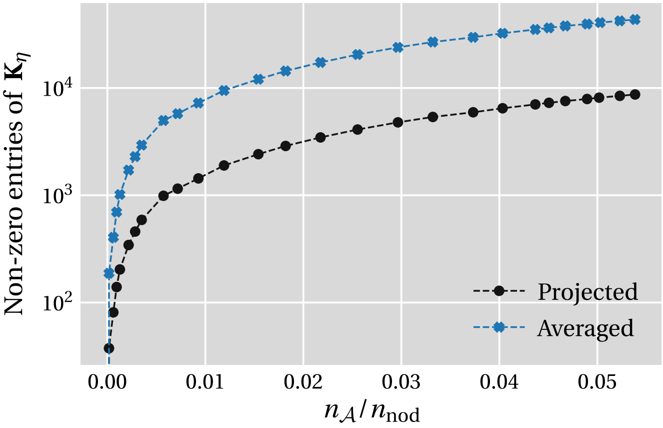
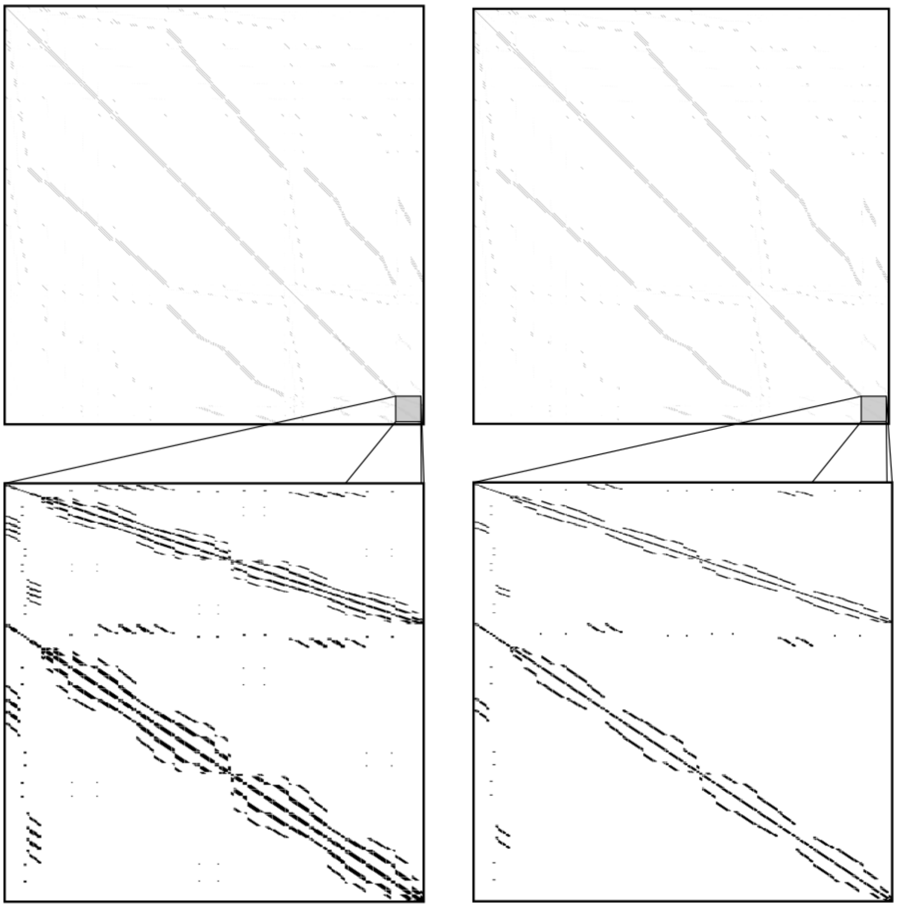


Figure 5.15: Half-torus – number of non-zero entries of the derivative of the unit normal vector $\mathscr{D}\boldsymbol{\eta} \equiv \mathbf{K}_\eta \Delta \mathbf{d}$.



(**a**) Averaged. (**b**) Projected.

Figure 5.16: Half-torus – exemplary sparsity pattern of the global stiffness matrix, where black pixels represent non-zero entries.

5.9.3 Piecewise linear interpolation - 3D punch

The last numerical example is focused on demonstrating the effect of the piecewise linear interpolation for quadratic dual mortar methods based on the Petrov-Galerkin technique in three dimensions. The problem is schematically represented in Figure 5.17 and consists of a deformable cuboid base being pressed against a rigid pin with rounded edges under frictionless conditions. The small fillet radius on the rigid punch naturally produces high local curvatures, especially for coarse discretisation, thus posing substantial convergence problems for algorithms that do not rely on strictly positive interpolation functions for the weighted gap calculation. The base material is characterised by the Neo-Hookean constitutive model. Both the bottom and lateral faces have their vertical displacement prescribed and are fixed along the remaining directions, such that no lateral movement is allowed (thus avoiding unstable configurations). The displacement is applied in 65 equally spaced increments, considering a convergence tolerance of $\varepsilon_r = 1 \times 10^{-6}$.

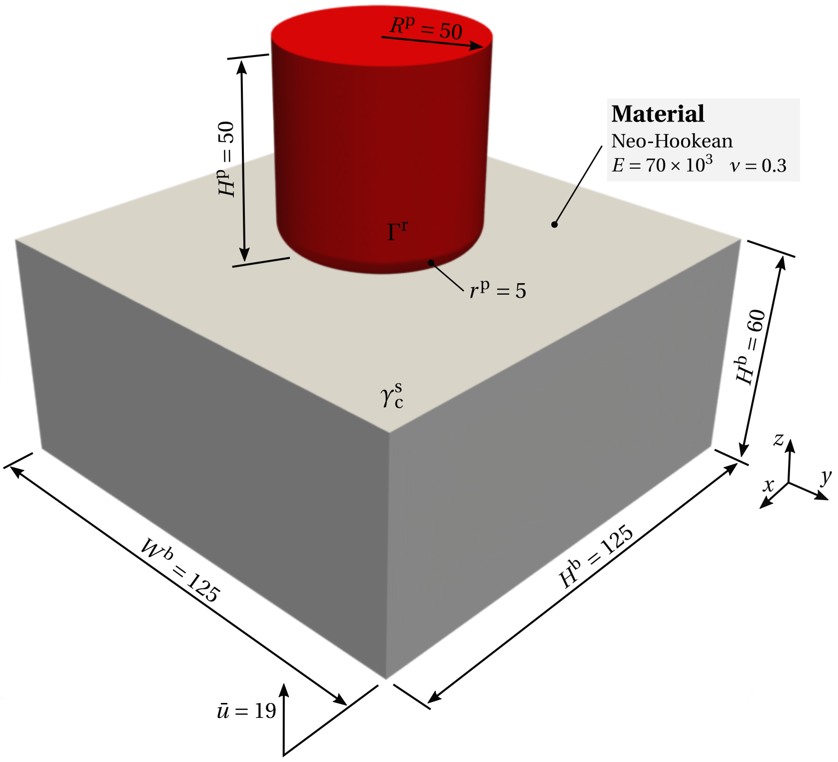


Figure 5.17: 3D punch – schematic representation of the problem setting.

The deformable base is discretised using 20-noded hexahedra with full integration, while the rigid pin is discretised with 8-noded quadrilateral elements. For the sake of integration accuracy, the mortar integrals are being evaluated using the segmentation method exclusively. The deformed configuration of the base at the end of the simulation is represented in Figure 5.18, including a contour plot of the displacement field along the vertical direction. The proposed algorithm converges without any spurious contact states or oscillations for both coarse and fine finite element meshes.

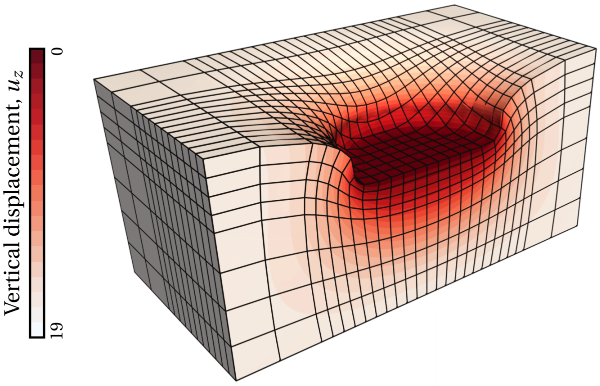


Figure 5.18: 3D punch – vertical displacement of the deformable base at the end of the simulation. Only half of the base is visualized.

In order to visualise more clearly the concepts of the projected frame and piecewise linear interpolation, a snippet of the integration cells at the final increment of the simulation is schematically represented in Figure 5.19. The unit normal vector associated with each non-mortar node is also indicated. Looking at the unit normals, one can see that they all point towards the rigid punch, which does not match the contour of the body for inactive regions of the contact interface (e.g. top left part of the left image in Figure 5.19). Nonetheless, as the distance between both surfaces decreases, they become almost identical and the normal vectors start capturing the contour of the non-mortar boundary very precisely. This aspect of the algorithm is also visualised when looking at the orientation of the integration cells, which are established based on the auxiliary plane defined from the continuous field of normal vectors. Lastly, in Figure 5.19 the division of the elements into sub-elements can also be identified when looking at the contour of the integration cells.

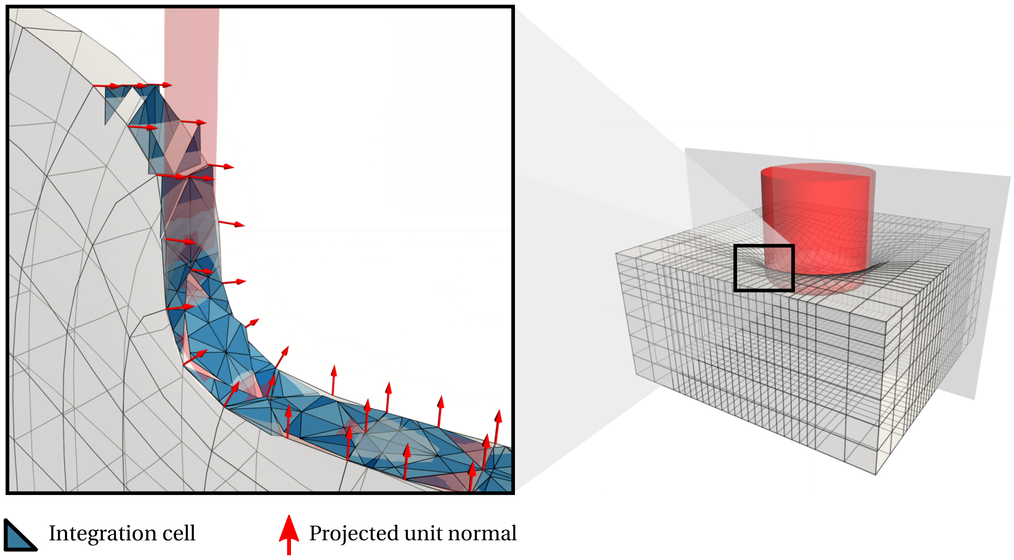


Figure 5.19: 3D punch – slice of the central section of the deformed configuration. The arrows represent the unit normal vector at each non-mortar node and the blue cells are the integration cells resulting from the segmentation algorithm.

Chapter 6

An Efficient Multi-scale Strategy to Predict the Real Contact Area

As mentioned in arguably every publication in the literature, micromechanical contact is a truly multi-physical problem. Numerous phenomena take place simultaneously at the contact interface, spanning over multiple temporal and physical scales. In reality, a strong coupling between all the physical properties and effects is expected, which proves to be one of the main challenges for computational modelling of micromechanical contact. In fact, one of the main arguments favouring numerical models is quite the opposite, i.e., the capability to isolate individual phenomena and, ultimately, access its underlying trend. In experiments, on the contrary, the ability to isolate individual effects proves to be a typical challenge. For this reason, no strategy has prevailed over the other and, over the last decades, both have been continuously developed in a combined effort to understand better what is still today an answered question: what are precisely the microscopic origins of friction and wear?

In engineering, one of the main aspects behind contact interaction regards the surface's topography. It is widely recognised that, in real surfaces, roughness has entities that span over multiple length scales. This is a challenge for numerical modelling of rough contact, as the computational complexity grows along with the considered bandwidth roughness spectra. The dangerous combination of the discretization requirements to properly assemble the smallest geometrical features, combined with the model size needed to encompass the largest ones, makes the application of methods like the FEM difficult. One of the approaches that tackle this issue is the multi-scale modelling of rough contact, which applies the concepts of homogenisation to the roughness itself.

This chapter starts with the definition of the RCE, the fundamental entity used to model the physics at the contact interface. The generation of rough boundaries using the power spectrum model is also presented, followed by the proposed multi-scale strategy. This includes the topography decomposition method, the process for generating rough topographies for each scale, and the new multiplicative homogenization scheme. The rest of the chapter is dedicated to numerical investigations. The multi-scale formulation is validated based on the comparison against the original full-size problem. Particular attention is given to the violation of the principle of scale separation. The impact of key aspects such as the discretization error, the number of scales, or bulk material constitutive behaviour are also investigated. Finally, the overall improvement in terms of computational complexity is quantitatively measured, including also a comparison with several solutions available in the literature.

6.1 Contact homogenization and the definition of a representative contact element

As mentioned in Chapter 1, in engineering, when thinking about roughness, one can recognise that the physical dimension of roughness features is considerably smaller than the characteristic size of the contacting bodies. This reassembles the principle of separation of scales and, thus, allows the establishment of a statistically representative element of the micro-scale features. To do so, the geometry and material properties for the contacting bodies must be specified, together with a set of admissible boundary conditions. Figure 6.1 shows a schematic representation of all the boundary conditions for a particular body being pressed against a rigid flat surface.

Remark. *The strategy developed assumes that the contact interaction occurs between a deformable body and a rigid flat base. In this case, the enforcement of contact constraints in the multi-scale algorithm becomes considerably simplified. Nonetheless, this is done just for the sake of simplicity, bearing in mind that the multi-scale formulation is equally applicable in the general scenario of unilateral contact between two deformable bodies.*

The RCE can be understood as a periodic unit cell containing a rough surface surrounded by boundaries under the periodic condition and an exterior boundary, subjected to a uniform external pressure p_0.[*] Furthermore, to simplify and improve the numerical stability of the model, the vertical displacement of the nodes at the exterior boundary is set to be the same (thus remaining horizontal throughout the deformation process). Appendix C provides additional details on the treatment of edges and vertices, as this detail is typically not addressed in the literature.

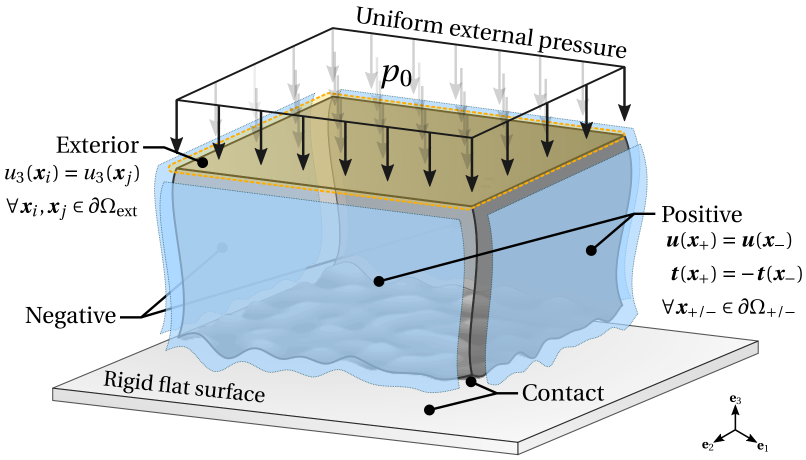


Figure 6.1: Numerical setup for the representative contact element.

---

[*] Within a fully coupled multi-scale framework, the external pressure establishes the fundamental link between both scales (Temizer and Wriggers, 2008). In this work, though, only the micro-scale is analysed and, therefore, the external pressure is freely imposed on the RCE.

6.2 Numerical generation of rough boundaries

One of the main aspects of the RCE geometry is its rough boundary. Typically, it can be either extracted from experimental measurement or artificially generated by employing some random process generation algorithm, see Y. Z. Hu and Tonder (1992) and Wu (2000) for exemplary strategies. The latter approach proves to be more convenient in the sense that it is less time-consuming, while also being suitable for integration within a scripted computational framework. Large numbers of rough topographies realizations with controlled statistical and spectral properties can be easily generated, thus promoting the execution of parametric studies.

Since the present work aims at establishing a fully automated framework for multi-scale analysis of rough contacts, the Gaussian topography generation algorithm proposed in Wu (2000), applicable both to random profiles and surfaces, has been implemented. The reader is referred to the original publication for further details and an informative flowchart presented in Couto Carneiro (2019) and Couto Carneiro *et al.* (2020). This algorithm is based on Fast Fourier Transforms and consists of synthesizing rough topographies as the superposition of waves with randomly generated phases and amplitudes, and dependent on the input Power Spectral Density (PSD) function. This function describes the frequency content of rough topographies in an unbiased and scale-independent way, and can arguably be regarded as the most important tool to characterise rough topographies (Persson, 2014; Jacobs *et al.*, 2017). Moreover, the PSD is related to both Root Mean Square (RMS) and fractal parameters, thus proving to be a powerful way to describe roughness (V. A. Yastrebov, Anciaux, *et al.*, 2015). In fact, the spectral content, together with a statistical description of the heights distributions, is capable of adequately describing a significant number of rough topographies in the context of engineering applications. Additionally, it should be noted that the topographies generated with this methodology are periodic, thus having the practical convenience of being naturally tailored for the construction of the periodic cell described in Section 6.1.

6.2.1 Roughness power spectrum model

Real engineering surfaces typically have fractal behaviour between a high-frequency cut-off k_s and a roll-off frequency k_r (Persson *et al.*, 2005). Moreover, between the roll-off frequency and a low-frequency cut-off k_l, a plateau of constant PSD is generally identified. However, for the sake of simplicity, in this work, the plateau is not considered ($k_r = k_l$). The PSD of an isotropic self-affine profile is mathematically defined as

$$\Phi_{\theta}^{\mathrm{PSD}}(k) = \begin{cases} C_0 \left(\dfrac{k_r}{k}\right)^{1+2H}, & k_l \leq k \leq k_s\,, \\ 0\,, & \text{elsewhere}\,, \end{cases} \tag{6.1}$$

see Figure 6.2, and for surfaces as

$$\Phi^{\mathrm{PSD}}(\boldsymbol{k}) = \begin{cases} C_0' \left(\dfrac{k_r}{\|\boldsymbol{k}\|}\right)^{2(H+1)}, & k_l \leq \|\boldsymbol{k}\| \leq k_s\,, \\ 0\,, & \text{elsewhere}\,. \end{cases} \tag{6.2}$$

Here, C_0 and C_0' are scale constants and $H \in [0,1]$ is the Hurst roughness exponent. In the end, the generation of a random self-affine topography requires the specification of a long cut-off wavelength $\lambda_l = 2\pi / k_l$, the roughness bandwidth λ_l / λ_s, the Hurst exponent H and the scale factor C_0 (for profiles) or C'_0 (for surfaces). This scale factor can be determined by fixing the RMS slope (defined as the square root of the profile spectral moment of order 2), which proves advantageous since this parameter can be controlled independently of the employed discretization (V. A. Yastrebov, Anciaux, *et al.*, 2015).

6.3 Multi-scale formulation

The methodology described so far can model contact interaction at roughness level and, essentially, establishes a general framework for contact homogenization. However, as already mentioned, roughness itself spans over multiple length scales, which ultimately may yield a cumbersome numerical model. To tackle this problem, in this work, a multi-scale approach for rough contact is proposed.

6.3.1 Topography decomposition

The basis for the multi-scale formulation is the decomposition of the surface spectrum. This idea has been originally proposed in Wagner, Wriggers, Klapproth, *et al.* (2015) and Wagner, Wriggers, Veltmaat, *et al.* (2017) within the context of rubber friction analysis, although it can be equally applied for the modelling of normal contact. As graphically represented in Figure 6.2, the topography is decomposed on the frequency domain by splitting the PSD function. Each part will correspond to a scale, with a frequency content smaller than the original topography. This decomposition must be embedded within a statistical description of the problem, to integrate all frequency contributions and recover the original problem.

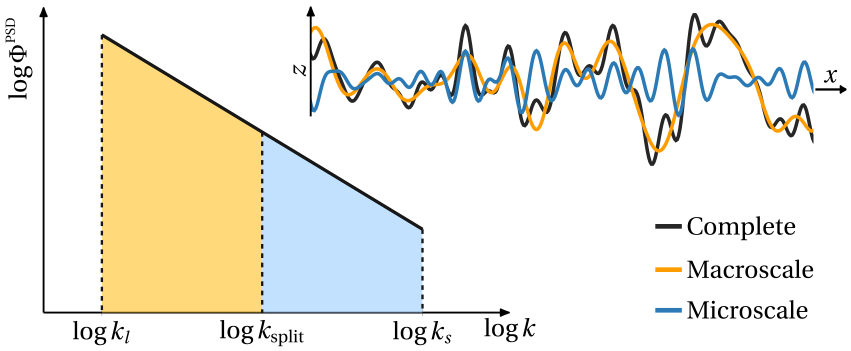


Figure 6.2: Surface decomposition.

6.3.2 Definition of splitting frequencies

At this stage, it is crucial to note that one needs a criterion for the definition of the splitting frequencies. In fact, in the original work of Wagner, Wriggers, Klapproth, *et al.* (2015), this problem is not clearly addressed and, to the author's knowledge, no other sources referring to similar issues have been found. Therefore, the idea proposed here consists of defining a constant bandwidth of $\zeta = \lambda_l / \lambda_s$ across all scales. This means that, for a generic decomposition of a given topography into n_s scales, the splitting frequency $\lambda_\mathrm{split}^{\{j\}}$ for each scale $\{j\}$ writes

$$\frac{\lambda_l}{\lambda_\mathrm{split}^{\{1\}}} = \frac{\lambda_\mathrm{split}^{\{1\}}}{\lambda_\mathrm{split}^{\{2\}}} = \cdots = \frac{\lambda_\mathrm{split}^{\{n_\mathrm{s}\}}}{\lambda_s} = \zeta\,. \tag{6.3}$$

The scale bandwidth ratio that fulfils Equation (6.3) is

$$\zeta = \sqrt[n_\mathrm{s}]{\frac{\lambda_l}{\lambda_s}}\,, \tag{6.4}$$

and, thus, all splitting frequencies $\lambda_\mathrm{split}^{\{j\}}$ can be obtained from

$$\lambda_\mathrm{split}^{\{j\}} = \frac{\lambda_l}{\zeta^j}, \quad \text{for } j = 1, \ldots, n_\mathrm{s} - 1\,. \tag{6.5}$$

The motivation for such a criterion is twofold. On the one hand, this ratio achieves maximum computational efficiency since all scales have similar discretization requirements (defined by the ratio between the maximum and minimum wavelengths). On the other hand, it is essential to note that the principle of separation of scales is not entirely fulfilled at roughness level—its various length scales cover a *continuous* spectrum. The smallest length scale at a given PSD range is the largest scale at the following one, which means that scales are not naturally separated. Nonetheless, as investigated in great detail in Section 6.4, the impact of the violation of the scale separation principle on the numerical accuracy of the method tends to vanish for sufficiently large roughness bandwidths. Therefore, the proposed decomposition strategy assures that the bandwidth across all scales (constant) is kept to the maximum, thus ultimately aiming at preserving the overall numerical accuracy.

6.3.3 Generation of rough topographies for each scale

An essential detail of the multi-scale framework regards the generation of the topography for each scale. For example, in Figure 6.2, the original profile is separated into two scales with the same length but with different, yet, consistent cut-off wavelengths. If these profiles were to be used directly in the numerical model, there would be no improvement in computational cost. Even though a coarse discretization could be employed at the macro-scale, the micro-scale would still require a mesh as fine as the original single scale problem (because its physical length is still the same). One way to circumvent this problem is by generating a full-length micro-scale topography and, then, truncating the profile. However, as shown below, such an approach distorts the micro-scale PSD and, on top of that, compromises the topography periodicity. The alternative method employed in this work is to directly use the random roughness generator, describing in the following the consequences of such a choice.

The frequency resolution of the power spectrum (i.e., the spacing between the discrete PSD points) is inversely proportional to the topography length and, together with the number of discrete points, dictates the maximum frequency available in the PSD. Using a generic random topography (complete) as a reference, Figure 6.3 represents its power spectral density, including the macro-scale and micro-scale contributions. The macro-scale topography, having the same length as the original one, can reproduce precisely the discrete PSD points of the complete topography (but only within its reduced bandwidth). The micro-scale, however, despite covering the expected frequency range, has a much less populated spectrum, owing to its shorter length. Nonetheless, as investigated in great detail in Couto Carneiro *et al.* (2020), a converging behaviour of the sample length is expected, which means that the impact on the results of increasing the topography length becomes gradually smaller. Within the multi-scale formulation, this means that if the bandwidth ζ is kept to reasonable values, it is expected to have no significant impact on the results due to the decrease in the number of discrete frequencies on the problem.

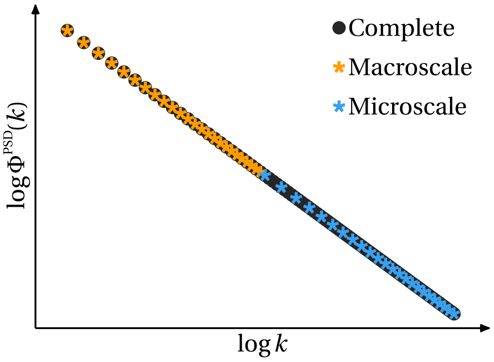


Figure 6.3: Power spectral density of a complete rough topography, together with the corresponding macro- and micro-scale contributions.

Besides the power spectrum, one can assess the micro-scale topography generation by looking at the RMS slope of the profile. As an example, consider a complete topography characterised by $H = 0.8$, $L = 1$, $\lambda_l = L/8$, $\lambda_s = L/128$ and RMS slope set to 0.2. This means that, by separating the profile into two scales, the largest wavelength of the micro-scale is $\lambda_l^{\{2\}} = L/32$, according to the previously defined rule. The full micro-scale (i.e., keeping the original length) is discretized with 1024 points and the generated micro-scale with 256. Figure 6.4 shows the statistical distribution of the RMS slope of the full, generated and truncated micro-scale topographies. It should be remarked that the RMS slope of the full and generated micro-scale topographies is only dictated by the prescribed discrete PSD, so they are represented by a Dirac Delta function centred at the respective value. In contrast, the RMS slope of the truncated topography depends on the region of the full micro-scale topography that is extracted, which originates a collection of RMS slope values following approximately a normal distribution centred at the prescribed value.

From a qualitative perspective, both the truncated and generated profiles look similar. A quantitative analysis of the statistical properties of the RMS slope, however, reveals that only the generated profile can reproduce very closely the value obtained with the full micro-scale. The truncated profile assumes a wide range of values depending on the extracted region and, therefore, does not resemble the statistical properties of the full micro-scale.[†]

[†] Note that for all cases, the RMS slope of the micro-scale is necessarily smaller than 0.2, as a conse-

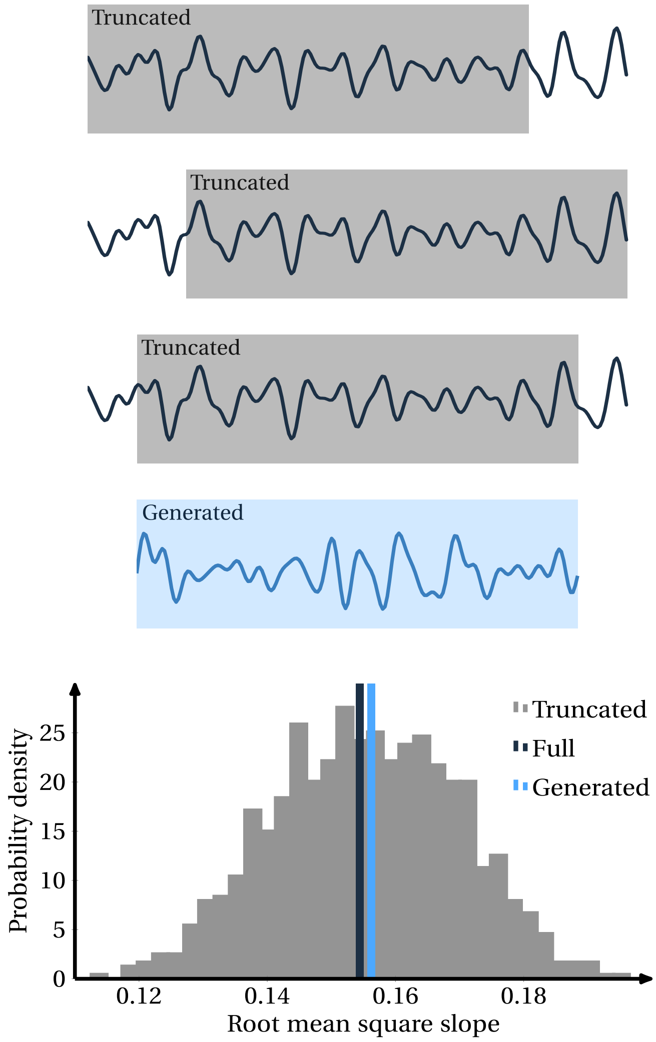


Figure 6.4: Comparison between the profile statistics of a directly generated topography and truncated one.

6.3.4 Multi-scale post-processing algorithm

Having specified the separation of the roughness scales and the corresponding RCEs, the contact area evolution curve for each scale can be *independently* obtained by performing a finite element simulation up to full contact conditions. With a complete description of the contact area at all load ranges for every scale at hand, the main objective is to establish a database of area-pressure pairs, which can then be post-processed according to several strategies.

The strategy proposed in Wagner, Wriggers, Veltmaat, *et al.* (2017), based on multiplicative homogenization, serves as the basis for the proposed methodology. It consists of passing, for a given load step at the macro-scale (scale 1), the mean contact normal pres-

---

quence of the scale separation (recall that the scale constants C_0 and C_0' in Equation (6.1) and Equation (6.2), respectively, are determined by fixing the RMS slope).

sure, $\bar{p}^{\eta\{1\}}$, to the following scale (scale 2). This contact pressure now acts as the uniform contact pressure, applied at the exterior boundary of scale 2. This downscaling process continues until the last scale is reached and, in the end, a list of real contact area fractions at all scales can be upscaled. The multi-scale solution for the real contact area is, then, computed in a multiplicative homogenization step, by taking the product between all contact area fractions of the said list. Mathematically, for a given scale $\{j\}$ it verifies

$$A_{\mathrm{r}}^{\mathsf{h}\{j\}} = A_{\mathrm{r}}^{\{j\}} \cdot \left[A_{\mathrm{r}}^{\mathsf{h}\{j+1\}} \left(\bar{p}^{\eta\{j\}} \right) \right] , \tag{6.6}$$

for $1 \leq j < n_{\mathsf{s}} - 1$, and assuming that, for the smallest scale, it reads

$$A_{\mathrm{r}}^{\mathsf{h}\{n_{\mathsf{s}}\}} = A_{\mathrm{r}}^{\{n_{\mathsf{s}}\}} . \tag{6.7}$$

Here, the index $(\bullet)^{\mathsf{h}}$ has been introduced to denote the homogenized quantity and the abbreviation

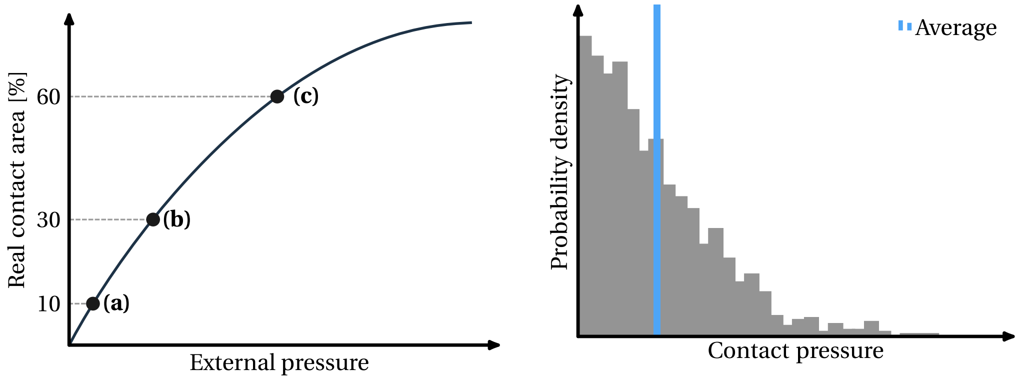
$$A_{\mathrm{r}} = \frac{A_{\mathrm{c}}}{A} \tag{6.8}$$

for the relative contact area. From a physical standpoint, this homogenization methodology originates from the fact that the variable of interest represents a *fraction* of some quantity. Considering that, at the macro-scale, the contact pressure field is generally distributed along many disconnected regions on the contact interface, the following scale setup can be thought of as being loaded—in an average sense—by the local pressure distribution, in every region.

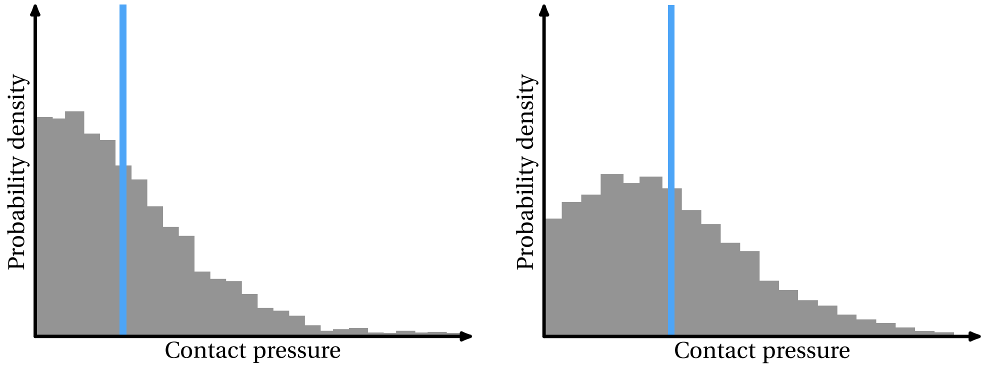
The scale transition of the methodology mentioned earlier can be categorized as a *zeroth-order* approximation, where only the average pressure is passed to the following scale. However, as investigated in great detail in Section 6.4.1, this approximation might be inappropriate in some situations, especially at later stages near full contact conditions. In order to understand this more clearly, in Figure 6.5, the typical contact pressure distribution at different stages for elastic materials is represented, including the corresponding average contact pressure. It can be verified that the contact pressure is non-Gaussian throughout the entire load range, especially for the early stages of lower contact area fractions, in which the contact pressure field reveals a very pronounced skewness. As the external pressure increases, the real contact area raises continuously, and the overall contact pressure distribution spreads along the pressure axis. The average contact pressure, although being able to measure the central tendency, compacts all the information available into a single value, thus being inadequate to describe the statistics of the contact pressure field.

6.3.5 Enhancement of the information passing scheme

To incorporate additional information in the formulation and capture more precisely the pressure distribution, a new information passing scheme is proposed. Starting at the macro-scale, for a given load step, the distribution of contact pressures is compacted into n_{b} bins, between 0 and the maximum—this process can be thought of as a discretization of the pressure field. As histograms, this gives a rough sense of the probability density of the underlying contact pressure distribution at the interface. In doing so, the question is: how can data binning be performed in a physically meaningful way? It should be kept in mind that the number of bins is strictly related to the compromise between probability



(a) Real contact area: 10 %



(b) Real contact area: 30 %

(c) Real contact area: 60 %

Figure 6.5: Typical contact pressure distribution for elastic materials at different real contact area fractions: (a) 10%, (b) 30% and (c) 60%.

density estimation and noise: the use of narrower bins gives precision to the density function, while wider bins lead to a reduction of the noise originated by the discretized nature of the problem.

The suggested approach is based on the relation between the physical lengths that can be associated with each scale. Assuming a uniformly structured discretization of the rough interface, each nodal contact pressure value can be thought to act over a region with size $\Delta x^{\{j\}}$, directly related to the mesh step (see Figure 6.6). Since each bin contains a given number of samples (pressure values), it can be linked with a physical length. Thus, rather than the trivial choice of evenly spaced bins, the pressure distribution is discretized using *equiprobable bins*. This assures an approximately equal number of samples (pressure value) in each bin, i.e., they are associated with a constant physical length. Fixing the length contained in each bin also means that the number of bins grows continuously with the contact area, as more nodes come into contact. As demonstrated in Section 6.3.6, the proposed methodology is expected to be particularly adequate for the region of high contact area fraction. The number of nodes in contact allows for a higher number of bins with narrower limits. Since the area of each bin is equal (recall that the probability of a bin is given by its area), another feature of this methodology is that it aims at capturing well gradients in the probability density function. This is, ranges of pressure with high probability are discretized more accurately, in exchange for a lower number of bins in regions with low probability.

Consider the rough contact problem at a generic scale $\{j\}$, modelled with a rough profile with length $L^{\{j\}}$. The extent of the contact area associated with each pressure bin is given by $n_{\mathrm{n}}^{\{j\}} \cdot \Delta x$, where $n_{\mathrm{n}}^{\{j\}}$ is the number of samples contained in each bin and Δx is the uniform mesh spacing. At each scale, one can define a reference or representative length scale $L_{\mathrm{Ref}}^{\{j\}}$, that can generally be expressed in proportion to the long cut-off wavelength of the scale, $\lambda_l^{\{j\}}$, see Couto Carneiro *et al.* (2020) and V. A. Yastrebov, Anciaux, *et al.* (2012). The number of samples contained in each bin, $n_{\mathrm{n}}^{\{j\}}$, is set such that it represents a fraction of the length $L^{\{j\}}$ equal to the ratio of between the reference lengths between the two scales, viz.

$$\frac{n_{\mathrm{n}}^{\{j\}} \Delta x^{\{j\}}}{L^{\{j\}}} \approx \frac{L_{\mathrm{Ref}}^{\{j+1\}}}{L_{\mathrm{Ref}}^{\{j\}}} \Leftrightarrow \frac{L_{\mathrm{Ref}}^{\{j\}}}{L^{\{j\}}} n_{\mathrm{n}}^{\{j\}} \Delta x^{\{j\}} \approx L_{\mathrm{Ref}}^{\{j+1\}} \,, \tag{6.9}$$

see Figure 6.6. This strategy ensures that the maximum number of discrete contact pressure values is equal to the number of reference micro-scale lengths that fit into the reference macro-scale length, in a full-contact condition. Moreover, Equation (6.9) assures proper scaling of the total number of nodes on each bin with the actual length $L^{\{j\}}$ of the RCE, in relation to its representative length $L_{\mathrm{Ref}}^{\{j\}}$. By expressing the representative length $L_{\mathrm{Ref}}^{\{j\}}$ of both scales as a function of the corresponding cut-off wavelength $\lambda_l^{\{j\}}$, and noting that the full length of the current scale is given by $L^{\{j\}} = n_{\mathrm{c}}^{\{j\}} \Delta x$, where $n_{\mathrm{c}}^{\{j\}}$ stands for the total number of interface nodes at the scale $\{j\}$, Equation (6.9) can also be written as

$$\frac{n_{\mathrm{n}}^{\{j\}}}{n_{\mathrm{c}}^{\{j\}}} \approx \frac{\lambda_l^{\{j+1\}}}{\lambda_l^{\{j\}}} \,. \tag{6.10}$$

In order to have the maximum number of bins and, thus, great discretization of the pressure field, the number of samples contained in each bin, $n_{\mathrm{n}}^{\{j\}}$ should be as low as possible, yet, without compromising its physical meaning given by Equation (6.10). It should also be remarked that even though Equation (6.10) has been idealised in a two-dimensional setting, it also seamlessly represents the three-dimension extension, thus providing a very concise and general definition of the algorithmic number of pressure bins.

Lastly, with the pressure field being discretized, each bin needs to be represented by a contact pressure value. One possible definition would be its centre value (i.e., the average of the limits of the bin). Nonetheless, the arithmetic mean of the actual contact pressure values inside each bin is suggested, to preserve as much information as possible from the packed data.[‡] The multi-scale solution is obtained by performing, once again, a multiplicative homogenization step, yet considering the *set* of contact pressure values being transmitted to the next scale. Recalling that each bin is associated with the same probability, the homogenized real contact area fraction for a given scale $\{j\}$ is obtained as

$$A_{\mathrm{r}}^{\mathrm{h}\{j\}} = A_{\mathrm{r}}^{\{j\}} \cdot \left[\frac{1}{n_{\mathrm{b}}} \sum_{b=1}^{n_{\mathrm{b}}} A_{\mathrm{r}}^{\mathrm{h}\{j+1\}} \left(p_b^{\eta\{j\}} \right) \right] \,, \tag{6.11}$$

[‡] Even though not documented (for the sake of brevity), the difference between both options has been investigated and shown to have a negligible impact on the results.

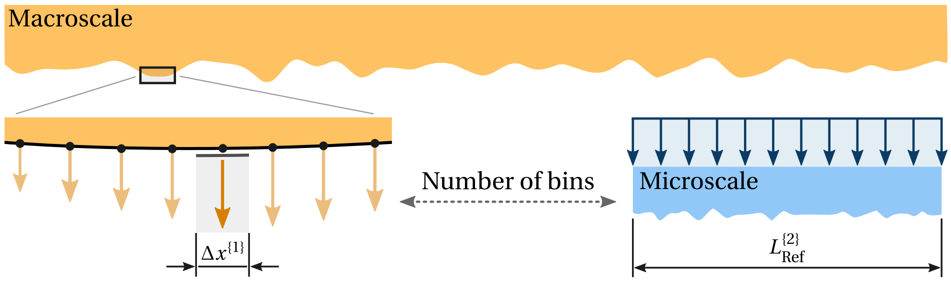


Figure 6.6: Physical relation used to establish the discretization of the contact pressure field. The bins are equiprobable, all containing $n_{\mathrm{n}}^{\{j\}}$ nodes, associated with the mesh step size $\Delta x^{\{j\}}$. The total length associated with each bin is set in relation to the length of the next scale, $L^{\{j+1\}}$.

for $1 \leq j < n_{\mathrm{s}} - 1$, and assuming that, for the smallest scale, once again it reads

$$A_{\mathrm{r}}^{\mathrm{h}\{n_{\mathrm{s}}\}} = A_{\mathrm{r}}^{\{n_{\mathrm{s}}\}} . \tag{6.12}$$

As more than one value is transmitted between scales, the resulting downscaling process has a recursive and cascading character, that grows to reach the last smallest scale. From a practical standpoint, the implementation of the algorithm can be viewed more easily from bottom to top. The procedure starts at the micro-scale, where the contact area evolution is known. Moving to the following upper scale, the discrete contact pressure values (for each bin) at each load step are identified. Then, the homogenized evolution at the current scale can be constructed by applying the homogenization step with the information available at the micro-scale (and interpolating the real contact area fractions, if needed). The process repeats for all scales until reaching the macro-scale, where the final homogenized solution can be obtained.

6.3.6 Combined multiplicative homogenization

As already mentioned, within the proposed information passing scheme, the contact pressure field is better discretized as the contact area grows (because it allows a higher number of bins with narrower limits). This means that, while suitable for high contact areas, for small contact area fractions, the contact pressure is not properly discretized and the overall methodology loses its physical interpretation. When just a few nodes are in contact, the change in the number of bins originates a jump on the homogenised result. In order to overcome this difficulty, a multiplicative homogenization scheme based on the combination of the two described methodologies is proposed. The idea consists of defining the homogenized solution as the affine combination between both strategies, using the real contact area fraction itself as a coefficient, i.e.

$$A_{\mathrm{r}}^{\mathrm{h}\{j\}} = A_{\mathrm{r}}^{\{j\}} \cdot \left[\left(1 - A_{\mathrm{r}}^{\{j\}}\right) \cdot A_{\mathrm{r}}^{\mathrm{h}\{j+1\}} \left(\bar{p}^{\eta\{j\}}\right) + \frac{A_{\mathrm{r}}^{\{j\}}}{n_{\mathrm{b}}} \sum_{b=1}^{n_{\mathrm{b}}} A_{\mathrm{r}}^{\mathrm{h}\{j+1\}} \left(p_b^{\eta\{j\}}\right) \right] . \tag{6.13}$$

Comparing with Equation (6.6) and Equation (6.11), this method essentially establishes a compromise between both. For small contact area fractions, the solution is mainly given by the average contact pressure. For near full contact conditions, the solution is dominated by the discretization of the pressure field with equiprobable bins. Both solutions are combined for values in between, and the transition between them occurs in a smooth fashion.

In Figure 6.7, a schematic representation for the simplest case of splitting the original problem into $n_\mathrm{s} = 2$ scales is given. After performing the topography decomposition (step A), for each value of external pressure at the macro-scale, it is necessary to evaluate the average contact pressure (step B) and to discretise the corresponding contact pressure distribution (step C). For the sake of simplicity, in the example, only $n_\mathrm{b} = 3$ bins are represented. Then, the average contact pressure and the representative value of each bin are transmitted to the micro-scale (step D). Finally, the homogenized response is obtained (step E) by combining both solutions:

$$A_\mathrm{r}^\mathrm{h} = A_\mathrm{r}^{\{1\}} \cdot \left[\left(1 - A_\mathrm{r}^{\{1\}}\right) \cdot A_\mathrm{r}^{\{2\}} \left(\bar{p}^{\eta\{1\}}\right) + \frac{A_\mathrm{r}^{\{1\}}}{n_\mathrm{b}} \sum_{b=1}^{n_\mathrm{b}} A_\mathrm{r}^{\{2\}} \left(p_b^{\eta\{1\}}\right) \right]. \tag{6.14}$$

For the general case of more than two scales, the problem has a recursive character. For any intermediate scale, each value of external pressure to be analysed requires both the average and the discretization into a given number of bins, that are downscaled to the following smaller scale.

As a final remark, it is important to mention that, many realizations are considered for each scale, due to the inherent statistical nature of roughness. Therefore, within the multi-scale framework, the response at each scale is defined as the average of all the corresponding realizations. Under these circumstances, Equation (6.10) should be interpreted on an ensemble basis, concatenating the information of all realisations.

6.3.7 Contact area and contact pressure field evaluation

In all discretization-based techniques, there is always an error associated with their grid/mesh, which affects the accuracy with which both the real contact area and contact pressure field are evaluated (V. A. Yastrebov, Anciaux, *et al.*, 2017a). For multi-scale algorithms, in particular, this topic should be carefully addressed, because the source of error can even be amplified due to the multiplicative nature of the homogenization process.

In the simplest way to approach the problem, two main definitions are typically employed. The first one consists of simply defining the real contact area based on the ratio between the number of active nodes and the total number of nodes. The second one is based on the sum of all the individual areas of all the active elements—here, defined as the individual cells of the grid in which all its corresponding nodes are active. The main difference between the two is related to the consideration of isolated nodes. The node-based definition can essentially be viewed as an upper bound and the element-based solution as a lower bound. The solution proposed here is to use the average between the two solutions since, as shown in more detail in Section 6.4.1.1, they both tend to converge to approximately the same value.

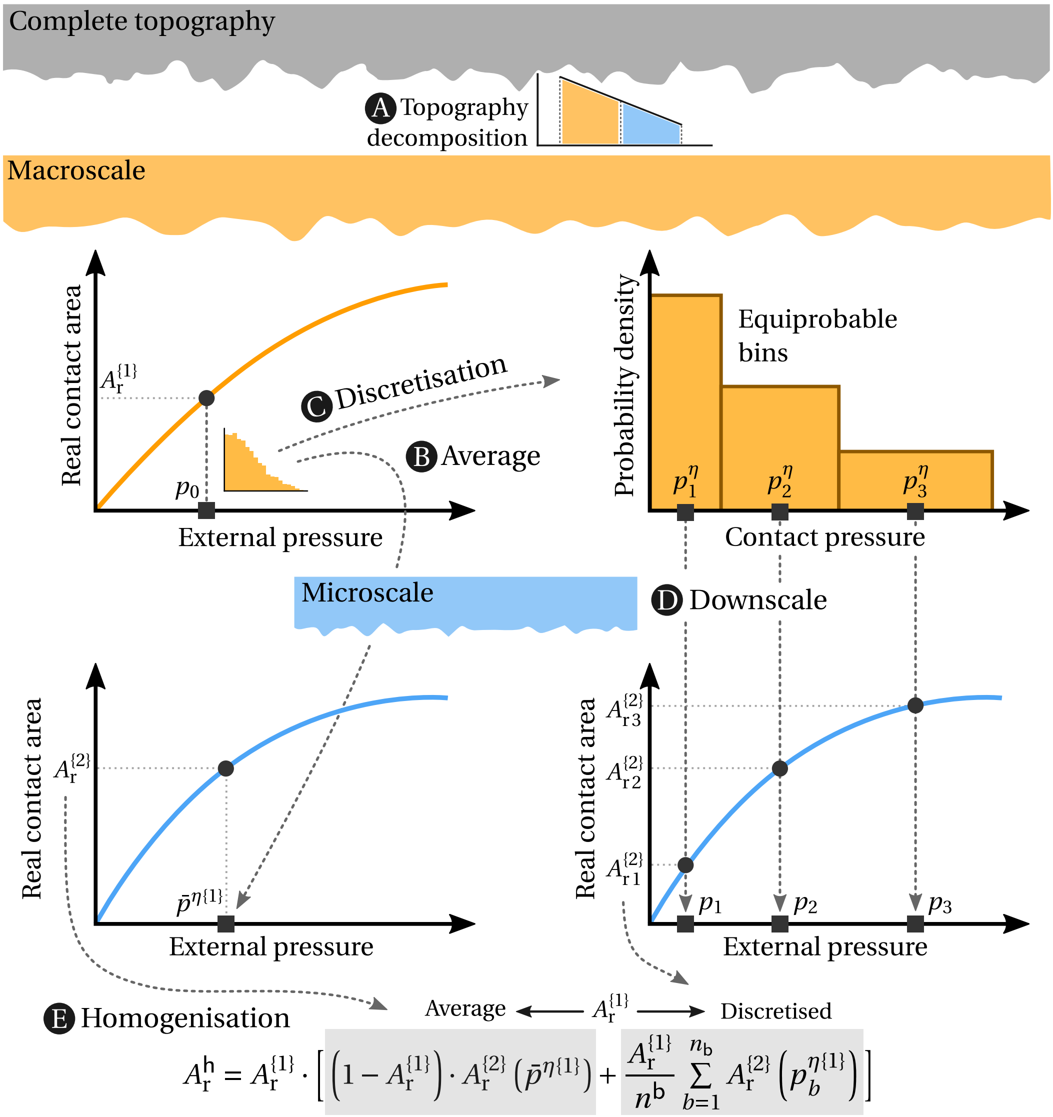


Figure 6.7: Illustration of the multi-scale algorithm for the simplest case of $n_\mathrm{s} = 2$ scales. The five main steps are: **(A)** Surface decomposition; **(B)** Evaluation of the average of the contact pressure; **(C)** discretization of the contact pressure distribution into n_b equiprobable bins; **(D)** Downscaling of all the contact pressure values; **(E)** Combined multiplicative homogenization.

6.4 Numerical results

To demonstrate the validity and efficiency of the proposed multi-scale rough contact algorithm, several numerical examples are presented and discussed in this section. Firstly, the algorithm is validated by comparing the results obtained with single-scale and multi-scale analyses. Additional aspects regarding mesh convergence and the effect of the number of scales are also analysed. Then, the validity of the method is further investigated within the context of elastoplastic rough contact. The computational performance of the multi-scale algorithm is compared with the single-scale model. In the end, based upon the conclusions obtained with two-dimensional simulations, a three-dimensional problem is analysed, comparing the results with several solutions available in the literature.

For the sake of simplicity, in all the examples, only the particular case of a rough body being pressed against a rigid flat surface is considered. The problem setup, properties and employed finite element discretization are schematically represented in Figure 6.8. Only a small portion of the finite element mesh is represented, showing the adopted refined layer near the rough interface and corresponding mesh transition scheme. In order to reduce volumetric locking effects, the block is discretized with the bilinear quadrilateral F-bar finite element. The external pressure p_0 is applied incrementally until full-contact.

Regarding the self-affine rough topography, the following properties are kept constant across all investigations: the long cut-off wavelength λ_l is fixed at 1 μm, the Hurst roughness exponent at $H = 0.8$ and the RMS slope at $\sqrt{m_2} = 0.2$. These are typical values found in the literature. The only property left to characterise the rough topography—the short cut-off wavelength, λ_s—is set as a free variable to investigate different roughness spectral bandwidths. The RCE dimensions are defined according to the strategy presented in Couto Carneiro *et al.* (2020), in which all the dimensions are obtained as a function of the topography roughness model. The length of the RCE, here termed as L, is expressed as a function of the long cut-off wavelength λ_l, whereas the mesh step size, Δx, is defined based on the small cut-off frequency λ_s. As shown in Couto Carneiro *et al.* (2020), the values $L/\lambda_l = 8$ and $\lambda_s/\Delta x = 8$ are the suggested starting values to discretise the problem properly. Additionally, both the total height of the substrate, H_{Sub}, and of the refined layer, H_{Ref}, are defined based on the value of the RMS height, σ_z. The value of $H_{\mathrm{Sub}} = 160\,\sigma_z$ is employed for the total height and the refined layer is set at $H_{\mathrm{Ref}} = 30\,\sigma_z$. Lastly, the material elastic behaviour is modelled by considering the Young's modulus $E = 210$ GPa and Poisson coefficient $\nu = 0.3$.

Remark. *The Hurst exponent is intrinsically associated with the contribution of higher frequencies on the topography. This means that for low pressures, the load-bearing capacity is directly affected and, therefore, one might expect a significant change in the behavior of the multi-scale algorithm. However, even though not documented, the numerical validation with two scales has also been carried out using different Hurst exponents and its effect on the results is negligible. Therefore, it is expected that all the conclusions obtained with* $H = 0.8$ *remain valid for other values of the Hurst exponent.*

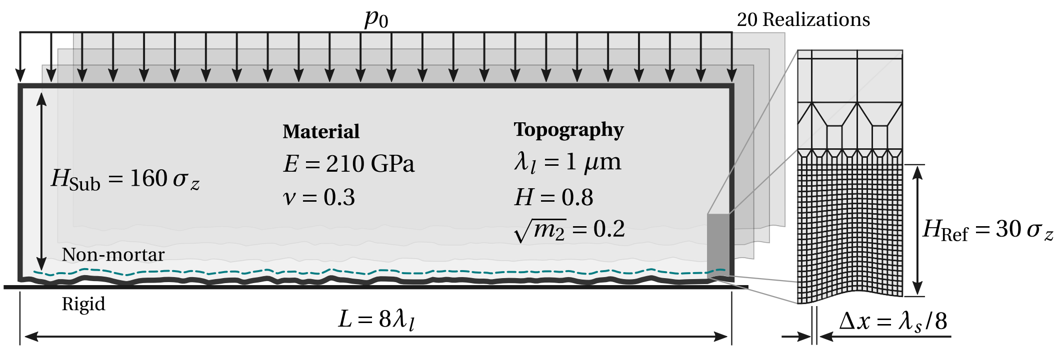


Figure 6.8: Problem setup and employed finite element discretization.

6.4.1 Numerical validation with two scales

In the first stage, the rough contact of an elastic body is investigated by employing two-dimensional RCEs containing random self-affine topographies with three roughness bandwidths:[§]

$$\lambda_l/\lambda_s = \{8, 32, 64\}\ .$$

To assess the validity and global accuracy of the proposed method, the evolution of the real contact area obtained with Direct Numerical Simulation (DNS)—i.e., by employing a single scale containing the original topography—is taken as reference and compared with the multi-scale algorithm proposed in this work and Wagner, Wriggers, Veltmaat, *et al.*, 2017.

The results are gathered in Figure 6.9, where, for each bandwidth, the entire range of contact area evolution and a snippet of the light contact region are represented. The background colours indicate the number of equiprobable bins used by the multi-scale method to discretise the contact pressure field. Overall, the proposed multi-scale algorithm is able to capture the original problem with an accuracy that, as expected, tends to improve with the increase in roughness bandwidth. This can be related to the violation of scale separation, which is expected to attenuate for increasing bandwidths. This is particularly noticeable in the region of low pressure, where contact spots occur at the top of some asperities. Therefore, the impact of changing its spectral content becomes more evident. Near the full contact condition, though, the algorithm reveals an accuracy that does not depend so much on the roughness bandwidth. In fact, the enhancement of the information passing scheme is mostly visualised in this region. Compared with the algorithm based on the average of the normal contact pressure, the real contact area is better estimated. The transition between both regimes is also correctly captured, proving that the combined multiplicative homogenization is able to couple both solutions in a smooth and accurate manner. Lastly, by analysing the number of bins, it is verified that it increases with both the contact area fraction and the roughness bandwidth. This result reflects the direct relation between the number of bins and the total number of active contract nodes, c.f. Equation (6.10).

[§] Having specified the spectral properties of the topography, the associated Nayak parameter α can be obtained, referring to Couto Carneiro *et al.* (2020) for the corresponding analytical expression.

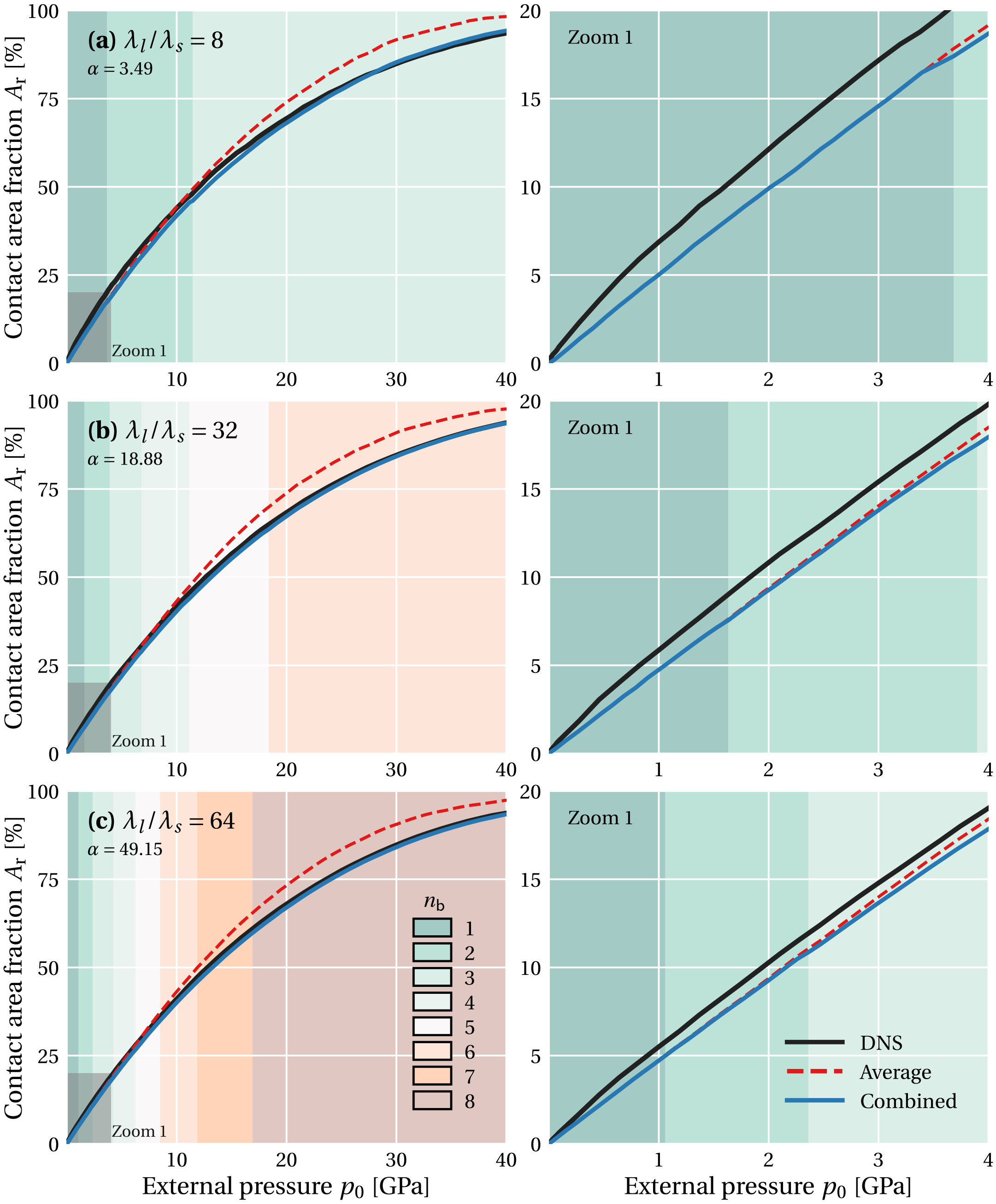


Figure 6.9: Evolution of the contact area fraction A_r for three different roughness bandwidths: (a) $\lambda_l/\lambda_s = 8$; (b) $\lambda_l/\lambda_s = 32$; (c) $\lambda_l/\lambda_s = 64$. The black line is the reference results obtained with DNS. The multi-scale algorithms based on the average contact pressure and the combined homogenization are represented by the red dashed and solid blue lines.

6.4.1.1 Mesh convergence

In order to evaluate the validity of the presented so far, a mesh convergence analysis focused on the contact area fraction and contact pressure field is carried out. Only the roughness bandwidth of $\lambda_l/\lambda_s = 8$ is considered, as the sensitivity to discretization errors tends to increase for topographies with narrower bandwidths (Couto Carneiro *et al.*, 2020). The main objective is to analyse the variation of the results, considering the same roughness profiles, yet with increasing mesh refinements:

$$\lambda_s/\Delta x = \{4, 8, 16, 32\} \ .$$

In Figure 6.10, the evolution of the contact area fraction for both DNS and multi-scale solutions is represented, containing once again snippets of different regions of interest. For the sake of clarity, only the multi-scale algorithm based on the combined multiplicative homogenization algorithm is presented, bearing in mind that similar behaviour is observed for the algorithm based on the average contact pressure.

As expected, the results show a global converging behaviour, with the discretizations $\lambda_s/\Delta x = 16$ and 32 being practically overlapped. A very similar progression is observed for both DNS and multi-scale solutions, meaning that the information passing scheme is not interfering with the expected discretized nature of the problem. The example with $\lambda_s/\Delta x = 4$ deviates the most from the converged solution, especially for the region of near full contact. This result is consistent with the conclusions presented in Couto Carneiro *et al.* (2020), where the minimum value of $\lambda_s/\Delta x = 8$ is recommended. This applies to both single-scale or multi-scale analyses.

Within the multi-scale algorithm, aside from the contact area fraction, the contact pressure field also plays a crucial role in the final homogenized result. Thus, instead of only looking at the contact area evolution throughout the entire load range, an additional mesh convergence investigation focused on the convergence of both the contact area fraction and average pressure for just the macro-scale at a given load step is presented. Figure 6.11 shows the average contact pressure and contact area fraction for the external pressure of $p_0 \approx 10.4087$ GPa. The shaded regions denote the range between the solution with and without isolated nodes for each of the 20 realizations. The solid lines represent their average (the values used within the multi-scale algorithm), indicating the associated standard deviation.

The average pressure and the real contact area converge with increasing mesh refinements. However, they have the opposite tendency: the contact area decreases as it converges, while the average contact pressure increases. Moreover, the discretization sensitivity is slightly more significant for the average contact pressure. For this particular example under analysis, the recommended discretization $\lambda_s/\Delta x = 8$ used so far is arguably the limit below what the results may be affected. As a final remark, it is essential to recall that the particular case under analysis is the smallest of the analysed roughness bandwidths. Even though not documented here, the sensitivity to both the discrete and statistical nature of random surface topographies tends to decrease for larger bandwidths (Couto Carneiro *et al.*, 2020). Ultimately, this means that the results shown in Figure 6.9 are properly discretized and, thus, validates the obtained conclusions regarding the accuracy of the multi-scale algorithm.

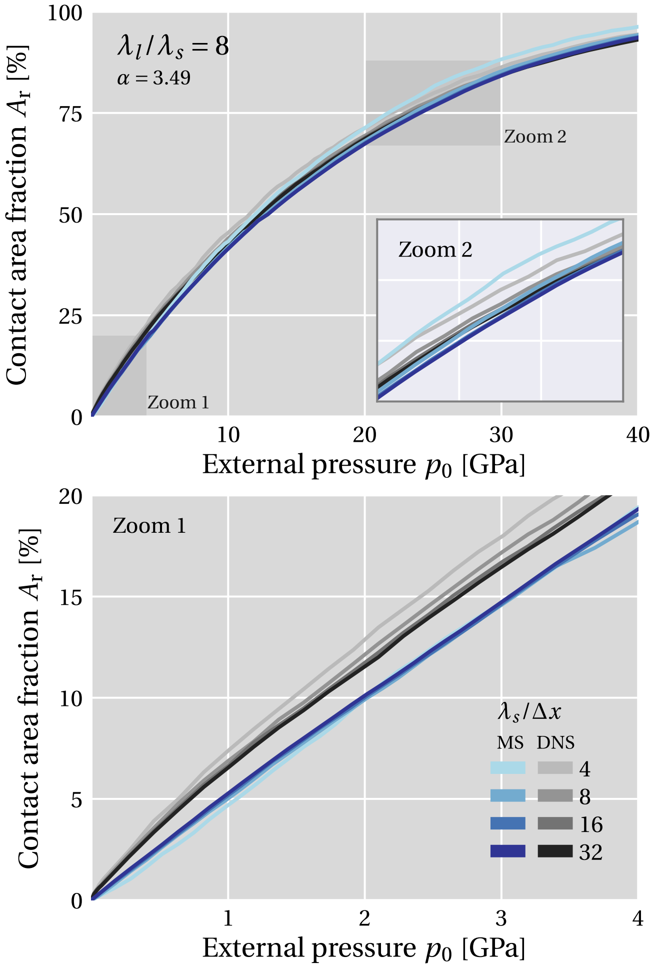


Figure 6.10: Mesh convergence for the roughness bandwidth of $\lambda_l/\lambda_s = 8$, considering a step mesh size of $\lambda_s/\Delta x = \{4, 8, 16, 32\}$. The black gradient represents the DNS solution and the multi-scale (MS) results are in blue.

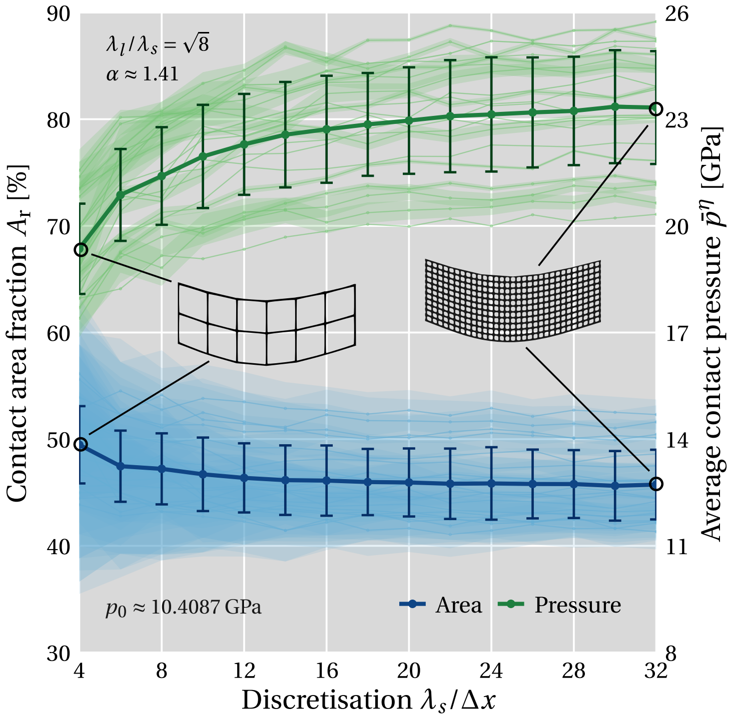


Figure 6.11: Contact area fraction and average contact pressure evolution for the macroscale at an external pressure of $p_0 \approx 10.4087$ GPa with increasing mesh refinement levels. The shaded regions denote the range between the solutions with and without isolated nodes for each of the 20 realizations. The solid lines denote their average, together with the associated standard deviation.

6.4.2 Analysis of the number of scales

One of the fundamental aspects of the multi-scale techniques discussed in this work is the number of scales into which the rough topography is decomposed. To understand how it impacts the homogenized solution and give insight into the recommended guidelines for its choice, this section presents an investigation in which the same rough topography is decomposed into an increasing number of scales. Since the bandwidth of each individual scale becomes gradually smaller, a rough topography with $\lambda_l/\lambda_s = 64$ is analysed (thus allowing for a great number of scale divisions). Figure 6.12 shows the reference solution obtained with DNS and both multi-scale algorithms with the following decompositions:

$$n_{\mathrm{s}} = \{2, 3, 4, 5\} \ .$$

The colour gradients represent the number of scales and, once again, different snippets of the contact area curve are highlighted.

Qualitatively, the main conclusion from the obtained results is that the accuracy of the multi-scale models is expected to decrease with the number of scales. Quantitatively, however, the behaviour of the multi-scale algorithms is different. The combined algorithm reveals to be more sensitive to the number of scales in the region of low to mid contact area fractions, with practically no impact on results at near full contact conditions. The average, while being more stable in the transition between low to mid contact area fractions, is very sensitive at later stages of near full compression, with a maximum variation that is greater than the one obtained with the combined algorithm. This means that the new proposed algorithm is less sensitive to the number of scales which, as a general rule, should be as low as possible to achieve the maximum accuracy potential.

The selection of the number of scales is mainly guided by the computational power available in each situation. For example, consider a rough contact problem involving a roughness spectrum so vast that it renders the direct numerical simulation completely intractable with the resources at hand. One can still obtain reasonably accurate results by employing the present multi-scale algorithm in a two-scale setup, as long as each scale can fit within the resources available. If this is not the case, the dimension of the problem at each scale can be further reduced by moving to a three-scale setup, yet sacrificing the overall accuracy, as shown in Figure 6.12. Ultimately, the number of scales should be set according to a trade-off between the desired accuracy and the computational power available.

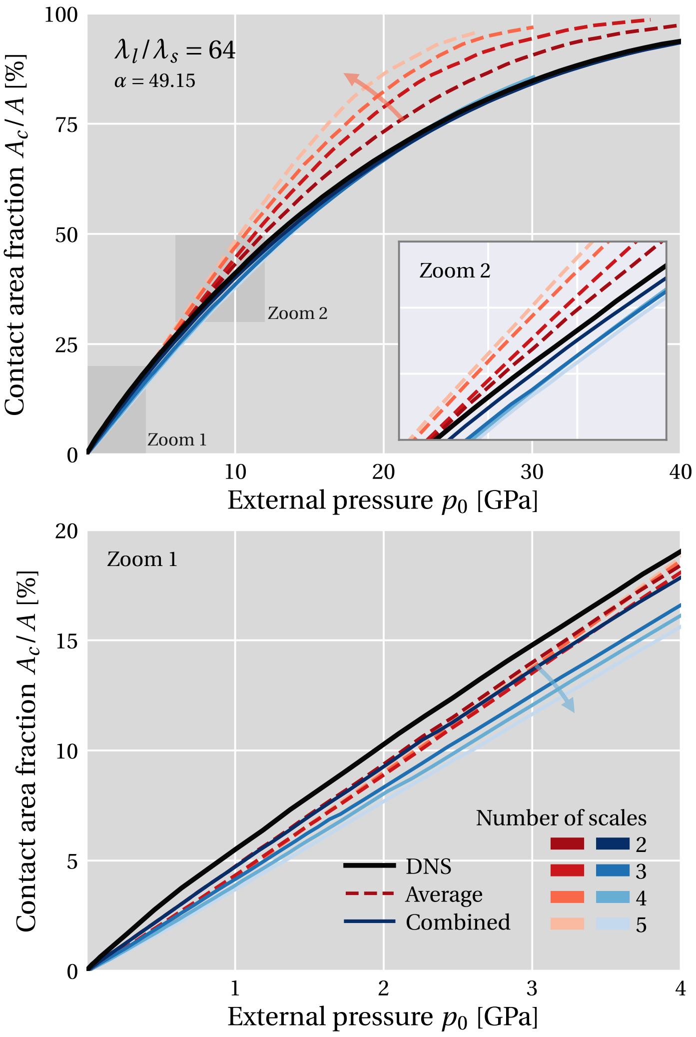


Figure 6.12: Contact area fraction for the different number of scales for the roughness bandwidth $\lambda_l/\lambda_s = 64$. The black line is the reference obtained with DNS, the dashed lines denote the multi-scale solution based on the contact average and the solid lines the combined algorithm. The associated colour gradient represents the number of scales.

6.4.3 Numerical validation for elastoplastic rough contact

In all the examples analysed so far, the material constitutive behaviour is assumed to be elastic. Locally at each contact spot, however, the material may undergo stress levels able to induce permanent plastic deformation on individual asperities. This has a great impact on the physical behaviour of the contact interface, for instance, changing the contact stress distribution and real contact area. In order to investigate its effect on the multi-scale algorithm, the numerical model complexity is further increased by modelling the elasto-plastic response of the material with a J_2 isotropic plasticity law. A validation study similar to the one presented in Section 6.4.1 is performed, considering three yield stresses:

$$\sigma_{\mathrm{y}} = \{2, 4, 10\}\ \mathrm{GPa}\,.$$

For numerical stability, 5% hardening is considered. For the sake of brevity, only the roughness bandwidth $\lambda_l / \lambda s = 32$ is analysed.

In the first step, the impact of plasticity on the contact stress field is investigated by looking at the contact pressure distribution at a given stage of the compression process. Figure 6.13 shows the results obtained for each material at the contact area fraction of $A_{\mathrm{r}} = 20\%$, also indicating the value of nominal pressure necessary to achieve that area fraction. For a more precise interpretation of the results, the deformed configuration of a single asperity is included, showing the accumulated plastic strain and nodal contact stresses.

The distributions show a relation between the admissible stress field on the material and the developed contact interaction at the interface. As the yield stress decreases, the contact pressure field range tends to get narrower. The distribution becomes skewed, with the pressure field being almost constant for the yield stress $\sigma_{\mathrm{y}} = 2$ GPa. This result is expected, as the stress state developed at the interface is deeply connected to the constitutive behaviour of the bulk material, both reflecting the yield stress limit. This type of stress saturation is also consistent with similar studies presented in Gao *et al.*, 2006; Pei *et al.*, 2005. Moreover, the nominal contact pressure required to achieve the contact area fraction of $A_{\mathrm{r}} = 20\%$ lowers with the yield strength, thus being already possible to expect accentuated differences in the contact area evolution curves. Finally, by inspecting the accumulated plastic strain and contact pressure nodal values, all the conclusions mentioned so far can be visualised. As the yield stress decreases, the plastic deformation spreads around the asperity (forming a typical pattern of an arc around it), and the nodal contact stresses range decreases.

The contact area curves for all the materials are shown in Figure 6.14, together with the results obtained with the elastic material (for reference). As expected, the material response has a significant impact on the real contact area. At the nominal pressure of $p_0 = 10$ GPa, while the contact area fraction is at $A_{\mathrm{r}} \approx 30\%$ for the elastic material, the elastoplastic material with $\sigma_{\mathrm{y}} = 2$ GPa is already at full contact. This comes as a natural consequence of the reduced load capacity of the materials with lower yield strength, which accentuates the rate of contact area fraction evolution. Regarding the accuracy of the multi-scale algorithms, similar conclusions obtained in the elastic case are verified in the elastoplastic material. However, the algorithm based on the average contact pressure becomes closer to the remaining solutions—the reason for that is due to the statistical properties of the contact stress field. Because the contact pressure becomes almost con-

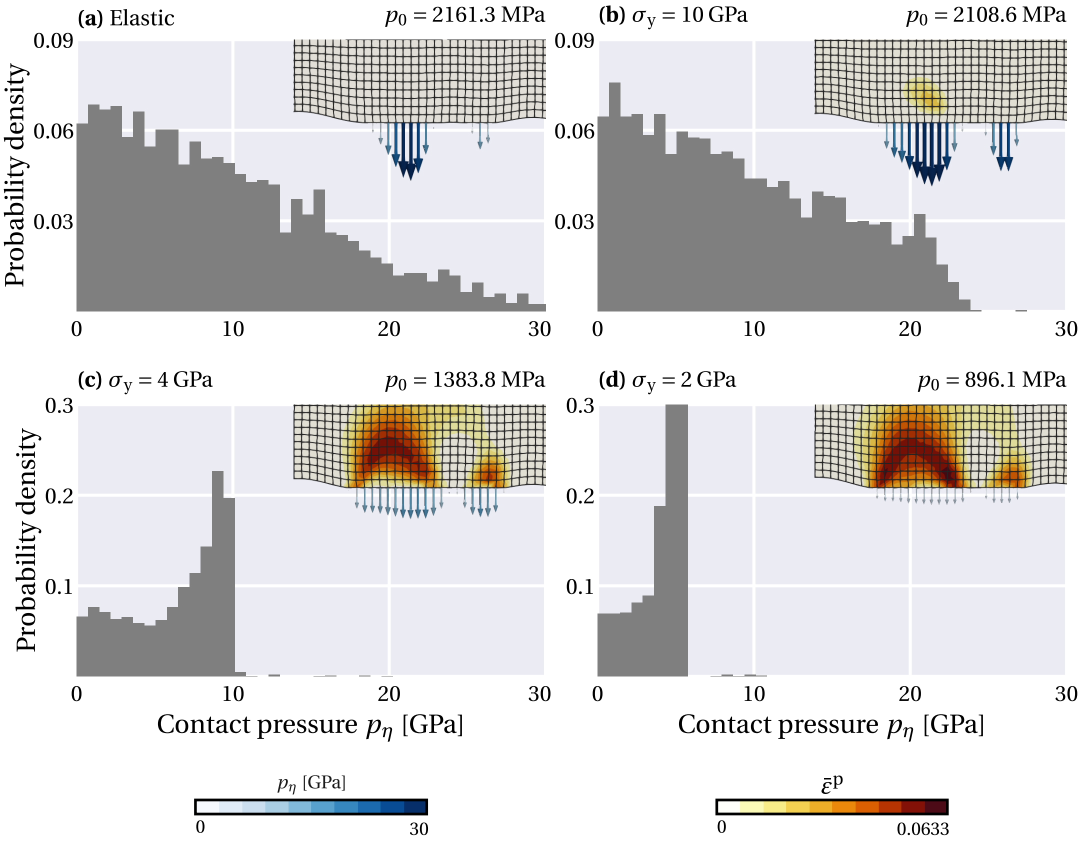


Figure 6.13: Contact stress distribution at the contact area fraction of $A_r = 20\%$ for (a) elastic material and elastoplastic materials with yield stresses: (b) $\sigma_y =$ 10 GPa; (c) $\sigma_y = 4$ GPa; (d) $\sigma_y = 2$ GPa. The nominal contact pressure necessary to achieve that contact area fraction is indicated, together with a snippet of the deformed configuration of a single asperity, showing the distribution of accumulated plastic strain and nodal contact stresses.

stant with the decrease in the yield strength, the discretization using equiprobable bins practically overlaps with the use of the arithmetic average. Nevertheless, the combined algorithm can keep a consistent level of accuracy between all the materials, particularly at the region of near full contact.

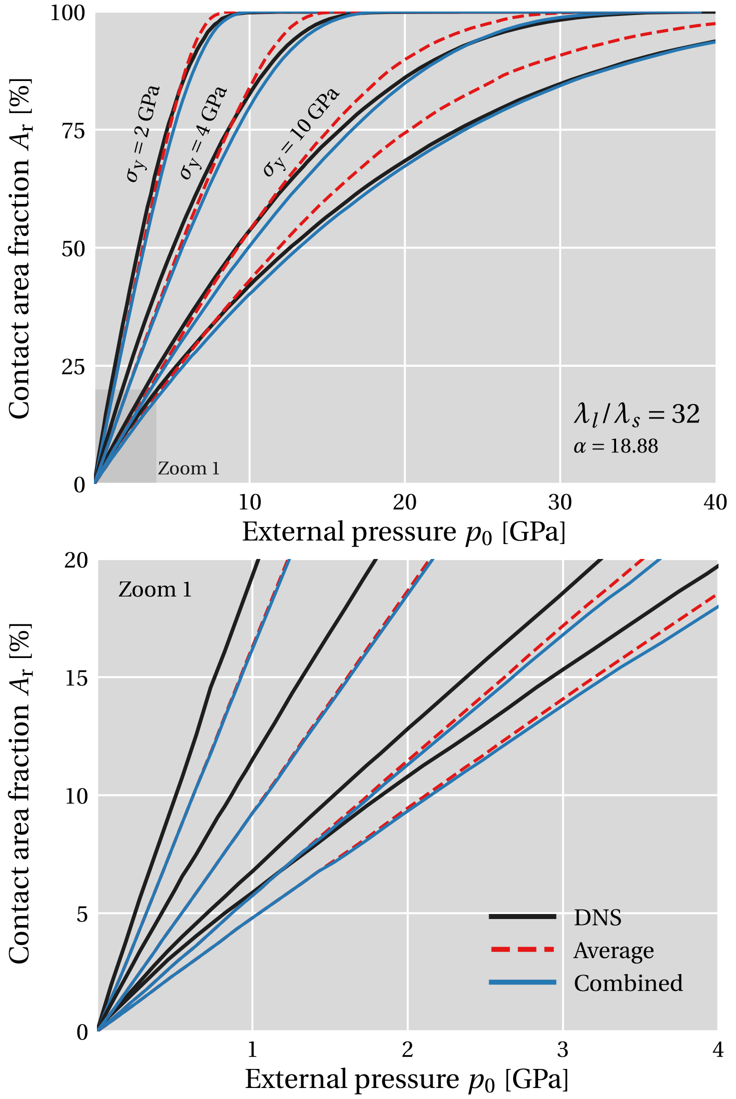


Figure 6.14: Evolution of the contact area fraction A_r for different materials and roughness bandwidth of $\lambda_l/\lambda_s = 32$. The black solid line is the reference result obtained with DNS, the red dashed line the multi-scale algorithm based on the average contact pressure and in solid blue the combined multi-scale algorithm.

6.4.4 Computational performance

This work focuses on the modelling of rough contact homogenization using the finite element method combined with the dual mortar contact algorithm. It establishes, arguably, one of the most versatile and accurate methodologies to carry this type of analysis. However, the main disadvantage regards its computational cost, which is expected to follow a power law with the number of degrees of freedom of the numerical model. The multi-scale strategies discussed throughout this work are one of the possible solutions to circumvent this problem. Having investigated, so far, the possible trade-off in terms of validation and accuracy, this section focuses on measuring the improvement in computational performance. To do so, the calculation time required to obtain the contact area curve for an elastic material with increasing roughness bandwidth is compared, showing in Figure 6.15 the results obtained with the DNS and combined multi-scale algorithm.

From the analysis of the evolution of the DNS method, it is possible to conclude that the expected power-law growth is observed. This takes into account that the number of degrees of freedom also follows approximately a power-law with the roughness bandwidth. The roughness bandwidth $\lambda_l/\lambda s = 62$ takes approximately 40 times the calculation time of the bandwidth $\lambda_l/\lambda s = 6$. The computation time of the combined multi-scale algorithm also increases more slowly than the single-scale model and, for the considered roughness bandwidth range, is at least one order of magnitude faster. Comparing the computational times between distinct number of scales, it can be seen that, despite having more RCEs to be analysed in total, the model with $n_\mathrm{s} = 3$ scales is computationally faster than the model with $n_\mathrm{s} = 2$ scales. This means that, more than the total number of simulations on the database, the main factor driving the computational time of the multi-scale scheme is the complexity of each individual scale. As a final remark, it is essential to mention that despite the multi-scale model adopting $n_\mathrm{s} = 3$ scales being faster, the accuracy of the results is not as good as the case with $n_\mathrm{s} = 2$ scales, c.f. Section 6.4.2. Therefore, as a general rule of thumb, more than 2 scales should only be employed when the computational complexity of the model requires.

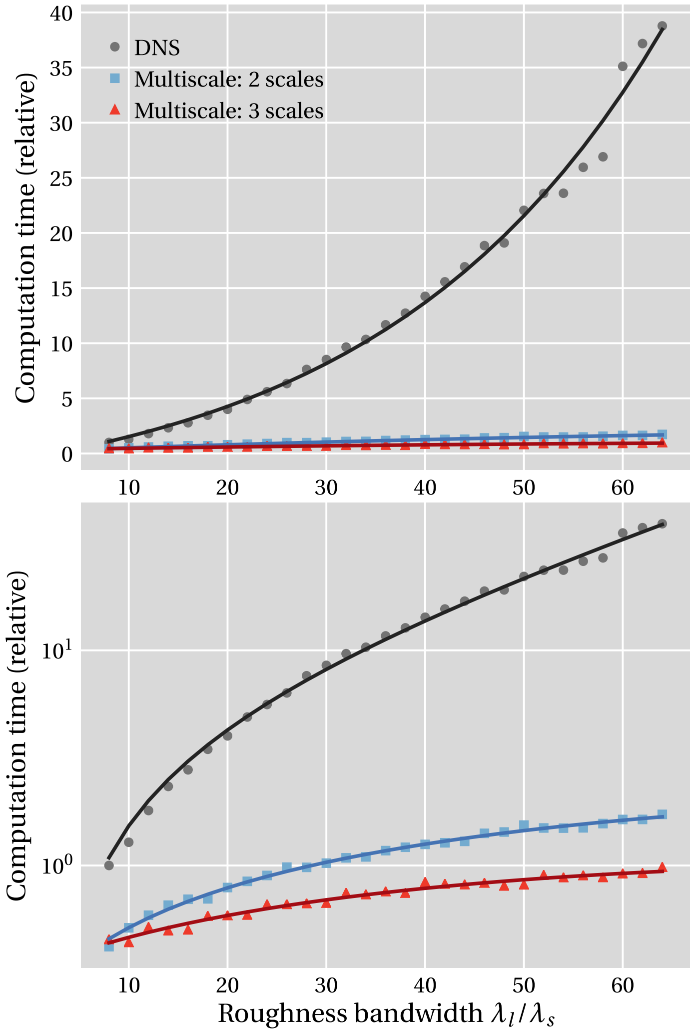


Figure 6.15: The computational time for the DNS and multi-scale solutions with two and three scales for a range of roughness bandwidth. The time of the DNS algorithm for the bandwidth of $\lambda_l/\lambda s = 6$ is taken as a reference. The solid line is the numerical fit, showing the results in linear and logarithmic scales on the left and right, respectively.

6.4.5 Comparison with other models

The last numerical study focuses on the comparison between the proposed multi-scale algorithm and other solutions available in the literature. Persson's contact theory (Persson, 2001b,a) is considered and two solutions within the class of multiasperity models: the asymptotic linear limit by Bush-Gibson-Thomas (BGT) (Bush *et al.*, 1975) and the Greenwood-Williamson simplified elliptic model (GW-SE) by Greenwood (2006). The numerical results obtained with the boundary element method provided in V. A. Yastrebov, Anciaux, *et al.* (2017b) are also represented.

In order to allow a fair direct comparison between the multi-scale algorithm and the remaining solutions, the numerical model described so far is extended to three dimensions. Due to the additional complexity of the numerical model, the DNS model would lead to impractical computation times and, thus, only the multi-scale solution is provided. Nonetheless, this is done bearing in mind the conclusions obtained with two-dimensional simulations. The roughness bandwidth is set to the value of $\lambda_l/\lambda_s = 64$, i.e., the largest bandwidth of the analysed set. It should be noted that the homogenization algorithm remains the same regardless of the number of spatial dimensions. Even though the contact stress field and contact area fraction might be evaluated using any physical model, the post-processing algorithm only needs input from the database of contact pressure and contact area fraction values. This aspect is an advantage of the proposed multi-scale algorithm and, therefore, is applied to Persson's model (since it gives a solution for both the area fraction and contact pressure field). This allows us also to investigate its behaviour beyond the realm of the finite element method.

An exemplary finite element mesh is represented in Figure 6.16, made up of 8-noded F-bar hexahedron mixed with 4-noded pyramidal and 6-noded wedge elements at the mesh transition layer. Compared with the 2D case, some discretization parameters have been relaxed, to obtain a problem that can still be solved under the 24 hours mark using a desktop workstation. The length of the RCE is reduced to $L = 4\lambda_l$, the mesh step size to $\Delta x = \lambda_s/6$ and the refined region near the rough topography is made of 4 layers of finite elements. It results in a finite element mesh of $\approx$ 200k nodes, with 200×200 elements along each direction at the interface. All the remaining properties are the same as in the 2D case, c.f. Figure 6.8. Each realization requires $\approx$ 20h to solve, using $\approx$ 100 increments that typically require 3-4 iterations of the semi-smooth Newton-Raphson iterative scheme (considering a convergence tolerance of 1×10^{-6} for the relative residual). Note that, even using the reduced parameters, the DNS simulation of the single scale problem would require a mesh with 1500×1500 elements at the interface, thus leading to an impractical problem with millions of degrees of freedom.

The evolution of the real contact area fraction is shown in Figure 6.17. Firstly looking at the region of light contact, all the numerical solutions are contained within limits provided by the GW-SE and asymptotic BGT. Overall, a good agreement between Persson's model and the FEM solution is found, with the BEM predicting a higher real contact area fraction. In fact, this result is observed throughout the entire load range. The FEM solution estimates a lower real contact area fraction at the middle region, between light and full contact. Another important result is that the proposed multi-scale algorithm can capture the original solution when applied to Persson's model, with an accuracy level consistent with the results obtained with finite elements. The homogenization based

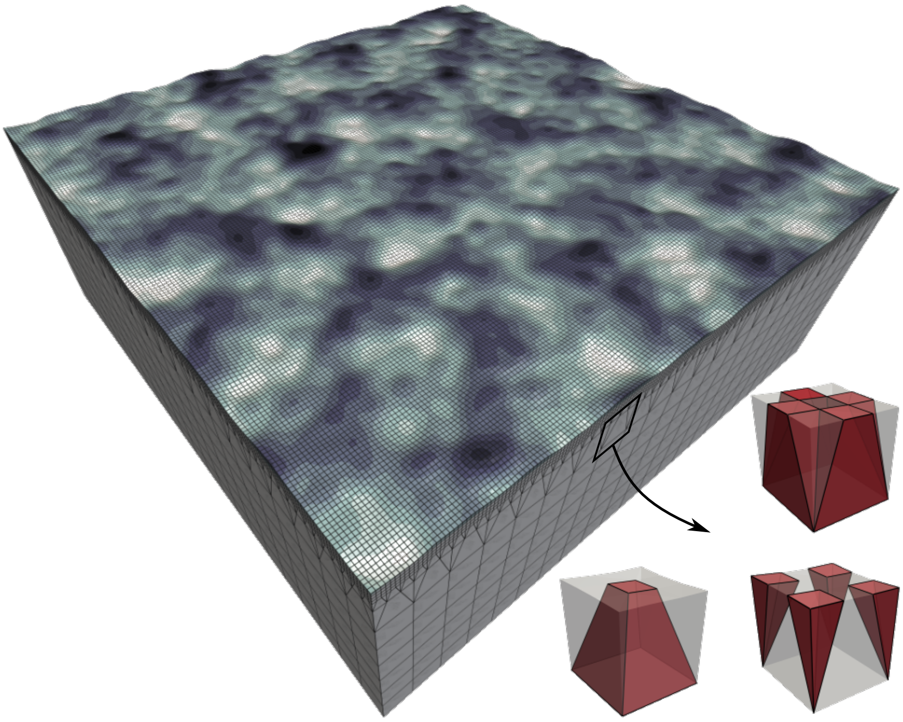


Figure 6.16: Finite element discretization and mesh transition scheme. The colours represent the initial height of the interface nodes.

on the average contact pressure tends to overestimate the real contact area fraction for the region of near full contact. This fact shows that the proposed algorithm is suitable to be applied to frameworks other than the FEM, thus revealing that the statistics of the contact pressure field are properly captured. Analysing the results obtained with the elastoplastic material, similar results as in Section 6.4.3 are obtained: the contact area fraction evolution rate increases, as a result of the load capacity limit introduced by the plastic deformation behaviour.

Other than the contact area fraction evolution, one can look at the morphology of the contact area regions. Figure 6.18 shows the active contact area for the elastic and elastoplastic materials at different stages of the compression process. The main observation is that the morphology of the active contact zone is practically similar. This leads to the conclusion that despite the evolution of the real contact area being highly affected by material behaviour—in this case, two materials with different load capacity—the shape of the contact clusters will be mostly identical. A possible explanation for this result is that both material models are isotropic, with the saturation character impacting only the distribution of contact stresses and the rate of change of the contact area fraction.

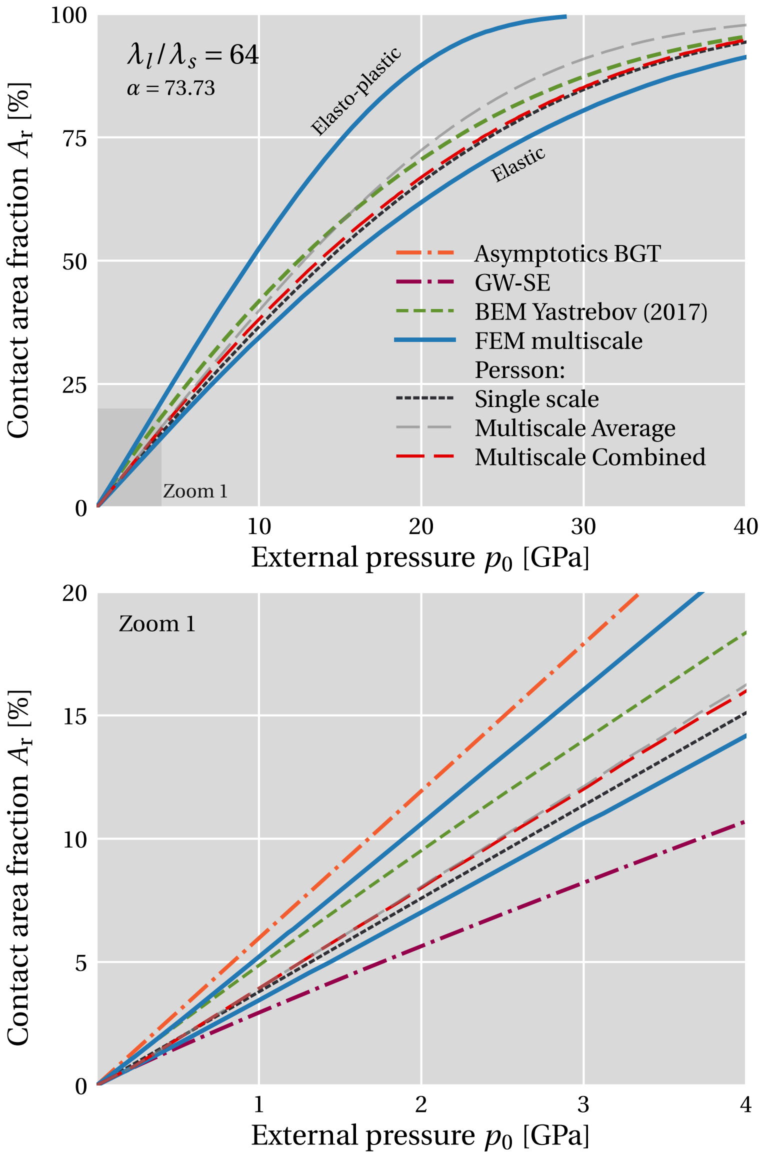


Figure 6.17: Contact area fraction for elastic rough contact according to different models: multi-scale (FEM and Persson), Persson, BEM, Greenwood-Williamson simplified elliptic (GW-SE) and the asymptotic linear limit by Bush-Gibson-Thomas (BGT). The solution obtained with the FEM multiscale algorithm for rough contact with plasticity is also included.

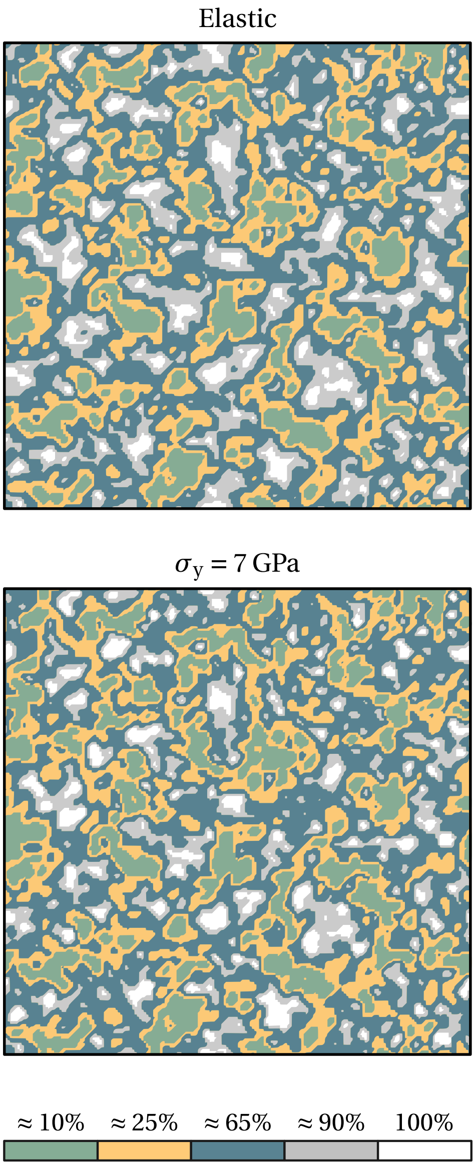


Figure 6.18: Contact area morphology at different stages of the compression process for the elastic and elastoplastic ($\sigma_y = 7$ GPa) materials.

Chapter 7

Summary and Outlook

The focus of this work is on the computational modelling of contact across scales using the FEM. After a brief historical note about tribology, an overview of recent developments with regard to computational contact mechanics is given. Both classical continuum mechanics and multi-scale methods are reviewed, giving particular emphasis on the class of mortar contact algorithms. The remaining introductory state-of-the-art focuses on the multi-scale modelling of contact, with special attention being given to computational contact homogenisation and numerical models based on the FEM.

The first part of this work is dedicated to an overview and subsequent implementation of the so-called dual mortar contact method. The choice for this algorithm is motivated by its solid mathematical formulation, accuracy and ability to deal with general, fully non-linear contact problems. The formulation is presented in detail, including the derivation of the strong and weak statements and the finite element approximation of the problem. Particular emphasis is given to the numerical evaluation of mortar integrals and the definition of the discrete Lagrange multiplier space. The employed primal-dual active set strategy is then discussed and the resulting semi-smooth Newton algorithm is introduced. Throughout the formulation, a suitable algebraic representation for the discrete entities is provided, including a detailed derivation of all the required consistent linearisation of mortar-based contact formulations. This stage of the work finishes with a discussion on the actual computational implementation of the algorithm, analysing several benchmarks and examples in order to validate the implementation.

Secondly, in preparation for the application of the dual mortar contact method to contact homogenisation, new concepts are proposed to optimise the algorithm for the particular case of Signorini contact, i.e., the contact of a deformable body against a rigid foundation. This is a problem configuration commonly found in contact homogenisation, which originates from a well-known result from contact mechanics: if there is no friction or adhesion between two rough surfaces and the surface slope is small, the elastic contact between two rough surfaces can be mapped to the contact between an equivalent single rough surface and a rigid flat (Johnson, 1987). The main motivation for this contribution is the considerable computational complexity of dual mortar methods, especially in three dimensions. The FEM itself is already regarded as an expensive alternative to model rough contact, which becomes an even more computationally demanding option when used in combination with mortar methods. Therefore, the simplifications associated with Signorini contact are exploited in order to simplify the algorithm while preserving its accuracy and flexibility—the main argument favouring the FEM. The first idea regards the variational formulation using a Petrov-Galerkin scheme. In the particular case of Signorini contact, this methodology is particularly attractive, as it eliminates the need to evaluate the dual basis during the simulation. The extension of this formulation to quadratic interpolation in three-dimensions is also carried out, proposing a piecewise linear interpolation for the variation of the Lagrange multipliers. The second idea is based on a new definition for the nodal orthonormal moving frame attached to each contact node, using the projection of the frame on the rigid side to the deformable body. When compared with the well-established method based on the averaged unit normal, this technique reduces the bandwidth of the derivatives associated with both the normal and tangential vectors, which now depend only on the degrees of freedom associated with the finite element node itself. This effect propagates throughout all the computations and, in the end, a significant reduction in the computation time of all the operations related to contact is achieved.

At the last stage of this thesis, a multi-scale algorithm for the evaluation of the real contact area fraction is proposed. It is based on the establishment of an RCE, which is then separated into multiple smaller scales by splitting the roughness spectrum. The statistics of the contact pressure field are incorporated into a new multiplicative homogenization scheme, combining two solutions based on: the average contact pressure field at the interface; the discretization of its PDF using equiprobable bins. The goal of the new formulation is to consistently integrate more information regarding the contact solution at the rough interface onto the homogenization algorithm, in order to achieve the robustness needed to deal with general scenarios and pressure distributions. Additional considerations of the method are also discussed, namely the definition of the splitting frequencies of the PSD function and the generation of the rough topography for the individual scales. The final algorithm can estimate the real contact area fraction from a given database of contact area fraction and pressure values, which can be obtained using any method. Numerical investigations on the violation of the principle of scale separation, the discretization error, the number of scales and the effect of bulk material are performed. The adopted strategy relies on the comparison of the results obtained with the multi-scale algorithm and the direct simulation of the entire rough topography. It is verified that the multi-scale algorithm can reproduce the original problem, with an accuracy that tends to increase with the roughness spectrum bandwidth. In terms of computational performance, it is demonstrated that the multi-scale algorithm is considerably faster than the original problem. For the analysed roughness spectra, the multi-scale algorithm is, at least, one order of magnitude faster. It is also shown that the number of scales should be as low as possible, to achieve better accuracy, i.e., it is limited by the computational complexity of the underlying model. The bulk material constitutive behaviour, as expected, has a significant impact on the problem, with the saturation character on the material stress response being reflected on the contact pressure, changing is contact pressure distribution, and the rate of change of the contact area fraction. Nonetheless, the accuracy of the multi-scale model keeps consistent even in this scenario. A comparison with several solutions available in the literature is also performed. Additionally, the multi-scale algorithm is applied to Persson's model, showing behaviour that is consistent with the numerical results obtained with the FEM.

7.1 Outlook and future challenges

Within the field of tribology, multi-scale computational contact mechanics based on the FEM, despite standing as an active research area over the last decade, is not so widespread as methodologies like analytical and phenomenological approaches. The main reason behind this trend regards the prohibitive computational complexity of full computational models. However, by employing alternative strategies like the one presented in this thesis, it has been shown that it is possible to circumvent this difficulty and still employ the FEM to model contact across scales. Furthermore, the field of computational sciences is growing rapidly, with new technologies and frameworks involving high-performance computing or data science making this argument weaker with time. The future is expected to be smarter, using technologies powered by embedded systems using advanced sensors and artificial intelligence. In this reality, one can expect the capacity to have high-fidelity large-scale numerical models feeding complex databases to become increasingly more important. In the following, further improvement of the proposed methods and computational algorithms is suggested, settling future research directions on multi-scale computational contact modelling.

Based on the methodology and framework presented in this thesis, the extension towards complex interface effects is easily ensured by the flexibility of the combination of finite deformation finite element technology and the dual mortar contact algorithm. The first obvious extension regards the modelling of frictional contact due to its practical and multidisciplinary importance. The developed homogenization framework is perfectly suitable to model frictional sliding, by considering an additional simulation step with prescribed external velocity. This type of approach has been already employed in the literature, e.g., in Stupkiewicz *et al.* (2014), and proved to be capable of accurately determining non-Coulomb frictional behaviour. Moreover, in Wagner, Wriggers, Veltmaat, *et al.* (2017), a multi-scale approach has been applied within the context of rubber friction, thus indicating the potential for the extension and further development of approximation multi-scale techniques. In this regard, the question of how to define individual smaller scales seems particularly interesting and challenging, especially considering non-Gaussian and anisotropic topographies.

The extension towards multi-physics systems, such as thermo-mechanics and electrochemistry, are possible directions building upon the developed framework. However, the challenge in this extension becomes the identification of the critical phenomena taking place at the interface. For instance, several modern technologies in tribology are strongly based on the metal/polymer interaction under dry friction conditions. These components typically offer low friction and, consequently, good wear performance, ensuring self-lubricating features that are critical for maintenance-free systems. The ability to model the complex contact interaction between different materials with very distinct properties enables the design of optimised solutions for specific applications. In this regard, the incorporation of effects like adhesion and thermo-mechanical contact within the current formulation is a logical step. Notwithstanding, the validation of the numerical models renders another big challenge. Experiments to access local interface states are not easy to accomplish, mainly due to the difficulty of performing *in situ* observations.

Page intentionally left blank

Appendix A

Dual mortar contact consistent linearisation

The consistent linearisation of mortar-based contact problems undergoing finite deformation requires finding directional derivatives for the first and second mortar coupling matrices, $[\mathbf{D}]$ and $[\mathbf{M}]$, respectively, and also for the discrete gap function g^{h}. To obtain these results, several elementary linearisations are needed. In the following, a comprehensive description of each of these terms is presented, including a suitable algebraic representation for the actual numerical implementation of the method. Firstly, two- and three-dimensional integration segments are treated individually and, at last, the discretised gap function and dual shape functions are derived.

A.1 Linearisation of mortar contact in two-dimensions

- Averaged normal and tangential vectors .... *Appendix A.1.1*
- Integration segment end coordinates .... *Appendix A.1.2*
- Integration segment Gauss points .... *Appendix A.1.3*
- Integration segment Jacobian determinant .... *Appendix A.1.4*

A.1.1 Averaged normal and tangential vectors

Linearisation of the averaged normal and tangential unit vectors is described in the following. Firstly, the fundamental steps for the evaluation of the directional derivatives of the averaged normal vector are presented and then, based on those results, the linearisation of the averaged tangential vector is obtained.

First of all, recall the definition of the averaged nodal unit normal $\tilde{\boldsymbol{\eta}}_j\left(\xi^{\mathrm{s}}_j\right)$ at a non-mortar node $j \in \mathcal{S}$,

$$\tilde{\boldsymbol{\eta}}_j\left(\xi^{\mathrm{s}}_j\right) \equiv \frac{\check{\boldsymbol{\eta}}_j}{\left\|\check{\boldsymbol{\eta}}_j\right\|}\,, \tag{A.1}$$

where $\check{\boldsymbol{\eta}}_j\left(\xi^{\mathrm{s}}_j\right)$ has been introduced as an abbreviation for the *non-unit* averaged normal vector at a non-mortar node j, obtained by adding all adjacent element unit normals, $\boldsymbol{\eta}^{\mathrm{a}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_j\right)$, $\mathrm{e} \in \{1, \ldots, n_{\mathrm{adj}}\}$, i.e.

$$\check{\boldsymbol{\eta}}_j\left(\xi^{\mathrm{s}}_j\right) \equiv \sum_{\mathrm{e}=1}^{n_{\mathrm{adj}}} \boldsymbol{\eta}^{\mathrm{a}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_j\right)\,. \tag{A.2}$$

Herein n_{adj} represents the total number of adjacent elements, which for two-dimensional analyses typically is $n_{\text{adj}} = 2$ (with the exception of interior nodes in the case of three-noded line segments used in quadratic finite element interpolation). Denoting the length of $\check{\boldsymbol{\eta}}_j$ as $l_{\check{\eta}_j} \equiv \|\check{\boldsymbol{\eta}}_j\|$, the directional derivative of Equation (A.1) leads to

$$\mathscr{D}\tilde{\boldsymbol{\eta}}_j = \mathscr{D}\left(\frac{\check{\boldsymbol{\eta}}_j}{l_{\check{\eta}_j}}\right) = \frac{\mathscr{D}\check{\boldsymbol{\eta}}_j}{l_{\check{\eta}_j}} - \frac{\left(\check{\boldsymbol{\eta}}_j \cdot \mathscr{D}\check{\boldsymbol{\eta}}_j\right)\check{\boldsymbol{\eta}}_j}{l^3_{\check{\eta}_j}} = \left(\frac{1}{l_{\check{\eta}_j}}\boldsymbol{I} - \frac{1}{l^3_{\check{\eta}_j}}\check{\boldsymbol{\eta}}_j \otimes \check{\boldsymbol{\eta}}_j\right)\mathscr{D}\check{\boldsymbol{\eta}}_j\,. \tag{A.3}$$

This means that linearisation of the non-unit normal vector $\check{\boldsymbol{\eta}}_j$ given in Equation (A.2) is required in order to further trace back the linearisation of the averaged unit normal vector to elementary nodal displacement increments. Therefore, introducing the *non-unit* normal $\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}(\xi^{\text{s}}_j)$ to the adjacent element e and the corresponding length $l_{\check{\eta}^{\text{a}}_{\text{e}}} \equiv \|\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}\|$, the directional derivative of Equation (A.2) gives[*]

$$\begin{gathered}\mathscr{D}\check{\boldsymbol{\eta}}_j = \mathscr{D}\left(\sum_{\text{e}=1}^{n_{\text{adj}}}\boldsymbol{\eta}^{\text{a}}_{\text{e}}\right) = \mathscr{D}\left(\sum_{\text{e}=1}^{n_{\text{adj}}}\frac{\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}}{l_{\check{\eta}^{\text{a}}_{\text{e}}}}\right)\\ = \sum_{\text{e}=1}^{n_{\text{adj}}}\left(\frac{\mathscr{D}\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}}{l_{\check{\eta}^{\text{a}}_{\text{e}}}} - \frac{\left(\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}\cdot\mathscr{D}\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}\right)\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}}{l^3_{\check{\eta}^{\text{a}}_{\text{e}}}}\right) = \sum_{\text{e}=1}^{n_{\text{adj}}}\left[\left(\frac{1}{l_{\check{\eta}^{\text{a}}_{\text{e}}}}\boldsymbol{I} - \frac{1}{l^3_{\check{\eta}^{\text{a}}_{\text{e}}}}\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}\otimes\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}\right)\mathscr{D}\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}}\right]\,.\end{gathered} \tag{A.4}$$

Finally, taking into account that

$$\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}} = \mathbf{e}_3 \times \check{\boldsymbol{\tau}}^{\text{a}}_{\text{e}}\,, \tag{A.5}$$

with $\check{\boldsymbol{\tau}}^{\text{a}}_{\text{e}}(\xi^{\text{s}}_j)$ representing the *non-unit* tangent vector to the adjacent non-mortar element e, i.e.

$$\check{\boldsymbol{\tau}}^{\text{a}}_{\text{e}}(\xi^{\text{s}}_j) \equiv \left\{\boldsymbol{x}^{\text{s}}_{,\xi}\right\}^{\text{h}} = \sum_{k=1}^{n^{\text{s}}_{\text{e}}} N^{\text{s}}_{k,\xi}(\xi^{\text{s}})\,\mathbf{x}^{\text{s}}_k\,, \tag{A.6}$$

the directional derivative given in Equation (A.4) can then be defined in terms of nodal displacement increments:

$$\mathscr{D}\check{\boldsymbol{\eta}}^{\text{a}}_{\text{e}} = \mathscr{D}\left(\mathbf{e}_3 \times \check{\boldsymbol{\tau}}^{\text{a}}_{\text{e}}\right) = \mathbf{e}_3 \times \left(\sum_{k=1}^{n^{\text{s}}_{\text{e}}} N^{\text{s}}_{k,\xi}(\xi^{\text{s}}_j)\Delta\mathbf{x}^{\text{s}}_k\right). \tag{A.7}$$

Herein n^{s}_{e} denotes the total number of nodes associated with non-mortar element e and $N^{\text{s}}_{k,\xi}$ the shape function derivative with respect to its local coordinate ξ^{s},

$$N^{\text{s}}_{k,\xi}(\xi^{\text{s}}) \equiv \frac{\mathrm{d}N^{\text{s}}_k(\xi^{\text{s}})}{\mathrm{d}\xi^{\text{s}}}\,. \tag{A.8}$$

With the directional derivative of the averaged nodal unit normal vector at hand, by taking into account the definition of the corresponding nodal averaged tangential vector,

$$\tilde{\boldsymbol{\tau}}_j = \tilde{\boldsymbol{\eta}}_j \times \mathbf{e}_3\,, \tag{A.9}$$

its linearisation becomes trivial:

$$\mathscr{D}\tilde{\boldsymbol{\tau}}_j = \mathscr{D}\tilde{\boldsymbol{\eta}}_j \times \mathbf{e}_3\,. \tag{A.10}$$

---

[*] Note that, for ease of notation, the subscript j in entities related to adjacent elements has been omitted, with the understanding that they are always associated with a given non-mortar node j.

A.1.1.1 Algebraic representation

To establish a suitable algebraic representation for the previous results, a set of matrices is firstly introduced. Starting by element-wise entities, the vector $\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\} \in \mathbb{R}^{2\cdot n^{\mathrm{s}}_{\mathrm{e}}}$ containing the nodal coordinates of all $n^{\mathrm{s}}_{\mathrm{e}}$ nodes of a given non-mortar element e is constructed as

$$\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\} \equiv \left\{ \begin{array}{c} \mathbf{x}^{\mathrm{s}}_{1} \\ \vdots \\ \mathbf{x}^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}} \end{array} \right\} . \tag{A.11}$$

Furthermore, the *element* matrix $\left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}(\xi^{\mathrm{s}})\right] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}_{\mathrm{e}}}$ denotes the discrete FE gradient operator, comprising the shape function derivatives evaluated at a given local coordinate ξ^{s}. For two-dimensional problems as

$$\left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}(\xi^{\mathrm{s}})\right] \equiv \left[\begin{array}{ccccc} N^{\mathrm{s}}_{1,\xi}(\xi^{\mathrm{s}}) & 0 & \cdots & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}},\xi}(\xi^{\mathrm{s}}) & 0 \\ 0 & N^{\mathrm{s}}_{1,\xi}(\xi^{\mathrm{s}}) & & 0 & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}},\xi}(\xi^{\mathrm{s}}) \end{array} \right] . \tag{A.12}$$

Then, the *global* discrete gradient operator $\left[\mathbf{B}^{\mathrm{s}}(\xi^{\mathrm{s}})\right] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}}$ also contains the shape function derivatives, but instead inserted into their corresponding columns (with the remaining blocks equal to zero), i.e.,

$$\left[\mathbf{B}^{\mathrm{s}}(\xi^{\mathrm{s}})\right] \equiv \left[\begin{array}{cccccccc} 0 & 0 & \cdots & N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}) & 0 & \cdots & 0 & 0 \\ 0 & 0 & & 0 & N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}) & & 0 & 0 \end{array} \right] . \tag{A.13}$$

Finally, the rotation matrix $[\mathbf{R}] \in \mathbb{R}^{2\times 2}$ is used to reproduce the cross product between a given vector and the basis vector $\mathbf{e}_3$, i.e.,

$$[\mathbf{R}] \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} , \tag{A.14}$$

such that for a generic vector $\boldsymbol{a} = \mathbf{e}_3 \times \boldsymbol{b}$ one has

$$\{\boldsymbol{a}\} = [\mathbf{R}]\{\boldsymbol{b}\} . \tag{A.15}$$

With the entities introduced above, the non-unit normal vector $\{\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{j})\} \in \mathbb{R}^2$ to a given adjacent non-mortar element e (associated with a non-mortar node of local coordinate ξ^{s}_{j} in that element) can be expressed as

$$\{\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{j})\} = [\mathbf{R}]\left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{j})\right]\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\} . \tag{A.16}$$

Moreover, its directional derivative vector $\{\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{j})\} \in \mathbb{R}^2$ can be computed via

$$\{\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{j})\} = \left[\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}\right]\{\Delta\mathbf{d}_{\mathcal{S}}\} , \tag{A.17}$$

where the global matrix $\left[\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}(\xi^{\mathrm{s}}_{j})\right] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}}$ is defined as

$$\left[\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}(\xi^{\mathrm{s}}_{j})\right] \equiv [\mathbf{R}]\left[\mathbf{B}^{\mathrm{s}}(\xi^{\mathrm{s}}_{j})\right] , \tag{A.18}$$

which yields

$$\left[\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}\left(\xi^{\mathrm{s}}_{j}\right)\right] = \left[\begin{array}{cccccccc} 0 & 0 & \cdots & 0 & -N^{\mathrm{s}}_{k,\xi}\left(\xi^{\mathrm{s}}_{j}\right) & \cdots & 0 & 0 \\ 0 & 0 & \cdots & N^{\mathrm{s}}_{k,\xi}\left(\xi^{\mathrm{s}}_{j}\right) & 0 & \cdots & 0 & 0 \end{array}\right] . \tag{A.19}$$

Once evaluated for each adjacent element the quantities described above, the non-unit averaged normal vector $\{\check{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} \in \mathbb{R}^2$ is determined by

$$\{\check{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} = \sum_{e=1}^{n_{\mathrm{adj}}} \left(\frac{1}{l_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}} \{\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\} \right) , \tag{A.20}$$

and its directional derivative vector $\{\mathscr{D}\check{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} \in \mathbb{R}^2$ by

$$\{\mathscr{D}\check{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} = [\mathbf{K}_{\check{\eta}_j}]\{\Delta\mathbf{d}_{\mathcal{S}}\} . \tag{A.21}$$

Here the global matrix $[\mathbf{K}_{\check{\eta}_j}(\xi^{\mathrm{s}}_j)] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}}$ has been introduced, which is obtained by the sum

$$[\mathbf{K}_{\check{\eta}_j}(\xi^{\mathrm{s}}_j)] \equiv \sum_{e=1}^{n_{\mathrm{adj}}} \left[\left(\frac{1}{l_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}} [\mathbf{I}] - \frac{1}{l^3_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}} \{\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\}\{\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\}^{\mathrm{T}} \right) [\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}] \right] , \tag{A.22}$$

with $[\mathbf{I}] \in \mathbb{R}^{2\times 2}$ denoting the identity matrix.

Finally, the averaged nodal unit normal vector $\{\tilde{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} \in \mathbb{R}^2$ can be expressed as

$$\{\tilde{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} = \frac{1}{l_{\check{\eta}_j}} \{\check{\boldsymbol{\eta}}_j\} . \tag{A.23}$$

Its linearisation is represented by the vector $\{\mathscr{D}\tilde{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} \in \mathbb{R}^2$ which, as in Equation (A.21), comes as

$$\{\mathscr{D}\tilde{\boldsymbol{\eta}}_j(\xi^{\mathrm{s}}_j)\} = [\mathbf{K}_{\tilde{\eta}_j}]\{\Delta\mathbf{d}_{\mathcal{S}}\} . \tag{A.24}$$

The tangent matrix $[\mathbf{K}_{\tilde{\eta}_j}(\xi^{\mathrm{s}}_j)] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}}$ is constructed as

$$[\mathbf{K}_{\tilde{\eta}_j}(\xi^{\mathrm{s}}_j)] \equiv \left(\frac{1}{l_{\check{\eta}_j}} [\mathbf{I}] - \frac{1}{l^3_{\check{\eta}_j}} \{\check{\boldsymbol{\eta}}_j\}\{\check{\boldsymbol{\eta}}_j\}^{\mathrm{T}} \right) [\mathbf{K}_{\check{\eta}_j}] . \tag{A.25}$$

With both averaged nodal unit normal vector and corresponding directional derivative at hand, the tangent vector $\{\tilde{\boldsymbol{\tau}}_j(\xi^{\mathrm{s}}_j)\} \in \mathbb{R}^2$ can be computed as

$$\{\tilde{\boldsymbol{\tau}}_j(\xi^{\mathrm{s}}_j)\} = [\mathbf{R}]^{\mathrm{T}}\{\tilde{\boldsymbol{\eta}}_j\} , \tag{A.26}$$

and its directional derivative $\{\mathscr{D}\tilde{\boldsymbol{\tau}}_j(\xi^{\mathrm{s}}_j)\} \in \mathbb{R}^2$ as

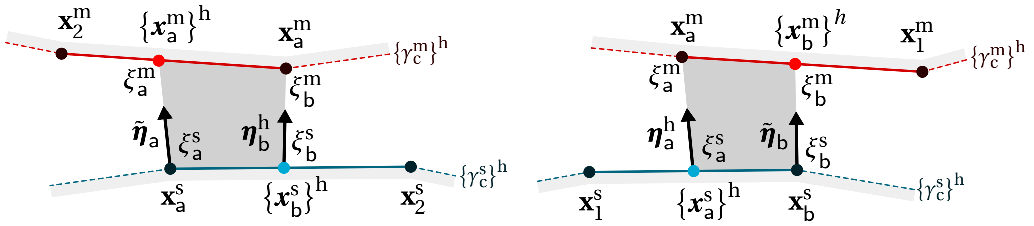
$$\{\mathscr{D}\tilde{\boldsymbol{\tau}}_j(\xi^{\mathrm{s}}_j)\} = [\mathbf{R}]^{\mathrm{T}}[\mathbf{K}_{\tilde{\eta}_j}]\{\Delta\mathbf{d}_{\mathcal{S}}\} . \tag{A.27}$$

A.1.2 Integration segment end coordinates

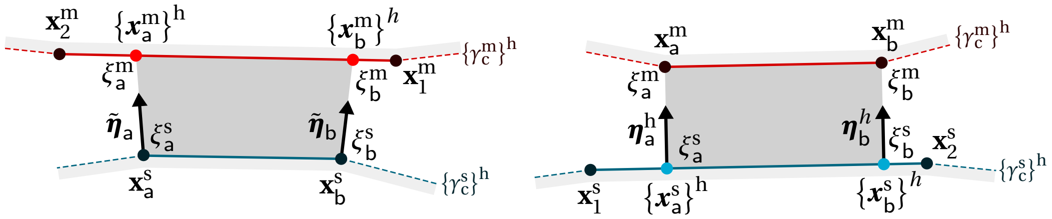
Since integration segments generation possibly evolves deformation-dependent projections of nodes from one boundary to another, its linearisation requires analysing different integration element configurations. A schematic depiction of all possible types of integration elements in two dimensions is given in Figure A.1. From it, one can easily conclude that there are two possibilities for the origin of the integration segment end coordinates $\xi^{\mathrm{s}}_{\mathrm{a}}$ and $\xi^{\mathrm{s}}_{\mathrm{b}}$ on the non-mortar side and $\xi^{\mathrm{m}}_{\mathrm{a}}$ and $\xi^{\mathrm{m}}_{\mathrm{b}}$ on the mortar side, respectively. They can either coincide with:

- **A finite element node of the corresponding surface:** the directional derivative of the segment end coordinate simply vanishes;
- **Projection of a finite element node from the respective other surface:** the directional derivative has to be computed by linearising the associated projection condition.

In what follows, two distinct cases—*Case 1* and *Case 2*, according to Figure A.1—are considered, which allow establishing directional derivatives for any possible mortar element type.



Case 1 Case 2



Case 3 Case 4

Figure A.1: Different combinations of integration element types.

A.1.2.1 Case 1: Projection of a non-mortar node onto a mortar element

First, let $\xi^{\mathrm{s}}_{\mathrm{a}}$ be coincident with a non-mortar node, i.e., $\mathbf{x}^{\mathrm{s}}_{\mathrm{a}} \equiv \mathbf{x}^{\mathrm{s}}_1$ and $\tilde{\boldsymbol{\eta}}_{\mathrm{a}} \equiv \tilde{\boldsymbol{\eta}}_1$. Additionally, it also follows that

$$\mathscr{D}\xi^{\mathrm{s}}_{\mathrm{a}} = 0\,, \tag{A.28}$$

and, at the same time, the projection $\xi_{\rm a}^{\rm m}$ of the non-mortar node $\mathbf{x}_{\rm a}^{\rm s}$ onto the mortar element e with nodes $\mathbf{x}_l^{\rm m}\,(l = 1,\ldots,n_{\rm e}^{\rm m})$ is obtained by solving

$$\left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\mathbf{x}_l^{\rm m}\right) - \mathbf{x}_{\rm a}^{\rm s}\right] \times \tilde{\boldsymbol{\eta}}_{\rm a} = \mathbf{0}\,, \tag{A.29}$$

with $\tilde{\boldsymbol{\eta}}_{\rm a}$ representing the averaged unit normal vector at the non-mortar node $\mathbf{x}_{\rm a}^{\rm s}$.

Starting by taking into account that, for two-dimensional problems, the cross product yields

$$\left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\mathrm{x}_{lx}^{\rm m}\right) - \mathrm{x}_{{\rm a}x}^{\rm s}\right] \tilde{\eta}_{{\rm a}y} - \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\mathrm{x}_{ly}^{\rm m}\right) - \mathrm{x}_{{\rm a}y}^{\rm s}\right] \tilde{\eta}_{{\rm a}x} = 0\,, \tag{A.30}$$

its linearisation reads

$$\begin{aligned}
&\left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_{l,\xi}^{\rm m}(\xi_{\rm a}^{\rm m})\,\mathscr{D}(\xi_{\rm a}^{\rm m})\,\mathrm{x}_{lx}^{\rm m}\right) + \sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\Delta\mathrm{x}_{lx}^{\rm m}\right) - \Delta\mathrm{x}_{{\rm a}x}^{\rm s}\right] \tilde{\eta}_{{\rm a}y} \\
&+ \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\,\mathrm{x}_{lx}^{\rm m}\right) - \mathrm{x}_{{\rm a}x}^{\rm s}\right] \mathscr{D}\tilde{\eta}_{{\rm a}y} \\
&- \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_{l,\xi}^{\rm m}(\xi_{\rm a}^{\rm m})\,\mathscr{D}(\xi_{\rm a}^{\rm m})\,\mathrm{x}_{ly}^{\rm m}\right) + \sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\Delta\mathrm{x}_{ly}^{\rm m}\right) - \Delta\mathrm{x}_{{\rm a}y}^{\rm s}\right] \tilde{\eta}_{{\rm a}x} \\
&- \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\,\mathrm{x}_{ly}^{\rm m}\right) - \mathrm{x}_{{\rm a}y}^{\rm s}\right] \mathscr{D}\tilde{\eta}_{{\rm a}x} = 0\,.
\end{aligned} \tag{A.31}$$

Thus, by evidencing the term $\mathscr{D}\xi_{\rm a}^{\rm m}$, the linearisation of the integration segment end coordinate $\xi_{\rm a}^{\rm m}$ can be expressed as

$$\mathscr{D}\xi_{\rm a}^{\rm m} = -\frac{\mathrm{C}_{\xi_{\rm a}^{\rm m}}^{\rm num}}{\mathrm{C}_{\xi_{\rm a}^{\rm m}}^{\rm den}}\,, \tag{A.32}$$

with

$$\begin{aligned}
\mathrm{C}_{\xi_{\rm a}^{\rm m}}^{\rm num} \equiv\; &\left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\Delta\mathrm{x}_{lx}^{\rm m}\right) - \Delta\mathrm{x}_{{\rm a}x}^{\rm s}\right] \tilde{\eta}_{{\rm a}y} - \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\Delta\mathrm{x}_{ly}^{\rm m}\right) - \Delta\mathrm{x}_{{\rm a}y}^{\rm s}\right] \tilde{\eta}_{{\rm a}x} \\
&+ \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\mathrm{x}_{lx}^{\rm m}\right) - \mathrm{x}_{{\rm a}x}^{\rm s}\right] \mathscr{D}\tilde{\eta}_{{\rm a}y} - \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_l^{\rm m}(\xi_{\rm a}^{\rm m})\mathrm{x}_{ly}^{\rm m}\right) - \mathrm{x}_{{\rm a}y}^{\rm s}\right] \mathscr{D}\tilde{\eta}_{{\rm a}x}
\end{aligned} \tag{A.33}$$

and

$$\mathrm{C}_{\xi_{\rm a}^{\rm m}}^{\rm den} \equiv \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_{l,\xi}^{\rm m}(\xi_{\rm a}^{\rm m})\mathrm{x}_{lx}^{\rm m}\right)\right] \tilde{\eta}_{{\rm a}y} - \left[\sum_{l=1}^{n_{\rm e}^{\rm m}} \left(N_{l,\xi}^{\rm m}(\xi_{\rm a}^{\rm m})\mathrm{x}_{ly}^{\rm m}\right)\right] \tilde{\eta}_{{\rm a}x}\,. \tag{A.34}$$

A.1.2.2 Algebraic representation

The algebraic representation of the previous result requires the introduction of additional entities. Thus, starting again by element-wise quantities, the matrix $\left[\mathbf{N}_{\rm e}^{\rm m}(\xi^{\rm m})\right] \in \mathbb{R}^{2\times 2\cdot n_{\rm e}^{\rm m}}$ contains the shape functions of the mortar element e, evaluated at a given point of local coordinate $\xi^{\rm m}$. In two dimensions comes as

$$\left[\mathbf{N}_{\rm e}^{\rm m}(\xi^{\rm m})\right] \equiv \begin{bmatrix} N_1^{\rm m}(\xi^{\rm m}) & 0 & \cdots & N_{n_{\rm e}^{\rm m}}^{\rm m}(\xi^{\rm m}) & 0 \\ 0 & N_1^{\rm m}(\xi^{\rm m}) & & 0 & N_{n_{\rm e}^{\rm m}}^{\rm m}(\xi^{\rm m}) \end{bmatrix}\,. \tag{A.35}$$

Likewise, the element discrete gradient operator $\left[\mathbf{B}_{\mathrm{e}}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right)\right] \in \mathbb{R}^{2\times 2\cdot n_{\mathrm{e}}^{\mathrm{m}}}$ contains the shape functions derivatives with respect to the local mortar element coordinate evaluated at a point ξ^{m}, i.e.,

$$\left[\mathbf{B}_{\mathrm{e}}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right)\right] \equiv \begin{bmatrix} N_{1,\xi}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) & 0 & \cdots & N_{n_{\mathrm{e}}^{\mathrm{m}},\xi}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) & 0 \\ 0 & N_{1,\xi}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) & \cdots & 0 & N_{n_{\mathrm{e}}^{\mathrm{m}},\xi}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) \end{bmatrix} . \tag{A.36}$$

The vector $\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\right\} \in \mathbb{R}^{2\cdot n_{\mathrm{e}}^{\mathrm{m}}}$ is constructed by gathering the mortar element nodal coordinates $\mathbf{x}_{l}^{\mathrm{m}}\left(l = 1,\ldots,n_{\mathrm{e}}^{\mathrm{m}}\right)$, viz.

$$\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\right\} \equiv \begin{Bmatrix} \mathbf{x}_{1}^{\mathrm{m}} \\ \vdots \\ \mathbf{x}_{n_{\mathrm{e}}^{\mathrm{m}}}^{\mathrm{m}} \end{Bmatrix} . \tag{A.37}$$

Another required entity is the vector $\left\{\mathbf{m}\left(\xi^{\mathrm{m}},\xi^{\mathrm{s}}\right)\right\} \in \mathbb{R}^{2}$ connecting a given point on the non-mortar boundary to another point on the opposing mortar counterpart, i.e.,

$$\left\{\mathbf{m}\left(\xi^{\mathrm{m}},\xi^{\mathrm{s}}\right)\right\} := \left\{\boldsymbol{x}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right)\right\}^{\mathrm{h}} - \left\{\boldsymbol{x}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right)\right\}^{\mathrm{h}} . \tag{A.38}$$

For the particular case under analysis, it connects the integration segment end points,

$$\left\{\mathbf{m}_{\mathrm{a}}\right\} = \left[\mathbf{N}_{\mathrm{e}}^{\mathrm{m}}\left(\xi_{\mathrm{a}}^{\mathrm{m}}\right)\right]\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\right\} - \left\{\mathbf{x}_{\mathrm{a}}^{\mathrm{s}}\right\} . \tag{A.39}$$

Focusing on globally assembled entities, the global shape function matrix $\left[\mathbf{N}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right)\right] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{m}}}$ is defined as

$$\left[\mathbf{N}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right)\right] \equiv \begin{bmatrix} 0 & 0 & \cdots & N_{k}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & N_{k}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) & \cdots & 0 & 0 \end{bmatrix} , \tag{A.40}$$

and the global discrete gradient operator $\left[\mathbf{B}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right)\right] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{m}}}$ as

$$\left[\mathbf{B}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right)\right] \equiv \begin{bmatrix} 0 & 0 & \cdots & N_{k,\xi}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & N_{k,\xi}^{\mathrm{m}}\left(\xi^{\mathrm{m}}\right) & \cdots & 0 & 0 \end{bmatrix} . \tag{A.41}$$

The global auxiliary matrix $\left[\breve{\mathbf{I}}_{k}^{\mathrm{s}}\right] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}}$ is introduced in order to multiply some quantity by the nodal displacement increments of a given non-mortar node k, coming in two dimensions as

$$\left[\breve{\mathbf{I}}_{k}^{\mathrm{s}}\right] \equiv \begin{bmatrix} 0 & 0 & \cdots & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \underbrace{0 \quad 1}_{\text{node } k} & & \cdots & 0 & 0 \end{bmatrix} . \tag{A.42}$$

The linearisation of the integration segment end coordinate $\xi_{\mathrm{a}}^{\mathrm{m}}$ can finally be written as

$$\mathscr{D}\xi_{\mathrm{a}}^{\mathrm{m}} = \left\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}}\right\} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.43}$$

where the *row* vector $\left\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}}\left(\xi_{\mathrm{a}}^{\mathrm{m}}\right)\right\} \in \mathbb{R}^{2\cdot\left(n^{\mathrm{m}}+n^{\mathrm{s}}\right)}$ is decomposed as follows:

$$\left\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}}\left(\xi_{\mathrm{a}}^{\mathrm{m}}\right)\right\} \equiv \left\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}\mathcal{M}} \, , \, \mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}\mathcal{S}}\right\} . \tag{A.44}$$

Here, the vector $\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}\mathcal{M}}(\xi_{\mathrm{a}}^{\mathrm{m}})\} \in \mathbb{R}^{2\cdot n^{\mathrm{m}}}$ contains the terms associated with mortar nodes and is calculated as

$$\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}\mathcal{M}}(\xi_{\mathrm{a}}^{\mathrm{m}})\} \equiv \mathrm{C}_{\xi_{\mathrm{a}}^{\mathrm{m}}} \{\tilde{\boldsymbol{\tau}}_{\mathrm{a}}\}^{\mathrm{T}} [\mathbf{N}^{\mathrm{m}}(\xi_{\mathrm{a}}^{\mathrm{m}})] \, . \tag{A.45}$$

Similarly, the vector $\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}\mathcal{S}}(\xi_{\mathrm{a}}^{\mathrm{m}})\} \in \mathbb{R}^{2\cdot n^{\mathrm{s}}}$ is related with non-mortar nodes and is given by

$$\{\mathbf{K}_{\xi_{\mathrm{a}}^{\mathrm{m}}\mathcal{S}}(\xi_{\mathrm{a}}^{\mathrm{m}})\} \equiv \mathrm{C}_{\xi_{\mathrm{a}}^{\mathrm{m}}} \left(\{\mathbf{m}_{\mathrm{a}}\}^{\mathrm{T}} [\mathbf{R}]^{\mathrm{T}} [\mathbf{K}_{\tilde{\eta}_{\mathrm{a}}}] - \{\tilde{\boldsymbol{\tau}}_{\mathrm{a}}\}^{\mathrm{T}} [\breve{\mathbf{I}}_{\mathrm{a}}^{\mathrm{s}}] \right) . \tag{A.46}$$

The constant $\mathrm{C}_{\xi_{\mathrm{a}}^{\mathrm{m}}}(\xi_{\mathrm{a}}^{\mathrm{m}})$ is computed as

$$\mathrm{C}_{\xi_{\mathrm{a}}^{\mathrm{m}}}(\xi_{\mathrm{a}}^{\mathrm{m}}) \equiv -\frac{1}{\{\tilde{\boldsymbol{\tau}}_{\mathrm{a}}\}^{\mathrm{T}} [\mathbf{B}_{\mathrm{e}}^{\mathrm{m}}(\xi_{\mathrm{a}}^{\mathrm{m}})] \{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\}} \, . \tag{A.47}$$

In the equations above, the averaged unit tangent vector $\{\tilde{\boldsymbol{\tau}}_{\mathrm{a}}\}$ can be obtained from Equation (A.26), written in this case as

$$\{\tilde{\boldsymbol{\tau}}_{\mathrm{a}}\}^{\mathrm{T}} = \{\tilde{\boldsymbol{\eta}}_{\mathrm{a}}\}^{\mathrm{T}} [\mathbf{R}] \, . \tag{A.48}$$

The matrix $[\mathbf{K}_{\tilde{\eta}_{\mathrm{a}}}]$, defined in Equation (A.25), is associated with the linearisation of the averaged unit normal vector $\{\tilde{\boldsymbol{\eta}}_{\mathrm{a}}\}$, here evaluated at the non-mortar node $\mathbf{x}_{\mathrm{a}}^{\mathrm{s}}$.

A.1.2.3 Case 2: Projection of a mortar node onto a non-mortar element

The second possible configuration for a integration segment end coordinate corresponds to the projection of a mortar node onto a non-mortar segment. Thus, as in *Case 2* of Figure A.1, let $\xi_{\mathrm{a}}^{\mathrm{m}}$ be coincident with mortar node, i.e., $\mathbf{x}_{\mathrm{a}}^{\mathrm{m}} = \mathbf{x}_{1}^{\mathrm{m}}$. Consequently, one obtains

$$\mathscr{D}\xi_{\mathrm{a}}^{\mathrm{m}} = 0 \, , \tag{A.49}$$

and the sought-after projection coordinate $\xi_{\mathrm{a}}^{\mathrm{s}}$ on the related non-mortar element e, with nodes $\mathbf{x}_{k}^{\mathrm{s}}$ $(k = 1, \dots, n_{\mathrm{e}}^{\mathrm{s}})$, can be determined by solving

$$\left[\mathbf{x}_{\mathrm{a}}^{\mathrm{m}} - \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k}^{\mathrm{s}}(\xi_{\mathrm{a}}^{\mathrm{s}}) \mathbf{x}_{k}^{\mathrm{s}} \right) \right] \times \left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k}^{\mathrm{s}}(\xi_{\mathrm{a}}^{\mathrm{s}}) \tilde{\boldsymbol{\eta}}_{k} \right) \right] = \mathbf{0} \, . \tag{A.50}$$

Starting once again by expanding the cross product operator,

$$\begin{aligned} &\left[\mathrm{x}_{\mathrm{a}x}^{\mathrm{m}} - \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k}^{\mathrm{s}}(\xi_{\mathrm{a}}^{\mathrm{s}}) \mathrm{x}_{kx}^{\mathrm{s}} \right) \right] \left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k}^{\mathrm{s}}(\xi_{\mathrm{a}}^{\mathrm{s}}) \tilde{\eta}_{ky} \right) \right] \\ - &\left[\mathrm{x}_{\mathrm{a}y}^{\mathrm{m}} - \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k}^{\mathrm{s}}(\xi_{\mathrm{a}}^{\mathrm{s}}) \mathrm{x}_{ky}^{\mathrm{s}} \right) \right] \left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k}^{\mathrm{s}}(\xi_{\mathrm{a}}^{\mathrm{s}}) \tilde{\eta}_{kx} \right) \right] = 0 \, , \end{aligned} \tag{A.51}$$

its linearisation yields

$$\begin{aligned}
&\left[\Delta \mathrm{x}^{\mathrm{m}}_{\mathsf{a}x} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathrm{x}^{\mathrm{s}}_{kx}\right) - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \Delta \mathrm{x}^{\mathrm{s}}_{kx}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \tilde{\eta}_{ky}\right)\right] \\
+ &\left[\mathrm{x}^{\mathrm{m}}_{\mathsf{a}x} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathrm{x}^{\mathrm{s}}_{kx}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}\left(\xi^{\mathrm{s}}_{\mathsf{a}}\right) \tilde{\eta}_{ky}\right) + \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}\tilde{\eta}_{ky}\right)\right] \\
- &\left[\Delta \mathrm{x}^{\mathrm{m}}_{\mathsf{a}y} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathrm{x}^{\mathrm{s}}_{ky}\right) - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \Delta \mathrm{x}^{\mathrm{s}}_{ky}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \tilde{\eta}_{kx}\right)\right] \\
- &\left[\mathrm{x}^{\mathrm{m}}_{\mathsf{a}y} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathrm{x}^{\mathrm{s}}_{ky}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}\left(\xi^{\mathrm{s}}_{\mathsf{a}}\right) \tilde{\eta}_{kx}\right) + \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}\tilde{\eta}_{kx}\right)\right] = 0\,.
\end{aligned} \tag{A.52}$$

Finally, by rearranging the previous equation, the linearisation of the integration segment end coordinate $\xi^{\mathrm{s}}_{\mathsf{a}}$ results in

$$\mathscr{D}\xi^{\mathrm{s}}_{\mathsf{a}} = -\frac{\mathrm{C}^{\mathrm{num}}_{\xi^{\mathrm{s}}_{\mathsf{a}}}}{\mathrm{C}^{\mathrm{den}}_{\xi^{\mathrm{s}}_{\mathsf{a}}}}\,, \tag{A.53}$$

where

$$\begin{aligned}
\mathrm{C}^{\mathrm{num}}_{\xi^{\mathrm{s}}_{\mathsf{a}}} \equiv \quad &\left[\Delta \mathrm{x}^{\mathrm{m}}_{\mathsf{a}x} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \Delta \mathrm{x}^{\mathrm{s}}_{kx}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \tilde{\eta}_{ky}\right)\right] \\
- &\left[\Delta \mathrm{x}^{\mathrm{m}}_{\mathsf{a}y} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \Delta \mathrm{x}^{\mathrm{s}}_{ky}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \tilde{\eta}_{kx}\right)\right] \\
+ &\left[\mathrm{x}^{\mathrm{m}}_{\mathsf{a}x} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \mathrm{x}^{\mathrm{s}}_{kx}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}\tilde{\eta}_{ky}\right)\right] \\
- &\left[\mathrm{x}^{\mathrm{m}}_{\mathsf{a}y} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \mathrm{x}^{\mathrm{s}}_{ky}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}})\, \mathscr{D}\tilde{\eta}_{kx}\right)\right]
\end{aligned} \tag{A.54}$$

and

$$\begin{aligned}
\mathrm{C}^{\mathrm{den}}_{\xi^{\mathrm{s}}_{\mathsf{a}}} \equiv - &\left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}}) \mathrm{x}^{\mathrm{s}}_{kx}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \tilde{\eta}_{ky}\right)\right] \\
+ &\left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}}) \mathrm{x}^{\mathrm{s}}_{ky}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \tilde{\eta}_{kx}\right)\right] \\
+ &\left[\mathrm{x}^{\mathrm{m}}_{\mathsf{a}x} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \mathrm{x}^{\mathrm{s}}_{kx}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}}) \tilde{\eta}_{ky}\right)\right] \\
- &\left[\mathrm{x}^{\mathrm{m}}_{\mathsf{a}y} - \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathsf{a}}) \mathrm{x}^{\mathrm{s}}_{ky}\right)\right] \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}(\xi^{\mathrm{s}}_{\mathsf{a}}) \tilde{\eta}_{kx}\right)\right]\,.
\end{aligned} \tag{A.55}$$

A.1.2.4 Algebraic representation

First, the auxiliary entities needed for the algebraic representation of Equation (A.53) are presented. The element matrix $\left[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\xi^{\mathrm{s}})\right] \in \mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}_{\mathrm{e}}}$ contains the non-mortar element shape functions, coming in the two-dimensional case as

$$\left[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}\right)\right] \equiv \begin{bmatrix} N^{\mathrm{s}}_{1}\left(\xi^{\mathrm{s}}\right) & 0 & \cdots & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}\left(\xi^{\mathrm{s}}\right) & 0 \\ 0 & N^{\mathrm{s}}_{1}\left(\xi^{\mathrm{s}}\right) & & 0 & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}\left(\xi^{\mathrm{s}}\right) \end{bmatrix}. \tag{A.56}$$

Moreover, the element vector $\{\tilde{\boldsymbol{\eta}}_\mathrm{e}\} \in \mathbb{R}^{2\cdot n^\mathrm{s}_\mathrm{e}}$ gathers all the averaged unit normal vectors $\tilde{\boldsymbol{\eta}}_k \left(k = 1,\ldots, n^\mathrm{s}_\mathrm{e}\right)$ of the non-mortar element, i.e.,

$$\{\tilde{\boldsymbol{\eta}}_\mathrm{e}\} \equiv \left\{ \begin{matrix} \tilde{\boldsymbol{\eta}}_1 \\ \vdots \\ \tilde{\boldsymbol{\eta}}_{n^\mathrm{s}_\mathrm{e}} \end{matrix} \right\} . \tag{A.57}$$

This vector can, for instance, be combined with the matrix $\left[\mathbf{N}^\mathrm{s}_\mathrm{e}(\xi^\mathrm{s})\right]$ to compute the interpolated *non-unit* normal vector $\{\check{\boldsymbol{\eta}}^\mathrm{h}_\mathrm{a}(\xi^\mathrm{s}_\mathrm{a})\} \in \mathbb{R}^2$ at the integration segment end coordinate $\xi^\mathrm{s}_\mathrm{a}$, viz.

$$\{\check{\boldsymbol{\eta}}^\mathrm{h}_\mathrm{a}(\xi^\mathrm{s}_\mathrm{a})\} \equiv \sum_{k=1}^{n^\mathrm{s}_\mathrm{e}} N^\mathrm{s}_k(\xi^\mathrm{s}_\mathrm{a})\tilde{\boldsymbol{\eta}}_k = \left[\mathbf{N}^\mathrm{s}_\mathrm{e}(\xi^\mathrm{s}_\mathrm{a})\right]\{\tilde{\boldsymbol{\eta}}_\mathrm{e}\} . \tag{A.58}$$

Recalling Equation (A.26), the corresponding *non-unit* tangent vector $\{\check{\boldsymbol{\tau}}^\mathrm{h}_\mathrm{a}(\xi^\mathrm{s}_\mathrm{a})\} \in \mathbb{R}^2$ can be defined as

$$\{\check{\boldsymbol{\tau}}^\mathrm{h}_\mathrm{a}(\xi^\mathrm{s}_\mathrm{a})\} = [\mathbf{R}]^\mathrm{T}\{\check{\boldsymbol{\eta}}^\mathrm{h}_\mathrm{a}\} . \tag{A.59}$$

The vector $\{\mathbf{m}_\mathrm{a}\}$, already introduced in Equation (A.38), for this case reads

$$\{\mathbf{m}_\mathrm{a}(\xi^\mathrm{s}_\mathrm{a})\} = \{\mathbf{x}^\mathrm{m}_l\} - \left[\mathbf{N}^\mathrm{s}_\mathrm{e}(\xi^\mathrm{s}_\mathrm{a})\right]\{\mathbf{x}^\mathrm{s}_\mathrm{e}\} . \tag{A.60}$$

Similarly to the matrix $[\check{\mathbf{I}}^\mathrm{s}_k]$ defined in Equation (A.42), the global matrix $[\check{\mathbf{I}}^\mathrm{m}_l] \in \mathbb{R}^{2\times 2\cdot n^\mathrm{m}}$ for this case comes as

$$[\check{\mathbf{I}}^\mathrm{m}_l] \equiv \left[\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \quad \cdots \quad \underbrace{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}}_{\text{node } l} \quad \cdots \quad \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \right] . \tag{A.61}$$

Finally, examining Equation (A.54) in more detail, it reveals helpful to construct a compact representation for the sum

$$\sum_{k=1}^{n^\mathrm{s}_\mathrm{e}} N^\mathrm{s}_k(\xi^\mathrm{s}_\mathrm{a})\, \mathscr{D}\tilde{\boldsymbol{\eta}}_k . \tag{A.62}$$

Therefore, the matrix $[\mathbf{K}_{\tilde{\eta}_\mathrm{e}}] \in \mathbb{R}^{(2\cdot n^\mathrm{s}_\mathrm{e})\times(2\cdot n^\mathrm{s})}$ is introduced, which is constructed by assembling line-by-line all the individual matrices $[\mathbf{K}_{\tilde{\eta}_k}]\left(k = 1,\cdots, n^\mathrm{s}_\mathrm{e}\right)$, defined in Equation (A.25), i.e.,

$$[\mathbf{K}_{\tilde{\eta}_\mathrm{e}}] \equiv \left[\begin{matrix} \mathbf{K}_{\tilde{\eta}_1} \\ \vdots \\ \mathbf{K}_{\tilde{\eta}_{n^\mathrm{s}_\mathrm{e}}} \end{matrix} \right] . \tag{A.63}$$

Each matrix block is related to the linearisation of the averaged unit normal vector at each non-mortar node and, more importantly, it makes it possible to write the sum Equation (A.62) as

$$\sum_{k=1}^{n^\mathrm{s}_\mathrm{e}} N^\mathrm{s}_k(\xi^\mathrm{s}_\mathrm{a})\, \mathscr{D}\tilde{\boldsymbol{\eta}}_k = \left[\mathbf{N}^\mathrm{s}_\mathrm{e}(\xi^\mathrm{s}_\mathrm{a})\right][\mathbf{K}_{\tilde{\eta}_\mathrm{e}}]\{\Delta\mathbf{d}_\mathcal{S}\} . \tag{A.64}$$

Making use of all the aforementioned matrices, the linearisation in Equation (A.53) can be arranged as

$$\mathscr{D}\xi^\mathrm{s}_\mathrm{a} = \{\mathbf{K}_{\xi^\mathrm{s}_\mathrm{a}}\} \left\{ \begin{matrix} \Delta\mathbf{d}_\mathcal{M} \\ \Delta\mathbf{d}_\mathcal{S} \end{matrix} \right\} , \tag{A.65}$$

with the *row* vector $\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}}(\xi^{\mathrm{s}}_{\mathrm{a}})\} \in \mathbb{R}^{2\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ constructed as

$$\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}}(\xi^{\mathrm{s}}_{\mathrm{a}})\} \equiv \{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{M}}\,,\, \mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{S}}\}\,. \tag{A.66}$$

The vector $\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{M}}(\xi^{\mathrm{s}}_{\mathrm{a}})\} \in \mathbb{R}^{2\cdot n^{\mathrm{m}}}$ is defined by

$$\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{M}}(\xi^{\mathrm{s}}_{\mathrm{a}})\} \equiv -\mathrm{C}_{\xi^{\mathrm{s}}_{\mathrm{a}}}\{\check{\boldsymbol{\tau}}^{\mathrm{h}}_{\mathrm{a}}\}^{\mathrm{T}}[\check{\mathbf{I}}^{\mathrm{m}}_{\mathrm{a}}]\,, \tag{A.67}$$

and the vector $\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{S}}(\xi^{\mathrm{s}}_{\mathrm{a}})\} \in \mathbb{R}^{2\cdot n^{\mathrm{s}}}$ as

$$\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{S}}(\xi^{\mathrm{s}}_{\mathrm{a}})\} \equiv \mathrm{C}_{\xi^{\mathrm{s}}_{\mathrm{a}}}\left(\{\mathbf{m}_{\mathrm{a}}\}^{\mathrm{T}}[\mathbf{R}]^{\mathrm{T}}[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{\mathrm{a}})][\mathbf{K}_{\tilde{\eta}_{\mathrm{e}}}] - \{\check{\boldsymbol{\tau}}^{\mathrm{h}}_{\mathrm{a}}\}^{\mathrm{T}}[\mathbf{N}^{\mathrm{s}}(\xi^{\mathrm{s}}_{\mathrm{a}})]\right)\,, \tag{A.68}$$

where the constant $\mathrm{C}_{\xi^{\mathrm{s}}_{\mathrm{a}}}(\xi^{\mathrm{s}}_{\mathrm{a}})$ is given by

$$\mathrm{C}_{\xi^{\mathrm{s}}_{\mathrm{a}}}(\xi^{\mathrm{s}}_{\mathrm{a}}) \equiv -\frac{1}{\{\mathbf{m}_{\mathrm{a}}\}^{\mathrm{T}}[\mathbf{R}]^{\mathrm{T}}[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{\mathrm{a}})]\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\} - \{\check{\boldsymbol{\tau}}^{\mathrm{h}}_{\mathrm{a}}\}^{\mathrm{T}}[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}(\xi^{\mathrm{s}}_{\mathrm{a}})]\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\}}\,. \tag{A.69}$$

A.1.3 Integration segment Gauss points

Due to the fact that the employed mortar integration scheme requires evaluating individual segment contributions, the linearisation of Gauss point coordinates $\mathscr{D}\xi^{\mathrm{s}}_{\mathrm{g}}$ and $\mathscr{D}\xi^{\mathrm{m}}_{\mathrm{g}}$ are typically non-zero as one or both segment end coordinates are possibly defined via deformation-dependent projections of mortar nodes. Recalling the mapping from segment Gauss point coordinate ζ_{g} to element Gauss point coordinate on the non-mortar side,

$$\xi^{\mathrm{s}}_{\mathrm{g}} = \frac{1}{2}(1-\zeta_{\mathrm{g}})\,\xi^{\mathrm{s}}_{\mathrm{a}} + \frac{1}{2}(1+\zeta_{\mathrm{g}})\,\xi^{\mathrm{s}}_{\mathrm{b}}\,, \tag{A.70}$$

its linearisation is relatively straightforward and yields

$$\mathscr{D}\xi^{\mathrm{s}}_{\mathrm{g}} = \frac{1}{2}(1-\zeta_{\mathrm{g}})\,\mathscr{D}\xi^{\mathrm{s}}_{\mathrm{a}} + \frac{1}{2}(1+\zeta_{\mathrm{g}})\,\mathscr{D}\xi^{\mathrm{s}}_{\mathrm{b}}\,. \tag{A.71}$$

The directional derivative of the Gauss point coordinate on the mortar side $\xi^{\mathrm{m}}_{\mathrm{g}}$ is more intricate, since it requires projecting the Gauss point $\xi^{\mathrm{s}}_{\mathrm{g}}$ from the non-mortar onto the mortar side, i.e., finding the coordinate $\xi^{\mathrm{m}}_{\mathrm{g}}$ satisfying

$$\left[\sum_{l=1}^{n^{\mathrm{m}}_{\mathrm{e}}}\left(N^{\mathrm{m}}_{l}(\xi^{\mathrm{m}}_{\mathrm{g}})\mathbf{x}^{\mathrm{m}}_{l}\right) - \mathbf{x}^{\mathrm{s}}_{\mathrm{g}}\right] \times \left[\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}}\left(N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathrm{g}})\tilde{\boldsymbol{\eta}}_{k}\right)\right] = \mathbf{0}\,. \tag{A.72}$$

For the sake of simplicity, in Equation (A.72), a slight abuse of notation is permitted[†] and $\mathbf{x}^{\mathrm{s}}_{\mathrm{g}}(\xi^{\mathrm{s}}_{\mathrm{g}})$ represents the position of the non-mortar element Gauss point, i.e.

$$\mathbf{x}^{\mathrm{s}}_{\mathrm{g}}(\xi^{\mathrm{s}}_{\mathrm{g}}) = \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k}(\xi^{\mathrm{s}}_{\mathrm{g}})\mathbf{x}^{\mathrm{s}}_{k}\,. \tag{A.73}$$

[†] According to the adopted notation, the exact representation would be $\{\boldsymbol{x}^{\mathrm{s}}_{\mathrm{g}}(\xi^{\mathrm{s}}_{\mathrm{g}})\}^{\mathrm{h}}$.

Once again, starting by expanding the cross product,

$$
\begin{aligned}
&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\mathrm{x}_{lx}^{\mathrm{m}}\right)-\mathrm{x}_{\mathrm{g}x}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{ky}\right)\right]\\
-&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\mathrm{x}_{ly}^{\mathrm{m}}\right)-\mathrm{x}_{\mathrm{g}y}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{kx}\right)\right]=0\,,
\end{aligned}
\tag{A.74}
$$

its linearisation yields

$$
\begin{aligned}
&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_{l,\xi}^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\,\mathscr{D}(\xi_{\mathrm{g}}^{\mathrm{m}})\,\mathrm{x}_{lx}^{\mathrm{m}}\right)+\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\Delta\mathrm{x}_{lx}^{\mathrm{m}}\right)-\mathscr{D}\mathrm{x}_{\mathrm{g}x}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{ky}\right)\right]\\
+&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\mathrm{x}_{lx}^{\mathrm{m}}\right)-\mathrm{x}_{\mathrm{g}x}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_{k,\xi}^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\tilde{\eta}_{ky}\right)+\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}\tilde{\eta}_{ky}\right)\right]\\
-&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_{l,\xi}^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\,\mathscr{D}(\xi_{\mathrm{g}}^{\mathrm{m}})\,\mathrm{x}_{ly}^{\mathrm{m}}\right)+\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\Delta\mathrm{x}_{ly}^{\mathrm{m}}\right)-\mathscr{D}\mathrm{x}_{\mathrm{g}y}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{kx}\right)\right]\\
-&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\mathrm{x}_{ly}^{\mathrm{m}}\right)-\mathrm{x}_{\mathrm{g}y}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_{k,\xi}^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\tilde{\eta}_{kx}\right)+\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}\tilde{\eta}_{kx}\right)\right]=0
\end{aligned}
\tag{A.75}
$$

and, consequently, $\mathscr{D}\xi_{\mathrm{g}}^{\mathrm{m}}$ can be expressed by

$$
\mathscr{D}\xi_{\mathrm{g}}^{\mathrm{m}}=-\frac{\mathrm{C}_{\xi_{\mathrm{g}}^{\mathrm{m}}}^{\mathrm{num}}}{\mathrm{C}_{\xi_{\mathrm{g}}^{\mathrm{m}}}^{\mathrm{den}}}\,,
\tag{A.76}
$$

with

$$
\begin{aligned}
\mathrm{C}_{\xi_{\mathrm{g}}^{\mathrm{m}}}^{\mathrm{num}}\equiv\;&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\Delta\mathrm{x}_{lx}^{\mathrm{m}}\right)-\mathscr{D}\mathrm{x}_{\mathrm{g}x}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{ky}\right)\right]\\
-&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\Delta\mathrm{x}_{ly}^{\mathrm{m}}\right)-\mathscr{D}\mathrm{x}_{\mathrm{g}y}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{kx}\right)\right]\\
+&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\mathrm{x}_{lx}^{\mathrm{m}}\right)-\mathrm{x}_{\mathrm{g}x}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_{k,\xi}^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\tilde{\eta}_{ky}\right)+\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}\tilde{\eta}_{ky}\right)\right]\\
-&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_l^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\mathrm{x}_{ly}^{\mathrm{m}}\right)-\mathrm{x}_{\mathrm{g}y}^{\mathrm{s}}\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_{k,\xi}^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\tilde{\eta}_{kx}\right)+\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\,\mathscr{D}\tilde{\eta}_{kx}\right)\right]
\end{aligned}
\tag{A.77}
$$

and

$$
\begin{aligned}
\mathrm{C}_{\xi_{\mathrm{g}}^{\mathrm{m}}}^{\mathrm{den}}\equiv\;&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_{l,\xi}^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\,\mathrm{x}_{lx}^{\mathrm{m}}\right)\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{ky}\right)\right]\\
-&\left[\sum_{l=1}^{n_{\mathrm{e}}^{\mathrm{m}}}\left(N_{l,\xi}^{\mathrm{m}}(\xi_{\mathrm{g}}^{\mathrm{m}})\,\mathrm{x}_{ly}^{\mathrm{m}}\right)\right]\left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}(\xi_{\mathrm{g}}^{\mathrm{s}})\tilde{\eta}_{kx}\right)\right].
\end{aligned}
\tag{A.78}
$$

Here the directional derivative $\mathscr{D}\xi_{\mathrm{g}}^{\mathrm{s}}$ can be obtained from Equation (A.71).

Note that when replacing the Gauss point subscript g by the segment end point index a, it is obvious the similar structure between the result above and the one given in Equation (A.32) for the projection of a non-mortar node onto a mortar segment. The additional complexity is that it involves the linearisation of not only the Gauss point $\mathbf{x}^{\mathrm{s}}_{\mathrm{g}}$, but all the averaged unit normal vectors associated with that non-mortar element as well. Nonetheless, the former is relatively straightforward and yields

$$\mathscr{D}\mathbf{x}^{\mathrm{s}}_{\mathrm{g}} = \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k,\xi}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right) \mathscr{D}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right) \mathbf{x}^{\mathrm{s}}_{k} \right) + \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} \left(N^{\mathrm{s}}_{k}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right) \Delta\mathbf{x}^{\mathrm{s}}_{k} \right) . \tag{A.79}$$

A.1.3.1 Algebraic representation

The linearisation of the non-mortar element Gauss point coordinate $\xi^{\mathrm{s}}_{\mathrm{g}}$ can be written as

$$\mathscr{D}\xi^{\mathrm{s}}_{\mathrm{g}} = \left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}}\right\} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.80}$$

with the *row* vector $\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}}\left(\zeta_{\mathrm{g}}\right)\right\} \in \mathbb{R}^{2\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ defined as

$$\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}}\left(\zeta_{\mathrm{g}}\right)\right\} \equiv \frac{1}{2}\left(1-\zeta_{\mathrm{g}}\right)\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}}\right\} + \frac{1}{2}\left(1+\zeta_{\mathrm{g}}\right)\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{b}}}\right\} . \tag{A.81}$$

Both vectors $\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}}\right\}$ and $\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{b}}}\right\}$ are constructed based on the non-mortar segment end coordinates linearisation (discussed in Appendix A.1.2), taking into account the node type. If it results from the projection of a mortar node onto the non-mortar element, one must recall Equation (A.66). If the point coincides with a non-mortar node, the associated vector vanishes. Following the separation between terms associated with non-mortar and mortar nodes, the vector $\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}}\right\}$ can be rearranged as follows:

$$\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}}\right\} = \left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}} \, , \, \mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\right\} . \tag{A.82}$$

Here, $\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}\left(\zeta_{\mathrm{g}}\right)\right\} \in \mathbb{R}^{2\cdot n^{\mathrm{m}}}$ and $\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\left(\zeta_{\mathrm{g}}\right)\right\} \in \mathbb{R}^{2\cdot n^{\mathrm{s}}}$ have a similar structure:

$$\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}\left(\zeta_{\mathrm{g}}\right)\right\} \equiv \frac{1}{2}\left(1-\zeta_{\mathrm{g}}\right)\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{M}}\right\} + \frac{1}{2}\left(1+\zeta_{\mathrm{g}}\right)\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{b}}\mathcal{M}}\right\} ; \tag{A.83}$$

$$\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\left(\zeta_{\mathrm{g}}\right)\right\} \equiv \frac{1}{2}\left(1-\zeta_{\mathrm{g}}\right)\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{a}}\mathcal{S}}\right\} + \frac{1}{2}\left(1+\zeta_{\mathrm{g}}\right)\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{b}}\mathcal{S}}\right\} . \tag{A.84}$$

Focusing now on the linearisation of the integration point on the mortar side $\xi^{\mathrm{m}}_{\mathrm{g}}$, firstly, an algebraic representation for the directional derivatives of the Gauss point coordinates on the non-mortar element $\mathbf{x}^{\mathrm{s}}_{\mathrm{g}}$ is obtained. This is done by introducing the vector $\left\{\mathscr{D}\mathbf{x}^{\mathrm{s}}_{\mathrm{g}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\} \in \mathbb{R}^{2}$, which follows as

$$\left\{\mathscr{D}\mathbf{x}^{\mathrm{s}}_{\mathrm{g}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\} = \left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}}\right] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.85}$$

with the matrix $\left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right] \in \mathbb{R}^{2\times 2\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ assembled as

$$\left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right] \equiv \left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}} \mid \mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\right] . \tag{A.86}$$

The block $\left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\in\mathbb{R}^{2\times 2\cdot n^{\mathrm{m}}}$ is defined by

$$\left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\equiv\left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\left\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\right\}\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}\right\}, \tag{A.87}$$

and contains the terms related to mortar nodes. The matrix $\left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\in\mathbb{R}^{2\times 2\cdot n^{\mathrm{s}}}$ combines with non-mortar nodes and is given as follows:

$$\left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\equiv\left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\left\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\right\}\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\right\}+\left[\mathbf{N}^{\mathrm{s}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]. \tag{A.88}$$

Additionally, the vector $\left\{\mathbf{m}_{\mathrm{g}}\right\}$ previously introduced in Equation (A.38) is used once again, although here connecting both non-mortar and mortar integration points, i.e.,

$$\left\{\mathbf{m}_{\mathrm{g}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}=\left[\mathbf{N}^{\mathrm{m}}_{\mathrm{e}}\left(\xi^{\mathrm{m}}_{\mathrm{g}}\right)\right]\left\{\mathbf{x}^{\mathrm{m}}_{\mathrm{e}}\right\}-\left[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\left\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\right\}. \tag{A.89}$$

Finally, the linearisation of the Gauss coordinate $\xi^{\mathrm{m}}_{\mathrm{g}}$ can be written as

$$\mathscr{D}\xi^{\mathrm{m}}_{\mathrm{g}}=\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}}\right\}\begin{Bmatrix}\Delta\mathbf{d}_{\mathcal{M}}\\ \Delta\mathbf{d}_{\mathcal{S}}\end{Bmatrix}, \tag{A.90}$$

where the *row* vector $\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}\in\mathbb{R}^{2\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ is assembled as follows

$$\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}\equiv\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}\mathcal{M}}\,,\,\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}\mathcal{S}}\right\}. \tag{A.91}$$

The vector $\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}\mathcal{M}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}\in\mathbb{R}^{2\cdot n^{\mathrm{m}}}$ is defined as

$$\begin{aligned}\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}\mathcal{M}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}\equiv C_{\xi^{\mathrm{m}}_{\mathrm{g}}}\Big[&\left\{\mathbf{m}_{\mathrm{g}}\right\}^{\mathrm{T}}\left[\mathbf{R}\right]^{\mathrm{T}}\left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\left\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\right\}\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}\right\}\\ &+\left\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\right\}^{\mathrm{T}}\left[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]^{\mathrm{T}}\left[\mathbf{R}\right]\Big(\left[\mathbf{N}^{\mathrm{m}}\left(\xi^{\mathrm{m}}_{\mathrm{g}}\right)\right]-\left[\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}\right]\Big)\Big]\end{aligned} \tag{A.92}$$

and the vector $\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}\mathcal{S}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}\in\mathbb{R}^{2\cdot n^{\mathrm{s}}}$ as

$$\begin{aligned}\left\{\mathbf{K}_{\xi^{\mathrm{m}}_{\mathrm{g}}\mathcal{S}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}\equiv C_{\xi^{\mathrm{m}}_{\mathrm{g}}}\Big[&\left\{\mathbf{m}_{\mathrm{g}}\right\}^{\mathrm{T}}\left[\mathbf{R}\right]^{\mathrm{T}}\Big(\left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\left\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\right\}\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\right\}+\left[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]\left[\mathbf{K}_{\tilde{\eta}_{\mathrm{e}}}\right]\Big)\\ &-\left\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\right\}^{\mathrm{T}}\left[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]^{\mathrm{T}}\left[\mathbf{R}\right]\left\{\mathbf{K}_{\xi^{\mathrm{s}}_{\mathrm{g}}\mathcal{S}}\right\}\Big]\end{aligned} \tag{A.93}$$

with constant $C_{\xi^{\mathrm{m}}_{\mathrm{g}}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)$ given by

$$C_{\xi^{\mathrm{m}}_{\mathrm{g}}}\left(\xi^{\mathrm{m}}_{\mathrm{g}},\xi^{\mathrm{s}}_{\mathrm{g}}\right)\equiv-\frac{1}{\left\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\right\}^{\mathrm{T}}\left[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right]^{\mathrm{T}}\left[\mathbf{R}\right]\left[\mathbf{B}^{\mathrm{m}}_{\mathrm{e}}\left(\xi^{\mathrm{m}}_{\mathrm{g}}\right)\right]\left\{\mathbf{x}^{\mathrm{m}}_{\mathrm{e}}\right\}}. \tag{A.94}$$

A.1.4 Integration segment Jacobian determinant

The employed numerical integration scheme, along with individual segment contributions, also requires the evaluation of the non-mortar element Jacobian and its directional derivative. Thus, recalling Equation (3.57), the non-mortar element Jacobian for a given Gauss point $\xi^{\mathrm{s}}_{\mathrm{g}}$ is evaluated as

$$J^{\mathrm{s}}_{\mathrm{e}}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\equiv\left\|\left\{\boldsymbol{x}^{\mathrm{s}}_{,\xi}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\right\}^{\mathrm{h}}\right\|=\left\|\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}}N^{\mathrm{s}}_{k,\xi}\left(\xi^{\mathrm{s}}_{\mathrm{g}}\right)\mathbf{x}^{\mathrm{s}}_{k}\right\| \tag{A.95}$$

and its linearisation yields

$$\mathscr{D}J_{\mathrm{e}}^{\mathrm{s}} = \frac{1}{J_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)} \left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k,\xi}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\mathbf{x}_{k}^{\mathrm{s}}\right)\right] \left[\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k,\xi\xi}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\mathscr{D}\xi_{\mathrm{g}}^{\mathrm{s}}\mathbf{x}_{k}^{\mathrm{s}}\right) + \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(N_{k,\xi}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\Delta\mathbf{x}_{k}^{\mathrm{s}}\right)\right]. \tag{A.96}$$

Note that the term $N_{k,\xi\xi}$ containing the second derivative, viz.

$$N_{k,\xi\xi} \equiv \frac{\mathrm{d}^2\, N_{k}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right)}{\left(\mathrm{d}\xi^{\mathrm{s}}\right)^2}, \tag{A.97}$$

only becomes relevant for second-order or higher-order finite element interpolation.

A.1.4.1 Algebraic representation

While the non-mortar element Jacobian determinant definition is relatively straightforward and can be written as

$$J_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right) = \|\left[\mathbf{B}_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right]\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{s}}\right\}\|, \tag{A.98}$$

its directional derivative requires introducing the matrix $\left[\mathbf{S}_{\mathrm{e}}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right)\right] \in \mathbb{R}^{2\cdot n_{\mathrm{e}}^{\mathrm{s}}}$, containing the shape functions second-order derivatives of a given non-mortar element e. In two dimensions it follows as

$$\left[\mathbf{S}_{\mathrm{e}}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right)\right] \equiv \begin{bmatrix} N_{1,\xi\xi}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right) & 0 & \cdots & N_{n_{\mathrm{e}}^{\mathrm{s}},\xi\xi}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right) & 0 \\ 0 & N_{1,\xi\xi}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right) & & 0 & N_{n_{\mathrm{e}}^{\mathrm{s}},\xi\xi}^{\mathrm{s}}\left(\xi^{\mathrm{s}}\right) \end{bmatrix}. \tag{A.99}$$

Then, the directional derivative of the non-mortar element Jacobian can be written as

$$\mathscr{D}J_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{m}},\xi_{\mathrm{g}}^{\mathrm{s}}\right) = \left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}\right\} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix}, \tag{A.100}$$

with the *row* vector $\left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}\left(\xi_{\mathrm{g}}^{\mathrm{m}},\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right\} \in \mathbb{R}^{2\cdot\left(n^{\mathrm{m}}+n^{\mathrm{s}}\right)}$ assembled as

$$\left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}\left(\xi_{\mathrm{g}}^{\mathrm{m}},\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right\} \equiv \left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}\mathcal{M}}\,,\, \mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}\mathcal{S}}\right\}. \tag{A.101}$$

The vector $\left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}\mathcal{M}}\left(\xi_{\mathrm{g}}^{\mathrm{m}},\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right\} \in \mathbb{R}^{2\cdot n^{\mathrm{m}}}$, related with mortar nodes, is defined as

$$\left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}\mathcal{M}}\right\} \equiv \frac{1}{J_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)} \left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{s}}\right\}^{\mathrm{T}} \left[\mathbf{B}_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right]^{\mathrm{T}} \left[\mathbf{S}_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right] \left\{\boldsymbol{x}_{\mathrm{e}}^{\mathrm{s}}\right\} \left\{\mathbf{K}_{\xi_{\mathrm{g}}^{\mathrm{s}}\mathcal{M}}\right\}, \tag{A.102}$$

and the vector $\left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}\mathcal{S}}\left(\xi_{\mathrm{g}}^{\mathrm{m}},\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right\} \in \mathbb{R}^{2\cdot n^{\mathrm{s}}}$, related with non-mortar nodes, is defined as

$$\left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}\mathcal{S}}\right\} \equiv \frac{1}{J_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)} \left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{s}}\right\}^{\mathrm{T}} \left[\mathbf{B}_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right]^{\mathrm{T}} \left(\left[\mathbf{S}_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right] \left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{s}}\right\} \left\{\mathbf{K}_{\xi_{\mathrm{g}}^{\mathrm{s}}\mathcal{S}}\right\} + \left[\mathbf{B}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right]\right). \tag{A.103}$$

Remark. *As mentioned, for first-order finite element interpolation, the second-order derivatives vanishes and, therefore, the linearisation of the non-mortar element Jacobian determinant reduces to*

$$\mathscr{D}J_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right) = \left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}\right\}\left\{\Delta\mathbf{d}_{\mathcal{S}}\right\}, \tag{A.104}$$

with

$$\left\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right\} = \frac{1}{J_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)} \left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{s}}\right\}^{\mathrm{T}} \left[\mathbf{B}_{\mathrm{e}}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right]^{\mathrm{T}} \left[\mathbf{B}^{\mathrm{s}}\left(\xi_{\mathrm{g}}^{\mathrm{s}}\right)\right]. \tag{A.105}$$

A.2 Linearisation of mortar contact in three-dimensions

A.2.1 Averaged normal and tangential vectors

The linearisation of the three-dimensional version of the averaged unit normal $\tilde{\boldsymbol{\eta}}_j$ and tangential vectors $\tilde{\boldsymbol{\tau}}^1_j$ and $\tilde{\boldsymbol{\tau}}^2_j$ at a non-mortar node $j \in \mathcal{S}$ can be treated in a very similar way to the two-dimensional counterpart addressed in Appendix A.1.1. Thus, the following paragraphs focus only on the additional complexities that emerge in three dimensions.

Firstly, the definition of the non-unit normal $\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}$ to the adjacent element e given Equation (A.5) is no longer valid. Nonetheless, in three dimensions, it can still be obtained as

$$\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right) \equiv \check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_1} \times \check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_2}\,. \tag{A.106}$$

Here, $\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)$ represents the *non-unit* tangent vector to the adjacent non-mortar element e along the isoparametric coordinate $\xi_i\;(i=1,2)$, which can be obtained as

$$\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right) \equiv \left\{\boldsymbol{x}^{\mathrm{s}}_{,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\}^{\mathrm{h}} = \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\mathbf{x}^{\mathrm{s}}_k\,, \tag{A.107}$$

Recall that, according to the notation used previously, $N^{\mathrm{s}}_{k,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)$ represents the shape function derivatives with respect to each local coordinate ξ_i, i.e.,

$$N^{\mathrm{s}}_{k,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}\right) \equiv \frac{\mathrm{d}N^{\mathrm{s}}_k\left(\boldsymbol{\xi}^{\mathrm{s}}\right)}{\mathrm{d}\xi^{\mathrm{s}}_i}\,. \tag{A.108a}$$

From Equation (A.106), the directional derivative of the non-unit normal $\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}$ follows as

$$\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}} = \mathscr{D}\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_1} \times \check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_2} + \check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_1} \times \mathscr{D}\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_2}\,, \tag{A.109}$$

and, recalling Equation (A.107), the linearisation of the non-unit tangent vector $\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\xi_i}$ yields

$$\mathscr{D}\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i} = \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\Delta\mathbf{x}^{\mathrm{s}}_k\,. \tag{A.110}$$

With the expressions introduced so far, the averaged unit normal $\tilde{\boldsymbol{\eta}}_j$ can be already obtained by following the expressions given in Appendix A.1.1 for two-dimensional problems. However, the definition of the associated nodal tangent vectors $\tilde{\boldsymbol{\tau}}^i_j\left(i=1,2\right)$ requires further investigation. While in two dimensions, the tangent vector is uniquely defined by using Equation (A.9), in three dimensions, the pair of tangent vectors $\tilde{\boldsymbol{\tau}}^i_j$ is not unique within the tangent plane to the non-mortar node j. The strategy adopted in this work to solve this problem is explained in detail in Chapter 3, and is based on considering the projection of the first edge connected with the non-mortar node j.

First of all, let one define $\boldsymbol{e}_i (i = 1, ..., n_e)$ as the edges connected to the non-mortar node j (i.e., vector connecting the node j to the other node of the edge); herein n_e stands for the total number of associated edges. By projecting one of the edges, for example $\boldsymbol{e}_1$, onto the plane $\mathcal{T}$ that passes the non-mortar node j and is orthogonal to its averaged unit normal vector $\tilde{\boldsymbol{\eta}}_j$, one can define the averaged unit tangent vector $\tilde{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})$ as

$$\tilde{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}}) \equiv \frac{\check{\boldsymbol{\tau}}_j^1}{\left\|\check{\boldsymbol{\tau}}_j^1\right\|}, \tag{A.111}$$

where $\check{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})$ represents the non-unit tangent vector obtained by the projection, i.e.,

$$\check{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}}) \equiv \left(\boldsymbol{I} - \tilde{\boldsymbol{\eta}}_j \otimes \tilde{\boldsymbol{\eta}}_j\right) \boldsymbol{e}_1 . \tag{A.112}$$

Denoting the length of the non-unit tangent vector $\check{\boldsymbol{\tau}}_j^1$ as $l_{\check{\tau}_j^1} \equiv \|\check{\boldsymbol{\tau}}_j^1\|$, the directional derivative of Equation (A.111) yields

$$\mathscr{D}\tilde{\boldsymbol{\tau}}_j^1 = \frac{\mathscr{D}\check{\boldsymbol{\tau}}_j^1}{l_{\check{\tau}_j^1}} - \frac{\left(\check{\boldsymbol{\tau}}_j^1 \cdot \mathscr{D}\check{\boldsymbol{\tau}}_j^1\right)\check{\boldsymbol{\tau}}_j^1}{l_{\check{\tau}_j^1}^3} = \left(\frac{1}{l_{\check{\tau}_j^1}}\boldsymbol{I} - \frac{1}{l_{\check{\tau}_j^1}^3}\check{\boldsymbol{\tau}}_j^1 \otimes \check{\boldsymbol{\tau}}_j^1\right)\mathscr{D}\check{\boldsymbol{\tau}}_j^1 . \tag{A.113}$$

This requires obtaining the directional derivative of the non-unit tangent vector $\check{\boldsymbol{\tau}}_j^1$ which, considering Equation (A.112), can be written as

$$\begin{aligned} \mathscr{D}\check{\boldsymbol{\tau}}_j^1 &= \left(\boldsymbol{I} - \tilde{\boldsymbol{\eta}}_j \otimes \tilde{\boldsymbol{\eta}}_j\right)\mathscr{D}\boldsymbol{e}_1 - \left(\mathscr{D}\tilde{\boldsymbol{\eta}}_j \otimes \tilde{\boldsymbol{\eta}}_j\right)\boldsymbol{e}_1 - \left(\tilde{\boldsymbol{\eta}}_j \otimes \mathscr{D}\tilde{\boldsymbol{\eta}}_j\right)\boldsymbol{e}_1 \\ &= \left(\boldsymbol{I} - \tilde{\boldsymbol{\eta}}_j \otimes \tilde{\boldsymbol{\eta}}_j\right)\mathscr{D}\boldsymbol{e}_1 - \left(\tilde{\boldsymbol{\eta}}_j \cdot \boldsymbol{e}_1 + \tilde{\boldsymbol{\eta}}_j \otimes \boldsymbol{e}_1\right)\mathscr{D}\tilde{\boldsymbol{\eta}}_j . \end{aligned} \tag{A.114}$$

Finally, the second averaged unit tangent vector $\tilde{\boldsymbol{\tau}}_j^2(\boldsymbol{\xi}_j^{\mathrm{s}})$ is obtained as

$$\tilde{\boldsymbol{\tau}}_j^2(\boldsymbol{\xi}_j^{\mathrm{s}}) \equiv \tilde{\boldsymbol{\eta}}_j \times \tilde{\boldsymbol{\tau}}_j^1 , \tag{A.115}$$

and its directional derivative follows as

$$\mathscr{D}\tilde{\boldsymbol{\tau}}_j^2 = \mathscr{D}\tilde{\boldsymbol{\eta}}_j \times \tilde{\boldsymbol{\tau}}_j^1 + \tilde{\boldsymbol{\eta}}_j \times \mathscr{D}\tilde{\boldsymbol{\tau}}_j^1 . \tag{A.116}$$

A.2.1.1 Algebraic representation

Starting by the derivation of the algebraic representation for the non-unit tangent vector to the adjacent element e defined in Equation (A.107), firstly, it proves convenient to define the three-dimensional version of the element and global discrete gradient operators previously introduced in Equations (A.12) and (A.13), respectively. On the one hand, the element matrix $\left[\mathbf{B}_{\mathrm{e},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\right] \in \mathbb{R}^{3\times 3\cdot n_{\mathrm{e}}^{\mathrm{s}}}$ contains the shape functions derivatives with respect to ξ_i at a given isoparametric coordinate $\boldsymbol{\xi}^{\mathrm{s}}$, i.e.,

$$\left[\mathbf{B}_{\mathrm{e},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\right] \equiv \begin{bmatrix} N_{1,\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 & 0 & & N_{n_{\mathrm{e}}^{\mathrm{s}},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 & 0 \\ 0 & N_{1,\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 & \cdots & 0 & N_{n_{\mathrm{e}}^{\mathrm{s}},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 \\ 0 & 0 & N_{1,\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) & & 0 & 0 & N_{n_{\mathrm{e}}^{\mathrm{s}},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) \end{bmatrix} . \tag{A.117}$$

On the other hand, its global counterpart is denoted as $\left[\mathbf{B}^{\mathrm{s}}_{\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ and comes as

$$\left[\mathbf{B}^{\mathrm{s}}_{\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}\right)\right] \equiv \begin{bmatrix} 0 & 0 & 0 & & N^{\mathrm{s}}_{k,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}\right) & 0 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & N^{\mathrm{s}}_{k,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}\right) & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 & N^{\mathrm{s}}_{k,\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}\right) & & 0 & 0 & 0 \end{bmatrix}. \tag{A.118}$$

With these definitions at hand, the non-unit tangent vector $\left\{\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} \in \mathbb{R}^3$ to the adjacent element e along the direction ξ^{s}_i can be expressed as

$$\left\{\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} = \left[\mathbf{B}^{\mathrm{s}}_{\mathrm{e},\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right]\left\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\right\}, \tag{A.119}$$

and its directional derivative vector $\left\{\mathscr{D}\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} \in \mathbb{R}^3$ gives

$$\left\{\mathscr{D}\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} = \left[\mathbf{B}^{\mathrm{s}}_{\xi_i}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right]\left\{\Delta\mathbf{d}_{\mathcal{S}}\right\}. \tag{A.120}$$

Recall that the definition of the element vector of nodal coordinates $\left\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\right\}$ is assembled as in Equation (A.11). The non-unit normal vector $\left\{\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} \in \mathbb{R}^3$ to a given adjacent non-mortar element e defined in Equation (A.106) follows as

$$\left\{\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} = \left[\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\right]_{\times}\left\{\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_2}\right\}. \tag{A.121}$$

Here, the matrix notation $\left[\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\right]_{\times} \in \mathbb{R}^{3\times 3}$ is a skew-symmetric matrix representation of the vector $\left\{\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_i}\right\}$, which establishes an alternative representation for the cross product between two vectors. For a generic vector $\left\{\boldsymbol{a}\right\} \in \mathbb{R}^3$, the matrix $\left[\boldsymbol{a}\right]_{\times} \in \mathbb{R}^{3\times 3}$ is assembled as

$$\left[\boldsymbol{a}\right]_{\times} \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}, \tag{A.122}$$

such that the cross product of $\left\{\boldsymbol{a}\right\}$ with other generic vector $\left\{\boldsymbol{b}\right\} \in \mathbb{R}^3$ can be written as

$$\left\{\boldsymbol{a}\right\} \times \left\{\boldsymbol{b}\right\} = \left[\boldsymbol{a}\right]_{\times}\left\{\boldsymbol{b}\right\} = \left[\boldsymbol{b}\right]^{\mathrm{T}}_{\times}\left\{\boldsymbol{a}\right\}. \tag{A.123}$$

The vector $\left\{\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} \in \mathbb{R}^3$ represents its directional derivative and, accordingly to Equation (A.110), can be written as

$$\left\{\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{a}}_{\mathrm{e}}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} = \left[\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}\right]\left\{\Delta\mathbf{d}_{\mathcal{S}}\right\} \tag{A.124}$$

with the global matrix $\left[\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ being

$$\left[\mathbf{K}_{\check{\eta}^{\mathrm{a}}_{\mathrm{e}}}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right] \equiv \left[\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_2}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right]^{\mathrm{T}}_{\times}\left[\mathbf{B}^{\mathrm{s}}_{\xi_1}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right] + \left[\check{\boldsymbol{\tau}}^{\mathrm{a}}_{\mathrm{e},\xi_1}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right]_{\times}\left[\mathbf{B}^{\mathrm{s}}_{\xi_2}\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right] \tag{A.125}$$

The remaining operations needed to obtain the averaged unit normal vector $\left\{\tilde{\boldsymbol{\eta}}_j\left(\boldsymbol{\xi}^{\mathrm{s}}_j\right)\right\} \in \mathbb{R}^3$ have already been introduced in Appendix A.1.1 and are equally applicable in three dimensions.

Focusing now on the algebraic representation of the first averaged unit tangent vector $\{\tilde{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} \in \mathbb{R}^3$ defined in (A.111), it follows that

$$\{\tilde{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} = \frac{1}{l_{\check{\tau}_j^1}}\{\check{\boldsymbol{\tau}}_j^1\}\,. \tag{A.126}$$

Here, the vector $\{\check{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} \in \mathbb{R}^3$ represents the non-unit tangent vector with length $l_{\check{\tau}_j^1}$ and, considering (A.112), it can be obtained as

$$\{\check{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} = \left([\mathbf{I}] - \{\tilde{\boldsymbol{\eta}}_j\}\{\tilde{\boldsymbol{\eta}}_j\}^{\mathrm{T}}\right)\{\boldsymbol{e}_1\}\,. \tag{A.127}$$

The vector $\{\boldsymbol{e}_1\} \in \mathbb{R}^3$ represents the first edge that passes the non-mortar node j, which, by denoting the second node of the edge by k, is defined as

$$\{\boldsymbol{e}_1\} \equiv \{\mathbf{x}_k^{\mathrm{s}}\} - \{\mathbf{x}_j^{\mathrm{s}}\}\,. \tag{A.128}$$

The directional derivative of the averaged tangent vector $\{\mathscr{D}\tilde{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} \in \mathbb{R}^3$ can be computed as

$$\{\mathscr{D}\tilde{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} = [\mathbf{K}_{\tilde{\tau}_j^1}]\{\Delta\mathbf{d}_{\mathcal{S}}\}\,. \tag{A.129}$$

with the matrix $[\mathbf{K}_{\tilde{\tau}_j^1}(\boldsymbol{\xi}_j^{\mathrm{s}})] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ defined as

$$[\mathbf{K}_{\tilde{\tau}_j^1}(\boldsymbol{\xi}_j^{\mathrm{s}})] \equiv \left(\frac{1}{l_{\check{\tau}_j^1}}[\mathbf{I}] - \frac{1}{l_{\check{\tau}_j^1}^3}\{\check{\boldsymbol{\tau}}_j^1\}\{\check{\boldsymbol{\tau}}_j^1\}^{\mathrm{T}}\right)[\mathbf{K}_{\check{\tau}_j^1}]\,. \tag{A.130}$$

Here, the matrix $[\mathbf{K}_{\check{\tau}_j^1}(\boldsymbol{\xi}_j^{\mathrm{s}})] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ represents the directional derivative of the non-unit tangent vector $\{\mathscr{D}\check{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} \in \mathbb{R}^3$, i.e.,

$$\{\mathscr{D}\check{\boldsymbol{\tau}}_j^1(\boldsymbol{\xi}_j^{\mathrm{s}})\} \equiv [\mathbf{K}_{\check{\tau}_j^1}]\{\Delta\mathbf{d}_{\mathcal{S}}\}\,, \tag{A.131}$$

which can be defined as

$$[\mathbf{K}_{\check{\tau}_j^1}(\boldsymbol{\xi}_j^{\mathrm{s}})] \equiv \left([\mathbf{I}] - \{\tilde{\boldsymbol{\eta}}_j\}\{\tilde{\boldsymbol{\eta}}_j\}^{\mathrm{T}}\right)[\mathbf{K}_{e_1^j}] - \left(\{\tilde{\boldsymbol{\eta}}_j\}^{\mathrm{T}}\{\boldsymbol{e}_1\} + \{\tilde{\boldsymbol{\eta}}_j\}\{\boldsymbol{e}_1\}^{\mathrm{T}}\right)[\mathbf{K}_{\tilde{\eta}_j}(\boldsymbol{\xi}_j^{\mathrm{s}})]\,. \tag{A.132}$$

In this equation, $[\mathbf{K}_{e_1}] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ stands for the directional derivative of the edge,

$$[\mathbf{K}_{e_1}] \equiv [\check{\mathbf{I}}_k^{\mathrm{s}}] - [\check{\mathbf{I}}_j^{\mathrm{s}}]\,, \tag{A.133}$$

with $[\check{\mathbf{I}}_k^{\mathrm{s}}] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ denoting the auxiliary matrix previously introduced in Equation (A.42), which in three dimensions reads

$$[\check{\mathbf{I}}_k^{\mathrm{s}}] \equiv \left[\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} \quad \cdots \quad \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{\text{node } k} \quad \cdots \quad \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}\right]. \tag{A.134}$$

Finally, once obtained the first averaged unit tangent vector $\{\tilde{\boldsymbol{\tau}}_j^1\}$, the second averaged unit tangent vector $\{\tilde{\boldsymbol{\tau}}_j^2(\boldsymbol{\xi}_j^{\mathrm{s}})\} \in \mathbb{R}^3$ follows as

$$\{\tilde{\boldsymbol{\tau}}_j^2(\boldsymbol{\xi}_j^{\mathrm{s}})\} = [\tilde{\boldsymbol{\eta}}_j]_\times \{\tilde{\boldsymbol{\tau}}_j^1\}\,. \tag{A.135}$$

Its directional derivative is represented by the vector $\{\mathscr{D}\tilde{\boldsymbol{\tau}}_j^2(\boldsymbol{\xi}_j^{\mathrm{s}})\} \in \mathbb{R}^3$, which comes

$$\{\mathscr{D}\tilde{\boldsymbol{\tau}}_j^2(\boldsymbol{\xi}_j^{\mathrm{s}})\} = [\mathbf{K}_{\tilde{\tau}_j^2}]\{\Delta\mathbf{d}_{\mathcal{S}}\}\,, \tag{A.136}$$

with the matrix $[\mathbf{K}_{\tilde{\tau}_j^2}(\boldsymbol{\xi}_j^{\mathrm{s}})] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ defined as

$$[\mathbf{K}_{\tilde{\tau}_j^2}(\boldsymbol{\xi}_j^{\mathrm{s}})] \equiv [\tilde{\boldsymbol{\tau}}_j^1]_\times^{\mathrm{T}}[\mathbf{K}_{\tilde{\eta}_j}] + [\tilde{\boldsymbol{\eta}}_j]_\times[\mathbf{K}_{\tilde{\tau}_j^1}]\,. \tag{A.137}$$

A.2.2 Auxiliary plane

The next stage in the linearisation of three-dimensional mortar-based contact algorithms is the treatment of the entities associated with the auxiliary plane $\mathcal{X}$. As explained in Section 3.7, the algorithmic approach for mortar integrals in three dimensions requires the establishment of piecewise flat segments along the non-mortar boundary. Underlying these segments is the definition of the auxiliary plane $\mathcal{X}$, which, as illustrated in Figure A.2, is based on the non-mortar element center $\{\boldsymbol{x}_\odot^{\mathrm{s}}\}^{\mathrm{h}}$ and the corresponding averaged unit normal vector $\boldsymbol{\eta}_\odot^{\mathrm{h}}$.

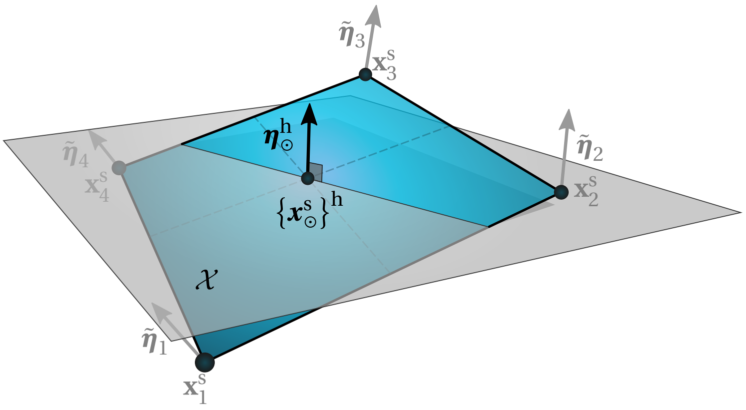


Figure A.2: Definition of the auxiliary plane for mortar integrals approximation.

On the one hand, due to the employed isoparametric discretisation scheme, the center of the element $\{\boldsymbol{x}_\odot^{\mathrm{s}}(\boldsymbol{\xi}_\odot^{\mathrm{s}})\}^{\mathrm{h}}$ can be obtained

$$\{\boldsymbol{x}_\odot^{\mathrm{s}}(\boldsymbol{\xi}_\odot^{\mathrm{s}})\}^{\mathrm{h}} \equiv \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} N_k^{\mathrm{s}}(\boldsymbol{\xi}_\odot^{\mathrm{s}})\,\mathbf{x}_k^{\mathrm{s}}\,. \tag{A.138}$$

For example, for quadrilateral elements, one has

$$\boldsymbol{\xi}_\odot^{\mathrm{s}} = \mathbf{0}\,. \tag{A.139}$$

The associated directional derivative yields

$$\{\mathscr{D}\boldsymbol{x}^{\mathrm{s}}_{\odot}\}^{\mathrm{h}} = \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\Delta\mathbf{x}^{\mathrm{s}}_{k}\,. \tag{A.140}$$

On the other hand, the averaged unit normal $\boldsymbol{\eta}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})$ can be obtained by FE interpolation as

$$\boldsymbol{\eta}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot}) \equiv \frac{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}}{\|\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}\|}\,, \tag{A.141}$$

with $\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})$ denoting the non-unit averaged unit normal vector, i.e.,

$$\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot}) \equiv \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\tilde{\boldsymbol{\eta}}^{\mathrm{s}}_{k}\,. \tag{A.142}$$

Its directional derivative follows as

$$\mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{\odot} = \left(\frac{1}{l_{\check{\eta}^{\mathrm{h}}_{\odot}}}\boldsymbol{I} - \frac{1}{l^{3}_{\check{\eta}^{\mathrm{h}}_{\odot}}}\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}\otimes\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}\right)\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot} \tag{A.143}$$

where $l_{\check{\eta}^{\mathrm{h}}_{\odot}} \equiv \|\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}\|$ has been introduced as the length of the non-unit interpolated normal vector $\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}$, with directional derivative

$$\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot} = \sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\,\mathscr{D}\tilde{\boldsymbol{\eta}}_{k}\,. \tag{A.144}$$

A.2.2.1 Algebraic representation

The algebraic treatment of the equations above is relatively straightforward, having the majority of the entities involved been introduced previously already. The only additional requirement is the three-dimensional version of the FE interpolation matrix $[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}})] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}_{\mathrm{e}}}$, which is now assembled as

$$[\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}})] \equiv \begin{bmatrix} N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 & 0 & & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 & 0 \\ 0 & N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 & \cdots & 0 & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}}) & 0 \\ 0 & 0 & N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}}) & & 0 & 0 & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}}) \end{bmatrix}. \tag{A.145}$$

Firstly, the center of the non-mortar element is represented by the vector $\{\boldsymbol{x}^{\mathrm{s}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\}^{\mathrm{h}} \in \mathbb{R}^{3}$, which can be computed as

$$\{\boldsymbol{x}^{\mathrm{s}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\}^{\mathrm{h}} = [\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})]\{\mathbf{x}^{\mathrm{s}}_{\mathrm{e}}\}\,, \tag{A.146}$$

and its directional derivative by the vector $\{\mathscr{D}\boldsymbol{x}^{\mathrm{s}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\}^{\mathrm{h}} \in \mathbb{R}^{3}$, which reads

$$\{\mathscr{D}\boldsymbol{x}^{\mathrm{s}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\}^{\mathrm{h}} = [\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})]\{\Delta\mathbf{d}_{\mathcal{S}}\}\,. \tag{A.147}$$

Lastly, the averaged unit normal vector $\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} \in \mathbb{R}^{3}$ follows as

$$\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} = \frac{1}{l_{\check{\eta}^{\mathrm{h}}_{\odot}}}\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}\}\,, \tag{A.148}$$

with the non-unit averaged normal vector $\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} \in \mathbb{R}^3$, with length $l_{\check{\eta}^{\mathrm{h}}_{\odot}}$, being obtained as

$$\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} = [\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})]\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\}\,. \tag{A.149}$$

Recall that the element vector $\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\}$ is assembled just as in Equation (A.57). Finally, its directional derivative, $\{\mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} \in \mathbb{R}^3$, is defined as

$$\{\mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} = [\mathbf{K}_{\eta^{\mathrm{h}}_{\odot}}]\{\Delta\mathbf{d}_{\mathcal{S}}\}\,, \tag{A.150}$$

with the matrix $[\mathbf{K}_{\eta^{\mathrm{h}}_{\odot}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ written as

$$[\mathbf{K}_{\eta^{\mathrm{h}}_{\odot}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})] \equiv \left(\frac{1}{l_{\check{\eta}^{\mathrm{h}}_{\odot}}}[\mathbf{I}] - \{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}\}\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}\right)[\mathbf{K}_{\check{\eta}^{\mathrm{h}}_{\odot}}]\,. \tag{A.151}$$

Here, the matrix $[\mathbf{K}_{\check{\eta}^{\mathrm{h}}_{\odot}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ denotes the directional derivative of the non-unit averaged normal vector, $\{\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} \in \mathbb{R}^3$, i.e.,

$$\{\mathscr{D}\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\odot}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})\} = [\mathbf{K}_{\check{\eta}^{\mathrm{h}}_{\odot}}]\{\Delta\mathbf{d}_{\mathcal{S}}\}\,, \tag{A.152}$$

and can be defined as

$$[\mathbf{K}_{\check{\eta}^{\mathrm{h}}_{\odot}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})] \equiv [\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\odot})][\mathbf{K}_{\tilde{\eta}_{\mathrm{e}}}]\,. \tag{A.153}$$

It should be reminded that the matrix $[\mathbf{K}_{\tilde{\eta}_{\mathrm{e}}}]$ in the equation above has been introduced in Equation (A.63), which gathers all the directional derivatives associated with the averaged unit normal vector at each node of the non-mortar element.

A.2.3 Integration cell vertices

As mentioned in Appendix A.2.2, the algorithm used to approximate mortar integrals is based on the subdivision of the active interface into piecewise flat segments—here termed as *integration cells*. These easy-to-integrate triangular subdomains result from a thorough division of the overlapping region between the non-mortar and mortar surfaces, which is based on suitable nodal projections, polygon clipping and Delaunay triangulation algorithms. Therefore, one must be aware of the different cases from which an integration cell vertex may have been originated. This can be visualized in Figure A.3, in which:

- The first and last integration cell vertices, $\mathbf{x}^{\mathrm{cell}}_1$ and $\mathbf{x}^{\mathrm{cell}}_3$, are the projection of a non-mortar and mortar node, accordingly, onto the auxiliary plane $\mathcal{X}$;

- the second node $\mathbf{x}^{\mathrm{cell}}_2$ results from the intersection of two projected non-mortar and mortar edges and, thus, is originated by the line clipping operation.

In order to be able to derive the consistent linearisation of the vertices, the different types of nodes must be appropriately identified and treated accordingly. In the following paragraphs, each case is discussed individually.

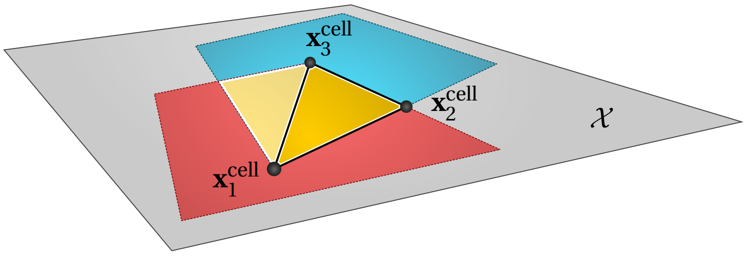


Figure A.3: Different cases for the definition of the integration cell vertices: $\mathbf{x}_1^{\mathrm{cell}}$ is a projected non-mortar node; $\mathbf{x}_2^{\mathrm{cell}}$ is generated by the line clipping algorithm; $\mathbf{x}_3^{\mathrm{cell}}$ is a projected mortar node.

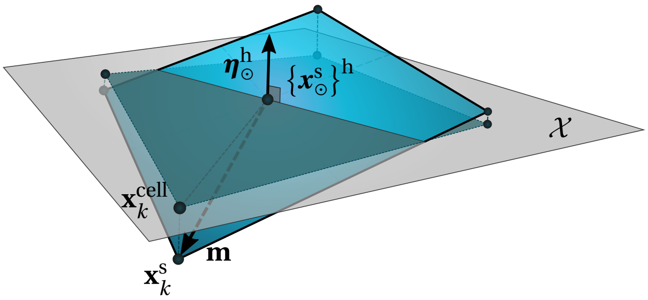


Figure A.4: Projection of a non-mortar node onto the auxiliary plane $\mathcal{X}$.

A.2.3.1 Case 1: Projection of a non-mortar/mortar node onto the auxiliary plane

The first case to be analysed is the projection of a non-mortar/mortar node onto the auxiliary plane $\mathcal{X}$. An exemplary projection of a non-mortar node is illustrated in Figure A.4. Nonetheless, it is important to note that, since the procedure is the same, either the element is non-mortar or mortar, the discussion throughout the following paragraphs is applicable for both cases.

By considering the projection formula of a node $\mathbf{x}_k^{\mathrm{s/m}}$ (either non-mortar or mortar) onto the auxiliary plane $\mathcal{X}$, it follows that

$$\begin{aligned}\mathbf{x}_k^{\mathrm{cell}} &= \{\boldsymbol{x}_{\odot}^{\mathrm{s}}\}^{\mathrm{h}} + \left(\boldsymbol{I} - \boldsymbol{\eta}_{\odot}^{\mathrm{h}} \otimes \boldsymbol{\eta}_{\odot}^{\mathrm{h}}\right)\mathbf{m} \\ &= \mathbf{x}_k^{\mathrm{s/m}} - \left(\boldsymbol{\eta}_{\odot}^{\mathrm{h}} \otimes \boldsymbol{\eta}_{\odot}^{\mathrm{h}}\right)\mathbf{m}\,,\end{aligned} \tag{A.154}$$

where the auxiliary vector $\mathbf{m}$ stands for the vector connecting the center of the non-mortar element, $\{\boldsymbol{x}_{\odot}^{\mathrm{s}}\}^{\mathrm{h}}$, and the node $\mathbf{x}_k^{\mathrm{s/m}}$ being projected, i.e.,

$$\mathbf{m} \equiv \mathbf{x}_k^{\mathrm{s/m}} - \{\boldsymbol{x}_{\odot}^{\mathrm{s}}\}^{\mathrm{h}}\,. \tag{A.155}$$

Its directional derivative results in

$$\mathscr{D}\mathbf{x}_k^{\mathrm{cell}} = \Delta\mathbf{x}_k^{\mathrm{s/m}} - \left(\boldsymbol{\eta}_{\odot}^{\mathrm{h}} \otimes \boldsymbol{\eta}_{\odot}^{\mathrm{h}}\right)\mathscr{D}\mathbf{m} - \left(\boldsymbol{\eta}_{\odot}^{\mathrm{h}} \cdot \mathbf{m} + \boldsymbol{\eta}_{\odot}^{\mathrm{h}} \otimes \mathbf{m}\right)\mathscr{D}\boldsymbol{\eta}_{\odot}^{\mathrm{h}}\,, \tag{A.156}$$

with

$$\mathscr{D}\mathbf{m} \equiv \Delta\mathbf{x}_k^{\mathrm{s/m}} - \{\mathscr{D}\boldsymbol{x}_{\odot}^{\mathrm{s}}\}^{\mathrm{h}}\,. \tag{A.157}$$

A.2.3.2 Algebraic representation

The algebraic representation of the projection operation is straightforward and the node $\{\mathbf{x}_k^{\text{cell}}\} \in \mathbb{R}^3$ is computed by

$$\{\mathbf{x}_k^{\text{cell}}\} = \{\mathbf{x}_k^{\text{s/m}}\} - \{\boldsymbol{\eta}_\odot^{\text{h}}\}\{\boldsymbol{\eta}_\odot^{\text{h}}\}^{\text{T}}\{\mathbf{m}\}\,. \tag{A.158}$$

Recall that the vector $\{\mathbf{m}\} \in \mathbb{R}^3$ can be directly computed by using Equation (A.155). Its directional derivative is represented by the vector $\{\mathscr{D}\mathbf{x}_k^{\text{cell}}\} \in \mathbb{R}^3$, which is defined as

$$\{\mathscr{D}\mathbf{x}_k^{\text{cell}}\} = [\mathbf{K}_{\text{x}_k^{\text{cell}}}] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix}\,. \tag{A.159}$$

Here, the matrix $[\mathbf{K}_{\text{x}_k^{\text{cell}}}] \in \mathbb{R}^{3\times 3\cdot(n^{\text{m}}+n^{\text{s}})}$ is assembled as

$$[\mathbf{K}_{\text{x}_k^{\text{cell}}}] \equiv [\,\mathbf{K}_{\text{x}_k^{\text{cell}}\mathcal{M}} \mid \mathbf{K}_{\text{x}_k^{\text{cell}}\mathcal{S}}\,]\,, \tag{A.160}$$

with the block $[\mathbf{K}_{\text{x}_k^{\text{cell}}\mathcal{M}}] \in \mathbb{R}^{3\times 3\cdot n^{\text{m}}}$, related with mortar degrees of freedom, being defined as

$$[\mathbf{K}_{\text{x}_k^{\text{cell}}\mathcal{M}}] \equiv \chi_{\mathcal{M}}\left([\mathbf{I}] - \{\boldsymbol{\eta}_\odot^{\text{h}}\}\{\boldsymbol{\eta}_\odot^{\text{h}}\}^{\text{T}}\right)[\breve{\mathbf{I}}_k^{\text{s/m}}]\,. \tag{A.161}$$

The block $[\mathbf{K}_{\text{x}_k^{\text{cell}}\mathcal{S}}] \in \mathbb{R}^{3\times 3\cdot n^{\text{s}}}$, related non-mortar degrees of freedom, follows as

$$\begin{aligned}[\mathbf{K}_{\text{x}_k^{\text{cell}}\mathcal{S}}] \equiv{}& \chi_{\mathcal{M}}\left([\mathbf{I}] - \{\boldsymbol{\eta}_\odot^{\text{h}}\}\{\boldsymbol{\eta}_\odot^{\text{h}}\}^{\text{T}}\right)[\breve{\mathbf{I}}_k^{\text{s/m}}] - \{\boldsymbol{\eta}_\odot^{\text{h}}\}\{\boldsymbol{\eta}_\odot^{\text{h}}\}^{\text{T}}[\mathbf{N}_{\text{e}}^{\text{s}}(\boldsymbol{\xi}_\odot^{\text{s}})] \\ &+ \left(\{\boldsymbol{\eta}_\odot^{\text{h}}\}^{\text{T}}\{\mathbf{m}\} + \{\boldsymbol{\eta}_\odot^{\text{h}}\}\{\mathbf{m}\}^{\text{T}}\right)[\mathbf{K}_{\breve{\eta}_\odot^{\text{h}}}]\,.\end{aligned} \tag{A.162}$$

In the expressions above, the characteristic function $\chi_{\mathcal{M}}$ is used to differentiate the type of node being projected, such that

$$\chi_{\mathcal{M}} \equiv \begin{cases} 1\,, & \text{if } k \in \mathcal{M}\,, \\ 0\,, & \text{if } k \in \mathcal{S}\,. \end{cases} \tag{A.163}$$

This means that if the node being projected is non-mortar, i.e., $k \in \mathcal{S}$ and thus $\chi_{\mathcal{M}} = 0$, one have that the auxiliary matrix $[\breve{\mathbf{I}}_k^{\text{s/m}}] \Rightarrow [\breve{\mathbf{I}}_k^{\text{s}}]$ is assembled as in Equation (A.42) and, consequently, derivative matrix $[\mathbf{K}_{\text{x}_k^{\text{cell}}}]$ contains only terms related to non-mortar degrees of freedom. On the other hand, if $k \in \mathcal{M}$, it follows that $\chi_{\mathcal{M}} = 1$, the auxiliary matrix $[\breve{\mathbf{I}}_k^{\text{s/m}}] \Rightarrow [\breve{\mathbf{I}}_k^{\text{m}}]$ is assembled as in Equation (A.61) and, consequently, the matrix $[\mathbf{K}_{\text{x}_k^{\text{cell}}}]$ contains terms related to both mortar and non-mortar degrees of freedom.

A.2.3.3 Case 2: Node generated by the line clipping algorithm

The second case to be analysed relates to the nodes generated by the intersection of lines within the polygon-clipping algorithm. In this work, the Sutherland-Hodgman algorithm (Sutherland and Hodgman, 1974) is employed, which is based on a divide-and-conquer strategy (e.g., refer to Foley *et al.* [1996] for additional details on the computational implementation of the algorithm). Of particular importance for the discussion is the fact the intersection of edges is carried out by the parametric Cyrus-Beck algorithm (Cyrus and Beck, 1978), based on which the intersection point $\mathbf{x}_k^{\text{cell}}$ can be written as

$$\mathbf{x}_k^{\text{cell}} = \hat{\mathbf{x}}_{\text{a}}^{\text{s}} + \mathsf{t}\left(\hat{\mathbf{x}}_{\text{b}}^{\text{s}} - \hat{\mathbf{x}}_{\text{a}}^{\text{s}}\right)\,. \tag{A.164}$$

The constant $\mathfrak{t}$, associated with the parametrization of the non-mortar edge, follows as

$$\mathfrak{t} = \frac{\left(\hat{\mathbf{x}}^{\mathrm{m}}_{\mathrm{a}} - \hat{\mathbf{x}}^{\mathrm{s}}_{\mathrm{a}}\right) \times \left(\hat{\mathbf{x}}^{\mathrm{m}}_{\mathrm{b}} - \hat{\mathbf{x}}^{\mathrm{m}}_{\mathrm{a}}\right) \cdot \boldsymbol{\eta}^{\mathrm{h}}_{\odot}}{\left(\hat{\mathbf{x}}^{\mathrm{s}}_{\mathrm{b}} - \hat{\mathbf{x}}^{\mathrm{s}}_{\mathrm{a}}\right) \times \left(\hat{\mathbf{x}}^{\mathrm{m}}_{\mathrm{b}} - \hat{\mathbf{x}}^{\mathrm{m}}_{\mathrm{a}}\right) \cdot \boldsymbol{\eta}^{\mathrm{h}}_{\odot}} \,. \tag{A.165}$$

As represented in Figure A.4, $\hat{\mathbf{x}}^{\mathrm{s/m}}_{\mathrm{a}}$ and $\hat{\mathbf{x}}^{\mathrm{s/m}}_{\mathrm{b}}$ are the starting and end points, accordingly, of the segments connecting the projected mortar/non-mortar nodes. It should be noticed that Equation (A.164) is no longer valid if the two edges become parallel with each other and, consequently, the denominator becomes zero. Nonetheless, such particular cases are adequately detected and treated by the line clipping algorithm, i.e., it reduces to Case 1 treated above.

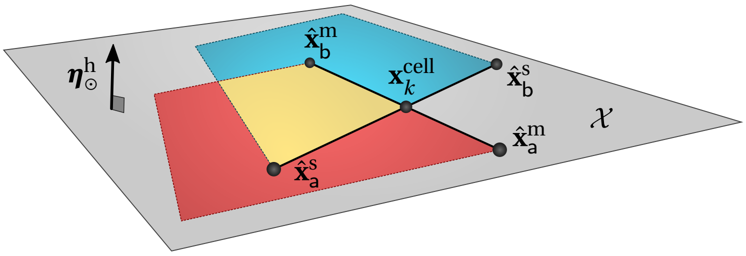


Figure A.5: Integration cell vertex originating from the line clipping algorithm.

The linearisation of the expression above, although quite laborious, is relatively straightforward. Thus, in order to compact the notation, it reveals convenient to introduce some abbreviations. Firstly, the mortar/non-mortar projected edge $\hat{\mathbf{e}}^{\mathrm{s/m}}$ is defined as

$$\hat{\mathbf{e}}^{\mathrm{s/m}} \equiv \hat{\mathbf{x}}^{\mathrm{s/m}}_{\mathrm{b}} - \hat{\mathbf{x}}^{\mathrm{s/m}}_{\mathrm{a}} \,, \tag{A.166}$$

and the vector $\hat{\mathbf{m}}_{\mathrm{a}}$, connecting the first nodes of the non-mortar and mortar projected edges, i.e.,

$$\hat{\mathbf{m}}_{\mathrm{a}} \equiv \hat{\mathbf{x}}^{\mathrm{m}}_{\mathrm{a}} - \hat{\mathbf{x}}^{\mathrm{s}}_{\mathrm{a}} \,. \tag{A.167}$$

These allow one to rewrite Equations (A.164) and (A.165) as

$$\mathbf{x}^{\mathrm{cell}}_{k} = \hat{\mathbf{x}}^{\mathrm{s}}_{\mathrm{a}} + \mathfrak{t}\hat{\mathbf{e}}^{\mathrm{s}} \,, \tag{A.168}$$

and

$$\mathfrak{t} = \frac{\hat{\mathbf{m}}_{\mathrm{a}} \times \hat{\mathbf{e}}^{\mathrm{m}} \cdot \boldsymbol{\eta}^{\mathrm{h}}_{\odot}}{\hat{\mathbf{e}}^{\mathrm{s}} \times \hat{\mathbf{e}}^{\mathrm{m}} \cdot \boldsymbol{\eta}^{\mathrm{h}}_{\odot}} \,, \tag{A.169}$$

respectively. The linearisation of the intersection point $\mathbf{x}^{\mathrm{cell}}_{k}$ yields

$$\mathscr{D}\mathbf{x}^{\mathrm{cell}}_{k} = \mathscr{D}\hat{\mathbf{x}}^{\mathrm{s}}_{\mathrm{a}} + \mathscr{D}\mathfrak{t}\,\hat{\mathbf{e}}^{\mathrm{s}} + \mathfrak{t}\,\mathscr{D}\hat{\mathbf{e}}^{\mathrm{s}} \tag{A.170}$$

with the derivative of the parameter $\mathfrak{t}$ coming as

$$\begin{aligned} \mathscr{D}\mathfrak{t} = \frac{1}{\left(\hat{\mathbf{e}}^{\mathrm{s}} \times \hat{\mathbf{e}}^{\mathrm{m}} \cdot \boldsymbol{\eta}^{\mathrm{h}}_{\odot}\right)^{2}} \Big\{ & \left[\hat{\mathbf{e}}^{\mathrm{s}} \times \hat{\mathbf{e}}^{\mathrm{m}} \cdot \boldsymbol{\eta}^{\mathrm{h}}_{\odot}\right] \left[\hat{\mathbf{e}}^{\mathrm{m}} \times \boldsymbol{\eta}^{\mathrm{h}}_{\odot} \cdot \mathscr{D}\hat{\mathbf{m}}_{\mathrm{a}} + \boldsymbol{\eta}^{\mathrm{h}}_{\odot} \times \hat{\mathbf{m}}_{\mathrm{a}} \cdot \mathscr{D}\hat{\mathbf{e}}^{\mathrm{m}} + \hat{\mathbf{m}}_{\mathrm{a}} \times \hat{\mathbf{e}}^{\mathrm{m}} \cdot \mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\right] \\ & - \left[\hat{\mathbf{m}}_{\mathrm{a}} \times \hat{\mathbf{e}}^{\mathrm{m}} \cdot \boldsymbol{\eta}^{\mathrm{h}}_{\odot}\right] \left[\hat{\mathbf{e}}^{\mathrm{m}} \times \boldsymbol{\eta}^{\mathrm{h}}_{\odot} \cdot \mathscr{D}\hat{\mathbf{e}}^{\mathrm{s}} + \boldsymbol{\eta}^{\mathrm{h}}_{\odot} \times \hat{\mathbf{e}}^{\mathrm{s}} \cdot \mathscr{D}\hat{\mathbf{e}}^{\mathrm{m}} + \hat{\mathbf{e}}^{\mathrm{s}} \times \hat{\mathbf{e}}^{\mathrm{m}} \cdot \mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\right] \Big\} \,. \end{aligned} \tag{A.171}$$

Here, the directional derivatives of the abbreviations are given by

$$\mathscr{D}\hat{\boldsymbol{e}}^{\mathrm{s/m}} = \mathscr{D}\hat{\mathbf{x}}_{\mathsf{b}}^{\mathrm{s/m}} - \mathscr{D}\hat{\mathbf{x}}_{\mathsf{a}}^{\mathrm{s/m}} , \tag{A.172}$$

and

$$\mathscr{D}\hat{\mathbf{m}}_{\mathsf{a}} = \mathscr{D}\hat{\mathbf{x}}_{\mathsf{a}}^{\mathrm{m}} - \mathscr{D}\hat{\mathbf{x}}_{\mathsf{a}}^{\mathrm{s}} . \tag{A.173}$$

Note that all these directional derivatives, since they are projected nodes, can be obtained based on the results derived in Case 1.

A.2.3.4 Algebraic representation

The integration cell vertex $\{\mathbf{x}_k^{\mathrm{cell}}\} \in \mathbb{R}^3$ can be simply written as

$$\{\mathbf{x}_k^{\mathrm{cell}}\} = \{\hat{\mathbf{x}}_{\mathsf{a}}^{\mathrm{s}}\} + \mathfrak{t}\,\{\hat{\boldsymbol{e}}^{\mathrm{s}}\} , \tag{A.174}$$

with the parameter $\mathfrak{t}$ computed as

$$\mathfrak{t} = \frac{\{\boldsymbol{\eta}_{\odot}^{\mathrm{h}}\}^{\mathrm{T}}[\hat{\mathbf{m}}_{\mathsf{a}}]_{\times}\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}}{\{\boldsymbol{\eta}_{\odot}^{\mathrm{h}}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{s}}]_{\times}\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}} . \tag{A.175}$$

The algebraic representation for the abbreviations is also clear, with the vector $\{\hat{\boldsymbol{e}}^{\mathrm{s/m}}\} \in \mathbb{R}^3$ following as

$$\{\hat{\boldsymbol{e}}^{\mathrm{s/m}}\} = \{\hat{\mathbf{x}}_{\mathsf{b}}^{\mathrm{s/m}}\} - \{\hat{\mathbf{x}}_{\mathsf{a}}^{\mathrm{s/m}}\} , \tag{A.176}$$

and the vector $\{\hat{\mathbf{m}}_{\mathsf{a}}\} \in \mathbb{R}^d$ as

$$\{\hat{\mathbf{m}}_{\mathsf{a}}\} = \{\hat{\mathbf{x}}_{\mathsf{a}}^{\mathrm{m}}\} - \{\hat{\mathbf{x}}_{\mathsf{a}}^{\mathrm{s}}\} . \tag{A.177}$$

Before continuing to the derivation of the algebraic representation for the linearisation of the integration cell vertex, the derivatives of the abbreviations are firstly presented. It should be noticed that, since both abbreviations have similar definitions, i.e., they are based on the difference between projected nodes, their linearisation is straightforward and can be carried out by using the expressions obtained for Case 1 treated before. Firstly, the linearisation of the projected mortar edge $\{\mathscr{D}\hat{\boldsymbol{e}}^{\mathrm{m}}\} \in \mathbb{R}^3$ is obtained from

$$\{\mathscr{D}\hat{\boldsymbol{e}}^{\mathrm{m}}\} = \left[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{M}} \mid \mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{S}}\right] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.178}$$

and, whereas the matrix $\left[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{M}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{m}}}$ relates with mortar degrees of freedom,

$$\left[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{M}}\right] = \left[\mathbf{K}_{\hat{\mathrm{x}}_{\mathsf{b}}^{\mathrm{m}}\mathcal{M}}\right] - \left[\mathbf{K}_{\hat{\mathrm{x}}_{\mathsf{a}}^{\mathrm{m}}\mathcal{M}}\right] , \tag{A.179}$$

the matrix $\left[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{S}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ relates with non-mortar terms,

$$\left[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{S}}\right] = \left[\mathbf{K}_{\hat{\mathrm{x}}_{\mathsf{b}}^{\mathrm{m}}\mathcal{S}}\right] - \left[\mathbf{K}_{\hat{\mathrm{x}}_{\mathsf{a}}^{\mathrm{m}}\mathcal{S}}\right] . \tag{A.180}$$

The directional derivative of the projected non-mortar edge, $\{\mathscr{D}\hat{\boldsymbol{e}}^{\mathrm{s}}\} \in \mathbb{R}^3$, can be written in a more compact way, as it contains only terms associated with non-mortar degrees of freedom, i.e.,

$$\{\mathscr{D}\hat{\boldsymbol{e}}^{\mathrm{s}}\} = \left[\mathbf{K}_{\hat{e}^{\mathrm{s}}}\right]\{\Delta\mathbf{d}_{\mathcal{S}}\} . \tag{A.181}$$

Here, the derivative matrix $\left[\mathbf{K}_{\hat{e}^{\mathrm{s}}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ reads

$$\left[\mathbf{K}_{\hat{e}^{\mathrm{s}}}\right] = \left[\mathbf{K}_{\hat{\mathrm{x}}^{\mathrm{s}}_{\mathrm{b}}}\right] - \left[\mathbf{K}_{\hat{\mathrm{x}}^{\mathrm{s}}_{\mathrm{a}}}\right] . \tag{A.182}$$

As mentioned, all the directional derivatives of the projected nodes associated with each edge can be obtained from Equations (A.161) and (A.162). Lastly, the linearisation of the auxiliary vector $\{\hat{\mathbf{m}}_{\mathrm{a}}\}$ connecting the first nodes of the mortar/non-mortar edges is represented by the vector $\{\mathscr{D}\hat{\mathbf{m}}_{\mathrm{a}}\} \in \mathbb{R}^3$, which can be obtained from

$$\{\mathscr{D}\hat{\mathbf{m}}_{\mathrm{a}}\} = \left[\mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{M}} \mid \mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{S}}\right] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} . \tag{A.183}$$

Here, and similarly to the linearisation of the projected edges vector, the matrix $\left[\mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{M}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{m}}}$ is defined as

$$\left[\mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{M}}\right] = \left[\mathbf{K}_{\hat{\mathrm{x}}^{\mathrm{m}}_{\mathrm{a}}\mathcal{M}}\right] - \left[\mathbf{K}_{\hat{\mathrm{x}}^{\mathrm{s}}_{\mathrm{a}}\mathcal{M}}\right] \tag{A.184}$$

and the matrix $\left[\mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{S}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ as

$$\left[\mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{S}}\right] = \left[\mathbf{K}_{\hat{\mathrm{x}}^{\mathrm{m}}_{\mathrm{a}}\mathcal{S}}\right] - \left[\mathbf{K}_{\hat{\mathrm{x}}^{\mathrm{s}}_{\mathrm{a}}\mathcal{S}}\right] . \tag{A.185}$$

Finally, the linearisation of the intersection point is represented by the vector $\{\mathscr{D}\mathbf{x}^{\mathrm{cell}}_k\} \in \mathbb{R}^3$, which is defined as

$$\{\mathscr{D}\mathbf{x}^{\mathrm{cell}}_k\} = \left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k}\right] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.186}$$

where the matrix $\left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k}\right] \in \mathbb{R}^{3\times 3\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ follows the structure

$$\left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k}\right] \equiv \left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k\mathcal{M}} \mid \mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k\mathcal{S}}\right] . \tag{A.187}$$

Here, the derivatives matrix $\left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k\mathcal{M}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{m}}}$ follows as

$$\left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k\mathcal{M}}\right] \equiv \{\hat{\boldsymbol{e}}^{\mathrm{s}}\}\{\mathbf{K}_{\mathsf{t}\mathcal{M}}\} , \tag{A.188}$$

and the matrix $\left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k\mathcal{S}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$ results in

$$\left[\mathbf{K}_{\mathrm{x}^{\mathrm{cell}}_k\mathcal{S}}\right] \equiv \left[\mathbf{K}_{\hat{\mathrm{x}}^{\mathrm{s}}_{\mathrm{a}}}\right] + \{\hat{\boldsymbol{e}}^{\mathrm{s}}\}\{\mathbf{K}_{\mathsf{t}\mathcal{S}}\} + \mathsf{t}\left[\mathbf{K}_{\hat{e}^{\mathrm{s}}}\right] . \tag{A.189}$$

The last term still to be defined is the linearisation of the parameter t, which follows as

$$\mathscr{D}\mathsf{t} = \{\mathbf{K}_{\mathsf{t}}\} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} . \tag{A.190}$$

Here, the matrix $\{\mathbf{K}_{\mathsf{t}}\} \in \mathbb{R}^{3\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ is assembled as

$$\{\mathbf{K}_{\mathsf{t}}\} \equiv \{\mathbf{K}_{\mathsf{t}\mathcal{M}} , \mathbf{K}_{\mathsf{t}\mathcal{S}}\} , \tag{A.191}$$

with the matrix $\{\mathbf{K}_{\mathsf{t}\mathcal{M}}\} \in \mathbb{R}^{3\cdot n^{\mathrm{m}}}$, related with non-mortar degrees of freedom, being computed as

$$\begin{aligned} \{\mathbf{K}_{\mathsf{t}\mathcal{M}}\} \equiv \mathrm{C}_{\mathsf{t}} \Big[& \Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{s}}]_{\times}\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}\Big)\Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{m}}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{M}}] + \{\hat{\mathbf{m}}_{\mathrm{a}}\}^{\mathrm{T}}[\boldsymbol{\eta}^{\mathrm{h}}_{\odot}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{M}}]\Big) \\ & - \Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\mathbf{m}}_{\mathrm{a}}]_{\times}\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}\Big)\Big(\{\hat{\boldsymbol{e}}^{\mathrm{s}}\}^{\mathrm{T}}[\boldsymbol{\eta}^{\mathrm{h}}_{\odot}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{M}}]\Big)\Big] . \end{aligned} \tag{A.192}$$

The matrix $\{\mathbf{K}_{\mathrm{t}\mathcal{S}}\} \in \mathbb{R}^{3\cdot n^{\mathrm{s}}}$ relates with mortar degrees of freedom and can be computed as follows:

$$\begin{aligned}\{\mathbf{K}_{\mathrm{t}\mathcal{S}}\} \equiv \mathrm{C}_{\mathrm{t}}\Big[&\Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{s}}]_{\times}\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}\Big)\Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{m}}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\hat{\mathrm{m}}_{\mathrm{a}}\mathcal{S}}]+\{\hat{\mathbf{m}}_{\mathrm{a}}\}^{\mathrm{T}}[\boldsymbol{\eta}^{\mathrm{h}}_{\odot}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{S}}]+\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}^{\mathrm{T}}[\hat{\mathbf{m}}_{\mathrm{a}}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\eta^{\mathrm{h}}_{\odot}}]\Big)\\ &-\Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\mathbf{m}}_{\mathrm{a}}]_{\times}\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}\Big)\Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{m}}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\hat{e}^{\mathrm{s}}}]+\{\hat{\boldsymbol{e}}^{\mathrm{s}}\}^{\mathrm{T}}[\boldsymbol{\eta}^{\mathrm{h}}_{\odot}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\hat{e}^{\mathrm{m}}\mathcal{S}}]+\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{s}}]^{\mathrm{T}}_{\times}[\mathbf{K}_{\eta^{\mathrm{h}}_{\odot}}]\Big)\Big]\,.\end{aligned} \tag{A.193}$$

The auxiliary constant C_{t} is defined as

$$\mathrm{C}_{\mathrm{t}} \equiv \frac{1}{\Big(\{\boldsymbol{\eta}^{\mathrm{h}}_{\odot}\}^{\mathrm{T}}[\hat{\boldsymbol{e}}^{\mathrm{s}}]_{\times}\{\hat{\boldsymbol{e}}^{\mathrm{m}}\}\Big)^{2}}\,. \tag{A.194}$$

A.2.4 Integration cell Gauss points

As explained in Section 3.7, on each integration cell, a standard triangular FE interpolation is employed by introducing the integration cell parameter space $\boldsymbol{\zeta}$, viz.

$$\boldsymbol{\zeta} = \big\{(\zeta_1,\zeta_2) \mid \zeta_1 \geq 0\,,\ \zeta_2 \geq 0\,,\ \zeta_1+\zeta_2 \leq 1\big\}\,, \tag{A.195}$$

for triangles and

$$\boldsymbol{\zeta} = \big\{(\zeta_1,\zeta_2) \mid -1 \leq \zeta_1 \leq 1\,,\ -1 \leq \zeta_2 \leq 1\big\}\,, \tag{A.196}$$

for quadrilaterals. In order to approximate the mortar integrals, an appropriate Gauss quadrature is employed on the integration cell domain and, then, the mortar and non-mortar Gauss points coordinates $\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}}$ and $\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}$, respectively, are obtained by projecting the integration cell Gauss points $\boldsymbol{\zeta}_{\mathrm{g}}$ back onto the mortar and non-mortar elements. This is illustrated in Figure A.6 and, in general, the projection requires the solution of a system of three non-linear equations (which can be achieved by a local Newton scheme). Considering the projection of the integration cell Gauss point $\boldsymbol{\zeta}_{\mathrm{g}}$ back onto the non-mortar element the relevant conditions are[‡]

$$\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\,\mathbf{x}^{\mathrm{s}}_{k} - \alpha\boldsymbol{\eta}^{\mathrm{h}}_{\odot} - \mathbf{x}^{\mathrm{cell}}_{\mathrm{g}} = \mathbf{0}\,, \tag{A.197}$$

where the parameter α represents the distance along with the normal direction between the two points and $\mathbf{x}^{\mathrm{cell}}_{\mathrm{g}}\big(\boldsymbol{\zeta}_{\mathrm{g}}\big)$ the integration cell Gauss point global coordinates, obtained by standard FE interpolation, i.e.,

$$\mathbf{x}^{\mathrm{cell}}_{\mathrm{g}}\big(\boldsymbol{\zeta}_{\mathrm{g}}\big) = \sum_{v=1}^{n^{\mathrm{cell}}_{\mathrm{e}}} N^{\mathrm{cell}}_{v}\big(\boldsymbol{\zeta}_{\mathrm{g}}\big)\,\mathbf{x}^{\mathrm{cell}}_{v}\,. \tag{A.198}$$

Here, $n^{\mathrm{cell}}_{\mathrm{e}}$ denotes the total number of nodes of the integration cell and $N^{\mathrm{cell}}_{v}(v = 1,\ldots,n_{\mathrm{cell}})$ stands for the standard linear shape functions in the reference domain.

Remark. *In what follows, for the sake of brevity, only the non-mortar Gauss points coordinates are analysed, given that the mortar Gauss points are absolutely analogous.*

[‡] As in Appendix A.1.3 for the two-dimensional case, the notation is slightly violated here. According to the adopted notation, the exact representation of $\mathbf{x}^{\mathrm{cell}}_{\mathrm{g}}$ would be $\{\boldsymbol{x}^{\mathrm{cell}}_{\mathrm{g}}(\boldsymbol{\zeta}_{\mathrm{g}})\}^{\mathrm{h}}$.

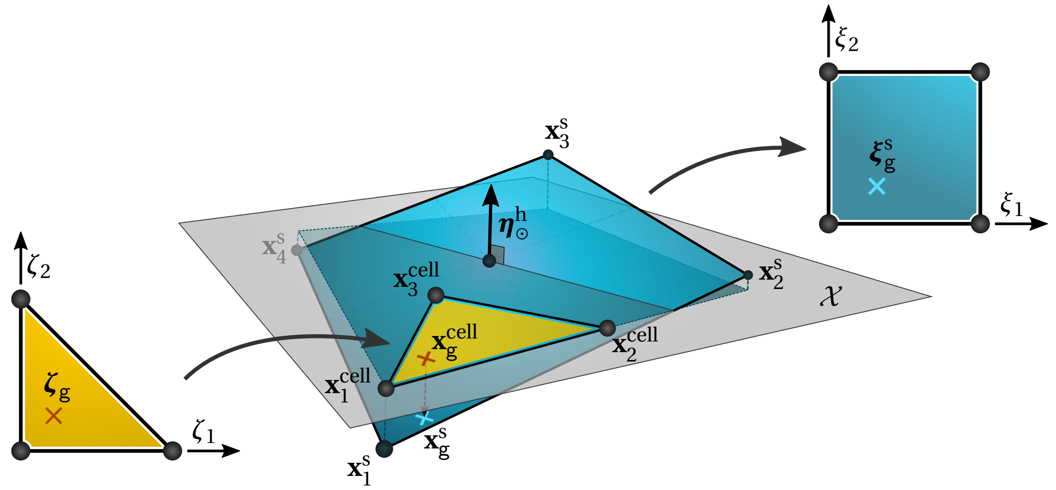


Figure A.6: Projection of Gauss points from the integration cell to the associated non-mortar element.

Linearisation of Equation (A.197) yields

$$\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left[N^{\mathrm{s}}_{k,\xi_1^{\mathrm{s}}}\left(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\right)\mathscr{D}\xi_1^{\mathrm{s}}\,\mathbf{x}_k^{\mathrm{s}}+N^{\mathrm{s}}_{k,\xi_2^{\mathrm{s}}}\left(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\right)\mathscr{D}\xi_2^{\mathrm{s}}\,\mathbf{x}_k^{\mathrm{s}}\right]+\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\left(N_k^{\mathrm{s}}\left(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\right)\Delta\mathbf{x}_k^{\mathrm{s}}\right)$$
$$-\mathscr{D}\alpha\,\boldsymbol{\eta}_{\odot}^{\mathrm{h}}-\alpha\,\mathscr{D}\boldsymbol{\eta}_{\odot}^{\mathrm{h}}-\mathscr{D}\mathbf{x}_{\mathrm{g}}^{\mathrm{cell}}=\mathbf{0}\,. \tag{A.199}$$

which, after being rearranged, allows to expose the directional derivatives of the non-mortar coordinates $\xi^{\mathrm{s}}_{\mathrm{g},1}$ and $\xi^{\mathrm{s}}_{\mathrm{g},2}$, i.e.,

$$\left\{\begin{matrix}\mathscr{D}\xi^{\mathrm{s}}_{\mathrm{g},1}\\ \mathscr{D}\xi^{\mathrm{s}}_{\mathrm{g},2}\\ \mathscr{D}\alpha\end{matrix}\right\}=\mathbf{W}^{-1}\left(\alpha\,\mathscr{D}\boldsymbol{\eta}_{\odot}^{\mathrm{h}}+\mathscr{D}\mathbf{x}_{\mathrm{g}}^{\mathrm{cell}}-\sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}N_k^{\mathrm{s}}\left(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\right)\Delta\mathbf{x}_k^{\mathrm{s}}\right)\,. \tag{A.200}$$

Here, the matrix $\mathbf{W}$ is defined as

$$\mathbf{W}\equiv\left[\;N^{\mathrm{s}}_{k,\xi_1^{\mathrm{s}}}\left(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\right)\mathbf{x}_k^{\mathrm{s}}\;\middle|\;N^{\mathrm{s}}_{k,\xi_2^{\mathrm{s}}}\left(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\right)\mathbf{x}_k^{\mathrm{s}}\;\middle|\;-\boldsymbol{\eta}_{\odot}^{\mathrm{h}}\;\right]\,, \tag{A.201}$$

and the linearisation of the integration cell Gauss point global coordinates results in

$$\mathscr{D}\mathbf{x}_{\mathrm{g}}^{\mathrm{cell}}=\sum_{v=1}^{3}N_v^{\mathrm{cell}}\left(\boldsymbol{\zeta}_{\mathrm{g}}\right)\mathscr{D}\mathbf{x}_v^{\mathrm{cell}}\,. \tag{A.202}$$

Note that all the individual contributions in the expressions above have been treated so far.

A.2.4.1 Algebraic representation

Firstly, the integration cell Gauss point global coordinates $\left\{\mathbf{x}_{\mathrm{g}}^{\mathrm{cell}}\left(\boldsymbol{\zeta}_{\mathrm{g}}\right)\right\}\in\mathbb{R}^3$ follows as

$$\left\{\mathbf{x}_{\mathrm{g}}^{\mathrm{cell}}\left(\boldsymbol{\zeta}_{\mathrm{g}}\right)\right\}=\left[\mathbf{N}_{\mathrm{e}}^{\mathrm{cell}}\left(\boldsymbol{\zeta}_{\mathrm{g}}\right)\right]\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{cell}}\right\}\,, \tag{A.203}$$

where the matrix $\left[\mathbf{N}_{\mathrm{e}}^{\mathrm{cell}}(\boldsymbol{\zeta}_{\mathrm{g}})\right] \in \mathbb{R}^{3\times 3\cdot n_{\mathrm{e}}^{\mathrm{cell}}}$ represents the FE interpolation discrete operator and the vector $\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{cell}}\right\} \in \mathbb{R}^{3\cdot n_{\mathrm{e}}^{\mathrm{cell}}}$ the vector of element coordinates. Its directional derivative is represented by the vector $\left\{\mathscr{D}\mathbf{x}_{\mathrm{g}}^{\mathrm{cell}}(\boldsymbol{\zeta}_{\mathrm{g}})\right\} \in \mathbb{R}^{3}$, which follows as

$$\left\{\mathscr{D}\mathbf{x}_{\mathrm{g}}^{\mathrm{cell}}(\boldsymbol{\zeta}_{\mathrm{g}})\right\} = \left[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{M}} \mid \mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{S}}\right] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.204}$$

where the matrix $\left[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{M}}(\boldsymbol{\zeta}_{\mathrm{g}})\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{m}}}$, related with mortar degrees of freedom, reads

$$\left[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{M}}(\boldsymbol{\zeta}_{\mathrm{g}})\right] \equiv \left[\mathbf{N}_{\mathrm{e}}^{\mathrm{cell}}(\boldsymbol{\zeta}_{\mathrm{g}})\right]\left[\mathbf{K}_{\mathrm{x}_{\mathrm{e}}^{\mathrm{cell}}\mathcal{M}}\right] , \tag{A.205}$$

and the matrix $\left[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{S}}(\boldsymbol{\zeta}_{\mathrm{g}})\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{s}}}$, associated with non-mortar terms, follows as

$$\left[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{S}}(\boldsymbol{\zeta}_{\mathrm{g}})\right] \equiv \left[\mathbf{N}_{\mathrm{e}}^{\mathrm{cell}}(\boldsymbol{\zeta}_{\mathrm{g}})\right]\left[\mathbf{K}_{\mathrm{x}_{\mathrm{e}}^{\mathrm{cell}}\mathcal{S}}\right] . \tag{A.206}$$

Here, the matrices $\left[\mathbf{K}_{\mathrm{x}_{\mathrm{e}}^{\mathrm{cell}}\mathcal{M}}\right] \in \mathbb{R}^{3\times 3\cdot n^{\mathrm{m}}}$ and $\left[\mathbf{K}_{\mathrm{x}_{\mathrm{e}}^{\mathrm{cell}}\mathcal{S}}\right] \in \mathbb{R}^{3\cdot n^{\mathrm{s}}}$ are obtained by assembling the directional derivatives of all the three nodes of the integration segment, i.e.,

$$\left[\mathbf{K}_{\mathrm{x}_{\mathrm{e}}^{\mathrm{cell}}\mathcal{M}}\right] \equiv \begin{bmatrix} \mathbf{K}_{\mathrm{x}_{1}^{\mathrm{cell}}\mathcal{M}} \\ \mathbf{K}_{\mathrm{x}_{2}^{\mathrm{cell}}\mathcal{M}} \\ \mathbf{K}_{\mathrm{x}_{3}^{\mathrm{cell}}\mathcal{M}} \end{bmatrix} , \qquad \left[\mathbf{K}_{\mathrm{x}_{\mathrm{e}}^{\mathrm{cell}}\mathcal{S}}\right] \equiv \begin{bmatrix} \mathbf{K}_{\mathrm{x}_{1}^{\mathrm{cell}}\mathcal{S}} \\ \mathbf{K}_{\mathrm{x}_{2}^{\mathrm{cell}}\mathcal{S}} \\ \mathbf{K}_{\mathrm{x}_{3}^{\mathrm{cell}}\mathcal{S}} \end{bmatrix} . \tag{A.207}$$

Finally, the linearisation of the first non-mortar isoparametric coordinate $\xi_{\mathrm{g},1}^{\mathrm{s}}$ can be written as

$$\mathscr{D}\xi_{\mathrm{g},1}^{\mathrm{s}}(\boldsymbol{\zeta}_{\mathrm{g}}) = \left[\mathbf{K}_{\xi_{\mathrm{g},1}^{\mathrm{s}}\mathcal{M}} \mid \mathbf{K}_{\xi_{\mathrm{g},1}^{\mathrm{s}}\mathcal{S}}\right] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.208}$$

and the second coordinate $\xi_{\mathrm{g},2}^{\mathrm{s}}$ as

$$\mathscr{D}\xi_{\mathrm{g},2}^{\mathrm{s}}(\boldsymbol{\zeta}_{\mathrm{g}}) = \left[\mathbf{K}_{\xi_{\mathrm{g},2}^{\mathrm{s}}\mathcal{M}} \mid \mathbf{K}_{\xi_{\mathrm{g},2}^{\mathrm{s}}\mathcal{S}}\right] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} . \tag{A.209}$$

The directional derivatives related to mortar degrees of freedom can be extracted from the system

$$\begin{Bmatrix} \mathbf{K}_{\xi_{\mathrm{g},1}^{\mathrm{s}}\mathcal{M}} \\ \mathbf{K}_{\xi_{\mathrm{g},2}^{\mathrm{s}}\mathcal{M}} \\ \mathbf{K}_{\alpha\mathcal{M}} \end{Bmatrix} = \mathbf{W}^{-1}\left(\alpha\mathbf{K}_{\eta_{\odot}^{\mathrm{h}}\mathcal{M}} + \mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{M}}\right) , \tag{A.210}$$

whereas the non-mortar terms from

$$\begin{Bmatrix} \mathbf{K}_{\xi_{\mathrm{g},1}^{\mathrm{s}}\mathcal{S}} \\ \mathbf{K}_{\xi_{\mathrm{g},2}^{\mathrm{s}}\mathcal{S}} \\ \mathbf{K}_{\alpha\mathcal{S}} \end{Bmatrix} = \mathbf{W}^{-1}\left(\alpha\mathbf{K}_{\eta_{\odot}^{\mathrm{h}}\mathcal{S}} + \mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{cell}}\mathcal{S}} - \left[\mathbf{N}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\right]\right) . \tag{A.211}$$

The matrix $\left[\mathbf{W}\right] \in \mathbb{R}^{3\times 3}$ follows as

$$\mathbf{W} \equiv \Big[\ \left[\mathbf{B}_{\mathrm{e},\xi_1}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\right]\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{s}}\right\} \ \Big|\ \left[\mathbf{B}_{\mathrm{e},\xi_2}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\right]\left\{\mathbf{x}_{\mathrm{e}}^{\mathrm{s}}\right\} \ \Big|\ -\left\{\boldsymbol{\eta}_{\odot}^{\mathrm{h}}\right\}\ \Big] . \tag{A.212}$$

A.2.5 Integration cell Jacobian determinant

The integration cells Jacobian determinant J^{cell} at a given Gauss point can be written as

$$J^{\text{cell}}(\boldsymbol{\zeta}_{\text{g}}) = \left\| \check{\boldsymbol{\tau}}_1^{\text{cell}} \times \check{\boldsymbol{\tau}}_2^{\text{cell}} \right\| . \tag{A.213}$$

Here, $\check{\boldsymbol{\tau}}_i^{\text{cell}}(\boldsymbol{\zeta}^{\text{cell}})$ represents the *non-unit* tangent vector along the isoparametric coordinate ζ_i $(i = 1, 2)$, which can be obtained as

$$\check{\boldsymbol{\tau}}_i^{\text{cell}}(\boldsymbol{\zeta}^{\text{cell}}) \equiv \sum_{k=1}^{n_{\text{e}}^{\text{cell}}} N_{k,\zeta_i}^{\text{cell}}(\boldsymbol{\zeta}^{\text{cell}})\, \mathbf{x}_k^{\text{cell}} , \tag{A.214}$$

The directional derivative of A.213 results in

$$\mathscr{D} J^{\text{cell}} = \frac{\check{\boldsymbol{\tau}}_1^{\text{cell}} \times \check{\boldsymbol{\tau}}_2^{\text{cell}}}{J^{\text{cell}}} \cdot \left(\mathscr{D}\check{\boldsymbol{\tau}}_1^{\text{cell}} \times \check{\boldsymbol{\tau}}_2^{\text{cell}} + \check{\boldsymbol{\tau}}_1^{\text{cell}} \times \mathscr{D}\check{\boldsymbol{\tau}}_2^{\text{cell}} \right) , \tag{A.215}$$

with the linearisation of each non-unit tangent vector obtained as

$$\mathscr{D}\check{\boldsymbol{\tau}}_i^{\text{cell}} = \sum_{k=1}^{n_{\text{e}}^{\text{cell}}} N_{k,\zeta_i}^{\text{cell}}(\boldsymbol{\zeta}_{\text{g}}^{\text{cell}})\, \mathscr{D}\mathbf{x}_k^{\text{cell}} . \tag{A.216}$$

Recall that the linearisation of the integration cell nodes is described in Appendix A.2.3. It is also noteworthy to mention that for three-noded integration cells, the standard shape functions derivatives yield a constant Jacobian, i.e., it is directly evaluated using the edges of the triangle.

A.2.5.1 Algebraic representation

The integration cell Jacobian determinant can be evaluated as

$$J^{\text{cell}} = \left\| \left[\check{\boldsymbol{\tau}}_1^{\text{cell}} \right]_{\times} \left\{ \check{\boldsymbol{\tau}}_2^{\text{cell}} \right\} \right\| , \tag{A.217}$$

with the non-unit tangent vectors $\left\{ \check{\boldsymbol{\tau}}_i^{\text{cell}} \right\} \in \mathbb{R}^3$ obtained as

$$\left\{ \check{\boldsymbol{\tau}}_i^{\text{cell}} \right\} = \left[\mathbf{B}_{\text{e},\zeta_i}^{\text{cell}}(\boldsymbol{\zeta}_{\text{g}}^{\text{cell}}) \right] \left\{ \mathbf{x}_{\text{e}}^{\text{cell}} \right\} . \tag{A.218}$$

The linearisation of the Jacobian determinant can be written as

$$\mathscr{D} J^{\text{cell}} = \left\{ \mathbf{K}_{J^{\text{cell}}} \right\} \begin{Bmatrix} \Delta \mathbf{d}_{\mathcal{M}} \\ \Delta \mathbf{d}_{\mathcal{S}} \end{Bmatrix} . \tag{A.219}$$

The matrix $\left\{ \mathbf{K}_{J^{\text{cell}}} \right\} \in \mathbb{R}^{3\cdot(n^{\text{m}}+n^{\text{s}})}$ is assembled as

$$\left\{ \mathbf{K}_{J^{\text{cell}}} \right\} \equiv \left\{ \mathbf{K}_{J^{\text{cell}}\mathcal{S}} \mid \mathbf{K}_{J^{\text{cell}}\mathcal{M}} \right\} , \tag{A.220}$$

with the block related with non-mortar degrees of freedom, $\left\{ \mathbf{K}_{J^{\text{cell}}\mathcal{S}} \right\} \in \mathbb{R}^{3\cdot n^{\text{s}}}$, coming as

$$\left\{ \mathbf{K}_{J^{\text{cell}}\mathcal{S}} \right\} \equiv \frac{\left\{ \check{\boldsymbol{\tau}}_2^{\text{cell}} \right\}^{\text{T}} \left[\check{\boldsymbol{\tau}}_1^{\text{cell}} \right]_{\times}^{\text{T}}}{J^{\text{cell}}} \left(\left[\check{\boldsymbol{\tau}}_2^{\text{cell}} \right]_{\times}^{\text{T}} \left[\mathbf{K}_{\check{\tau}_1^{\text{cell}}\mathcal{S}} \right] + \left[\check{\boldsymbol{\tau}}_1^{\text{cell}} \right]_{\times} \left[\mathbf{K}_{\check{\tau}_2^{\text{cell}}\mathcal{S}} \right] \right) , \tag{A.221}$$

and the block $\{\mathbf{K}_{J^{\text{cell}}\mathcal{M}}\} \in \mathbb{R}^{3\cdot n^{\text{m}}}$, related with mortar degrees of freedom, as

$$\{\mathbf{K}_{J^{\text{cell}}\mathcal{M}}\} \equiv \frac{\{\check{\boldsymbol{\tau}}_2^{\text{cell}}\}^{\text{T}}[\check{\boldsymbol{\tau}}_1^{\text{cell}}]_\times^{\text{T}}}{J^{\text{cell}}}\left([\check{\boldsymbol{\tau}}_2^{\text{cell}}]_\times^{\text{T}}[\mathbf{K}_{\check{\tau}_1^{\text{cell}}\mathcal{M}}] + [\check{\boldsymbol{\tau}}_1^{\text{cell}}]_\times[\mathbf{K}_{\check{\tau}_2^{\text{cell}}\mathcal{M}}]\right) . \tag{A.222}$$

In Equations (A.221) and (A.222), the directional derivative of the non-unit tangent vectors is given by

$$[\mathbf{K}_{\check{\tau}_i^{\text{cell}}\mathcal{M}}] = [\mathbf{B}_{,\zeta_i}^{\text{cell}}(\boldsymbol{\zeta}_{\text{g}}^{\text{cell}})][\mathbf{K}_{\text{x}_{\text{e}}^{\text{cell}}\mathcal{M}}] , \tag{A.223}$$

$$[\mathbf{K}_{\check{\tau}_i^{\text{cell}}\mathcal{S}}] = [\mathbf{B}_{,\zeta_i}^{\text{cell}}(\boldsymbol{\zeta}_{\text{g}}^{\text{cell}})][\mathbf{K}_{\text{x}_{\text{e}}^{\text{cell}}\mathcal{S}}] . \tag{A.224}$$

These derivatives are traced back to the type of node of the integration cell, cf. Appendix A.2.3.

A.3 Discretised gap function

Remark. *In what follows, without any loss of generality, contact modelling in three dimensions is assumed. This means that the parameter space for FE interpolation of contact terms is* $\boldsymbol{\xi}^i = (\xi_1^i, \xi_2^i)$*. For the two-dimensional version, it reduces to* ξ^i*.*

The linearisation of the discretised version of the gap function, g^{h}, appears along with the determination of the directional derivative of the weighted gap function $\tilde{g}$. Thus, by considering that at each Gauss point one have[§]

$$g^{\text{h}}(\boldsymbol{\xi}_{\text{g}}^{\text{m}}, \boldsymbol{\xi}_{\text{g}}^{\text{s}}) \equiv -\boldsymbol{\eta}_{\text{g}}^{\text{h}} \cdot \left(\mathbf{x}_{\text{g}}^{\text{s}} - \mathbf{x}_{\text{g}}^{\text{m}}\right) , \tag{A.225}$$

its directional derivative reads

$$\mathscr{D}g^{\text{h}} = -\mathscr{D}\boldsymbol{\eta}_{\text{g}}^{\text{h}} \cdot \left(\mathbf{x}_{\text{g}}^{\text{s}} - \mathbf{x}_{\text{g}}^{\text{m}}\right) - \boldsymbol{\eta}_{\text{g}}^{\text{h}} \cdot \left(\mathscr{D}\mathbf{x}_{\text{g}}^{\text{s}} - \mathscr{D}\mathbf{x}_{\text{g}}^{\text{m}}\right) . \tag{A.226}$$

Examining in more detail the expression above, two additional terms are yet undefined and, therefore, must be derived. First, the Gauss point coordinate on the mortar side $\mathbf{x}_{\text{g}}^{\text{m}}$, which, analogously to the term $\mathbf{x}_{\text{g}}^{\text{s}}$, can be written as

$$\mathbf{x}_{\text{g}}^{\text{m}}(\boldsymbol{\xi}_{\text{g}}^{\text{m}}) \equiv \sum_{l=1}^{n_{\text{e}}^{\text{m}}} N_l^{\text{m}}(\boldsymbol{\xi}_{\text{g}}^{\text{m}})\,\mathbf{x}_l^{\text{m}} , \tag{A.227}$$

and its directional derivative

$$\mathscr{D}\mathbf{x}_{\text{g}}^{\text{m}} = \sum_{l=1}^{n_{\text{e}}^{\text{m}}}\left(N_{l,\xi_1}^{\text{m}}(\boldsymbol{\xi}_{\text{g}}^{\text{m}})\,\mathscr{D}\xi_1^{\text{m}}\,\mathbf{x}_l^{\text{m}} + N_{l,\xi_2}^{\text{m}}(\boldsymbol{\xi}_{\text{g}}^{\text{m}})\,\mathscr{D}\xi_2^{\text{m}}\,\mathbf{x}_l^{\text{m}} + N_l^{\text{m}}(\xi_{\text{g}}^{\text{m}})\Delta\mathbf{x}_l^{\text{m}}\right) . \tag{A.228}$$

Second, the interpolated *unit* normal vector $\boldsymbol{\eta}_{\text{g}}^{\text{h}}$ accounts for a more elaborate treatment and, thus, comes described in detail below.

[§] For simplicity, a minor abuse of notation is permitted again and, in this case, the representation for the interpolated Gauss point position on the mortar side, $\{\boldsymbol{x}_{\text{g}}^{\text{m}}(\boldsymbol{\xi}_{\text{g}}^{\text{m}})\}^{\text{h}}$, is simplified to $\mathbf{x}_{\text{g}}^{\text{m}}$.

A.3.1 The interpolated unit normal vector

Based on the averaged unit normal vector $\tilde{\boldsymbol{\eta}}_j$ discussed in Appendix A.1.1, the continuous field of normals over the non-mortar boundary is obtained through finite element interpolation. Within a given non-mortar element e, the approximation at a Gauss point coordinate $\xi_{\mathrm{g}}^{\mathrm{s}}$ follows as

$$\boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) = \frac{\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}}{\big\|\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}\big\|}\,, \tag{A.229}$$

with $\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)$, already introduced in Equation (A.58), representing the abbreviation for the *non-unit* interpolated normal vector, viz.

$$\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) \equiv \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} N_k^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\tilde{\boldsymbol{\eta}}_k\,. \tag{A.230}$$

The linearisation of the interpolated normal $\boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}$ can be handled identically to the first step of the linearisation of the unit averaged normal $\tilde{\boldsymbol{\eta}}_j$ presented in Appendix A.1.1, i.e.,

$$\mathscr{D}\boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) = \left(\frac{1}{l_{\check{\eta}_{\mathrm{g}}^{\mathrm{h}}}}\boldsymbol{I} - \frac{1}{l_{\check{\eta}_{\mathrm{g}}^{\mathrm{h}}}^{3}}\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}\otimes\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}\right)\mathscr{D}\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}\,, \tag{A.231}$$

where $l_{\check{\eta}_{\mathrm{g}}^{\mathrm{h}}}$ has been introduced as the length of the non-unit interpolated normal vector $\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}$, with directional derivative

$$\mathscr{D}\check{\boldsymbol{\eta}}_{\mathrm{g}}^{\mathrm{h}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big) = \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}}\Big(N_{k,\xi_1}^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\,\mathscr{D}\xi_1^{\mathrm{s}}\,\tilde{\boldsymbol{\eta}}_k + N_{k,\xi_2}^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\,\mathscr{D}\xi_2^{\mathrm{s}}\,\tilde{\boldsymbol{\eta}}_k + N_k^{\mathrm{s}}\big(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}\big)\,\mathscr{D}\tilde{\boldsymbol{\eta}}_k\Big)\,. \tag{A.232}$$

A.3.2 Algebraic representation

Before establishing a suitable algebraic representation for the linearisation of the discretised gap function g^{h}, one must first define the necessary entities to represent the directional derivative of the Gauss point coordinate on the mortar side $\mathbf{x}_{\mathrm{g}}^{\mathrm{m}}$. Thus, by denoting this term by the vector $\{\mathscr{D}\mathbf{x}_{\mathrm{g}}^{\mathrm{m}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} \in \mathbb{R}^d$, it can be written that

$$\{\mathscr{D}\mathbf{x}_{\mathrm{g}}^{\mathrm{m}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} = [\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}}]\begin{Bmatrix}\Delta\mathbf{d}_{\mathcal{M}}\\ \Delta\mathbf{d}_{\mathcal{S}}\end{Bmatrix}\,, \tag{A.233}$$

which, following the line of thought adopted in the remaining deviations, assumes the matrix $[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \in \mathbb{R}^{d\times d\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ assembled as

$$[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \equiv [\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}\mathcal{M}} \mid \mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}\mathcal{S}}]\,. \tag{A.234}$$

The structure of both matrices $[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}\mathcal{M}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \in \mathbb{R}^{d\times n^{\mathrm{m}}\cdot d}$ and $[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}\mathcal{S}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \in \mathbb{R}^{d\times n^{\mathrm{s}}\cdot d}$ is similar to the ones related with the linearisation of the Gauss point on the non-mortar side, viz.

$$[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}\mathcal{M}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \equiv [\mathbf{B}_{\mathrm{e},\xi_1}^{\mathrm{m}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}})]\{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\}\{\mathbf{K}_{\xi_1^{\mathrm{m}}\mathcal{M}}\} + [\mathbf{B}_{\mathrm{e},\xi_2}^{\mathrm{m}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}})]\{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\}\{\mathbf{K}_{\xi_2^{\mathrm{m}}\mathcal{M}}\} + [\mathbf{N}^{\mathrm{m}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}})]\,, \tag{A.235}$$

$$[\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}\mathcal{S}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \equiv [\mathbf{B}_{\mathrm{e},\xi_1}^{\mathrm{m}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}})]\{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\}\{\mathbf{K}_{\xi_1^{\mathrm{m}}\mathcal{S}}\} + [\mathbf{B}_{\mathrm{e},\xi_2}^{\mathrm{m}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}})]\{\mathbf{x}_{\mathrm{e}}^{\mathrm{m}}\}\{\mathbf{K}_{\xi_2^{\mathrm{m}}\mathcal{S}}\}\,. \tag{A.236}$$

Additionally to the Gauss point coordinate on the mortar side, an algebraic representation for the interpolated unit normal and corresponding directional derivative must be defined as well. First, let the vector $\{\boldsymbol{\eta}^{\mathrm{h}}_{\mathrm{g}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \in \mathbb{R}^{d}$ represent the interpolated unit normal vector at a given Gauss point of local coordinate $\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}$, i.e.,

$$\{\boldsymbol{\eta}^{\mathrm{h}}_{\mathrm{g}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \equiv \frac{1}{l_{\check{\eta}^{\mathrm{h}}}}\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\mathrm{g}}\}\,, \tag{A.237}$$

recalling Equation (A.58) for the definition of the non-unit interpolated normal vector $\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\mathrm{g}}\}$, here evaluated at the Gauss position $\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}$ instead. The linearisation of the unit normal vector, $\{\mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{\mathrm{g}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \in \mathbb{R}^{d}$, follows as

$$\{\mathscr{D}\boldsymbol{\eta}^{\mathrm{h}}_{\mathrm{g}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \equiv [\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}}] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix}\,, \tag{A.238}$$

with the matrix $[\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})] \in \mathbb{R}^{d\times d\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ constructed as

$$[\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})] \equiv [\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}\mathcal{M}} \mid \mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}\mathcal{S}}]\,. \tag{A.239}$$

Once again, the terms related with mortar and non-mortar nodes are isolated and the matrix $[\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}\mathcal{M}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})] \in \mathbb{R}^{d\times n^{\mathrm{m}}\cdot d}$ comes as

$$[\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}\mathcal{M}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})] \equiv \left(\frac{1}{l_{\check{\eta}^{\mathrm{h}}_{\mathrm{g}}}}[\mathbf{I}] - \frac{1}{l^{3}_{\check{\eta}^{\mathrm{h}}_{\mathrm{g}}}}[\check{\mathbf{W}}^{\mathrm{h}}_{\mathrm{g}}]\right) \tag{A.240}$$

$$\left([\mathbf{B}^{\mathrm{s}}_{\mathrm{e},\xi_1}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})]\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\}\{\mathbf{K}_{\xi^{\mathrm{s}}_{1}\mathcal{M}}\} + [\mathbf{B}^{\mathrm{s}}_{\mathrm{e},\xi_2}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})]\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\}\{\mathbf{K}_{\xi^{\mathrm{s}}_{2}\mathcal{M}}\}\right)\,, \tag{A.241}$$

and the matrix $[\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}\mathcal{S}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})] \in \mathbb{R}^{d\times n^{\mathrm{s}}\cdot d}$ as

$$[\mathbf{K}_{\eta^{\mathrm{h}}_{\mathrm{g}}\mathcal{S}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})] \equiv \left(\frac{1}{l_{\check{\eta}^{\mathrm{h}}_{\mathrm{g}}}}[\mathbf{I}] - \frac{1}{l^{3}_{\check{\eta}^{\mathrm{h}}_{\mathrm{g}}}}\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\mathrm{g}}\}\{\check{\boldsymbol{\eta}}^{\mathrm{h}}_{\mathrm{g}}\}^{\mathrm{T}}\right) \tag{A.242}$$

$$\left([\mathbf{B}^{\mathrm{s}}_{\mathrm{e},\xi_1}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})]\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\}\{\mathbf{K}_{\xi^{\mathrm{s}}_{1}\mathcal{S}}\} + [\mathbf{B}^{\mathrm{s}}_{\mathrm{e},\xi_2}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})]\{\tilde{\boldsymbol{\eta}}_{\mathrm{e}}\}\{\mathbf{K}_{\xi^{\mathrm{s}}_{2}\mathcal{S}}\} + [\mathbf{N}^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})][\mathbf{K}_{\tilde{\eta}_{\mathrm{e}}}]\right)\,. \tag{A.243}$$

Finally, after introducing all the necessary entities, the linearisation of the discretised gap function can be written as

$$\mathscr{D}g^{\mathrm{h}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}) = \{\mathbf{K}_{g^{\mathrm{h}}}\} \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix}\,, \tag{A.244}$$

with the *row* vector $\{\mathbf{K}_{g^{\mathrm{h}}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \in \mathbb{R}^{d\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ divided into

$$\{\mathbf{K}_{g^{\mathrm{h}}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \equiv \{\mathbf{K}_{g^{\mathrm{h}}\mathcal{M}}\,,\,\mathbf{K}_{g^{\mathrm{h}}\mathcal{S}}\}\,. \tag{A.245}$$

Here both vectors $\{\mathbf{K}_{g^{\mathrm{h}}\mathcal{M}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \in \mathbb{R}^{d\cdot n^{\mathrm{m}}}$ and $\{\mathbf{K}_{g^{\mathrm{h}}\mathcal{S}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \in \mathbb{R}^{d\cdot n^{\mathrm{s}}}$ have similar definitions, with the only difference being the type of nodes involved, namely

$$\{\mathbf{K}_{g^{\mathrm{h}}\mathcal{M}}(\boldsymbol{\xi}^{\mathrm{m}}_{\mathrm{g}},\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}})\} \equiv \{\mathbf{m}_{\mathrm{g}}\}^{\mathrm{T}}[\mathbf{K}_{\eta_{\mathrm{g}}\mathcal{M}}] - \{\boldsymbol{\eta}^{\mathrm{h}}_{\mathrm{g}}\}^{\mathrm{T}}\left([\mathbf{K}_{\mathrm{x}^{\mathrm{s}}_{\mathrm{g}}\mathcal{M}}] - [\mathbf{K}_{\mathrm{x}^{\mathrm{m}}_{\mathrm{g}}\mathcal{M}}]\right) \tag{A.246}$$

and

$$\{\mathbf{K}_{\mathrm{g}^{\mathrm{h}}\mathcal{S}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{m}},\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} \equiv \{\mathbf{m}_{\mathrm{g}}\}^{\mathrm{T}}[\mathbf{K}_{\eta_{\mathrm{g}}\mathcal{S}}] - \{\boldsymbol{\eta}_{\mathrm{g}}^{\mathrm{h}}\}^{\mathrm{T}}\left([\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{s}}\mathcal{S}}] - [\mathbf{K}_{\mathrm{x}_{\mathrm{g}}^{\mathrm{m}}\mathcal{S}}]\right) \tag{A.247}$$

In the equations above, the vector $\{\mathbf{m}_{\mathrm{g}}\}$ is used once again, recalling Equation (A.89) for its calculation.

A.4 Dual shape functions

As discussed in Section 3.2, the dual shape functions Φ_j within a non-mortar element e are constructed as a linear combination of the element's standard shape functions N_k^{s}. Thus, by recalling Equation (3.9), the dual shape function Φ_j at a given Gauss point $\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}(\boldsymbol{\zeta}_{\mathrm{g}})$ is defined as

$$\Phi_j(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}) = \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \mathrm{a}_{jk}^{\Phi} N_k^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\,, \tag{A.248}$$

and its linearisation yields the sum

$$\mathscr{D}\Phi_j(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}) = \sum_{k=1}^{n_{\mathrm{e}}^{\mathrm{s}}} \left(\mathscr{D}\mathrm{a}_{jk}^{\Phi} N_k^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}}) + \mathrm{a}_{jk}^{\Phi} N_{k,\xi_1}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\mathscr{D}\xi_1^{\mathrm{s}} + \mathrm{a}_{jk}^{\Phi} N_{k,\xi_2}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\mathscr{D}\xi_2^{\mathrm{s}}\right). \tag{A.249}$$

At this stage, linearisation of the linear coefficients a_{jk}^{Φ} must be performed, given that they are deformation-dependent for non-constant non-mortar element Jacobian determinants. Within the current formulation, these linear coefficients are grouped in the matrix $[\mathbf{A}_{\mathrm{e}}^{\Phi}] = [\mathrm{a}_{jk}^{\Phi}]$ which, recalling Equation (3.13), is defined as

$$[\mathbf{A}_{\mathrm{e}}^{\Phi}] = [\mathbf{D}_{\mathrm{e}}^{\Phi}][\mathbf{M}_{\mathrm{e}}^{\Phi}]^{-1}\,. \tag{A.250}$$

Thus, linearisation of the expression above results in

$$[\mathscr{D}\mathbf{A}_{\mathrm{e}}^{\Phi}] = [\mathscr{D}\mathbf{D}_{\mathrm{e}}^{\Phi}][\mathbf{M}_{\mathrm{e}}^{\Phi}]^{-1} - [\mathbf{A}_{\mathrm{e}}^{\Phi}][\mathscr{D}\mathbf{M}_{\mathrm{e}}^{\Phi}][\mathbf{M}_{\mathrm{e}}^{\Phi}]^{-1}\,, \tag{A.251}$$

which means that further linearisation of the matrices $[\mathbf{D}_{\mathrm{e}}^{\Phi}]$ and $[\mathbf{M}_{\mathrm{e}}^{\Phi}]$ is required. According to the individual entries given in Equations (3.14) and (3.15), their linearisation yields

$$\mathscr{D}\mathrm{d}_{jk} = \delta_{jk} \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} w_{\mathrm{g}} N_k^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})\mathscr{D}J_{\mathrm{e}}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})\,, \tag{A.252}$$

$$\mathscr{D}\mathrm{m}_{jk} = \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} w_{\mathrm{gp}} N_j^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}}) N_k^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})\mathscr{D}J_{\mathrm{e}}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})\,. \tag{A.253}$$

It is important to stress out that throughout the derivation of the expressions above, the evaluation of the matrices $[\mathbf{D}_{\mathrm{e}}^{\Phi}]$ and $[\mathbf{M}_{\mathrm{e}}^{\Phi}]$ is *independent* of any other Gauss quadrature the resulting dual shape functions Φ_j might be involved. This means the Gauss point coordinates $\xi_{\mathrm{gp}}^{\mathrm{s}}$ in Equations (A.252) and (A.253) are completely unrelated with, for instance, the Gauss coordinates $\xi_{\mathrm{g}}^{\mathrm{s}}$ aforementioned in Equations (A.248) and (A.249). Moreover, for non-mortar elements entirely integrated, their computation does not require the application of any segmentation strategy whatsoever, which means that integration points coordinates $\xi^{\mathrm{s}}_{\mathrm{gp}}$ do not depend on the displacement field and, consequently, their linearisation vanishes. Therefore, the directional derivatives of matrices $[\mathbf{D}^{\Phi}_{\mathrm{e}}]$ and $[\mathbf{M}^{\Phi}_{\mathrm{e}}]$ only require linearisation of the non-mortar element Jacobian, which, following the same line of thought as in Appendix A.1.4, results in

$$\mathscr{D}J^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}}) = \frac{1}{J^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}})}\left(\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k,\xi}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}})\mathbf{x}^{\mathrm{s}}_{k}\right)\left(\sum_{k=1}^{n^{\mathrm{s}}_{\mathrm{e}}} N^{\mathrm{s}}_{k,\xi}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}})\Delta\mathbf{x}^{\mathrm{s}}_{k}\right). \tag{A.254}$$

Once again, since integration Gauss points do not depend on the displacements, when compared with Equation (A.96), this expression becomes simplified.

A.4.1 Algebraic representation

In order to establish an algebraic representation suitable to handle dual shape functions, some element-wise matrices are introduced first. The auxiliary matrix $[\check{\mathbf{D}}^{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}})] = [\check{\mathrm{d}}^{\Phi}_{jk}] \in \mathbb{R}^{n^{\lambda}_{\mathrm{e}} \times n^{\mathrm{s}}_{\mathrm{e}}}$, used to define the matrix $[\mathbf{D}^{\Phi}_{\mathrm{e}}]$, is defined as

$$\check{\mathrm{d}}^{\Phi}_{jk}(\boldsymbol{\xi}^{\mathrm{s}}) \equiv \delta_{jk} N^{\mathrm{s}}_{j}(\boldsymbol{\xi}^{\mathrm{s}}). \tag{A.255}$$

For the particular case where $n^{\lambda}_{\mathrm{e}} = n^{\mathrm{s}}_{\mathrm{e}}$, it yields the trivial diagonal matrix structure

$$[\check{\mathbf{D}}^{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}})] \equiv \begin{bmatrix} N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}}) \end{bmatrix}. \tag{A.256}$$

Analogously, the auxiliary matrix $[\check{\mathbf{M}}^{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}})] = [\check{\mathrm{m}}^{\Phi}_{jk}] \in \mathbb{R}^{n^{\mathrm{s}}_{\mathrm{e}} \times n^{\mathrm{s}}_{\mathrm{e}}}$ comes as

$$\check{\mathrm{m}}^{\Phi}_{jk}(\boldsymbol{\xi}^{\mathrm{s}}) \equiv N^{\mathrm{s}}_{j}(\boldsymbol{\xi}^{\mathrm{s}}) N^{\mathrm{s}}_{k}(\boldsymbol{\xi}^{\mathrm{s}}), \tag{A.257}$$

which yields the square matrix

$$[\check{\mathbf{M}}^{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}})] \equiv \begin{bmatrix} N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}})N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}}) & \cdots & N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}})N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}}) \\ \vdots & \ddots & \vdots \\ N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}})N^{\mathrm{s}}_{1}(\boldsymbol{\xi}^{\mathrm{s}}) & \cdots & N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}})N^{\mathrm{s}}_{n^{\mathrm{s}}_{\mathrm{e}}}(\boldsymbol{\xi}^{\mathrm{s}}) \end{bmatrix}. \tag{A.258}$$

By using these notations, the matrices $[\mathbf{D}^{\Phi}_{\mathrm{e}}]$ and $[\mathbf{M}^{\Phi}_{\mathrm{e}}]$ can be conveniently computed as

$$[\mathbf{D}^{\Phi}_{\mathrm{e}}] = \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} w_{\mathrm{gp}} [\check{\mathbf{D}}^{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}})]\, J^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}}), \tag{A.259}$$

$$[\mathbf{M}^{\Phi}_{\mathrm{e}}] = \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} w_{\mathrm{gp}} [\check{\mathbf{M}}^{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}})]\, J^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}}). \tag{A.260}$$

Recall that the non-mortar element Jacobian $J^{\mathrm{s}}_{\mathrm{e}}(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}})$ is given by Equation (A.98). The directional derivatives of the matrix $[\mathbf{D}^{\Phi}_{\mathrm{e}}]$ and $[\mathbf{M}^{\Phi}_{\mathrm{e}}]$, derived in Equations (A.252) and (A.253), can be constructed as

$$[\mathscr{D}\mathbf{D}_{\mathrm{e}}^{\Phi}] = \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} w_{\mathrm{gp}} [\breve{\mathbf{D}}_{\mathrm{e}}^{\Phi}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})] \left(\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})\}\{\Delta\mathbf{d}_{\mathcal{S}}\} \right) , \tag{A.261}$$

$$[\mathscr{D}\mathbf{M}_{\mathrm{e}}^{\Phi}] = \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} w_{\mathrm{gp}} [\breve{\mathbf{M}}_{\mathrm{e}}^{\Phi}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})] \left(\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})\}\{\Delta\mathbf{d}_{\mathcal{S}}\} \right) , \tag{A.262}$$

where the tangent vector $\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}\}$ has already been introduced in Equation (A.105). At last, using the algebraic notations above, the directional derivative of the coefficients matrix $[\mathbf{A}_{\mathrm{e}}^{\Phi}]$ given in Equation (A.251) is represented by the vector $[\mathscr{D}\mathbf{A}_{\mathrm{e}}^{\Phi}] \in \mathbb{R}^{n_{\mathrm{e}}^{\lambda} \times n_{\mathrm{e}}^{\mathrm{s}}}$, defined as

$$[\mathscr{D}\mathbf{A}_{\mathrm{e}}^{\Phi}] = \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} w_{\mathrm{gp}} \left([\breve{\mathbf{D}}_{\mathrm{e}}^{\Phi}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})][\mathbf{M}_{\mathrm{e}}^{\Phi}]^{-1} - [\mathbf{A}_{\mathrm{e}}^{\Phi}][\breve{\mathbf{M}}_{\mathrm{e}}^{\Phi}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})][\mathbf{M}_{\mathrm{e}}^{\Phi}]^{-1} \right) \left(\{\mathbf{K}_{J_{\mathrm{e}}^{\mathrm{s}}}(\boldsymbol{\xi}_{\mathrm{gp}}^{\mathrm{s}})\}\{\Delta\mathbf{d}_{\mathcal{S}}\} \right) . \tag{A.263}$$

With the directional derivatives of the linear coefficients matrix $[\mathbf{A}_{\mathrm{e}}^{\Phi}]$ at hand, attention is now shifted towards linearisation of the dual shape functions Φ_j. As already explained in Section 3.2, the dual shape functions Φ_j within a non-mortar element can be handled simultaneously by using the vector $\{\boldsymbol{\Phi}_{\mathrm{e}}\}$, defined in Equation (3.10). Thus, by recalling Equation (3.12),

$$\{\boldsymbol{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} = [\mathbf{A}_{\mathrm{e}}^{\Phi}]\{\mathbf{n}_{\mathrm{e}}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} , \tag{A.264}$$

the linearisation of the dual shape functions vector $\{\boldsymbol{\Phi}_{\mathrm{e}}\}$ yields

$$\{\mathscr{D}\boldsymbol{\Phi}_{\mathrm{e}}\} = [\mathscr{D}\mathbf{A}_{\mathrm{e}}^{\Phi}]\{\mathbf{n}_{\mathrm{e}}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} + [\mathbf{A}_{\mathrm{e}}^{\Phi}]\{\mathscr{D}\mathbf{n}_{\mathrm{e}}^{\mathrm{s}}\} . \tag{A.265}$$

Here, the vector $\{\mathscr{D}\mathbf{n}_{\mathrm{e}}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} \in \mathbb{R}^{n_{\mathrm{e}}^{\mathrm{s}}}$, containing the directional derivatives of the standard shape functions vector $\{\mathbf{n}_{\mathrm{e}}^{\mathrm{s}}\}$, is defined as

$$\{\mathscr{D}\mathbf{n}_{\mathrm{e}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} = \left(\{\mathbf{b}_{\mathrm{e},\xi_1}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\}\{\mathbf{K}_{\xi_1^{\mathrm{s}}}\} + \{\mathbf{b}_{\mathrm{e},\xi_2}^{\mathrm{s}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\}\{\mathbf{K}_{\xi_2^{\mathrm{s}}}\} \right) \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} , \tag{A.266}$$

where the vector $\{\mathbf{b}_{\mathrm{e},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\} \in \mathbb{R}^{n_{\mathrm{e}}^{\mathrm{s}}}$ contains the standard shape functions derivatives with respect to the local isoparametric coordinate ξ_i^{s}, i.e.,

$$\{\mathbf{b}_{\mathrm{e},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}})\} \equiv \begin{Bmatrix} N_{1,\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) \\ \vdots \\ N_{n_{\mathrm{e}}^{\mathrm{s}},\xi_i}^{\mathrm{s}}(\boldsymbol{\xi}^{\mathrm{s}}) \end{Bmatrix} . \tag{A.267}$$

Finally, by gathering all the introduced matrix notations, Equation (A.265) can be rewritten as

$$\{\mathscr{D}\boldsymbol{\Phi}_{\mathrm{e}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})\} = [\mathbf{K}_{\Phi_{\mathrm{e}}}] \begin{Bmatrix} \Delta\mathbf{d}_{\mathcal{M}} \\ \Delta\mathbf{d}_{\mathcal{S}} \end{Bmatrix} \tag{A.268}$$

where the tangent matrix $[\mathbf{K}_{\Phi_{\mathrm{e}}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \in \mathbb{R}^{n_{\mathrm{e}}^{\lambda} \times d\cdot(n^{\mathrm{m}}+n^{\mathrm{s}})}$ is constructed as

$$[\mathbf{K}_{\Phi_{\mathrm{e}}}(\boldsymbol{\xi}_{\mathrm{g}}^{\mathrm{s}})] \equiv [\mathbf{K}_{\Phi_{\mathrm{e}}\mathcal{M}} \mid \mathbf{K}_{\Phi_{\mathrm{e}}\mathcal{S}}] . \tag{A.269}$$

Here, the matrix $\left[\mathbf{K}_{\Phi_{\mathrm{e}}\mathcal{S}}\left(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\right)\right] \in \mathbb{R}^{n^{\lambda}_{\mathrm{e}} \times d \cdot n^{\mathrm{s}}}$ contains the terms related with non-mortar degrees of freedom and is defined as

$$\begin{aligned}\left[\mathbf{K}_{\Phi_{\mathrm{e}}\mathcal{S}}\right] \equiv \sum_{\mathrm{gp}=1}^{n_{\mathrm{gp}}} \Big[w_{\mathrm{gp}} \Big(\big[\breve{\mathbf{D}}^{\Phi}_{\mathrm{e}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}}\big)\big] \big[\mathbf{M}^{\Phi}_{\mathrm{e}}\big]^{-1} - \big[\mathbf{A}^{\Phi}_{\mathrm{e}}\big] \big[\breve{\mathbf{M}}^{\Phi}_{\mathrm{e}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}}\big)\big] \big[\mathbf{M}^{\Phi}_{\mathrm{e}}\big]^{-1} \Big) \big\{\mathbf{n}^{\mathrm{s}}_{\mathrm{e}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} \big\{\mathbf{K}_{J^{\mathrm{s}}_{\mathrm{e}}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}}\big)\big\} \Big] \\ + \big[\mathbf{A}^{\Phi}_{\mathrm{e}}\big] \big\{\mathbf{b}^{\mathrm{s}}_{\mathrm{e},\xi_1}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} \big\{\mathbf{K}_{\xi^{\mathrm{s}}_1\mathcal{S}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} + \big[\mathbf{A}^{\Phi}_{\mathrm{e}}\big] \big\{\mathbf{b}^{\mathrm{s}}_{\mathrm{e},\xi_2}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} \big\{\mathbf{K}_{\xi^{\mathrm{s}}_2\mathcal{S}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} .\end{aligned} \tag{A.270}$$

Likewise, the matrix $\left[\mathbf{K}_{\Phi_{\mathrm{e}}\mathcal{M}}\left(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\right)\right] \in \mathbb{R}^{n^{\lambda}_{\mathrm{e}} \times d \cdot n^{\mathrm{m}}}$ relates with mortar degrees of freedom and yields

$$\left[\mathbf{K}_{\Phi_{\mathrm{e}}\mathcal{M}}\right] \equiv \big[\mathbf{A}^{\Phi}_{\mathrm{e}}\big] \big\{\mathbf{b}^{\mathrm{s}}_{\mathrm{e},\xi_1}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} \big\{\mathbf{K}_{\xi^{\mathrm{s}}_1\mathcal{M}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} + \big[\mathbf{A}^{\Phi}_{\mathrm{e}}\big] \big\{\mathbf{b}^{\mathrm{s}}_{\mathrm{e},\xi_2}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} \big\{\mathbf{K}_{\xi^{\mathrm{s}}_2\mathcal{M}}\big(\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}\big)\big\} . \tag{A.271}$$

It is important to note that in Equation (A.270), the distinction between the Gauss coordinate $\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{gp}}$, associated with the numerical integration of the coefficients matrix $\left[\mathbf{A}^{\Phi}_{\mathrm{e}}\right]$, and the Gauss coordinate $\boldsymbol{\xi}^{\mathrm{s}}_{\mathrm{g}}$, emerging from the segmentation scheme used to approximate mortar integrals.

Appendix B

Mapping operations between quadratic finite elements and linear sub-elements

The application of the piecewise interpolation requires the establishment of proper mappings, herein denoted by Λ, from the parent element space, $\boldsymbol{\xi}$, to the sub-element space, $\boldsymbol{\xi}^{\mathrm{sub}}$, viz.

$$\boldsymbol{\xi}^{\mathrm{sub}} = \Lambda(\boldsymbol{\xi}) \, . \tag{B.1}$$

Conversely, the inverse mapping from the sub-element space $\boldsymbol{\xi}^{\mathrm{sub}}$ back to the parent element reads

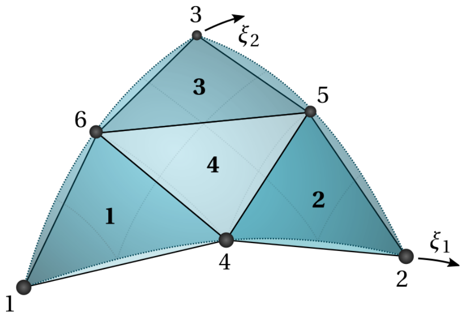
$$\boldsymbol{\xi} = \Lambda^{-1}(\boldsymbol{\xi}^{\mathrm{sub}}) \, . \tag{B.2}$$

These mappings can be derived by employing simple linear transformations between both domains based on geometrical considerations. Table B.1 shows the transformations associated with quadratic 6-noded triangles, followed in Tables B.2 and B.3 by the 8-noded and 9-noded quadrilaterals, respectively. Besides the expressions for the mappings, linearisation of all deformation-dependent terms in mortar integrals requires taking into account the mapping in the derivative chain rule. Therefore, one also needs to derive the Jacobian matrices associated with the mapping operations, which are represented by the matrices $[\boldsymbol{\Xi}] \in \mathbb{R}^{(d-1)\times(d-1)}$. The construction of these matrices is as follows:

$$[\boldsymbol{\Xi}] \equiv \frac{\partial \boldsymbol{\xi}^{\mathrm{sub}}}{\partial \boldsymbol{\xi}^{\mathrm{s}}} = \begin{bmatrix} \xi^{\mathrm{sub}}_{1,\xi_1} & \xi^{\mathrm{sub}}_{1,\xi_1} \\ \xi^{\mathrm{sub}}_{2,\xi_1} & \xi^{\mathrm{sub}}_{2,\xi_2} \end{bmatrix} . \tag{B.3}$$

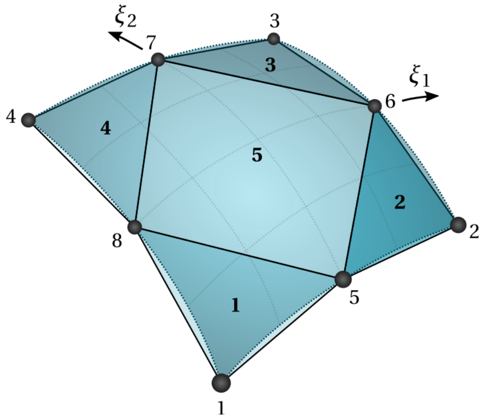
Their inverse read

$$[\boldsymbol{\Xi}]^{-1} = \frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{\xi}^{\mathrm{sub}}} = \begin{bmatrix} \xi_{1,\xi^{\mathrm{sub}}_1} & \xi_{1,\xi^{\mathrm{sub}}_1} \\ \xi_{2,\xi^{\mathrm{sub}}_1} & \xi_{2,\xi^{\mathrm{sub}}_2} \end{bmatrix} . \tag{B.4}$$



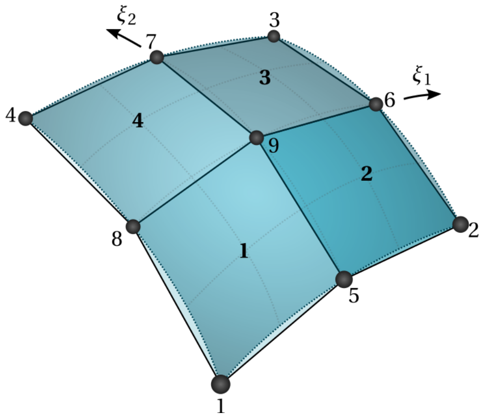
Sub-element	Order	$\Lambda(\boldsymbol{\xi})$	Ξ	$\Lambda^{-1}(\boldsymbol{\xi}^{\text{sub}})$	Ξ^{-1}
1	1, 4, 6	$\xi_1^{\text{sub}} = 2\xi_1$ $\xi_2^{\text{sub}} = 2\xi_2$	$\begin{bmatrix} +2 & 0 \\ 0 & +2 \end{bmatrix}$	$\xi_1 = 1/2\,\xi_1^{\text{sub}}$ $\xi_2 = 1/2\,\xi_2^{\text{sub}}$	$\begin{bmatrix} +1/2 & 0 \\ 0 & +1/2 \end{bmatrix}$
2	4, 2, 5	$\xi_1^{\text{sub}} = 2\xi_1 - 1$ $\xi_2^{\text{sub}} = 2\xi_2$	$\begin{bmatrix} +2 & 0 \\ 0 & +2 \end{bmatrix}$	$\xi_1 = 1/2\,(\xi_1^{\text{sub}} + 1)$ $\xi_2 = 1/2\,\xi_2^{\text{sub}}$	$\begin{bmatrix} +1/2 & 0 \\ 0 & +1/2 \end{bmatrix}$
3	6, 5, 3	$\xi_1^{\text{sub}} = 2\xi_1$ $\xi_2^{\text{sub}} = 2\xi_2 - 1$	$\begin{bmatrix} +2 & 0 \\ 0 & +2 \end{bmatrix}$	$\xi_1 = 1/2\,\xi_1^{\text{sub}}$ $\xi_2 = 1/2\,(\xi_2^{\text{sub}} + 1)$	$\begin{bmatrix} +1/2 & 0 \\ 0 & +1/2 \end{bmatrix}$
4	5, 6, 4	$\xi_1^{\text{sub}} = -2\xi_1 + 1$ $\xi_2^{\text{sub}} = -2\xi_2 + 1$	$\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$	$\xi_1 = 1/2\,(-\xi_1^{\text{sub}} + 1)$ $\xi_2 = 1/2\,(-\xi_2^{\text{sub}} + 1)$	$\begin{bmatrix} -1/2 & 0 \\ 0 & -1/2 \end{bmatrix}$

Table B.1: Mapping between interface elements and the associated sub-elements for the quadratic 6-noded triangle.



Sub-element	Order	$\Lambda(\boldsymbol{\xi})$	Ξ	$\Lambda^{-1}(\boldsymbol{\xi}^{\text{sub}})$	Ξ^{-1}
1	1, 5, 8	$\xi_1^{\text{sub}} = \xi_1 + 1$ $\xi_2^{\text{sub}} = \xi_2 + 1$	$\begin{bmatrix} +1 & 0 \\ 0 & +1 \end{bmatrix}$	$\xi_1 = \xi_1^{\text{sub}} - 1$ $\xi_2 = \xi_2^{\text{sub}} - 1$	$\begin{bmatrix} +1 & 0 \\ 0 & +1 \end{bmatrix}$
2	2, 6, 5	$\xi_1^{\text{sub}} = \xi_2 + 1$ $\xi_2^{\text{sub}} = -\xi_1 + 1$	$\begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}$	$\xi_1 = -\xi_2^{\text{sub}} + 1$ $\xi_2 = \xi_1^{\text{sub}} - 1$	$\begin{bmatrix} 0 & -1 \\ +1 & 0 \end{bmatrix}$
3	3, 7, 6	$\xi_1^{\text{sub}} = -\xi_1 + 1$ $\xi_2^{\text{sub}} = -\xi_2 + 1$	$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\xi_1 = -\xi_1^{\text{sub}} + 1$ $\xi_2 = -\xi_2^{\text{sub}} + 1$	$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$
4	4, 8, 7	$\xi_1^{\text{sub}} = -\xi_2 + 1$ $\xi_2^{\text{sub}} = \xi_1 + 1$	$\begin{bmatrix} 0 & -1 \\ +1 & 0 \end{bmatrix}$	$\xi_1 = \xi_2^{\text{sub}} - 1$ $\xi_2 = -\xi_1^{\text{sub}} + 1$	$\begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}$
5	5, 6, 7, 8	$\xi_1^{\text{sub}} = \xi_2 + \xi_1$ $\xi_1^{\text{sub}} = \xi_2 - \xi_1$	$\begin{bmatrix} +1 & +1 \\ -1 & +1 \end{bmatrix}$	$\xi_1 = {}^1/_2\,(\xi_1^{\text{sub}} - \xi_2^{\text{sub}})$ $\xi_2 = {}^1/_2\,(\xi_1^{\text{sub}} + \xi_2^{\text{sub}})$	$\begin{bmatrix} +{}^1/_2 & -{}^1/_2 \\ +{}^1/_2 & +{}^1/_2 \end{bmatrix}$

Table B.2: Mapping between interface elements and the associated sub-elements for the quadratic 8-noded quadrilateral.



Sub-element	Order	$\Lambda(\boldsymbol{\xi})$	Ξ	$\Lambda^{-1}(\boldsymbol{\xi}^{\text{sub}})$	Ξ^{-1}
1	1, 5, 9, 8	$\xi_1^{\text{sub}} = 2\xi_1 + 1$ $\xi_2^{\text{sub}} = 2\xi_2 + 1$	$\begin{bmatrix} +2 & 0 \\ 0 & +2 \end{bmatrix}$	$\xi_1 = 1/2\,(\xi_1^{\text{sub}} - 1)$ $\xi_2 = 1/2\,(\xi_2^{\text{sub}} - 1)$	$\begin{bmatrix} +1/2 & 0 \\ 0 & +1/2 \end{bmatrix}$
2	5, 2, 6, 9	$\xi_1^{\text{sub}} = 2\xi_2 - 1$ $\xi_2^{\text{sub}} = 2\xi_1 + 1$	$\begin{bmatrix} +2 & 0 \\ 0 & +2 \end{bmatrix}$	$\xi_1 = 1/2\,(\xi_2^{\text{sub}} + 1)$ $\xi_2 = 1/2\,(\xi_1^{\text{sub}} - 1)$	$\begin{bmatrix} +1/2 & 0 \\ 0 & +1/2 \end{bmatrix}$
3	9, 6, 3, 7	$\xi_1^{\text{sub}} = 2\xi_1 - 1$ $\xi_2^{\text{sub}} = 2\xi_2 - 1$	$\begin{bmatrix} +2 & 0 \\ 0 & +2 \end{bmatrix}$	$\xi_1 = 1/2\,(\xi_1^{\text{sub}} + 1)$ $\xi_2 = 1/2\,(\xi_2^{\text{sub}} + 1)$	$\begin{bmatrix} +1/2 & 0 \\ 0 & +1/2 \end{bmatrix}$
4	8, 9, 7, 4	$\xi_1^{\text{sub}} = 2\xi_2 + 1$ $\xi_1^{\text{sub}} = 2\xi_2 - 1$	$\begin{bmatrix} +2 & 0 \\ 0 & +2 \end{bmatrix}$	$\xi_1 = 1/2\,(\xi_1^{\text{sub}} - 1)$ $\xi_2 = 1/2\,(\xi_2^{\text{sub}} + 1)$	$\begin{bmatrix} +1/2 & 0 \\ 0 & +1/2 \end{bmatrix}$

Table B.3: Mapping between interface elements and the associated sub-elements for the quadratic 9-noded quadrilateral.

Appendix C

RCE Boundary Conditions at Edges and Vertices

The type of boundary conditions the RCE is subjected to is relatively standard, referring to references such as Stupkiewicz (2007) for discussions on the topic. However, a small detail that usually is not addressed regards the intersections between boundaries (edges and vertices). The employed solution to avoid overconstraints applies to discretization methods that use regular meshes and is based on master-slave conditions (see Figure C.1).

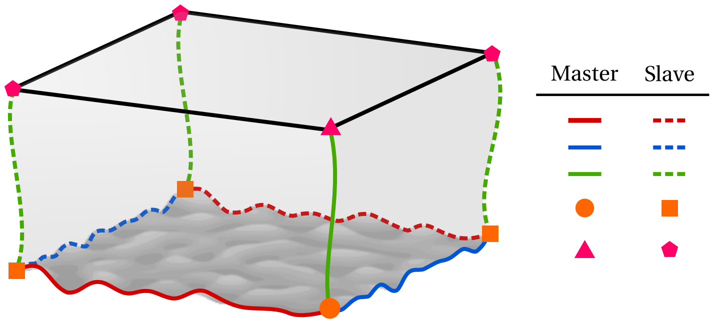


Figure C.1: Master-slave conditions of the RCE.

Starting by the edges at the intersections between periodic and contact boundaries (red and blue lines), the idea is to impose the contact constraints only at the master edges (solid lines), and then setting the matching points at the opposite edge as slaves (dashed lines). The same idea applies to the edges that result from the intersection between the periodic faces of the RCE (green lines), with the only difference being that only one edge is set as master, with the remaining three being slaves.

As for the vertices, two groups can be identified: the vertices at the exterior (top) boundary and the ones at the contact (bottom) boundary. The idea is to set one of them as master—for convenience, the one at the master edge—and the remaining three as slaves. The vertex at the contact boundary satisfies the contact constraints and the one at the exterior boundary a combination of boundary conditions: periodic displacement except along the direction defined by the normal to the exterior boundary, where the displacement is the same as the remaining nodes at the exterior boundary.

Page intentionally left blank

References

Afferrante, L., Carbone, G., and Demelio, G. (2012). “Interacting and Coalescing Hertzian Asperities: A New Multiasperity Contact Model”. In: *Wear* 278–279, pp. 28–33.

Akarapu, S., Sharp, T., and Robbins, M. O. (2011). “Stiffness of Contacts between Rough Surfaces”. In: *Physical Review Letters* 106 (20), p. 204301.

Alart, P. and Curnier, A. (1991). “A Mixed Formulation for Frictional Contact Problems Prone to Newton like Solution Methods”. In: *Computer Methods in Applied Mechanics and Engineering* 92 (3), pp. 353–375.

Anciaux, G. and Molinari, J. F. (2013). “A Molecular Dynamics and Finite Elements Study of Nanoscale Thermal Contact Conductance”. In: *International Journal of Heat and Mass Transfer* 59, pp. 384–392.

Badia, S., Martín, A. F., and Principe, J. (2018). “FEMPAR: An Object-Oriented Parallel Finite Element Framework”. In: *Archives of Computational Methods in Engineering* 25 (2), pp. 195–271.

Bandeira, A. A., Pimenta, P. M., and Wriggers, P. (2008). “A 3D Contact Investigation of Rough Surfaces Considering Elastoplasticity”. In: *Exacta* 6 (1), pp. 109–118.

Bandeira, A. A., Wriggers, P., and de Mattos Pimenta, P. (2004). “Numerical Derivation of Contact Mechanics Interface Laws Using a Finite Approach for Large 3D Deformation”. In: *International Journal for Numerical Methods in Engineering* 59 (2), pp. 173–195.

Bathe, K. J. (2006). *Finite Element Procedures.* New Jersey: Prentice Hall.

Bathe, K. J. and Chaudhary, A. (1985). “A Solution Method for Planar and Axisymmetric Contact Problems”. In: *International Journal for Numerical Methods in Engineering* 21 (1), pp. 65–88.

Bazeley, G. P. *et al.* (1966). “Triangular Elements in Plate Bending - Conforming and Non-Conforming Solutions”. In: *Proceedings of the First Conference on Matrix Methods in Structural Mechanics.* Vol. AFFDL-TR-6, pp. 547–576.

Belgacem, F. B. (1999). “The Mortar Finite Element Method with Lagrange Multipliers”. In: *Numerische Mathematik* 84 (2), pp. 173–197.

Belgacem, F., Hild, P., and Laborde, P. (1998). “The Mortar Finite Element Method for Contact Problems”. In: *Mathematical and Computer Modelling* 28 (4-8), pp. 263–271.

Belhachmi, Z. and Bernardi, C. (1994). “Resolution of Fourth-Order Problems by the Mortar Element Method”. In: *Computer Methods in Applied Mechanics and Engineering* 116 (1-4), pp. 53–58.

Belytschko, T. *et al.* (2013). *Nonlinear Finite Elements for Continua and Structures.* John Wiley & Sons.

Ben Belgacem, F. and Maday, Y. (1994). “A Spectral Element Methodology Tuned to Parallel Implementations”. In: *Computer Methods in Applied Mechanics and Engineering* 116 (1-4), pp. 59–67.

Benson, D. J. and Hallquist, J. O. (1990). “A Single Surface Contact Algorithm for the Post-Buckling Analysis of Shell Structures”. In: *Computer Methods in Applied Mechanics and Engineering* 78 (2), pp. 141–163.

Bernardi, C., Maday, Y., and Patera, A. T. (1993). “Domain Decomposition by the Mortar Element Method”. In: *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters.* Dordrecht: Springer Netherlands, pp. 269–286.

Blum, H. *et al.* (2016). “Semi-Smooth Newton Methods for Mixed FEM Discretizations of Higher-Order for Frictional, Elasto-Plastic Two-Body Contact Problems”. In: *Computer Methods in Applied Mechanics and Engineering* 309, pp. 131–151.

Bonet, J. and Wood, R. D. (2008). *Nonlinear Continuum Mechanics for Finite Element Analysis.* Cambridge University Press.

Bowden, F. P. and Tabor, D. (2001). *The Friction and Lubrication of Solids.* Oxford: Oxford University Press.

Brunssen, S. *et al.* (2007). “A Fast and Robust Iterative Solver for Nonlinear Contact Problems Using a Primal-Dual Active Set Strategy and Algebraic Multigrid”. In: *International Journal for Numerical Methods in Engineering* 69 (3), pp. 524–543.

Bush, A. W., Gibson, R. D., and Thomas, T. R. (1975). “The Elastic Contact of a Rough Surface”. In: *Wear* 35 (1), pp. 87–111.

Campañá, C. and Müser, M. H. (2006). “Practical Green’s Function Approach to the Simulation of Elastic Semi-Infinite Solids”. In: *Physical Review B* 74 (7), p. 075420.

Carbone, G. and Bottiglione, F. (2008). “Asperity Contact Theories: Do They Predict Linearity between Contact Area and Load?” In: *Journal of the Mechanics and Physics of Solids* 56 (8), pp. 2555–2572.

Carbone, G., Lorenz, B., *et al.* (2009). “Contact Mechanics and Rubber Friction for Randomly Rough Surfaces with Anisotropic Statistical Properties”. In: *The European Physical Journal E* 29 (3), pp. 275–284.

Cavalieri, F. J. and Cardona, A. (2013). “Three-Dimensional Numerical Solution for Wear Prediction Using a Mortar Contact Algorithm”. In: *International Journal for Numerical Methods in Engineering* 96 (8), pp. 467–486.

Christensen, P. W. *et al.* (1998). “Formulation and Comparison of Algorithms for Frictional Contact Problems”. In: *International Journal for Numerical Methods in Engineering* 42 (1), pp. 145–173.

Cichosz, T. and Bischoff, M. (2011). “Consistent Treatment of Boundaries with Mortar Contact Formulations Using Dual Lagrange Multipliers”. In: *Computer Methods in Applied Mechanics and Engineering* 200 (9-12), pp. 1317–1332.

Couto Carneiro, A. M., Pinto Carvalho, R., and Andrade Pires, F. M. (2020). “Representative Contact Element Size Determination for Micromechanical Contact Analysis of Self-Affine Topographies”. In: *International Journal of Solids and Structures* 206, pp. 262–281.

Couto Carneiro, A. M. (2019). *Modeling and Analysis of Rough Contact by Computational Homogenization.* Tech. rep. Faculty of Engineering, University of Porto.

Cowper, G. (1973). “Gaussian Quadrature Formulas for Triangles”. In: *International Journal for Numerical Methods in Engineering* 7 (3), pp. 405–408.

Crisfield, M. A. (2000). “Re-Visiting the Contact Patch Test”. In: *International Journal for Numerical Methods in Engineering* 48 (3), pp. 435–449.

Cyrus, M. and Beck, J. (1978). “Generalized Two- and Three-Dimensional Clipping”. In: *Computers & Graphics* 3 (1), pp. 23–28.

De Lorenzis, L., Temizer, İ., *et al.* (2011). “A Large Deformation Frictional Contact Formulation Using NURBS-based Isogeometric Analysis”. In: *International Journal for Numerical Methods in Engineering* (March), n/a–n/a.

De Lorenzis, L. and Wriggers, P. (2013). “Computational Homogenization of Rubber Friction on Rough Rigid Surfaces”. In: *Computational Materials Science* 77, pp. 264–280.

De Lorenzis, L., Wriggers, P., and Hughes, T. J. (2014). “Isogeometric Contact: A Review”. In: *GAMM-Mitteilungen* 37 (1), pp. 85–123.

de Souza Neto, E. A. *et al.* (1996). “Design of Simple Low Order Finite Elements for Large Strain Analysis of Nearly Incompressible Solids”. In: *International Journal of Solids and Structures* 33 (20-22), pp. 3277–3296.

Derjaguin, B. V., Muller, V. M., and Toporov, Y. U. P. (1975). “Effect of Contact Deformation on the Adhesion of Particles.” In: *Journal of colloid and interface science* 53 (2), pp. 314–326.

Dieterich, J. H. and Kilgore, B. D. (1994). “Direct Observation of Frictional Contacts: New Insights for State-Dependent Properties”. In: *Pure and Applied Geophysics PAGEOPH* 143 (1-3), pp. 283–302.

Dittmann, M. *et al.* (2014). “Isogeometric Analysis and Thermomechanical Mortar Contact Problems”. In: *Computer Methods in Applied Mechanics and Engineering* 274, pp. 192–212.

Doca, T. and Andrade Pires, F. (2017). "Finite Element Modeling of Wear Using the Dissipated Energy Method Coupled with a Dual Mortar Contact Formulation". In: *Computers & Structures* 191, pp. 62–79.

Dowson, D. (1991). *History of Tribology*. New York: Wiley.

Euler, L. (1750a). "Sur La Diminution de La Résistance Du Frottement". In: *Mémoires de l'académie des sciences de Berlin* 4, pp. 133–148.

Euler, L. (1750b). "Sur Le Frottement Des Corps Solides". In: *Memoires de l'academie des sciences de Berlin* 4, pp. 122–132.

Farah, P., Gitterle, M., *et al.* (2016). "Computational Wear and Contact Modeling for Fretting Analysis with Isogeometric Dual Mortar Methods". In: *Key Engineering Materials* 681, pp. 1–18.

Farah, P., Popp, A., and Wall, W. A. (2015). "Segment-Based vs. Element-Based Integration for Mortar Methods in Computational Contact Mechanics". In: *Computational Mechanics* 55 (1), pp. 209–228.

Farah, P., Wall, W. A., and Popp, A. (2017). "An Implicit Finite Wear Contact Formulation Based on Dual Mortar Methods". In: *International Journal for Numerical Methods in Engineering* 111 (4), pp. 325–353.

Fischer, K. A. and Wriggers, P. (2005). "Frictionless 2D Contact Formulations for Finite Deformations Based on the Mortar Method". In: *Computational Mechanics* 36 (3), pp. 226–244.

Fischer, K. A. and Wriggers, P. (2006). "Mortar Based Frictional Contact Formulation for Higher Order Interpolations Using the Moving Friction Cone". In: *Computer Methods in Applied Mechanics and Engineering* 195 (37-40), pp. 5020–5036.

Flemisch, B. and Wohlmuth, B. I. (2007). "Stable Lagrange Multipliers for Quadrilateral Meshes of Curved Interfaces in 3D". In: *Computer Methods in Applied Mechanics and Engineering* 196 (8), pp. 1589–1602.

Foley, J. D. *et al.* (1996). *Computer Graphics: Principles and Practice*. Addison-Wesley Professional.

Francavilla, A. and Zienkiewicz, O. C. (1975). "A Note on Numerical Computation of Elastic Contact Problems". In: *International Journal for Numerical Methods in Engineering* 9 (4), pp. 913–924.

Fuller, K. N. G. and Tabor, D. (1975). "The Effect of Surface Roughness on the Adhesion of Elastic Solids". In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 345 (1642), pp. 327–342.

Gamma, E. *et al.* (1995). *Design Patterns: Elements of Reusable Object-Oriented Software*. Boston: Addison-Wesley.

Gao, Y. *et al.* (2006). "The Behavior of an Elastic–Perfectly Plastic Sinusoidal Surface under Contact Loading". In: *Wear* 261 (2), pp. 145–154.

Ghosh, S. and Moorthy, S. (1998). “Particle Fracture Simulation in Non-Uniform Microstructures of Metal–Matrix Composites”. In: *Acta Materialia* 46 (3), pp. 965–982.

Gitterle, M. *et al.* (2010). “Finite Deformation Frictional Mortar Contact Using a Semi-Smooth Newton Method with Consistent Linearization”. In: *International Journal for Numerical Methods in Engineering* 84 (5), pp. 543–571.

Greenwood, J. A. (1966). “Constriction Resistance and the Real Area of Contact”. In: *British Journal of Applied Physics* 17 (12), pp. 1621–1632.

Greenwood, J. A. (2006). “A Simplified Elliptic Model of Rough Surface Contact”. In: *Wear* 261 (2), pp. 191–200.

Greenwood, J. A. and Williamson, J. B. P. (1966). “Contact of Nominally Flat Surfaces”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 295 (1442), pp. 300–319.

Gurtin, M. E., Fried, E., and Anand, L. (2010). *The Mechanics and Thermodynamics of Continua.* Cambridge: Cambridge University Press.

Hager, C. and Wohlmuth, B. I. (2009). “Nonlinear Complementarity Functions for Plasticity Problems with Frictional Contact”. In: *Computer Methods in Applied Mechanics and Engineering* 198 (41), pp. 3411–3427.

Hallquist, J. O. (1979). *NIKE2D: An Implicit, Finite Deformation, Finite Element Code for Analysing of the Static and Dynamic Response of Two-Dimensional Solids.* Tech. rep. California.

Hallquist, J., Goudreau, G., and Benson, D. (1985). “Sliding Interfaces with Contact-Impact in Large-Scale Lagrangian Computations”. In: *Computer Methods in Applied Mechanics and Engineering* 51 (1-3), pp. 107–137.

Haraldsson, A. and Wriggers, P. (2000). “A Strategy for Numerical Testing of Frictional Laws with Application to Contact between Soil and Concrete”. In: *Computer Methods in Applied Mechanics and Engineering* 190 (8-10), pp. 963–977.

Hartmann, S., Brunssen, S., *et al.* (2007). “Unilateral Non-Linear Dynamic Contact of Thin-Walled Structures Using a Primal-Dual Active Set Strategy”. In: *International Journal for Numerical Methods in Engineering* 70 (8), pp. 883–912.

Hartmann, S., Oliver, J., *et al.* (2009). “A Contact Domain Method for Large Deformation Frictional Contact Problems. Part 2: Numerical Aspects”. In: *Computer Methods in Applied Mechanics and Engineering* 198 (33-36), pp. 2607–2631.

Hertz, H. (1881). “Über Die Berührung Fester Elasticher Körper”. In: *J. reine und angewandte Mathematik* 92, pp. 156–171.

Hesch, C. and Betsch, P. (2009). “A Mortar Method for Energy-Momentum Conserving Schemes in Frictionless Dynamic Contact Problems”. In: *International Journal for Numerical Methods in Engineering* 77 (10), pp. 1468–1500.

Hild, P. (2000). “Numerical Implementation of Two Nonconforming Finite Element Methods for Unilateral Contact”. In: *Computer Methods in Applied Mechanics and Engineering* 184 (1), pp. 99–123.

Hill, R. (1963). “Elastic Properties of Reinforced Solids: Some Theoretical Principles”. In: *Journal of the Mechanics and Physics of Solids* 11 (5), pp. 357–372.

Hintermüller, M., Ito, K., and Kunisch, K. (2002). “The Primal-Dual Active Set Strategy as a Semismooth Newton Method”. In: *SIAM Journal on Optimization* 13 (3), pp. 865–888.

Holzapfel, G. A. (2000). *Nonlinear Solid Mechanics: A Continuum Approach for Engineering.* Wiley.

Hu, Y. Z. and Tonder, K. (1992). “Simulation of 3-D Random Rough Surface by 2-D Digital Filter and Fourier Analysis”. In: *International Journal of Machine Tools and Manufacture.* Proceedings of the 5th International Conference on Metrology and Properties of Engineering Surfaces 32 (1), pp. 83–90.

Hüeber, S. and Wohlmuth, B. (2009). “Thermo-Mechanical Contact Problems on Non-Matching Meshes”. In: *Computer Methods in Applied Mechanics and Engineering* 198 (15-16), pp. 1338–1350.

Hüeber, S. (2008). “Discretization Techniques and Efficient Algorithms for Contact Problems”. PhD thesis.

Hüeber, S., Stadler, G., and Wohlmuth, B. I. (2008). “A Primal-Dual Active Set Algorithm for Three-Dimensional Contact Problems with Coulomb Friction”. In: *SIAM Journal on Scientific Computing* 30 (2), pp. 572–596.

Hüeber, S. and Wohlmuth, B. I. (2005). “A Primal–Dual Active Set Strategy for Non-Linear Multibody Contact Problems”. In: *Computer Methods in Applied Mechanics and Engineering* 194 (27-29), pp. 3147–3166.

Hughes, T. J. (2000). *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis.* Dover Publications.

Hughes, T. J., Taylor, R. L., Sackman, J. L., and Kanoknukulchai, W. (1976). *Finite Element Formulation and Solution of Contact-Impact Problems in Continuum Mechanics.* Tech. rep. California.

Hughes, T. J., Taylor, R. L., Sackman, J. L., Curnier, A., *et al.* (1976). “A Finite Element Method for a Class of Contact-Impact Problems”. In: *Computer Methods in Applied Mechanics and Engineering* 8 (3), pp. 249–276.

Hutchings, I. M. (2016). “Leonardo Da Vinci’s Studies of Friction”. In: *Wear* 360–361, pp. 51–66.

Hyun, S. *et al.* (2004). “Finite-Element Analysis of Contact between Elastic Self-Affine Surfaces”. In: *Physical Review E* 70 (2), pp. 026117–026117.

Jacobs, T. D. B., Junge, T., and Pastewka, L. (2017). “Quantitative Characterization of Surface Topography Using Spectral Analysis”. In: *Surface Topography: Metrology and Properties* 5 (1), p. 013001.

Johnson, K. L., Kendall, K., and Roberts, A. D. (1971). “Surface Energy and the Contact of Elastic Solids”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 324 (1558), pp. 301–313.

Johnson, K. L. (1987). *Contact Mechanics.* Cambridge University Press.

Jost, H. P. (1966). *Lubrication: ; Education and Research ; Report on the Present Position and Industry’s Needs (Submitted to the Department of Education and Science by the Lubrication Engineering and Research) Working Group.* London: Her Majesty’s Stationery Office.

Kikuchi, N. and Oden, J. T. (1988). *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods.* SIAM.

Kindo, T. M., Laursen, T. A., and Dolbow, J. E. (2014). “Toward Robust and Accurate Contact Solvers for Large Deformation Applications: A Remapping/Adaptivity Framework for Mortar-Based Methods”. In: *Computational Mechanics* 54 (1), pp. 53–70.

Konter, A. W. A. (2006). *Advanced Finite Element Contact Benchmarks.* Tech. rep. 1-874-376-13-1. Hamilton.

Konyukhov, A. and Schweizerhof, K. (2013). “Computational Contact Mechanics: Geometrically Exact Theory for Arbitrary Shaped Bodies”. In: *Lecture Notes in Applied and Computational Mechanics.* New York: Springer.

Koziara, T. and Bićanić, N. (2008). “Semismooth Newton Method for Frictional Contact between Pseudo-Rigid Bodies”. In: *Computer Methods in Applied Mechanics and Engineering* 197 (33-40), pp. 2763–2777.

Kubiak, K. J., Liskiewicz, T. W., and Mathia, T. G. (2011). “Surface Morphology in Engineering Applications: Influence of Roughness on Sliding and Wear in Dry Fretting”. In: *Tribology International* 44 (11), pp. 1427–1432.

Lamichhane, B. P., Stevenson, R. P., and Wohlmuth, B. I. (2005). “Higher Order Mortar Finite Element Methods in 3D with Dual Lagrange Multiplier Bases”. In: *Numerische Mathematik* 102 (1), pp. 93–121.

Laursen, T. A. (1994). “The Convected Description in Large Deformation Frictional Contact Problems”. In: *International Journal of Solids and Structures* 31 (5), pp. 669–681.

Laursen, T. A. (2013). *Computational Contact and Impact Mechanics: Fundamentals of Modeling Interfacial Phenomena in Nonlinear Finite Element Analysis.* Springer Science & Business Media.

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”. In: *International Journal for Numerical Methods in Engineering* 36 (20), pp. 3451–3485.

Luenberger, D. G. and Ye, Y. (2008). *Linear and Nonlinear Programming.* Springer Science & Business Media.

Manners, W. and Greenwood, J. A. (2006). “Some Observations on Persson’s Diffusion Theory of Elastic Contact”. In: *Wear* 261 (5), pp. 600–610.

Marsden, J. E. and Hughes, T. J. R. (2012). *Mathematical Foundations of Elasticity.* Courier Corporation.

Matouš, K. *et al.* (2017). “A Review of Predictive Nonlinear Theories for Multiscale Modeling of Heterogeneous Materials”. In: *Journal of Computational Physics* 330, pp. 192–220.

McDevitt, T. W. and Laursen, T. A. (2000). “A Mortar-Finite Element Formulation for Frictional Contact Problems”. In: *International Journal for Numerical Methods in Engineering* 48 (10), pp. 1525–1547.

Metcalf, M., Reid, J., and Cohen, M. (2011). *Modern Fortran Explained.* Oxford: Oxford University Press.

Michel, J., Moulinec, H., and Suquet, P. (1999). “Effective Properties of Composite Materials with Periodic Microstructure: A Computational Approach”. In: *Computer Methods in Applied Mechanics and Engineering* 172 (1-4), pp. 109–143.

Mikic, B. B. and Rohsenow, W. M. (1966). *Thermal Contact Resistance.* Technical Report. Cambridge, Mass. : M.I.T. Dept. of Mechanical Engineering, [1966].

Moulinec, H. and Suquet, P. (1998). “A Numerical Method for Computing the Overall Response of Nonlinear Composites with Complex Microstructure”. In: *Computer Methods in Applied Mechanics and Engineering* 157 (1-2), pp. 69–94.

Mróz, Z. and Stupkiewicz, S. (1994). “An Anisotropic Friction and Wear Model”. In: *International Journal of Solids and Structures* 31 (8), pp. 1113–1131.

Neto, E. A. d. S., Peric, D., and Owen, D. R. J. (2011). *Computational Methods for Plasticity: Theory and Applications.* John Wiley & Sons.

Nitsche, R. (2011). “A Multiscale Projection Method for Contact on Rough Surfaces”. PhD thesis. Hannover: Institut für Kontinuumsmechanik.

Ogden, R. W. (1997). *Non-Linear Elastic Deformations.* Courier Corporation.

Oliver, J. *et al.* (2009). “A Contact Domain Method for Large Deformation Frictional Contact Problems. Part 1: Theoretical Basis”. In: *Computer Methods in Applied Mechanics and Engineering* 198 (33-36), pp. 2591–2606.

Paggi, M. and Ciavarella, M. (2010). “The Coefficient of Proportionality κ between Real Contact Area and Load, with New Asperity Models”. In: *Wear* 268 (7), pp. 1020–1029.

Papadopoulos, P. and Taylor, R. L. (1992). “A Mixed Formulation for the Finite Element Solution of Contact Problems”. In: *Computer Methods in Applied Mechanics and Engineering* 94 (3), pp. 373–389.

Pastewka, L. and Robbins, M. O. (2014). “Contact between Rough Surfaces and a Criterion for Macroscopic Adhesion”. In: *Proceedings of the National Academy of Sciences* 111 (9), pp. 3298–3303.

Pei, L. *et al.* (2005). “Finite Element Modeling of Elasto-Plastic Contact between Rough Surfaces”. In: *Journal of the Mechanics and Physics of Solids* 53 (11), pp. 2385–2409.

Pereira, K., Yue, T., and Abdel Wahab, M. (2017). “Multiscale Analysis of the Effect of Roughness on Fretting Wear”. In: *Tribology International* 110 (February), pp. 222–231.

Persson, B. N. J. (2001a). “Elastoplastic Contact between Randomly Rough Surfaces”. In: *Physical Review Letters* 87 (11), p. 116101.

Persson, B. N. J. (2001b). “Theory of Rubber Friction and Contact Mechanics”. In: *The Journal of Chemical Physics* 115 (8), pp. 3840–3861.

Persson, B. N. J. (2014). “On the Fractal Dimension of Rough Surfaces”. In: *Tribology Letters* 54 (1), pp. 99–106.

Persson, B. N. J. *et al.* (2005). “On the Nature of Surface Roughness with Application to Contact Mechanics, Sealing, Rubber Friction and Adhesion”. In: *Journal of Physics: Condensed Matter* 17 (1), R1.

Pinto Carvalho, R., Couto Carneiro, A. M., Andrade Pires, F. M., and Doca, T. (2022). “An Efficient Multiscale Strategy to Predict the Evolution of the Real Contact Area between Rough Surfaces”. In: *Tribology International* 165, p. 107255.

Pinto Carvalho, R., Couto Carneiro, A. M., Andrade Pires, F. M., and Popp, A. (2022). “An Efficient Contact Algorithm for Rigid/Deformable Interaction Based on the Dual Mortar Method”. In: *arXiv:2201.01165 [cs]*.

Pinto Carvalho, R., Rodrigues Lopes, I. A., and Andrade Pires, F. M. (2018). “Prediction of the Yielding Behaviour of Ductile Porous Materials through Computational Homogenization”. In: *Engineering Computations* 35 (2), pp. 604–621.

Polonsky, I. A. and Keer, L. M. (1999). “A Numerical Method for Solving Rough Contact Problems Based on the Multi-Level Multi-Summation and Conjugate Gradient Techniques”. In: *Wear* 231 (2), pp. 206–219.

Popp, A. (2012). “Mortar Methods for Computational Contact Mechanics and General Interface Problems”. PhD thesis.

Popp, A. (2018). “State-of-the-Art Computational Methods for Finite Deformation Contact Modeling of Solids and Structures”. In: *Contact Modeling for Solids and Particles.* Ed. by A. Popp and P. Wriggers. CISM International Centre for Mechanical Sciences. Cham: Springer International Publishing, pp. 1–86.

Popp, A., Gee, M. W., and Wall, W. A. (2009). “A Finite Deformation Mortar Contact Formulation Using a Primal-Dual Active Set Strategy”. In: *International Journal for Numerical Methods in Engineering* 79 (11), pp. 1354–1391.

Popp, A., Gitterle, M., *et al.* (2010). “A Dual Mortar Approach for 3D Finite Deformation Contact with Consistent Linearization”. In: *International Journal for Numerical Methods in Engineering* 83 (11), pp. 1428–1465.

Popp, A., Seitz, A., *et al.* (2013). “Improved Robustness and Consistency of 3D Contact Algorithms Based on a Dual Mortar Approach”. In: *Computer Methods in Applied Mechanics and Engineering* 264, pp. 67–80.

Popp, A., Wohlmuth, B. I., *et al.* (2012). “Dual Quadratic Mortar Finite Element Methods for 3D Finite Deformation Contact”. In: *SIAM Journal on Scientific Computing* 34 (4), B421–B446.

Puso, M. A. (2004). “A 3D Mortar Method for Solid Mechanics”. In: *International Journal for Numerical Methods in Engineering* 59 (3), pp. 315–336.

Puso, M. A. and Laursen, T. A. (2004a). “A Mortar Segment-to-Segment Contact Method for Large Deformation Solid Mechanics”. In: *Computer Methods in Applied Mechanics and Engineering* 193 (6-8), pp. 601–629.

Puso, M. A. and Laursen, T. A. (2004b). “A Mortar Segment-to-Segment Frictional Contact Method for Large Deformations”. In: *Computer Methods in Applied Mechanics and Engineering* 193 (45-47), pp. 4891–4913.

Puso, M. A., Laursen, T. A., and Solberg, J. (2008). “A Segment-to-Segment Mortar Contact Method for Quadratic Elements and Large Deformations”. In: *Computer Methods in Applied Mechanics and Engineering* 197 (6-8), pp. 555–566.

Puso, M. A. and Laursen, T. A. (2002). “A 3D Contact Smoothing Method Using Gregory Patches”. In: *International Journal for Numerical Methods in Engineering* 54 (8), pp. 1161–1194.

Qi, L. and Sun, J. (1993). “A Nonsmooth Version of Newton’s Method”. In: *Mathematical Programming* 58 (1-3), pp. 353–367.

Reddy, J. N. (2014). *An Introduction to Nonlinear Finite Element Analysis: With Applications to Heat Transfer, Fluid Mechanics, and Solid Mechanics.* Second. Oxford: Oxford University Press.

Rouson, D., Xia, J., and Xu, X. (2011). *Scientific Software Design: The Object-Oriented Way.* Cambridge: Cambridge University Press.

Rumbaugh, J., Jacobson, I., and Booch, G. (1999). *Unified Modeling Language Reference Manual.* New York: Addison-Wesley.

Schweizerhof, K. and Konyukhov, A. (2005). “Covariant Description for Frictional Contact Problems”. In: *Computational Mechanics* 35 (3), pp. 190–213.

Scott, L. R. and Zhang, S. (1990). “Finite Element Interpolation of Nonsmooth Functions Satisfying Boundary Conditions”. In: *Mathematics of Computation* 54 (190), pp. 483–483.

Seitz, A., Farah, P., *et al.* (2016). “Isogeometric Dual Mortar Methods for Computational Contact Mechanics”. In: *Computer Methods in Applied Mechanics and Engineering* 301, pp. 259–280.

Seitz, A., Popp, A., and Wall, W. A. (2015). “A Semi-Smooth Newton Method for Orthotropic Plasticity and Frictional Contact at Finite Strains”. In: *Computer Methods in Applied Mechanics and Engineering* 285, pp. 228–254.

Seitz, A., Wall, W. A., and Popp, A. (2018). “A Computational Approach for Thermo-Elasto-Plastic Frictional Contact Based on a Monolithic Formulation Using Non-Smooth Nonlinear Complementarity Functions”. In: *Advanced Modeling and Simulation in Engineering Sciences* 5 (1), pp. 5–5.

Simo, J. C. and Hughes, T. J. R. (2006). *Computational Inelasticity.* Springer Science & Business Media.

Simo, J. C. and Laursen, T. A. (1992). “An Augmented Lagrangian Treatment of Contact Problems Involving Friction”. In: *Computers & Structures* 42 (1), pp. 97–116.

Simo, J. C., Wriggers, P., and Taylor, R. L. (1985). “A Perturbed Lagrangian Formulation for the Finite Element Solution of Contact Problems”. In: *Computer Methods in Applied Mechanics and Engineering* 50 (2), pp. 163–180.

Sitzmann, S., Willner, K., and Wohlmuth, B. I. (2016). “Variationally Consistent Quadratic Finite Element Contact Formulations for Finite Deformation Contact Problems on Rough Surfaces”. In: *Finite Elements in Analysis and Design* 109, pp. 37–53.

Stadler, M. and Holzapfel, G. A. (2004). “Subdivision Schemes for Smooth Contact Surfaces of Arbitrary Mesh Topology in 3D”. In: *International Journal for Numerical Methods in Engineering* 60 (7), pp. 1161–1195.

Stanley, H. M. and Kato, T. (1997). “An FFT-Based Method for Rough Surface Contact”. In: *Journal of Tribology* 119 (3), pp. 481–485.

Strömberg, N. (1997). “An Augmented Lagrangian Method for Fretting Problems”. In: *European journal of mechanics. A, Solids* 16 (4), pp. 573–593.

Strömberg, N., Johansson, L., and Klarbring, A. (1996). “Derivation and Analysis of a Generalized Standard Model for Contact, Friction and Wear”. In: *International Journal of Solids and Structures* 33 (13), pp. 1817–1836.

Stupkiewicz, S. (2007). *Micromechanics of Contact and Interphase Layers.* Lecture Notes in Applied and Computational Mechanics v. 30. Berlin ; London: Springer.

Stupkiewicz, S., Lewandowski, M. J., and Lengiewicz, J. (2014). “Micromechanical Analysis of Friction Anisotropy in Rough Elastic Contacts”. In: *International Journal of Solids and Structures* 51 (23), pp. 3931–3943.

Sutherland, I. E. and Hodgman, G. W. (1974). “Reentrant Polygon Clipping”. In: *Commun. ACM* 17 (1), pp. 32–42.

Taylor, R. L. and Papadopoulos, P. (1991). “On a Patch Test for Contact Problems in Two Dimensions”. In: *Nonlinear Computational Mecahnics*. Berlin: Springer-Verlag, pp. 690–702.

Taylor, R. L., Simo, J. C., *et al.* (1986). “The Patch Test—a Condition for Assessing FEM Convergence”. In: *International Journal for Numerical Methods in Engineering* 22 (1), pp. 39–62.

Temizer, İ. (2011). “Thermomechanical Contact Homogenization with Random Rough Surfaces and Microscopic Contact Resistance”. In: *Tribology International* 44 (2), pp. 114–124.

Temizer, İ. (2014). “Multiscale Thermomechanical Contact: Computational Homogenization with Isogeometric Analysis”. In: *International Journal for Numerical Methods in Engineering* 97 (8), pp. 582–607.

Temizer, İ. (2016). “Sliding Friction across the Scales: Thermomechanical Interactions and Dissipation Partitioning”. In: *Journal of the Mechanics and Physics of Solids* 89, pp. 126–148.

Temizer, İ. and Wriggers, P. (2008). “A Multiscale Contact Homogenization Technique for the Modeling of Third Bodies in the Contact Interface”. In: *Computer Methods in Applied Mechanics and Engineering* 198 (3), pp. 377–396.

Temizer, İ. and Wriggers, P. (2010a). “Inelastic Analysis of Granular Interfaces via Computational Contact Homogenization”. In: *International Journal for Numerical Methods in Engineering* 84 (8), pp. 883–915.

Temizer, İ. and Wriggers, P. (2010b). “Thermal Contact Conductance Characterization via Computational Contact Homogenization: A Finite Deformation Theory Framework”. In: *International Journal for Numerical Methods in Engineering* (March), n/a–n/a.

Temizer, İ., Wriggers, P., and Hughes, T. (2011). “Contact Treatment in Isogeometric Analysis with NURBS”. In: *Computer Methods in Applied Mechanics and Engineering* 200 (9-12), pp. 1100–1112.

Temizer, İ., Wriggers, P., and Hughes, T. (2012). “Three-Dimensional Mortar-Based Frictional Contact Treatment in Isogeometric Analysis with NURBS”. In: *Computer Methods in Applied Mechanics and Engineering* 209–212, pp. 115–128.

Temizer, İ. (2013). “Granular Contact Interfaces with Non-Circular Particles”. In: *Tribology International* 67, pp. 229–239.

Timoshenko, S. and Goodier, J. N. (1951). *Theory of Elasticity*. New York: McGraw-Hill.

Tur, M., Fuenmayor, F., and Wriggers, P. (2009). “A Mortar-Based Frictional Contact Formulation for Large Deformations Using Lagrange Multipliers”. In: *Computer Methods in Applied Mechanics and Engineering* 198 (37-40), pp. 2860–2873.

Tur, M., Giner, E., *et al.* (2012). “2D Contact Smooth Formulation Based on the Mortar Method”. In: *Computer Methods in Applied Mechanics and Engineering* 247–248, pp. 1–14.

Tworzydlo, W. W. *et al.* (1998). “Computational Micro- and Macroscopic Models of Contact and Friction: Formulation, Approach and Applications”. In: *Wear* 220 (2), pp. 113–140.

Tzanakis, I. *et al.* (2012). “Future Perspectives on Sustainable Tribology”. In: *Renewable and Sustainable Energy Reviews* 16 (6), pp. 4126–4140.

Vakis, A. *et al.* (2018). “Modeling and Simulation in Tribology across Scales: An Overview”. In: *Tribology International* 125, pp. 169–199.

Wagner, P., Wriggers, P., Klapproth, C., *et al.* (2015). “Multiscale FEM Approach for Hysteresis Friction of Rubber on Rough Surfaces”. In: *Computer Methods in Applied Mechanics and Engineering* 296, pp. 150–168.

Wagner, P., Wriggers, P., Veltmaat, L., *et al.* (2017). “Numerical Multiscale Modelling and Experimental Validation of Low Speed Rubber Friction on Rough Road Surfaces Including Hysteretic and Adhesive Effects”. In: *Tribology International* 111, pp. 243–253.

Wiesner, T. A. *et al.* (2018). “Algebraic Multigrid Methods for Dual Mortar Finite Element Formulations in Contact Mechanics”. In: *International Journal for Numerical Methods in Engineering* 114 (4), pp. 399–430.

Wilking, C. and Bischoff, M. (2017). “Alternative Integration Algorithms for Three-Dimensional Mortar Contact”. In: *Computational Mechanics* 59 (2), pp. 203–218.

Wilking, C., Bischoff, M., and Ramm, E. (2018). “Improved Contact Stress Recovery for Mortar-Based Contact Formulations”. In: Computational Methods in Applied Sciences 46. Ed. by E. Oñate *et al.*, pp. 395–415.

Williams, J. R. and O’Connor, R. (1999). “Discrete Element Simulation and the Contact Problem”. In: *Archives of Computational Methods in Engineering* 6 (4), pp. 279–304.

Wohlmuth, B. I. (2000). “A Mortar Finite Element Method Using Dual Spaces for the Lagrange Multiplier”. In: *SIAM Journal on Numerical Analysis* 38 (3), pp. 989–1012.

Wohlmuth, B. I. (2011). “Variationally Consistent Discretization Schemes and Numerical Algorithms for Contact Problems”. In: *Acta Numerica* 20 (April 2011), pp. 569–734.

Wohlmuth, B. I. (2012). *Discretization Methods and Iterative Solvers Based on Domain Decomposition.* Springer Science & Business Media.

Wohlmuth, B. I. and Krause, R. H. (2003). “Monotone Multigrid Methods on Nonmatching Grids for Nonlinear Multibody Contact Problems”. In: *SIAM Journal on Scientific Computing* 25 (1), pp. 324–347.

Wohlmuth, B. I., Popp, A., *et al.* (2012). “An Abstract Framework for a Priori Estimates for Contact Problems in 3D with Quadratic Finite Elements”. In: *Computational Mechanics* 49 (6), pp. 735–747.

Wriggers, P., Krstulovic-Opara, L., and Korelc, J. (2001). "Smooth C1-interpolations for Two-dimensional Frictional Contact Problems". In: *International Journal for Numerical Methods in Engineering* 51 (12), pp. 1469–1495.

Wriggers, P. (2006). *Computational Contact Mechanics*. Second. Berlin Heidelberg: Springer-Verlag.

Wriggers, P. and Reinelt, J. (2009). "Multi-Scale Approach for Frictional Contact of Elastomers on Rough Rigid Surfaces". In: *Computer Methods in Applied Mechanics and Engineering* 198 (21-26), pp. 1996–2008.

Wriggers, P., Vu Van, T., and Stein, E. (1990). "Finite Element Formulation of Large Deformation Impact-Contact Problems with Friction". In: *Computers & Structures* 37 (3), pp. 319–331.

Wu, J.-J. (2000). "Simulation of Rough Surfaces with FFT". In: *Tribology International* 33 (1), pp. 47–58.

Yang, B. and Laursen, T. A. (2008). "A Contact Searching Algorithm Including Bounding Volume Trees Applied to Finite Sliding Mortar Formulations". In: *Computational Mechanics* 41 (2), pp. 189–205.

Yang, B. and Laursen, T. A. (2009). "A Mortar-Finite Element Approach to Lubricated Contact Problems". In: *Computer Methods in Applied Mechanics and Engineering* 198 (47-48), pp. 3656–3669.

Yang, B., Laursen, T. A., and Meng, X. (2005). "Two Dimensional Mortar Contact Methods for Large Deformation Frictional Sliding". In: *International Journal for Numerical Methods in Engineering* 62 (9), pp. 1183–1225.

Yang, C., Tartaglino, U., and Persson, B. N. (2006). "A Multiscale Molecular Dynamics Approach to Contact Mechanics". In: *The European Physical Journal E* 19 (1), pp. 47–58.

Yastrebov, V. (2011). "Computational Contact Mechanics: Geometry, Detection and Numerical Techniques". PhD thesis. École Nationale Supérieure des Mines de Paris.

Yastrebov, V. A., Anciaux, G., and Molinari, J.-F. (2017a). "On the Accurate Computation of the True Contact-Area in Mechanical Contact of Random Rough Surfaces". In: *Tribology International* 114, pp. 161–171.

Yastrebov, V. A., Anciaux, G., and Molinari, J.-F. (2012). "Contact between Representative Rough Surfaces". In: *Physical Review E* 86 (3), p. 035601.

Yastrebov, V. A., Anciaux, G., and Molinari, J.-F. (2015). "From Infinitesimal to Full Contact between Rough Surfaces: Evolution of the Contact Area". In: *International Journal of Solids and Structures* 52, pp. 83–102.

Yastrebov, V. A., Anciaux, G., and Molinari, J.-F. (2017b). "The Role of the Roughness Spectral Breadth in Elastic Contact of Rough Surfaces". In: *Journal of the Mechanics and Physics of Solids* 107, pp. 469–493.

Yastrebov, V. A., Cailletaud, G., *et al.* (2015). “Three-Level Multi-Scale Modeling of Electrical Contacts Sensitivity Study and Experimental Validation”. In: *2015 IEEE 61st Holm Conference on Electrical Contacts (Holm),* pp. 414–422.

Yastrebov, V. A., Durand, J., *et al.* (2011). “Rough Surface Contact Analysis by Means of the Finite Element Method and of a New Reduced Model”. In: *Comptes Rendus Mécanique*. Surface Mechanics : Facts and Numerical Models 339 (7), pp. 473–490.

Zavarise, G. and Wriggers, P. (1998). “A Segment-to-Segment Contact Strategy”. In: *Mathematical and Computer Modelling* 28 (4-8), pp. 497–515.

Zavarise, G. and De Lorenzis, L. (2009). “A Modified Node-to-Segment Algorithm Passing the Contact Patch Test”. In: *International Journal for Numerical Methods in Engineering* 79 (4), pp. 379–416.

Zhi-Hua, z. and Nilsson, L. (1990). “A Contact Searching Algorithm for General 3-D Contact-Impact Problems”. In: *Computers & Structures* 34 (2), pp. 327–335.

Zhu, D. and Hu, Y.-Z. (2001). “A Computer Program Package for the Prediction of EHL and Mixed Lubrication Characteristics, Friction, Subsurface Stresses and Flash Temperatures Based on Measured 3-D Surface Roughness”. In: *Tribology Transactions* 44 (3), pp. 383–390.

Zienkiewicz, O. C., Taylor, R. L., and Fox, D. D. (2014). *The Finite Element Method for Solid and Structural Mechanics.* Amsterdam: Elsevier.