Durham E-Theses

Continuum mechanics and implicit material point method to underpin the modelling of drag anchors for cable risk assessment

PRETTI, GIULIANO

How to cite:

PRETTI, GIULIANO (2024) *Continuum mechanics and implicit material point method to underpin the modelling of drag anchors for cable risk assessment*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/15374/

Use policy

Continuum mechanics and implicit material point method to underpin the modelling of drag anchors for cable risk assessment

Giuliano Pretti

In the desperate yet strenuous hope that *"manuscripts don't burn."**

---

*M. Bulgakov, *The master and Margarita*. London: Collins and Harvill, 1967.

Continuum mechanics and implicit material point method to underpin the modelling of drag anchors for cable risk assessment
Giuliano Pretti

Abstract

The last years have seen an extraordinary expansion of the wind offshore industry towards new markets. With this development, wind farms are being pushed further offshore, leading to new challenges in maintaining their infrastructures. Data have revealed that one of the assets most susceptible to risks are power cables transporting the electricity generated offshore to the onshore transmission system. Whenever there is no alternative to shielding the cables, these are buried in the seabed to ward off the threat of drag anchors. Understanding the embedment process of drag anchors deployed by ships for mooring purposes is critical to determining the appropriate burial depth of cables. In turn, this protection strategy must include the influence exerted by the seabed conditions.

Historically, studies have focused on lab and field tests, whose results have been recently called into question. The lack of an appropriate scientific tool to investigate the anchor-cable interaction has recently fostered this process examination via numerical methods. Among these, the Material Point Method (MPM) is well-placed to master large deformation mechanics without mesh distortion and retains all of the advantages of a Lagrangian method. As a matter of fact, the MPM is the main object of investigation of this thesis, whose ultimate goal is to:

1. include inertia forces in the context of finite strain elasto-plasticity for solid materials;

2. expand the above point to handle the presence of water in the porous seabed; and

3. model the friction between the anchor and the surrounding soil.

This work has dedicated particular attention to the compliance of the MPM discretised algorithms with the underlying continuum formulation. In this sense, this work's primary contributions comprise:

- a conservation law consistent MPM algorithm for solid mechanics;

- a constitutive relationship for porous materials respectful of the solid mass conservation; and

- a rigorous assessment of the frictional contact formulations available in the MPM literature.

Supervisors: Prof. William M. Coombs and Prof. Charles E. Augarde

Acknowledgements

This work is dedicated to all those who contributed to my growth.

To my parents, Elio and Fernanda, who granted me everything, including the privilege of an education.

To my sisters, Elena and Monica, who helped me feel less lonely.

To my grandmother, who contributed so much to my early upbringing.

To those whom I consider friends and who consider themselves friends in return, it has been a pleasure to share a part of the path with you. Please do not be offended if I do not list you, but that would unequivocally exclude some, and you would not deserve it. If you are unsure whether this thesis is also meant for you, trust me, it is.

To Charles and Will (in no particular order), who have supported and tolerated me on this PhD journey, teaching me patiently what the art of research is all about.

To Anna, who despite knowing me for most of my life, not only sees the person I currently am, but possesses the pride and strength to tell me, typically lethargic, what potential I have.

I apologise if the thoughts I have expressed above appear banal or superficial. What all of you are to me is certainly beyond words.

Contents

Declaration

The work in this thesis is based on research carried out at the Department of Engineering, Durham University, England. No part of this thesis has been submitted elsewhere for any other degree or qualification, and it is the sole work of the author unless referenced to the contrary in the text.

Some of the work presented in this thesis has been published in journals and conference proceedings - the relevant publications are listed below.

Conference oral presentations

G. Pretti, W.M. Coombs, C.E. Augarde, "A displacement-controlled arc-length solution scheme," *Engineering Mechanics Institute Conference*, 22-24/03/2021, Online.

G. Pretti, W.M. Coombs, C.E. Augarde, "A comparison of implicit time-discretisation techniques for the material point method," *United Kingdom Association for Computational Mechanics*, 14-16/04/2021, Online.

G. Pretti, W.M. Coombs, C.E. Augarde, "A comparison of approaches for modelling poromechanics in the material point method," *United Kingdom Association for Computational Mechanics*, 20-22/04/2022, University of Nottingham. Paper submitted to proceedings.

Conference Papers

G. Pretti, W.M. Coombs, C.E. Augarde, M.M. Puigvert, J.A.R. Gutiérrez, L. Cross, "A review of drag anchor penetration models to inform cable burial risk assessment," in *4th International Symposium on Frontiers in Offshore Geotechnics*, American Society of Civil Engineers, 2020, pp. 1586-1595.

Publications

Publications towards this thesis

G. Pretti, W.M. Coombs, C.E. Augarde, B. Sims, M.M. Puigvert, J.A.R. Gutiérrez, "A conservation law consistent updated Lagrangian material point method for dynamic analysis", *Journal of Computational Physics*, vol. 485, p. 112 075, 2023.

G. Pretti, W.M. Coombs, C.E. Augarde, M.M. Puigvert, J.A.R. Gutiérrez, L. Cross, "Imposing constrained porosity values in a bi-phase elasto-plastic material under Terzaghi's effective stress principle," *Submitted to Mechanics of Materials*, 2023.

Publications separate from this thesis

G. Pretti, W.M. Coombs, C.E. Augarde, "A displacement-controlled arc-length solution scheme," *Computers & Structures*, vol. 258, p. 106674, 2022.

N.D. Gavin, G. Pretti, W. M. Coombs, J.C. Brigham, C.E. Augarde, "On the implementation of a material point-based arc-length method," *International Journal for Numerical Methods in Engineering*, 2024.

List of Figures

List of Tables

Acronyms

ALE Arbitrary Lagrangian-Eulerian

ALM Augmented Lagrangian method

AMPLE A Material Point Learning Environment

APIC Affine Particle in Cell

BC Boundary Condition

CBRA Cable Burial Risk Assessment

CEL Coupled Eulerian-Lagrangian

CLC Conservation Law Consistent

CFL Courant-Friedrichs-Lewy

CPDI Convected Particle Domain Interpolation

CPDI2 Second-order Convected Particle Domain Interpolation

cpGIMPM contiguous particle Generalised Interpolation Material Point Method

CVM Contact Velocity-based Method

DDMP Dual Domain Material Point

DEM Discrete Element Method

DoF Degree of Freedom

EFG Element-Free Galerkin

FEM Finite Element Method

FLIP FLuid Implicit Particle

GIMPM Generalised Interpolation Material Point Method

G2P Grid-to-point

KKT Karush-Kuhn-Tucker

LBB Ladyzhenskaya-Babuška-Brezzi

LHS Left-Hand Side

LM Lagrange Multiplier

LMM Lagrange Multiplier Method

MLPG Meshless Local Petrov Galerkin

MMS Method of Manufactured Solutions

MP Material Point

MPM Material Point Method

NR Newton-Raphson

P2G Point-to-grid

PIC Particle In Cell

PM Penalty Method

PolyPIC Polynomial Particle In Cell

PPP Polynomial Pressure Projection

RHS Right-Hand Side

SPH Smoothed Particle Hydrodynamics

TL Total Lagrangian

uGIMPM unchanged Generalised Interpolation Material Point Method

UL Updated Lagrangian

XPIC eXtended Particle In Cell

Nomenclature

Throughout this work, the standards of notation in the computational mechanics community were adopted. In particular:

- non-bold quantities represent scalars;

- bold symbols indicate vectors, tensors or matrices;

Standard scalar, vectorial and tensorial operators (such as inner product $(\bullet) \cdot (\bullet)$, double contraction $(\bullet) : (\bullet)$, Euclidean norm $\|(\bullet)\|$, dyadic product $(\bullet) \otimes (\bullet)$, transposition $(\bullet)^T$ and inversion $(\bullet)^{-1}$) are not explained in this work. Their use should be clear to the reader sufficiently familiar with this algebra. The use of upright (bold and non-bold) quantities was introduced to specify quantities computed at the mesh level, e.g., A_{IJ}, $\mathbf{A}$, or α_{IJ}, $\boldsymbol{\alpha}$.

The above notation has been used in conjunction with the index notation, which is preferred where the former may have been unclear. *Einstein* index notation over repeated indices is always implied unless otherwise stated, as in Sections 3.6.6 and 4.6.5. To make these operators as unambiguous as possible, different classes of indices were adopted according to these rules:

- lower case indices are referred to the Euclidean $\mathcal{E}$ dimensions of the space n^{dim}, $i, j, h, \cdots = 1 \ldots n^{dim}$;

- sans serif font indices are referred to the mesh nodes, with number of nodes N^{nds}, so that $\mathsf{A}, \mathsf{B}, \mathsf{C}, \cdots = 1 \ldots N^{nds}$;

- upper case indices are referred to the mesh Degrees of Freedom (DoFs), i.e., $I, J, H, \cdots = 1 \ldots n^{dim} \times N^{nds}$. These mesh DoFs can be obtained by $I = \mathsf{A} \times i$. The use of one notation over the other depends on the context. Sometimes, when different variables are considered, mesh DoFs can also use upper case calligraphic font such as $\mathcal{I}, \mathcal{J}$.

A particular notation was necessary in Section 5.4.4. This consists in indicating $(\bullet)\big|^{ij}$ the entry correspondent to the i-th row and j-th column of the considered matrix $(\bullet)$.

Identifiers

$^{h}(\bullet)$ . . . . finite-dimensional part of $(\bullet)$

Scalars

Vectors

Operators

$\boldsymbol{P}_{B(\boldsymbol{n},s)}(\bullet)$ . . . . friction projection operator associated with to direction $\boldsymbol{n}$ and threshold s

$\mathrm{sym}(\bullet)$ . . . . symmetric part of $(\bullet)$

$\mathrm{tr}(\bullet)$ . . . . trace operator

δ_{ij} . . . . Krönecker delta

ϵ_{ijk} . . . . Levi-Civita symbol

Chapter 1

Introduction

The energy industry ranks as a strategic asset in any government's schedule that needs continuous planning and development. On top of these constant requirements, several other factors are emerging. Among these new driving imperatives, Cerfontaine *et al.* [1] identified the growth of the population demanding more energy supplies and the need for decarbonisation to counter the climate emergency while decoupling the energy production from geopolitical events. To meet these necessities, the UK government has set the ambitious target of producing 50 GW by 2030 via the offshore wind industry, indicating an average deployment rate of 4.5 GW per year from 2023 (see [2]).

In this particular sector, cables connecting the wind farm to the onshore grid have proven to be a weak spot. Data (see [3, 4]) indicate that approximately $70-80\%$ of all financial losses and insurance claims are attributed to power cable failures. Warnock *et al.* [5] estimated the rate of power cable failure to be 0.00299 failures/km/year based on alternate current transmission connections of European wind farms. Spreading this value over the estimated lifespan of an offshore wind farm (between 20 to 25 years), each kilometre of cable could experience between $\approx 6\%$ and $\approx 7.5\%$ of failure probability. This datum should then be compared with the $3,834$ km of marine cables in the UK (see [6]) or the nearly $40,000$ km expected worldwide by 2030 (see [7]). All of these factors paint a rather worrying picture for cable risk management.

Literature (see [8, 9]) agrees that the most common threat to subsea power cables is represented by fishing gear or drag anchors (illustrated in Figure 1.1) deployed by ships. The common procedure to minimise this risk consists in avoiding this possibility in the first place, i.e., rerouting these facilities away from areas of high marine traffic or designated anchorage. However, when this practice is not applicable, cable burial is required, so the soil acts as a primary protection from the striking of the anchors or fishing gears[†].

Standard common practices in the industry developed in the 1980s stipulated a standard burial depth of 0.6 m (see Allan [11]). The Burial Protection Index (BPI) methodology (see Mole *et al.* [12]) was the first to introduce a differentiation between different soil

---

[†]Carter *et al.* [8] considers the damage caused by fishing gear highly unlikely if cables are buried more than 60 cm. Since anchors can penetrate deeper into the soil (see Section 1.1.2), it can be concluded that anchors striking represents the worst case scenario for the safety of cables.
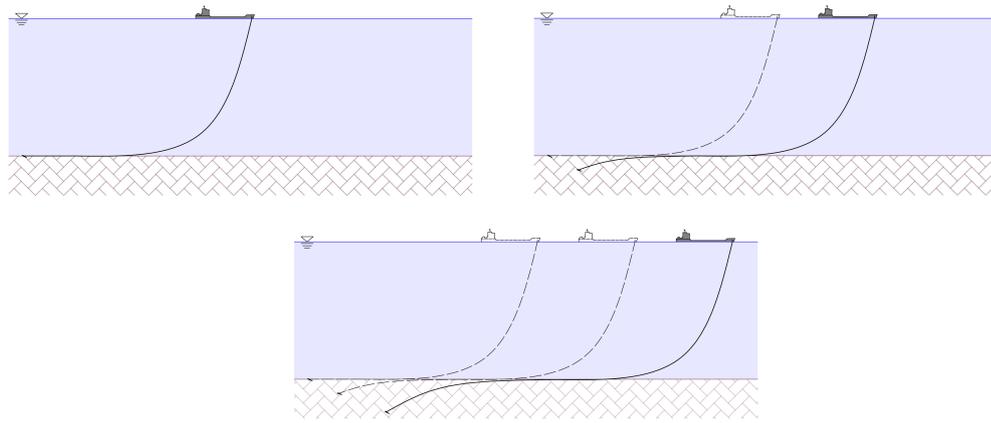
(a) Illustration of the progressive embedment process.
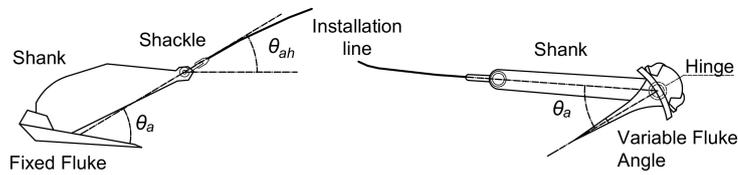
(b) Illustration of the main components and of some key parameters of a drag anchor.
Figure reproduced from Pretti *et al.* [10].

Figure 1.1: Anchor embedment process (top rows) and a couple of anchor typologies with relative nomeclature (bottom row).

conditions. The standard procedure employed nowadays is the Cable Burial Risk Assessment (CBRA) by the Carbon Trust [13]. This method is designed as a probabilistic approach to define an optimum burial depth while balancing the installation costs. Pivotal in the information process of such a method is the penetration depth of drag anchors in different seabed conditions and for different anchor typologies and weights. It can be understood that estimating the penetration depth of anchors is vital to selecting the appropriate burial depth of power cables. So far, the burial depth considered by the CBRA has been informed by anchor field trials run in the 1980s [14]. However, doubts concerning these tests have arisen in various directions about the representativeness and, thus, the validity of these tests. On one hand, more recent publications (see [13, 15, 16]) have raised a general concern about the soundness and consistency of those guidelines, questioning their over-conservatism regarding the suggested burial depth. On the other hand, the above-discussed percentage of failure per kilometre suggests that the burial depth could be insufficient in other situations. Overall, a rather uncertain picture emerges from the lack of unambiguous and sufficiently reliable indications. This is confirmed by the words of Zheng *et al.* [17] in a recent work:

> ''At present, there is no unified standard for the determination of
> burial depth of submarine cables, which is only based on operation
> experience or general definition.''

The following sections are intended to illustrate and discuss how, besides the above-mentioned

tests, the kinematics of the embedment process have been studied in the literature, distinguishing between analytical and numerical methods[*]. However, the work cited below will confirm how the literature is scant on extensive parametric studies to inform the depth of penetration required for the CBRA procedure.

1.1 Review of the methods employed in the study of drag anchor kinematics

The following Sections 1.1.1 and 1.1.2 heavily drawn on Pretti *et al.* [10].

1.1.1 Analytical methods

From a historical perspective, analytical methods have been the most investigated tool so far, primarily due to the possibility of developing them via calculations by hand. Four main approaches can be tracked back from the literature: limit equilibrium methods, yield envelope formulations, limit analysis approaches and kinematic enforcement methods.

The limit equilibrium method [18–25] describes the soil surrounding the anchorage under failure conditions. The resisting mechanisms offered by the soil in different directions are considered as separate ultimate resistances. However, these bearing capacity factors in different directions are problematic to estimate *a priori*, as they depend on the geometry of the anchor which, in turn, affects its unknown trajectory.

The yield envelope formulations [26–29] are based on plasticity concepts, where a combination of forces acting on the anchor and leading to soil failure must be determined. The locus determined by this combination of forces is called the yield envelope. Given this surface, the anchor motion can be evaluated using concepts inherited from continuum plasticity, such as the associated flow rule (giving the direction of the anchor movement) and the orthogonality condition (providing the magnitude of this displacement). However, the coefficients coupling the combination of forces in the locus surface must be determined via several numerical analyses, which are tailored to the specific case under consideration (i.e., soil resistance and anchor geometry). This is also why this method is difficult to generalise to pressure-independent soils, as considering this factor varying with depth would make the number of numerical analyses even larger.

The workgroup formed by Aubery, Kim and Murff established an implementable tool to predict anchor kinematics valid for clays (see [30–34]). Based on the limit analysis method, the internal and external powers generated by the motion of the anchor is calculated. The internal power is computed considering anchor rigid rotations about an unknown centre. This power must be equal to the external power supplied by the traction at the shackle (provided by the installation line), which varies depending on the angle between the direction of this traction and the horizontal line (θ_{ah} in Figure 1.1(b)). The graphical intersection

---

[*]Field and lab tests are not included in this work, whose main focus concerns computational techniques. However, the reader interested in a review of these tests is referred to Pretti *et al.* [10].

of these two powers (internal and external) selects the value of θ_{ah} compatible with both. This angle value can be related to the centre of rotation of the anchor which, in turn, describes an incremental displacement. This method is particularly effective for anchors with a bulky shank. In addition, the reason why this technique can be applied only to clays is that the limit loads for this soil can (in some circumstances) be independent of depth.

In the case of the kinematic enforcement method proposed by Liu *et al.* [35], the motion of the anchor is chosen *a priori* to be a cycloid, i.e., a curve traced by a point on the circumference of a circle whose centre is moving at a constant speed along a straight line. This model can be applied to both sand and clay soils. However, this model strongly depends on two other parameters (θ_{ah} illustrated in Figure 1.1(b) and the final burial depth of the anchor), which can be calculated only via other analytical or numerical methods.

As can be guessed by the description of the analytical methods, their simplicity constitute, at the same time, their strength and weakness; these methods can be applied only to specific cases and, in some instances, they rely on data (mainly from numerical analyses) to work correctly. Therefore, in recent years, more attention has been paid to numerical methods to model drag anchor embedment process. These are addressed in the following section.

1.1.2 Numerical modelling of drag anchor kinematics

The relatively recent (2010) inclusion of the Coupled Eulerian-Lagrangian (CEL) method[*] in the commercial software Abaqus (see [36]) has enabled the modelling of a solid under considerable deformations in just a few clicks, thus permitting the exploration of anchor trajectory.

To the author's best knowledge, Liu and Zhao's work [37] played a significant role as a forerunner in the study of this topic. These authors considered the whole chain-anchor system, modelled as rigid bodies. In particular, cylindrical rigid bodies allowing for relative rotation among them represented the chain segments, while rectangular or wedge-shape forms modelled anchors with a fixed fluke. Liu and Zhao's simulation considered a linear elastic-perfectly plastic (von Mises) material in the small strain regime for the soil, with a Poisson's ratio close to 0.5 to approximate the undrained conditions. The interface between the rigid bodies and the surrounding soil was governed by Coulomb's friction law. These studies investigated the role of the prescribed velocity (applied at the free end of the chain) on the penetration depth, highlighting that faster speeds affect only the initial phases of the anchor embedment, not its final penetration depth.

Grabe *et al.* [38], Osthoff *et al.* [39] and Grabe and Heins [40] expanded the field of investigation, considering a vast range of partially drained simulations in sands (from very loose to dense sands, in particular) coupled with an anchor with a moving fluke (model AC14). As for Liu and Zhao [37], this work modelled the chain and the anchor as rigid bodies, while the sand was modelled with a hypoplastic material capable of representing different features, such as nonlinear and inelastic behaviour, dilatancy, or dependence of

---

[*]A qualitative description of the CEL method and other numerical tools is provided in Section 1.1.3.

the stiffness parameters on pressure and void ratio. These analyses [38–40] found that the drained conditions correspond to the most profound anchor embedment scenario, while, in the case of undrained conditions, it could also be possible that the anchor cannot penetrate the soil. Sand density level plays a part too, with a deeper embedment found in the case of loose sands. These numerical analyses also emphasised the role played by the presence of the cable in the embedment process, revealing three scenarios: the anchor drags the cable if this is buried superficially (≈ 0.3 m); an indirect contact in the form of stresses applied on the cable is found if its burial depth is intermediate (≈ 0.5 m); and no interaction between the cable and the anchor kinematics takes place if the burial depth of the former is considerable (≈ 1.5 m). The results obtained by these authors were also confirmed by real-scale tests (see [41]), attesting the considerable achievement of their numerical setup.

A setup very similar (in terms of geometry and analyses) to those proposed in [38–40] was considered by Grabe and Wu [42] to model the penetration of the AC 14 anchor (see [43]) in clay. To properly represent this kind of soil, a visco-hypoplastic model was adopted, and undrained conditions were considered. A periodic penetration process is observed in these numerical analyses, with the anchor embedding and resurfacing at regular intervals. These results agree with observed behaviour of ship anchors in very soft clay (see [43]). The sizes and the weights of the AC 14 anchors varied from $12,500$ t to $50,000$ t, and a full penetration (slightly more than 2 m depth) could be appreciated only in the case of the heaviest anchor. Further analyses were performed by varying the dragging speed from a minimum of 0.5 m s^{-1} to a maximum of 2 m s^{-1}. Results showed that no penetration occurs when the speed exceeds 1.25 m s^{-1}, regardless of the soil strength (this spans from 5 kPa to 20 kPa).

As it emerges from the above literature, all the listed publications employed the CEL method implemented in Abaqus [36]. While using this software satisfies a pragmatic approach, it can provide the user only with a certain degree of flexibility. Hence, to deliver these studies by choosing the techniques selected by the analyst's best awareness, the appropriate numerical tool must be wisely chosen and developed accordingly. The following section details the numerical models that might be employed when dealing with soils undergoing large deformations (as the installation process suggests; see Figure 1.1(a)), assessing their strengths and weaknesses.

1.1.3 A survey on available numerical techniques

No numerical study can disregard that the Finite Element Method (FEM) is the foremost method when approaches based on the hypothesis of continuum material are considered. However, the embedment process of the anchor comes with considerable deformations of the surrounding soil. Since FEMs are based on the persistent presence of a unique grid, this mesh can be significantly distorted when deformations and rotations (such as the one impressed by the anchor) might occur. This distortion can invalidate the results, thus making this technique unsuitable for studying anchor kinematics.

As the limitation of FEMs for this particular purpose comes from the mesh distortion, it seems natural to substitute the mesh when distorted (i.e., *remeshing*) and, therefore, move

the historical variables from the old to the new position (i.e., *remapping*). This core idea led to Arbitrary Lagrangian-Eulerian (ALE) methods, whose name comes from the similarity between the remeshing process and the convection in Eulerian analysis involving material flow. Augarde *et al.* [44] classify the remeshing methods into two main strategies, i.e., those preserving mesh connectivity (nodes are rearranged to mitigate mesh distortion) and those that introduce the entire mesh anew. As for the remapping techniques, Mota *et al.* [45] distinguish between stress recovery techniques (based on a multi-field variational principle) and mapping of internal variables from one mesh to another (via differently designed transfer operators). As testified by the proliferation of techniques in both the remeshing and remapping algorithms (see Augarde *et al.* [44] and references therein), there is no consensus on which method is the best, but the analyst must determine the adequate technique according to the circumstances.

As evidenced by the publications reported in Section 1.1.2, the CEL method constitutes a valid alternative to the ALE techniques. While the roots of the CEL method can be found in the 1960s (see [46]), its popularity has increased since its implementation in commercial codes (Abaqus [36] as previously mentioned, but also LS-DYNA [47]). This method stipulates that two domains are considered, and their description takes place according to different methodologies. A domain representing a softer material is described following an Eulerian approach, while a second domain, modelling a stiffer (or rigid) material, is tracked by the use of a Lagrangian description*. The interaction between these two domains occurs via a contact formulation, sometimes considering friction between them. The presence of an Eulerian description increases the computational costs since the convected terms appearing in the material derivatives must be considered.

Smoothed Particle Hydrodynamics (SPH) is the oldest mesh-less technique and bases the discretisation of the strong form of continuum equations only on particles, thus requiring specific search algorithms to locate the nearest neighbours for a given particle. In SPH, required material properties are spread over the spacing distance between the particles (called the smoothing length, hence the method's name). While attractive for problems in astrophysics (see, for instance, [48]) and for applications with large deformations (see, e.g., [49]), applying boundary conditions in the SPH is particularly troublesome, as documented in Soga *et al.* [50] and Augarde *et al.* [44]. Moreover, the commonly adopted point-wise integration can cause instabilities due to under-integration issues.

Other meshless techniques are represented by the Element-Free Galerkin (EFG) and the Meshless Local Petrov Galerkin (MLPG) methods, where nodes (with no elements connecting them) constitute the only discretisation available. Contrary to the SPH, these methods are based on the weak form of the continuum equations. However, while the EFG method does not qualify entirely as a mesh-free technique, the MLPG does. The main reason lies in the necessity of integration points for solving the global weak form in the case of the EFG method. In contrast, the MLPG, as the name suggests, solves the local *Petrov Galerkin* weak form, thus not requiring elements or meshes for interpolation or integration purposes. As claimed by several sources (see Soga *et al.* [50] and Augarde *et al.* [44], for instance), the

---

*Definitions of Lagrangian and Eulerian descriptions are given in Chapter 4.

computational costs, especially for small to medium problems, provide a critical barrier in the proliferation of these methods.

Beyond continuous methods, O'Sullivan [51] sustains that the Discrete Element Method (DEM) represents the most popular discontinuum-based approach. The versatility of the DEM is also testified by its adoption in extremely different contexts, such as geotechnical applications [52], masonry structure [53], up to human traffic flow for evacuations [54]. Since the force acting on a single particle is computed by the interaction exerted by the particles coming into contact with it, contact detection algorithms play a crucial role in the DEM. While the vast literature has recognised the potential of the DEM, its high computational costs, especially for field scale simulations, indicate that continuum-based approaches are preferable (see, for instance, Augarde *et al.* [44] and Coetzee *et al.* [55]).

Last but not least, the Material Point Method (MPM) (see [56, 57]) must be included among the methods based on the continuum hypothesis. The MPM is entirely based on a Lagrangian description of the material, represented by a cloud of points (Material Points (MPs)) and assisted by a computational grid, which is disposed of and renewed at every time-step. This method discretises the weak forms of equations and proved to be very successful when modelling solids undergoing extreme deformations and rotations. Since the MPM is the numerical tool adopted in this work, its details are given later, with an overview on the method presented in Chapter 2. However, a brief comparison of this method with the above-listed techniques must be given at this stage to, if not motivate, at least present the strengths that justify its adoption. The MPM does not suffer from the mesh distortion happening in FEM simulations with extreme deformations and, unlike the ALE method, presents a consistent algorithm to generate and dispose of a simple (not body-fitted) mesh at every time-step. At the same time, the MPM can use all the constitutive relationships mainly developed for a FEM framework. If the implicit version of the MPM is considered, its similarity with the FEM have been well-expressed by Guilkey and Weiss [58], where the MPM is described as

> ``... a finite element method where the integration points (material points) are allowed to move independently of the mesh.''

Using a computational grid in the MPM resolves the problem of the nearest particle search (contrary to the SPH method and the other mesh-less techniques mentioned above), and contact detection (contrary to the DEM). Moreover, the grid in the MPM moves with the MPs in the time-step, so that the considered primary equations can be set in a Lagrangian description, differently from the CEL method. In some specific cases (described and discussed in multiple sections of this work), the MPM can apply Dirichlet Boundary Conditions (BCs) on the mesh without reconstructing it,which makes it more convenient than, for instance, SPH. The main drawbacks of the MPM are introduced and discussed in detail later in this work, when the narrative demands it.

1.2 Outline and novelties of this work

1.2.1 Thesis outline

The previous section has established the suitability of the MPM for modelling drag anchor kinematics during the embedment process. As such, the MPM constitutes the main object of investigation of this manuscript. Having stipulated these driving factors, it is now possible to introduce an outline of this work. This thesis consists of the following chapters, which can be divided based on their content into three main sections:

- the first of these sections constitutes the introduction for the later parts, in particular:

  Chapter 1 has provided the rationale and necessity that support this investigation (i.e., the modelling of drag anchor kinematics), comparing the analytical and numerical tools that are already available in the literature or might be suitable for this purpose. In particular, it was found that the MPM might be the right numerical method to describe anchor kinematics while retaining significant similarities with the FEM (in particular, constitutive models for soils could be straightforwardly reused in the MPM framework).

  As such, **Chapter 2** focuses on this numerical method. This chapter provides an overview of the method and its algorithm, while offering the different options available in the MPM literature;

- the second section deals with modelling the principal phenomena which can be significant in the anchor embedment process in the seabed. All of the following chapters present the investigation of the continuum-based formulations of these phenomena from the principles of classical mechanics. These continuum principles are subsequently discretised in the MPM algorithms, which, as far as possible, are desired to inherit the properties of their continuum formulations. Numerical examples provide the basis for discussing the developed algorithms, highlighting the difference between those already available in the literature and those proposed here.

  Chapter 3 details the dynamics of a solid material undergoing large elasto-plastic strains and rotations.

  Chapter 4 extends the dynamics of a continuum body (in the conditions described in Chapter 3) to the case of porous materials. These materials are characterised by the presence of an interstitial fluid whose pressure plays a significant role in the mechanical behaviour of the comprehensive macroscopic material.

  Chapter 5 describes the mechanics of frictional contact. Different methods available in the literature for the FEM and the MPM are unified in a single framework and assessed thoroughly.

  It should be understood that Chapter 3 is preliminary for Chapter 4. This latter aims to describe the mechanical behaviour of the seabed, covering all the drainage conditions available. Chapter 5 deals with the friction between the anchor's surfaces and the

surrounding soil. Since the embedment process is a relatively low-frequency problem, this motivates the adoption of the implicit version of the MPM for all of these chapters, while favouring longer time-steps; and

- the last section (which consists only of **Chapter 6**) presents the conclusions relative to the work developed in the previous chapters and provides a few recommendations on the possible directions for future work.

The appendices at the end of this work provide some details on tensor algebra and report cumbersome calculations that might decrease the text readability.

1.2.2   Thesis novelties and relative publications

The parallel study of continuum mechanics formulations and their resulting application to MPM algorithms has brought to light how discrepancies can arise between these two parts. In particular, it emerged that the MPM algorithms might not respect all of the continuum equations not expressly discretised. While still maintaining the study and development of the physical phenomena necessary to describe the anchor kinematics as a background objective, this work set the primary goal of designing MPM algorithms inheriting the desired properties from the continuous equations of mechanics. These two purposes drove most of the novelties of the current manuscript:

- the literature review of drag anchors and their study is partially contained in Chapter 1 and its extended version has been presented in Pretti *et al.* [10];

- Chapter 3 describes the first Updated Lagrangian (UL) implicit MPM algorithm for the dynamics of solid materials devised to comply with the conservation of momenta and Clausius-Duhem inequality. The design of this algorithm allowed the assessment of Particle In Cell (PIC) and FLuid Implicit Particle (FLIP) mapping techniques. In turn, this evaluation led to a rigorous analysis which clarifies the dissipative or conserving nature of these mappings. An approximated Courant-Friedrichs-Lewy (CFL) condition accounting for the spatial variability in MPM analyses was introduced to guide implicit dynamic simulations. This chapter and Pretti *et al.* [59] share these novelties;

- Chapter 4 introduces a free energy function for porous material describing the effective stress-strain constitutive relationship conserving the solid mass constituent. The introduction of this free energy function serves the above-explained purpose of designing MPM discrete algorithms sufficiently informed by their continuum physics. However, this free energy function could be directly included in FEM formulations. The development of this energy function is outlined also in Pretti *et al.* [60]. Other secondary novelties of this chapter concern $\boldsymbol{u} - p^{(f)}$ formulations. For these algorithms, some clarifications on whether or not mapping the pressure field are qualitatively discussed. The need for boundary stabilisations is also introduced adapting a technique from the cut-FEM community. Similar findings of interest to the MPM community regards the

Polynomial Pressure Projection (PPP) stabilisation, where it is emphasised that this method might be ineffective for the type of considered equations; and

- Chapter 5 unifies some of the different algorithms available in the MPM and the FEM in a unique framework. In particular, sources of dissipation (when present) for these algorithms are made clear. As for some of the above novelties, these clarifications could be particularly effective for the part of the MPM community more dedicated to code development.

Chapter summary

This chapter has detailed why drag anchors are a matter of interest for cables protection in the offshore wind industry, showing that no standard practice exists on cable burial depth. A summary of the analytical and numerical methods used in the literature to investigate drag anchors has been given. Other possible numerical techniques available for simulating a material undergoing extreme distortions and deformations have been briefly discussed. This analysis has shown that the MPM is well-placed to investigate the drag anchor embedment process. As the MPM constitute the main object of investigation of this work, an outline and the original contributions of the thesis have been described.

References

[1] B. Cerfontaine, D. White, K. Kwa, S. Gourvenec, J. Knappett and M. Brown, 'Anchor geotechnics for floating offshore wind: Current technologies and future innovations,' *Ocean Engineering*, vol. 279, p. 114 327, 2023.

[2] British Energy Security Strategy, Deparment for Business, Energy & Industrial Strategy, Department for Energy Security & Net Zero, Apr. 2022. [Online]. Available: `http://https://www.gov.uk/government/publications/british-energy-security-strategy/british-energy-security-strategy` (visited on 12/09/2022).

[3] C. Strang-Moran. 'Offshore wind leaders urged to collaborate for radical advances to reduce subsea cable failures.' (Nov. 2020), [Online]. Available: `https://ore.catapult.org.uk/press-releases/world-first-data-collection-platform-to-reduce-subsea-cable-failures/` (visited on 12/09/2023).

[4] J. McKenna. 'Experience is key to reduce cable claims.' (Jun. 2012), [Online]. Available: `https://www.windpowermonthly.com/article/1133682/experience-key-reduce-cable-claims` (visited on 13/09/2023).

[5] J. Warnock, D. McMillan, J. Pilgrim and S. Shenton, 'Failure rates of offshore wind transmission systems,' *Energies*, vol. 12, no. 14, p. 2682, 2019.

[6] C. Strang-Moran, 'Subsea cable management: Failure trending for offshore wind,' *Wind Energy Science Discussions*, pp. 1–11, 2020.

[7] L. Burley, V. Terente, L. Hewitson and H. Stuart, 'Cable integrity risk assessment (CIRA),' OSIG Conference:Innovative Geotechnologies for Energy Transition, September 2023.

[8] L. Carter, D. Burnett, S. Drew, G. Marle, L. Hagadorn, D. Bartlett-Ncneil and N. Irvine, 'Submarine cables and the oceans: Connecting the world (no. 31),' *UNEP/Earthprint*, 2010.

[9] G. M. Insights, *Offshore Wind Cable Manufacturing and Installation Forecast 2018–2028*. RenewableUK, 2018.

[10] G. Pretti, W. Coombs, C. Augarde, M. Puigvert, J. Gutierrez and L. Cross, 'A review of drag anchor penetration models to inform cable burial risk assessment,' in *4th International Symposium on Frontiers in Offshore Geotechnics*, American Society of Civil Engineers, 2020, pp. 1586–1595.

[11] P. Allan, 'Selecting appropriate cable burial depths a methodology,' in *A Methodology IBC conference on Submarine Communication. The Future of Network Infrastructure, France*, Citeseer, 1998, pp. 1–12.

[12] P. Mole, I. Featherstone and S. Winter, 'Cable protection: Solutions through new installation and burial approaches: Les liaisons optiques sous-marines,' *REE. Revue de l'électricité et de l'électronique*, no. 5, pp. 34–39, 1997.

[13] C. Trust. 'Cable Burial Risk Assessment Methodology.' (Feb. 2015), [Online]. Available: `https://ctprodstorageaccountp.blob.core.windows.net/prod-drupal-files/documents/resource/public/cable-burial-risk-assessment-guidance.pdf` (visited on 22/09/2023).

[14] N. C. E. L. (Hueneme), *Drag embedment anchors for navy moorings*. Naval Civil Engineering Laboratory, 1987, vol. 83.

[15] D. Luger and M. Harkes, 'Anchor tests German Bight: Test setup and results,' *Deltares*, 2013.

[16] M. Junagade, 'Study of methods to assess the risk of damage for buried offshore power cables,' Ph.D. dissertation, University of Texas, 2022.

[17] X. Zheng, Y. Ge, Z. Lu, C. Cao, P. Zhou, S. Li and J. Chen, 'Study on buried depth protection index of submarine cable based on physical and numerical modeling,' *Journal of Marine Science and Engineering*, vol. 10, no. 2, p. 137, 2022.

[18] S. Neubecker and M. Randolph, 'The performance of drag anchor and chain systems in cohesive soil,' *Marine Georesources & Geotechnology*, vol. 14, no. 2, pp. 77–96, 1996.

[19] C. Thorne, 'Penetration and load capacity of marine drag anchors in soft clay,' *Journal of Geotechnical and Geoenvironmental engineering*, vol. 124, no. 10, pp. 945–953, 1998.

[20] R. Ruinen, 'Penetration analysis of drag embedment anchors in soft clays,' in *The Fourteenth International Offshore and Polar Engineering Conference*, International Society of Offshore and Polar Engineers, 2004.

[21] H. Liu, Y. Li, H. Yang, W. Zhang and C. Liu, 'Analytical study on the ultimate embedment depth of drag anchors,' *Ocean Engineering*, vol. 37, no. 14-15, pp. 1292–1306, 2010.

[22] W. Zhang, H. Liu, Y. Zhao and Y. Yue, 'Interactional properties between drag anchor and installation line,' *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 140, no. 2, p. 04 013 018, 2013.

[23] W. Zhang, H. Liu, X. Li, Q. Li and J. Cao, 'An analytical method for positioning drag anchors in seabed soils,' *China Ocean Engineering*, vol. 29, no. 1, pp. 49–64, 2015.

[24] H. Liu, W. Zhang, C. Liu and C. Hu, 'Movement direction of drag anchors in seabed soils,' *Applied Ocean Research*, vol. 34, pp. 78–95, 2012.

[25] L. Wang, K. Shen, L. Li and Z. Guo, 'Integrated analysis of drag embedment anchor installation,' *Ocean Engineering*, vol. 88, pp. 149–163, 2014.

[26] M. O'Neill, M. Bransby and M. Randolph, 'Drag anchor fluke soil interaction in clays,' *Canadian Geotechnical Journal*, vol. 40, no. 1, pp. 78–94, 2003.

[27] S. Elkhatib and M. Randolph, 'The effect of interface friction on the performance of drag-in plate anchors,' in *Proc. Int. Symp. on Frontiers in Offshore Geotechnics, IS-FOG05*, Taylor & Francis Perth, Western Australia, 2005, pp. 171–177.

[28] Y. Tian, M. Randolph and M. Cassidy, 'Analytical solution for ultimate embedment depth and potential holding capacity of plate anchors,' *Géotechnique*, vol. 65, no. 6, pp. 517–530, 2015.

[29] Y. Tian, M. Cassidy, C. Gaudin and M. Peng, 'Improving force resultant model for anchors in clays from large deformation finite element analysis,' *Marine Georesources & Geotechnology*, pp. 1–9, 2019.

[30] B. Kim, 'Upper bound analysis for drag anchors in soft clay,' Ph.D. dissertation, Texas A&M University, 2007.

[31] C. Aubeny, B. Kim and J. Murff, 'Proposed upper bound analysis for drag embedment anchors in soft clay,' *Frontiers in Offshore Geotechnics. London*, Taylor and F. Group, Eds., pp. 179–183, 2005.

[32] C. Aubeny, J. Murff and B. Kim, 'Prediction of anchor trajectory during drag embedment in soft clay,' *International Journal of Offshore and Polar Engineering*, vol. 18, no. 04, 2008.

[33] C. Aubeny and C. Chi, 'Mechanics of drag embedment anchors in a soft seabed,' *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 136, no. 1, pp. 57–68, 2009.

[34] C. Aubeny and C. Chi, 'Trajectory prediction for drag embedment anchors under out of plane loading,' in *Frontiers in Offshore Geotechnics II*, CRC Press, 2010, pp. 717–722.

[35] H. Liu, C. Liu, H. Yang, Y. Li, W. Zhang and Z. Xiao, 'A novel kinematic model for drag anchors in seabed soils,' *Ocean Engineering*, vol. 49, pp. 33–42, 2012.

[36] G. Abaqus, 'Abaqus 6.11,' *Dassault Systemes Simulia Corporation, Providence, RI, USA*, p. 3, 2011.

[37] H. Liu and Y. Zhao, 'Numerical study of the penetration mechanism and kinematic behavior of drag anchors using a coupled Eulerian-Lagrangian approach,' *Geotechnical Engineering Journal of the SEAGS & AGSSEA*, vol. 45, no. 4, pp. 29–39, 2014.

[38] J. Grabe, G. Qiu and L. Wu, 'Numerical simulation of the penetration process of ship anchors in sand,' *Geotechnik*, vol. 38, no. 1, pp. 36–45, 2015.

[39] D. Osthoff, E. Heins and J. Grabe, 'Impact on submarine cables due to ship anchor–soil interaction,' *Geotechnik*, vol. 40, no. 4, pp. 265–270, 2017.

[40] J. Grabe and E. Heins, 'Penetration of ship anchors and the influence of submarine cables,' in *International Conference on Offshore Mechanics and Arctic Engineering*, American Society of Mechanical Engineers, vol. 51302, 2018.

[41] D. Luger and M. Harkes, 'Anchor tests German Bight, Test set-up and results,' Deltares, Tech. Rep., 2013.

[42] J. Grabe and L. Wu, 'Coupled Eulerian-Lagrangian simulation of the penetration and braking behaviour of ship anchors in clay,' *Geotechnik*, vol. 39, no. 3, pp. 168–174, 2016.

[43] 'The guide to anchoring,' Vryhof Anchors, Tech. Rep., 2010.

[44] C. E. Augarde, S. J. Lee and D. Loukidis, 'Numerical modelling of large deformation problems in geotechnical engineering: A state-of-the-art review,' *Soils and Foundations*, vol. 61, no. 6, pp. 1718–1735, 2021.

[45] A. Mota, W. Sun, J. T. Ostien, J. W. Foulk and K. N. Long, 'Lie-group interpolation and variational recovery for internal variables,' *Computational Mechanics*, vol. 52, pp. 1281–1299, 2013.

[46] W. F. Noh, 'CEL: A time-dependent, two-space-dimensional, coupled Eulerian-Lagrange code,' Lawrence Radiation Lab., Univ. of California, Livermore, Tech. Rep., 1963.

[47] J. O. Hallquist *et al.*, 'LS-DYNA theory manual,' *Livermore software Technology corporation*, vol. 3, pp. 25–31, 2006.

[48] V. Springel, 'Smoothed particle hydrodynamics in astrophysics,' *Annual Review of Astronomy and Astrophysics*, vol. 48, pp. 391–430, 2010.

[49] C. Augarde and C. Heaney, 'The use of meshless methods in geotechnics,' in *Computational Geomechanics: COMGEO I: Proceedings of the 1st International Symposium om Computational Geomechanics (COMGEO I)*, 2009, pp. 311–320.

[50] K. Soga, E. Alonso, A. Yerro, K. Kumar and S. Bandara, 'Trends in large-deformation analysis of landslide mass movements with particular emphasis on the material point method,' *Géotechnique*, vol. 66, no. 3, pp. 248–273, 2016.

[51] C. O'Sullivan, 'Particle-based discrete element modeling: Geomechanics perspective,' *International Journal of Geomechanics*, vol. 11, no. 6, pp. 449–464, 2011.

[52] Y. U. Sharif, M. J. Brown, B. Cerfontaine, C. Davidson, M. O. Ciantia, J. A. Knappett, J. D. Ball, A. Brennan, C. Augarde, W. Coombs *et al.*, 'Effects of screw pile installation on installation requirements and in-service performance using the discrete element method,' *Canadian Geotechnical Journal*, vol. 58, no. 9, pp. 1334–1350, 2021.

[53] J. V. Lemos, 'Discrete element modeling of masonry structures,' *International Journal of Architectural Heritage*, vol. 1, no. 2, pp. 190–213, 2007.

[54] D. Helbing, I. Farkas and T. Vicsek, 'Simulating dynamical features of escape panic,' *Nature*, vol. 407, no. 6803, pp. 487–490, 2000.

[55] C. Coetzee, A. Basson and P. Vermeer, 'Discrete and continuum modelling of excavator bucket filling,' *Journal of terramechanics*, vol. 44, no. 2, pp. 177–186, 2007.

[56] D. Sulsky, Z. Chen and H. Schreyer, 'A particle method for history-dependent materials,' *Computer Methods in Applied Mechanics and Engineering*, vol. 118, no. 1-2, pp. 179–196, 1994.

[57] D. Sulsky, S.-J. Zhou and H. L. Schreyer, 'Application of a particle-in-cell method to solid mechanics,' *Computer Physics Communications*, vol. 87, no. 1-2, pp. 236–252, 1995.

[58]   J. E. Guilkey and J. A. Weiss, 'Implicit time integration for the material point method: Quantitative and algorithmic comparisons with the finite element method,' *International Journal for Numerical Methods in Engineering,* vol. 57, no. 9, pp. 1323–1338, 2003.

[59]   G. Pretti, W. M. Coombs, C. E. Augarde, B. Sims, M. M. Puigvert and J. A. R. Gutiérrez, 'A conservation law consistent updated Lagrangian material point method for dynamic analysis,' *Journal of Computational Physics*, vol. 485, p. 112 075, 2023.

[60]   G. Pretti, W. M. Coombs, C. E. Augarde, M. M. Puigvert and J. A. R. Gutiérrez, 'Imposing constrained porosity values in a bi-phase elasto-plastic material under Terzaghi's effective stress principle,' *Submitted to Mechanics of Materials*, 2023.

Chapter 2

Overview of the Material Point Method (MPM)

The MPM was first conceived and published by Sulsky *et al.* [1, 2]. It originates from an extension of FLuid Implicit Particle (FLIP) [3, 4], which, in turn, is a variation of Particle In Cell (PIC) [5–8]. While PIC and FLIP[†] methods were successfully applied to fluids (see, for instance, [3–8]) and plasma (see, e.g., [9–11]) simulations, MPM's main application today is to solid mechanics.

The scope of this chapter is to provide an overview of the MPM under different points of view. The different kinds of MPM formulations are briefly introduced, explaining the advantages and disadvantages of these techniques. This subject is specifically addressed in Section 2.1. Regardless of the differences these techniques might present, tracing an essential structure of the MPM algorithm is possible. The different steps of the algorithm are discussed in Section 2.2 from a general point of view (i.e., without going into the mathematical details). Provided the main algorithm, it is more straightforward to understand the MPM's main strengths and challenges ahead. These are listed and motivated in Section 2.3.

Reproducing the vast literature regarding the MPM is a significant challenge beyond this work's scope. The publications considered in this chapter and the rest of this work are cited when considered helpful for explaining or directing the reader to further studies. However, recent publications of long articles (see Jiang *et al.* [12], de Vaucorbeil *et al.* [13] and Sołowski *et al.* [14]) and books (see Fern *et al.* [15] Nguyen *et al.* [16]) practically cover the published literature on the MPM so far.

2.1   MPM formulations

The original MPM (see [1, 2]) uses linear polynomial shape functions and assumes that Material Points (MPs) have no domain length, resulting in a rather robust method. How-

---

[†]When mentioned in a general context, PIC and FLIP are employed as a synecdoche, i.e., they indicate the whole numerical method. However, when these terms are used in the context of mapping procedures (see Section 3.6.6), they are related to specific formulae which operate the mapping of required information from the grid nodes to the particles.

ever, as transitions between adjacent elements are governed by non-smooth functions, gradients of shape functions abruptly change signs, causing the so-called cell-crossing instability and particularly affecting the stress response. The Generalised Interpolation Material Point Method (GIMPM)(see [17]) presents an improvement of the algorithm in this respect. Introducing a finite domain length and performing a convolution integral to compute the shape functions permits quadratic polynomials in the transition zones between adjacent elements, thus smoothening the gradients. The assumption of a domain of finite length questions how to update it during the analysis. The unchanged Generalised Interpolation Material Point Method (uGIMPM) stipulates that no update of the characteristic function should occur, while contiguous particle Generalised Interpolation Material Point Method (cpGIMPM) describes the evolution of this function as a rectangle or a cuboid. The uGIMPM has proved (see [36]) to be prone to the issue of domain overlapping or separation as the analysis proceeds, even under mono-dimensional loading. While free from this problem when the simulation is aligned to the axes, cpGIMPM cannot cope with the material rotating (see [36]). This last issue and the representation of the curved geometry with smooth edges is solved by the Convected Particle Domain Interpolation (CPDI) (see [28]), where MP domains can deform as parallelograms. This result is achieved by assuming basis functions based on a linear approximation between the standard linear polynomial shape function values at each corner. Despite this procedure and its consequential improvement, CPDI volumes are still susceptible to creating gaps or overlaps. The Second-order Convected Particle Domain Interpolation (CPDI2) resolves this issue by tracking the corners of MP domains, thus allowing the deformation into any quadrilateral shape. However, neither CPDI nor CPDI2 achieves second-order convergence (see Sołowski *et al.* [14]). This constraint fostered the use of higher-order shape functions. While the GIMPM increases the degree of the shape function locally to mitigate cell-crossing instability, B-splines functions achieve the same result by increasing the degree over the whole domain. However, only structured rectangular grids can be used if B-splines are considered. Regardless, the string of publications (see, for instance, [18–22]) in this area suggests that using B-splines for MPM is still a prolific field of investigation. The use of Powell-Sabin splines as basis functions overcomes this constraint (see [23]), thus allowing for the use of unstructured grids. Another approach designed to reduce the cell-crossing instability is the Dual Domain Material Point (DDMP) [24], which retains linear polynomial shape functions while modifying their gradient to guarantee continuity at the interface between adjacent elements. Despite this advantage, research (see [25]) has found that the DDMP requires more MPs to produce accurate results than other MPM formulations.

While the different types of MPM present benefits or weaknesses depending on the specific situation, this work considers only the original MPM and the GIMPM shape functions (see details in Sections 3.6.2 and 4.6.2).

2.2 MPM core algorithm

The workflow of an MPM algorithm is represented in Figure 2.1. Based on this figure, it can be seen how substeps on its Left-Hand Side (LHS) involve mainly the MPs, while substeps on the Right-Hand Side (RHS) are primarily related to the presence of a mesh (or grid). Substeps in the middle represent the interchange between the MPs and the grid. These two sets (mesh and MPs) represent the essential items upon which the MPM algorithm is structured. Broadly speaking, the MPs guarantee the simulation's continuity, while the mesh is used for calculation purposes. Depending on the ontology, the MPs and the grid can be seen as two different communicating discretisations or as a unique one in which the two groups play different roles*. Regardless of this dispute, the strength of the MPM lies in the presence and non-trivial communication between these two groups (MPs and grid).

Back to Figure 2.1, it can be seen that the salient points of an MPM algorithm are as follows:

- Introduction of the initial MPs setup;

- (A) Introduction of the computational grid;

- (B) Point-to-grid (P2G) mapping;

- (C) Primary equation assembling;

- (D) Determining the solution;

- (E) Grid-to-point (G2P) mapping;

- (F) Grid discharge;

- Production of post-processing data.

The substeps denoted by a letter in the above list (A) − (F) are required to complete an MPM (time-)step. The initial setup is given once for a simulation, while the production of post-processing data is not mandatory but useful to visualise and compare the results.

The subsections 2.2.1-2.2.8 of this chapter follow exactly the list above and Figure 2.1.

2.2.1 Initial MPs setup introduction

The starting point is the initial setup, which involves the initialisation of a cloud of MPs, representing the initial geometry of the considered body. Computational quantities and historical variables (such as stress, mass, elastic moduli, etc.) are initialised on the MPs that will carry these kinds of information, appropriately updated, throughout the simulation. Initial conditions on primary variables (such as displacements or velocities) can be specified

---

*This work adopts the two discretisations terminology, emphasising that discrete representations of the (Euclidean) space and the considered fields exist based on the MPs and the grid. However, the author acknowledges that this nomenclature is not shared by the entire community (see, for instance, Coombs and Augarde [26].).
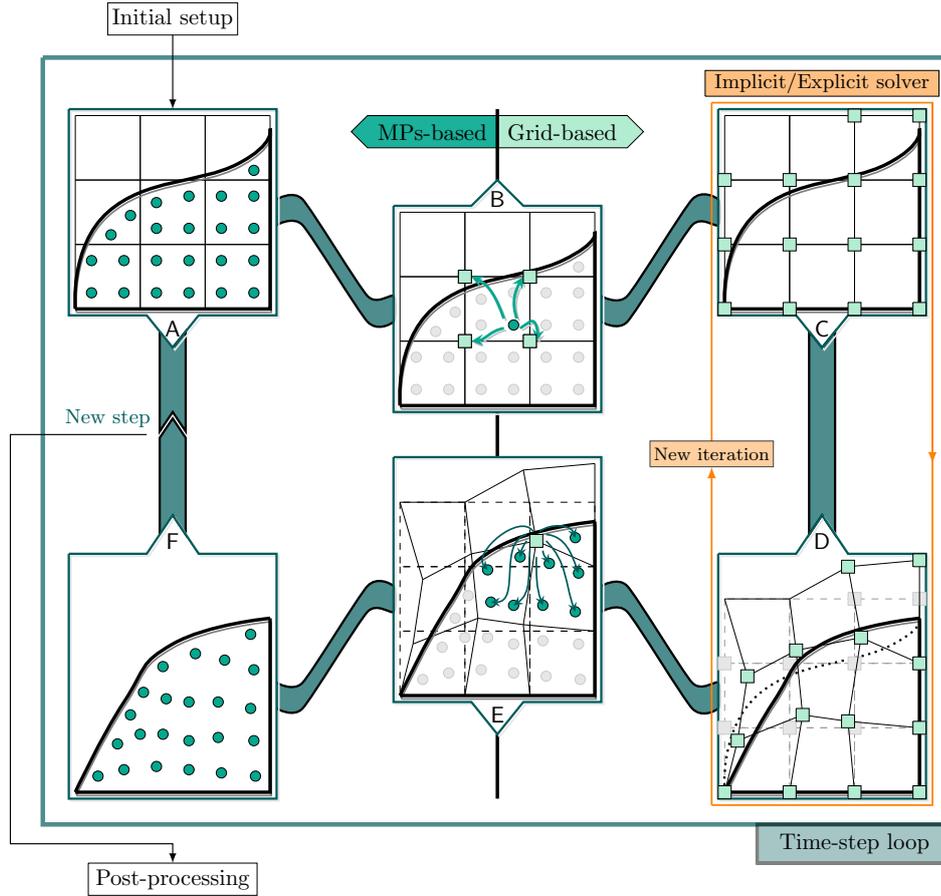
Figure 2.1: Illustration of an MPM algorithm. The initial setup is fed into the time-step loop, which is run until the simulation is complete at time T. On the left-hand side, there are the substeps involving the MPs, while, on the right-hand side, the substeps using (mainly) the grid-based calculations. A single time-step loop comprises several substeps: (A) immersion of the MPs into the computational grid, (B) P2G information mapping, (C) equilibrium equations formulation at the grid nodes, (D) primary equations solution, (E) G2P mapping and (F) MPs update and distorted grid disposal. (C) and (D) can be run multiple times in the case of an implicit solver. At the end of every time-step after (F), the required data are saved for post-processing purposes. Figure modified from Pretti *et al.* [27].

indiscriminately on the grid (yet to be introduced) or on the MPs. This indifference is due to substep (B) (see Section 2.2.3), which entails mapping this information from the MPs to the grid.

As previously mentioned, a critical input in any MPM algorithm is the domain characteristic length of the MPs, which differentiates the original MPM from the GIMPM (first proposed by Bardenhagen and Kober [17])*. In the former case, this length is condensed into a point (null in terms of width but representing a non-zero scalar value), whereas the characteristic domain is a length in all respects in the GIMPM. A few details on this difference and its consequences are given in Section 3.6.2. Since the continuum quantities are sampled at the

---

*This work does not consider other formulations than the original MPM or the GIMPM. For the *Convected Particle Domain Interpolation* (CPDI) or *Dual Domain* MPM (DDMPM) the reader is referred to their original papers (Sadeghirad *et al.* [28, 29] and Zhang *et al.* [24]). Other MPM variants, such as the *Moving Least Squares* (MLS) MPM [30] or the *B-splines* MPM [31] are occasionally mentioned in this work, but their treatment is not developed.

MPs' locations and averaged over the characteristic domains, the MP-based discretisation introduces an initial error.

The length of the simulation and its (initial) segmentation must also be specified. Adaptations of the time-step length are possible (see, in the MPM literature, [27, 32, 33], among others). Discussions on time discretisations are included in Sections 3.6.2.2, 4.6.2.2 and 5.3.2 of this work.

2.2.2   (A) **Introduction of the computational grid**

At the beginning of each time-step, the MP-based discretisation is immersed into a computational grid. This work adopts a regular Cartesian mesh, but this choice is arbitrary, as indicated by the use of different grids in other works. Among these, de Koster *et al.* [23] and Wang *et al.* [34] proposed grids based on simplex elements, while Gao *et al.* [35] presented a Cartesian mesh adaptive algorithm to coarsen or refine the mesh in specific locations.

The superposition of the grid onto the cloud of MPs gives rise to a convolution integral between the grid-defined function (satisfying the partition of unity property as well) and the characteristic function. The result of this integral defines the basis and the shape functions discretising the primary field variables and the geometry of the body. These shape functions inherit the partition of unity property from their original grid and characteristic functions. With the original MPM, this convolution integral leaves the grid functions unchanged. On the other hand, in the case of the GIMPM, a smoothening and an increase in the stencil of the grid-based functions is obtained. Accordingly, the Degrees of Freedom (DoFs) relative to a single MP are those of the element in which the MP is situated for the original MPM. The GIMPM stipulates that more DoFs may be active if the MP considered is close enough to the edge of the element, such that its domain interacts with multiple elements. Given the shape functions, their spatial derivatives are accordingly computed for each active node.

Grid introduction and relative assumptions are explored in Sections 3.6.2.1, 4.6.2.1 and 5.3.1 of this work.

2.2.3   (B) **Point-to-grid (P2G) mapping**

Having introduced the calculation grid, it is (sometimes) necessary to initialise the mesh to reproduce specific characteristics inherited from the MP-based discretisation. This transfer might be necessary since the equations are solved in incremental form and, as the grid is newly introduced, it might not be representative of what happened in the previous time-steps.

A first distinction regarding the mapping of information from MPs to the grid is the information itself that is required. The choice is not unique and depends on the type of problem under consideration, and the characteristics to be preserved during transfer. For example, if a quasi-static formulation of a solid material is considered, there is no need to map any

kind of information between the MPs and the grid at the beginning of each calculation step (see, for instance, Coombs and Augarde [26]).

Second, when required, the P2G mapping must be performed according to some specific criteria*, as the algorithm is desired to have some properties so that certain conservation features are guaranteed among different time-steps.

From a computational perspective, the P2G mapping is usually performed by looping over the MPs using the shape functions† to map the considered quantities from the MPs to the grid nodes.

This work rigorously assesses the mapping for a solid body in dynamic conditions in Section 3.6.6.1. Some less stringent considerations are introduced for a porous material in dynamic conditions in Section 4.6.5.1.

2.2.4   (C) **Primary equation assembling**

Having introduced both fundamental elements for an MPM algorithm (i.e., the MPs in Section 2.2.1 and the grid in 2.2.2) and appropriately initialised the mesh for the beginning of each time-step (Section 2.2.3), it is possible to evaluate the primary equations under consideration. The standard MPM formulation stipulates that these equations and their relative unknowns are assembled at the grid nodes, while the MPs act as integration points‡.

2.2.5   (D) **Determining the solution**

Given the equations, the nodal solution to these must consequently be found. Depending on the chosen solver (explicit versus implicit), the assembling process, substep (C), and the solution determination, substep (D), can be repeated once (explicit solver) or multiple times (implicit solver).

In the explicit case, the solution at the current time-step can be directly found by solving the assembled equations for the primary unknowns. This straightforward solution has historically favoured a matrix-free implementation of the explicit MPM. Depending on when the constitutive relationships at the MP-level are updated compared to the resolution of the primary equations, a distinction is made among the methods *Update Stress Last* (which, in the literature, goes under the acronym USL, see, for instance, [1]), *Modified Update Stress Last* (MUSL, see, for instance, [2]) and *Update Stress First* (USF, see, for instance, [37]). While attractive for some applications (see a comparison between explicit and implicit MPM

---

*Any quantity can always be naively mapped from the MPs' locations to the grid nodes and back. However, this approach introduces a diffusion that artificially dumps the simulation and scatters the mapped data unless adequately addressed (see, in this regard, the discussion in Brackbill *et al.* [4]).

†Most of the mappings provided in the literature use only shape functions to map quantities from MPs to the grid. An isolated case is the work of Wallstedt and Guilkey [36], where the derivatives of form functions are also used for P2G mapping. However, this aspect is not addressed further in this thesis.

‡The original MPM [1, 2] formulation proposed the use of the MPs as integration points. However, a few variants exist (see, for instance, Sulsky and Gong [30] or Gan *et al.* [21]) where MPs are used to store historical variables from which the considered integration points draw.

formulations in Sołowski *et al.* [14]), this work does not consider explicit MPM formulations. The reader can refer to Buzzi *et al.*'s work [38] comparing the above-mentioned explicit implementations available in the MPM.

In the implicit case, i.e., when implicit solvers are considered, the solution is iteratively found, and iterations are stopped when a specific arrest criterion is met. Both matrix-free (see Sulsky and Kaul [39]) and matrix-storing (see, for instance, Guilkey and Weiss [40]) variations have been tried for the implicit version of the MPM. Consistent linearisations are provided in Sections 3.6.4, 4.6.3, 5.4.1, and Appendix B for the primary equations under considerations. A linearisation of a specific non-linear stress-strain relationship is provided by Appendix C.

2.2.6 (E) **Grid-to-point (G2P) mapping**

Having obtained the nodal solution to the primary equations, this information must be transferred to the MPs before the grid is discharged. Contrary to the P2G mapping, this G2P mapping is always necessary since MPs are responsible for providing continuity across the different time-steps in a simulation. If this G2P mapping did not take place, the nodal solution obtained at the current time-step would be irretrievably lost. In the literature, it has been often pointed out that the number of MPs is usually higher than that of grid nodes, implying that the P2G mapping constitutes the bottleneck (or filter) in the simulation (see Burgess *et al.* [41], Wallstedt and Guilkey [36] and Hammerquist and Nairn [42]). However, the G2P mapping, while representing the same challenges in terms of properties desired to be conserved, also represents an opportunity to remove the effects of the P2G mapping.

Similarly to the P2G initial mapping, the G2P mapping is operated by the shape functions[*], moving the nodal solution back to the MPs' locations. This mapping is rigorously assessed in Section 3.6.6.2 for a solid material in dynamic conditions, while Section 4.6.5.2 briefly and less stringently summarises the procedure for a porous material.

As previously described, based on the updated values of the solution at the MPs' location, the characteristic domain is amended to be representative of the strain state, if cpGIMPM is considered. Otherwise, as the uGIMPM dictates, volume are not updated. This work considers the cpGIMPM and updates it according to the method proposed by Charlton *et al.* [44][†].

2.2.7 (F) **Grid discharge**

Having moved the nodal solution from the grid to the MPs at substep (E), the grid happens to be partially distorted and essentially superfluous, since all relevant pieces of information

---

[*]In the so-called Polynomial Particle In Cell (PolyPIC) (proposed by Fu *et al.* [43]), *ah hoc* polynomial functions (different from the ones used in the P2G mapping and in the assembling the primary equations at the grid nodes) are built to perform the G2P mapping.

[†]Even in the case of cpGIMPM, there exist several ways to update the MP characteristic domain (the reader can refer to Coombs *et al.* [45] for a comparison and implementational details).

now lie at the MPs' location. Hence, it is usual in an MPM algorithm to dispose the grid before starting a new time-step. In this fashion, one of the MPM's main advantages, namely being effective with large deformation and rotations, can be fully exploited, as a significant mesh distortion cannot take place. Furthermore, even though the current substep is not mandatory, its absence would undermine substeps (A), (B), and (E), and the whole time-step loop would resemble a Finite Element Method (FEM) with a poor choice of the quadrature scheme.

2.2.8 Production of post-processing data

While not part of MPM's essential substeps, the production of data for post-processing is an ordinary part of any numerical method. As the grid supports only the calculation phase in the MPM, it is routine to post-process data only at the MPs' level. Nevertheless, for debugging purposes, for example, nothing prevents the grid data from being displayed. If this is the case, depending on whether these data are considered per every time-step or every iteration (in the case of implicit formulations), this substep should be moved before (F) in the former case and after (D) in the latter.

2.3 MPM advantages and challenges

The different roles played by the grid (where the equations are assembled) and the MPs (integration point and historical data storage) and their interaction constitute, possibly, the ultimate essence of the MPM. Consequently, these two objects (i.e., the grid and the MPs) dictate not only the MPM's main advantages but also its weaknesses, which constitute a challenge for the MPM community.

Among the strengths, it is worthy of mentioning:

S.1 the primary equations are solved on a grid, avoiding all of the typical issues of meshless methods (e.g., neighbouring particles search);

S.2 MPs store historical variables guaranteeing their continuity (with no diffusion) during the simulation;

S.3 when coexisting, the grid and the MPs move and deform without permitting relative displacement, i.e., the equations are Lagrangian;

S.4 mesh distortion can be avoided by its disposal at the end of the time-step;

S.5 an automatic no-slip contact algorithm occurring at the grid level is pre-built in the method;

S.6 convertibility between voxel-based images and the MP-based discretisation can be straightforward (no body fitting is required);

S.7 MPM can be largely parallelisable; and

S.8 similarities shared with the FEM permit to incorporate the knowledge developed for the FEM into the MPM.

As for the MPM challenges, these can be considered:

C.1 cell-crossing instability;

C.2 more significant integration error compared to the FEM (which might inhibit higher-order convergence) and small-cut instability;

C.3 absence of an error estimate (causing difficulty in adaptive techniques);

C.4 treatment of phenomena considering the boundaries (e.g.,Boundary Conditions (BCs));

C.5 volumetric locking for nearly incompressible materials;

C.6 issues in moving quantities from the sampling point without considerable diffusion (i.e., mapping-relating problems);

C.7 null-space error;

C.8 higher (compared to the FEM) computational costs; and

C.9 drop in the Courant-Friedrichs-Lewy (CFL) condition due to partially filled elements;

In particular, challenges **C.1**, **C.2**, **C.4**, **C.6**, and **C.9** are mentioned or addressed throughout this work.

—— *Chapter summary*

This chapter has given a synopsis of the MPM, describing its different formulations. The MPM core algorithm has been qualitatively described, and a list of advantages and drawbacks of the method has been provided.

References

[1] D. Sulsky, Z. Chen and H. Schreyer, 'A particle method for history-dependent materials,' *Computer Methods in Applied Mechanics and Engineering*, vol. 118, no. 1-2, pp. 179–196, 1994.

[2] D. Sulsky, S.-J. Zhou and H. L. Schreyer, 'Application of a particle-in-cell method to solid mechanics,' *Computer Physics Communications*, vol. 87, no. 1-2, pp. 236–252, 1995.

[3] J. U. Brackbill and H. M. Ruppel, 'FLIP: A method for adaptively zoned, particle-in-cell calculations of fluid flows in two dimensions,' *Journal of Computational Physics*, vol. 65, no. 2, pp. 314–343, 1986.

[4] J. U. Brackbill, D. B. Kothe and H. M. Ruppel, 'Flip: A low-dissipation, particle-in-cell method for fluid flow,' *Computer Physics Communications*, vol. 48, no. 1, pp. 25–38, 1988.

[5] M. W. Evans, F. H. Harlow and E. Bromberg, 'The particle-in-cell method for hydrodynamic calculations,' Los Alamos National Lab NM, Tech. Rep., 1957.

[6] F. H. Harlow, 'Hydrodynamic problems involving large fluid distortions,' *Journal of the ACM*, vol. 4, no. 2, pp. 137–142, 1957.

[7] F. H. Harlow, D. O. Dickman, D. E. Harris Jr and R. E. Martin, 'Two-dimensional hydrodynamic calculations,' Apr. 1959.

[8] F. H. Harlow, 'The particle-in-cell computing method for fluid dynamics,' *Methods in Computational Physics*, vol. 3, pp. 319–343, 1964.

[9] T. Tajima, J. Leboeuf and J. Dawson, 'A magnetohydrodynamic particle code with force free electrons for fluid simulations,' *Journal of Computational Physics*, vol. 38, no. 2, pp. 237–250, 1980.

[10] J. Brackbill, 'Flip mhd: A particle-in-cell method for magnetohydrodynamics,' *Journal of Computational Physics*, vol. 96, no. 1, pp. 163–192, 1991.

[11] V. Olshevsky, F. Bacchini, S. Poedts and G. Lapenta, 'Slurm: Fluid particle-in-cell code for plasma modeling,' *Computer Physics Communications*, vol. 235, pp. 16–24, 2019.

[12] C. Jiang, C. Schroeder, J. Teran, A. Stomakhin and A. Selle, 'The material point method for simulating continuum materials,' in *ACM siggraph 2016 courses*, 2016, pp. 1–52.

[13] A. de Vaucorbeil, V. P. Nguyen, S. Sinaie and J. Y. Wu, 'Material point method after 25 years: Theory, implementation, and applications,' in *Advances in Applied Mechanics*, vol. 53, Elsevier, 2020, pp. 185–398.

[14] W. T. Sołowski, M. Berzins, W. M. Coombs, J. E. Guilkey, M. Möller, Q. A. Tran, T. Adibaskoro, S. Seyedan, R. Tielen and K. Soga, 'Material point method: Overview and challenges ahead,' *Advances in Applied Mechanics*, vol. 54, pp. 113–204, 2021.

[15] J. Fern, A. Rohe, K. Soga and E. Alonso, *The material point method for geotechnical engineering: a practical guide*. CRC Press, 2019.

[16] V. P. Nguyen, A. de Vaucorbeil and S. Bordas, *The Material Point Method: Theory, Implementations and Applications*. Springer Nature, 2023.

[17] S. G. Bardenhagen and E. M. Kober, 'The generalized interpolation material point method,' *Computer Modeling in Engineering and Sciences*, vol. 5, no. 6, pp. 477–496, 2004.

[18] S. Andersen and L. Andersen, 'Analysis of spatial interpolation in the material-point method,' *Computers & Structures*, vol. 88, no. 7-8, pp. 506–518, 2010.

[19] Y. G. Motlagh and W. M. Coombs, 'An implicit high-order material point method,' *Procedia Engineering*, vol. 175, pp. 8–13, 2017.

[20] R. Tielen, E. Wobbes, M. Möller and L. Beuth, 'A high order material point method,' *Procedia Engineering*, vol. 175, pp. 265–272, 2017.

[21] Y. Gan, Z. Sun, Z. Chen, X. Zhang and Y. Liu, 'Enhancement of the material point method using B-spline basis functions,' *International Journal for Numerical Methods in Engineering*, vol. 113, no. 3, pp. 411–431, 2018.

[22] Y. Yamaguchi, S. Moriguchi and K. Terada, 'Extended B-spline-based implicit material point method,' *International Journal for Numerical Methods in Engineering*, vol. 122, no. 7, pp. 1746–1769, 2021.

[23] P. de Koster, R. Tielen, E. Wobbes and M. Möller, 'Extension of B-spline material point method for unstructured triangular grids using powell–sabin splines,' *Computational Particle Mechanics*, vol. 8, no. 2, pp. 273–288, 2021.

[24] D. Z. Zhang, X. Ma and P. T. Giguere, 'Material point method enhanced by modified gradient of shape function,' *Journal of Computational Physics*, vol. 230, no. 16, pp. 6379–6398, 2011.

[25] T. R. Dhakal and D. Z. Zhang, 'Material point methods applied to one-dimensional shock waves and dual domain material point method with sub-points,' *Journal of Computational Physics*, vol. 325, pp. 301–313, 2016.

[26] W. M. Coombs and C. E. Augarde, 'AMPLE: A material point learning environment,' *Advances in Engineering Software*, vol. 139, p. 102 748, 2020.

[27] G. Pretti, W. M. Coombs, C. E. Augarde, B. Sims, M. M. Puigvert and J. A. R. Gutiérrez, 'A conservation law consistent updated Lagrangian material point method for dynamic analysis,' *Journal of Computational Physics*, vol. 485, p. 112 075, 2023.

[28] A. Sadeghirad, R. M. Brannon and J. Burghardt, 'A convected particle domain interpolation technique to extend applicability of the material point method for problems involving massive deformations,' *International Journal for Numerical Methods in Engineering*, vol. 86, no. 12, pp. 1435–1456, 2011.

[29] A. Sadeghirad, R. M. Brannon and J. Guilkey, 'Second-order convected particle domain interpolation (CPDI2) with enrichment for weak discontinuities at material interfaces,' *International Journal for Numerical Methods in Engineering*, vol. 95, no. 11, pp. 928–952, 2013.

[30] D. Sulsky and M. Gong, 'Improving the material-point method,' *Innovative Numerical Approaches for Multi-Field and Multi-Scale Problems: In Honor of Michael Ortiz's 60th Birthday*, pp. 217–240, 2016.

[31] M. Steffen, R. M. Kirby and M. Berzins, 'Analysis and reduction of quadrature errors in the material point method (MPM),' *International Journal for Numerical Methods in Engineering*, vol. 76, no. 6, pp. 922–948, 2008.

[32] Y. Fang, Y. Hu, S.-M. Hu and C. Jiang, 'A temporally adaptive material point method with regional time stepping,' in *Computer Graphics Forum*, Wiley Online Library, vol. 37, 2018, pp. 195–204.

[33] Y. Sun, T. Shinar and C. Schroeder, 'Effective time step restrictions for explicit MPM simulation,' in *Computer Graphics Forum*, Wiley Online Library, vol. 39, 2020, pp. 55–67.

[34] L. Wang, W. M. Coombs, C. E. Augarde, M. Cortis, M. J. Brown, A. J. Brennan, J. A. Knappett, C. Davidson, D. Richards and D. J. White, 'An efficient and locking-free material point method for three-dimensional analysis with simplex elements,' *International Journal for Numerical Methods in Engineering*, vol. 122, no. 15, pp. 3876–3899, 2021.

[35] M. Gao, A. P. Tampubolon, C. Jiang and E. Sifakis, 'An adaptive generalized interpolation material point method for simulating elastoplastic materials,' *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–12, 2017.

[36] P. Wallstedt and J. Guilkey, 'Improved velocity projection for the material point method,' *Computer Modeling in Engineering and Sciences*, vol. 19, no. 3, p. 223, 2007.

[37] S. Bardenhagen, 'Energy conservation error in the material point method for solid mechanics,' *Journal of Computational Physics*, vol. 180, no. 1, pp. 383–403, 2002.

[38] O. Buzzi, D. M. Pedroso, A. Giacomini *et al.*, 'Caveats on the implementation of the generalized material point method,' *Computer Modeling in Engineering and Sciences*, vol. 1, no. 1, pp. 1–21, 2008.

[39] D. Sulsky and A. Kaul, 'Implicit dynamics in the material-point method,' *Computer Methods in Applied Mechanics and Engineering*, vol. 193, no. 12-14, pp. 1137–1170, 2004.

[40] J. E. Guilkey and J. A. Weiss, 'Implicit time integration for the material point method: Quantitative and algorithmic comparisons with the finite element method,' *International Journal for Numerical Methods in Engineering*, vol. 57, no. 9, pp. 1323–1338, 2003.

[41] D. Burgess, D. Sulsky and J. Brackbill, 'Mass matrix formulation of the FLIP particle-in-cell method,' *Journal of Computational Physics*, vol. 103, no. 1, pp. 1–15, 1992.

[42] C. C. Hammerquist and J. A. Nairn, 'A new method for material point method particle updates that reduces noise and enhances stability,' *Computer Methods in Applied Mechanics and Engineering*, vol. 318, pp. 724–738, 2017.

[43] C. Fu, Q. Guo, T. Gast, C. Jiang and J. Teran, 'A polynomial particle-in-cell method,' *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–12, 2017.

[44] T. Charlton, W. Coombs and C. Augarde, 'IGIMP: An implicit generalised interpolation material point method for large deformations,' *Computers & Structures*, vol. 190, pp. 108–125, 2017.

[45] W. M. Coombs, C. E. Augarde, A. J. Brennan, M. J. Brown, T. J. Charlton, J. A. Knappett, Y. G. Motlagh and L. Wang, 'On Lagrangian mechanics and the implicit material point method for large deformation elasto-plasticity,' *Computer Methods in Applied Mechanics and Engineering*, vol. 358, p. 112 622, 2020.

Chapter 3

Dynamics of solid materials

When approaching numerical simulations (but not limited to these) to model a specific situation, it is pretty natural to wonder which phenomena play a major role, so that the produced algorithm can represent the given circumstances with a good margin of acceptability. If considering the anchor installation process, the speed of the anchor (between 1 and 5 knots[†], as reported by Grabe and Wu [1]) and the masses of the system involved (i.e., mass of the anchor and the displaced soil) are particularly worthy of attention. It follows that installing an anchor cannot be regarded as a *quasi-static* process since the involved inertia forces seem anything but secondary. Furthermore, if not devoting attention to some specific parts of the algorithm, the anchor-soil system's kinetic energy can be lost via spurious undesired numerical dissipation. This chapter, therefore, serves a threefold purpose: the first consists of studying the equations governing a deformable continuous body in the context of *dynamics*, while designing (second goal) an algorithm that dissipates energy according to the second principle of thermodynamics; the third is familiarising with finite strain mechanics (to maximise one of the strengths of the Material Point Method (MPM)).

Given these purposes, it seems rather logical to introduce the first hypotheses (labelled with the letter **H** and progressive numbers) upon which this chapter is built, namely:

H.3.1 the solid body $\mathcal{B}$ is considered at a scale to which it appears to be continuous (as required by the MPM), consisting in an infinite number of material particles. The body is a subset of the Euclidean space $\mathcal{E}$, and each particle occupies a position in the body subset; and

H.3.2 the material under consideration undergoes finite strains, i.e., there is a clear distinction between the original and the current configuration.

Sections 3.1-3.5 of this chapter represents classical subjects in continuum mechanics, which include finite strain (elasto-plastic) kinematics 3.1, mass conservation for a solid body 3.2, balance of momenta 3.3, basic thermodynamics principles 3.4 and their application to the study of constitutive relationships 3.5. Section 3.6 discusses the discretisation of the equa-

---

[†]A speed between 1 and 5 knots corresponds to 0.5 and 2.5 m s^{-1}.

tions given in Sections 3.1-3.5 within the framework of an MPM algorithm. Particular attention is paid to ensuring that such an algorithm respects the aforementioned relationships outlined in Sections 3.1-3.5. Details of how to devise stable and rigorous MPM algorithms are also given in Section 3.6, focusing on the primary grid-based equations and the mapping procedures. Numerical simulations and critical assessment concerning both the entire chapter and the MPM for solid dynamics are given in Sections 3.7 and 3.8.

To keep the discussion as concise as possible, it was not possible on the author's side to deepen most of the topics with sufficient diligence. The reader interested in learning more about solid continuum mechanics in the context of finite strains for dynamic analyses can refer to the works of Oden [2], Ogden [3] and Gurtin *et al.* [4], amongst others. These references, together with de Souza Neto *et al.* [5], are the main inspirations for Sections 3.1-3.5.

3.1 Kinematics

The first stage that is classically introduced for the description of deformable solid bodies is kinematics, i.e. how such a body moves and deforms in space in different instants of the time.

3.1.1 Motion, displacement and material time derivatives

Let the deformable body $\mathcal{B}$ occupy an initial volume Ω (part of the Euclidean space $\mathcal{E}$) with boundary Γ in the reference configuration (i.e., at the time $t = 0$). As is traditionally the case for solid mechanics, this chapter is entirely based on a *Lagrangian* description. While a possible definition of the Lagrangian description is reported in the next Chapter, Section 4.1.1, it is here sufficient to recall that its main interest lies in tracking particles belonging to the body $\mathcal{B}$ in position and time. For this purpose, let a generic particle in the original configuration being defined by its position $\boldsymbol{X}$. Let the *motion* being a function $\boldsymbol{\varphi} : \Omega \times [0, T] \to \mathcal{E}$ such that the particle initially in $\boldsymbol{X}$ is defined by a position $\boldsymbol{x} = \boldsymbol{\varphi}\left(\boldsymbol{X}, t\right)$ in the current configuration at the time t. In regards to the motion, let us assume that:

H.3.3 $\boldsymbol{\varphi}$ is a smooth* function; and

H.3.4 the mapping is *admissible*, meaning that it is invertible.

The definition of motion allows also to define each point in the current configuration and, as a consequence, the whole current volume of the body can be described by the motion. Let $\omega = \boldsymbol{\varphi}\left(\Omega, t\right)$ denote the volume occupied by the body $\mathcal{B}$ in the current configuration, while γ is the current boundary of this volume.

---

*The term smooth is assumed to avoid a mathematical statement on the regularity of the function. This is a terminology which is widely used in continuum mechanics (see, for instance, de Souza Neto *et al.* [5]), and implies that all the manipulations applied to the considered smooth functions exist and are properly defined.
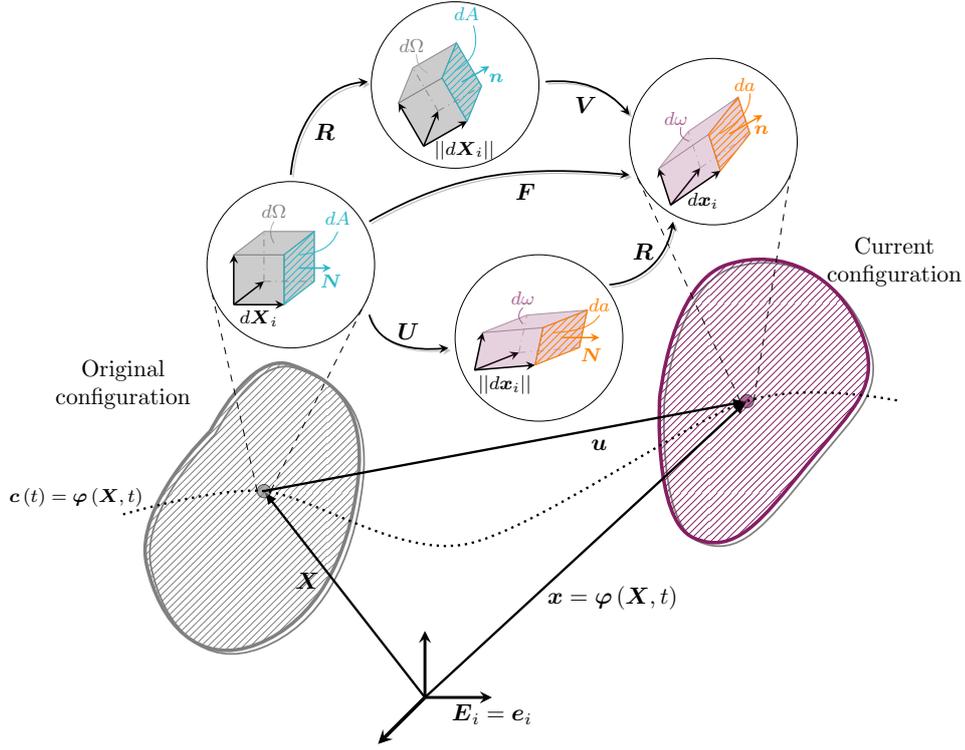
Figure 3.1: Illustration of initial and current configurations. The zoomed regions describe the different roles played in the polar decomposition (3.12). The curve $c\,(t) = \varphi\,(X, t)$ defines the *trajectory* of the particle initially in X.

The description of the two introduced positions X and x can take place only via some base vectors. To this end, let us introduce two *fixed Cartesian* coordinate systems, one in the reference configuration (with base vectors E_i), and one in the current configuration (with base vectors e_i) so that

$$X = X_i\,E_i, \qquad x = x_i\,e_i, \tag{3.1}$$

with $i = 1 \ldots n^{dim}$ being the dimensions of the Euclidean space. Practically, these coordinates systems have coincident base vectors, but considering them as separate helps to understand the description of some quantities which will be considered below*.

Owing to **H.3.4** the motion φ must describe a one-to-one function, implying that its inverse can be computed, i.e., if it exists the mapping

$$x = \varphi\,(X, t)\,, \tag{3.2}$$

then its inverse exists too

$$X = \varphi^{-1}\,(x, t)\,. \tag{3.3}$$

From a physical perspective, an invertible mapping prohibits material interpenetration.

The presence of the original and the current configurations defines two different Lagrangian formulations, these being the Total Lagrangian (TL) and the Updated Lagrangian (UL) formulations. In Belytschko's *et al.*'s [6] words

---

*For the rest of this work, when not explicit stated, the Cartesian coincident base vectors are implied.

``... the major difference between the two formulations is in the
point of view: in the total Lagrangian formulation variables are
described in the original configuration, in the updated Lagrangian
formulation in the current configuration.''

It must also be noted that, for the continuum case, the difference between the two formulations is superficial[*], since they are both fully defined by the mapping φ, as it is presumed for Lagrangian descriptions.

The introduction of the motion defining the position x of the particle initially in X allows also the introduction of the *displacement* field, defined by the vector

$$u\left(X,t\right) := \varphi\left(X,t\right) - X. \tag{3.4}$$

In turn, the *velocity* and *acceleration* of the particle initially in X are defined by the first and second the material derivatives[†] of the displacement field, i.e.,

$$v\left(X,t\right) := \frac{du\left(X,t\right)}{dt}; \tag{3.5}$$

$$a\left(X,t\right) := \frac{d^2 u\left(X,t\right)}{dt^2} = \frac{dv\left(X,t\right)}{dt}, \tag{3.6}$$

where it is assumed that the vague assumption **H.3.3** allows for the existence of the above derivatives. In this chapter, the material derivatives $\frac{d(\bullet)}{dt}$ of the general quantity $(\bullet)$ are also often concisely denoted as $(\dot{\bullet})$.

3.1.2 Deformation gradient

While the previous section dealt with the motion of a particle within a deformable body, no quantity introduced so far actually describes the deformation of the body under consideration. For this purpose, the most fundamental quantity is the *deformation gradient F* (see Figure 3.1 for a condensed representation of the deformation gradient and some of the related quantities of interest), defined as

$$F := \frac{\partial \varphi\left(X,t\right)}{\partial X}, \tag{3.7}$$

where, once again, it is assumed that **H.3.3** guarantees the existence of the above derivative. The deformation gradient is responsible for mapping infinitesimal line elements in the original configuration dX to its current counterpart dx. This is performed by using Eqs. (3.2) and (3.7) and ignoring higher order terms, i.e.,

$$dx = \varphi\left(X + dX,t\right) - \varphi\left(X,t\right) = F dX. \tag{3.8}$$

---

[*]In the discretised case for the MPM, as it will be shown in Section 3.6.2, the two formulations are still communicating, even if the TL appear to be, in some sense, more artificial then the UL.

[†]Privileging a Lagrangian description, all the time derivatives appearing in this chapter are material derivatives, in the sense that the time derivation is carried on with respect to all of the arguments of the considered function.

Since the deformation gradient is responsible for mapping the infinitesimal line from one configuration to the other, it can straightforwardly be used to compute the change of the original infinitesimal volume $d\Omega$ into the current one $d\omega$

$$
\begin{aligned}
d\omega &= \left((dx^{(1)})_i \, (dx^{(2)})_j (dx^{(3)})_k \right) \epsilon_{ijk} \\
&= \left(F_{i\alpha} \, (dX^{(1)})_\alpha \, F_{j\beta} \, (dX^{(2)})_\beta \, F_{k\gamma} (dX^{(3)})_\gamma \right) \epsilon_{ijk} \\
&= det(\boldsymbol{F}) \epsilon_{\alpha\beta\gamma} (dX^{(1)})_\alpha \, (dX^{(2)})_\beta (dX^{(3)})_\gamma \\
&= \underbrace{det(\boldsymbol{F})}_{:=J>0} \, d\Omega,
\end{aligned}
\tag{3.9}
$$

where J defines the so-called *Jacobian* and $\epsilon_{\alpha\beta\gamma}$ is the *Levi-Civita symbol*. From a mathematical perspective, the Jacobian must differ from zero to allow the deformation gradient to be invertible (as highlighted by Eq. (3.3)), while, physically speaking, the Jacobian defines a relationship between two infinitesimal volumes, which excludes negative values.

The deformation gradient is also useful when computing the change between the original and the current configuration of the outward normal attached to an infinitesimal surface. Denoting $\boldsymbol{n}$ ($\boldsymbol{N}$) and da (dA) the current (original) normal and infinitesimal area, it follows that

$$
n_i \, da = (dx^{(1)})_j \, (dx^{(2)})_k \, \epsilon_{ijk} = F_{j\alpha} \, (dX^{(1)})_\alpha \, F_{k\beta} (dX^{(2)})_\beta \, \epsilon_{ijk},
\tag{3.10}
$$

which, when rearranged, goes in the literature under the name of *Nanson's formula*, i.e.,

$$
\begin{aligned}
F_{i\gamma} \, n_i \, da &= F_{j\alpha} \, (dX^{(1)})_\alpha \, F_{k\beta} (dX^{(2)})_\beta \, \epsilon_{ijk} \, F_{i\gamma} \\
&= J \, \epsilon_{\alpha\beta\gamma} \, (dX^{(1)})_\alpha \, (dX^{(2)})_\beta \\
&= J \, N_\gamma \, dA.
\end{aligned}
\tag{3.11}
$$

3.1.3 Stretch and strain measures

3.1.3.1 Polar decompositions

The deformation gradient $\boldsymbol{F}$ include two pieces of information, i.e., it expresses how much the infinitesimal line is stretched and how much it (rigidly) rotates with respect to its original configuration. These two different aspects can be captured by applying the polar decomposition theorem[*] to the deformation gradient, i.e.,

$$
\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} = \boldsymbol{V}\boldsymbol{R},
\tag{3.12}
$$

where the rotation $\boldsymbol{R}$ is the *proper orthogonal* rotation tensor[†], while $\boldsymbol{U}$ and $\boldsymbol{V}$ are the symmetric *right* and *left stretch tensors*. It can be seen that the relationship between the left and right stretch tensors is provided by Eq. (3.12) itself, i.e.,

$$
\boldsymbol{V} = \boldsymbol{R}\boldsymbol{U}\boldsymbol{R}^T, \qquad \boldsymbol{U} = \boldsymbol{R}^T \boldsymbol{V}\boldsymbol{R}.
\tag{3.13}
$$

---

[*]The existence of the *polar decomposition theorem* to the deformation gradient $\boldsymbol{F}$ is guaranteed by $J > 0$.

[†]The definition of an *orthogonal* tensor dictates that $\boldsymbol{R}^T = \boldsymbol{R}^{-1}$. From the definition of orthogonality it follows that $\det \boldsymbol{R} = \pm 1$. A *proper* tensor in an orthogonal tensor defined by $\det \boldsymbol{R} = +1$.

Using the algebra introduced outlined in Appendix A, it can also be shown that $\boldsymbol{U}$ and $\boldsymbol{V}$ share the same eigenvalues, named the *principal stretches* and denoted by $(\lambda_1, \lambda_2, \lambda_3)$, see Eq. (A.4). The eigenvectors of $\boldsymbol{U}$ and $\boldsymbol{V}$ (named the *principal directions*) are related via the rotation $\boldsymbol{R}$, which can be straightforwardly shown. To this end, let $\hat{\boldsymbol{N}}_i$ and $\hat{\boldsymbol{n}}_i$ denote the eigenvectors of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively. Pre-multiplying by $\boldsymbol{R}$ the eigen-decomposition problem for $\boldsymbol{U}$, the eigen-decomposition for $\boldsymbol{V}$ follows, i.e.,

$$\boldsymbol{R}\left(\boldsymbol{U} - \lambda_i \boldsymbol{I}^{(2)}\right)\hat{\boldsymbol{N}}_i = \left(\boldsymbol{R}\boldsymbol{U} - \lambda_i \boldsymbol{R}\right)\hat{\boldsymbol{N}}_i = \left(\boldsymbol{V}\boldsymbol{R} - \lambda_i \boldsymbol{R}\right)\hat{\boldsymbol{N}}_i =$$
$$= \left(\boldsymbol{V} - \lambda_i \boldsymbol{I}^{(2)}\right)\underbrace{\boldsymbol{R}\hat{\boldsymbol{N}}_i}_{=\hat{\boldsymbol{n}}_i} = \boldsymbol{0} \quad (3.14)$$

The above equations show also that the rotation tensor is responsible for mapping the original principal axes to their current counterpart, i.e.,

$$\hat{\boldsymbol{n}}_i = \boldsymbol{R}\,\hat{\boldsymbol{N}}_i. \quad (3.15)$$

The tensor product of the above equation with $\hat{\boldsymbol{N}}_j$ gives

$$\boldsymbol{R} = \hat{\boldsymbol{n}}_i \otimes \hat{\boldsymbol{N}}_j, \quad (3.16)$$

which is justified if $\hat{\boldsymbol{N}}_i \otimes \hat{\boldsymbol{N}}_j = \delta_{ij}$, i.e., the base vectors are orthogonal (compare with Eq. (A.7)). Eqs. (3.14) and (3.16) yield to

$$\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} = \boldsymbol{R}\sum_{i=1}^{3}\lambda_i \hat{\boldsymbol{N}}_i \otimes \hat{\boldsymbol{N}}_i = \sum_{i=1}^{3}\lambda_i \hat{\boldsymbol{n}}_i \otimes \hat{\boldsymbol{N}}_i, \quad (3.17)$$

meaning that, for both $\boldsymbol{R}$ and $\boldsymbol{F}$, the asymmetric character of the tensor is driven by their representation in different (original and current) base vectors. In the literature, this feature is often referred to as *two-point* character of a tensor (see, for instance, Marsden and Hughes [7]), implying that the tensor is not entirely defined in one configuration.

3.1.3.2 Strain tensors

As stated above, the deformation gradient describes the change of a line segment in direction and magnitude. If only this latter feature is considered (i.e., the magnitude), the change in the length of the infinitesimal line element is given by

$$dl = \sqrt{dx_i\,dx_i} = \sqrt{(F_{i\alpha}dX_\alpha)(F_{i\beta}dX_\beta)} = \left(dX_\alpha\,\underbrace{F_{i\alpha}\,F_{i\beta}}_{:=C_{\alpha\beta}}\,dX_\beta\right)^{\frac{1}{2}} \quad (3.18)$$

where the symmetric *right Cauchy-Green strain* $\boldsymbol{C}$ naturally appears in the above equation. This tensor possesses some useful properties. In particular, $\boldsymbol{C}$ (as well as the tensors originating from it) is defined on the initial configuration, as it can be seen by making its base vectors explicit, i.e.,

$$\boldsymbol{C} = \frac{\partial x_k}{\partial X_i}\,\frac{\partial x_k}{\partial X_j}\,\boldsymbol{E}_i \otimes \boldsymbol{E}_j. \quad (3.19)$$

It can also be demonstrated that $\boldsymbol{C}$ is symmetric and does not contain any rotation. For this purpose, let us write the following chain of equations

$$\boldsymbol{C} = \boldsymbol{F}^T \boldsymbol{F} = \boldsymbol{U}^T \boldsymbol{R}^T \boldsymbol{R} \boldsymbol{U} = \boldsymbol{U} \boldsymbol{U} = \boldsymbol{U}^2, \tag{3.20}$$

where the symmetry of $\boldsymbol{U}$ and the properties of the rotation $\boldsymbol{R}$ are exploited. The independence of $\boldsymbol{C}$ from the rotations could also be physically understood by Eq. (3.18), as the change in length of a line element cannot depend on rotations.

Since the deformation gradient can be inverted (as assumed in **H.3.4**), it could be worthy of investigation considering also the inverse of Eq. (3.18), i.e.,

$$dL = \sqrt{dX_i \, dX_i} = \sqrt{\left(F_{i\alpha}^{-1} dx_\alpha\right)\left(F_{i\beta}^{-1} dx_\beta\right)} = \left(dx_\alpha \underbrace{F_{i\alpha}^{-1} F_{i\beta}^{-1}}_{:=b_{\alpha\beta}^{-1}} dx_\beta\right)^{\frac{1}{2}}, \tag{3.21}$$

where, in this case, the (inverse of the) *left Cauchy-Green strain* $\boldsymbol{b} := \boldsymbol{F} \boldsymbol{F}^T$ appears naturally in the above equation. Following the same steps taken for its right counterpart, it can be seen that $\boldsymbol{b}$ is a defined in the current configuration

$$\boldsymbol{b} = \frac{\partial x_i}{\partial X_k} \frac{\partial x_j}{\partial X_k} \boldsymbol{e}_i \otimes \boldsymbol{e}_j \tag{3.22}$$

As for $\boldsymbol{C}$, the left Cauchy-Green strain tensor $\boldsymbol{b}$ is symmetric and rotation free

$$\boldsymbol{b} = \boldsymbol{V} \boldsymbol{R} \boldsymbol{R}^T \boldsymbol{V}^T = \boldsymbol{V} \boldsymbol{V} = \boldsymbol{V}^2. \tag{3.23}$$

Based on $\boldsymbol{U}$ and $\boldsymbol{V}$, two families of strain tensors can be defined on the initial and the current configuration, these being

$$\boldsymbol{E}^{(m)} := \begin{cases} \frac{1}{m}\left(\boldsymbol{U}^m - \boldsymbol{I}^{(2)}\right) = \sum_{i=1}^3 \frac{1}{m}\left(\lambda^m - 1\right) \hat{\boldsymbol{N}}_i \otimes \hat{\boldsymbol{N}}_i & \text{if} \quad m \neq 0; \\ \ln\left(\boldsymbol{U}\right) = \sum_{i=1}^3 \ln \lambda_i \, \hat{\boldsymbol{N}}_i \otimes \hat{\boldsymbol{N}}_i & \text{if} \quad m = 0, \end{cases} \tag{3.24}$$

and

$$\boldsymbol{\epsilon}^{(m)} := \begin{cases} \frac{1}{m}\left(\boldsymbol{V}^m - \boldsymbol{I}^{(2)}\right) = \sum_{i=1}^3 \frac{1}{m}\left(\lambda^m - 1\right) \hat{\boldsymbol{n}}_i \otimes \hat{\boldsymbol{n}}_i & \text{if} \quad m \neq 0; \\ \ln\left(\boldsymbol{V}\right) = \sum_{i=1}^3 \ln \lambda_i \, \hat{\boldsymbol{n}}_i \otimes \hat{\boldsymbol{n}}_i & \text{if} \quad m = 0, \end{cases} \tag{3.25}$$

where the notation $\ln\left(\bullet\right)$ denotes the tensor logarithm of the second-order tensor $\left(\bullet\right)$. Having used the spectral decomposition (A.8) on the Right-Hand Side (RHS) of the above equations emphasises the collinearity among the tensors of the family $\boldsymbol{E}^{(m)}$ (defined in the original configuration). Similarly, the family of tensors $\boldsymbol{\epsilon}^{(m)}$, defined in the current configuration, exhibits collinearity. On top of this, the polar decomposition on the RHS of Eq. (3.25) highlights a very straightforward way to compute the tensor logarithm via the logarithm of the principal stretches (i.e., by computing the logarithm of scalar values).

Among the above-defined strain tensors, worthy of particular attention is the *logarithmic strain* $\boldsymbol{\epsilon}$ stemming from Eq. (3.25), namely

$$\boldsymbol{\epsilon} := \boldsymbol{\epsilon}^{(0)} = \ln \boldsymbol{V} = \frac{1}{2} \ln \boldsymbol{b}. \tag{3.26}$$

This measure of strain is particularly attractive since, as proven by Simo [8], it mimics the small-strain additive decomposition of elastic and plastic strain in the case of *isotropic* plasticity (see Section 3.5 for further details). As for any second-order tensor, the above logarithmic strain can be decomposed into a *spheric* and a deviatoric part, i.e.,

$$\boldsymbol{\epsilon} = \frac{1}{3} \underbrace{\epsilon_{ii}}_{:=\epsilon_v} \boldsymbol{I}^{(2)} + \underbrace{\operatorname{dev}(\boldsymbol{\epsilon})}_{:=\boldsymbol{e}} \tag{3.27}$$

where the notation $\operatorname{dev}(\bullet) := \left(\bullet_{ij} - \frac{1}{3} \bullet_{ij} \delta_{ij} \right)$ is adopted. The above equation also defines the volumetric logarithmic strain ϵ_v and the deviatoric part of the logarithmic strain $\boldsymbol{e}$. The volumetric logarithmic strain ϵ_v can be directly linked to the Jacobian J via a relationship which will be widely used in this work (especially in Chapter 4). Owing to Eqs. (A.4), (3.12) and (3.17), it can be seen that

$$\det \boldsymbol{F} = \lambda_1 \lambda_2 \lambda_3 = \det \boldsymbol{V}. \tag{3.28}$$

Considering the definitions (3.25) and (3.27), it also follows that

$$\ln (\det \boldsymbol{V}) = \ln (\lambda_1 \lambda_2 \lambda_3) = \ln \lambda_1 + \ln \lambda_2 + \ln \lambda_3 = \epsilon_{ii} = \epsilon_v, \tag{3.29}$$

which, combined with (3.28), results in

$$\epsilon_v = \ln J. \tag{3.30}$$

3.1.4 Analysis of motion

While the deformation gradient introduced in the previous section describes the change in space of the line segment, it provides no information on how fast this change is occurring. To describe this, the quantity of interest is the *velocity gradient*, which is defined as

$$\boldsymbol{l} := \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}}. \tag{3.31}$$

Applying the chain rule to the above equation, it follows that

$$\boldsymbol{l} = \frac{\partial}{\partial \boldsymbol{X}} \left(\frac{d\boldsymbol{\varphi}(\boldsymbol{X}, t)}{dt} \right) \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{x}} = \frac{d}{dt} \left(\frac{\partial \boldsymbol{\varphi}(\boldsymbol{X}, t)}{\partial \boldsymbol{X}} \right) \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{x}} = \dot{\boldsymbol{F}} \boldsymbol{F}^{-1}, \tag{3.32}$$

which provides a relationship between the velocity gradient and the deformation gradient (and its rate).

Further quantities of interest can be defined by decomposing the velocity gradient $\boldsymbol{l}$ into a symmetric part, named the *stretching tensor $\boldsymbol{d}$*, and a skew-symmetric part, the *spin tensor $\boldsymbol{m}$*, which is expressed by the following relationship

$$\boldsymbol{l} = \underbrace{\operatorname{sym}(\boldsymbol{l})}_{:=\boldsymbol{d}} + \underbrace{(\boldsymbol{l} - \operatorname{sym}(\boldsymbol{l}))}_{:=\boldsymbol{m}}, \tag{3.33}$$

where the classical notation $\operatorname{sym}(\bullet) = \left((\bullet) + (\bullet)^T \right) / 2$ applied to the generic second-order tensor $(\bullet)$ is used.

Another useful relationship relates the material derivative of the Jacobian J with d_{ii} (i.e., the trace of the stretching tensor). This can be seen by noticing that

$$\dot{J} = \frac{d}{dt}\left(J\left(\boldsymbol{F}\right)\right) = \frac{\partial J}{\partial \boldsymbol{F}} : \frac{d}{dt}\left(\boldsymbol{F}\right) = J\boldsymbol{F}^{-T} : \dot{\boldsymbol{F}}, \tag{3.34}$$

where the last equation above follows from the property $\frac{\partial \det(\boldsymbol{A})}{\partial \boldsymbol{A}} = \det\left(\boldsymbol{A}\right) \boldsymbol{A}^{-T}$. Eqs. (3.32) and (3.34) give

$$\dot{J} = J F_{ji}^{-1} \dot{F}_{ij} = J\dot{F}_{ij} F_{ji}^{-1} = J\dot{F}_{ij} F_{jk}^{-1} \delta_{ik} = J\, l_{ik}\, \delta_{ik} = J\frac{\partial v_i}{\partial x_k} \delta_{ik} = J\, d_{ik}\, \delta_{ik}. \tag{3.35}$$

The rate of the Jacobian can also be expressed via the rate of the volumetric logarithmic strain. This is achieved by the time derivative of Eq. (3.30), which results in

$$\dot{J} = \exp\left(\epsilon_v\right)\dot{\epsilon}_v = J\,\dot{\epsilon}_v. \tag{3.36}$$

3.1.5 Kinematics for multiplicative plasticity

As plasticity plays a big role in soil mechanical behaviour (especially if undergoing large deformations), the basic equations useful to define finite strain plasticity are briefly summarised. The reader interested in further details can refer to de Souza Neto *et al.* [5], which presents details on the theory and implementations for finite strain elasto-plasticity.

Elasto-plastic kinematics is based on the local multiplicative decomposition of the deformation gradient[*], this being

$$\boldsymbol{F}\left(\boldsymbol{X},t\right) = \boldsymbol{F}^e\left(\boldsymbol{X},t\right) \boldsymbol{F}^p\left(\boldsymbol{X},t\right) = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}^p}\frac{\partial \boldsymbol{x}^p}{\partial \boldsymbol{X}}, \tag{3.37}$$

where $\boldsymbol{x}^p$ defines the position of the particle at the *intermediate* or *stress-free* configuration (see Figure 3.2 for all the mappings involved in elasto-plasticity and their relative configurations). The polar decomposition theorem (Eq. (3.12)) can be applied separately to the elastic and the plastic deformation gradients, resulting in

$$\boldsymbol{F}^e = \boldsymbol{V}^e\,\boldsymbol{R}^e, \qquad \boldsymbol{F}^p = \boldsymbol{V}^p\,\boldsymbol{R}^p, \tag{3.38}$$

where the definitions and properties of the quantity appearing above trivially follows those introduced in Section 3.1.3. As for the fully elastic case explained above, the elastic left Cauchy-Green strain tensor can be defined as

$$\boldsymbol{b}^e := \boldsymbol{F}^e\left(\boldsymbol{F}^e\right)^T := \left(\boldsymbol{V}^e\right)^2. \tag{3.39}$$

The elastic part of the logarithmic strain is defined as a function of the above tensor as follows

$$\boldsymbol{\epsilon}^e = \ln \boldsymbol{V}^e = \frac{1}{2}\ln \boldsymbol{b}^e. \tag{3.40}$$

---

[*]The multiplicative decomposition of the deformation gradient into an elastic and plastic part has been postulated separately by Kröner [9], Lee [10] and Mandel [11]. This assumption is supported by micromechanical observations of a single crystal in the case of metal plasticity (see, for instance, Rice [12] for a thorough explanation).
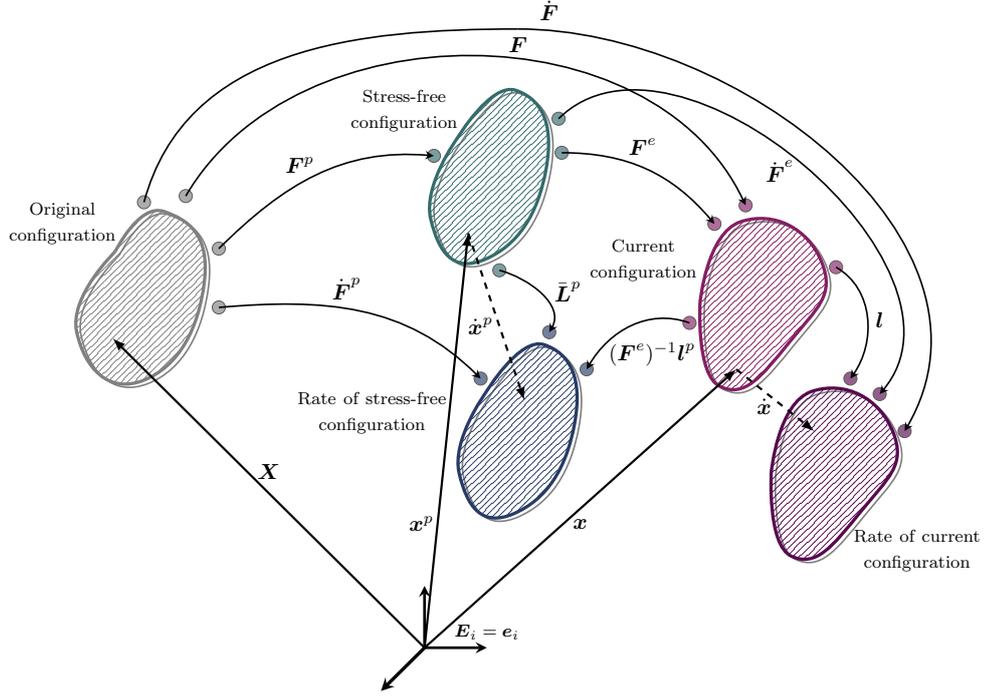
Figure 3.2: Illustration of configurations and relative mappings considered in the case of elasto-plastic kinematics.

Using Eq. (3.37), it can also be appreciated that the elasto-plastic Jacobian is decomposed as follows

$$J = \det\left(\boldsymbol{F}\right) = \det\left(\boldsymbol{F}^e\,\boldsymbol{F}^p\right) = \underbrace{\det\left(\boldsymbol{F}^e\right)}_{:=J^e}\underbrace{\det\left(\boldsymbol{F}^p\right)}_{:=J^p}. \tag{3.41}$$

Owing to the above formulae and to the definition of logarithmic strain (3.26), the following equations hold separately for the elastic and plastic Jacobians

$$J^e = \exp\epsilon_v^e, \qquad J^p = \exp\epsilon_v^p. \tag{3.42}$$

3.1.5.1 Analysis of elasto-plastic motion

Using Eqs. (3.32) and (3.37), the velocity gradient can be additively decomposed into an elastic and a plastic part as follows

$$
\begin{aligned}
l_{ij} &= \dot{F}_{i\alpha}F^{-1}_{\alpha j}\\
&= \left(\frac{d}{dt}\left(F^e_{i\beta}F^p_{\beta\alpha}\right)\right)\left((F^p)^{-1}_{\alpha\gamma}\,(F^e)^{-1}_{\gamma j}\right)\\
&= \underbrace{\dot{F}^e_{i\beta}\,(F^e)^{-1}_{\beta j}}_{:=l^e_{ij}} + F^e_{i\beta}\,\underbrace{\dot{F}^p_{\beta\alpha}\,(F^p)^{-1}_{\alpha\gamma}}_{\bar{L}^p_{\beta\gamma}}\,(F^e)^{-1}_{\gamma j}\\
&= l^e_{ij} + \underbrace{F^e_{i\beta}\,\bar{L}^p_{\beta\gamma}\,(F^e)^{-1}_{\gamma j}}_{:=l^p_{ij}}
\end{aligned}
$$

$$= l_{ij}^e + l_{ij}^p$$
$$= d_{ij}^e + m_{ij}^e + d_{ij}^p + m_{ij}^p. \tag{3.43}$$

In the above chain of formulae, the following quantities have been defined:

- the elastic velocity gradient $\boldsymbol{l}^e$, divided into its symmetric $\boldsymbol{d}^e$ and skew-symmetric part $\boldsymbol{m}^e$;

- the plastic velocity gradient $\boldsymbol{l}^p$, with $\boldsymbol{d}^p$ and $\boldsymbol{m}^p$ being its symmetric and skew-symmetric parts; and

- the velocity gradient in the intermediate configuration $\bar{\boldsymbol{L}}^P$. This last tensor can be, in turn, additively decomposed into a symmetric and skew-symmetric part as follows

$$\bar{\boldsymbol{L}}^P = \underbrace{\mathrm{sym}(\bar{\boldsymbol{L}}^P)}_{:=\bar{\boldsymbol{D}}^P} + \underbrace{\left(\bar{\boldsymbol{L}}^P - \mathrm{sym}(\bar{\boldsymbol{L}}^P)\right)}_{:=\bar{\boldsymbol{M}}^P}. \tag{3.44}$$

Eq. (3.44) is particularly important as the hypothesis of plastic *isotropy* (which will be introduced in 3.5) implies that $\bar{\boldsymbol{M}}^P = \boldsymbol{0}^*$. From this assumption, it straightforwardly follows that the skew-symmetric part of the plastic velocity gradient is null, i.e., $\boldsymbol{m}^p = \boldsymbol{0}$.

3.2 Mass conservation

While the previous section has discussed the motion and the deformation of a body, this section briefly states one of the fundamental principles of continuum mechanics for solids. To begin with, a first principle is

H.3.5 the postulate of conservation of mass, whose mathematical statement is as follows

$$\frac{d}{dt}\int_\omega \rho\,dv = 0. \tag{3.45}$$

with $\rho(\boldsymbol{X}, t)$ being the current *density* of the particle initially in $\boldsymbol{X}$.

Using Eqs. (3.9) and (3.35), the above equation can be stated point-wisely as

$$0 = \frac{d\rho}{dt} + \rho\frac{\partial v_i}{\partial x_i}$$
$$= \dot{\rho} + \rho\frac{\dot{J}}{J}.$$

Integrating the above rates over time leads to

$$\frac{\rho}{\rho_0} = \frac{1}{J}, \tag{3.46}$$

where $\rho_0 = \rho(\boldsymbol{X}, 0)$ is the initial density of the particle initially in $\boldsymbol{X}$. Eq. (3.46) highlights how a local increment in the density is compensated by a local decrease in the volume of particle.

---

*The extension of finite strain plasticity to the anistropic case has been provided by Dafalias [13].

3.3 Momentum balance

In this section, the type of forces acting on the considered body are described and their integration into *Euler's laws* (sometimes referred to as *momentum principles*, which are equivalent to the first and second Newton's laws of motion) are outlined.

To this end, let us assume the following statements to be valid:

H.3.6 let us consider a non-polar continuum, i.e., a *Cauchy continuum*, where only two types of finite external forces can be applied to the body under consideration. These are the body forces $\rho \boldsymbol{f}$ (which are forces per unit volume) and the boundary forces $\boldsymbol{t}$ (forces per unit area), defined as

$$d\boldsymbol{f} = \rho \, \boldsymbol{f} d\omega, \qquad d\boldsymbol{t} = \boldsymbol{t} \, da; \tag{3.47}$$

H.3.7 let us also regard as valid *Cauchy's axiom*, which states that that boundary forces at a given position $\boldsymbol{x}$ at a given time t are only a function of the outward normal $\boldsymbol{n}$ to the surface da, i.e.,

$$\boldsymbol{t} = \boldsymbol{t}\left(\boldsymbol{x}, t, \boldsymbol{n}\right); \tag{3.48}$$

H.3.8 Euler's laws, i.e., the balance of rate of linear and angular momentum are valid. When applied to a body where **H.3.6** is true, these become

$$\frac{d}{dt}\left(\int_{\omega} \rho \, \boldsymbol{v} dv\right) = \int_{\omega} \rho \, \boldsymbol{f} dv + \int_{\gamma} \boldsymbol{t} \, da; \tag{3.49}$$

$$\frac{d}{dt}\left(\int_{\omega} \boldsymbol{x} \times \rho \boldsymbol{v} \, dv\right) = \int_{\omega} \boldsymbol{x} \times \rho \, \boldsymbol{f} dv + \int_{\gamma} \boldsymbol{x} \times \boldsymbol{t} \, da. \tag{3.50}$$

The application of the above principles on the *Cauchy tetrahedron*** allows to state the *Cauchy theorem*, i.e.,

$$\boldsymbol{\sigma} \, \boldsymbol{n} = \boldsymbol{t}, \qquad \text{with} \quad \boldsymbol{\sigma} = \boldsymbol{\sigma}^{T}, \tag{3.51}$$

where $\boldsymbol{\sigma}$ is the symmetric second-order tensor defined as the *Cauchy stress*.

When applied to opposite sides of a single surface, it can be seen that Cauchy's axiom **H.3.7** and the balance of rate of linear momentum (3.49) implies *action-reaction* principle, this being equivalent to Newton's third law.

3.3.1 Local balances of rate of linear momentum

Some manipulations of the balance of rate of linear momentum is possible if mass conservation (3.45) and Cauchy's stress theorem (3.51) are considered. When these equations are substituted into the balance of rate of linear momentum (3.49), it yields

$$\int_{\omega} \rho \, \boldsymbol{a} \, dv = \int_{\omega} \rho \, \boldsymbol{f} dv + \int_{\gamma} \boldsymbol{\sigma} \boldsymbol{n} \, da. \tag{3.52}$$

**This proof is omitted in this work, but the interested reader is referred to Gurtin *et al.* [4], for instance.*

The *divergence* (or *Gauss-Green*) theorem can be applied on the outer surface, resulting in

$$\int_\omega \rho \, a_i \, dv = \int_\omega \rho \, f_i dv + \int_\omega \frac{\partial \sigma_{ij}}{\partial x_j} \, dv. \tag{3.53}$$

Since all of the integrals in the above equation are evaluated on the (current) volume and no particular assumptions have been made in this regard, it follows that the above equation is valid locally (point-wisely) on the whole current volume ω, i.e.,

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \rho \, f_i = \rho \, a_i. \tag{3.54}$$

However, while the above equation is locally valid for the continuous body under consideration, it will not be strictly satisfied by its discretised counterpart. This is due to the choice of the equations discretised by, for instance, the MPM, which takes the weak form (or variational form) of equations into account.

3.3.2 Stress measures and Total Lagrangian (TL) balance of rate of linear momentum

As can be seen from Eqs. (3.52) or (3.53), the particular form of rate of linear momentum is expressed in the current configuration, i.e., those equations are set in an UL description. However, the same equation can be stated in the original configuration, using a TL framework.

For this purpose, let us consider, in the current configuration, a force t acting on an infinitesimal area da. If the resultant force is the same between original and current configuration, it implies that $t \, da = t_0 \, dA$, where t_0 and dA are the initial force and initial infinitesimal area. If the Cauchy stress theorem (3.51) is considered, the above equality can be expanded as follows

$$t \, da = t_0 \, dA = \boldsymbol{\sigma} \boldsymbol{n} \, da = J \boldsymbol{\sigma} \boldsymbol{F}^{-T} \boldsymbol{N} dA, \tag{3.55}$$

where Nanson's formula (3.11) has been applied to obtain the RHS of the above chain of formulae. Let us define the *first Piola-Kirchhoff* stress as the quantity appearing on the RHS of the above equation, i.e.,

$$\boldsymbol{P} := J \boldsymbol{\sigma} \boldsymbol{F}^{-T}. \tag{3.56}$$

It is quite clear that this stress measure is not symmetric, i.e., this tensor exhibits a two-point nature. Physically speaking, while the Cauchy stress is defined as the stress acting on the current area da defined by the normal $\boldsymbol{n}$, the first Piola-Kirchhoff stress acts on the original area dA defined by the original normal $\boldsymbol{N}$.

The *second Piola-Kirchhoff* stress can be regarded as the symmetrised version of its first counterpart, namely

$$\boldsymbol{S} := J \boldsymbol{F}^{-1} \boldsymbol{\sigma} \boldsymbol{F}^{-T} = \boldsymbol{S}^T. \tag{3.57}$$

Another convenient measure of stress (which is particularly attractive for some problems, see 3.5.1) is the *Kirchhoff stress*, this being given by

$$\boldsymbol{\tau} := J \boldsymbol{\sigma}. \tag{3.58}$$

Following its definition, it can be seen that the Kirchhoff stress inherits symmetry from the Cauchy stress.

Having introduced different measures of stress, it is possible to state the TL formulation of the balance of rate of linear momentum. Using Eqs. (3.9), (3.46), (3.55) and the divergence theorem, the balance of rate of linear momentum Eq. (3.52) can be expressed as follows

$$\int_\Omega \rho_0 \, a_i dV - \int_\Omega \rho_0 \, f_i dV = \int_\gamma \sigma_{ij} n_j \, da = \int_\Gamma P_{ij} N_j \, dA = \int_\Omega \frac{\partial}{\partial X_j} P_{ij} dV.$$

The local form of the above reads

$$\frac{\partial}{\partial X_j} P_{ij} + \rho_0 \, f_i = \rho_0 \, a_i. \tag{3.59}$$

As already stated in Section 3.1.1 and highlighted by Eq. (3.59), the difference between an UL and a TL approach is superficial. As such, the two forms of the same equation are exploited depending on their convenience.

3.4 Fundamental principles of thermodynamics

The fundamental principles of continuum mechanics do not end with the principles of conservation of mass and momenta. While the former law states how mass redistributes spatially in a given instant (but that its totality cannot change over time), the latter laws (in the form of conservation of linear and angular momentum) state how loads applied to the body are internally balanced. However, neither of these laws establishes what ties the displacements (or strains) to internal body forces (or stresses). Describing this kind of relationships is the scope of *constitutive laws*. This is where the first and second principles of thermodynamics come into play, allowing the conservation and/or dissipation of energy to be established. In particular, the union of these two principles (called the *Clausius-Duhem inequality*) makes it possible to establish when a transformation does not conform to these laws. Hence, the Clausius-Duhem inequality acts as a watershed between constitutive laws that comply with the laws of thermodynamics and those that violate them. In Feynman's words [14]

```
''... we see that a substance's properties must be limited in a
certain way; one cannot make up anything he wants, or he would be
able to invent a substance which he could use to produce more than
the maximum allowable work when he carried it around a reversible
cycle. This principle, this limitation, is the only real rule that
comes out of the thermodynamics.''
```

In this sense, the laws of thermodynamics do not add any insight to the constitutive laws, but only set specific boundaries for them.

3.4.1 First principle of thermodynamics

H.3.9 Feynman [14] defines the first principle of thermodynamics as

> ''... the conservation of energy: if one has a system and puts
> heat into it, and does work on it, then its energy is increased
> by the heat put in and the work done.''

In the continuous body of interest for this chapter, there are only two types of energy that can be developed, i.e., the kinetic energy and the strain energy. If K and U denote the kinetic and the strain energies for the whole considered body, the first principle takes the following form

$$\frac{d}{dt}\left(K + U\right) = P^{ext} + Q,\tag{3.60}$$

where P^{ext} is the externally supplied power and Q denotes the heat introduced in the system.

Assumption **H.3.6** implies also that the external power supplied to this particular kind of body can only be performed by body forces $\rho \boldsymbol{f}$ and boundary forces $\boldsymbol{t}$. Integral definitions of the quantities appearing in Eq. (3.60) are listed below

$$K := \frac{1}{2}\int_\omega \rho\,\boldsymbol{v}^2\,dv;\tag{3.61}$$

$$U := \int_\omega \rho\,\check{e}\,dv;\tag{3.62}$$

$$P^{ext} = P^{ext}_f + P^{ext}_t := \int_\omega \rho\,\boldsymbol{f}\cdot\boldsymbol{v}\,dv; + \int_\gamma \boldsymbol{t}\cdot\boldsymbol{v}\,da\tag{3.63}$$

$$Q := -\int_\gamma \boldsymbol{q}^{(T)}\cdot\boldsymbol{n}\,da,\tag{3.64}$$

where $\check{e}$ indicates the internal energy per unit current mass, while $\boldsymbol{q}^{(T)}$ is the heat flow vector pointing outwards to the domain boundary. It can be seen from Eq. (3.64) that this thesis does not take volume sources of heat into consideration as part of the heat power. Mass conservation (3.45) allows expressing the time derivatives appearing in Eq. (3.60) as

$$\frac{dK}{dt} = \frac{1}{2}\int_\omega \frac{d}{dt}\left(\rho\,J\,\boldsymbol{v}^2\right)dV = \int_\omega \rho\,\boldsymbol{a}\cdot\boldsymbol{v}\,dv;\tag{3.65}$$

$$\frac{dU}{dt} = \int_\omega \frac{d}{dt}\left(\rho\,J\,\check{e}\right)dV = \int_\omega \rho\,\frac{d\check{e}}{dt}\,dv.\tag{3.66}$$

Furthermore, applying the divergence theorem to the power externally supplied by the boundary forces gives

$$P^{ext}_t = \int_\gamma \left(\boldsymbol{\sigma}\,\boldsymbol{n}\right)\cdot\boldsymbol{v}\,da = \int_\omega \frac{\partial}{\partial\boldsymbol{x}}\cdot\left(\boldsymbol{\sigma}\,\boldsymbol{v}\right)dv = \int_\omega \left(\frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{x}}\cdot\boldsymbol{v} + \boldsymbol{\sigma}:\boldsymbol{l}\right)dv,\tag{3.67}$$

where the gradient of velocity on the RHS of the above equality is expressed using Eq. (3.31). If Eqs. (3.65)-(3.67) are substituted in the statement of the first principle of thermodynamics (3.60), it follows that

$$\int_\omega \underbrace{\left(\rho\left(\boldsymbol{a} - \boldsymbol{f}\right) - \frac{\partial\boldsymbol{\sigma}}{\partial\boldsymbol{x}}\right)}_{=\boldsymbol{0}}\cdot\boldsymbol{v}\,dv + \int_\omega \rho\,\frac{d\check{e}}{dt}\,dv = \int_\omega \left(\boldsymbol{\sigma}:\boldsymbol{l}\right)dv - \int_\gamma \boldsymbol{q}^{(T)}\cdot\boldsymbol{n}\,da,\tag{3.68}$$

where the gathered terms are null due to the balance of rate of linear momentum Eq. (3.54). Owing to the symmetry of the Cauchy stress tensor, the power performed by the stress and the velocity gradient can also be expressed as $\boldsymbol{\sigma} : \boldsymbol{l} = \boldsymbol{\sigma} : \boldsymbol{d}$. The use of the divergence theorem on the heat flux term allows to express all of the terms in Eq. (3.68) on the considered volume, which, in turn, can be eliminated to express the above equation locally as

$$\rho \frac{d\check{e}}{dt} = \boldsymbol{\sigma} : \boldsymbol{l} - \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{q}^{(T)}. \tag{3.69}$$

Eq. (3.69) constitutes a local statement of the first principle of thermodynamics per unit volume for the considered body under all of the hypotheses introduced so far.

3.4.2 Second principle of thermodynamics

Feynman [14] concisely states the second principle of thermodynamics as follows

```
''... the entropy of the universe is always increasing.''
```

H.3.10 The mathematical statement of the above idea is expressed, for the body under consideration, by the following inequality

$$\frac{d}{dt} \int_{\omega} \rho \check{s} \, dv \geq - \int_{\gamma} \frac{1}{T} \left(\boldsymbol{q}^{(T)} \cdot \boldsymbol{n} \right) da, \tag{3.70}$$

where $\check{s}$ denotes the entropy per unit current mass and T the absolute temperature*.

Following similar steps to those adopted for the first principle, mass conservation Eq. (3.45) and divergence theorem permits the following statement of the second principle of thermodynamics

$$\int_{\omega} \frac{d}{dt} \left(\rho \, J \, \check{s} \right) dV \geq - \int_{\omega} \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\frac{1}{T} \boldsymbol{q}^{(T)} \right) dv, \tag{3.71}$$

or, in a local form,

$$\rho \dot{\check{s}} + \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\frac{1}{T} \boldsymbol{q}^{(T)} \right) \geq 0, \tag{3.72}$$

From the above inequality, it can be seen that an increment of entropy produces a divergence (i.e., a spatial variation) in temperature and heat flux.

3.4.3 Clausius-Duhem inequality

As will be detailed in Section 4.4.3, for the fluid part of a porous material, there is no unique way to define energy in a system, as this can be described via different state variables. Among the forms of interest describing the energy of a system, this thesis has introduced

---

*It can be noted that the second principle of thermodynamics described by Eq. (3.70) does not give a description of what entropy is, but it describes how it changes. For a definition of entropy (and not its variation, which appears in (3.70)) the so-called *third law of thermodynamics* is necessary (see, for instance, Feynman [14]). Moreover, a definition of temperature independent from the definition of entropy is possible only if the *zeroth law of thermodynamics* is introduced (see, again, Feynman [14]).

internal energy $\check{e}$ in the first principle of thermodynamics (see Eq. (3.62)). However, via the *Legendre* transformation (see, for instance, Puzrin and Houlsby [15]), the *Helmholtz free energy* $\check{\psi}$ can be introduced as follows

$$\check{\psi} := \check{e} - \check{s}\,T. \tag{3.73}$$

If the rate of the above free energy is substituted into the second principle of thermodynamics (3.72), it results in

$$\frac{\rho}{T}\left(\dot{\check{e}} - \check{s}\dot{T} - \dot{\check{\psi}}\right) + \frac{\partial}{\partial \boldsymbol{x}}\left(\frac{1}{T}\right) \cdot \boldsymbol{q}^{(T)} + \frac{1}{T}\frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{q}^{(T)} \geq 0. \tag{3.74}$$

The rate of the internal energy $\check{e}$ can be expressed via the local form of the first principle of thermodynamics, i.e., Eq. (3.69). Substituting this law into (3.74) (this being the local form of the second principle) gives the *Clausius-Duhem inequality*, i.e.,

$$\frac{1}{T}\left(\boldsymbol{\sigma} : \boldsymbol{l} - \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{q}^{(T)}\right) - \frac{\rho}{T}\left(\check{s}\dot{T} + \dot{\check{\psi}}\right) - \frac{1}{T^2}\frac{\partial}{\partial \boldsymbol{x}}(T) \cdot \boldsymbol{q}^{(T)} + \frac{1}{T}\frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{q}^{(T)} \geq 0. \tag{3.75}$$

Since the absolute temperature is positive by definition, the above equation can be simplified as follows

$$\boldsymbol{\sigma} : \boldsymbol{l} - \rho\,\check{s}\,\dot{T} - \rho\dot{\check{\psi}} - \frac{1}{T}\frac{\partial T}{\partial \boldsymbol{x}} \cdot \boldsymbol{q}^{(T)} \geq 0. \tag{3.76}$$

Expressing the power performed by the stress vector via Kirchhoff stress (3.58) and using Eq. (3.46) to relate original and current density, the above inequality becomes

$$\mathcal{D} := \boldsymbol{\tau} : \boldsymbol{l} - \rho_0\,\check{s}\,\dot{T} - \rho_0\dot{\check{\psi}} - \frac{J}{T}\frac{\partial T}{\partial \boldsymbol{x}} \cdot \boldsymbol{q}^{(T)} \geq 0, \tag{3.77}$$

with $\mathcal{D}$ indicating the dissipation taking place in the considered body. From the above formula, it can be seen how the power resulting from the stress and its conjugate variable ($\boldsymbol{l}$ in the above formula, or, equally $\boldsymbol{d}$) gives more than the sum of increase in temperature, heat flux and strain (stored) energy function. Hence, it can be seen that including a Helmholtz free energy function giving more power than that obtained by the couple $\boldsymbol{\tau} : \boldsymbol{d}$ is not only physically questionable, but it would imply issues in the algorithmic stability too, as the energy of the system would not be upper-bounded.

3.5 Constitutive relationship

So far, what has been described is how the continuous body of interest moves and deforms (Section 3.1), how external loads are distributed and balanced (Section 3.3), and how thermodynamics allows only specific changes for the energy in the body (Section 3.4). However, all of these pieces of information are not sufficient to characterise a constitutive behaviour. To this end, the following general assumptions about constitutive modelling of the body under consideration are introduced*

---

* Constitutive modelling represents a study field which, for completeness, would require an entire treatise on the argument itself. However, for reasons of brevity, only the ready-to-use hypotheses are considered in the above list. Other basic rules or principles (such as the principle of equipresence or determinism), which this work takes for granted, can be found, for instance, in Truesdell and Noll [16] or Oden [2].

H.3.11 thermal effects do not give mechanical stress;

H.3.12 rate effects are excluded;

H.3.13 the material under consideration is elastically and plastically isotropic; and

H.3.14 the principle of maximum dissipation governs the evolution equations for plasticity.

Assumption **H.3.11** excludes the role of temperature in the considered relationship. However, from a physical perspective, it must be accounted that the role of thermal effects cannot be entirely ruled out (as stated by the second principle of thermodynamics 3.4.2), since the irreversible changes taking place in the solid are transformed into heat. With the above assumption, it is meant that the thermal effects are of secondary importance when compared to the other phenomena taking place in the solid. This assumption allows simplification of the Clausius-Duhem inequality (4.75) into the *Clausius-Planck* inequality (see Bennet *et al.* [17], for instance, in regards to the use of these terms), this being

$$\mathcal{D} := \boldsymbol{\tau} : \boldsymbol{l} - \dot{\Psi} \geq 0, \tag{3.78}$$

having defined the Helmholtz free energy function per unit initial volume as $\Psi := \rho_0 \check{\psi}$.

Hypothesis **H.3.12** permits a description of the classical rate-independent plasticity. Hence, when rates are considered, these are regarded as increments with respect to a pseudo-time.

Assumption **H.3.13** excludes anisotropic behaviour for the material under consideration, i.e., the free energy function and the plastic flow rule can be entirely defined via invariants.

The principle of maximum dissipation **H.3.14** implies a simplification for the direction of the plastic flow, as detailed below in Section 3.5.2.

3.5.1   Free energy function

In (3.78), the free energy function has been kept as general[*] as possible so far. One of the most fundamental requirements concerning the free energy function is that it must be *objective* (see, in this regard, Simo [18]), meaning that the (scalar) free energy function must remain constant under rotations and translations and rotations of the reference frame[†]. Hence, while it makes physical sense that the free energy is a function of the deformation gradient $\boldsymbol{F}$, objectivity requires that this dependency must take place via the right Cauchy-Green strain $\boldsymbol{C}$. If, on top of this, assumption **H.3.13** is considered too, the free energy function can be equally expressed as a function of the left Cauchy-Green strain $\boldsymbol{b}$.

---

[*]Strictly speaking, deriving a constitutive stress-strain relationship from an energy function excludes another type of materials, named *hypo-elastic* (see Simo [18] for a definition of these materials).

[†] The objectivity defined above is labelled objectivity for *isometric transformations* of the reference frame (see, for instance, Truesdell and Noll [16]). However, there exists also another concept of objectivity under general changes of the reference frame, named *covariance* (see, for instance, Marsden and Hughes [7]). For the purposes of this work, the fist type of objectivity is sufficient.

To extend the characterisation of the material under consideration to the elasto-plastic case, this work considers the free energy function to be dependent on the elastic left Cauchy-Green strain $\boldsymbol{b}^e$ and on a set of internal variables, which are generically denoted as $\boldsymbol{\alpha}$. This is expressed via the following formula, i.e.,

$$\Psi = \hat{\Psi}\left(\boldsymbol{b}^e, \boldsymbol{\alpha}\right). \tag{3.79}$$

Furthermore, if a *decoupled* material is considered (see discussion in Section 4.5.1 for coupled materials), the above dependency simplifies into the sum

$$\Psi = \hat{\Psi}_1\left(\boldsymbol{b}^e\right) + \hat{\Psi}_2\left(\boldsymbol{\alpha}\right), \tag{3.80}$$

i.e., it is assumed that hardening mechanisms are not coupled with the elastic strains.

In light of all the introduced considerations, the Clausius-Planck inequality (3.78) becomes

$$\mathcal{D} = \boldsymbol{\tau} : \boldsymbol{l} - \dot{\hat{\Psi}}\left(\boldsymbol{b}^e, \boldsymbol{\alpha}\right) \geq 0, \tag{3.81}$$

or, making explicit the rate of the free energy function and using (3.43) to additively decoupling the elastic and plastic parts of the velocity gradient, we have

$$\mathcal{D} = \boldsymbol{\tau} : \left(\boldsymbol{l}^e + \boldsymbol{l}^p\right) - \frac{\partial\hat{\Psi}_1}{\partial\boldsymbol{b}^e}\dot{\boldsymbol{b}}^e - \frac{\partial\hat{\Psi}_2}{\partial\boldsymbol{\alpha}} * \dot{\boldsymbol{\alpha}} \geq 0. \tag{3.82}$$

In the above formula, $*$ denotes the appropriate operator between the conjugate powers $\frac{\partial\hat{\Psi}_2}{\partial\boldsymbol{\alpha}}$ and $\dot{\boldsymbol{\alpha}}$. Following the definition of the elastic left Cauchy-Green strain (3.39), its rate can be written as

$$\dot{\boldsymbol{b}}^e = \dot{\boldsymbol{F}}^e\left(\boldsymbol{F}^e\right)^T + \boldsymbol{F}^e\left(\dot{\boldsymbol{F}}^e\right)^T = \boldsymbol{l}^e\boldsymbol{F}^e\left(\boldsymbol{F}^e\right)^T + \boldsymbol{F}^e\left(\boldsymbol{F}^e\right)^T\left(\boldsymbol{l}^e\right)^T = \boldsymbol{l}^e\boldsymbol{b}^e + \boldsymbol{b}^e\left(\boldsymbol{l}^e\right)^T, \tag{3.83}$$

where the definition of $\boldsymbol{l}^e$ has been used. Acknowledging the symmetry of the tensors appearing in inequality (3.82), it becomes

$$\mathcal{D} = \left(\boldsymbol{\tau} - 2\frac{\partial\hat{\Psi}_1}{\partial\boldsymbol{b}^e}\boldsymbol{b}^e\right) : \boldsymbol{d}^e + \boldsymbol{\tau} : \boldsymbol{d}^p - \frac{\partial\hat{\Psi}_2}{\partial\boldsymbol{\alpha}} * \dot{\boldsymbol{\alpha}} \geq 0. \tag{3.84}$$

Since the above inequality should hold for any rate of strain internal variables, standard arguments, which go under the name of the *Coleman-Noll* procedure (see, for instance, Coleman and Noll [19], and Coleman and Gurtin [20]), allow rewriting as follows

$$\begin{cases} \boldsymbol{\tau} = 2\dfrac{\partial\hat{\Psi}_1}{\partial\boldsymbol{b}^e}\boldsymbol{b}^e; & \text{(3.85a)} \\[2ex] \mathcal{D} = \boldsymbol{\tau} : \boldsymbol{d}^p + \boldsymbol{q} * \dot{\boldsymbol{\alpha}} \geq 0, & \text{(3.85b)} \end{cases}$$

where the thermodynamic forces

$$\boldsymbol{q} := -\frac{\partial\hat{\Psi}_1}{\partial\boldsymbol{\alpha}} \tag{3.86}$$

have been introduced. Eq. (3.85a) describes a constitutive relationship between stress state and elastic strain. Owing to the nature of elastic strain (i.e., entirely recoverable), it makes sense that Eq. (3.85a) does not contribute to the system's entropy increase (3.85b). This growth of dissipation is purely given by the power that the stresses perform for the plastic strains and the conjugate powers $\boldsymbol{q} * \dot{\boldsymbol{\alpha}}$.

The form of $\hat{\Psi}_1$ defined so far is still quite generic. This chapter considers a *Hencky* material (see Hencky [21] for the original formulation or de Souza Neto *et al.* [5] for further details), which is defined as follows

$$\hat{\Psi}_1\left(\epsilon_v^e, \epsilon_q^e\right) = \frac{\bar{K}}{2}\left(\epsilon_v^e\right)^2 + \frac{3}{2}G\left(\epsilon_q^e\right)^2 \tag{3.87}$$

with $\bar{K}$ and G being the *bulk* and *shear* moduli, and the *von Mises equivalent strain* $\epsilon_q := \sqrt{\frac{2}{3}\,e:e}$, with e defined in Eq. (3.27). The reason why Hencky materials are adopted in this chapter is that their energy is a function of the logarithmic strains, these defined by Eq. (3.26). As described originally by Simo [8], the use of the logarithmic strain in the context of finite strain elasto-plasticity reproduces the additive structure of the small strain setting. Furthermore, it can be seen that, for this specific material, the constitutive relationship Eq. (3.85a) becomes linear, i.e.,

$$\boldsymbol{\tau} = 2\frac{\partial\hat{\Psi}_1}{\partial\boldsymbol{b}^e}\boldsymbol{b}^e = 2\frac{\partial\hat{\Psi}_1}{\partial\boldsymbol{\epsilon}^e}\frac{\partial\boldsymbol{\epsilon}^e}{\partial\boldsymbol{b}^e}\boldsymbol{b}^e = \frac{\partial\hat{\Psi}_1}{\partial\boldsymbol{\epsilon}^e} = \boldsymbol{\mathcal{D}}^e:\boldsymbol{\epsilon}^e. \tag{3.88}$$

where the (constant) elastic moduli tensor $\boldsymbol{\mathcal{D}}^e$ appearing in the above equation is given by

$$\boldsymbol{\mathcal{D}}^e := \bar{K}\,\boldsymbol{I}^{(2)}\otimes\boldsymbol{I}^{(2)} + 2\,G\,\boldsymbol{I}^{(4),\,dev}. \tag{3.89}$$

A definition of $\boldsymbol{I}^{(4),\,dev}$ in *Voigt notation* is given by Eq. (C.8) in Appendix C.

3.5.2 Plastic evolution laws

While the elastic behaviour has been fully described in the previous section, plasticity still needs to be characterised. In this regards, two items are necessary, namely the *yield function* and the *flow rule*. The former assesses when the stresses (and the thermodynamic forces $\boldsymbol{q}$, as suggested by the inequality (3.85b)) are such that the material starts exhibiting incipient plasticity (or not), while the latter describes the direction of plastic deformations $\boldsymbol{d}^p$ and internal variables $\boldsymbol{\alpha}$ (when/if they occur). It is also well-known that the plastic behaviour must obey the Karush-Kuhn-Tucker (KKT) conditions, these being

$$\begin{cases} \dot{\gamma}^p \geq 0; \\ \Phi\left(\boldsymbol{\tau},\boldsymbol{q}\right) \leq 0; \\ \dot{\gamma}^p\,\Phi\left(\boldsymbol{\tau},\boldsymbol{q}\right) = 0, \end{cases} \tag{3.90}$$

where γ^p is the (scalar) plastic multiplier and $\Phi\left(\boldsymbol{\tau},\boldsymbol{q}\right)$ is the (scalar) yield function. The plastic multiplier is responsible for describing the intensity of plastic deformations $\boldsymbol{d}^p$ and internal variables $\boldsymbol{\alpha}$, while further considerations should be made to compute their directions. Contrary to the description which is developed in Chapter 4 (where more complex models will be introduced), this chapter considers assumption **H.3.14**, namely the principle of maximum dissipation (see, for the original contributions, von Mises [22], Hill [23], and Mandel [24]). Under this hypothesis, the yield function and the direction of the flow rule are described by the same function $\Phi\left(\boldsymbol{\tau},\boldsymbol{q}\right)$. Hence, recalling the hypothesis of isotropic

plasticity **H.3.13** (implying $\boldsymbol{m}^p = \mathbf{0}$, as pointed out in Section 3.1.5.1), the plastic strains and the internal variables are given by

$$\boldsymbol{d}^p = \dot{\gamma}^p \frac{\partial \Phi}{\partial \boldsymbol{\tau}}; \tag{3.91}$$

$$\dot{\boldsymbol{\alpha}} = \dot{\gamma}^p \frac{\partial \Phi}{\partial \boldsymbol{q}}. \tag{3.92}$$

The above set of equations is particularly convenient since, as described by Simo [8], it extends the small-strain additive decomposition for elasto-plasticity to the case of finite strains when $\bar{\boldsymbol{L}}^p$ is exponentially integrated.

3.6 The discrete displacement-based formulations for the Material Point Method (MPM)

So far, all the introduced considerations and equations aimed at a description of the continuum. However, the continuum equations are not compatible with programming on finite machines since, by definition, a continuum deformable body has infinite Degrees of Freedom (DoFs). On the other hand, discrete formulations are suitable for this purpose.

The standard formulation for simulating the dynamics of solids in the MPM consists of discretising the (weak form of) the balance of rate of linear momentum as the primary equation, with primary variables being the displacement field (or, equally, velocities or accelerations). These displacements give rise to strains via the kinematic relationships explained in Section 3.1. The strains, in turn, cause stresses according to the constitutive relationships outlined in Section 3.5.

Discretising only the balance of rate of linear momentum implies that the discrete versions of other conservation laws (i.e., balance of rate of angular momentum and Clausius-Planck inequality) are not guaranteed to be conserved*. The possible violation of these laws justifies the choice of appropriate time discretisation in Section 3.6.2.2 and the modification of the internal force vector in Section 3.6.3. These modified algorithms are labelled as Conservation Law Consistent (CLC) MPM formulations, as they guarantee mass, momenta and energy conservation laws. The first TL MPM formulation of these algorithms was firstly described in Love and Sulsky [25], while this work (drawn to Pretti *et al.* [26]) details its UL version. To explain how the CLC UL internal forces vector is obtained from its TL counterpart (provided by the existing literature), the strong forms of the balance of rate of linear momentum for both of these two Lagrangian formulations are given below

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \rho \left(b_i - \dot{v}_i \right) = \mathbf{0}; \tag{3.93}$$

$$\frac{\partial P_{ij}}{\partial X_j} + \rho_0 \left(b_i - \dot{v}_i \right) = \mathbf{0}. \tag{3.94}$$

In the following subsection 3.6.1, both of these strong forms are discretised, and more details on the CLC UL formulation are given in the following parts of this chapter.

---

*A separate discussion must be devoted to mass conservation, which, as it will be outlined in 3.6.2.1, is automatically conserved for an MPM discretisation insofar as a solid material is considered.
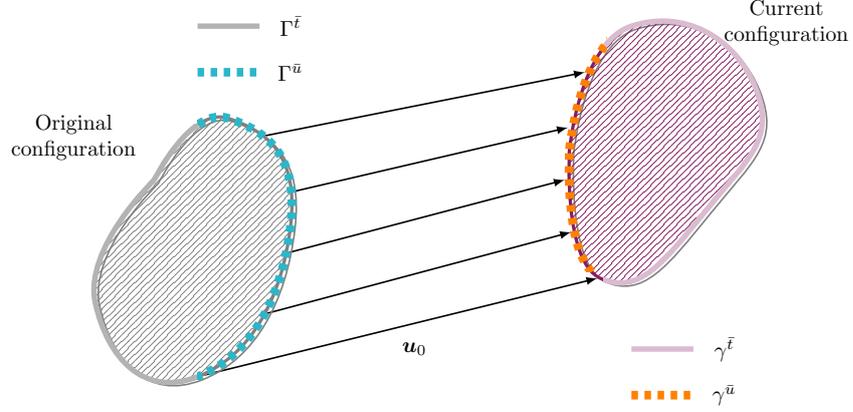
Figure 3.3: Relationship between the BCs applied on the original and the current configuration.

3.6.1 Weak forms

To fully describe a boundary value problem and its weak formulation, the introduction of conditions on the boundary (named Boundary Conditions (BCs)) is necessary. In the case of displacement-based formulations, these BCs describe the displacements (Dirichlet) and boundary forces (Neumann) prescribed on the body. For this purpose, let the current boundary γ be partitioned as follows

$$\gamma = \gamma^{\bar{u}} \cup \gamma^{\bar{t}}; \tag{3.95}$$

$$\gamma^{\bar{u}} \cap \gamma^{\bar{t}} = \emptyset, \tag{3.96}$$

and let the BCs defined on these surfaces be

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \qquad \text{on } \gamma^{\bar{u}}; \tag{3.97}$$

$$\boldsymbol{\sigma}\,\boldsymbol{n} = \bar{\boldsymbol{t}} \qquad \text{on } \gamma^{\bar{t}}. \tag{3.98}$$

If it is desired to apply equivalent BCs on the initial boundary Γ, a few minor modifications must be introduced. These consist in noticing that the current and the initial partitions of the boundaries are related by the motion mapping, i.e.,

$$\gamma^{\bar{t}} = \boldsymbol{\varphi}\left(\Gamma^{\bar{t}}, t\right); \tag{3.99}$$

$$\gamma^{\bar{u}} = \boldsymbol{\varphi}\left(\Gamma^{\bar{u}}, t\right). \tag{3.100}$$

As assumed by hypothesis **H.3.4** (i.e., one-to-one relationship for the motion mapping), the initial boundaries $\Gamma^{\bar{t}}$ and $\Gamma^{\bar{u}}$ inherit the same properties given by Eqs. (3.95) and (3.96) from their current counterparts. BCs equivalent to (3.97) and (3.98) in the case of a TL formulation are given by

$$\boldsymbol{u} = \bar{\boldsymbol{u}}_0 = \boldsymbol{u}_0 + \bar{\boldsymbol{u}} \qquad \text{on } \Gamma^{\bar{u}}; \tag{3.101}$$

$$\boldsymbol{P}\,\boldsymbol{N} = \bar{\boldsymbol{t}}_0 \qquad \text{on } \Gamma^{\bar{t}}, \tag{3.102}$$

where $\boldsymbol{u}_0$ are the applied displacements between the initial and the current configurations (see Figure 3.3). Equivalence between current and initial tractions (expressed by Eq. (3.55)) has been used for Neumann BCs.

A further component necessary for weak formulations are the spaces of admissible trial and weighting functions. In the case of a UL formulation, these are given by

$$\mathscr{V}^{\boldsymbol{u}} = \left\{ \boldsymbol{u} \in \left[H^1\left(\omega\right) \right]^{n^{dim}} \mid J > 0 \,\wedge\, \boldsymbol{u} = \bar{\boldsymbol{u}} \text{ on } \gamma^{\bar{u}} \right\}; \tag{3.103}$$

$$\mathscr{W}^{\boldsymbol{w}} = \left\{ \boldsymbol{w} \in \left[H^1\left(\omega\right) \right]^{n^{dim}} \mid \boldsymbol{w} = \boldsymbol{0} \text{ on } \gamma^{\bar{u}} \right\}; \tag{3.104}$$

with $H^1\left(\omega\right)$ denoting the *Sobolev* space of degree one on ω (see, for instance, Brenner and Scott [27] for a rigorous definition of this space). To obtain the spaces necessary for the TL formulation, the considered domain Ω, the Dirichlet boundary and the BCs should be accommodated as described above.

The procedure to obtain the weak form consists in multiplying the equilibrium equation by the weighting function $\boldsymbol{w} \in \mathscr{W}^{\boldsymbol{w}}$ and integrating over the volumes (current in the case of UL formulation, original for the TL case). The fundamental theorem of calculus, the divergence theorem, and the Cauchy stress theorem (3.51) applied to the boundary $\gamma^{\bar{t}}$ (according to the Neumann BC (3.98)) allow us to state the *weak forms* as follows: find $\boldsymbol{u} \in \mathscr{V}^{\boldsymbol{u}}$ such that, for $t \in [0, T]$,

$$\int_\omega \frac{\partial w_i}{\partial x_j} \sigma_{ij}\, dv - \int_\omega w_i \rho \left(b - \dot{v}_i\right)\, dv - \int_{\gamma_{\bar{t}}} w_i \bar{t}_i\, da = 0, \quad \forall \boldsymbol{w} \in \mathscr{W}^w; \tag{3.105}$$

given the initial conditions $\boldsymbol{u}\left(t = 0\right) = \boldsymbol{u}_0$ and $\dot{\boldsymbol{u}}\left(t = 0\right) = \boldsymbol{v}_0$. For a TL case, a similar weak formulation can be found by

$$\int_\Omega \frac{\partial w_i}{\partial X_j} P_{ij}\, dV - \int_\Omega \rho_0 w_i \left(b_i - \dot{v}_i\right)\, dV - \int_{\Gamma_{\bar{t}}} w_i \bar{t}_i^0\, dA = 0, \quad \forall \boldsymbol{w} \in \mathscr{W}^w, \tag{3.106}$$

where the reader is asked to overlook the abuse of notation in the weighting and trial functions in above TL formula, as their function spaces are, as explained above, slightly different.

It must also be noted that both Eqs. (3.105) and (3.106) make use of the zero value of the weighting functions on their Dirichlet boundary, so that the only applied tractions are those on the Neumann boundaries (i.e., reaction forces on the Dirichlet boundary are set as zero owing to the weighting functions). While this information does not seem to be particularly important at this stage, clarification is given in Section 3.6.2.

Some additional considerations can be made regarding the spaces for weighting and trial functions. In a loose sense, the Sobolev space of degree one is linked to a space of functions almost everywhere one time differentiable. While this (weak) one-time differentiability guarantees the existence of quantities connected to the first derivatives of the displacement field (such as the strain and stress), the first degree of the Sobolev space is not enough to guarantee the continuity of these first derivatives. However, this continuity of the first

derivatives is particularly valuable in the MPM, as Material Points (MPs) can move, across different time-steps, from one mesh element to an adjacent one. Roughly speaking, if linear polynomial function spaces are considered, MPs can experience positive (in one element) and negative (in the adjacent element) derivatives along the normal to the elements' faces (see, in this regard, Figure 3.4). In the MPM literature, this phenomenon goes under the name of *cell-crossing instability* (see, for instance, Bardenhagen and Kober [28] or González-Acosta *et al.* [29]). Therefore, while Eqs. (3.105) and (3.106) do not require spaces that provide greater continuity than $H^1(\omega)$, this greater continuity is desired to have results that, in terms of the strain and stress state, are physically satisfactory, i.e., smoother. This is the reason why this work considers both the original MPM basis functions (linear polynomials), but also the Generalised Interpolation Material Point Method (GIMPM) basis functions. Further details of these are given in Bardenhagen and Kober [28] or Charlton *et al.* [30], while this work addresses these functions briefly in Section 3.6.2.

3.6.2 Discretisation processes

The scope of this section is three-fold: a spatial discretisation is introduced (for the mesh and for the MPs), finite-dimensional spaces for weighting and trial functions are described, and time-discretisation consistent with the conservation laws is motivated and introduced.

3.6.2.1 Space and field discretisation

Let a regular Cartesian grid (or mesh) discretise the Euclidean space $\mathcal{E}$. This results in a discretisation of the body $\mathcal{B}$ too, where, in particular, the discretised volume is given by $^h\omega = \cup_{e=1}^{N^{els}} \omega^e$, with $e = 1 \ldots N^{els}$ indicating each active element of the grid. Let the global nodes of these elements be denoted by $\mathsf{A}, \mathsf{B}, \mathsf{C} = 1 \ldots N^{nds}$ and their relative (vectorial) DoFs be $I, J, K = 1 \ldots n^{dim} \times N^{nds}$. This work adopts the following linear polynomial functions defined on the Cartesian grid for the one-dimensional case*

$$\mathrm{N}(\xi) = \begin{cases} 1 + \xi/h, & \text{if} \quad -h < \xi \le 0; \\ 1 - \xi/h, & \text{if} \quad 0 < \xi \le h, \end{cases} \tag{3.107}$$

where $\xi \in (-h, h]$ denotes the local coordinate and h is the distance (constant) between two nodes of the mesh. Higher-dimensional functions for the Cartesian mesh are built by the tensor product of Eq. (3.107) in the different considered directions.

In the MPM, however, the mesh discretisation does not entirely define the body under consideration. A second discretisation for the body $\mathcal{B}$ is given by a set of MPs so that $^{MP}\omega = \cup_{mp=1}^{N^{mps}} v^{mp}$, where v^{mp} is the current volume of a single MP. This second discretisation is pivotal in the MPM as it guarantees continuity between time-steps, when the

---

*Higher-order functions are also available in the literature for the MPM, such as B-splines (see, for instance, Gan *et al.* [31]) or *Local Maximum-Entropy* shape functions (see, for instance, Molinos *et al.* [32]). However, as the GIMPM serves the purpose of having more continuity than $H^1(\omega)$ at the interfaces across the elements (see Eq. (3.109)), these higher-order functions are not considered in this work.

grid is decommissioned. In the literature, the functions operating the MP-based discretisation are called the *characteristic functions*. The original MPM uses *Dirac delta distributions* while the GIMPM adopts step functions (piecewise constant polynomials, defined over a characteristic length whose length is $2\,l_p$) for the same purposes.

Basis and shape functions discretising the finite-dimensional weighting and trial spaces and the considered geometry are given by the convolution of the functions defined on the grid with the characteristic functions described above[*]. This results in the following basis functions for the MPM (in $1D$)

$$
\mathrm{N}\left(\xi\right) =
\begin{cases}
1 + \xi/h, & \text{if} \quad -h < \xi \le 0; \\
1 - \xi/h, & \text{if} \quad 0 < \xi \le h,
\end{cases}
\tag{3.108}
$$

and for the GIMPM

$$
\mathrm{S}\left(\xi\right) =
\begin{cases}
\left(h + l_p + \xi\right)^2 / \left(4hl_p\right), & \text{if} \quad -h - l_p < \xi \le -h + l_p; \\
1 + \xi/h, & \text{if} \quad -h + l_p < \xi \le -l_p; \\
1 - \xi^2 / \left(2hl_p\right) + l_p^2 / \left(2hl_p\right), & \text{if} \quad -l_p < \xi \le l_p; \\
1 - \xi/h, & \text{if} \quad l_p < \xi \le h - l_p, \\
\left(h + l_p - \xi\right)^2 / \left(4hl_p\right), & \text{if} \quad h - l_p < \xi \le h + l_p.
\end{cases}
\tag{3.109}
$$

It can be seen that the convolution of Eq. (3.107) with the Dirac delta distribution leaves the function defined on the Cartesian mesh unaffected, while the same procedure involving the constant characteristic function produces second and first order piecewise polynomials (see Figure 3.4 for a qualitative illustration of the basis functions and their stencils). The choice to favour GIMPM basis functions over the original MPM ones is justified by the reasons given above with the discussion on the Sobolev spaces.

In the vast majority of the cases, the grid-based discretisation is bigger than the MP-based one, i.e., $^h\omega \supseteq {}^{MP}\omega$. This motivates another assumption which, although restrictive, proves effective. Let us formulate that the grid-discretisation is *conforming* to the MP-based discretisation at least on the Dirichlet boundary, i.e., $^h\gamma^{\bar{u}} = {}^{MP}\gamma^{\bar{u}}$. Two major simplifications come with this assumption: Dirichlet BCs can be applied directly on the mesh, and reaction tractions arising on the Dirichlet boundary $\gamma^{\bar{u}}$ are not included in the weak formulation thanks to the zero value of the weighting functions on that part of the grid (compare with Eq. (3.104)). The reader is referred to the discussion in Section 3.8 for existing techniques to apply non-conforming BCs.

The above assumption imposes some restrictions on the Dirichlet boundary, which is the easiest boundary to track during a simulation, particularly if homogeneous Dirichlet BCs are applied. If non-homogeneous Dirichlet BCs are prescribed, the grid can be moved to accommodate the above hypothesis of conforming Dirichlet boundary. However, displacement-free (Neumann) boundaries are more challenging, as their shape is guided by their (unknown) displacements. This leads to two challenges: tracking the MP-based Neumann boundary

---

[*]The reader interested in this integration process can refer to Bardenhagen and Kober [28].
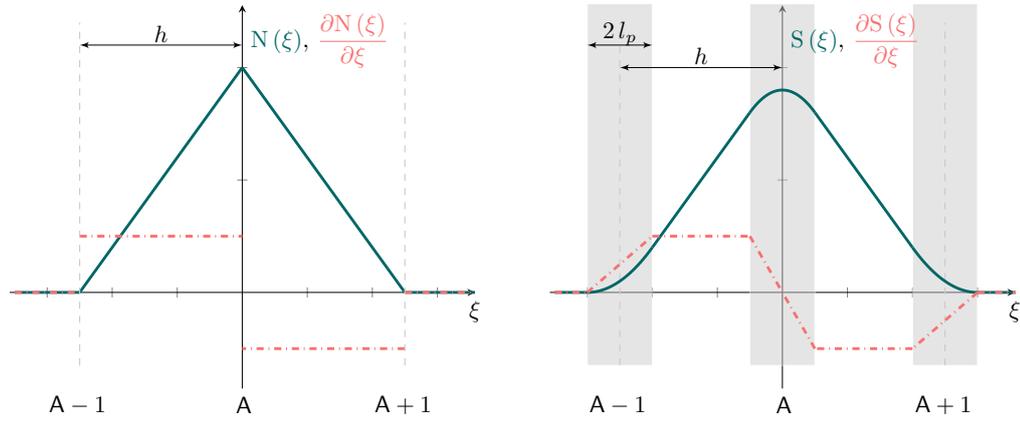
(a) Linear polynomial basis function defined by Eq. (3.108).

(b) Convolution of linear polynomial function with step characteristic function defined by Eq. (3.109).
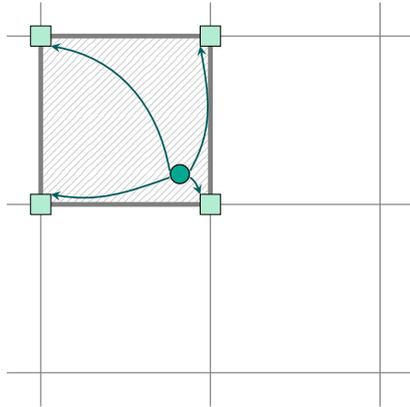
(c) A MP (teal coloured) with relative active elements and nodes for the MPM.
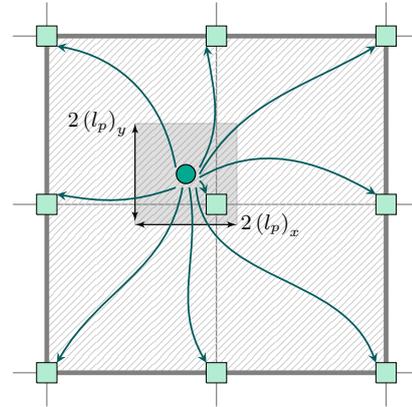
(d) A MP (teal coloured) with relative active elements and nodes for the GIMPM. The characteristic domain is also illustrated.

Figure 3.4: Top figures: Qualitative illustration of the different basis functions (and derivatives)used in MPM and GIMPM formulations. Basis functions are centred at the A-th node of the mesh.
Bottom figures: illustration of the MPM and GIMPM stencils with relative active elements and nodes.

is non-trivial (see Bing *et al.* [33], Yamaguchi *et al.* [34] or Remmerswaal [35] for procedures that serve this purpose), and boundary elements can be poorly populated by MPs. In this latter scenario, the situation can be particularly exacerbated if the ratio between the volumes of the MPs and the volume of the element is small. In the literature, this issue goes under the name of the *small-cut* problem (see, for instance, Sticko *et al.* [36] for the use of this nomenclature). Consequences of poorly-integrated elements are manifested by ill-conditioned matrices, as the entries relative to these elements are small compared to adequately-integrated elements. Since it is not possible to invert ill-conditioned matrices without encountering numerical difficulties, Section 4.6.4.1 reports a possible solution to this kind of issue.

Having introduced the basis functions, it is possible to define the finite-dimensional trial

and weighting function spaces as follows

$$
{}^{h}\mathscr{V}^{\boldsymbol{u}} = \left\{ {}^{h}u_i\left(\boldsymbol{X}, t\right) = \mathrm{N}^u_{iI}\left(\boldsymbol{X}\right) \mathrm{u}_I\left(t\right) \mid {}^{h}u_i = \bar{u}_i\left(\mathrm{x}_I\right) \text{ on } {}^{h}\gamma^{\bar{u}} \right\}; \tag{3.110}
$$

$$
{}^{h}\mathscr{W}^{\boldsymbol{w}} = \left\{ {}^{h}w_i\left(\boldsymbol{X}\right) = \mathrm{N}^u_{iI}\left(\boldsymbol{X}\right) \mathrm{w}_I \mid {}^{h}w_i = \boldsymbol{0} \text{ on } {}^{h}\gamma^{\bar{u}} \right\} \tag{3.111}
$$

where the basis functions $\mathbf{N}^u$ can be either given by Eq. (3.108) for the MPM or by Eq. (3.109) for the GIMPM.

The discrete weak form of the strong Eqs. (3.105) and (3.106) is given by finding ${}^{h}\boldsymbol{u} \in {}^{h}\mathscr{V}^{\boldsymbol{u}}$ such that, for $t \in [0, T]$,

$$
\int_{{}^{h}\omega} \frac{\partial \mathrm{N}^u_{Ii}}{\partial x_j}\,\sigma_{ij}\,dv - \int_{{}^{h}\omega} \rho\,\mathrm{N}^u_{Ii}\left(b_i - \mathrm{N}^u_{iK}\,\ddot{\mathrm{u}}_{\mathrm{K}}\right)\,dv - \int_{{}^{h}\gamma_{\bar{t}}} \mathrm{N}^u_{Ii}\,\bar{t}_i\,da \approx \boldsymbol{0}, \tag{3.112}
$$

or, equivalently,

$$
\int_{{}^{h}\Omega} \frac{\partial \mathrm{N}^u_{Ii}}{\partial X_j}\,F_{ip}S_{pj}\,dV - \int_{{}^{h}\Omega} \rho_0\,\mathrm{N}^u_{Ii}\left(b_i - \mathrm{N}^u_{iK}\,\ddot{\mathrm{u}}_{\mathrm{K}}\right)\,dV - \int_{{}^{h}\Gamma_{\bar{t}}} \mathrm{N}^u_{Ii}\,\bar{t}^0_i\,dA \approx \boldsymbol{0}, \tag{3.113}
$$

given the initial conditions ${}^{h}\boldsymbol{u}\left(t = 0\right) = \boldsymbol{u}_0$ and ${}^{h}\dot{\boldsymbol{u}}\left(t = 0\right) = \boldsymbol{v}_0$. It can be seen from the above equation that the nodal values of the weighting functions have already eliminated. The above equations can be concisely written in a matrix-vector system whose entries are as follows

$$
\mathrm{r}_I\left(\mathbf{u}\right) = \mathrm{f}^{int}_I + \mathrm{M}_{IK}\,\ddot{\mathrm{u}}_{\mathrm{K}} - \mathrm{f}^{ext}_I \approx \boldsymbol{0}; \tag{3.114}
$$

$$
\mathrm{R}_I\left(\mathbf{u}\right) = \mathrm{F}^{int}_I + \mathrm{M}_{IK}\,\ddot{\mathrm{u}}_{\mathrm{K}} - \mathrm{F}^{ext}_I \approx \boldsymbol{0}, \tag{3.115}
$$

where the the *internal force vector*, the *external force vector*, the *consistent mass matrix* and the *nodal acceleration vector* have been introduced. Detailed definitions of the above quantities are given below

$$
\mathrm{f}^{int}_I := \int_{{}^{h}\omega} \frac{\partial \mathrm{N}^u_{Ii}}{\partial x_j}\,\sigma_{ij}dv \approx \sum_{mp}^{N^{mps}} v^{mp} \frac{\partial \mathrm{N}^u_{Ii}\left(\boldsymbol{x}^{mp}\right)}{\partial x_j}\,\sigma_{ij}\left(\boldsymbol{x}^{mp}\right); \tag{3.116}
$$

$$
\mathrm{F}^{int}_I := \int_{{}^{h}\Omega} \frac{\partial \mathrm{N}^u_{Ii}}{\partial X_j}\,F_{ip}S_{pj}\,d^hV \approx \sum_{mp}^{N^{mps}} V^{mp} \frac{\partial \mathrm{N}^u_{Ii}\left(\boldsymbol{x}^{mp}\right)}{\partial X_j}\,F_{ip}\left(\boldsymbol{x}^{mp}\right)S_{pj}\left(\boldsymbol{x}^{mp}\right); \tag{3.117}
$$

$$
\mathrm{M}_{IK} := \int_{{}^{h}\omega} \rho\,\mathrm{N}^u_{Ii}\mathrm{N}^u_{iK}dv = \int_{{}^{h}\Omega} \rho_0\,\mathrm{N}^u_{Ii}\mathrm{N}^u_{iK}dV \approx \sum_{mp}^{N^{mps}} \bar{m}^{mp}\mathrm{N}^u_{Ii}\left(\boldsymbol{x}^{mp}\right)\mathrm{N}^u_{iK}\left(\boldsymbol{x}^{mp}\right); \tag{3.118}
$$

$$
\mathrm{f}^{ext}_I := \int_{{}^{h}\omega} \rho\,\mathrm{N}^u_{Ii}\,b_i\,dv + \int_{{}^{h}\gamma^{\bar{t}}} \mathrm{N}^u_{Ii}\,\bar{t}_i\,dA \approx \sum_{mp}^{N^{mps}} \bar{m}^{mp}\,\mathrm{N}^u_{Ii}\left(\boldsymbol{x}^{mp}\right)b_i + \int_{{}^{h}\gamma^{\bar{t}}} \mathrm{N}^u_{Ii}\,\bar{t}_i\,da; \tag{3.119}
$$

$$
\mathrm{F}^{ext}_I := \int_{{}^{h}\Omega} \rho_0\,\mathrm{N}^u_{Ii}\,b_i\,dV + \int_{{}^{h}\Gamma^{\bar{t}}} \mathrm{N}^u_{Ii}\,\bar{t}^0_i\,dA \approx \sum_{mp}^{N^{mps}} \bar{m}^{mp}\mathrm{N}^u_{Ii}\left(\boldsymbol{x}^{mp}\right)b_i + \int_{{}^{h}\Gamma_0} \mathrm{N}^u_{Ii}\,\bar{t}^0_i\,dA,
$$
$$
\tag{3.120}
$$

where $\bar{m}$ is the mass, while the other quantities have been previously defined in this chapter. From the above set of equations, it can be seen that their RHSs represent the (approxim-

ated) numerical integrals in the standard MPM, where MPs are used as quadrature points*. Quantities appearing in Eqs. (3.116)-(3.120) and denoted by the superscript $(\bullet)^{mp}$ are meant to be computed at the MPs' locations.

Several other observations can be made from the above list of equations. First, Eqs. (3.119) and (3.120) are not entirely numerically integrated. The reason underlying this missing explicit discretisation lies in the explanation provided above for tracking Neumann boundaries. Second, the internal and external force vectors have been denoted differently depending on the considered configuration ($\mathbf{f}^{int}$ and $\mathbf{f}^{ext}$ in the UL formulation and $\mathbf{F}^{int}$ and $\mathbf{F}^{ext}$ in the TL one) to distinguish them more conveniently. However, as demonstrated for the case of their strong formulations, there is a unique correspondence between these vectors. Third, owing to the definition of the consistent mass matrix (3.118) and by its numerical integration via MPs, MPM algorithms inherently conserve mass in the case of solid dynamics insofar as volume and densities are related via Eq. (3.46). Since MPs' masses are constant, they are algorithmically conserved across different time-steps. The summation of all of these MPs' masses returns the total mass of the body $\mathcal{B}$.

3.6.2.2 Time discretisation

So far, we have considered the continuum body of interest in a temporal interval $[0, T]$, without any further characterisation. As standard in any temporal discretisation, the total time is divided into smaller time-steps, whose generic length is Δt. At any of these time-steps, the solution (and all of the related variables) are known at the previously converged time t (the $n-$th time-step), while the solution at the current time $t+\Delta t$ (the $n+1-$th time-step) remains unknown . Depending on the considered approach (implicit or explicit, or a combination of these), this solution is calculated and the relevant variables updated so that a new time-step can subsequently be undertaken. As previously mentioned in Chapter 2, this work considers an implicit approach.

Subdividing the time under consideration in smaller time-steps is not sufficient when introducing a time-dependent problem, as it is the case of solid dynamics. A further decision must be taken on how to relate the temporal derivatives to each other. As Simo and Tarnow [38] demonstrated, not all choices are equal if it is desired to respect the conservation laws (linear and angular momenta, and Clausius-Planck inequality). Since the balance of rate of linear momentum is the primary discretised equation, it can be understood that this equation is always satisfied, independently from how time derivatives are related. However, the same conclusion cannot be drawn for the conservation of angular momentum, as this equation is not explicitly discretised. Among all the possible time discretisations available, Simo and Tarnow [38] have shown that the *mid-point* rule is the only one that conserves angular momentum for a solid body undergoing finite strains in the dynamic context. As

---

*In the MPM literature, some other approximations of the integrals are available. Among these, it is worthy of mention the *improved* MPM by Sulsky and Gong [37], where the integrals are approximated using a moving-least-square procedure. Another method has been provided by Gan *et al.* [31], where MPs are used to guarantee continuity among the time-steps, and integrand functions are moved from MPs' positions (where they lie) to Gauss Points locations via B-splines.
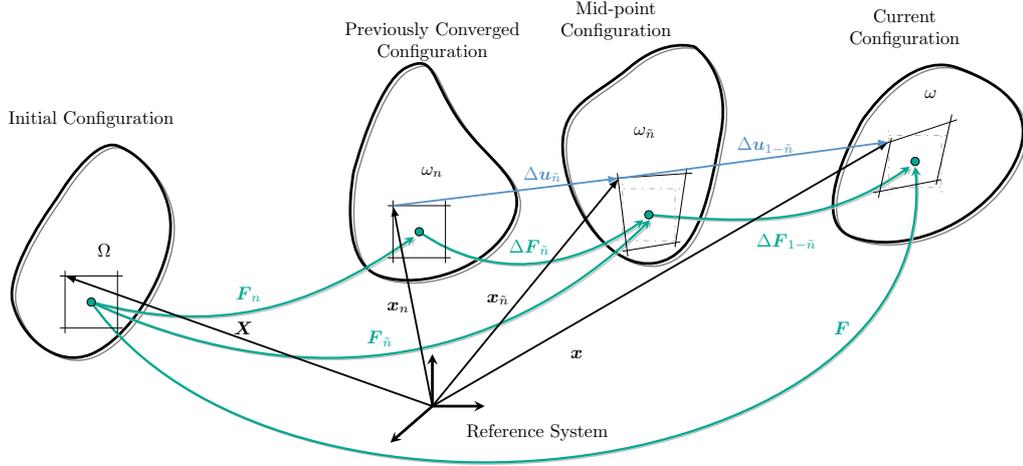
Figure 3.5: Configurations taken into account by the current MPM formulation. Grid nodal positions, grid (incremental) displacements, and MP deformation gradients are shown in each configuration. Colours relative to the different configuration has been removed for the sake of clarity. Figure modified from Pretti *et al.* [26].

such, the mid-point rule is the adopted temporal discretisation in this chapter. This method stipulates that a fourth configuration, denoted at the $\tilde{n}-$th, is added to the standard three[*] introduced so far, as illustrated in Figure 3.5. This mid-point configuration is defined by the following displacement field

$$
\begin{aligned}
\mathbf{u}_{\tilde{n}} &:= \frac{1}{2}\left(\mathbf{u}_n + \mathbf{u}_{n+1}\right) \\
&= \frac{\mathbf{u}_{n+1}}{2} \\
&= \Delta\mathbf{u}_{\tilde{n}},
\end{aligned}
\tag{3.121}
$$

where the second and third lines of the above equation hold since the mesh is reset at each time-step in the MPM, i.e., $\mathbf{u}_n = \mathbf{0}$. Furthermore, it is assumed that the velocity and the acceleration defined on the same configuration are given by

$$
\mathbf{v}_{\tilde{n}} := \frac{\mathbf{u}_{n+1} - \mathbf{u}_n}{\Delta t};
\tag{3.122}
$$

$$
\mathbf{a}_{\tilde{n}} := \frac{\mathbf{v}_{n+1} - \mathbf{v}_n}{\Delta t}.
\tag{3.123}
$$

The total deformation gradient and its time-discrete versions follow the usual chain rule, giving

$$
\boldsymbol{F} := \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}} = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}_{\tilde{n}}} \frac{\partial \boldsymbol{x}_{\tilde{n}}}{\partial \boldsymbol{x}_n} \frac{\partial \boldsymbol{x}_n}{\partial \boldsymbol{X}} := \Delta\boldsymbol{F}_{1-\tilde{n}} \,\Delta\boldsymbol{F}_{\tilde{n}} \,\boldsymbol{F}_n,
\tag{3.124}
$$

where, following Eq. (3.121), the mid-point position is given by

$$
\boldsymbol{x}_{\tilde{n}} = \boldsymbol{x}_n + \mathbf{u}_{\tilde{n}}.
\tag{3.125}
$$

---

[*]For the sake of clarity, the three configurations introduced so far are the original $(0-th$ time-step), the current $(n+1$-th time-step), and the previously converged ones ($n-$th time-step).

It should also be emphasised that, generally speaking, the notation $\Delta\left(\bullet\right)$ (without any further subscript) implies $\left(\bullet\right)_{n+1}-\left(\bullet\right)_{n}$, with the exception of the deformation gradients, which refers to the above relationship.

Simo and Tarnow [38] also proved that no time-discretisation in the context of finite strain dynamics can implicitly satisfy the Clausius-Planck inequality. A modification of the internal force vector is therefore required, this being detailed in Section 3.6.3 for the UL formulation.

It should also be noted that the UL formulation appears less artificial than its TL counterpart in the MPM. As can be appreciated from Figure 3.5, the grid in the initial configuration does not exist in any of the other three configurations. Conversely, the grid present in these configurations (the previously converged, the mid-point and the current configurations) is the same, which deforms progressively. Thus, while it is still possible to map all the necessary quantities back to the original configuration, it seems rather fictitious.

Adaptive time-step length base on the CFL condition   Regardless of the type of temporal discretisation and whether it is explicit or implicit, choosing the time-step size of a simulation involves a compromise between stability and efficiency. In the explicit Finite Element Method (FEM), the maximum time-step size is governed by the well-known Courant-Friedrichs-Lewy (CFL) condition*

$$\Delta t^{CFL} = \min_{i}\left(h_{i}^{h}\right)\sqrt{\frac{\rho}{M}}, \tag{3.126}$$

where h_{i}^{h} is the grid size in the $i-$th direction $(\in \mathbb{R}^{n^{dim}})$, while ρ and M are the density and the P-wave modulus of the material being analysed. This equation is known to be true under the hypothesis of small strain theory or for an undeformed Hencky materials in the context of finite strain theory. While the adaptation of this formula to a deformed body undergoing finite strain has been proposed by Sun *et al.* [40] (i.e., the change in material properties is considered), it is here remarked (based on the work of Pretti *et al.* [26]) how the CFL condition is not a steady quantity in the MPM due to geometrical reasons. As the analysis proceeds, the way that the MP-based discretisation intersects the grid changes, leading to differently partially filled elements. It can be understood that these partially filled elements offer a different discretisation to entirely filled elements, both in terms of density and mechanical properties (volumetric stiffness and density) of the material. In particular, the grid-averaged values of these quantities coming from partially filled elements are smaller than those coming from entirely filled elements, leading to different grid-averaged values of the CFL condition. Similar considerations have already been recognised in the context of cut-FEM† (see, for instance, Sticko *et al.* [36]), where partially filled elements reduce the time-step size given by Eq. (3.126).

---

*The original CFL contribution (see [39]) referred to the finite-difference method. However, it can be easily extended to the FEM.

†The cut-FEM method is an embedded/immersed domain approach, combining the fictitious domain method with finite elements. Its basic idea is to immerse the considered body into a simple unfitted structured mesh with finite element basis functions. The reader interested in this subject is referred to, for instance, Schillinger and Ruess [41].

While Sun *et al.* [40] have considered the material variation and Sticko *et al.* [36] the geometrical change in rigorous ways (these works both rely on explicit solution techniques), this work suggests an approximation for the CFL condition which could act as a rule of thumb, especially for implicit solvers, given by

$$\Delta t^{CFL} \approx \min_i \left(h_i^h \right) \min_{\mathsf{A}} \left(\sqrt{\frac{m^h}{v^h \, M^h}} \right)_{\mathsf{A}}, \tag{3.127}$$

where the above grid quantities are given by their sum over MPs as follows

$$m_{\mathsf{A}}^h = \sum_{mp}^{\mathrm{N}^{mps}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp} \right) m^{mp}, \qquad v_{\mathsf{A}}^h = \sum_{mp}^{\mathrm{N}^{mps}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp} \right) v^{mp}, \qquad M_{\mathsf{A}}^h = \sum_{mp}^{\mathrm{N}^{mps}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp} \right) M^{mp}. \tag{3.128}$$

In particular, Eq. (3.127) considers the geometrical variability during a simulation in an approximate way, and allows the evolving CFL condition to be efficiently approximated throughout the analysis, thus providing time-step adaptability.

The CFL condition is not necessary to guarantee method stability for implicit temporal discretisations, but it can be employed to provide an adaptive time step length for efficiency purposes. Quantifying (even approximately) how this condition varies spatially during the simulation is therefore, if not necessary, at least helpful for adapting the time step length. Overall, Eq. (3.127) has to be regarded as a rule of thumb more than as a stringent rule, making it nonetheless helpful for adapting the time-step in implicit schemes.

3.6.3 Conservation Law Consistent (CLC) internal force vector

This section heavily draws on Pretti *et al.* [26], and considers the modification of the internal force vector proposed by Simo and Tarnow [38] for a TL formulation, i.e.,

$$\tilde{\mathbf{F}}_{\tilde{n}}^{int} = \int_{\Omega} \frac{\partial \left(\mathbf{N}^u \right)^T}{\partial \boldsymbol{X}} : \left(\boldsymbol{F}_{\tilde{n}} \, \boldsymbol{S}^{mp} \right) \, dV, \tag{3.129}$$

where the modified second Piola-Kirchhoff tensor $\boldsymbol{S}^{mp}$ will be detailed below.

To write a UL formulation equivalent to the above internal force vector, the following relationships given by the chain rule are useful

$$\frac{\partial \mathbf{N}}{\partial \boldsymbol{X}} = \frac{\partial \mathbf{N}}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}}; \tag{3.130}$$

$$\boldsymbol{F}_{\tilde{n}} = \frac{\partial \boldsymbol{x}_{\tilde{n}}}{\partial \boldsymbol{X}} = \frac{\partial \boldsymbol{x}_{\tilde{n}}}{\partial \boldsymbol{x}_n} \frac{\partial \boldsymbol{x}_n}{\partial \boldsymbol{X}} = \Delta \boldsymbol{F}_{\tilde{n}} \, \boldsymbol{F}_n. \tag{3.131}$$

According to Eqs. (3.56) and (3.58), the modified Kirchhoff stress relates to the modified second Piola-Kirchhoff stress according to

$$\boldsymbol{S}^{mp} = \left(\boldsymbol{F} \right)^{-1} \boldsymbol{\tau}^{mp} \left(\boldsymbol{F} \right)^{-T}. \tag{3.132}$$

Using Eqs. (3.130)-(3.132), the modified internal force vector (3.129) can be described in an UL formulation as follows

$$\tilde{\mathbf{f}}_{\tilde{n}}^{int} = \int_{\Omega} \frac{\partial \left(\mathbf{N}^u \right)^T}{\partial \boldsymbol{x}} \underbrace{\Delta \boldsymbol{F}_{\tilde{n}} \, \boldsymbol{F}_n \left(\boldsymbol{F} \right)^{-1}}_{= \left(\Delta \boldsymbol{F}_{1-\tilde{n}} \right)^{-1}} \boldsymbol{\tau}^{mp} dV. \tag{3.133}$$

It should be noted that the above internal force vector contains a non-symmetric measure of stress, this being $(\Delta \boldsymbol{F}_{1-\tilde{n}})^{-1} \boldsymbol{\tau}^{mp}$. This is the result of the discrepancy between where the equilibrium is imposed (i.e., the intermediate configuration) and where the modified Kirchhoff stress is defined (the current configuration). In other words, the stress $(\Delta \boldsymbol{F}_{1-\tilde{n}})^{-1} \boldsymbol{\tau}^{mp}$ is a two-point tensor. It should be also noted that nothing prevents Eq. (3.132) from referring to a Kirchhoff-type stress mapped to the mid-point configuration. However, the decision to push the modified stress to the current configuration seems more streamlined, especially since the linearisation of $\boldsymbol{\tau}^{mp}$ is independent from the considered intermediate configuration, which, if time integration schemes different from the mid-point rule are considered, can vary between the previously converged and the current one.

3.6.3.1 Stress-strain relationship modification

While Simo and Tarnow [38] proposed a specific stress-strain relationship for the modified stress, this work adopts the more general definition proposed by Gonzalez [42], including a term to include plasticity added by Meng and Laursen [43], which results in the following formula*

$$\boldsymbol{S}^{mp} = \frac{1}{2}\left(\boldsymbol{S}_n + \boldsymbol{S}_{n+1}\right) + \frac{\Delta \boldsymbol{C}}{\|\Delta \boldsymbol{C}\|^2} \underbrace{\left(2\left(\Delta \Psi + \Delta \mathcal{D}\right) - \frac{1}{2}\left(\boldsymbol{S}_n + \boldsymbol{S}_{n+1}\right) : \Delta \boldsymbol{C}\right)}_{:=\Delta \mathscr{H}^{int}}, \qquad (3.134)$$

where it can be seen that the quantity $\Delta \mathscr{H}^{int}$ is built so that the power performed by the stress-strain power conjugates is equal to the time-difference of plastic dissipation and free energy function (compare with Eq. (3.78)).

While a relationship between the modified second Piola-Kirchhoff stress $\boldsymbol{S}^{mp}$ and the modified Kirchhoff stress $\boldsymbol{\tau}^{mp}$ has been introduced by Eq. (3.132), a relationship between their respective time-discrete conjugate kinematic quantities is missing. This can be accomplished by recalling how the discrete powers over the time-step relates, i.e.,

$$\frac{1}{2}\boldsymbol{S}^{mp} : \Delta \boldsymbol{C} = \boldsymbol{\tau}^{mp} : \bar{\boldsymbol{d}}, \qquad (3.135)$$

where $\bar{\boldsymbol{d}}$ indicates the symmetric averaged (over the time-step) part of the velocity gradient. The use of Eq. (3.132) defines $\bar{\boldsymbol{d}}$ as follows

$$\bar{\boldsymbol{d}} := \frac{1}{2}\boldsymbol{F}^{-T} \Delta \boldsymbol{C} \boldsymbol{F}^{-1}. \qquad (3.136)$$

The time-difference in internal dissipation $\Delta \mathcal{D}$ is the time-discrete version of (3.85b). If *associative flow rules* are considered (see Eqs. (3.91) and (3.92)), $\Delta \mathcal{D}$ can be expressed by

$$\Delta \mathcal{D} = \Delta \gamma^p \left(\boldsymbol{\tau}_{n+1} : \frac{\partial \Phi}{\partial \boldsymbol{\tau}_{n+1}} + \boldsymbol{q}_{n+1} * \frac{\partial \Phi}{\partial \boldsymbol{q}_{n+1}}\right). \qquad (3.137)$$

---

*The modified stress-strain relationship (3.134) fulfils the properties of *directionality* and *consistency* between continuous and discrete derivatives. For the meaning and verification of these properties, the reader is referred to Gonzalez [42], Armero and Romero [44], and Love and Sulsky [45].

If Eq. (3.132) is substituted into Eq. (3.134), the modified Kirchhoff stress is given by

$$
\begin{aligned}
\boldsymbol{\tau}^{mp} &= \boldsymbol{F}\,\boldsymbol{S}^{mp}\,\boldsymbol{F}^T \\
&= \frac{1}{2}\boldsymbol{F}\,(\boldsymbol{S}_n + \boldsymbol{S}_{n+1})\,\boldsymbol{F}^T + \boldsymbol{F}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^T\left(\frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\right) \\
&= \frac{1}{2}\underbrace{\boldsymbol{F}\,\boldsymbol{S}_n\,\boldsymbol{F}^T}_{=\Delta\boldsymbol{F}\,\boldsymbol{\tau}_n\,\Delta\boldsymbol{F}^T} + \frac{1}{2}\boldsymbol{\tau}_{n+1} + \boldsymbol{F}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^T\left(\frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\right) \\
&= \bar{\boldsymbol{\tau}} + \boldsymbol{F}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^T\left(\frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\right),
\end{aligned}
\tag{3.138}
$$

where the trapezoidal stress $\bar{\boldsymbol{\tau}}$ is defined by

$$
\bar{\boldsymbol{\tau}} := \frac{1}{2}\boldsymbol{F}\,(\boldsymbol{S}_n + \boldsymbol{S}_{n+1})\,\boldsymbol{F}^T = \frac{1}{2}\left(\Delta\boldsymbol{F}\,\boldsymbol{\tau}_n\,\Delta\boldsymbol{F}^T + \boldsymbol{\tau}_{n+1}\right) = \frac{1}{2}\left(\boldsymbol{\tau}_n^{PF} + \boldsymbol{\tau}_{n+1}\right),
\tag{3.139}
$$

with the stress $\boldsymbol{\tau}_n^{PF}$ (i.e., the previously converged Kirchhoff stress mapped into the reference configuration) given by

$$
\boldsymbol{\tau}_n^{PF} = \Delta\boldsymbol{F}\,\boldsymbol{\tau}_n\,\Delta\boldsymbol{F}^T.
\tag{3.140}
$$

Another kinematic quantity of interest appearing in Eq. (3.138) can be further analysed

$$
\begin{aligned}
F_{ih}\,(\Delta C)_{hk}\,F_{jk} &= F_{ih}\,(C_{hk} - (C_n)_{hk})\,F_{jk} \\
&= F_{ih}\,F_{\alpha h}\,F_{\alpha k}\,F_{jk} - \Delta F_{i\beta}\,(F_n)_{\beta h}\,(F_n)_{\alpha h}\,(F_n)_{\alpha k}\,\Delta F_{j\gamma}\,(F_n)_{\gamma k} \\
&= b_{i\alpha}\,b_{\alpha j} - \Delta F_{i\beta}\,(b_n)_{\beta\alpha}\,(b_n)_{\alpha\gamma}\,\Delta F_{j\gamma} := \Delta^{PF}b_{ij},
\end{aligned}
\tag{3.141}
$$

where, when the reference configuration is omitted, the quantities are referred to the current time $t+\Delta t$ ($n+1$-th time-step). The above equation expresses the quantity $\boldsymbol{F}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^T$ as a difference between the squared current left Cauchy-Green strain tensor ($\boldsymbol{b}^2$) and the same quantity evaluated at the previous step ($\boldsymbol{b}_t^2$) which has been pushed forward to the current configuration. Thus, the modified Kirchhoff stress can be concisely written as follows

$$
\boldsymbol{\tau}^{mp} = \bar{\boldsymbol{\tau}} + \Delta^{PF}\boldsymbol{b}\left(\frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\right).
\tag{3.142}
$$

The above equation, when plugged into Eq. (3.133), provides a Conservation Law Consistent (CLC) Updated Lagrangian (UL) formulation of the internal force vector.

It can also be proved that the modified stress-strain relationships respects the Clausius-Planck inequality (3.81), whose time-discrete form is given by

$$
\frac{1}{2}\,\boldsymbol{S}^{mp} : \Delta\boldsymbol{C} = \boldsymbol{\tau}^{mp} : \bar{\boldsymbol{d}} = \Delta\Psi + \Delta\mathcal{D} \geq \Delta\mathcal{D}.
\tag{3.143}
$$

From Eqs. (3.136) and (3.138), it can be obtained that

$$
\begin{aligned}
\boldsymbol{\tau}^{mp} : \bar{\boldsymbol{d}} &= \left(\frac{1}{2}\left(\boldsymbol{\tau}_n^{PF} + \boldsymbol{\tau}_{n+1}\right) + \boldsymbol{F}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^T\left(\frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\right)\right) : \frac{1}{2}\left(\boldsymbol{F}^{-T}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^{-1}\right) \\
&= \frac{1}{4}\left(\boldsymbol{\tau}_n^{PF} + \boldsymbol{\tau}_{n+1}\right) : \left(\boldsymbol{F}^{-T}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^{-1}\right) + \frac{1}{2}\frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\left(\boldsymbol{F}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^T : \boldsymbol{F}^{-T}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^{-1}\right).
\end{aligned}
\tag{3.144}
$$

The following simplifications can be made

$$\boldsymbol{F}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^T : \boldsymbol{F}^{-T}\,\Delta\boldsymbol{C}\,\boldsymbol{F}^{-1} = \Delta\boldsymbol{C} : \Delta\boldsymbol{C}, \tag{3.145}$$

$$\begin{aligned}
\frac{1}{2}\left(\boldsymbol{S}_n + \boldsymbol{S}_{n+1}\right) : \Delta\boldsymbol{C} &= \left(\boldsymbol{S}_n + \boldsymbol{S}_{n+1}\right) : \boldsymbol{F}^T\,\bar{\boldsymbol{d}}\,\boldsymbol{F} \\
&= \left(\Delta\boldsymbol{F}\,\boldsymbol{\tau}_n\,\Delta\boldsymbol{F}^T + \boldsymbol{\tau}_{n+1}\right) : \bar{\boldsymbol{d}} \\
&= 2\,\bar{\boldsymbol{\tau}} : \bar{\boldsymbol{d}}.
\end{aligned} \tag{3.146}$$

Therefore, the quantity $\Delta\mathscr{H}^{int}$ in the Eq. (3.142) can be rewritten as

$$\Delta\mathscr{H}^{int} = 2\left(\Delta\Psi + \Delta\mathcal{D}\right) - \frac{1}{2}\left(\boldsymbol{S}_n + \boldsymbol{S}_{n+1}\right) : \Delta\boldsymbol{C} = 2\left(\Delta\Psi + \Delta\mathcal{D}\right) - 2\,\bar{\boldsymbol{\tau}} : \bar{\boldsymbol{d}},$$

and inequality (3.144) can be proved

$$\begin{aligned}
\boldsymbol{\tau}^{mp} : \bar{\boldsymbol{d}} &= \frac{1}{2}\left(\boldsymbol{\tau}_n^{PF} + \boldsymbol{\tau}_{n+1}\right) : \bar{\boldsymbol{d}} + \frac{1}{2\,\|\Delta\boldsymbol{C}\|^2}\left(2\left(\Delta\Psi + \Delta\mathcal{D}\right) - 2\,\bar{\boldsymbol{\tau}} : \bar{\boldsymbol{d}}\right)\Delta\boldsymbol{C} : \Delta\boldsymbol{C} \\
&= \bar{\boldsymbol{\tau}} : \bar{\boldsymbol{d}} + \left(\Delta\Psi + \Delta\mathcal{D}\right) - \bar{\boldsymbol{\tau}} : \bar{\boldsymbol{d}} \\
&= \Delta\Psi + \Delta\mathcal{D} \geq \Delta\Psi.
\end{aligned} \tag{3.147}$$

3.6.4 Linearisation

Since the (modified) internal force vector is non-linear in terms of displacements and the mid-point is a (semi-)implicit discretisation, the grid displacement field $\mathbf{u}_{n+1}$ must be computed in an iterative way. For this purpose, let us recall the UL matrix-vector system (3.114) where the modified internal force vector given by Eq. (3.133) has been used

$$\left(\mathrm{r}_{\tilde{n}}\right)_I(\mathbf{u}) = \left(\tilde{\mathrm{f}}_{\tilde{n}}^{int}\right)_I + \mathrm{M}_{IK}\,\left(\mathrm{a}_{\tilde{n}}\right)_K - \left(\mathrm{f}_{\tilde{n}}^{ext}\right)_I \approx \mathbf{0}. \tag{3.148}$$

The Newton-Raphson (NR) method requires the writing of Taylor expansion of the above residual $\mathbf{r}_{\tilde{n}}$ in correspondence with the previous iteration (denoted by $(\bullet)^{(k)}$) and to set that approximation equal to zero, i.e.,

$$\mathbf{r}_{\tilde{n}}^{(k+1)} \approx \mathbf{r}_{\tilde{n}}^{(k)} + \underbrace{\frac{\partial\mathbf{r}_{\tilde{n}}^{(k)}}{\partial\mathbf{u}_{n+1}}}_{:=\mathbf{J}}\underbrace{\left(\mathbf{u}_{n+1}^{(k+1)} - \mathbf{u}_{n+1}^{(k)}\right)}_{:=\delta\mathbf{u}} \approx \mathbf{0}, \tag{3.149}$$

where $\mathbf{J}$ defines the *Jacobian matrix* of the system, which, for the mid-point rule, is given by

$$\mathbf{J} = \frac{2}{\Delta t^2}\,\mathbf{M} + \mathbf{K}, \tag{3.150}$$

with the *stiffness matrix* given by $\mathbf{K} := \frac{\partial\tilde{\mathbf{f}}_{\tilde{n}}^{int}}{\partial\mathbf{u}_{n+1}}$. The full expression of the stiffness matrix for an (elasto-plastic) Hencky material described by the CLC stress-strain relationship Eq. (3.142) is reported in Appendix B.1.

Eq. (3.148) is solved for $\mathbf{u}_{n+1}^{(k+1)}$, which provides the new solution to compute a new residual. This procedure is repeated until a certain measure of error is deemed satisfactory (see, in this regard, Section 3.7).

3.6.5 Boundary stabilisation

As previously discussed when introducing the space discretisations in Section 3.6.2.1, Neumann boundaries are mostly non-matching between the MP-based and grid discretisations. This non-conformity can lead to the small-cut problem, especially when the ratio between the MP volumes to the volume of the element in which these MPs are in is small[*].

The literature offers a few techniques to mitigate this issue. The *aggregated* FEM (see Badia *et al.* [46]) requires that partially filled boundary elements are agglomerated to fully populated elements in the interior of the material. In this fashion, partially filled elements cannot contribute with extremely low entries in the matrices, avoiding ill-conditioning problems. To the author's best knowledge, this method has not been applied to the MPM. The gist of the *extended B-splines* (originally proposed by Höllig *et al.* [47] and extended to the MPM by Yamaguchi *et al.* [34]) is very similar to the one of aggregated FEM, in the sense that degenerate basis functions arising in the proximity of the boundaries are replaced by linear combinations of stable basis functions lying in the interior of the domain. The *ghost penalty stabilisation* was first proposed by Burman [48] in the context of cut-FEM, and its face penalty version[†] has been extended by Coombs [49] to the MPM. Since this last version of the ghost method has already been adapted to linear polynomials and GIMPM shape functions and has also already been tested in the MPM, this work adopts it.

The main idea of the face penalty ghost stabilisation method lies in the addition of a penalty term on the faces shared between partially filled elements (lying on the boundary) and fully populated elements in the interior of the material. This additional term penalises the jump between normal gradients (up to the order of approximation) of weighting and trial functions. In the case of MPM or GIMPM basis functions (defined by Eqs. (3.108) and (3.109)), this penalisation is as follows

$$(\mathrm{J}_G)_{IK} \coloneqq \frac{h^3}{3} \int_{\gamma_{(G)}} \left(\frac{\partial \mathrm{N}_{Ii}^{u;\,+}}{\partial\,(x_n)_j}\, n_j^h - \frac{\partial \mathrm{N}_{Ii}^{u;\,-}}{\partial\,(x_n)_j}\, n_j^h \right) \left(\frac{\partial \mathrm{N}_{iK}^{u;\,+}}{\partial\,(x_n)_k}\, n_k^h - \frac{\partial \mathrm{N}_{iK}^{u;\,-}}{\partial\,(x_n)_k}\, n_k^h \right) da, \quad (3.151)$$

where the superscripts of the shape functions $(\bullet)^{\pm}$ refer to different (positive or negative) elements relative to a considered face and $\boldsymbol{n}^h$ is the normal to the grid faces belonging to $\gamma_{(G)}$. The selection of the boundary $\gamma_{(G)}$ for the above integral as well as other details are thoroughly discussed in Coombs [49]. As it can be noticed from the gradient relative to the previously converged position $\boldsymbol{x}_n$, this works consider the above matrix constant in a given time-step[‡].

The term expressed by Eq. (3.151), appropriately scaled, is usually multiplied by the nodal displacements and added to the the discretised balance of rate of linear momentum Eq.

---

[*]Since the small-cut problem is a situation which can occur, it is recommended to apply a boundary stabilisation method in most scenarios. However, as it can be seen in the numerical analyses in Sections 3.7, 4.7 and 5.5, this stabilisation is not applied to all of the cases, particularly those where displacements are relatively small and MPs are not expected to migrate substantially from their original element.

[†]The original work by Burman [48] proposed a bulk penalty term and a face penalty term.

[‡]The normal to the considered grid face $\boldsymbol{n}^h$ should also change direction, moving in compliance with the displacement field. However, if ghost stabilisation is interpreted as an artifice for numerical purposes, it can be considered constant in the time-step (see Coombs [49]).

(3.148), adding a stabilising stiffness-like component, defined as

$$\mathbf{K}_G := \gamma^K \, \mathbf{J}_G. \tag{3.152}$$

The parameter γ^K is user-selected which, from the literature (see, again, Sticko *et al.* [36] or Coombs [49]), can vary considerably.

However, other stabilisations than that expressed by Eq. (3.152) could be required for multiple reasons. As can be seen from the Jacobian matrix (3.150), stabilising the stiffness parameter may not be sufficient, especially for low-frequency problems, where the mass matrix plays a bigger role as term in the Jacobian matrix (see Eq. (3.150)). Furthermore, an MPM algorithm describing the dynamic for a solid body requires mapping of the initial velocity from MPs to the grid at the beginning of the step (as is more clearly expressed in Section 3.6.6). This step requires inversion of the mass matrix which can suffer from the same small-cut deficiency of the stiffness. Owing to these reasons, a stabilised mass-like term must be considered too

$$\mathbf{M}_G := \gamma^M \, \mathbf{J}_G, \tag{3.153}$$

with γ^M being a second user-defined parameter.

Given the stabilisations (3.152) and (3.153), balance of rate of linear momentum Eq. (3.148) becomes

$$(\tilde{\mathrm{r}}_{\tilde{n}})_I \, (\mathbf{u}_{n+1}) = \left(\mathrm{f}_{\tilde{n}}^{int}\right)_I + (\mathrm{K}_G)_{IK} \, (\mathrm{u}_{\tilde{n}})_K + (\mathrm{M}_{IK} + \, (\mathrm{M}_G)_{IK}) \, (\mathrm{a}_{\tilde{n}})_K - \left(\mathrm{f}_{\tilde{n}}^{ext}\right)_I \approx \mathbf{0}. \tag{3.154}$$

It must be considered that the addition of the ghost stabilisations to the equilibrium equation has consequences on its compliance with the principles of thermodynamics since the equilibrium equation now includes energetically spurious terms. However, the coefficients γ^K and γ^M can be chosen to be a few orders of magnitude below the elastic parameters (in the case of γ^K) or the mass (in the case of γ^M). Hence, while an assessment of the introduction of the stabilisations is not quantified, it is recognised that its role is marginal due to the choice of the parameters.

3.6.6 Mapping processes: Particle In Cell (PIC) vs FLuid Implicit Particle (FLIP)

While the previously illustrated CLC formulation conserves the quantity of interest during the time-step, this is not sufficient to guarantee that these are preserved in an MPM algorithm. This condition is due to the introduction and decommission of the computational grid at each time-step. Thus, it is also necessary to check that the mapping processes at each time-step's beginning (Point-to-grid (P2G)) and end (Grid-to-point (G2P)) are equally conservative. This section is aimed specifically at this purpose. The Einstein index notation is suspended for this subsection to make the summations as clear as possible.

In the case of an elasto-plastic solid mechanics problem where kinetic energy plays a role, there are two forms of energy which should be conserved (see, in this regard, the first principle of thermodynamics, Eq. (3.60) and Clausius-Plank inequality (3.85)), i.e., the (elastic)

strain and the kinetic energies. As for the standard displacement-based MPM formulation, the former is not moved from the MPs, while the latter is mapped back and forth from MPs to the grid. Therefore, it is on kinetic energy that the calculations focus. However, this is not sufficient to guarantee the compliance of the mapping processes with the underlying continuum principles, as it is desired for linear and angular momentum to be conserved according to their conservation laws. To perform the calculations, let us define all of the above-mentioned quantities on the grid and the MPs, these being:

- grid linear momentum

$$\mathbf{L}^h := \sum_{\mathsf{A}}^{N^{nds}} \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} \, \mathbf{v}_{\mathsf{B}} \right); \tag{3.155}$$

- MPs' linear momentum

$$\boldsymbol{L}^{mp} := \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}^{mp}; \tag{3.156}$$

- grid angular momentum

$$\mathbf{J}^h := \sum_{\mathsf{A}}^{N^{nds}} \mathbf{x}_{\mathsf{A}} \times \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} \, \mathbf{v}_{\mathsf{B}} \right); \tag{3.157}$$

- MPs' angular momentum

$$\boldsymbol{J}^{mp} := \sum_{mp}^{N^{mps}} \boldsymbol{x}^{mp} \times \bar{m}^{mp} \boldsymbol{v}^{mp}; \tag{3.158}$$

- grid kinetic energy

$$\mathrm{K}^h := \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} \mathbf{v}_{\mathsf{A}} \cdot \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} \, \mathbf{v}_{\mathsf{B}} \right); \tag{3.159}$$

- MPs' kinetic energy

$$K^{mp} := \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \left\| \boldsymbol{v}^{mp} \right\|^2. \tag{3.160}$$

The above equations use the *effective* mass matrix, given by

$$\tilde{\mathbf{M}} := (1 - \epsilon) \, \mathbf{M} + \epsilon \, \check{\mathbf{M}} \qquad \text{with} \quad \epsilon = [0, 1], \tag{3.161}$$

where the *lumped mass matrix* $\check{\mathbf{M}}$ introduced above is defined as follows

$$\check{\mathbf{M}}_{IK} := \mathrm{diag}_I \left(\int_{{}^h\Omega} \rho \, \mathrm{N}^u_K \, dv \right) \approx \mathrm{diag}_I \left(\sum_{mp=1}^{N^{mps}} \bar{m}^{mp} \, \mathrm{N}^u_K \left(\boldsymbol{x}^{mp} \right) \right). \tag{3.162}$$

Using Eq. (3.161) in lieu of Eq. (3.118) allows us to evaluate at the same time a wider range of mass matrices (see, for instance, the orginal MPM paper [50] , which makes use of a lumped mass matrix).

3.6.6.1   Point-to-grid (P2G) mapping

The P2G mapping at the beginning of the step can be generalised by the following equation (see Love and Sulsky [25] or Pretti *et al.* [26])

$$\sum_{B}^{N^{nds}} \tilde{\mathbf{M}}_{AB} \, (\mathbf{v}_n)_B = \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \mathrm{N}_A^u \, (\boldsymbol{x}^{mp}) \, \boldsymbol{v}_n^{mp}. \tag{3.163}$$

However, as detailed in Section 3.6.5, it is not always possible to numerically invert the mass matrix*, due to its poor condition number, caused by the small-cuts between MPs' domain and grid cells. Stabilisation (3.153) is usually added to the consistent mass matrix (3.118), so that its inversion becomes possible. Nonetheless, for the sake of clarity, this stabilising addition is not considered in this section, even though it is recognised that it would cause some losses proportional to the user-defined value γ_M in Eq. (3.153).

Assessment of linear momentum conservation   From definitions (3.155) and (3.156), it can be seen that the mapping 3.163 conserves linear momentum

$$\boldsymbol{L}_n^{mp} - \mathbf{L}_n^h = \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_n^{mp} - \sum_{A}^{N^{nds}} \left(\sum_{B}^{N^{nds}} \tilde{\mathbf{M}}_{AB} \, (\mathbf{v}_n)_B \right)$$

$$= \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_n^{mp} - \sum_{A}^{N^{nds}} \left(\sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \mathrm{N}_A^u \, (\boldsymbol{x}^{mp}) \, \boldsymbol{v}_n^{mp} \right) = \mathbf{0}. \tag{3.164}$$

Assessment of angular momentum conservation   To compute the difference between grid and MPs' angular momentum, definitions (3.157) and (3.158) are re-called. It can be seen that mapping (3.163) conserves angular momentum too

$$\boldsymbol{J}_n^{mp} - \mathbf{J}_n^h = \sum_{mp}^{N^{mps}} \boldsymbol{x}_n^{mp} \times \bar{m}^{mp} \boldsymbol{v}_n^{mp} - \sum_{A}^{N^{nds}} (\mathbf{x}_n)_A \times \left(\sum_{B}^{N^{nds}} \tilde{\mathbf{M}}_{AB} \, (\mathbf{v}_n)_B \right)$$

$$= \sum_{mp}^{N^{mps}} \boldsymbol{x}_n^{mp} \times \bar{m}^{mp} \boldsymbol{v}_n^{mp} - \sum_{A}^{N^{nds}} (\mathbf{x}_n)_A \times \left(\sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \mathrm{N}_A^u \, (\boldsymbol{x}^{mp}) \, \boldsymbol{v}_n^{mp} \right)$$

$$= \sum_{mp}^{N^{mps}} \boldsymbol{x}_n^{mp} \times \bar{m}^{mp} \boldsymbol{v}_n^{mp} - \sum_{mp}^{N^{mps}} \sum_{A}^{N^{nds}} \mathrm{N}_A^u \, (\boldsymbol{x}^{mp}) \, (\mathbf{x}_n)_A \times (\bar{m}^{mp} \, \boldsymbol{v}_n^{mp}) = \mathbf{0}.$$

$$\tag{3.165}$$

Assessment of kinetic energy conservation   Using definitions (3.159) and (3.160), it can be noted that, regardless of the considered mass matrix, the difference between MPs' and grid kinetic energy at the beginning of the step is given by

$$K_n^{mp} - \mathrm{K}_n^h = \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_n^{mp} \cdot \boldsymbol{v}_n^{mp} - \frac{1}{2} \sum_{A}^{N^{nds}} (\mathbf{v}_n)_A \cdot \left(\sum_{B}^{N^{nds}} \tilde{\mathbf{M}}_{AB} \, (\mathbf{v}_n)_B \right)$$

---

*While the lumped mass matrix is always invertible, the small cut problem causes other troubles. If Eq. (3.163) is considered with a poorly-integrated lumped mass matrix, the grid velocity computed via this equation would be incredibly high at the nodes of the poorly-integrated elements, undermining the whole simulation.

$$= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_n^{mp} \cdot \boldsymbol{v}_n^{mp} - \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_n)_{\mathsf{A}} \cdot \left(\sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \mathrm{N}_A^u \left(\boldsymbol{x}^{mp} \right) \boldsymbol{v}_n^{mp} \right)$$

$$= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \left(\boldsymbol{v}_n^{mp} - \underbrace{\sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_A^u \left(\boldsymbol{x}^{mp} \right) (\mathbf{v}_n)_{\mathsf{A}}}_{:= \tilde{\boldsymbol{v}}_n^{mp}} \right) \cdot \boldsymbol{v}_n^{mp}. \tag{3.166}$$

In the above equation, the grid velocity directly mapped to the MPs has been denoted by $\tilde{\boldsymbol{v}}_n^{mp}$, while the difference between this quantity and the MPs' velocity is non-zero, i.e.,

$$\boldsymbol{v}_n^{mp} := \tilde{\boldsymbol{v}}_n^{mp} + (\boldsymbol{v}')_n^{mp}. \tag{3.167}$$

Substituting $(\boldsymbol{v}')_n^{mp}$ into Eq. (3.166) clearly shows how $(\boldsymbol{v}')_n^{mp}$ contributes to the energy of the system. However, as this value of velocity does not appear in Eq. (3.164), its contribution does not affect the linear momentum balance, i.e.,

$$\sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \mathrm{N}_{\mathsf{A}}^u \, (\boldsymbol{v}')_n^{mp} = \mathbf{0}. \tag{3.168}$$

The sign of this additive velocity $(\boldsymbol{v}')_n^{mp}$ in Eq. (3.167) is debatable, but it has no consequence, since it is usually squared (as it will be shown below) due to its use in the computation of the kinetic energy. In terms of physical meaning, as suggested by Burgess *et al.* [51], the velocity $(\boldsymbol{v}')_n^{mp}$ is a velocity field which can be seen only at the particle level, while, from the grid's perspective, this velocity and the velocity $\boldsymbol{v}_n^{mp}$ are the same. This discrepancy is due to the number of MPs usually being higher than the number of grid nodes, i.e., the grid filters some (high) frequency modes due to its coarseness as compared to the MPs.

Using Eqs. (3.167) and (3.168), the MPs' kinetic energy is given by

$$K_n^{mp} = \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_n^{mp} \cdot \boldsymbol{v}_n^{mp}$$

$$= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \left(\tilde{\boldsymbol{v}}_n^{mp} + (\boldsymbol{v}')_n^{mp} \right) \cdot \left(\tilde{\boldsymbol{v}}_n^{mp} + (\boldsymbol{v}')_n^{mp} \right)$$

$$= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \tilde{\boldsymbol{v}}_n^{mp} \cdot \tilde{\boldsymbol{v}}_n^{mp} + \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, (\boldsymbol{v}')_n^{mp} \cdot (\boldsymbol{v}')_n^{mp} + \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \tilde{\boldsymbol{v}}_n^{mp} \cdot (\boldsymbol{v}')_n^{mp}$$

$$= \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_n)_{\mathsf{A}} \cdot \left(\sum_{\mathsf{B}}^{N^{nds}} \mathbf{M}_{\mathsf{AB}} \, (\mathbf{v}_n)_{\mathsf{B}} \right) + \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, (\boldsymbol{v}')_n^{mp} \cdot (\boldsymbol{v}')_n^{mp}$$

$$= \mathrm{K}_n^h + \frac{\epsilon}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_n)_{\mathsf{A}} \cdot \left(\sum_{\mathsf{B}}^{N^{nds}} \left(\check{\mathbf{M}}_{\mathsf{AB}} - \mathbf{M}_{\mathsf{AB}} \right) (\mathbf{v}_n)_{\mathsf{B}} \right) + \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, (\boldsymbol{v}')_n^{mp} \cdot (\boldsymbol{v}')_n^{mp}, \tag{3.169}$$

where definition (3.159) is employed too. Since it can be proved (see Love and Sulksy [25]) that the matrix $\left(\check{\mathbf{M}}_{\mathsf{AB}} - \mathbf{M}_{\mathsf{AB}} \right)$ is positive semi-definite, it follows that

$$K_n^{mp} - \mathrm{K}_n^h \geq 0. \tag{3.170}$$

Another way to re-write the difference shown in Eq. (3.166) is as follows

$$K_n^{mp} - \mathrm{K}_n^h = \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_n^{mp} \cdot \boldsymbol{v}_n^{mp} - \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_n)_{\mathsf{A}} \cdot \left(\sum_{\mathsf{B}}^{N^{nds}} \mathbf{M}_{\mathsf{AB}} \, (\mathbf{v}_n)_{\mathsf{B}} \right)$$

$$- \frac{\epsilon}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_n)_{\mathsf{A}} \cdot \left(\sum_{\mathsf{B}}^{N^{nds}} \left(\check{\mathbf{M}}_{\mathsf{AB}} - \mathbf{M}_{\mathsf{AB}} \right) (\mathbf{v}_n)_{\mathsf{B}} \right). \qquad (3.171)$$

If Eqs. (3.169) and (3.171) are considered, it can be seen that the difference between MPs' and grid kinetic energy decreases if the consistent mass matrix is used (i.e., setting $\epsilon = 0$), even though it cannot be entirely cancelled out due to the energy contribution of the velocities $(\boldsymbol{v}')_n^{mp}$.

A summary of the differences between grid and MPs' quantities at the beginning of the step due to mapping (3.163) is given in Table 3.1.

Table 3.1: Summary of the differences between grid and MPs' quantities of interest at the beginning of the step using the mapping defined by Eq. (3.163). Table reproduced from Pretti *et al.* [26].

Quantities	Difference	
Linear momentum	$\boldsymbol{L}_n^{mp} - \mathbf{L}_n^h$	$= \mathbf{0}$
Angular momentum	$\boldsymbol{J}_n^{mp} - \mathbf{J}_n^h$	$= \mathbf{0}$
Kinetic energy	$K_n^{mp} - \mathrm{K}_n^h$	≥ 0

3.6.6.2 Grid-to-point (G2P) mapping

While in the literature the initial mappings available can be summarised by Eq. (3.163), the G2P mappings have received much more attention, as evidenced by the profusion of techniques on the topic. To the author's best knowledge, these are PIC [52–55], FLIP [51, 56, 57], Affine Particle in Cell (APIC) [58], eXtended Particle In Cell (XPIC) [59], and Polynomial Particle In Cell (PolyPIC) [60]. This work evaluates only PIC and FLIP, clarifying why the former has been regarded as dissipative while the latter has not.

The PIC and FLIP mappings are defined by updating the MPs' velocity at the end of the step in the following manners

$$\text{PIC:} \qquad \boldsymbol{v}_{n+1}^{mp} = \sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp} \right) (\mathbf{v}_{n+1})_{\mathsf{A}} \, ; \qquad (3.172)$$

$$\text{FLIP:} \qquad \boldsymbol{v}_{n+1}^{mp} = \boldsymbol{v}_n^{mp} + \sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp} \right) (\mathbf{v}_{n+1} - \mathbf{v}_n)_{\mathsf{A}} \, . \qquad (3.173)$$

From the above formulae, it can be seen how the velocity at the end of the step is passed from the grid to the MPs for the PIC procedure, while the difference in the velocity (i.e., an acceleration over the given time step) is mapped in the case of FLIP.

The update of the MP current position at the end of the step is unique and is given by

$$\boldsymbol{x}_{n+1}^{mp} = \boldsymbol{x}_n^{mp} + \sum_A^{N^{nds}} \mathrm{N}_A^u(\boldsymbol{x}^{mp}) (\mathbf{u}_{n+1})_A = \sum_A^{N^{nds}} \mathrm{N}_A^u(\boldsymbol{x}^{mp}) (\mathbf{x}_{n+1})_A. \qquad (3.174)$$

Assessment of linear momentum conservation   Owing to Eq. (3.164) (which ensures that no linear momentum is lost in the P2G mapping), checking the time-difference of linear momentum at the end of the steps corresponds to evaluating the MPs' and grid difference only at the end of the step.

This difference is first evaluated for PIC (3.172)

$$\begin{aligned}
\Delta \boldsymbol{L}^{mp} - \Delta \mathbf{L}^h &= \boldsymbol{L}_{n+1}^{mp} - \mathbf{L}_{n+1}^h \\
&= \sum_{mp}^{N^{mps}} \bar{m}^{mp} \boldsymbol{v}_{n+1}^{mp} - \sum_A^{N^{nds}} \left(\sum_B^{N^{nds}} \tilde{\mathbf{M}}_{AB} (\mathbf{v}_{n+1})_B \right) \\
&= \sum_{mp}^{N^{mps}} \bar{m}^{mp} \left(\sum_A^{N^{nds}} \mathrm{N}_A^u(\boldsymbol{x}^{mp}) (\mathbf{v}_{n+1})_A \right) \\
&\quad - (1-\epsilon) \sum_{mp}^{N^{mps}} \bar{m}^{mp} \underbrace{\sum_A^{N^{nds}} \mathrm{N}_A^u(\boldsymbol{x}^{mp})}_{=1} \sum_B^{N^{nds}} \mathrm{N}_B^u(\boldsymbol{x}^{mp}) (\mathbf{v}_{n+1})_B \\
&\quad - \epsilon \sum_{mp}^{N^{mps}} \bar{m}^{mp} \sum_B^{N^{nds}} \mathrm{N}_B^u(\boldsymbol{x}^{mp}) (\mathbf{v}_{n+1})_B = \mathbf{0}, \qquad (3.175)
\end{aligned}$$

which produces a linear momentum conserving mapping.

The same procedure can be carried on for FLIP (3.173)

$$\begin{aligned}
\Delta \boldsymbol{L}^{mp} - \Delta \mathbf{L}^h &= \boldsymbol{L}_{n+1}^{mp} - \mathbf{L}_{n+1}^h \\
&= \sum_{mp}^{N^{mps}} \bar{m}^{mp} \boldsymbol{v}_{n+1}^{mp} - \sum_A^{N^{nds}} \left(\sum_B^{N^{nds}} \tilde{\mathbf{M}}_{AB} (\mathbf{v}_{n+1})_B \right) \\
&= \sum_{mp}^{N^{mps}} \bar{m}^{mp} \left(\boldsymbol{v}_n^{mp} + \sum_A^{N^{nds}} \mathrm{N}_A^u(\boldsymbol{x}^{mp}) (\Delta \mathbf{v})_A \right) - \sum_A^{N^{nds}} \left(\sum_B^{N^{nds}} \tilde{\mathbf{M}}_{AB} (\mathbf{v}_{n+1})_B \right) \\
&= \sum_{mp}^{N^{mps}} \bar{m}^{mp} \sum_A^{N^{nds}} \mathrm{N}_A^u(\boldsymbol{x}^{mp}) (\mathbf{v}_{n+1})_A - \sum_A^{N^{nds}} \left(\sum_B^{N^{nds}} \tilde{\mathbf{M}}_{AB} (\mathbf{v}_{n+1})_B \right) = \mathbf{0},
\end{aligned}$$
$$(3.176)$$

resulting in a linear momentum conserving algorithm for FLIP too.

Assessment of angular momentum conservation   As for the linear momentum case, assessing the time-difference in the angular momentum is the same as computing the MPs' and grid difference at the end of the step. This was proved by Eq. (3.165).

In the case of PIC (3.172), this difference gives

$$
\Delta \boldsymbol{J}^{mp} - \Delta \mathbf{J}^h = \boldsymbol{J}_{n+1}^{mp} - \mathbf{J}_{n+1}^h
$$

$$
= \sum_{mp}^{N^{mps}} \boldsymbol{x}_{n+1}^{mp} \times \bar{m}^{mp} \boldsymbol{v}_{n+1}^{mp} - \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} (\mathbf{v}_{n+1})_{\mathsf{B}} \right)
$$

$$
= \sum_{mp}^{N^{mps}} \left(\sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_{\mathsf{A}}^u (\boldsymbol{x}^{mp}) (\mathbf{x}_{n+1})_{\mathsf{A}} \right) \times \left(\bar{m}^{mp} \sum_{\mathsf{B}}^{N^{nds}} \mathrm{N}_{\mathsf{B}}^u (\boldsymbol{x}^{mp}) (\mathbf{v}_{n+1})_{\mathsf{B}} \right)
$$

$$
- \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} (\mathbf{v}_{n+1})_{\mathsf{B}} \right)
$$

$$
= \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \sum_{\mathsf{B}}^{N^{nds}} (\mathbf{M}_{\mathsf{AB}} - \tilde{\mathbf{M}}_{\mathsf{AB}}) (\mathbf{v}_{n+1})_{\mathsf{B}}
$$

$$
= -\epsilon \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \sum_{\mathsf{B}}^{N^{nds}} \left(\check{\mathbf{M}}_{\mathsf{AB}} - \mathbf{M}_{\mathsf{AB}} \right) (\mathbf{v}_{n+1})_{\mathsf{B}} . \tag{3.177}
$$

The above equation describes two scenarios. If the consistent mass matrix is used (setting $\epsilon = 0$), PIC mapping conserves angular momentum. If, on the other hand, an effective mass matrix is taken into account, this difference is non-zero and it is proportional to the grid velocity at the end of the step $\mathbf{v}_{n+1}$. If the difference of the Euclidean norms is considered $(\|\Delta \boldsymbol{J}^{mp}\| - \|\Delta \mathbf{J}^h\|)$, it can also be seen that there is no way to predict if it would result in values bigger or lower than zero. This implies that the difference in magnitude of MPs' and grid angular momenta can be either decreasing or increasing.

Let us proceed to compute the difference of angular momenta in the case of FLIP mapping (3.173)

$$
\Delta \boldsymbol{J}^{mp} - \Delta \mathbf{J}^h = \boldsymbol{J}_{n+1}^{mp} - \mathbf{J}_{n+1}^h
$$

$$
= \sum_{mp}^{N^{mps}} \boldsymbol{x}_{n+1}^{mp} \times \bar{m}^{mp} \boldsymbol{v}_{n+1}^{mp} - \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} (\mathbf{v}_{n+1})_{\mathsf{B}} \right)
$$

$$
= \sum_{mp}^{N^{mps}} \left(\sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_{\mathsf{A}}^u (\boldsymbol{x}^{mp}) (\mathbf{x}_{n+1})_{\mathsf{A}} \right) \times \bar{m}^{mp} \left(\boldsymbol{v}_n^{mp} + \sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_{\mathsf{A}}^u (\boldsymbol{x}^{mp}) (\mathbf{v}_{n+1} - \mathbf{v}_n)_{\mathsf{A}} \right)
$$

$$
- \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} (\mathbf{v}_{n+1})_{\mathsf{B}} \right)
$$

$$
= \sum_{\mathsf{A},\mathsf{B}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \tilde{\mathbf{M}}_{\mathsf{AB}} (\mathbf{v}_n)_{\mathsf{B}} + \sum_{\mathsf{A},\mathsf{B}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \mathbf{M}_{\mathsf{AB}} (\Delta \mathbf{v})_{\mathsf{B}}
$$

$$
- \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \left(\sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}} (\mathbf{v}_{n+1})_{\mathsf{B}} \right)
$$

$$
= -\epsilon \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{x}_{n+1})_{\mathsf{A}} \times \left(\sum_{\mathsf{B}}^{N^{nds}} \left(\check{\mathbf{M}}_{\mathsf{AB}} - \mathbf{M}_{\mathsf{AB}} \right) (\Delta \mathbf{v})_{\mathsf{B}} \right) . \tag{3.178}
$$

As for PIC, FLIP mapping conserves angular momentum if the consistent mass matrix is considered. In the opposite case, (i.e., taking an effective mass matrix into account), the error generated by FLIP is proportional to the time-step difference of grid velocity. Comparing this error with that computed for PIC (Eq. (3.177)), it can be appreciated that the magnitude of $\Delta \mathbf{v}$ should be, in the majority of cases, much smaller than the value of the velocity at the end of the step $\mathbf{v}_{n+1}$. As a general rule, in the effective mass matrix case, FLIP preserves more angular momentum than PIC. The discussion made in the case of PIC regarding the norms of error in angular momentum also applies to the case of FLIP for the same reasons.

Assessment of kinetic energy conservation   The evaluation of the time difference in kinetic energy between MPs' and grid values cannot take advantage of the same considerations made for linear and angular momenta, as is clear from the error in the initial mapping Eq. (3.171). Let us proceed then with the evaluation of the kinetic energy difference in the case of PIC mapping (3.172)

$$
\begin{aligned}
K_{n+1}^{mp} - \mathrm{K}_{n+1}^{h} &= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp}\, \boldsymbol{v}_{n+1}^{mp} \cdot \boldsymbol{v}_{n+1}^{mp} - \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_{n+1})_{\mathsf{A}} \cdot \sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}}\, (\mathbf{v}_{n+1})_{\mathsf{B}} \\
&= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \left(\sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_{\mathsf{A}}^{u}\, (\boldsymbol{x}^{mp})\, (\mathbf{v}_{n+1})_{\mathsf{A}} \right) \cdot \left(\sum_{\mathsf{B}}^{N^{nds}} \mathrm{N}_{\mathsf{B}}^{u}\, (\boldsymbol{x}^{mp})\, (\mathbf{v}_{n+1})_{\mathsf{B}} \right) \\
&\quad - \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_{n+1})_{\mathsf{A}} \cdot \sum_{\mathsf{B}}^{N^{nds}} \tilde{\mathbf{M}}_{\mathsf{AB}}\, (\mathbf{v}_{n+1})_{\mathsf{B}} \\
&= \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_{n+1})_{\mathsf{A}} \cdot \sum_{\mathsf{B}}^{N^{nds}} \left(\mathbf{M}_{\mathsf{AB}} - \tilde{\mathbf{M}}_{\mathsf{AB}} \right) (\mathbf{v}_{n+1})_{\mathsf{B}} \\
&= -\frac{\epsilon}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_{n+1})_{\mathsf{A}} \cdot \sum_{\mathsf{B}}^{N^{nds}} \left(\check{\mathbf{M}}_{\mathsf{AB}} - \mathbf{M}_{\mathsf{AB}} \right) (\mathbf{v}_{n+1})_{\mathsf{B}} \leq 0, \quad\quad (3.179)
\end{aligned}
$$

which results in a null difference at the end of the step if the consistent mass matrix is used. Therefore, the above equation clearly states that PIC *per se* is not dissipative in the case of a consistent mass matrix. However, the overall error is not only that given by the difference at the end of the step, as the error in the initial mapping must be taken into account too. Eqs. (3.179) and (3.171) are useful to compute this overall time-difference in MPs and grid kinetic energy

$$
\begin{aligned}
\Delta K^{mp} - \Delta \mathrm{K}^{h} &= \underbrace{K_{n+1}^{mp} - \mathrm{K}_{n+1}^{h}}_{\leq 0}\, \underbrace{- \left(K_{n}^{mp} - \mathrm{K}_{n}^{h} \right)}_{\leq 0} \\
&= -\frac{\epsilon}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_{n+1})_{\mathsf{A}} \cdot \sum_{\mathsf{B}}^{N^{nds}} \left(\check{\mathbf{M}}_{\mathsf{AB}} - \mathbf{M}_{\mathsf{AB}} \right) (\mathbf{v}_{n+1})_{\mathsf{B}} - \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp}\, \boldsymbol{v}_{n}^{mp} \cdot \boldsymbol{v}_{n}^{mp} \\
&\quad + \frac{1}{2} \sum_{\mathsf{A}}^{N^{nds}} (\mathbf{v}_{n})_{\mathsf{A}} \cdot \left(\sum_{\mathsf{B}}^{N^{nds}} \mathbf{M}_{\mathsf{AB}}\, (\mathbf{v}_{n})_{\mathsf{B}} \right)
\end{aligned}
$$

$$+ \frac{\epsilon}{2} \sum_A^{N^{nds}} (\mathbf{v}_n)_A \cdot \left(\sum_B^{N^{nds}} \left(\check{\mathbf{M}}_{AB} - \mathbf{M}_{AB} \right) (\mathbf{v}_n)_B \right)$$

$$= -\frac{\epsilon}{2} \sum_A^{N^{nds}} (\Delta\mathbf{v})_A \cdot \sum_B^{N^{nds}} \left(\check{\mathbf{M}}_{AB} - \mathbf{M}_{AB} \right) (\mathbf{v}_{n+1} + \mathbf{v}_n)_B - K_n^{mp}$$

$$+ \frac{1}{2} \sum_A^{N^{nds}} (\mathbf{v}_n)_A \cdot \left(\sum_B^{N^{nds}} \mathbf{M}_{AB} \, (\mathbf{v}_n)_B \right)$$

$$= -\frac{\epsilon}{2} \sum_A^{N^{nds}} (\Delta\mathbf{v})_A \cdot \sum_B^{N^{nds}} \left(\check{\mathbf{M}}_{AB} - \mathbf{M}_{AB} \right) (\mathbf{v}_{n+1} + \mathbf{v}_n)_B - \left(K_n^{mp} - K_n^h \right).$$

(3.180)

Several conclusions can be drawn from the above chain of formulae. First of all, it can be appreciated how the kinetic energy in the time-step should always decrease in the case of PIC mapping. This mechanism is exacerbated if the effective mass matrix is considered, i.e., $\epsilon \neq 0$. Second, the error made in the initial mapping and expressed by Eq. (3.171) is not cancelled out in this overall error. Hence, Eqs. (3.179) and (3.180) allow us to state that PIC is not dissipative if used alone, but the combined use of initial mapping (3.163) and PIC (this being defined by (3.172)) dissipates kinetic energy.

Let us now proceed in the same calculations for the case of FLIP mapping (3.173). If the difference at the end of the step is considered, this is given by

$$K_{n+1}^{mp} - K_{n+1}^h = \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_{n+1}^{mp} \cdot \boldsymbol{v}_{n+1}^{mp} - \frac{1}{2} \sum_A^{N^{nds}} (\mathbf{v}_{n+1})_A \cdot \sum_B^{N^{nds}} \tilde{\mathbf{M}}_{AB} (\mathbf{v}_{n+1})_B$$

$$= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \left(\boldsymbol{v}_n^{mp} + \sum_A^{N^{nds}} \mathrm{N}_A^u (\boldsymbol{x}^{mp}) (\Delta\mathbf{v})_A \right) \cdot \left(\boldsymbol{v}_n^{mp} + \sum_B^{N^{nds}} \mathrm{N}_B^u (\boldsymbol{x}^{mp}) (\Delta\mathbf{v}_{n+1})_B \right)$$

$$- \frac{1}{2} \sum_A^{N^{nds}} (\mathbf{v}_{n+1})_A \cdot \sum_B^{N^{nds}} \tilde{\mathbf{M}}_{AB} (\mathbf{v}_{n+1})_B$$

$$= \frac{1}{2} \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \boldsymbol{v}_n^{mp} \cdot \boldsymbol{v}_n^{mp} + \frac{1}{2} \sum_{A,B}^{N^{nds}} \Delta\mathbf{v}_A \mathbf{M}_{AB} \Delta\mathbf{v}_B + \frac{1}{2} \sum_{A,B}^{N^{nds}} \Delta\mathbf{v}_A \tilde{\mathbf{M}}_{AB} (\mathbf{v}_n)_B$$

$$- \frac{1}{2} \sum_A^{N^{nds}} (\mathbf{v}_{n+1})_A \cdot \sum_B^{N^{nds}} \tilde{\mathbf{M}}_{AB} (\mathbf{v}_{n+1})_B$$

$$= K_n^{mp} - K_n^h + \frac{1}{2} \sum_{A,B}^{N^{nds}} \Delta\mathbf{v}_A \underbrace{\left(\mathbf{M}_{AB} - \tilde{\mathbf{M}}_{AB} \right)}_{= -\epsilon \left(\check{\mathbf{M}}_{AB} - \mathbf{M}_{AB} \right)} \Delta\mathbf{v}_B. \qquad (3.181)$$

However, as for the PIC mapping, the difference of interest is given by the accumulation of this error with that coming from the initial mapping. Considering Eqs. (3.171) and (3.181) results in an overall error for FLIP expressed by the following formula

$$\Delta K^{mp} - \Delta \mathrm{K}^h = K_{n+1}^{mp} - \mathrm{K}_{n+1}^h - \left(K_n^{mp} - \mathrm{K}_n^h \right)$$

$$= K_n^{mp} - K_n^h - \frac{\epsilon}{2} \sum_{A,B}^{N^{nds}} \Delta\mathbf{v}_A \left(\check{\mathbf{M}}_{AB} - \mathbf{M}_{AB} \right) \Delta\mathbf{v}_B - \left(K_n^{mp} - K_n^h \right)$$

$$= -\frac{\epsilon}{2} \sum_{A,B}^{N^{nds}} \Delta\mathbf{v}_A \left(\check{\mathbf{M}}_{AB} - \mathbf{M}_{AB} \right) \Delta\mathbf{v}_B \leq 0. \tag{3.182}$$

The equation above yields conclusions for FLIP that are diametrically opposed to those drawn in the PIC case: FLIP is dissipative *per se* as a standalone technique (Eq. (3.173)), but FLIP cancels out the error inherited from the initial mapping (3.163).

These findings for PIC and FLIP mappings differ from those drawn in some of the literature (see, for instance, [59], which focuses exclusively on the G2P mapping): even when the consistent mass matrix is employed, the energy bottleneck is represented by the initial mapping, quantified by Eq. (3.166). However, the error in energy committed by the FLIP mapping (quantified by Eq. (3.181)) compensates for the initial one, given zero net error (see Eq. (3.182)). This cancellation does not happen with PIC, as the error energy for this technique is null (see Eq. (3.179)), and the initial error accumulates (given by Eq. (3.180)). These conclusions and those relative to the conservation of linear and angular momentum over the step are summarised in Table 3.2.

Table 3.2: Résumé of time-step differences between and grid and MPs' quantities of interest. Table reproduced from Pretti *et al.* [26].

Differences	Method	Results
$\Delta\boldsymbol{L}^{mp} - \Delta\mathbf{L}^h$	Eq. (3.163) + PIC:	$= \mathbf{0}$
	Eq. (3.163) + FLIP:	$= \mathbf{0}$
$\Delta\boldsymbol{J}^{mp} - \Delta\mathbf{J}^h$	Eq. (3.163) + PIC:	$= -\epsilon \sum_A^{N^{nds}} (\mathbf{x}_{n+1})_A \times \sum_B^{N^{nds}} \left(\check{\mathbf{M}} - \mathbf{M} \right)_{AB} (\mathbf{v}_{n+1})_B$
	Eq. (3.163) + FLIP:	$= -\epsilon \sum_A^{N^{nds}} (\mathbf{x}_{n+1})_A \times \sum_B^{N^{nds}} \left(\check{\mathbf{M}} - \mathbf{M} \right)_{AB} \Delta\mathbf{v}_B$
$\Delta K^{mp} - \Delta\mathrm{K}^h$	Eq. (3.163) + FLIP:	$= -\frac{\epsilon}{2} \sum_A^{N^{nds}} \Delta\mathbf{v}_A \cdot \sum_B^{N^{nds}} \left(\check{\mathbf{M}} - \mathbf{M} \right)_{AB} \Delta\mathbf{v}_B$
	Eq. (3.163) + PIC:	$= -\frac{\epsilon}{2} \sum_A^{N^{nds}} \Delta\mathbf{v}_A \cdot \sum_B^{N^{nds}} \left(\check{\mathbf{M}} - \mathbf{M} \right)_{AB} (\mathbf{v}_{n+1} + \mathbf{v}_n)_B$
		$\quad - K_n^{mp} + \frac{1}{2} \sum_A^{N^{nds}} (\mathbf{v}_n)_A \cdot \mathbf{M}_{AB} (\mathbf{v}_n)_B.$

3.6.7 Implementation details

Since the equations considered so far in the chapter are numerous, this section specifies which play a greater part implementation-wise for developing the CLC UL implicit dynamic code used for (most of) the examples in Section 3.7. This code stems from the base version of AMPLE (see Coombs and Augarde [61]), expanded to model dynamic conditions. Figure 3.6 illustrates how these considered equations fit into the MPM algorithmic overview given in Chapter 2 and graphically shown in Figure 2.1. From a comparison of Figure 2.1 with Figure 3.6, it can be seen that only substeps from (B) to (E) are considered in the latter
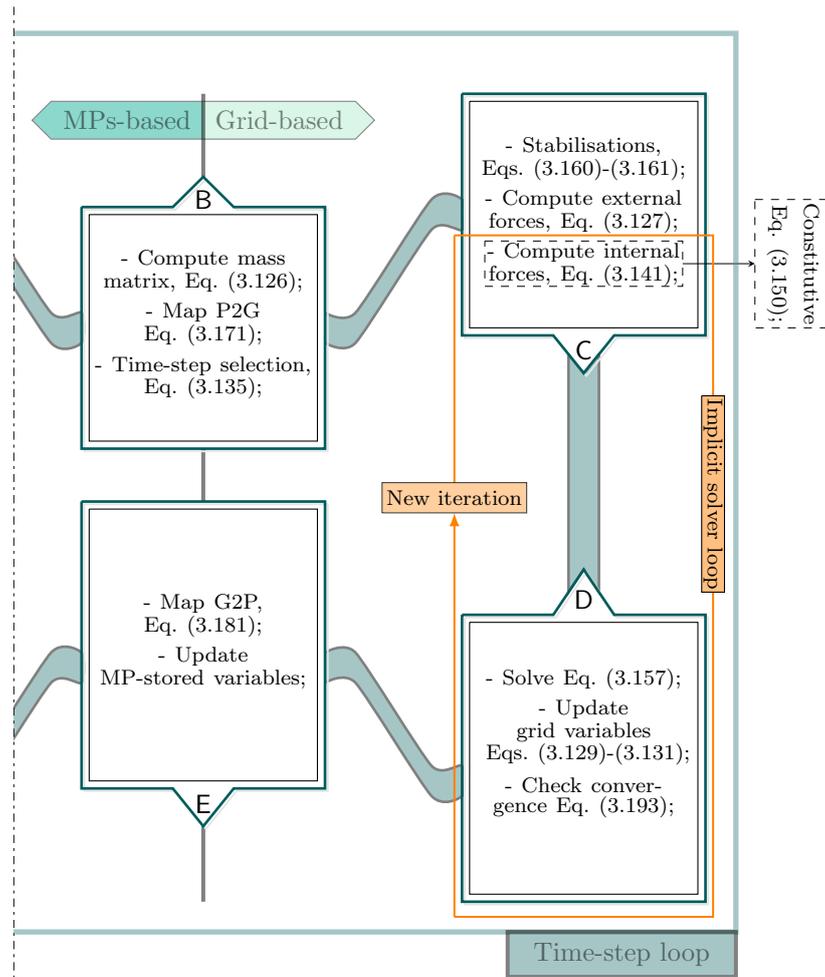
Figure 3.6: Main equations and relative use in the structure detailing the MPM algorithmic substeps. The considered equations are employed to design the UL CLC implicit dynamic code and necessary boundary stabilisations.

(and, consequently, in this section too). The initial setup, the grid introduction in substep (A) and its disposal in (F), and the post-processing phase are not addressed below since they are not specific to a code modelling solid dynamics.

The designed code constructs the consistent mass matrix, Eq. (3.118), as a first task in substep (B). Conversely to the FEM, the MPM requires the assembly of this matrix at every time-step, since the MPs move their position with respect to the newly introduced grid, and they can also cross elements, thus changing the active nodes. Necessary for the construction of this matrix is the presence of the shape functions, which can be computed only once per time-step since there is no relative displacement between the MPs and the grid in a single time-step*. The P2G mapping via Eq. (3.163) is performed via the mass matrix. For those

---

*The derivatives of the shape functions, conversely, must be updated at every iteration. This applies to the derivatives of these functions both with respect to the initial (appearing in the TL formulation, Eq. (3.117)) and current positions (necessary for the UL formulation, Eq. (3.116)). The main reason is that the continuous introduction and disposal of the grid allows only the computation of the derivatives of the shape functions with respect to the previously converged step (see Charlton *et al.* [30] or Coombs *et al.* [62] for more details).

cases where the small-cut issue could give rise to problems in the inversion of Eq. (3.163), the stabilisation for the mass matrix (3.153) should be assembled too. In this case, the performed mapping is provided by the following equation (summation explicitly written)

$$\sum_{\mathsf{B}}^{N^{nds}} \left(\mathbf{M}_{AB} + (\mathbf{M}_G)_{\mathsf{AB}}\right) (\mathbf{v}_n)_{\mathsf{B}} = \sum_{mp}^{N^{mps}} \bar{m}^{mp} \, \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp}\right) \boldsymbol{v}_n^{mp}, \tag{3.183}$$

in place of Eq. (3.163). Since the penalty values γ^M in front of the ghost penalty (3.153) are minimal, it is to be expected that the stabilisation affects energy conservation to a very small degree. It is customary to compute the stabilisation of the mass matrix for the P2G mapping at every time-step or to use it only when required (i.e., when the mass matrix is significantly ill-conditioned). This work considers the latter option to have even less impact on the conservation of kinetic energy. Selecting the time step using Eq. (3.127) shifts the focus to what is happening on the grid, i.e., substep (C).

Outside of the NR loop, it is possible to build the stabilisation matrices Eqs. (3.152) and (3.153) (if not already assembled). The same applies to the part of the external force vector considering body forces, i.e., the first term in Eq. (3.119). The NR iterative scheme is then started, using the following predictors as initial guess

$$\mathbf{u}_{\tilde{n}} = \mathbf{v}_{\tilde{n}} = \mathbf{a}_{\tilde{n}} = \mathbf{0}, \tag{3.184}$$

which are in compliance with the P2G mapping (i.e., $\mathbf{v}_n \neq \mathbf{0}$ and $\mathbf{u}_n = \mathbf{a}_n = \mathbf{0}$) and the mid-point time-discretisation Eqs. (3.121)-(3.123). While Eq. (3.184) is not a unique choice, the NR procedure can find the zeros of the considered function provided that the above predictors are close enough to the solution (see, for instance, Quarteroni *et al.* [63]). As standard for MPM implicit schemes (see, for a comparison, Coombs and Augarde [61]), the first iteration is employed to assemble the Jacobian matrix, Eq. (3.150). The inversion of the Jacobian matrix in Eq. (3.149) provides an updated displacement nodal solution (substep (D)), which, in turn, updates the velocity and acceleration variables via Eqs. (3.121)-(3.123). The nodal solution is thus used to update the terms in the residual Eq. (3.154), where particular attention must be paid to the considered internal force vector, Eq. (3.133). The computation of this vector is based on the considered stress-strain relationship, which, in the case of the CLC formulation, is given by Eq. (3.142). The convergence of the NR procedure is checked based on the residual and the iterative solution, Eq. (3.185). The choice of this equation as a convergence criterion is motivated in Section 3.7. If this condition is not met, another Jacobian is computed, and the loop is repeated between substeps (C) and (D).

When the convergence criterion is satisfied, the G2P mapping in substep (E) must be performed before the grid disposal. This work adopts the FLIP mapping (3.173) for the same reasons which justify the use of the consistent mass matrix and has been motivated by the conclusions drawn in Section 3.6.6. Other MP-stored variables, such as deformation gradient, stress, and other variables necessary for post-processing, are also updated.

3.7 Numerical simulations

This section presents six numerical examples, run (mostly) with the CLC method outlined in Section 3.6.3 and the FLIP mapping at the end of the step. When not explicitly stated otherwise, the stress-strain relationship used is that given by Eq. (3.142), where the Helmholtz free energy function describes a Hencky material (Eq. (3.87)). Example 3.7.1 compares the CLC method with other kinds of time discretisations. Example 3.7.2 tests the (small-strain) elasto-plastic implementation of Eq. (3.142) via the Method of Manufactured Solutions (MMS). Example 3.7.3 compares the CLC MPM algorithm for compressible and nearly-incompressible Hencky material. Example 3.7.4 and simulation (A) of Example 3.7.5 share the same set-up, with the difference that the former considers elasticity, while elasto-plastic strains are analysed in the second. Plasticity play a role for simulations (A) and (B) of Example 3.7.5 and Example 3.7.6. Examples 3.7.3-3.7.5 are plane-strain simulations, while 3.7.6 is fully three-dimensional.

Each time-step in the analyses is considered converged when the tolerance $tol = 10^{-11}$ at the $k-$th current iteration satisfies the inequality

$$\frac{|\mathbf{r}^{(k)}| \cdot |\delta\mathbf{u}^{(k)}|}{|\mathbf{r}^{(1)}| \cdot |\delta\mathbf{u}^{(1)}|} < tol, \tag{3.185}$$

with $|(\bullet)|$ being the absolute value of the vector $(\bullet)$. The above error measure can be seen as a good compromise to track at the same time the reduction in the nodal residual vector $\mathbf{r}^{(k)}$ and the progressive decrease of the nodal incremental displacements $\delta\mathbf{u}^{(k)}$.

3.7.1 1D elastic column under self-weight

(a) Illustration of the elastic column under self-weight.

(b) Applied load time variation.
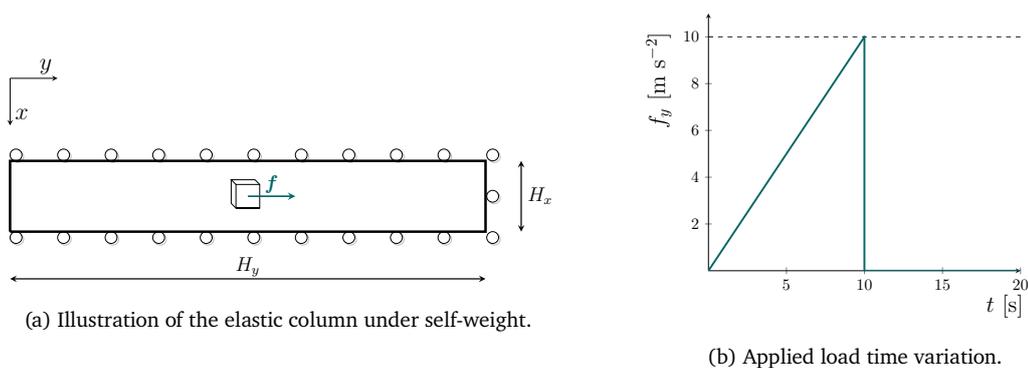
Figure 3.7: The 1D elastic column under self-weight.

Example scope The goal of this example (illustrated in Figure 3.7(a)) is two-fold: on the one hand, it compares the CLC formulation outlined in Section 3.6.3 with other time-schemes; on the other, it is shown that the CLC shows good agreement with the analytical stress distribution in very specific circumstances (detailed below).

As for the first goal, the generalised-α method [64] the Newmark scheme [65], the trapezoidal rule and the mid-point rule (without the internal force modification) are compared with the

Table 3.3: Résumé of the necessary equations for the generalised-α and CLC methods.

	generalised-α method (Family)	CLC method
Equilibrium Equations (Unstabilised)	$\mathbf{M}\mathbf{a}_{n+\alpha_m} + \mathbf{f}^{int}_{n+\alpha_f} - \mathbf{f}^{ext}_{n+\alpha_f} = \mathbf{0};$ $\mathbf{f}^{int}_{n+\alpha_f} = \mathbf{f}^{int}\left(\mathbf{u}_{n+\alpha_f}\right);$ $\mathbf{f}^{ext}_{n+\alpha_f} = \mathbf{f}^{ext}_n + \alpha_f\left(\mathbf{f}^{ext}_{n+1} - \mathbf{f}^{ext}_n\right);$	$\mathbf{M}\mathbf{a}_{\tilde{n}} + \tilde{\mathbf{f}}^{int}_{\tilde{n}} - \mathbf{f}^{ext}_{\tilde{n}} = \mathbf{0};$ $\tilde{\mathbf{f}}^{int}_{\tilde{n}} = \tilde{\mathbf{f}}^{int}(\boldsymbol{\tau}^{mp});$ $\mathbf{f}^{ext}_{\tilde{n}} = \frac{1}{2}\left(\mathbf{f}^{ext}_{n+1} + \mathbf{f}^{ext}_n\right);$
Kinematic Variables	$\mathbf{v}_{n+1} = \frac{\tilde{\gamma}}{\tilde{\beta}\,\Delta t}\mathbf{u}_{n+1} + \mathbf{v}_n\left(1 - \frac{\tilde{\gamma}}{\tilde{\beta}}\right);$ $\mathbf{a}_{n+1} = \frac{1}{\tilde{\beta}\,\Delta t^2}\left(\mathbf{u}_{n+1} - \Delta t\,\mathbf{v}_n\right);$ $\mathbf{u}_{n+\alpha_f} = \alpha_f\,\mathbf{u}_{n+1};$ $\mathbf{a}_{n+\alpha_m} = \alpha_m\,\mathbf{a}_{n+1};$	$\mathbf{v}_{n+1} = \frac{2}{\Delta t}\mathbf{u}_{n+1} - \mathbf{v}_n;$ $\mathbf{a}_{n+1} = \frac{4}{\Delta t^2}\left(\mathbf{u}_{n+1} - \Delta t\,\mathbf{v}_n\right);$
Constitutive Relationship	Classical Hencky elastic material $\boldsymbol{\tau}$ define by Eq. (3.85a);	Modification of the constituive relationship $\boldsymbol{\tau}^{mp}$ defined by Eq. (3.142).

CLC formulation. Since the generalised-α method is more general, its balance of rate of linear momentum, the relationship among time-discretised kinematic variables and the used (elastic) constitutive relationships are listed in in Table 3.3. The same table compares the above-mentioned relationships to the equivalents used for the CLC formulation. The relationships among the coefficients for the different time schemes and the chosen values for this particular simulation are reported in Table 3.4.

While a stress analytical solution to this problem exists for the quasi-static case (see Charlton *et al.* [30] for its derivation), this can be extended to specific situations to the dynamic case when the acceleration is close to zero, i.e.,

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \rho\,b_i = \rho\,\dot{v}_i \approx \mathbf{0}. \tag{3.186}$$

On top of this, the maximum value of the strain energy obtained from the dynamic case should match the one for the quasi-static simulation, since the externally supplied power (via the application of body forces) is the same for both simulations.

Setup A linear Hencky material constitutes the column, with an Young's modulus of $\bar{E} = 1$ MPa and a Poisson's ratio of $\nu = 0$. Similarly, the initial density of the material is $\rho_0 = 80$ kg m^{-3}. No boundary stabilisations (as the one suggested in Section 3.6.5) have been considered (in this way the CLC method is expected not to dissipate), and the GIMPM functions are adopted. The dimensions of the column are $H_x = 0.5$ m, $H_y = 50$ m, and it is discretised by square elements of side length 0.5 m. Each element has initially two MPs per direction, obtained by dividing the length of the cell into equal parts and positioning each MP in the centre of these split parts. As the problem is *de facto* mono-dimensional, this number of MPs per cell can be regarded sufficient. The chosen time increment is $\Delta t = 0.05$

Table 3.4: Coefficient values for the different methods belonging to the generalised-α scheme.

	Generalised-α Method	Newmark Method	Trapezoidal Rule	Mid-point Rule
Relationships among Coefficients	$\alpha_m = \frac{2-\rho_\infty}{1+\rho_\infty};$ $\alpha_f = \frac{1}{1+\rho_\infty};$ $\tilde{\beta} = \frac{1}{4}\left(1+\alpha_m-\alpha_f\right)^2;$ $\tilde{\gamma} = \frac{1}{2} + \alpha_m - \alpha_f;$	$\alpha_m = 1;$ $\alpha_f = 1;$ $\tilde{\beta}, \tilde{\gamma}$ free;	$\alpha_m = 1;$ $\alpha_f = 1;$ $\tilde{\beta} = \frac{1}{4};$ $\tilde{\gamma} = \frac{1}{2};$	$\alpha_m = \frac{1}{2};$ $\alpha_f = \frac{1}{2};$ $\tilde{\beta} = \frac{1}{4};$ $\tilde{\gamma} = \frac{1}{2};$
Choices for the Analysis	‡$\rho_\infty = 0.818;$	†$2\tilde{\beta} = \tilde{\gamma} = \frac{1}{3};$	—	—

‡ Values taken from Kontoe *et al.* [66].
† The idea of considering a linear acceleration $\tilde{\beta} = \frac{1}{6}$ is taken from Nøst [67], whereas the relationship $2\tilde{\beta} = \tilde{\gamma}$ comes from the need for algorithmic stability (see, for instance, Géradin and Rixen [68]).

s so that the entire simulation consists of 400 time-steps. The parameters considered above and the BCs follow those provided by Charlton *et al.* [30] for the quasi-static case.

The mono-dimensional column is subjected to an increase in self-weight (illustrated in Figure 3.7(b)), which is linearly ramped from zero at time $t_0 = 0$ s to a value of $\boldsymbol{f} = f_y = 10$ m s^{-2} at time $t_1 = 10$ s. At this point ($t_1 = 10$ s), the self-weight is suddenly removed and the column is free to oscillate around its equilibrium position (which matches the undeformed configuration) until the analysis is stopped at $T = 20$ s.

(a) Strain Energy time plot.

(b) (Grid) Kinetic Energy time plot.

(c) Total Energy time plot.
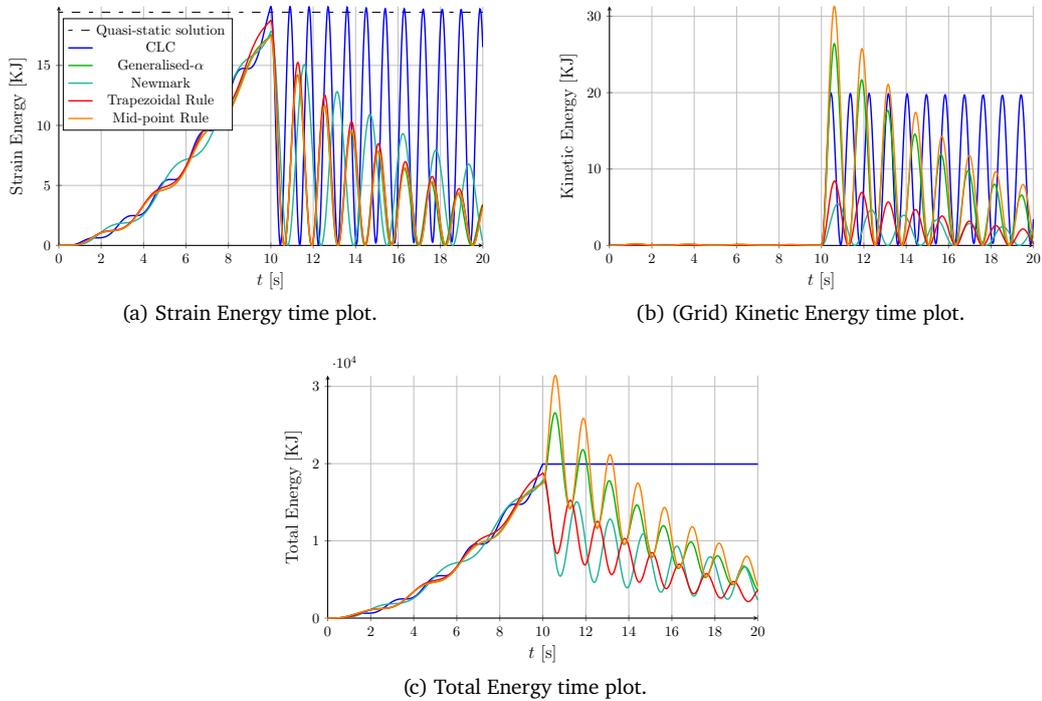
Figure 3.8: Comparison of the different energy behaviours for the 1D column under self-weight.
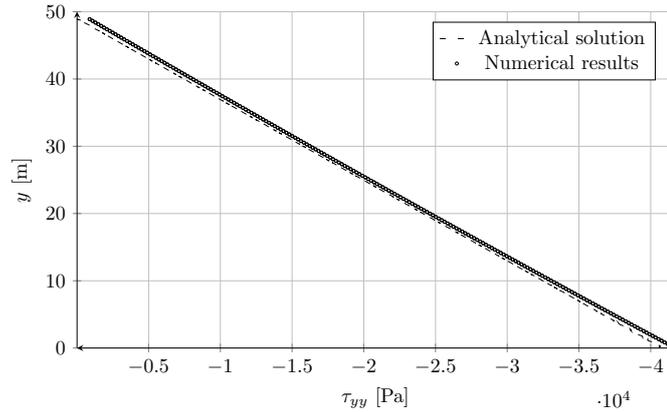
Figure 3.9: Comparison between analytical solution and numerical results for the Kirchhoff stresses τ_{yy} plotted along depth of the column at $t = 10$ s.

Results discussion   Figure 3.8 reports the results in terms of energy values for the different considered time discretisations. It can be observed how using different time schemes leads to different oscillations periods of the column. Comparison of the strain energy (Figure 3.8(a)) shows good agreement between the quasi-static solution (obtained via monotonic increase of the loading) and the CLC dynamic formulation, with the difference (CLC exhibits a little more strain energy) due to the time-sampling of the CLC method. From the same figure it can observed how other methods start to dissipate during the loading phase, while showing very similar dissipative trends when the load is removed.

Figure 3.8(b) highlights how the generalised-α method and the mid-point rule overestimate substantially the kinetic energy. The ratio between the highest value of the kinetic energy of the generalised-α method and the mid-point rule with the same value for the CLC analysis is considerable $\approx 133\%$ and $\approx 157\%$. On the other hand, the Newmark method and the trapezoidal rule dissipate kinetic energy, with highest values of kinetic energy being $\approx 27\%$ and $\approx 42\%$ of the highest value from the CLC analysis.

Figure 3.8(c) combines the energy trends coming from the previous two figures. It is interesting to see how, even if the energy values differs quite substantially when the load is removed, the final values of the total energy are similar, being in the interval between $\approx 12\%$ (represented by generalised-α scheme) and $\approx 21\%$ (in the case of the mid-point rule) when compared to the CLC constant value. This is representative of the dissipative behaviour shown by any method but the CLC, which make these methods questionable, as their results can change significantly depending on when the simulation is stopped.

Figure 3.9 compares the values of the vertical stresses along the depth of the column obtained by this numerical simulation at the time $t = 10$ s and with the analytical solution. Good agreement is shown between the two sets of values, with the numerical results exhibiting slightly higher values of stresses. This is due to the time-sampling of the step, with the accelerations marginally contributing as further body load to the equilibrium equations.

3.7.2 Comparison with manufactured solution for the elasto-plastic Conservation Law Consistent (CLC) material
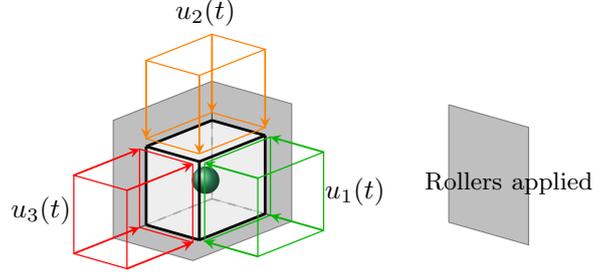
Figure 3.10: Illustration of the unit cube containing one MP and applied BCs for the MMS problem.

Example scope While the previous example has tested the validity of the elastic constitutive relationship, the current example aims at verifying elasto-plastic behaviour. Moreover, since the CLC formulation requires a non-trivial modification of the stress-strain relationship (see Eq. (3.142)), it was decided to introduce an intermediate test to verify that the code was truly capable of representing an elasto-plastic material. The MMS appears adequate to comply with this purpose. This method consists (in brief) of assuming an infinitely differentiable solution (i.e., a displacement field, in this case) and analytically computing the source terms (i.e., the body forces or the applied tractions/stresses) by solving the partial differential equation backwards. These terms are then used as input in the code to check code results with the predefined solution.

In particular, one of the verification tests proposed by Kamojjala *et al.* [69] can verify the implementation of an elasto-plastic material.

Setup This simulation (illustrated in Figure 3.10) considers a material whose shear modulus is $G = 79$ GPa and its Poisson's ratio is $\nu = 1/3$. *Von Mises* yield function is considered, i.e.,[*]

$$\Phi\left(\boldsymbol{\tau}\right) = \frac{\sqrt{2\,J_2}}{\rho_y} - 1 \leq 0, \tag{3.187}$$

with $J_2 = \frac{s_{ij}\,s_{ji}}{2}$, $s_{ij} = \tau_{ij} - \frac{\tau_{kk}}{3}\delta_{ij}$.

This analysis consists in prescribing a (small) strain history to a single element populated by one MP and to verify its Cauchy stress response. Since this example cannot suffer from cell-crossing instability nor the small-cut issue, MPM shape functions and no boundary stabilisation have been considered. To match the solution provided for this MMS example, the inertia effects are neglected. The driving strains start from a null value at $t = 0$ and are later prescribed as follows

$$\epsilon_{11} = \begin{cases} -0.003 & \text{if} \quad 0 < t \leq 1; \\ -0.0103923 & \text{if} \quad 1 < t \leq 2; \end{cases} \tag{3.188}$$

---

[*]No appreciable difference between Kirchhoff stress and Cauchy stress, since this example is set in the small strain context.
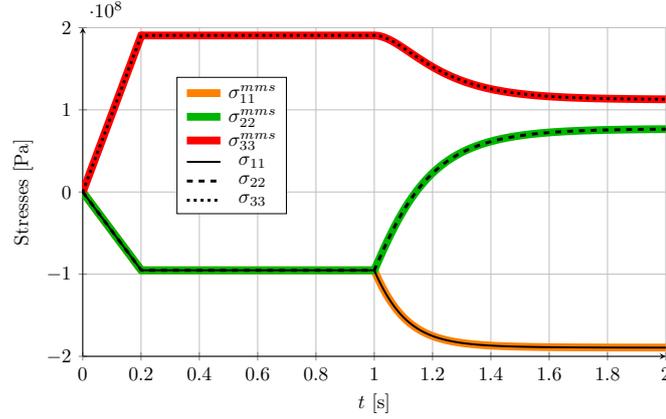
Figure 3.11: Stress paths comparison between results from the MMS (denoted by the superscript $(\bullet)^{mms}$) and numerical simulation (with no superscript).

$$\epsilon_{22} = \begin{cases} -0.003 & \text{if} \quad 0 < t \leq 1; \\ 0 & \text{if} \quad 1 < t \leq 2; \end{cases} \tag{3.189}$$

$$\epsilon_{33} = \begin{cases} 0.006 & \text{if} \quad 0 < t \leq 1; \\ 0.0103923 & \text{if} \quad 1 < t \leq 2, \end{cases} \tag{3.190}$$

where the time is expressed in seconds. The physical interpretation of this strain path and the setup of the problem can be found in Kamojjala *et al.* [69]. Time-varying stresses under constant strain rate (as described by Eqs. (3.188)-(3.190)) can be computed in correspondence to constant strain rates (see, for instance, Krieg and Krieg [70] for the described procedure), resulting in the following stress paths

$$\sigma_{11} = \begin{cases} -474\,t & \text{if} \quad 0 < t \leq 0.201; \\ -95.26 & \text{if} \quad 0.201 < t \leq 1; \\ \frac{189.4 + 0.1704\,\sqrt{a} - 0.003242\,a}{1 + 0.00001712\,a} & \text{if} \quad 1 < t \leq 2; \\ 189.4 & \text{if} \quad t > 2; \end{cases} \tag{3.191}$$

$$\sigma_{22} = \begin{cases} -474\,t & \text{if} \quad 0 < t \leq 0.201; \\ -95.26 & \text{if} \quad 0.201 < t \leq 1; \\ \frac{76.87 + 1.443\,\sqrt{a} - 0.001316\,a}{1 + 0.00001712\,a} & \text{if} \quad 1 < t \leq 2; \\ 76.87 & \text{if} \quad t > 2; \end{cases} \tag{3.192}$$

$$\sigma_{33} = \begin{cases} 948\,t & \text{if} \quad 0 < t \leq 0.201; \\ 190.5 & \text{if} \quad 0.201 < t \leq 1; \\ \frac{-112.5 + 1.272\,\sqrt{a} - 0.001926\,a}{1 + 0.00001712\,a} & \text{if} \quad 1 < t \leq 2; \\ 112.5 & \text{if} \quad t > 2, \end{cases} \tag{3.193}$$

where $a = \exp(12.33\,t)$. The time is expressed in seconds, while the stresses are given in MPa.

Results discussion   Figure 3.11 shows that results coming from the numerical solutions match those provided by Eqs. (3.191)-(3.193), confirming the correct implementation in the case of (small strain) linear elasticity and a von Mises yield function. This example better reproduces the solutions provided by Kamojjala *et al.* [69] (compared to Example 3.7.1, particularly to Figure 3.9) because of the choice of selecting the inertia effects for this case. Hence, while the internal force vector for this example follows the CLC implementation, the inertia term in the equilibrium equations is set as zero.

3.7.3   2D elastic cantilever beam

(a) Geometrical features and discretisation.

(b) Applied load time variation.
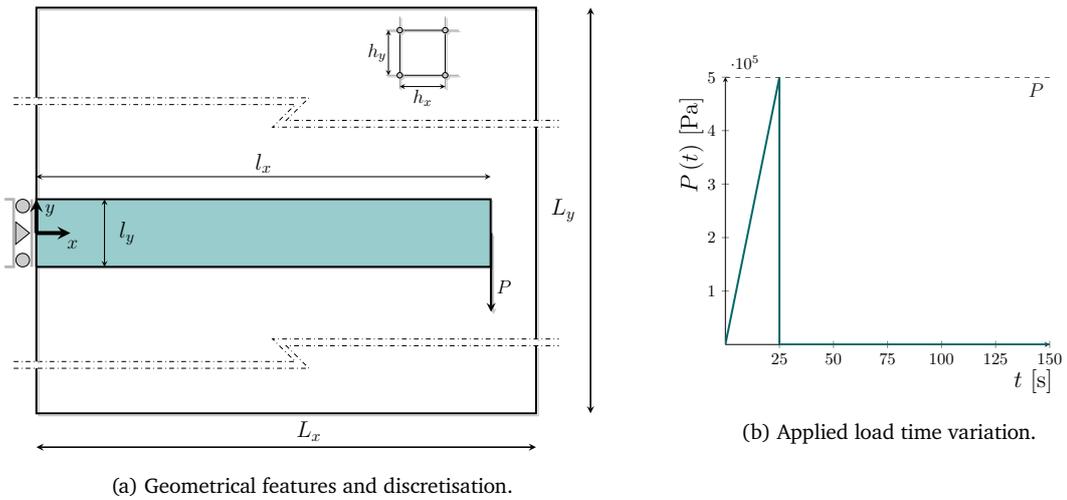
Figure 3.12: The cantilever beam problem. Figures reproduced from Pretti *et al.* [26].

Example scope   This example considers a bi-dimensional elastic beam represented in Figure 3.12(a), with the goal of assessing the periodic behaviour of a low-frequency structure for a fully bi-dimensional example. On top of this, two analyses investigate the applicability of the CLC to different materials: the former analysis, i.e., (A) considers a compressible material, while the latter simulation, i.e., (B), takes a nearly-incompressible material into account. Linearisation of the NR algorithm is also checked.

Setup   The material parameters, beam geometry and discretisation for the current analyses (A) and (B) are summarised in Table 3.5. The mechanical parameters are arbitrary, but Young's moduli and Poisson's ratios are thought to represent materials storing similar amounts of strain energy.

The time increments Δt shown in Table 3.5 correspond to 15 times the CFL time-step approximation defined at the beginning of the simulation by Eq. (3.126). From the same table it can be noticed how the number of MPs per element is particularly high. The reason for such an unusual number lies in the attempt to compensate the absence of the ghost stabilisation outlined in Section 3.6.5, which was not considered to fully assess the conservation properties (see, in this regards, the discussion in Section 3.6.6). The ghost stabilisation

plays an essential role especially when a grid element is active only due to few misplaced (from an integration position perspective) MPs. As one might predict, this situation is less likely to occur if the mesh is sufficiently coarse and the number of MPs per element is high. The chance of experiencing this issue increases with the order of the shape functions, as their relative stencil is larger*.

The analyses were run using GIMPM shape functions. As illustrated in Figure 3.12(b), the load, applied at the end of the free edge, is monotonically increased from 0 to its highest value P and then suddenly removed. The peak value of the load is reached at the time $t_1 = 25$ s, while the whole simulation carries on until time $T = 150$ s. The effects of gravity were neglected. The external load vector is represented by the point load, which is split between the two end MPs in the proximity of the x-axis (for a more detailed description of this loading conditions, see Charlton *et al.* [30]).

Table 3.5: Summary of the parameters considered in the analyses of the $2D$ elastic cantilever beam. Table reproduced from Pretti *et al.* [26].

Parameter Settings		Analysis (A)	Analysis (B)
Material Parameters	$\bar{E}$	10^9 Pa	$3 \cdot 10^8$ Pa
	ν	0.2	0.49
	ρ_0	‡$7750 \cdot 10^3$ kg/m^3	
Geometry, Load and Timings	l_x, l_y	10, 1 m	
	P	500 kN	
	t_0, t_1, T	0, 25, 150 s	
Analysis Parameters	L_x, L_y	11, 20 m	
	h_x, h_y	0.5 m	
	†mmp	12	
	Δt	≈ 0.6264 s	≈ 0.2194 s

‡ The density of the beam has been artificially increased to create a low-frequency problem.
† mmp is the number of MPs per direction per element.

Results discussion   The periodic behaviour of the structure's energies is reported in Figure 3.13(a) for analysis (A), where total, MPs' kinetic, and strain energies are considered. The time-steps denoted by the letters $A \ldots D$ in the same figure correspond to sensitive configurations: the last time-step where the load is applied before its removal ($A \approx 24.43$ s, 39th time-step), the maximum value of the strain energy ($B \approx 29.44$ s, 47th time-step, and $D \approx 55.12$ s, 88th time-step), and the maximum value of the kinetic energy ($C \approx 42.59$ s, 68th time-step). As shown in Figures 3.13(b) and 3.13(c), after the load is removed (just after time-step A), the beam continues to move downwards until the inertia is wholly

---

*A qualitative description of the above idea can be understood comparing the domains of influence in Figures 3.4(c) and 3.4(d).
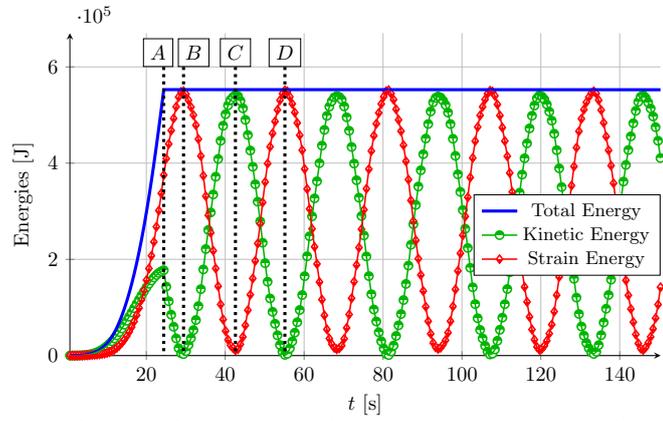
(a) Time evolution of total, kinetic and strain energy for the cantilever beam problem, analysis (A).

(b) Strain Energy, time-step A, analysis (A).

(c) Kinetic Energy, time-step A, analysis (A).

(d) Strain Energy, time-step B, analysis (A).

(e) Kinetic Energy, time-step B, analysis (A).

(f) Strain Energy, time-step C, analysis (A).

(g) Kinetic Energy, time-step C, analysis (A).

(h) Strain Energy, time-step D, analysis (A).

(i) Kinetic Energy, time-step D, analysis (A).
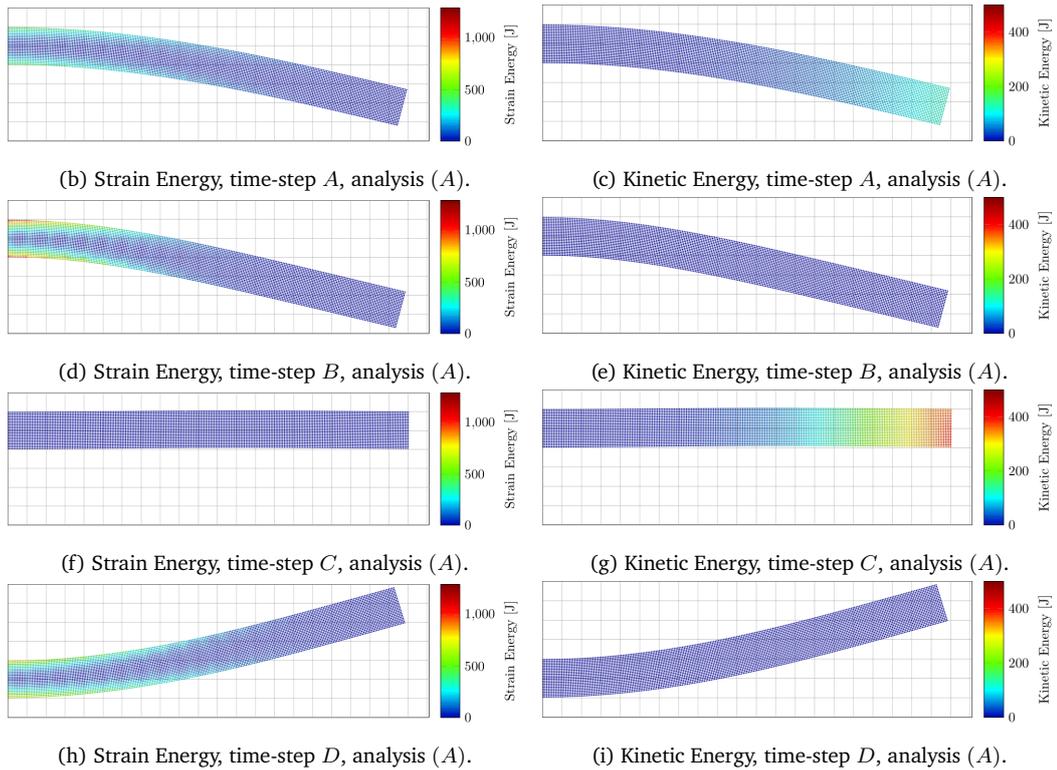
Figure 3.13: Time plot of the energies during the simulation (top row) and deflection shapes representing the strain (left) and kinetic (right) energies for time-steps $A \dots D$. Graphs refer all to analysis (A). Figures reproduced from Pretti *et al.* [26].
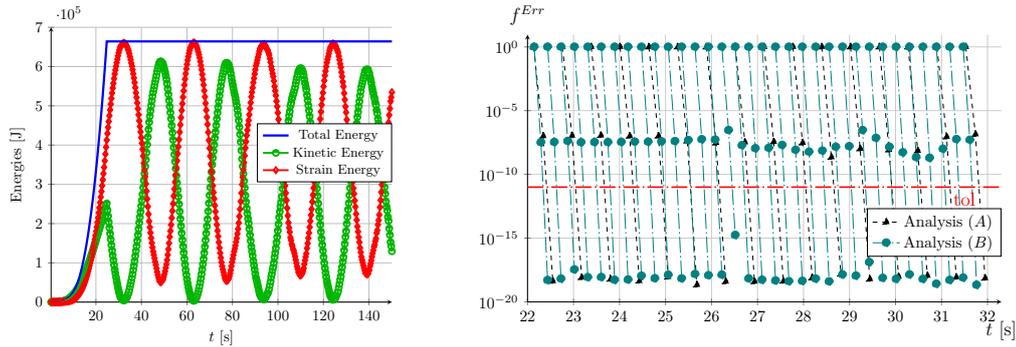
(a) Time evolution of total, kinetic and strain energy for the cantilever beam problem, analysis (B).

(b) Convergence of the Newton-Raphson scheme for time between 22 s and 32 s.

Figure 3.14: Time plot of the energies for simulation (B) (on the left-hand side), and convergence comparison for both the simulations (on the right-hand side). Figures reproduced from Pretti *et al.* [26].

converted into strain energy (Figures 3.13(d) and 3.13(e), time-step B). As expected, Figure 3.13(d) shows that a considerable amount of strain energy is located at the external fibres in the fixed end. On the other hand, the beam is unstrained in its original configuration at time-step C (Figure 3.13(f)), while the kinetic energy (Figure 3.13(g)) is distributed with a gradient along the x-axis, independent of the MPs' position along the vertical direction. Time-step D represents another maximum point of the strain energy, with D's deflection shape horizontally mirroring B's. However, when comparing the strain energy of the two configurations (Figure 3.13(h) and 3.13(d), respectively), it can be seen how the external fibres in B are more heavily loaded than in D. With a closer look, it can be noted how the displacements at the free end of the beam in points B and D do not coincide. This inaccuracy is due to time sampling, which does not accurately capture the peak strain energy. Therefore, points B and D are not temporally spaced as the proper period of the beam would predict.

The energies time-response of simulation (B) is plotted in Figure 3.14(a). The structure presents periodic behaviour with a lower frequency time compared to simulation (A). It can be seen how the total energy for this simulation is slightly higher than for analysis (A), which is due to a higher kinetic energy accumulated in the loading phase for simulation (B). Comparing the behaviours in correspondence of the minimum values of strain energy for simulation (A) (Figure 3.13(a)) and (B) (Figure 3.14(a)), it can be seen that this latter simulation presents less smooth behaviour. This is due to the numerical errors accumulated in the proximity of the boundaries of the body (and their likelihood of occurring or not, as explained above) due to the absence of ghost stabilisation, which appear to be more evident for simulation (B). This is possibly due to a slightly more stiff material for simulation (B), as testified by the higher value of energy (compare Figure 3.13(a) with 3.14(a)), which implies a material more capable of storing energy (and relative errors too), under the same conditions.

Figure 3.14(b) shows the convergences of the NR algorithm for both the analyses, consider-

ing the time-steps from ≈ 22 s to ≈ 32 s. Even though these steps take difficult stages of the simulations into account (with the release of the load happening at $t = 25$ s), it can be seen how the algorithms converge smoothly within three iterations for both the simulations.

3.7.4 Collision of elastic cylinders
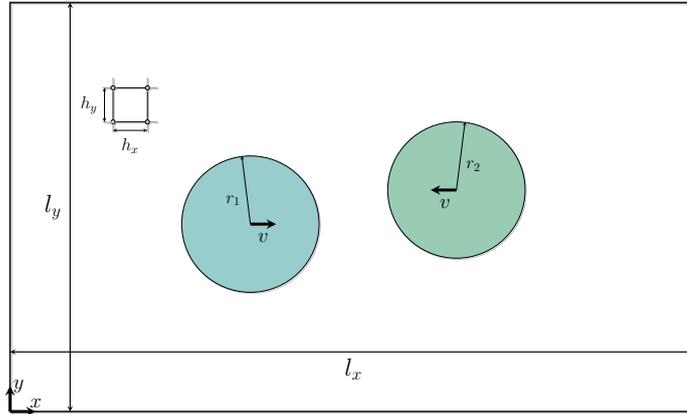
Figure 3.15: Illustration of the initial conditions of the cylinder impact problem. Figure reproduced from Pretti *et al.* [26].

Example scope This fourth example considers the collision of two elastic cylinders, evaluating the energy conservation in the case of a low-frequency elastic collision. This example bears similarities with the numerical examples in other MPM papers (see [25, 50, 71, 72]).

Table 3.6: Summary of the parameters considered in the analysis of the elastic cylinder collision. Table reproduced from Pretti *et al.* [26].

Parameter Settings		
Material Parameters	$\bar{E}_1$, $\bar{E}_2$	100 Pa
	ν_1, ν_2	0.3
	$(\rho_0)_1$, $(\rho_0)_2$	5 kg/m^3
Geometry, Velocities and Timings	l_x, l_y	20, 12 m
	r_1, r_2	2 m
	(x_1, y_1), (x_2, y_2)	(5.8, 5.5) m, (14.2, 6.5) m
	$\|v_1\|$, $\|v_2\|$	0.75 m/s
	T	8 s
Analysis Parameters	h_x, h_y	2 m
	†mmp	35
	Δt_0	≈ 0.0573 s
	$\min(\Delta t)$, $\max(\Delta t)$	≈ 0.0562, ≈ 0.0655 s

† mmp is the number of MPs per direction per element.
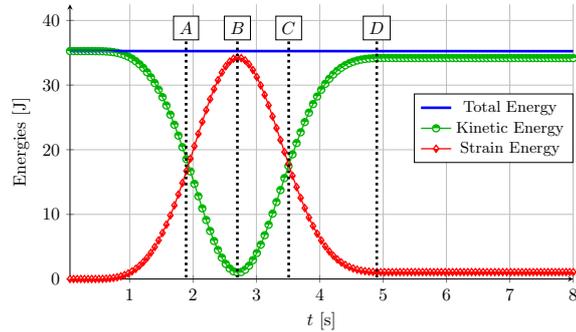
(a) Time evolution of total, kinetic and strain energy for the elastic cylinders collision.

(b) Strain Energy, time-step A.

(c) Kinetic Energy, time-step A.

(d) Strain Energy, time-step B.

(e) Kinetic Energy, time-step B.

(f) Strain Energy, time-step C.

(g) Kinetic Energy, time-step C.

(h) Strain Energy, time-step D.

(i) Kinetic Energy, time-step D.
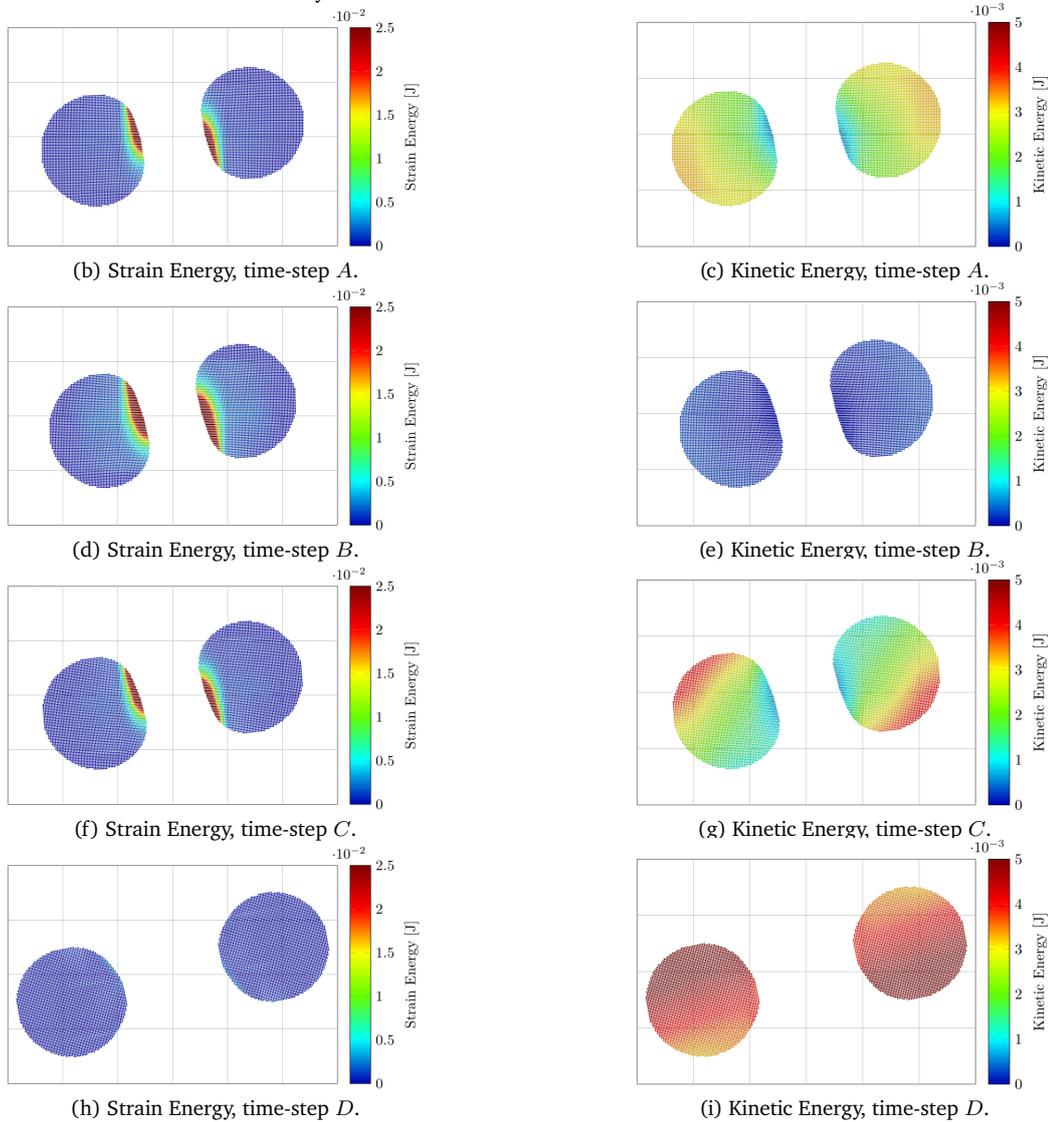
Figure 3.16: Time plot of the energies during the simulation (top row) and deflection shapes representing the strain (left) and kinetic (right) energies for time-steps $A \ldots D$. Figures reproduced from Pretti *et al.* [26].

Setup The parameters chosen for this simulation are listed in Table 3.6, and they are selected to represent a low-frequency problem. As also shown in Table 3.6, the time-step size was not constant during the simulation. The selected time-step length is computed at every new step and is five times more than that defined by Eq. (3.127). This time-step length (a few times above the CFL condition) has been selected as a compromise between computational effort and NR algorithm stability. Nonetheless, the variation of such size is minimal, and this is due to the grid size. As for the previous example, the grid lengths and number of MPs per direction per element (mmp) are both significant, as ghost stabilisation has also not been enforced in the simulation. The even higher number mmp can be justified by the collision occurring, which, compared to the previous example, might overlap the MPs and the mesh in more unpredictable and small-cut prone ways. MPM shape functions are used to run this example, and gravity effects are neglected.

Results discussion Figure 3.16(a) shows the time evolution of the energies for four selected time-steps corresponding to $A \approx 1.89$ s, 32nd time-step; $B \approx 2.7$ s, 45th time-step; $C \approx 3.51$ s, 58th time-step; and $D \approx 4.9$ s, $80^{t}h$ time-step. In particular, A and C correspond to similar levels of energies (both kinetic and strain energy), with the former being more uniformly spread on MPs (Figures 3.16(c), 3.16(g)), and the latter being more concentrated on the MPs in the proximity of the nodes where contact* between cylinders is taking place (Figures 3.16(b), 3.16(f)). However, even if the overall amount of strain and kinetic energies in A and C appear to be very similar, these steps present different distributions of energies on MPs, as is clear from a comparison of Figures 3.16(b), and 3.16(c) for time-step A, with Figures 3.16(f), and 3.16(g) for C. The minimum kinetic energy (Figure 3.16(e)) corresponding to a maximum of the strain energy (Figure 3.16(d)) occurs at time-step B. From this point on, kinetic energy starts to be recovered until time-step D is reached. This time-step marks the end of contact between cylinders. As expected from elastic bodies, the deflection shapes of the cylinders (Figures 3.16(h), and 3.16(i)) recover circular sections, even though some elastic waves slightly modify their shapes.

3.7.5 Elasto-plastic impact of cylinders

Example scope While extending the investigation of the results from the previous example to an elasto-plastic impact, the following examples investigate the role of different mesh sizes and shape function stencils. Similar simulations have been discussed by Meng and Laursen [43], and Love and Sulsky [45].

Setup Analysis (A) (run using the original MPM functions) makes use of all of the parameters in Table 3.6, with the exception that the stress-strain relationship is elasto-plastic. While elasticity is given by an Hencky model (whose elastic moduli are available in Table 3.6), the von Mises yield function (see Eq. (3.187)) governs the plasticity.

---

*In this chapter and its relative simulations, by contact it is meant the no-slip standard automatic algorithm in the MPM (see, for a more detailed explanation, Love and Sulsky [25]), which differs substantially in the implementation from the (frictional) contact outlined in Chapter 5.
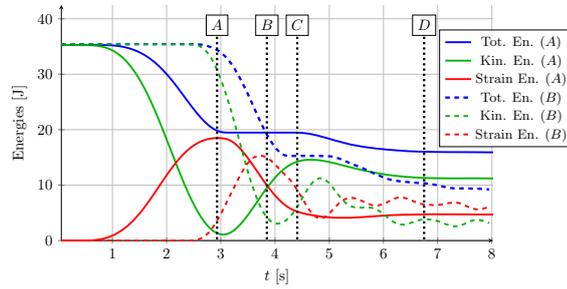
(a) Time evolution of total, kinetic and strain energy for the elasto-plastic cylinders collision.
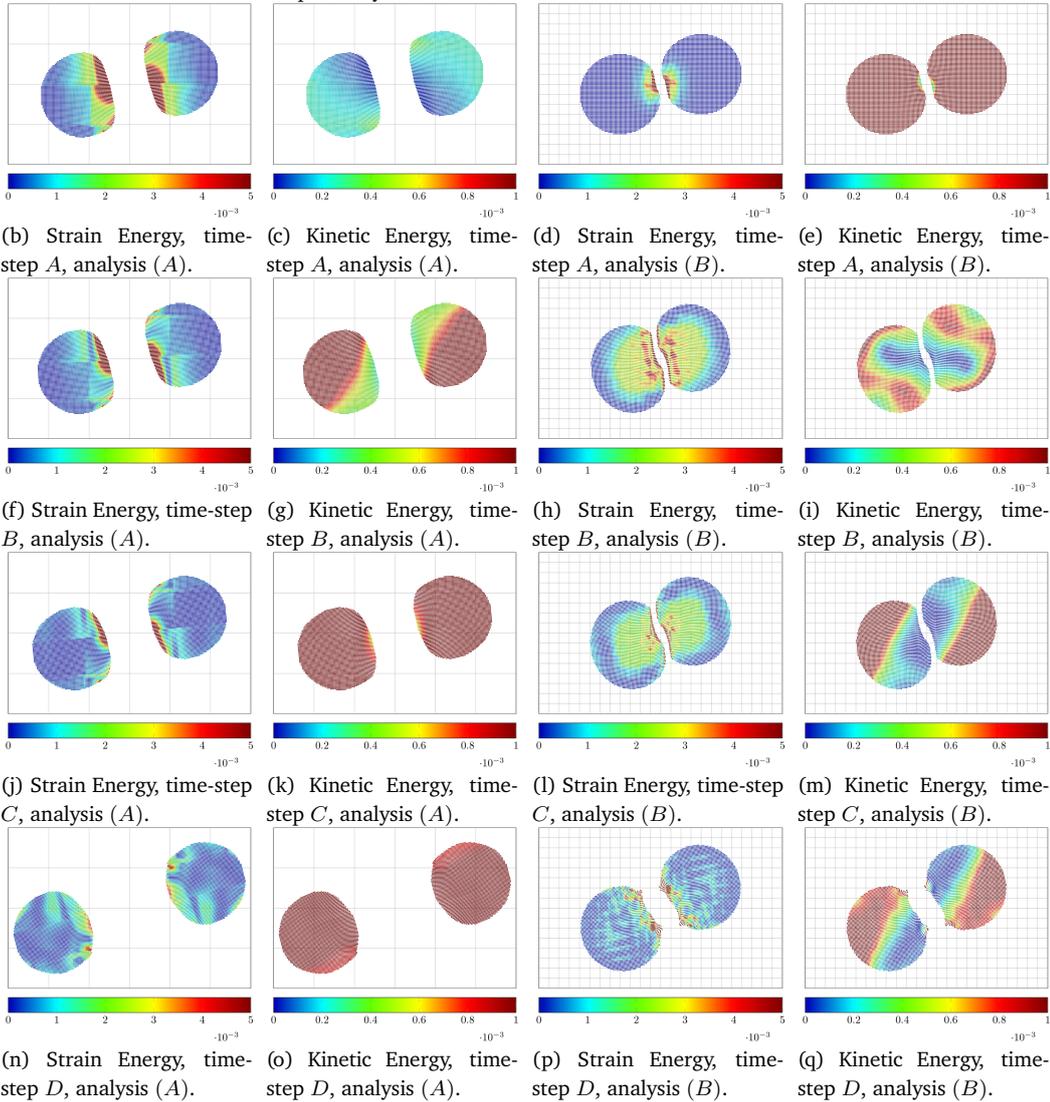
(b) Strain Energy, time-step A, analysis (A).

(c) Kinetic Energy, time-step A, analysis (A).

(d) Strain Energy, time-step A, analysis (B).

(e) Kinetic Energy, time-step A, analysis (B).

(f) Strain Energy, time-step B, analysis (A).

(g) Kinetic Energy, time-step B, analysis (A).

(h) Strain Energy, time-step B, analysis (B).

(i) Kinetic Energy, time-step B, analysis (B).

(j) Strain Energy, time-step C, analysis (A).

(k) Kinetic Energy, time-step C, analysis (A).

(l) Strain Energy, time-step C, analysis (B).

(m) Kinetic Energy, time-step C, analysis (B).

(n) Strain Energy, time-step D, analysis (A).

(o) Kinetic Energy, time-step D, analysis (A).

(p) Strain Energy, time-step D, analysis (B).

(q) Kinetic Energy, time-step D, analysis (B).

Figure 3.17: Time plot of the energies during the simulation (top row) and deflection shapes representing the simulation (A) (first two columns) and (B) (last two columns). Strain energy (first and third column) and kinetic energy (second and fourth) column are considered for time-steps $A \dots D$. Physical dimension of contours are in Joule [J]. Figures reproduced from Pretti *et al.* [26].

On the other hand, analysis (B) (run using the GIMPM functions) presents a mesh 4 times finer than the one used for (A), with $mmp = 8$. The parameters (concerning both simulations (A) and (B)) that differ from those used in Table 3.6 are presented below:

- $(\rho_y)_1 = (\rho_y)_2 = 10$ Pa, being the yield stress;

- $\min(\Delta t) \approx 0.0571$ s, and ≈ 0.06020 s for analyses (A) and (B), respectively (being 5 times the minimum time-step size defined by Eq. (3.127)); and

- $\max(\Delta t) \approx 0.0666$ s, and ≈ 0.06023 s for analyses (A) and (B), respectively (being 5 times the maximum time-step size defined by Eq. (3.127)).

Ghost stabilisation has not been considered for these simulations.

Results discussion   Four time-steps were selected, and their relative deflection shapes and energies are represented in Figures 3.17(b)-3.17(o): A at ≈ 2.9 s, B at 3.9 s, C at 4.4 s, and D at 6.7 s. From Figure 3.17(a), it can be noticed how the start of the collision, governed by the mesh size and the different stencils of the shape functions, is delayed for simulation (B). In this sense, it can be seen how time-step A, which corresponds to a minimum for the kinetic energy for simulation (A) (see Figure 3.17(c)), defines instead the initial stages of contact for analysis (B) (Figures 3.17(d) and 3.17(e)). In turn, time-step B is representative of the same levels of kinetic and strain energies for simulation (A) (Figures 3.17(f) and 3.17(g)), while it is particularly close to the peak of contact for analyses (B) (maximum in strain energy, Figure 3.17(h)). A second yielding takes place starting from time-step C and, although with a decreasing trend, continues until the end of the simulation (A). The cause beyond this second yielding can be found in the waves propagating through the bodies after the collision, whose deviatoric part is progressively damped by the von Mises yield function. Time-step D defines the end of contact for simulation (A) (constant total energy, Figures 3.17(n) and 3.17(o)), while it can be appreciated from Figures 3.17(p) and 3.17(q) how contact is still persistent for analysis (B) at this stage.

From a comparison of the strain energies distributions of analysis (A) (Figures 3.17(b), 3.17(f), 3.17(j), and 3.17(n)) with their counterparts of simulation (B) (Figures 3.17(d), 3.17(h), 3.17(l), and 3.17(p)), it is clear that the mesh refinement and the use of different shape functions do not entirely eliminate the dependency of the strain energies from the mesh. From a closer look, it can be appreciated how such strain energy localisations take place in correspondence of element edges or corners. This can be explained by the cell-crossing instability (moderated though in simulation (B) by the use of GIMPM shape functions), which is also aggravated by the absence of the ghost stabilisation for the boundary cells (see, for similar results, Coombs [49]).

3.7.6   Impact of a Taylor bar

Example scope   This example (see Figure 3.18) is intended to compare the results obtained from this three-dimensional analysis with those in the literature (for FEM analyses

Table 3.7: Summary of the parameters considered in the analysis of the Taylor bar impact. Table reproduced from Pretti *et al.* [26].

Parameter Settings		
Material	$\bar{K}$, G	130, 43.3 GPa
Parameters	ρ_0	8930 kg/m^3
Geometry,	h_0, r_0	32.2, 3.2 mm
Velocities	L_x, L_y, L_z	7.2, 32.2, 7.2 mm
and Timings	$\|\boldsymbol{v}\| = v_y$,	-227 m/s
	T	80 μs
Analysis	h_x, h_y, h_z	0.8, 0.648, 0.8 mm
Parameters	†mmp	2
	Δt_0	≈ 0.375 μs
	$\min{(\Delta t)}$, $\max{(\Delta t)}$	≈ 0.371, ≈ 0.375 μs

† mmp is the number of MPs per direction per element.
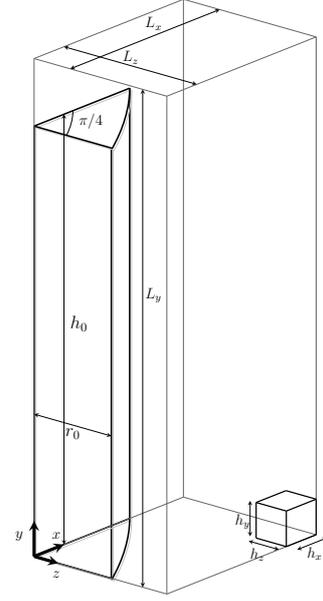
Figure 3.18: Illustration of the initial conditions of the Taylor bar problem. Figure reproduced from Pretti *et al.* [26].

see [43, 73–76], while see [45] for MPM) for an elasto-plastic impact at high frequencies. In this case, a hardening rule was also considered. This example constitutes a benchmark for transient dynamic codes.

Setup   The list of parameters (inherited from the literature) necessary to run the analysis is presented in Table 3.7. GIMPM functions are used. To compare the results with the literature, the von Mises yield function exhibiting linear hardening is considered, this being

$$\Phi\left(\boldsymbol{\tau}, q\right) = \frac{\sqrt{2\,J_2}}{\sigma_y} - \sqrt{\frac{2}{3}}\left(1 - \frac{q}{\sigma_y}\right) \leq 0, \tag{3.194}$$

where the value of the yield stress is $\sigma_y = 0.4$ GPa. The assumed hardening parameter is $H = 0.1$ GPa, which relates the power-conjugates for the linear isotropic hardening case as follows, i.e., $q = -H\,\alpha$. The time-step length is kept 15 times the value given by Eq. (3.127). Given the lower number of MPs per element originally chosen, this simulation uses the ghost stabilisation to avoid the problems described in the previous simulations. The chosen parameters for stabilising the stiffness and the mass matrix are, respectively $\gamma^K = \frac{9\,\bar{K}\,G}{3\,\bar{K}+G} = 10^{-6}$ and $\gamma^M = \rho_0 \cdot 10^{-6}$. As outlined by Coombs [49], the face ghost method is not particularly sensitive to the penalty values. Thus, the values used in this analysis have been chosen to modify the primary equations and the stabilised P2G mapping (3.183) as little as possible.
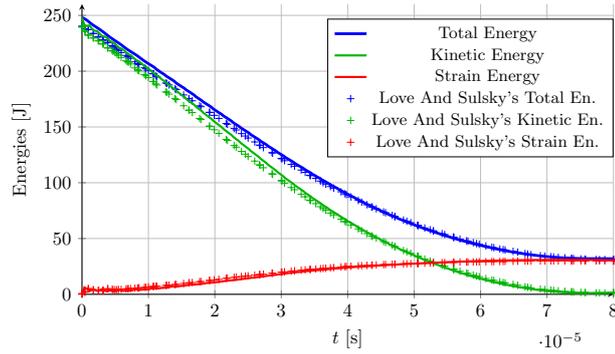
(a) Time evolution of total, kinetic and strain energy for the elasto-plastic cylinders collision.
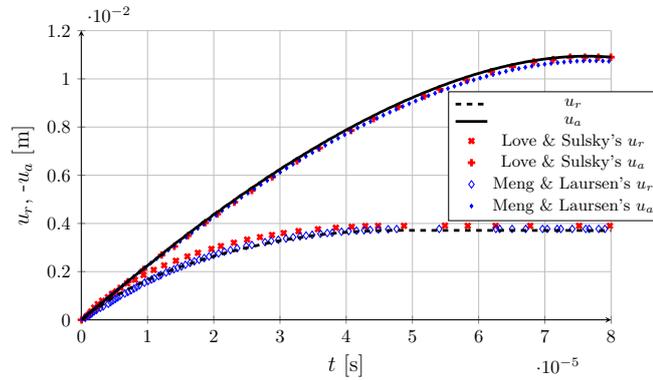
(b) Displacement history of the maximum radial and axial displacements.

Figure 3.19: Time plots of energies and displacements for the Taylor bar problem. Results from Love and Sulsky [45] are obtained via the MPM, while those from Meng and Laursen [43] via the FEM. Figure adapted from Pretti *et al.* [26].

Results discussion The time-history of the energies is represented in Figure 3.19(a)), and the values are compared with the results coming from the MPM analysis coming from [45]. The results are mostly overlapping, with a slight difference in the initial values of the energies. From a closer look, it can be appreciated how this dissimilar values possibly due to initial different discretisations, as it can be seen from the value at time $t = 0$ s of the total energies.

As for the displacement values illustrated in Figure 3.19(b), the results of the current analysis are compared with those coming from Love and Sulsky [45] for the MPM and Meng and Laursen [43] for the FEM. Despite the initial different discretisations discussed above between this analysis and that from [45], it can be seen that the radial displacements match among the three simulations. As for the radial displacements, it can be seen that the current analysis is more similar to the results coming from the FEM analysis [43] than those from [45] for the MPM. This difference might be due to the initial discretisation error again.

The contours of the displacement along the y-direction and the quantity $\sqrt{2J_2}$ are reproduced in Figure 3.20. As can appreciated from Figure 3.20(a), the displacement field appears to be very regular, while cell-crossing instability partially affects the considered meas-
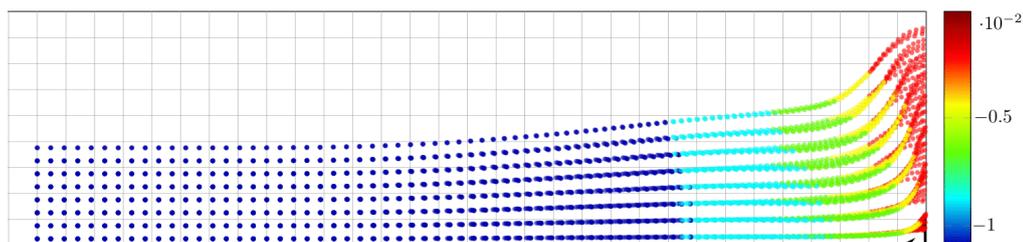
(a) Deflection shape in the zy-plane. Contours represent the displacement along the y-direction (expressed in [m]).
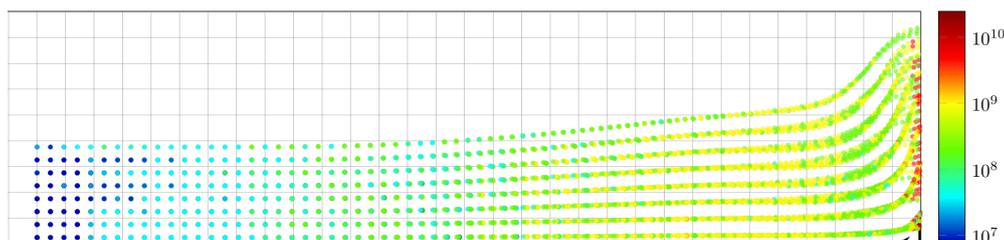
(b) Values of $\sqrt{2J_2}$ plotted in logarithmic scale in the zy-plane. Contours represent the stress values in [Pa].

Figure 3.20: Contour values for the Taylor bar problem plotted on the deflected shape at the end of the simulation.

ure of stress (see Figure 3.20), despite the use of GIMPM shape functions.

3.8 Observations

Some final conclusions are now drawn based on the considerations raised during the chapter and the provided examples.

The CLC method provides an effective formulation that complies with all of the underlying principles of continuum mechanics, making the simulations not only physically meaningful but energetically stable. In particular, this chapter has detailed the UL novel version of such an algorithm. While attractive because of its compliance with the underlying continuum laws, the method comes with the price of modifying the chosen constitutive relationship. As the bulky linearisation in Appendix B.1 emphasises, these modifications come with a particular burdensome cost regarding components appearing in the stiffness matrix (see Eq. (B.34)). As for the mapping procedures, their sources of dissipation have been thoroughly investigated. It was concluded that the P2G initial mapping introduces an error between the energy stored at the MPs' location and the grid one. Contrary to a share of the literature, it was proved that the PIC mapping as a standalone technique is not dissipative, while the FLIP method *per se* is. However, the error introduced by the initial mapping is cleverly cancelled by an equal and opposite error when the FLIP is employed. Since the PIC procedure does not introduce any error, the initial error accumulates in this case. The rigorous quantification of these phenomena in terms of energy (but also in terms of conservation of momenta) and their induced conclusions constitute a novelty in the literature.

As for the numerical analyses, Example 3.7.5 shows how the cell-crossing instability is not

entirely resolved (but certainly mitigated) by GIMPM functions. In this regard, using B-splines (as proposed, for instance, by Yamaguchi *et al.* [34]) on the grid (in conjunction with a Dirac delta distribution as a characteristic function) should be encouraged, since it would increase the continuity of the functions above $H^1(\omega)$ spaces. However, this type of approach has two disadvantages: the B-splines must be locally modified to directly enforce the Dirichlet BCs on the grid (even in the case of a conforming grid), and the size of stencil increases for these functions. The former issue could be treated with the use of specific techniques developed in the context of mesh-free methods (see Fernández-Méndez and Huerta [77], for instance). The latter issue entails a higher computational cost (larger bandwidth of the matrices to be inverted), without unfortunately guaranteeing an improvement in convergence as expected in the case of the FEM (see, in this regards, the examples shown by Charlton *et al.* [30] and Coombs *et al.* [62]). An increase in the stencils would also make boundary stabilisation more necessary. To avoid stabilising more layers of faces in correspondence of the external boundaries, Yamaguchi *et al.* [34] has opted to modify the basis functions when considering B-splines to mitigate the small-cut issue.

In this respect, the ghost method briefly described in Section 3.6.5 stabilises the small-cut issue, proving its validity in challenging simulations as Example 3.7.6. However, it must be considered that a specific technique must be employed to use the ghost method in the case of higher-order shape functions, such as B-splines. Moreover, since the robustness of a numerical method also lies in the existence of different ways of tackling the same problem (an approach can be more or less effective depending on the situation), other techniques already present in the context of the cut-FEM (see the discussion in Section 3.6.5) could be incorporated into MPM in the future.

Insofar as the same rationale is applied to the application of non-conforming BCs (which is beyond the scope of this work), only few techniques are available in the literature. To the author's best knowledge, the works describing the application of Dirichlet BCs are Cortis *et al.* [78], and Liu and Sun [79]. While these papers provide a fundamental basis to tackle the application of non-conforming essential BCs in the MPM, they present some limitations. In particular, Cortis *et al.* [78] can be applied only to straight boundaries, while Liu and Sun [79] use the penalty method, i.e., an approximated method. The reader is referred to Section 5.4.3 for a discussion of the meaning of the term approximated in the context of the penalty method. If Neumann BCs are considered, Liang *et al.* [80] presented an interesting technique for the MPM. These authors offered a technique based on treating the stress on the boundaries (i.e., the boundary forces, via Cauchy stress theorem Eq. (3.51)) as arising from a virtual body exterior to the real discretised one. However, this technique can only be applied in specific cases of virtual stress expressions. Another study on the application of Neumann BCs was provided by Remmerswaal [35]. The gist of the method proposed by Remmerswaal [35] consists in body-fitting the MP-based discretisation (based on level-sets) via composite Bézier curve. Bing *et al.* [33] and Yamaguchi *et al.* [34] took both Dirichlet and Neuman BCs into account. Bing *et al.* [33] proposed a method similar to Remmerswaal [35], but B-splines functions are used to model the boundaries which are tracked by boundary points (these are progressively moved during the analysis with the displacement field). On

these reconstructed surfaces, Neumann BCs can be straightforwardly applied (using Gauss-Legendre quadrature points), while a technique similar to that proposed by Cortis *et al.* [78] and extended to the case of curved surfaces is employed for Dirichlet BCs. In this case, when non-homogeneous conditions vary spatially, a global minimisation problem must be solved to compute the nodal displacement values. Moreover, since this technique is adapted from the *implicit boundary method* (see [81]), its convergence rate is sub-optimal, as documented by the literature (see, for instance, [81] and [82]). Yamaguchi *et al.* [34] provided a method for imposing both Dirichlet (using Nitsche's weak imposition of non-conforming BCs) and Neumann BCs via extended B-splines and tracking all surfaces via boundary points. While the method presents some extremely interesting points (e.g., there is no need for boundary stabilisations), it should be noted that none of the examples provided by these authors considers non-conforming Dirichlet BCs. Moreover, the examples considering Neumann BCs in both Bing *et al.* [33] and Yamaguchi *et al.* [34] do not present significant distortion of these surfaces. This is due to the chosen technique for tracking the boundary, which cannot cope with the creation of new surfaces for a body undergoing extreme distortion conditions. Given the above, it is clear that there is still room for research to provide more choices in the application of boundary conditions.

As standard for the MPM, the integration error associated with using MPs as integration points remains an issue more significant than for standard FEM. A few attempts to solve this issue have been made (see, for instance, Sulsky and Gong [37] and Gan *et al.* [31]), but this issue is extremely challenging as it entails moving historical variables. The development of versatile techniques to improve this type of error would be a significant shift in the future growth of the MPM.

It should also be pointed out that the approximated CFL condition described by Eq. (3.127) scales inversely proportional to the square of MPs in one cell. While this factor still falls in favour of safety (as the time-step is always scaled down by it), it introduces an undesired dependency which could be eliminated by setting

$$\Delta t^{CFL} \approx \min_i \left(h_i^h \right) \min_{\mathsf{A}} \left(\sqrt{\frac{\rho^h}{M^h}} \right)_{\mathsf{A}}, \tag{3.195}$$

with

$$\rho_{\mathsf{A}}^h = \sum_{mp}^{\mathrm{N}^{mps}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp} \right) \rho^{mp}, \qquad M_{\mathsf{A}}^h = \sum_{mp}^{\mathrm{N}^{mps}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp} \right) M^{mp}. \tag{3.196}$$

In this fashion, the number of MPs in a cell appears at the numerator and denominator in the square root of Eq. (3.195) and it cancels out.

Chapter summary

 This chapter has introduced the Lagrangian description of the kinematics necessary to describe a solid material undergoing finite elasto-plastic strain. Other fundamental laws, such as mass conservation, rate of momenta balances and the first two principles of thermodynamics, have been described. The compliance of the considered constitutive law with the Clausis-Planck inequality via the Colemann-Noll procedure has been detailed. These underlying principles of the continuum mechanics have also been integrated with a discrete MPM algorithm. Discussions concerning where the MPM algorithm could present numerical dissipation have followed. Numerical examples have backed up such discussions.

References

[1] J. Grabe and L. Wu, 'Coupled Eulerian-Lagrangian simulation of the penetration and braking behaviour of ship anchors in clay,' *Geotechnik*, vol. 39, no. 3, pp. 168–174, 2016.

[2] J. T. Oden, *Finite elements of nonlinear continua*. Courier Corporation, 2006.

[3] R. W. Ogden, *Non-linear elastic deformations*. Courier Corporation, 1997.

[4] M. E. Gurtin, E. Fried and L. Anand, *The mechanics and thermodynamics of continua*. Cambridge University Press, 2010.

[5] E. A. de Souza Neto, D. Peric and D. R. Owen, *Computational methods for plasticity: theory and applications*. John Wiley & Sons, 2011.

[6] T. Belytschko, W. K. Liu, B. Moran and K. Elkhodary, *Nonlinear finite elements for continua and structures*. John wiley & sons, 2014.

[7] J. E. Marsden and T. J. Hughes, *Mathematical foundations of elasticity*. Courier Corporation, 1994.

[8] J. C. Simo, 'A framework for finite strain elastoplasticity based on maximum plastic dissipation and the multiplicative decomposition: Part I. continuum formulation,' *Computer Methods in Applied Mechanics and Engineering*, vol. 66, no. 2, pp. 199–219, 1988.

[9] E. Kröner, 'Allgemeine kontinuumstheorie der versetzungen und eigenspannungen,' *Archive for Rational Mechanics and Analysis*, vol. 4, no. 1, pp. 273–334, 1959.

[10] E. H. Lee, 'Elastic-plastic deformation at finite strains,' *Journal of Applied Mechanics*, vol. 36, no. 1, pp. 1–6, 1969.

[11] J. Mandel, *Plasticité Classique et Viscoplasticité* (CISM Courses and Lectures). Springer, 1971, vol. 97.

[12] J. R. Rice, 'Inelastic constitutive relations for solids: An internal-variable theory and its application to metal plasticity,' *Journal of the Mechanics and Physics of Solids*, vol. 19, no. 6, pp. 433–455, 1971.

[13] Y. F. Dafalias, 'The plastic spin,' 1985.

[14] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman lectures on physics, Vol. I: The new millennium edition: mainly mechanics, radiation, and heat*. Basic books, 2011, vol. 1.

[15] G. T. Houlsby and A. M. Puzrin, *Principles of hyperplasticity: an approach to plasticity theory based on thermodynamic principles*. Springer Science & Business Media, 2007.

[16] C. Truesdell, W. Noll, C. Truesdell and W. Noll, *The non-linear field theories of mechanics*. Springer, 2004.

[17] K. Bennett, R. Regueiro and R. Borja, 'Finite strain elastoplasticity considering the Eshelby stress for materials undergoing plastic volume change,' *International Journal of Plasticity*, vol. 77, pp. 214–245, 2016.

[18] J. Simo, 'Numerical analysis and simulation of plasticity,' *Handbook of numerical analysis,* vol. 6, pp. 183–499, 1998.

[19] B. Coleman and W. Noll, 'The thermodynamics of elastic materials with heat conduction,' *Archive for Rational Mechanics and Analysis*, vol. 13, p. 167, 1963.

[20] B. D. Coleman and M. E. Gurtin, 'Thermodynamics with internal state variables,' *The Journal of Chemical Physics*, vol. 47, no. 2, pp. 597–613, 1967.

[21] H. Hencky, 'The elastic behavior of vulcanized rubber,' *Journal of Applied Mechanics*, no. 1, pp. 45–48, 1933.

[22] R. von Mises, 'Mechanik der plastischen formänderung von kristallen,' *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 8, no. 3, pp. 161–185, 1928.

[23] R. Hill, 'A variational principle of maximum plastic work in classical plasticity,' *The Quarterly Journal of Mechanics and Applied Mathematics*, vol. 1, no. 1, pp. 18–28, 1948.

[24] J. Mandel, 'Contribution théorique à l'étude de l'écrouissage et des lois de l'écoulement plastique,' in *Applied Mechanics: Proceedings of the Eleventh International Congress of Applied Mechanics Munich (Germany) 1964*, Springer, 1966, pp. 502–509.

[25] E. Love and D. L. Sulsky, 'An unconditionally stable, energy–momentum consistent implementation of the material-point method,' *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 33-36, pp. 3903–3925, 2006.

[26] G. Pretti, W. M. Coombs, C. E. Augarde, B. Sims, M. M. Puigvert and J. A. R. Gutiérrez, 'A conservation law consistent updated Lagrangian material point method for dynamic analysis,' *Journal of Computational Physics*, vol. 485, p. 112 075, 2023.

[27] S. C. Brenner, L. R. Scott and L. R. Scott, *The mathematical theory of finite element methods*. Springer, 2008, vol. 3.

[28] S. G. Bardenhagen and E. M. Kober, 'The generalized interpolation material point method,' *Computer Modeling in Engineering and Sciences*, vol. 5, no. 6, pp. 477–496, 2004.

[29] J. L. González Acosta, P. J. Vardon, G. Remmerswaal and M. A. Hicks, 'An investigation of stress inaccuracies and proposed solution in the material point method,' *Computational Mechanics*, vol. 65, pp. 555–581, 2020.

[30] T. Charlton, W. Coombs and C. Augarde, 'IGIMP: An implicit generalised interpolation material point method for large deformations,' *Computers & Structures*, vol. 190, pp. 108–125, 2017.

[31] Y. Gan, Z. Sun, Z. Chen, X. Zhang and Y. Liu, 'Enhancement of the material point method using B-spline basis functions,' *International Journal for Numerical Methods in Engineering*, vol. 113, no. 3, pp. 411–431, 2018.

[32] M. Molinos, P. Navas, D. Manzanal and M. Pastor, 'Local maximum entropy material point method applied to quasi-brittle fracture,' *Engineering Fracture Mechanics*, vol. 241, p. 107 394, 2021.

[33] Y. Bing, M. Cortis, T. Charlton, W. Coombs and C. Augarde, 'B-spline based boundary conditions in the material point method,' *Computers & Structures*, vol. 212, pp. 257–274, 2019.

[34] Y. Yamaguchi, S. Moriguchi and K. Terada, 'Extended B-spline-based implicit material point method,' *International Journal for Numerical Methods in Engineering*, vol. 122, no. 7, pp. 1746–1769, 2021.

[35] G. Remmerswaal, 'Development and implementation of moving boundary conditions in the material point method,' M.S. thesis, TU Delft, 2017.

[36] S. Sticko, G. Ludvigsson and G. Kreiss, 'High-order cut finite elements for the elastic wave equation,' *Advances in Computational Mathematics*, vol. 46, no. 3, pp. 1–28, 2020.

[37] D. Sulsky and M. Gong, 'Improving the material-point method,' *Innovative Numerical Approaches for Multi-Field and Multi-Scale Problems: In Honor of Michael Ortiz's 60th Birthday*, pp. 217–240, 2016.

[38] J. Simo and N. Tarnow, 'The discrete energy-momentum method. Conserving algorithms for nonlinear elastodynamics,' *Zeitschrift für angewandte Mathematik und Physik ZAMP*, vol. 43, no. 5, pp. 757–792, 1992.

[39] R. Courant, K. Friedrichs and H. Lewy, 'On the partial difference equations of mathematical physics,' *IBM journal of Research and Development*, vol. 11, no. 2, pp. 215–234, 1967.

[40] Y. Sun, T. Shinar and C. Schroeder, 'Effective time step restrictions for explicit MPM simulation,' in *Computer Graphics Forum*, Wiley Online Library, vol. 39, 2020, pp. 55–67.

[41] D. Schillinger and M. Ruess, 'The Finite Cell Method: A review in the context of higher-order structural analysis of CAD and image-based geometric models,' *Archives of Computational Methods in Engineering*, vol. 22, pp. 391–455, 2015.

[42] O. Gonzalez, 'Exact energy and momentum conserving algorithms for general models in nonlinear elasticity,' *Computer Methods in Applied Mechanics and Engineering*, vol. 190, no. 13-14, pp. 1763–1783, 2000.

[43] X. Meng and T. Laursen, 'Energy consistent algorithms for dynamic finite deformation plasticity,' *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 15-16, pp. 1639–1675, 2002.

[44] F. Armero and I. Romero, 'On the formulation of high-frequency dissipative time-stepping algorithms for nonlinear dynamics. Part I: Low-order methods for two model problems and nonlinear elastodynamics,' *Computer Methods in Applied Mechanics and Engineering*, vol. 190, no. 20-21, pp. 2603–2649, 2001.

[45] E. Love and D. L. Sulsky, 'An energy-consistent material-point method for dynamic finite deformation plasticity,' *International Journal for Numerical Methods in Engineering*, vol. 65, no. 10, pp. 1608–1638, 2006.

[46]   S. Badia, F. Verdugo and A. F. Martín, 'The aggregated unfitted finite element method for elliptic problems,' *Computer Methods in Applied Mechanics and Engineering*, vol. 336, pp. 533–553, 2018.

[47]   K. Höllig, U. Reif and J. Wipper, 'Weighted extended B-spline approximation of Dirichlet problems,' *SIAM Journal on Numerical Analysis*, vol. 39, no. 2, pp. 442–462, 2001.

[48]   E. Burman, 'Ghost penalty,' *Comptes Rendus Mathematique*, vol. 348, no. 21-22, pp. 1217–1220, 2010.

[49]   W. M. Coombs, 'Ghost stabilisation of the material point method for stable quasi-static and dynamic analysis of large deformation problems,' *International Journal for Numerical Methods in Engineering*, pp. 1–35, 2023.

[50]   D. Sulsky, Z. Chen and H. Schreyer, 'A particle method for history-dependent materials,' *Computer Methods in Applied Mechanics and Engineering*, vol. 118, no. 1-2, pp. 179–196, 1994.

[51]   D. Burgess, D. Sulsky and J. Brackbill, 'Mass matrix formulation of the FLIP particle-in-cell method,' *Journal of Computational Physics*, vol. 103, no. 1, pp. 1–15, 1992.

[52]   M. W. Evans, F. H. Harlow and E. Bromberg, 'The particle-in-cell method for hydro-dynamic calculations,' Los Alamos National Lab NM, Tech. Rep., 1957.

[53]   F. H. Harlow, 'Hydrodynamic problems involving large fluid distortions,' *Journal of the ACM*, vol. 4, no. 2, pp. 137–142, 1957.

[54]   F. H. Harlow, D. O. Dickman, D. E. Harris Jr and R. E. Martin, 'Two-dimensional hydrodynamic calculations,' Apr. 1959.

[55]   F. H. Harlow, 'The particle-in-cell computing method for fluid dynamics,' *Methods in Computational Physics*, vol. 3, pp. 319–343, 1964.

[56]   J. U. Brackbill and H. M. Ruppel, 'FLIP: A method for adaptively zoned, particle-in-cell calculations of fluid flows in two dimensions,' *Journal of Computational Physics*, vol. 65, no. 2, pp. 314–343, 1986.

[57]   J. U. Brackbill, D. B. Kothe and H. M. Ruppel, 'Flip: A low-dissipation, particle-in-cell method for fluid flow,' *Computer Physics Communications*, vol. 48, no. 1, pp. 25–38, 1988.

[58]   C. Jiang, C. Schroeder, A. Selle, J. Teran and A. Stomakhin, 'The affine particle-in-cell method,' *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, pp. 1–10, 2015.

[59]   C. C. Hammerquist and J. A. Nairn, 'A new method for material point method particle updates that reduces noise and enhances stability,' *Computer Methods in Applied Mechanics and Engineering*, vol. 318, pp. 724–738, 2017.

[60]   C. Fu, Q. Guo, T. Gast, C. Jiang and J. Teran, 'A polynomial particle-in-cell method,' *ACM Transactions on Graphics (TOG)*, vol. 36, no. 6, pp. 1–12, 2017.

[61]   W. M. Coombs and C. E. Augarde, 'AMPLE: A material point learning environment,' *Advances in Engineering Software*, vol. 139, p. 102 748, 2020.

[62] W. M. Coombs, C. E. Augarde, A. J. Brennan, M. J. Brown, T. J. Charlton, J. A. Knappett, Y. G. Motlagh and L. Wang, 'On Lagrangian mechanics and the implicit material point method for large deformation elasto-plasticity,' *Computer Methods in Applied Mechanics and Engineering*, vol. 358, p. 112 622, 2020.

[63] A. Quarteroni, R. Sacco and F. Saleri, *Numerical mathematics*. Springer Science & Business Media, 2010, vol. 37.

[64] J. Chung and G. Hulbert, 'A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized-α method,' *Journal of Applied Mechanics*, vol. 60, no. 2, pp. 371–375, 1993.

[65] N. M. Newmark, 'A method of computation for structural dynamics,' *Journal of the Engineering Mechanics Division*, vol. 85, no. 3, pp. 67–94, 1959.

[66] S. Kontoe, L. Zdravkovic and D. M. Potts, 'An assessment of time integration schemes for dynamic geotechnical problems,' *Computers and Geotechnics*, vol. 35, no. 2, pp. 253–264, 2008.

[67] H. A. Nøst, 'Iteratively coupled implicit dynamic MPM-FVM,' 2019.

[68] M. Géradin and D. J. Rixen, *Mechanical vibrations: theory and application to structural dynamics*. John Wiley & Sons, 2014.

[69] K. Kamojjala, R. Brannon, A. Sadeghirad and J. Guilkey, 'Verification tests in solid mechanics,' *Engineering with Computers*, vol. 31, pp. 193–213, 2015.

[70] R. D. Krieg and D. Krieg, 'Accuracies of numerical solution methods for the elastic-perfectly plastic model,' 1977.

[71] F. Armero, 'Formulation and finite element implementation of a multiplicative model of coupled poro-plasticity at finite strains under fully saturated conditions,' *Computer Methods in Applied Mechanics and Engineering*, vol. 171, no. 3-4, pp. 205–241, 1999.

[72] C. J. Coetzee, 'The modelling of granular flow using the particle-in-cell method,' Ph.D. dissertation, Stellenbosch: University of Stellenbosch, 2004.

[73] J. C. Simo, 'Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory,' *Computer Methods in Applied Mechanics and Engineering*, vol. 99, no. 1, pp. 61–112, 1992.

[74] B. Engelmann and J. Hallquist, 'Nike2d user manual,' *University of California-Lawrence Livermore National Laboratory, UCRL-MA-105413*, vol. 199, 1991.

[75] M. S. Gadala, 'Recent trends in ALE formulation and its applications in solid mechanics,' *Computer Methods in Applied Mechanics and Engineering*, vol. 193, no. 39-41, pp. 4247–4275, 2004.

[76] B. Maker, R. Ferencz and H. JO, 'Nike3d user manual,' *Lawrence Livermore National Laboratory: Livermore, CA*, vol. 199, 1995.

[77] S. Fernández-Méndez and A. Huerta, 'Imposing essential boundary conditions in mesh-free methods,' *Computer Methods in Applied Mechanics and Engineering*, vol. 193, no. 12-14, pp. 1257–1275, 2004.

[78] M. Cortis, W. Coombs, C. Augarde, M. Brown, A. Brennan and S. Robinson, 'Imposition of essential boundary conditions in the material point method,' *International Journal for Numerical Methods in Engineering*, vol. 113, no. 1, pp. 130–152, 2018.

[79] C. Liu and W. Sun, 'Shift boundary material point method: An image-to-simulation workflow for solids of complex geometries undergoing large deformation,' *Computational Particle Mechanics*, vol. 7, pp. 291–308, 2020.

[80] Y. Liang, J. Given and K. Soga, 'The imposition of nonconforming Neumann boundary condition in the material point method without boundary representation,' *Computer Methods in Applied Mechanics and Engineering*, vol. 404, p. 115 785, 2023.

[81] R. K. Burla and A. V. Kumar, 'Implicit boundary method for analysis using uniform B-spline basis and structured grid,' *International journal for numerical methods in engineering*, vol. 76, no. 13, pp. 1993–2028, 2008.

[82] A. V. Kumar, S. Padmanabhan and R. Burla, 'Implicit boundary method for finite element analysis using non-conforming mesh or grid,' *International Journal for Numerical Methods in Engineering*, vol. 74, no. 9, pp. 1421–1447, 2008.

Chapter 4

Dynamics of porous materials

While the interest of the previous chapter has lied in solid materials, this chapter extends the previous theory to account for porous materials, i.e., it investigates how the presence of a fluid constituent changes the material behaviour. As a matter of fact, in the process of embedding in the seabed, the anchor encounters a material (i.e., the soil) that consists of a solid phase (the very particles that constitute the soil skeleton) and a fluid phase (the water present in the interstices among these particles).

The anchor installation process can occur in different conditions, depending on the soil's permeability and the anchor's velocity. At the extreme of these circumstances are the fully drained and undrained conditions, meaning that the water either can or cannot escape from its original volume, and exceed pore pressures either instantly dissipate or accumulate. When the installation process can be regarded as an undrained process, this can be simulated via incompressible solid dynamics. This can be representative for clays with a particularly low values of permeability, especially if the installation process is quite fast. Considering opposite values on the permeability spectrum (and/or slower anchor installation too), the embedment process can take place in fully drained conditions. Soils emblematic of such a process can be well-sorted gravels or sands. From a modelling point of view, this situation can be simulated by considering the dynamics of a compressible solid body. As just mentioned, these extreme scenarios, i.e., with extremely low and very high values of permeability, can be represented by solid dynamics, although the soil is physically composed of two phases. This type of modelling can be explained with regard to how the total stress tensor is balanced entirely by excess fluid pressures (in the case of extremely low permeability) or entirely by effective stresses and hydrostatic pressure (in the case of extremely high permeability)[†]. All intermediate situations of permeability (causing partially drained conditions), where the total stress tensor is partially balanced by fluid pressure and partially by the effective stress, require considering the soil as a material composed of a solid and a liquid phase, which is the subject of poromechanics. This study field has been

---

[†]The author recognises that the concepts of effective stress and fluid pressure have not been introduced yet, as a detailed explanation of the total stress tensor decomposition is given in Section 4.5.1. However, for the readers familiar with this basic ideas of geotechnics, this interpretation justifies from the very beginning the need of the current chapter.

described at its best in the words of one of the most significant authors in the field; quoting Coussy [1]

```
''We define poromechanics as the study of porous materials whose
mechanical behaviour is significantly influenced by the pore fluid.''
```

While the investigation of such a field study goes back to the 19-th century (see, in this regards, the historical reconstruction proposed by de Boer and Ehlers [2]), the foundations of poromechanics were laid by Biot in a series of seminal papers [3–5], extending the work carried out by Terzaghi [6, 7] (see [8] for a collection of materials in English). Since then, poromechanics has been applied in a plethora of engineering fields, which span from environmental to bioengineering, from mining and petroleum to seismic engineering (see [9–13] for examples of these applications). The reader interested in updated bibliographies on poromechanics can refer to Selvadurai and Suvorov [14] and Zhang [15] and references therein. Bowen [16], Zienkiewicz *et al.* [17], Lewis *et al.* [18], and Coussy [1] constitute well-acknowledged and accepted state-of-the art books. As for the Material Point Method (MPM) applied to porous materials, Nøst [19] and Zheng *et al.* [20, 21] provided a thorough and updated list of references, categorising the published works according to the sets of Material Points (MPs) used in the analyses and the primary equations considered (see Section 4.6 of this work for a more detailed explanation on this subject).

The outline of the current chapter describing poromechanics consists of the introduction of different ways to describe the kinematics of the two phases constituting the porous material in Section 4.1. While the previous chapter on solid dynamics has required only to consider a Lagrangian description for its kinematics, a porous material imposes a need of the Eulerian description too. Mass conservation of both of the phases in different descriptions are addressed in Section 4.2. Some of the consequences of these conservation laws are unfolded in Section 4.4, where the principles of thermodynamics applied to a porous material are detailed. These are necessary to consider thermodynamically-consistent constitutive laws governing different power-conjugates terms of the material under consideration. The statics of the porous material are given in Section 4.3, in the form of balance of rates of linear and angular momenta. Section 4.5 discusses the constitutive relationships necessary to describe a porous material, accounting for both the constituents and their connection. In particular, the introduction of an effective stress-strain relationship to conserve the mass of the solid constituent is introduced and is addressed in Section 4.5.1.1. This model constitutes a novelty compared to those previously available in the literature and is aimed at making the discretized equations more informed by its continuum physical formulation. While all of the aspects considered so far takes the continuum formulation into account, Section 4.6 introduces an MPM discretisation of the necessary equations. The different stabilisations required to run MPM simulations are outlined in the same section, as well as the employed mapping processes. Numerical analyses are considered in Section 4.7, where a detailed discussion of the results is provided together. Section 4.8 reports observations on this chapter and some plausible future work.

Due to the variety of possible assumptions which can be made with regards to porous ma-

terials, this chapters introduces progressively the assumptions necessary to develop the continuum formulation and, consequently, the MPM algorithm. This choice was driven by the intention to be as clear as possible in the evolving explanation so that the consequences of the assumptions are revealed right after their introductions.

4.1 Kinematics of the porous medium

The porous material of interest is hereafter treated as a continuous body, in which both a solid and a fluid phase are present. To define the former phase, the terms solid skeleton or solid matrix are sometimes employed, while the latter is also referred to as the fluid phase or interstitial fluid. In the context of geomechanics, if the voids of the porous body are filled by a unique liquid phase, this is referred to as *fully saturated* conditions. In such materials, it may happen that voids within the solid matrix may not be communicating with each other. In this work, when the term porosity is used, it does not take unconnected voids into account, but it is always assumed that fluid can pass from a certain portion of the porous material to another.

To deal with the porous material under consideration by means of a continuous body description meeting the conditions of interest for this work, certain hypotheses are required:

H.4.1 the fully saturated porous material is treated as superposition of two continua, i.e., the solid and the fluid; and

H.4.2 as assumed in Chapter 3 for a solid material, the porous material undergoes finite strains.

Hypothesis **H.4.1** implies that the microscopic medium is ignored*, and the analysis focuses only on a macroscopic scale at which the physical phenomena of interest can be modelled. This implies that any infinitesimal particle of the porous material contains at the same time a solid and a fluid phase.

The simultaneous adoption of assumptions **H.4.1** and **H.4.2** has essential consequences on the kinematics of the material considered (as detailed below in Sections 4.1.2 and 4.1.3). The continuous superposition of the different phases implies that the fully saturated porous material remains so during the evolution of the finite strains. However, this raises questions on which phase (solid or fluid) should be kept track of (see Figure 4.1) and how to describe it (in a Lagrangian or an Eulerian approach). As one of the most significant advantages of the MPM is the use of a Lagrangian approach for the (hydro)-mechanical description, tracking in a Lagrangian sense the solid phase is the ultimate desirable choice.

---

*At a microscopic scale, a distinction can be drawn between a volume occupied by the solid skeleton or the fluid phase (see, for instance, Coussy *et al.* [22]).

4.1.1 Lagrangian and Eulerian descriptions

In the previous chapter on the mechanics of solids there was no need to distinguish between a Lagrangian and an Eulerian description, since solid mechanics is mostly described in the former of the two ways. However, as motivated below, an Eulerian description will be unavoidable at several points in the following discussion of the chapter. To avoid any confusion, it is for the best to define the two terms, since both descriptions will be helpful and used according to need and convenience. As stated by Cengel and Cimbala [23] for the description of fluid, a Lagrangian description

> ``... requires us to track the position and velocity of each
> individual fluid parcel, which we refer to as a fluid particle, and
> take to be a parcel of fixed identity.''

Quoting the same authors, in an Eulerian description

> ``... of fluid flow, a finite volume called a flow domain or control
> volume is defined, through which fluid flows in and out. Instead of
> tracking individual fluid particles, we define field variables,
> functions of space and time, within the control volume.''

Details on the Lagrangian descriptions are given in Sections 4.1.2 and 4.1.3, while Eulerian representation is discussed in Section 4.1.4. The relationships between these descriptions are detailed in Section 4.1.5.

As demonstrated in the previous chapter, the distinction between Total Lagrangian (TL) and Updated Lagrangian (UL) formulations is superficial. Hence, in this chapter, passing from one formulation to the other is performed without further clarification. Ultimately, the UL formulation is used in the current chapter for the same reasons provided in Section 3.6.2.2.

4.1.2 Lagrangian solid phase kinematics

To proceed with a Lagrangian description of kinematics relative to the solid constituent, let the porous deformable body $\mathcal{B}$ occupy an initial volume Ω (at the time $t = 0$) of the Euclidean space $\mathcal{E}$. Let the original boundary of this body be denoted by Γ. According to hypothesis **H.4.1**, an infinitesimal initial volume of the porous material $d\Omega$ is given by the union of an infinitesimal solid part and and infinitesimal fluid part, i.e., $d\Omega = d\Omega^{(sk)} \cup d\Omega^{(f)}$.

Let us now introduce the current configuration (at the generic time t) of the same body $\mathcal{B}$, which now occupies a volume ω whose boundaries are γ. As illustrated by Figure 4.1, phases occupying fractions of the same initial volume $d\Omega$ evolve differently over time. Focusing only on the solid fraction, let the solid skeleton $d\Omega^{(sk)}$ partially occupy a position (in the same initial configuration) defined by $\boldsymbol{X}$. Let the same solid phase partially occupy a position in the current configuration indicated by $\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X}, t)$, where the *motion of the solid phase is defined* via the mapping $\boldsymbol{\varphi} : \Omega \times [0, T] \rightarrow \mathcal{E}$ (see Figure 4.2 for an illustration of the mappings and the different involved quantities).
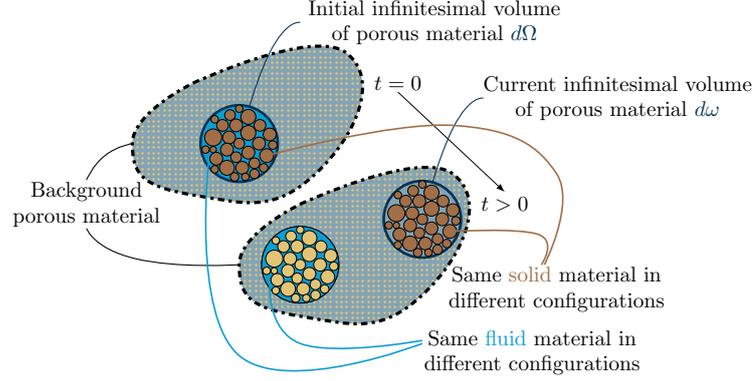
Figure 4.1: Time evolution of initial porous volume $d\Omega$ and of its constituents. For the sake of simplicity, the infinitesimal volume taken into account do not present any deformation but only translates over time.

Being the motion of the solid phase equal in all aspects to the same mapping introduced in the Chapter 3 for the solid kinematics, all the measures of rotations, deformations and strains introduced in that chapter are valid here to describe the Lagrangian kinematics of the solid phase. Let us recall the *deformation gradient* relative to the solid phase be defined as $\boldsymbol{F} := \frac{\partial\boldsymbol{\varphi}(\boldsymbol{X},t)}{\partial\boldsymbol{X}}$. It must be noted that the determinant of this deformation gradient ($J := \det\boldsymbol{F}$), while following the motion of the solid phase, must describe, according to assumption **H.4.1**, a whole porous infinitesimal volume, i.e., $d\omega = J\,d\Omega$, with $d\omega = d\omega^{(sk)} \cup d\omega^{(f)}$. Let us also list some helpful relationships relative to the transformations between current and initial position of the solid phase, these being:

- Nanson's formula (whose calculation follows the same steps given in Eqs. (3.10) and (3.11)), to describe a change in the normal attached to a infinitesimal area:

$$\boldsymbol{n}\,da = J\boldsymbol{F}^{-1}\boldsymbol{N}dA, \tag{4.1}$$

  with $\boldsymbol{N}$ (or $\boldsymbol{n}$) being the original (or current) normal outward vector, and dA (or da) the original (or current) infinitesimal mixture area;

- if $\boldsymbol{V}$ is a generic vector defined in the original configuration and $\boldsymbol{v}$ its current counterpart, related in this fashion $\boldsymbol{v}\cdot\boldsymbol{n}\,da = \boldsymbol{V}\cdot\boldsymbol{N}\,dA$, then, the following equation holds

$$V_i = JF_{ij}^{-1}v_j; \tag{4.2}$$

- from Eq. (4.2), the so-called *Piola transformation* (see, for instance, Marsden and Hughes [24]) follows, this being

$$\frac{\partial V_i}{\partial X_i} = J\frac{\partial v_i}{\partial x_i}. \tag{4.3}$$

To complete the set of Lagrangian kinematic variables, let the solid phase displacement be given by

$$\boldsymbol{u}\left(\boldsymbol{X},t\right) := \boldsymbol{\varphi}\left(\boldsymbol{X},t\right) - \boldsymbol{X}, \tag{4.4}$$
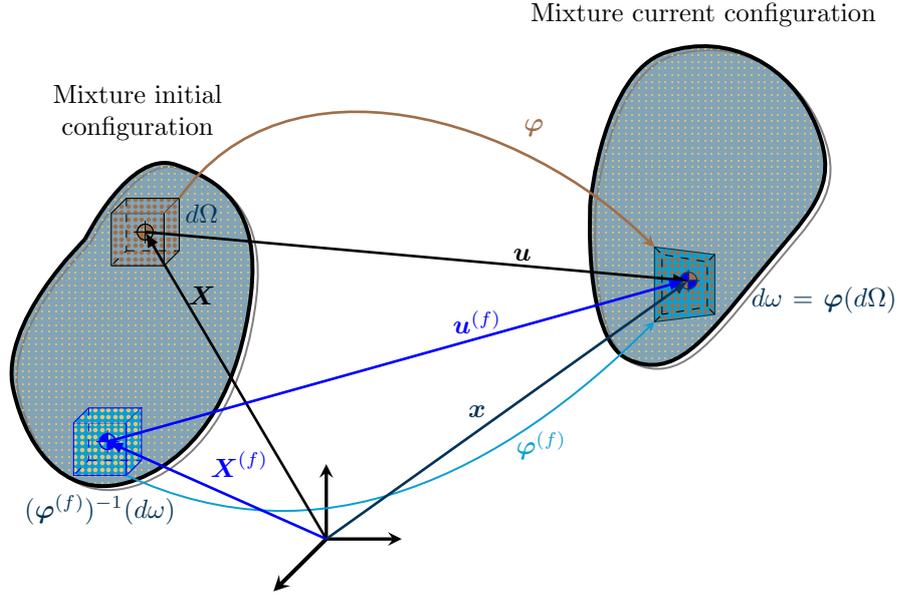
Figure 4.2: Illustration of mixture configurations and relative Lagrangian mappings.

and its velocity being

$$\boldsymbol{v}\left(\boldsymbol{X}, t\right) := \frac{d\boldsymbol{u}\left(\boldsymbol{X}, t\right)}{dt}\bigg|_X . \tag{4.5}$$

The acceleration of the solid phase particle is defined as the material derivative with respect to the solid phase of the solid velocity, i.e.,

$$\boldsymbol{a}\left(\boldsymbol{X}, t\right) := \frac{d\boldsymbol{v}\left(\boldsymbol{X}, t\right)}{dt}\bigg|_X = \frac{d^2\boldsymbol{u}\left(\boldsymbol{X}, t\right)}{dt^2}\bigg|_X . \tag{4.6}$$

As it can be deduced from the above equations, the notation for the material derivatives with respect to the solid phase is given by $\frac{d(\bullet)}{dt}\big|_X$ for the general quantity $(\bullet)$.

4.1.3   Lagrangian fluid phase kinematics

While the previous section introduced an unambiguous way to determine how the solid phase evolves, nothing has been specified about the evolution of the liquid phase. In Section 4.1.2, two fluid volumes have been introduced, namely an initial volume $d\Omega^{(f)}$ and a current volume $d\omega^{(f)}$. In the most general case, the fluid contained in these volumes does not match, since the volumes $d\Omega^{(f)}$ and $d\omega^{(f)}$ have been defined as fractions of the mixture volumes $d\Omega$ and $d\omega$, these being defined via the mapping of the solid phase $\boldsymbol{\varphi}\left(\boldsymbol{X}, t\right)$.

Owing to this difference in the considered fluid parts, it is legitimate to wonder which fluid volumes to keep track of: on the one hand, as in the case of the solid constituent, the evolution of the fluid occupying the initial volume $d\Omega^{(f)}$ could be retained; on the other hand, the fluid occupying the current volume $d\omega^{(f)}$ could be retroactively mapped to its initial configuration. Since an Eulerian description (especially of the fluid phase) will be required (see Section 4.1.4) and such description makes no reference to an initial configuration, the latter volume $d\omega^{(f)}$ is considered.

Thus, let this current fraction of fluid phase $d\omega^{(f)}$ be related to the initial configuration of the whole porous material $d\Omega$ via the *Lagrangian porosity* $\phi\left(\boldsymbol{X}, t\right)$ such that $d\omega^{(f)} = \phi\, d\Omega$, with $\phi > 0$. Let also the initial value of the Lagrangian porosity be denoted as $\phi_0 := \phi\left(\boldsymbol{X}, 0\right)$. However, it can be seen that is convenient to relate the current volume of fluid phase $d\omega^{(f)}$ to the current porous material volume $d\omega$. In particular, this relationship is obtained by introducing the *Eulerian porosity* $n\left(\boldsymbol{x}, t\right)$ such that $d\omega^{(f)} = n\, d\omega$. This measure of porosity is particularly attractive since its values are always bounded between zero and one, i.e., $0 < n < 1^*$. The initial value of the Eulerian porosity is given by $n_0 = n\left(\boldsymbol{x}, 0\right)$. Owing to their definitions, the two measures of porosity are related as follows

$$\phi = J\, n, \tag{4.7}$$

which, for the initial values (in the undeformed configuration), simplifies into $\phi_0 = n_0$.

It can be noted that, *e contrario*, the introduced measures of porosities allow also to write the fractions of the solid volume in the different configurations, i.e.,

$$d\Omega^{(sk)} = \left(1 - \phi_0\right) d\Omega; \tag{4.8}$$

$$d\omega^{(sk)} = \left(1 - n\right) d\omega. \tag{4.9}$$

Having decided to keep track of the current fluid volume in a Lagrangian sense, let the infinitesimal fluid phase of volume $d\omega^{(f)}$ partially occupy the position $\boldsymbol{x}$. This position is also mapped via the *motion of the fluid phase* $\boldsymbol{x} = \boldsymbol{\varphi}^{(f)}\left(\boldsymbol{X}^{(f)}, t\right)$, with $\boldsymbol{\varphi}^{(f)} : \Omega \times [0, T] \rightarrow \mathcal{E}$ (see, again, Figure 4.2). As for the solid constituent, the above mapping allows the introduction of the displacement of the fluid phase as

$$\boldsymbol{u}^{(f)}\left(\boldsymbol{X}^{(f)}, t\right) := \boldsymbol{\varphi}\left(\boldsymbol{X}^{(f)}, t\right) - \boldsymbol{X}^{(f)}, \tag{4.10}$$

and its fluid velocity as

$$\boldsymbol{v}^{(f)}\left(\boldsymbol{X}^{(f)}, t\right) := \left.\frac{d\boldsymbol{u}^{(f)}\left(\boldsymbol{X}^{(f)}, t\right)}{dt}\right|_{X^{(f)}}. \tag{4.11}$$

The acceleration of the fluid phase particle is given by the material derivative of its velocity with respect to the fluid phase, i.e.,

$$\boldsymbol{a}^{(f)}\left(\boldsymbol{X}^{(f)}, t\right) := \left.\frac{d\boldsymbol{v}^{(f)}\left(\boldsymbol{X}, t\right)}{dt}\right|_{X^{(f)}} = \left.\frac{d^2\boldsymbol{u}^{(f)}\left(\boldsymbol{X}^{(f)}, t\right)}{dt^2}\right|_{X^{(f)}}, \tag{4.12}$$

where the notation for the fluid material derivatives of the general quantity $(\bullet)$ is given by $\left.\frac{d(\bullet)}{dt}\right|_{X^{(f)}}$.

4.1.4 Eulerian kinematics

As can be deduced from Figures 4.1 and 4.2, the Lagrangian descriptions of the solid and liquid phases have been established, assuming both phases occupy the same volume $d\omega$ in

---

*According to the assumption **H.4.1**, porous materials present both a fluid and a solid component. Hence, the extreme values of Eulerian porosity ($n = 0$ corresponding to only solid material, and $n = 1$ corresponding to only fluid material) are excluded.

(a) (Updated) Lagrangian description relative to the solid phase.

(b) (Updated) Lagrangian description relative to the fluid phase.

(c) Eulerian description. Reference volume is highlighted in red.
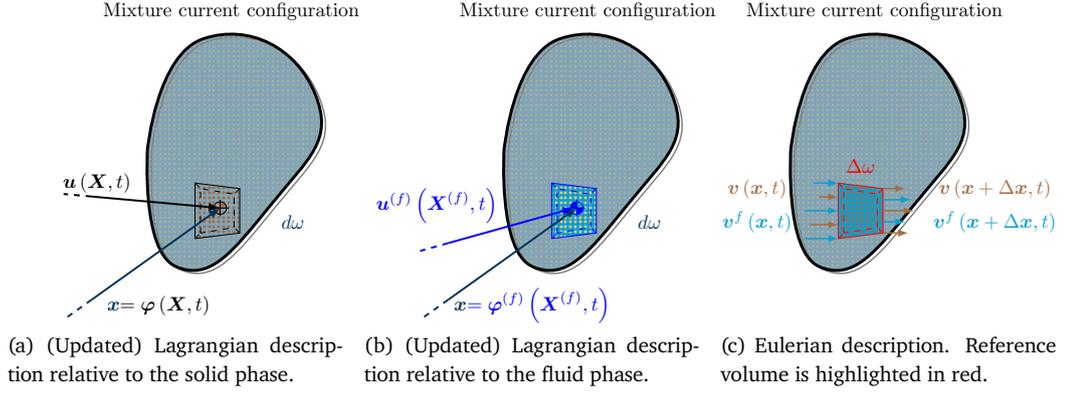
Figure 4.3: Illustration of the different descriptions used in this chapter to model the time-evolution of porous materials.

the current configuration at the time t. However, exactly as these phases come from initially different positions ($\boldsymbol{X}$ the solid phase, $\boldsymbol{X}^{(f)}$ the liquid phase), they can occupy positions that are no longer overlapping at the infinitesimal instant following the current configuration, this being $t + dt$. Thus, a Lagrangian description at instant $t + dt$ must select whether to keep track of the infinitesimal volume occupied at the time t by the solid phase $d\omega^{(sk)}$ or by the liquid phase $d\omega^{(f)}$, leading to different descriptions of the kinematics.

However, there exists a way to have these descriptions communicating. This approach consists in first describing in a Lagrangian manner and, therefore, separately, the solid and liquid volumes at the instant $t + dt$. These descriptions relate to a unique Eulerian description of the phenomenon via a finite control volume chosen to contain both of these phases. This unique Eulerian description can finally be related back to the Lagrangian description following the solid phase using a relationship inverse to the one just used (i.e., the one that correlates the solid Lagrangian description to the unique Eulerian one). Therefore, it can be understood how different descriptions and their relationships are essential requirement to complete this step. This idea dates back originally to Thorpe [25]. A graphical representation of these descriptions and their relationships is sketched in Figure 4.3.

While a Lagrangian description refers to an initial configuration, an Eulerian approach has no interest in tracking how the considered body evolves over time with respect to this original configuration, but focuses only on the current configuration at time t. This kind of description focuses on a finite mixture volume $\Delta\omega$ (the above-mentioned *control volume*). The primary quantities of interest for the Eulerian description are the velocity fields flowing through this volume at a given point $\boldsymbol{x}$, at a given time t, i.e., $\boldsymbol{v}^{(ph)}(\boldsymbol{x}, t)$ for the generic phase $(ph) = (sk)$ or $(ph) = (f)$. These velocities, although equal to the definitions introduced by the Eqs. (4.5) and (4.11) for solid and fluid particles, do not imply the knowledge of the different motions $\boldsymbol{\varphi}^{(ph)}\left(\boldsymbol{X}^{(ph)}, t\right)$.

4.1.5 Relationships between Lagrangian and Eulerian descriptions

Let us consider a generic field with Eulerian specifications $(\bullet)(\boldsymbol{x}, t)$. This same field can be given with Lagrangian specifications of the generic phase (ph), i.e., $(\bullet)\left(\boldsymbol{\varphi}^{(ph)}\left(\boldsymbol{X}^{(ph)}, t\right) t\right)$. Using the chain rule (see, for instance, Batchelor [26]), the relationship between a Lagrangian (material) and and Eulerian description descriptions is given by

$$\frac{d(\bullet)\left(\boldsymbol{\varphi}^{(ph)}\left(\boldsymbol{X}^{(ph)}, t\right) t\right)}{dt}\Bigg|_{X^{(ph)}} = \frac{\partial(\bullet)}{\partial\boldsymbol{x}}\underbrace{\frac{\partial\boldsymbol{x}}{\partial t}}_{:=\boldsymbol{v}^{(ph)}} + \frac{\partial(\bullet)}{\partial t}, \tag{4.13}$$

where the notation $\frac{\partial(\bullet)}{\partial t}$ indicates specifically the partial derivative of $(\bullet)$ with respect to time (i.e., the position $\boldsymbol{x}$ is kept fixed), opposed to the material derivatives $\frac{d(\bullet)}{dt}\big|_{X^{(ph)}}$.

Derivatives of the current volume   While the above relationship connects point-wisely the Lagrangian and the Eulerian description, it would be also useful to express the same relationship in an integral sense. For this scope, it is required to understand how the material derivatives of the current mixture volume $d\omega$ can be computed when following the solid or fluid phase. This can be achieved considering a line element $d\boldsymbol{x}$. The material velocity of this element (following the phase (ph)) at any given time can be related via

$$\frac{d}{dt}\Big|_{X^{(ph)}}(d\boldsymbol{x}) = \frac{d}{dt}\Big|_{X^{(ph)}}(\boldsymbol{x} + d\boldsymbol{x} - \boldsymbol{x}) = \boldsymbol{v}^{(ph)}(\boldsymbol{x} + d\boldsymbol{x}, t) - \boldsymbol{v}^{(ph)}(\boldsymbol{x}, t)$$

$$= \boldsymbol{v}^{(ph)}(\boldsymbol{x}, t) + \left(\frac{\partial}{\partial\boldsymbol{x}}\cdot\boldsymbol{v}^{(ph)}\right)d\boldsymbol{x} - \boldsymbol{v}^{(ph)}(\boldsymbol{x}, t) = \left(\frac{\partial}{\partial\boldsymbol{x}}\cdot\boldsymbol{v}^{(ph)}\right)d\boldsymbol{x}. \tag{4.14}$$

Recalling the relationship between the current infinitesimal volume and infinitesimal vectors describing it, i.e., $d\omega = \left(d\boldsymbol{x}^{(1)}\times d\boldsymbol{x}^{(2)}\right)\cdot d\boldsymbol{x}^{(3)}$, the above equation permits writing what follows (in a similar fashion to Eq. (3.35))

$$\frac{dJ}{dt}\Big|_{X^{(ph)}} d\Omega = \frac{\partial}{\partial\boldsymbol{x}}\cdot\boldsymbol{v}^{(ph)}d\omega. \tag{4.15}$$

Reynolds transportation theorem   The introduction of Eqs. (4.13) and (4.15), allows to relate the material time derivatives of integrated quantities to an integral Eulerian description. This procedure is named the *Reynolds transportation theorem*, this being

$$\frac{d}{dt}\Big|_{X^{(ph)}}\left(\int_{\omega}(\bullet)\,dv\right) = \int_{\omega}\frac{d}{dt}\Big|_{X^{(ph)}}((\bullet)\,dv)$$

$$= \int_{\omega}\left(\frac{\partial(\bullet)}{\partial\boldsymbol{x}}\cdot\boldsymbol{v}^{(ph)} + \frac{\partial(\bullet)}{\partial t}\right)dv + \int_{\omega}\left((\bullet)\,\frac{\partial}{\partial\boldsymbol{x}}\cdot\boldsymbol{v}^{(ph)}\right)dv$$

$$= \int_{\omega}\left(\frac{\partial(\bullet)}{\partial t} + \frac{\partial}{\partial\boldsymbol{x}}\cdot\left((\bullet)\,\boldsymbol{v}^{(ph)}\right)\right)dv$$

$$= \int_{\omega}\frac{\partial(\bullet)}{\partial t}dv + \int_{\gamma}\left((\bullet)\,\boldsymbol{v}^{(ph)}\right)\cdot\boldsymbol{n}\,da, \tag{4.16}$$

with $\boldsymbol{n}$ being the outward normal to the surface γ. To obtain the above formula, the divergence theorem has also been applied. The first term on the last Right-Hand Side (RHS)

of Eq. (4.16) is the time rate of change of the ($\bullet$) within the control volume ω (fixed, as assumed in Eulerian descriptions), while the second term represents the flux of ($\bullet$) flowing through the surface γ of the control volume ω.

4.2 Mass conservation

Up to this point, several relationships involving the kinematic behaviour of the porous material have been introduced and detailed. However, to proceed in considering the mass balance equations, two further hypotheses are necessary:

H.4.3 the postulate of conservation of mass is valid for the porous material. Moreover, the fluid and solid phases do not exchange mass; and

H.4.4 the porosity network across the porous medium is considered to be homogeneous from a geometrical point of view.

While the former of the above assumptions implies that the mass conservation can be written separately on the two phases, assumption **H.4.4** allows the treatment of the hydraulic phenomena in the porous material in an unified manner. Physically speaking, this latter hypothesis excludes those natural materials (such as rocks, for instance) whose porosity can vary in diameter and shape (rounded or flat pores) across the sample.

Let $\rho^{(sk)}(\boldsymbol{x}, t)$ and $\rho^{(f)}(\boldsymbol{x}, t)$ be the current mesoscopic (or intrisic) densities of the solid and fluid phases, so that $\rho^{(f)} n \, d\omega$ and $\rho^{(sk)}(1 - n) \, d\omega$ are the fluid and solid mass contained in the current volume $d\omega$. According to the notation used so far, let the initial values of these densities being denoted as $\rho_0^{(ph)} := \rho^{(ph)}(\boldsymbol{x}, 0)$.

4.2.1 Continuity equation

According to hypothesis **H.4.3**, the mass balance can be written separately on each phases in the current configuration as follows

$$\frac{d}{dt}\Big|_X \left(\int_\omega \rho^{(sk)}(1 - n) \, dv \right) = 0; \tag{4.17}$$

$$\frac{d}{dt}\Big|_{X^{(f)}} \left(\int_\omega \rho^{(f)} n \, dv \right) = 0. \tag{4.18}$$

Having made no assumption on the current volume ω, Eq. (4.17) can also be expressed locally as

$$\frac{d}{dt}\Big|_X \left(\rho^{(sk)}(1 - n) \, dv \right) = 0, \tag{4.19}$$

which implies

$$n = 1 - \frac{1}{J}(1 - n_0)\frac{\rho_0^{(sk)}}{\rho^{(sk)}}. \tag{4.20}$$

This equation provides a kinematic relationship between the Eulerian porosity and the Jacobian, which constitutes the basis for some detailed discussions in Section 4.5.1.1.

The Eulerian counterpart of Eq. (4.18) can be expressed using Eq. (4.13) and considering that the control volume is generic. Thus, the mass balance also holds in a point-wise sense as follows

$$\frac{\partial}{\partial t}\left(\rho^{(f)}n\right) + \frac{\partial}{\partial \boldsymbol{x}}\cdot\left(\rho^{(f)}n\boldsymbol{v}^{(f)}\right) = 0. \tag{4.21}$$

Following the procedure described in Section 4.1.4, the mass conservation equation for the fluid phase in the above Eulerian form can be expressed more concisely in a Lagrangian description of the solid skeleton by introducing the *relative flow of fluid mass vector* $\boldsymbol{q}^{(f)}$

$$\boldsymbol{q}^{(f)} := n\,\rho^{(f)}\left(\boldsymbol{v}^{(f)} - \boldsymbol{v}\right), \tag{4.22}$$

and the *fluid mass content* $m^{(f)}$ per unit of initial mixture volume $d\Omega$

$$m^{(f)} := \rho^{(f)}\,\phi. \tag{4.23}$$

Using the above-introduced definitions and Eqs. (4.13) and (4.15), the Eulerian fluid mass conservation Eq. (4.21) can be expressed in the Lagrangian solid description as follows

$$\frac{\partial}{\partial t}\left(\rho^{(f)}n\right) + \frac{\partial}{\partial \boldsymbol{x}}\cdot\left(\rho^{(f)}n\boldsymbol{v}^{(f)}\right) \pm \frac{\partial}{\partial \boldsymbol{x}}\cdot\left(\rho^{(f)}n\boldsymbol{v}\right) = 0;$$

$$\frac{d}{dt}\Big|_X\left(\rho^{(f)}\,n\right) + \frac{\partial}{\partial \boldsymbol{x}}\cdot\boldsymbol{q}^{(f)} + \rho^{(f)}n\frac{\partial}{\partial \boldsymbol{x}}\cdot\boldsymbol{v} = 0;$$

$$\frac{d}{dt}\Big|_X\left(m^{(f)}\right) + \frac{\partial}{\partial \boldsymbol{X}}\cdot\boldsymbol{Q}^{(f)} = 0, \tag{4.24}$$

where $\boldsymbol{Q}^{(f)}$ is the material counterpart of $\boldsymbol{q}^f$, which follows the Piola identity (4.3) and it is defined as

$$\boldsymbol{Q}^{(f)} := \phi\,\rho^{(f)}\boldsymbol{F}^{-1}\left(\boldsymbol{v}^{(f)} - \boldsymbol{v}\right). \tag{4.25}$$

Eq. (4.24) thus describes the fluid mass (per unit original volume $d\Omega$) flowing through the porous particle described in a Lagrangian frame following the solid phase.

4.3 Momentum balance

While previous sections have focused on the kinematics of deformable porous materials and the conservation of mass of both their phases, this one focuses on discussing the forces that cause the motion (and deformations) and how applied forces are distributed in the porous body.

Similarly to the deformable solid case, to fully characterise which and how external loads distribute in the porous material under consideration, further assumptions are required:

H.4.5 the non-polar case is considered, i.e., a *Cauchy continuum*, where only two types of external forces can be applied to porous bodies, namely, the body forces $\rho\boldsymbol{f}$ and the boundary forces $\boldsymbol{t}$, such that

$$d\boldsymbol{f} = \rho\,\boldsymbol{f}d\omega; \tag{4.26}$$

$$d\boldsymbol{t} = \boldsymbol{t}\, da, \tag{4.27}$$

where the (current) *density of the porous material* is defined as

$$\rho := \rho^{(sk)} \left(1 - n\right) + \rho^{(f)} n; \tag{4.28}$$

H.4.6 *Cauchy's axiom* holds, i.e., implying that contact forces at a given position $\boldsymbol{x}$ at a given time t are only a function of the outward normal $\boldsymbol{n}$ to the surface da, i.e.,

$$\boldsymbol{t} = \boldsymbol{t}\left(\boldsymbol{x}, t, \boldsymbol{n}\right); \tag{4.29}$$

H.4.7 Euler's laws, i.e., the balance of rate of linear and angular momenta are valid on the mixture material. When the forces described in **H.4.5** are applied on the porous material considered so far, these laws become

$$\frac{d}{dt}\Big|_{X} \left(\int_{\omega} \rho^{(sk)} \left(1 - n\right) \boldsymbol{v}\, dv \right) + \frac{d}{dt}\Big|_{X^{(f)}} \left(\int_{\omega} \rho^{(f)} n \boldsymbol{v}^{(f)}\, dv \right) = \int_{\omega} \rho\, \boldsymbol{f}\, dv + \int_{\gamma} \boldsymbol{t}\, da, \tag{4.30}$$

$$\frac{d}{dt}\Big|_{X} \left(\int_{\omega} \boldsymbol{x} \times \rho^{(sk)} \left(1 - n\right) \boldsymbol{v}\, dv \right) + \frac{d}{dt}\Big|_{X^{(f)}} \left(\int_{\omega} \boldsymbol{x} \times \rho^{(f)} n \boldsymbol{v}^{(f)}\, dv \right)$$
$$= \int_{\omega} \boldsymbol{x} \times \rho\, \boldsymbol{f}\, dv + \int_{\gamma} \boldsymbol{x} \times \boldsymbol{t}\, da; \tag{4.31}$$

and

H.4.8 viscous effects are not considered for both the constituents.

Owing to the above hypotheses, the *Cauchy stress theorem* holds, i.e.,

$$\boldsymbol{\sigma}\, \boldsymbol{n} = \boldsymbol{t}, \qquad \text{with} \quad \boldsymbol{\sigma} = \boldsymbol{\sigma}^{T}. \tag{4.32}$$

Cauchy's axiom and balance of rate of linear momentum implies, in turn, the action reaction principle on two infinitesimal surfaces da defined by opposite normals, i.e.,

$$\boldsymbol{t}\left(\boldsymbol{x}, t, \boldsymbol{n}\right)\, da = -\boldsymbol{t}\left(\boldsymbol{x}, t, -\boldsymbol{n}\right)\, da \tag{4.33}$$

It can be seen that none of the above assumptions considers any decomposition between the phases of the Cauchy stress, even though some information is provided by **H.4.8**. This hypothesis plays two roles: on the one hand, it allows to consider the stress tensor relative to the fluid phase to be constituted only by a spherical part (i.e., implying that a *perfect* fluid is considered); on the other hand, it excludes rate effects for the (effective) stress-strain relationship (this will be detailed in Section 4.5). From a technical point of view, this dual role would suggest decoupling the hypothesis **H.4.8** (between a part relative to the statics and a part concerning constitutive relationships). However, for the sake of simplicity, this work considers it as a single hypothesis, while recognising its double role.

4.3.1 Mixture local balance of rate of linear momentum

The material derivatives appearing in Eq. (4.30) can be moved into the integrals and simplified via the mass balances Eqs. (4.17) and (4.18). Owing to these and to the definition of accelerations for each phase, Eq. (4.30) can be rewritten as

$$\int_\omega \rho^{(sk)} (1-n) \, \boldsymbol{a} \, dv + \int_\omega \rho^{(f)} n \, \boldsymbol{a}^{(f)} \, dv = \int_\omega \rho \, \boldsymbol{f} dv + \int_\gamma \boldsymbol{t} \, da. \tag{4.34}$$

Making use of the Cauchy stress theorem (4.32) and the divergence theorem, it follows that Eq. (4.34) can be stated in a local form, i.e.,

$$\frac{\partial}{\partial x_j} \sigma_{ij} + \rho^{(sk)} (1-n) (f_i - a_i) + \rho^{(f)} n \left(f_i - a_i^{(f)} \right) = \boldsymbol{0}. \tag{4.35}$$

The above equation will constitute one of the two primary equations underlying the so-called $\boldsymbol{u} - p^f$ formulation presented in Section 4.6, which also addresses the diverse possibilities available for modelling poromechanical problems in the MPM.

The local form of Eq. (4.31) is not reported since, as already mentioned in the previous Chapter, its local form ensures only the symmetry of the Cauchy stress tensor $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$.

4.3.2 Local balances of rate of linear momentum for the phases

The balance equations introduced above have been stated for the porous material as a whole, making no distinction between what happens in the solid skeleton and what in the interstitial fluid. To do so, it is necessary to introduce an assumption about how the total Cauchy stress is distributed between the solid and fluid phases. However, there is no agreed consensus in the literature (see, for instance, Karrech *et al.* [12]) on how the total stress tensor should be partitioned. One way could be to consider the total (Cauchy) stress as a sum of the average stresses acting on the solid matrix and on the fluid. Thus, let us assume that

H.4.9 the total stress tensor can be given by a part relative to the solid skeleton $\boldsymbol{\sigma}^{(sk)}$ and a part relative to the interstitial fluid $\boldsymbol{\sigma}^{(f)*}$, i.e.,

$$\boldsymbol{\sigma} = (1-n) \, \boldsymbol{\sigma}^{(sk)} + n \, \boldsymbol{\sigma}^{(f)}. \tag{4.36}$$

In compliance with neglecting the viscous effect (assumption **H.4.8**), the fluid cannot bear any deviatoric stress, i.e.,

$$\boldsymbol{\sigma}^{(f)} = -p^{(f)} \boldsymbol{I}^{(2)}. \tag{4.37}$$

It also can be seen that Eq. (4.36) implies the additive decomposition of the tractions vector

$$\boldsymbol{t} = \boldsymbol{t}^{(sk)} + \boldsymbol{t}^{(f)} \tag{4.38}$$

via the Cauchy theorem (4.32) applied to each phase, i.e.,

$$\boldsymbol{t}^{(sk)} = (1-n) \, \boldsymbol{\sigma}^{(sk)} \, \boldsymbol{n}; \tag{4.39}$$

---

*While the decomposition given by Eq. (4.36) helps write the rates of linear momenta separately for the phases, this work adopts a different stress decomposition (see Section 4.5.1).

$$t^{(f)} = n\,\boldsymbol{\sigma}^{(f)}\,\boldsymbol{n}. \tag{4.40}$$

The decomposition (4.36) allows to write the rate of linear momentum equation on each phase separately as follows

$$\frac{\partial}{\partial x_j}\left((1-n)\,\sigma_{ij}^{(sk)}\right) + \rho^{(sk)}\,(1-n)\,(f_i - a_i) + f_i^{\rightarrow(sk)} = \mathbf{0}; \tag{4.41}$$

$$\frac{\partial}{\partial x_j}\left(n\,\sigma_{ij}^{(f)}\right) + \rho^{(f)}\,n\left(f_i - a_i^{(f)}\right) + f_i^{\rightarrow(f)} = \mathbf{0}, \tag{4.42}$$

where the forces of one phase acting on the other one obey to the action-reaction principle, i.e.,

$$\boldsymbol{f}^{\rightarrow(sk)} = -\boldsymbol{f}^{\rightarrow(f)}. \tag{4.43}$$

It can also be noticed that this principle is necessary so that the sum of Eq. (4.41) with Eq. (4.42) returns Eq. (4.35).

If Eq. (4.37) is considered, the rate of linear momentum of the fluid phases (4.42) can also be written as

$$-\frac{\partial}{\partial x_i}\left(n\,p^{(f)}\right) + \rho^{(f)}\,n\left(f_i - a_i^{(f)}\right) + f_i^{\rightarrow(f)} = \mathbf{0}. \tag{4.44}$$

4.4 Fundamental principles of thermodynamics

As for solid mechanics in Chapter 3, the thermodynamics principles allow us to understand how transformations of different forms of energy can occur in porous materials. Hence, this knowledge enables the exclusion of the constitutive relationships violating the permitted transformations.

4.4.1 First principle of thermodynamics

The first principle of thermodynamics deals with the conservation of energy. In the porous material under consideration, let K denote the kinetic energy and U the internal energy for the whole body. Let also the power supplied by the external forces (body forces and applied tractions) be P^{ext}, and the heat Q.

H.4.10 The first principle of thermodynamics states that the time variation of kinetic and internal energy must be equal to sum of power supplied by external forces and heat, i.e.,

$$\frac{d}{dt}\left(K + U\right) = P^{ext} + Q \tag{4.45}$$

In particular, for a porous material, the above quantities are given by (compare with Eqs. (3.61)-(3.64) for a solid material)

$$K = K^{(sk)} + K^{(f)} := \frac{1}{2}\int_\omega \rho^{(sk)}\,(1-n)\,\boldsymbol{v}^2\,dv + \frac{1}{2}\int_\omega \rho^{(f)}\,n\left(\boldsymbol{v}^{(f)}\right)^2\,dv; \tag{4.46}$$

$$U = U^{(sk)} + U^{(f)} := \int_{\omega} \rho^{(sk)} \, (1-n) \, \check{e}^{(sk)} \, dv + \int_{\omega} \rho^{(f)} \, n \check{e}^{(f)} \, dv; \qquad (4.47)$$

$$P^{ext} = P_f^{ext} + P_t^{ext} := \int_{\omega} \rho^{(sk)} \, (1-n) \, \boldsymbol{f} \cdot \boldsymbol{v} \, dv + \int_{\omega} \rho^{(f)} \, n \, \boldsymbol{f} \cdot \boldsymbol{v}^{(f)} dv$$
$$+ \int_{\gamma} \boldsymbol{t}^{(sk)} \cdot \boldsymbol{v} \, da + \int_{\gamma} \boldsymbol{t}^{(f)} \cdot \boldsymbol{v}^{(f)} da; \quad (4.48)$$

$$Q := - \int_{\gamma} \boldsymbol{q}^{(T)} \cdot \boldsymbol{n} \, da. \qquad (4.49)$$

In particular, $\check{e}^{(ph)}$ represents the density of internal energy per unit current mass of the considered phase, while $\boldsymbol{q}^{(T)}$ is the outward heat flux through the domain boundary. In the definition of the thermal power externally supplied, the volume sources of heat have been neglected.

Owing to mass conservation (Eqs. (4.17) and (4.18)) and using the definitions of acceleration for each phase (Eqs. (4.6) and (4.12)), the time derivatives of the kinetic energies in their respective frames can be written as

$$\left. \frac{dK^{(sk)}}{dt} \right|_X = \int_{\omega} \rho^{(sk)} \, (1-n) \, \boldsymbol{a} \cdot \boldsymbol{v} \, dv; \qquad (4.50)$$

$$\left. \frac{dK^{(f)}}{dt} \right|_{X^{(f)}} = \int_{\omega} \rho^{(f)} \, n \, \boldsymbol{a}^{(f)} \cdot \boldsymbol{v}^{(f)} \, dv. \qquad (4.51)$$

On the other hand, the power supplied by the applied tractions becomes (via the divergence theorem and the Fundamental Theorem of Calculus)

$$P_t^{ext} = \int_{\gamma} (1-n) \left(\boldsymbol{\sigma}^{(sk)} \, \boldsymbol{n} \right) \cdot \boldsymbol{v} \, da + \int_{\gamma} n \left(\boldsymbol{\sigma}^{(f)} \, \boldsymbol{n} \right) \cdot \boldsymbol{v}^{(f)} da$$
$$= \int_{\omega} \frac{\partial}{\partial \boldsymbol{x}} \cdot \left((1-n) \, \boldsymbol{\sigma}^{(sk)} \, \boldsymbol{v} \right) dv + \int_{\omega} \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(n \, \boldsymbol{\sigma}^{(f)} \, \boldsymbol{v}^{(f)} \right) dv$$
$$= \int_{\omega} \left(\frac{\partial}{\partial \boldsymbol{x}} \left((1-n) \, \boldsymbol{\sigma}^{(sk)} \right) \right) \cdot \boldsymbol{v} \, dv + \int_{\omega} (1-n) \, \boldsymbol{\sigma}^{(sk)} : \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} \, dv$$
$$+ \int_{\omega} \left(\frac{\partial}{\partial \boldsymbol{x}} \left(n \, \boldsymbol{\sigma}^{(f)} \right) \right) \cdot \boldsymbol{v}^{(f)} dv + \int_{\omega} n \, \boldsymbol{\sigma}^{(f)} : \frac{\partial \boldsymbol{v}^{(f)}}{\partial \boldsymbol{x}} dv. \qquad (4.52)$$

In the above equation, the velocity gradient of each phase has been concisely denoted (as in Eq. (3.31) for a solid material) as $l_{ij}^{(ph)} := \frac{\partial v_i^{(ph)}}{\partial x_j}$. Using the symmetry of the Cauchy stress tensor, it also follows that the stress power can be written as $\boldsymbol{\sigma}^{(ph)} : \boldsymbol{l}^{(ph)} = \boldsymbol{\sigma}^{(ph)} : \boldsymbol{d}^{(ph)}$, with $\boldsymbol{d}^{(ph)} := \frac{1}{2} \left(\boldsymbol{l}^{(ph)} + \left(\boldsymbol{l}^{(ph)} \right)^T \right)$.

Using balance of rate of linear momentum on each phase (Eqs. (4.41) and (4.42)) the power supplied by the external forces Eq. (4.52) minus the time derivatives of the kinetic energies Eqs. (4.50) and (4.51), becomes

$$P^{ext} - \frac{dK}{dt} = \int_{\omega} \left((1-n) \, \boldsymbol{\sigma}^{(sk)} : \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} + n \, \boldsymbol{\sigma}^{(f)} : \frac{\partial \boldsymbol{v}^{(f)}}{\partial \boldsymbol{x}} + \boldsymbol{f}^{\rightarrow (f)} \cdot \left(\boldsymbol{v} - \boldsymbol{v}^{(f)} \right) \right) dv$$
$$= \int_{\omega} \left((1-n) \, \boldsymbol{\sigma}^{(sk)} : \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}} + n \, \boldsymbol{\sigma}^{(f)} : \frac{\partial \boldsymbol{v}^{(f)}}{\partial \boldsymbol{x}} \right.$$

$$+ \left(\frac{\partial}{\partial \boldsymbol{x}} \left(n\, p^{(f)} \boldsymbol{I}^{(2)} \right) - \rho^{(f)}\, n \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right) \right) \cdot \left(\boldsymbol{v} - \boldsymbol{v}^{(f)} \right) \right) dv$$

$$= \int_{\omega} \left(\underbrace{\left((1-n)\, \boldsymbol{\sigma}^{(sk)} - n\, p^{(f)} \right)}_{=\boldsymbol{\sigma}} : \boldsymbol{d} - \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\frac{p^{(f)}}{\rho^{(f)}} \boldsymbol{q}^{(f)} \right) + \boldsymbol{q}^{(f)} \cdot \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right) \right) dv,$$

$$(4.53)$$

where also Eq. (4.44) has been used to express the force $\boldsymbol{f}^{\rightarrow (f)}$ which the solid phase applied to the fluid one.

To proceed in the process of re-writing the first principle of thermodynamics for porous materials, it is convenient to introduce the density of internal energy per unit mixture current volume e as follows

$$e := \rho^{(sk)} (1-n)\, \check{e}^{(sk)} + \rho^{(f)}\, n\, \check{e}^{(f)}. \tag{4.54}$$

In this fashion, the time derivatives of the internal energies becomes

$$\frac{dU}{dt} = \frac{d}{dt} \Big|_X \int_{\omega} \rho^{(sk)} (1-n)\, \check{e}^{(sk)}\, dv + \frac{d}{dt} \Big|_{X^{(f)}} \int_{\omega} \rho^{(f)}\, n \check{e}^{(f)}\, dv$$

$$= \underbrace{\int_{\omega} \frac{d}{dt} \Big|_X \left(\rho^{(sk)} (1-n) \right) \check{e}^{(sk)}\, dv + \int_{\Omega} \rho^{(sk)} (1-n)\, \check{e}^{(sk)} \frac{dJ}{dt} \Big|_X dV}_{=0}$$

$$+ \int_{\omega} \rho^{(sk)} (1-n) \frac{d\check{e}^{(sk)}}{dt} \Big|_X dv + \int_{\omega} \rho^{(f)}\, n \frac{d\check{e}^{(f)}}{dt} \Big|_{X^{(f)}} dv$$

$$+ \int_{\Omega} \check{e}^{(f)} \underbrace{\frac{d\rho^{(f)}\, n\, J}{dt} \Big|_{X^{(f)}}}_{=0} dv, \tag{4.55}$$

where the above-introduced zero quantities are exploited to properly develop the calculations. As qualitatively described in Section 4.1.4, the time derivative of the fluid density of internal energy with respect to the fluid Lagrangian description can be moved to the solid Lagrangian description using twice Eq. (4.13), i.e.,

$$\frac{d\check{e}^{(f)}}{dt} \Big|_{X^{(f)}} = \frac{\partial \check{e}^{(f)}}{\partial t} + \frac{\partial \check{e}^{(f)}}{\partial \boldsymbol{x}} \cdot \boldsymbol{v}^{(f)} = \frac{d\check{e}^{(f)}}{dt} \Big|_X - \frac{\partial \check{e}^{(f)}}{\partial \boldsymbol{x}} \cdot \boldsymbol{v} + \frac{\partial \check{e}^{(f)}}{\partial \boldsymbol{x}} \cdot \boldsymbol{v}^{(f)}. \tag{4.56}$$

Fluid mass conservation, which is part of the null last term in Eq. (4.55), can be also expressed in the solid reference frame by making use of Eq. (4.24) . Thus, if the volume material time derivative (Eq. (4.15)) is used with Eq. (4.56), then Eq. (4.55) becomes

$$\frac{dU}{dt} = \frac{d}{dt} \Big|_X \int_{\omega} \rho^{(sk)} (1-n)\, \check{e}^{(sk)}\, dv + \int_{\omega} \rho^{(sk)} (1-n) \frac{d\check{e}^{(sk)}}{dt} \Big|_X dv$$

$$+ \int_{\omega} \rho^{(f)}\, n \left(\frac{d\check{e}^{(f)}}{dt} \Big|_X - \frac{\partial \check{e}^{(f)}}{\partial \boldsymbol{x}} \cdot \boldsymbol{v} + \frac{\partial \check{e}^{(f)}}{\partial \boldsymbol{x}} \cdot \boldsymbol{v}^{(f)} \right) dv$$

$$+ \int_{\omega} \rho^{(sk)} (1-n)\, \check{e}^{(sk)} \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{v}\, dv$$

$$+ \int_{\omega} \check{e}^{(f)} \left(\frac{d}{dt} \Big|_X \left(\rho^{(f)}\, n \right) + \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{q}^{(f)} + \rho^{(f)}\, n \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{v} \right) dv$$

$$= \int_{\omega} \left(\frac{de}{dt} \Big|_X + e \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{v} + \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\check{e}^{(f)} \boldsymbol{q}^{(f)} \right) \right) dv \tag{4.57}$$

Given all the introduced Eqs. (4.49), (4.53), and (4.57), the first principle of thermodynamics generically expressed by Eq. (4.45) can particularised to a porous material as follows

$$
\int_{\omega} \left(\frac{de}{dt} \Big|_X + e \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{v} + \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\check{e}^{(f)} \boldsymbol{q}^{(f)} \right) \right) dv
$$
$$
= \int_{\omega} \left(\boldsymbol{\sigma} : \boldsymbol{d} - \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\frac{p^{(f)}}{\rho^{(f)}} \boldsymbol{q}^{(f)} \right) + \boldsymbol{q}^{(f)} \cdot \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right) \right) dv - \int_{\gamma} \boldsymbol{q}^{(T)} \cdot \boldsymbol{n} \, da, \quad (4.58)
$$

or, since no assumption has been made on the considered volume, this equation holds locally, i.e.,

$$
\frac{de}{dt} \Big|_X + e \frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{v}
$$
$$
= \boldsymbol{\sigma} : \boldsymbol{d} - \frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\left(\check{e}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}} \right) \boldsymbol{q}^{(f)} + \boldsymbol{q}^{(T)} \right) + \boldsymbol{q}^{(f)} \cdot \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right). \quad (4.59)
$$

For general purposes, it can be convenient to express the first principle of thermodynamics of a porous material in the original solid configuration. If we consider the density of energy per unit reference volume E to obeys to the law

$$
E := e \, J, \tag{4.60}
$$

then Eq. (4.59) reads

$$
\frac{dE}{dt} \Big|_X = \frac{1}{2} \boldsymbol{S} : \dot{\boldsymbol{C}} - \frac{\partial}{\partial \boldsymbol{X}} \cdot \left(\left(\check{e}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}} \right) \boldsymbol{Q}^{(f)} + \boldsymbol{Q}^{(T)} \right) + \boldsymbol{F} \boldsymbol{Q}^{(f)} \cdot \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right), \quad (4.61)
$$

with $\boldsymbol{Q}^{(T)}$ being the outward heat flux in the original configuration, which transforms according to Eq. (4.2).

4.4.2 Second principle of thermodynamics

Depending on the different statements, the second principle of thermodynamics deals with irreversibility of some changes, or, from a different perspective, with the work which can be extracted from a certain system.

H.4.11 For a porous material, the mathematical statement of the second principle of thermodynamics is given by

$$
\frac{d}{dt} \Big|_X \int_{\omega} \rho^{(sk)} (1 - n) \, \check{s}^{(sk)} \, dv + \frac{d}{dt} \Big|_{X^{(f)}} \int_{\omega} \rho^{(f)} n \check{s}^{(f)} \, dv \geq - \int_{\gamma} \frac{1}{T} \left(\boldsymbol{q}^{(T)} \cdot \boldsymbol{n} \right) da, \tag{4.62}
$$

with $\check{s}^{(ph)}$ being the specific entropy per unit current mass of the specific phase and T the absolute temperature.

Let the entropy per unit mixture volume as

$$
s := \rho^{(sk)} (1 - n) \, \check{s}^{(sk)} + \rho^{(f)} n \check{s}^{(f)}; \tag{4.63}
$$
$$
S := s \, J. \tag{4.64}
$$

Having introduced the above definitions and following steps similar to those for the first principle of thermodynamics, the local form of the second principle can be stated both per unit current volume

$$\frac{ds}{dt}\Big|_X + s\frac{\partial}{\partial \boldsymbol{x}} \cdot \boldsymbol{v} \geq -\frac{\partial}{\partial \boldsymbol{x}} \cdot \left(\frac{\boldsymbol{q}^{(T)}}{T} + \check{s}^{(f)}\boldsymbol{q}^{(f)} \right),$$ (4.65)

and per unit original volume

$$\frac{dS}{dt}\Big|_X \geq -\frac{\partial}{\partial \boldsymbol{X}} \cdot \left(\frac{\boldsymbol{Q}^{(T)}}{T} + \check{s}^{(f)}\boldsymbol{Q}^{(f)} \right).$$ (4.66)

4.4.3 Fluid state variables

To deal with the thermodynamics of the fluid phase in an efficient manner, it is advantageous to express the internal energy in a fluid as a function of different state variables, i.e., of other independent variables describing the fluid state.

By definition (see, for instance, Karrech *et al.* [12]) the rate of fluid internal energy is defined as

$$d\check{e}^{(f)} := -p^{(f)} d\left(\frac{1}{\rho^{(f)}} \right) + T\, d\check{s}^{(f)},$$ (4.67)

with, clearly, $\check{e}^{(f)}\left(\frac{1}{\rho^{(f)}}, \check{s}^{(f)} \right)$. Hence, the increment of $\check{e}^{(f)}\left(\frac{1}{\rho^{(f)}}, \check{s}^{(f)} \right)$ can be written as

$$d\check{e}^{(f)}\left(\frac{1}{\rho^{(f)}}, \check{s}^{(f)} \right) = \frac{\partial \check{e}^{(f)}}{\partial \left(\frac{1}{\rho^{(f)}} \right)} d\left(\frac{1}{\rho^{(f)}} \right) + \frac{\partial \check{e}^{(f)}}{\partial \check{s}^{(f)}} d\check{s}^{(f)},$$ (4.68)

so that, from the above equations, it follows immediately that

$$p^{(f)} = -\frac{\partial \check{e}^{(f)}}{\partial \left(\frac{1}{\rho^{(f)}} \right)};$$ (4.69)

$$T = \frac{\partial \check{e}^{(f)}}{\partial \check{s}^{(f)}}.$$ (4.70)

Other forms of fluid specific energy, named *enthalpy* $\check{h}^{(f)}$, *Helmholtz free energy* $\check{\psi}^{(f)}$, and *free enthalpy* $\check{\mu}^{(f)}$ can be introduced by a series of *Legendre transformation* (see Hously and Puzrin [27] for further details). These are defined, respectively, as

$$\check{h}^{(f)}\left(p, \check{s}^{(f)} \right) = \check{e}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}};$$ (4.71)

$$\check{\psi}^{(f)}\left(\frac{1}{\rho^{(f)}}, T \right) = \check{e}^{(f)} - T\, \check{s}^{(f)};$$ (4.72)

$$\check{\mu}^{(f)}\left(p^{(f)}, T \right) = \check{\psi}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}} = \check{h}^{(f)} - T\, \check{s}^{(f)}.$$ (4.73)

Which of the quantities listed above should be taken into account depends on the context, provided that these forms of energy are all related to each other.

4.4.4 Clausius-Duhem inequality

To determine which constitutive models are thermodynamically admissible, it is necessary to plug the first principle of thermodynamics into the second principle. When thermal effects are not excluded, this inequality goes under the name of the *Clausius-Duhem* inequality, while it is called *Clausius-Planck* when the temperature is treated as constant (see assumption **H.3.11** in Chapter 3).

To proceed with this substitution, the material time derivative in the solid frame of the internal energy (per unit reference volume) E can be expressed as a function of the Helmholtz free energy function, temperature and entropy (using a definition similar to Eq. (4.72) extended to the whole porous material). This gives

$$\frac{dE}{dt}\Big|_X = \frac{d\Psi}{dt}\Big|_X + S\frac{dT}{dt}\Big|_X + T\frac{dS}{dt}\Big|_X. \tag{4.74}$$

The term on the Left-Hand Side (LHS) of the above equation can be computed via the first principle of thermodynamics Eq. (4.61), while the last term on the RHS can be re-written using the second principle of thermodynamics (4.66). In this way, Eq. (4.74) can be written in the following form

$$\frac{1}{2}\boldsymbol{S} : \dot{\boldsymbol{C}} - \frac{\partial}{\partial \boldsymbol{X}} \cdot \left(\left(\check{e}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}} \right) \boldsymbol{Q}^{(f)} + \boldsymbol{Q}^{(T)} \right) + \boldsymbol{F}\,\boldsymbol{Q}^{(f)} \cdot \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right) - \frac{d\Psi}{dt}\Big|_X - S\frac{dT}{dt}\Big|_X$$
$$\geq -T\frac{\partial}{\partial \boldsymbol{X}} \cdot \left(\frac{\boldsymbol{Q}^{(T)}}{T} + \check{s}^{(f)}\boldsymbol{Q}^{(f)} \right). \tag{4.75}$$

Owing to the definition of free enthalpy Eq. (4.73), the following equation holds

$$-\frac{\partial}{\partial \boldsymbol{X}} \cdot \left(\left(\check{e}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}} \right) \boldsymbol{Q}^{(f)} \right) + T\frac{\partial}{\partial \boldsymbol{X}} \cdot \left(\check{s}^{(f)}\boldsymbol{Q}^{(f)} \right) = -\boldsymbol{Q}^{(f)} \cdot \frac{\partial \check{\mu}^{(f)}}{\partial \boldsymbol{X}}\Big|_T - \check{\mu}^{(f)}\frac{\partial}{\partial \boldsymbol{X}}\Big|_T \cdot \boldsymbol{Q}^{(f)}, \tag{4.76}$$

where $\frac{\partial(\bullet)}{\partial \boldsymbol{X}}\Big|_T$ is the spatial gradient of $(\bullet)$ with respect to the solid original configuration keeping the temperature T as constant.

Let also the following definitions of *dissipations* $\mathcal{D}$ be introduced

$$\mathcal{D}^{(sk)} := \frac{1}{2}\boldsymbol{S} : \dot{\boldsymbol{C}} - \check{\mu}^{(f)}\frac{\partial}{\partial \boldsymbol{X}}\Big|_T \cdot \boldsymbol{Q}^{(f)} - \frac{d\Psi}{dt}\Big|_X - S\frac{dT}{dt}\Big|_X; \tag{4.77}$$

$$\mathcal{D}^{(f)} := \left(-\frac{\partial \check{\mu}^{(f)}}{\partial \boldsymbol{X}}\Big|_T + \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right)\boldsymbol{F} \right) \cdot \boldsymbol{Q}^{(f)}; \tag{4.78}$$

$$\mathcal{D}^{(th)} := -\frac{\boldsymbol{Q}^{(T)}}{T} \cdot \frac{\partial T}{\partial \boldsymbol{X}}. \tag{4.79}$$

Hence, (4.75) can be concisely written as

$$\mathcal{D}^{(sk)} + \mathcal{D}^{(f)} + \mathcal{D}^{(th)} \geq 0. \tag{4.80}$$

Since it has been possible to identify and separate which mechanisms underlie the different forms of dissipation, it makes sense to decouple the single Clausius-Duhem inequality into three different inequalities, i.e.,

$$\begin{cases} \mathcal{D}^{(sk)} \geq 0; \\ \mathcal{D}^{(f)} \geq 0; \\ \mathcal{D}^{(th)} \geq 0. \end{cases} \tag{4.81}$$

The sections below are based on the introduced decomposition of the dissipation into different inequalities and aim to write these dissipations so that the introduction of the necessary constitutive relations is facilitated.

4.4.4.1 Solid skeleton dissipation

Since the Helmholtz free energy in the solid skeleton dissipation in Eq. (4.77) refers to the whole porous material, it can be additively decomposed into parts relative to each phase as follows

$$\Psi = \Psi^{(sk)} + m^{(f)} \, \check{\psi}^{(f)}. \tag{4.82}$$

The same rationale can be applied to the entropy, i.e.,

$$S = S^{(sk)} + m^{(f)} \, \check{s}^{(f)}. \tag{4.83}$$

Owing to the above decompositions and making use of the fluid state variables introduced Section 4.4.3, the RHS of Eq. (4.77) can be re-written as

$$
\begin{aligned}
\mathcal{D}^{(sk)} =& \frac{1}{2} \boldsymbol{S} : \dot{\boldsymbol{C}} + \left(\check{\psi}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}} \right) \frac{dm^{(f)}}{dt} \Big|_X - \frac{d}{dt} \Big|_X \left(\Psi^{(sk)} + m^{(f)} \, \check{\psi}^{(f)} \right) \\
& - \left(S^{(sk)} + m^{(f)} \, \check{s}^{(f)} \right) \frac{dT}{dt} \Big|_X \\
=& \frac{1}{2} \boldsymbol{S} : \dot{\boldsymbol{C}} + \frac{p^{(f)}}{\rho^{(f)}} \frac{dm^{(f)}}{dt} \Big|_X - \frac{d\Psi^{(sk)}}{dt} \Big|_X - S^{(sk)} \frac{dT}{dt} \Big|_X \\
& - m^{(f)} \frac{d\check{e}^{(f)}}{dt} \Big|_X - m^{(f)} T \frac{d\check{s}^{(f)}}{dt} \Big|_X \\
=& \frac{1}{2} \boldsymbol{S} : \dot{\boldsymbol{C}} + p^{(f)} \frac{d\phi}{dt} \Big|_X - \frac{d\Psi^{(sk)}}{dt} \Big|_X - S^{(sk)} \frac{dT}{dt} \Big|_X .
\end{aligned}
\tag{4.84}
$$

It can be seen that the first term on the last RHS of the above chain of formula is the expression of (total) stress-(rate of) strain power conjugates per unit reference volume. Very similar to the case of a solid material are also the last two terms on the last RHS of Eq. (4.84) (compare with (3.77)). It is clear that the only term that differs substantially from the same principle applied to purely solid bodies is the second term, which links fluid pressure to the Lagrangian porosity variation. How this term also interrelates with the power developed by the total stress (and its strain rate) depends on the assumption of compressibility or incompressibility of the solid phase. This topic, and how to obtain a Helmholtz energy functions respectful of bounded values of the Eulerian porosity, have been extensively addressed in Nedjar [28, 29] for the compressible case. This work (based on Pretti *et al.* [30], where the reader is referred for further details) briefly describes the same topic for the incompressible case. In this work, only the second case (i.e., an incompressible solid matrix) is discussed, which (as explained below in Section 4.5.1.1) allows for an exceptionally convenient decomposition of the total stress tensor.

However, Eq. (4.84) implicitly carries information about the fluid part not only through the term $p^{(f)} \frac{d\phi}{dt} \Big|_X$. To visualise this more efficiently, a quantity giving null contribution in terms of power can be subtracted from Eq. (4.84), giving

$$\mathcal{D}^{(sk)} = \frac{1}{2} \boldsymbol{S} : \dot{\boldsymbol{C}} + p^{(f)} \frac{d\phi}{dt} \Big|_X - \frac{d\Psi^{(sk)}}{dt} \Big|_X - S^{(sk)} \frac{dT}{dt} \Big|_X$$

$$- m^{(f)} \left(\frac{d\check{\psi}^{(f)}}{dt} \Big|_X + p^{(f)} \frac{d}{dt} \Big|_X \frac{1}{\rho^{(f)}} + \check{s}^{(f)} \frac{dT}{dt} \Big|_X \right)$$

$$= \frac{1}{2} \boldsymbol{S} : \dot{\boldsymbol{C}} + p^{(f)} \frac{d\phi}{dt} \Big|_X - \frac{d\Psi^{(sk)}}{dt} \Big|_X - m^{(f)} \left(\frac{d\check{\mu}^{(f)}}{dt} \Big|_X - \frac{1}{\rho^{(f)}} \frac{dp^{(f)}}{dt} \Big|_X \right) - S \frac{dT}{dt} \Big|_X .$$

(4.85)

While the first three terms on the RHS of Eqs. (4.84) and (4.85) are equal, the mixture entropy S in Eq. (4.85) is considered in lieu of the solid one $S^{(sk)}$ in Eq. (4.84). This change allows to explicitly show an extra term (i.e., the fourth on the RHS of Eq. (4.85)). Needless to say, the use of Eq. (4.84) rather than Eq. (4.85) is merely a matter of convenience as these equation are identical.

4.4.4.2 Fluid dissipation

Before introducing a constitutive relationship for the fluid flow, it is convenient to express the dissipation of the fluid part Eq. (4.78) per unit mixed current volume, making use again of the state variables for the fluid phase (described in Section 4.4.3). Using also Eqs. (4.2) and (4.3), Eq. (4.78) becomes

$$\frac{1}{J} \mathcal{D}^{(f)} = \frac{1}{J} \left(-\frac{\partial}{\partial \boldsymbol{X}} \Big|_T \left(\check{e}^{(f)} - T \check{s}^{(f)} + \frac{p^{(f)}}{\rho^{(f)}} \right) + \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right) \boldsymbol{F} \right) \cdot \boldsymbol{Q}^{(f)}$$

$$= \left(-\frac{1}{\rho^{(f)}} \frac{\partial p^{(f)}}{\partial \boldsymbol{x}} \Big|_T + \left(\boldsymbol{f} - \boldsymbol{a}^{(f)} \right) \right) \cdot \boldsymbol{q}^{(f)}$$

(4.86)

As mentioned above for the solid skeleton, how the power conjugates appearing on the RHS of the above equations are related depends on further constitutive assumptions, given below in Section 4.5.2.

4.4.4.3 Thermal dissipation

Thermal dissipation can be also expressed per unit current mixed volume (using Eqs. (4.2) and (4.3)) as follows

$$\frac{1}{J} \mathcal{D}^{(th)} := -\frac{\boldsymbol{q}^{(T)}}{T} \cdot \frac{\partial T}{\partial \boldsymbol{x}} .$$

(4.87)

4.5 Constitutive laws

At this stage, it is not possible to further characterise the material without making further assumptions concerning how power conjugates appearing in the dissipations (4.85), (4.86) and (4.87) are related. Some general assumptions on the material behaviour are reported here*:

H.4.12 thermal effects are neglected;

---

*As in Chapter 3, the above list of assumptions follows the same rationale explained in the footnote at page 45.

H.4.13 the porous body is considered to be an *isotropic* medium;

H.4.14 the solid phase (i.e., the particles constituting the solid matrix) is considered incompressible;

H.4.15 multiplicative decomposition of the deformation gradient (proposed separately by Kröner [31], Lee [32] and Mandel [33]) into an elastic and a plastic part; and

H.4.16 the fluid flow in the porous medium can be seen as laminar.

Further assumptions of a more specific nature are given below when required by the discussion.

The reason why the discussion of heat dissipation has been particularly compact lies in assumption **H.4.12**. By this, it is still intended to recognise that irreversible dissipations occurring in both the solid matrix and the fluid phase lead to a change in temperature, but this variation is considered irrelevant for modelling purposes. As already mentioned above, when this assumption is considered, the Clausius-Duhem inequality (4.81) is replaced by the Clausius-Planck inequality, becoming

$$
\begin{cases}
\mathcal{D}^{(sk)} = \dfrac{1}{2}\boldsymbol{S}:\dot{\boldsymbol{C}} + p^{(f)}\dot{\phi} - \dot{\Psi}^{(sk)} - m^{(f)}\left(\dot{\mu}^{(f)} - \dfrac{1}{\rho^{(f)}}\dot{p}^{(f)}\right) \geq 0 & \text{(4.88a)} \\[4mm]
\mathcal{D}^{(f)} = J\left(-\dfrac{1}{\rho^{(f)}}\dfrac{\partial p^{(f)}}{\partial \boldsymbol{x}} + \left(\boldsymbol{f} - \boldsymbol{a}^{(f)}\right)\right)\cdot \boldsymbol{q}^{(f)} \geq 0; & \text{(4.88b)} \\[4mm]
\mathcal{D}^{(th)} = J\dfrac{\boldsymbol{q}^{(T)}}{T}\cdot\dfrac{\partial T}{\partial \boldsymbol{x}} = 0, & \text{(4.88c)}
\end{cases}
$$

where the compact form $(\dot{\bullet}) = \frac{d(\bullet)}{dt}\big|_X$ for the material derivative with respect to the solid phase has been used. As all equations are expressed in the solid phase Lagrangian reference frame, this notation $(\dot{\bullet})$ replaces the more cumbersome $\frac{d(\bullet)}{dt}\big|_X$ from here to the end of the current chapter.

It must be highlighted that, strictly speaking, **H.4.15** is a kinematic assumption, which has consequences on the constitutive relationship of the material under consideration. However, since the kinematics relative to the solid phase is inherited from Chapter 3 (as assumed by Eq. 3.37), it was chosen to recall this hypothesis at this stage for this chapter.

4.5.1 Effective constitutive equation and fluid constitutive equation

The content of the this section heavily draws on the discussion available in Pretti *et al.* [30]. The reader interested in deepening the discussion about how the compressibility or incompressibility (this latter assumed by **H.4.14**) of the solid phase combines with bounded values of the Eulerian porosity is referred to Pretti *et al.* [30].

Generally speaking, when considering a volume of mixed material, three mechanisms allow the volume of such a sample to vary (see, in this regard, Figure 4.4): the compressibility of the two constituents and the net flow of fluid across the volume's surfaces (this most general case is represented in Figure 4.4(b)). While varying fluid flow depends on the boundary

(a) Initial sample

(b) Option 1: Compressible solid matrix

(c) Option 2: Incompressible solid matrix
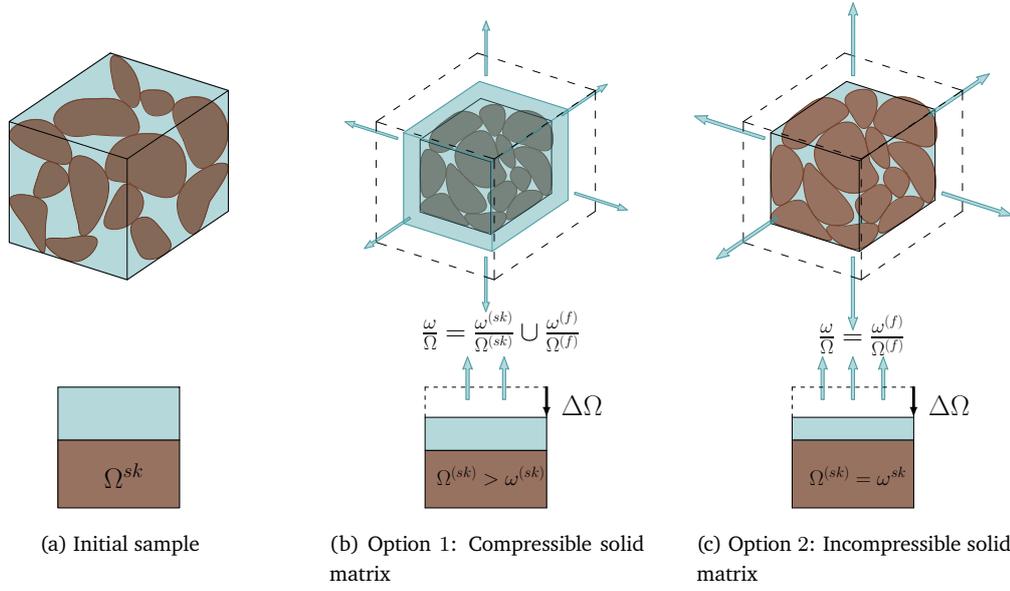
Figure 4.4: Comparison of drained volumetric compression in three dimensions with grains (top row) and idealised bi-dimensional boxes (bottom row). From the initial configuration (a), the same drained volumetric compression is applied to a compressible solid matrix sample (b) and to an incompressible solid matrix sample (c). Figure adapted from Pretti *et al.* [30].

conditions of the problem under consideration (i.e., drained or undrained cases), choosing between the compressibility or incompressibility of the phases is up to the analyst, informed by the physics of the problem. As assumed above in **H.4.14**, this work considers the solid phase as incompressible (case illustrated in Figure 4.4(c)), thus excluding one of the three possibilities mentioned above. This assumption is generally accepted in the context of soil mechanics, especially when the soils under consideration are particularly loose, and the volume change is primary due to water exchange.

If the assumption of incompressibility of the solid constituent is introduced, it can appreciated that the solid mass conservation in the form of Eq. (4.20) becomes

$$n = 1 - \frac{1}{J}\left(1 - n_0\right). \tag{4.89}$$

The above equation introduces a one-to-one kinematic relationship between the Jacobian J and the Eulerian porosity n, provided that the initial porosity n_0 is a constant value. Knowing the relationship between the Eulerian and Lagrangian porosities, and also recalling Eq. (3.30) from the previous chapter, Eq. (4.89) can also be expressed as

$$J = \exp\left(\epsilon_v\right) = 1 + \phi - \phi_0, \tag{4.90}$$

where $\phi_0 = n_0$ has also been employed. The introduction of incompressibility of the solid matrix (expressed either via Eq. (4.89) or Eq. (4.90)) has implications on the treatment of porous materials. In particular, the physical constraints on the values of the Eulerian porosity between $(0, 1)$ become, in turn, constraints on the Jacobian J (or, equally, on the

volumetric part of the logarithmic strain). Thus, inequalities

$$\begin{cases} n > 0; \\ n < 1, \end{cases} \tag{4.91}$$

become, respectively,

$$\begin{cases} \epsilon_v > \ln(1 - n_0); & \text{(4.92a)} \\ \dfrac{1}{\exp(\epsilon_v)}(1 - n_0) > 0. & \text{(4.92b)} \end{cases}$$

While the second of the above is always satisfied by $J > 0$ (see Eq. (3.9)), the same cannot be said for the first. Inequality (4.92a) expresses a constraint on the negative (in compression) volumetric part of the logarithmic strain. This inequality is inherited from the constraint $n > 0$. Physically speaking, it implies that, when the fluid constituent is progressively expelled from the considered mixture volume (i.e., the porous material becomes closer and closer to be constituted only by its solid phase), the porous material shows increasing incompressible behaviour. Another consequence of the assumed incompressibility of the solid phase is that the free energy function of the solid matrix $\Psi^{(sk)}$ cannot be a function depending separately on the deformation gradient $\boldsymbol{F}$ (or one of its related strain measures, as objectivity requires*) and the porosity, as these are not independent variables.

Furthermore, as assumed by hypothesis **H.4.15**, the multiplicative decomposition of the deformation gradient into an elastic and a plastic part implies $J = J^e J^p$. Considering the rate of Eq. (4.90), this gives (using also Eqs. (3.35) and (3.36))

$$\dot{J} = J\,\dot{\epsilon}_v = \dot{\phi} = J\,\boldsymbol{I}^{(2)} : \boldsymbol{d}, \tag{4.93}$$

i.e., a change in (infinitesimal) volume of mixed material must correspond to a change in the fluid phase, regardless of this being due to a compression of the fluid constituent or to its flow through the infinitesimal particle of volume $d\omega$. However, Eqs. (4.90) and (4.93) do not imply any one-to-one correspondence between elastic or plastic Jacobian and Lagrangian porosity counterparts, either in terms of finite quantities or in their rate forms, i.e.,

$$\phi^e \neq J^e, \qquad \phi^p \neq J^p; \tag{4.94}$$

$$\dot{\phi}^e \neq \dot{J}^e, \qquad \dot{\phi}^p \neq \dot{J}^p, \tag{4.95}$$

meaning that the porosity cannot be decomposed into an elastic and plastic part. Hence, to take the kinematic relationship Eq. (4.89) (or, equally, Eq. (4.90)) between the Jacobian and the Eulerian porosity and the impossibility of decomposing the latter into an elastic and plastic part into account, the free Helmholtz free energy function cannot depend only from $\boldsymbol{F}^e$. A possible form of this energy function respectful of the above considerations can be expressed as a function of the whole left Cauchy-Green strain tensor $\boldsymbol{b}$ (thus, assuming **H.4.13**, i.e., isotropy) as follows

$$\Psi^{(sk)} = \hat{\Psi}^{(sk)}_{inc}(\boldsymbol{b}^e, \boldsymbol{b}^p, \boldsymbol{\alpha}), \tag{4.96}$$

---

*The importance and the meaning of objectivity have already been introduced and explained in Section 3.5.1.

with $\boldsymbol{\alpha}$ being the set of internal state variables. Eq. (4.96) shows how the dependency from the whole deformation gradient $\boldsymbol{F}$ (via $\boldsymbol{b}$) is necessary to describe a material respectful of the kinematic relationship Eq. (4.89).

If the rate of Eq. (4.96), the power conjugate equation $\frac{1}{2}\boldsymbol{S} : \dot{\boldsymbol{C}} = \boldsymbol{\tau} : \boldsymbol{d}$ and the relationship Eq. (4.93) are substituted into (4.88a), it leads to

$$\mathcal{D}^{(sk)} = \boldsymbol{\tau} : \boldsymbol{d} + p^{(f)}\dot{J} - \dot{\hat{\Psi}}_{inc}^{(sk)}\left(\boldsymbol{b}^e, \boldsymbol{b}^p, \boldsymbol{\alpha}\right) - m^{(f)}\left(\dot{\breve{\mu}}^{(f)} - \frac{1}{\rho^{(f)}}\dot{p}^{(f)}\right) \geq 0. \qquad (4.97)$$

Introducing the multiplicative decomposition of the deformation gradient, allows one to re-write the above inequality as

$$\mathcal{D}^{(sk)} = \left(\boldsymbol{\tau} + J\,p^{(f)}\boldsymbol{I}^{(2)} - 2\frac{\partial\hat{\Psi}_{inc}^{sk}}{\partial\boldsymbol{b}^e}\boldsymbol{b}^e\right) : \boldsymbol{d}^e + \left(\boldsymbol{\tau} + J\,p^{(f)}\boldsymbol{I}^{(2)}\right) : \boldsymbol{d}^p - 2\frac{\partial\hat{\Psi}_{inc}^{sk}}{\partial\boldsymbol{b}^p}\dot{\boldsymbol{b}}^p$$

$$- \frac{\partial\hat{\Psi}_{inc}^{sk}}{\partial\boldsymbol{\alpha}} * \dot{\boldsymbol{\alpha}} - m^{(f)}\left(\frac{\partial\breve{\mu}^{(f)}}{\partial p^{(f)}} - \frac{1}{\rho^{(f)}}\right)\dot{p}^{(f)} \geq 0, \quad (4.98)$$

with $*$ being the appropriate product operator between the (rate of) set of internal variables $\dot{\boldsymbol{\alpha}}$ and their power conjugates $\frac{\partial\hat{\Psi}_{inc}^{sk}}{\partial\boldsymbol{\alpha}}$. In the above equation, accordingly with the assumption of neglecting the thermal effects, the free enthalpy of the fluid part $\breve{\mu}^{(f)}$ has been only expressed as a function of the fluid pressure (see Eq. (4.73)). The above equation also makes use of the relationship $\dot{\boldsymbol{b}}^e = \boldsymbol{l}^e\boldsymbol{b}^e + \boldsymbol{b}^e\left(\boldsymbol{l}^e\right)^T$ (see Eq. (3.83)), with $\boldsymbol{l}^e := \dot{\boldsymbol{F}}^e\left(\boldsymbol{F}^e\right)^{-1}$ and $\boldsymbol{d}^e := \left(\boldsymbol{l}^e + \left(\boldsymbol{l}^e\right)^T\right)/2$ (both defined in Eq. (3.43)).

The so-called Coleman-Noll procedure (Coleman and Noll [34], and Coleman and Gurtin [35]) permits the constitutive equations to be written, excluding the parts of the above inequality that do not contribute to the dissipation. This procedure is usually considered valid for any rate of measure of the elastic strain ($\boldsymbol{d}^e$, in this case) and for any rate of fluid pressure $p^{(f)}$. However, especially in regards to the former of these rates, it can be seen that the increment of the elastic strain cannot be random, but (4.92a) must be taken into account. Hence, if the Coleman-Noll procedure and (4.92a) are together applied, (4.98) becomes

$$\begin{cases} \boldsymbol{\tau}' = 2\dfrac{\partial\hat{\Psi}_{inc}^{sk}}{\partial\boldsymbol{b}^e}\boldsymbol{b}^e; & (4.99a) \\[2mm] \dfrac{\partial\breve{\mu}^{(f)}}{\partial p^{(f)}} = \dfrac{1}{\rho^{(f)}}; & (4.99b) \\[2mm] \mathcal{D}^{(sk)} = \boldsymbol{\tau}' : \boldsymbol{d}^p - 2\dfrac{\partial\hat{\Psi}_{inc}^{sk}}{\partial\boldsymbol{b}^p}\dot{\boldsymbol{b}}^p - \dfrac{\partial\hat{\Psi}_{inc}^{sk}}{\partial\boldsymbol{\alpha}} * \dot{\boldsymbol{\alpha}} \geq 0; & (4.99c) \\[2mm] \epsilon_v > \ln\left(1 - n_0\right), & \end{cases}$$

where it has been convenient to introduce the *effective stress tensor*, defined as

$$\boldsymbol{\tau}' := \boldsymbol{\tau} + J\,p^{(f)}\boldsymbol{I}^{(2)}. \qquad (4.100)$$

This decomposition of the total stress tensor into an effective part (depending only on the elastic strain) and a fluid pressure part is attributed to Terzaghi [8], and it was derived under the same incompressibility assumption by Borja and Alarcón [36] in the context of finite strain theory.

Eq. (4.99a) expresses a constitutive relationship relative to the effective part of the stress which retains multiple similarities with that computed for non-porous materials (see Eq. (3.85a)). Constitutive relationship Eq. (4.99b) can either describe a *barotropic* fluid, i.e., a fluid whose density depends only on the fluid pressure, or an *incompressible* fluid, i.e., a fluid whose density is constant. The residual part of the dissipation (4.99c) can be described in a thermodynamically consistent way via the framework of hyperplasticity (see, for instance, Collins and Houlsby [37] or Houlsby and Puzrin [27]). A brief discussion on this subject is given in Section 4.5.1.2, where hyperplasticity is considered in the context of finite strain mechanics. The last of the above-written inequalities is a repetition of (4.92a) (and, as such, is not numbered above). However, it must to be included in the considered formulation as strains (and, consequently, their rates) are constrained. One possible way of considering this constraint could be its inclusion in the weak form of the primary equations, via methods such as the Penalty Method (PM) and the Lagrange Multiplier Method (LMM) (which are addressed in detail in Chapter 5 for the contact problem). However, these methods add terms to the primary equations (and primary unknowns too, in the case of the LMM), which make them less attractive especially in the case where the primary equations are linearised and implicitly solved, as for this work. A more appealing way to include (4.92a) consists in modifying the free energy function (and, as a consequence, the effective stress-strain relationship). This solution is particularly easy to apply to those materials whose free energy function can be decomposed into a volumetric and deviatoric part (these materials being first described by Flory [38]). In this case, including the considered constraint can be ensured if the volumetric part of the free energy function satisfies the following features:

- it presents a vertical asymptote for volumetric strain approaching (from above) the values excluded by (4.92a), i.e., $\lim_{\epsilon_v \to \ln(1-n_0)^+} \hat{\Psi}^{sk}_{inc}(\epsilon_v) = +\infty$; and

- for the values of the volumetric strain excluded by (4.92a), the free energy function is not defined, i.e., $\hat{\Psi}^{sk}_{inc}(\epsilon_v) : \epsilon_v \in (\ln(1-n_0), +\infty) \to \mathbb{R}$.

In some aspects, the need to explicitly enforce solid mass conservation presents similarities to the discussion in Section 3.6.2.2 about conservation laws. All of the relationships which are not unequivocally discretised cannot be expected to inherit the properties of their continuum counterpart automatically. Owing to this, discrete algorithms should (must?) be designed for this purpose. With this in mind, the considerations covered in Sections 3.6.2.2, 3.6.3 and 3.6.6 of Chapter 3, Section 4.5.1.1 of this chapter and Section 5.4 of Chapter 5 become even more important.

4.5.1.1 Skeleton free energy function

This section presents a free energy function satisfying the features described in Section 4.5.1, while preserving the elasto-plastic setting derived for small-strain theory. Giving these two features, the new material is named *improved Hencky* material, and its free energy function is as follows

$$\hat{\Psi}^{sk}_{inc}(\boldsymbol{\epsilon}, \boldsymbol{\alpha}) = \frac{\tilde{K}}{2\,n}(\epsilon^e_v)^2 + \frac{3}{2}G(\epsilon^e_q)^2 + \tilde{\Psi}^{sk}_{inc}(\boldsymbol{\alpha}), \tag{4.101}$$

(a) Volumetric part of the free energy function.

(b) Effective Kirchhoff pressure.

(c) Tangent elastic bulk modulus.

(d) Tangent Elastic Poisson's ratio.
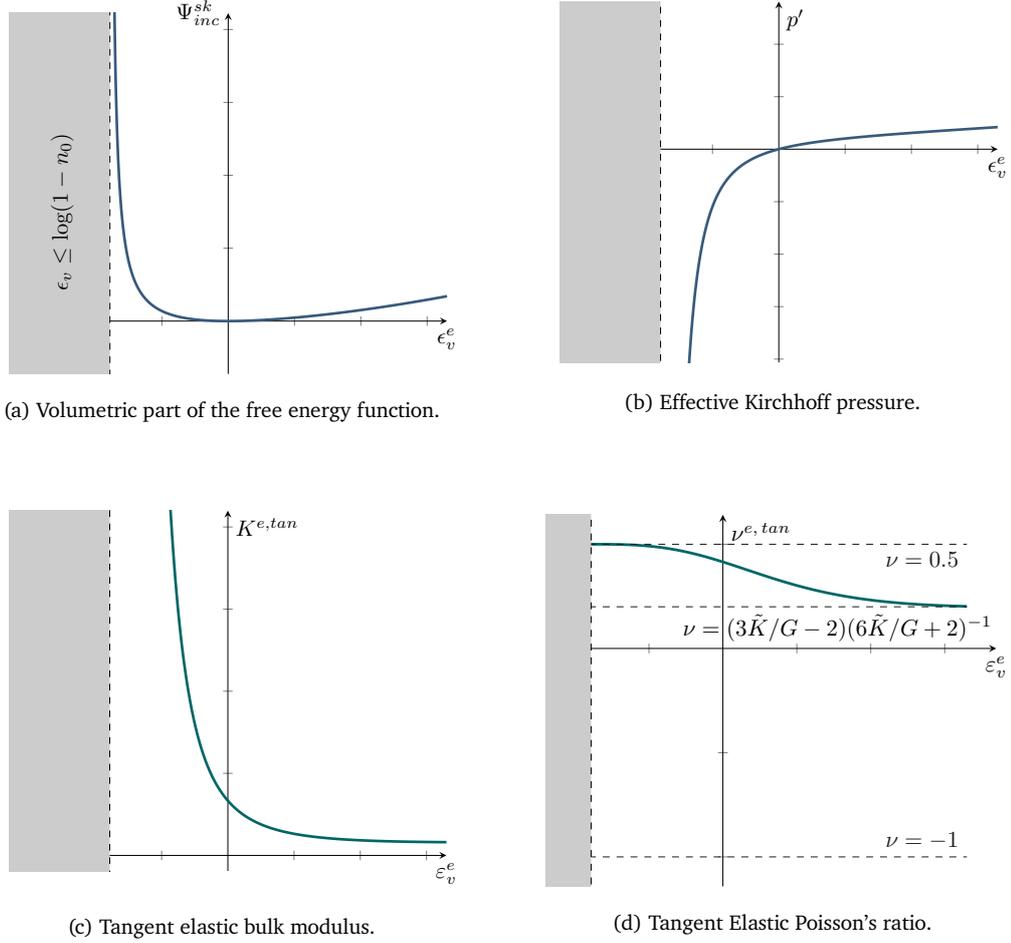
Figure 4.5: Bi-dimensional plots for the elastic case. Grey-shaded areas indicate where (4.92a) is not respected. Figure reproduced from Pretti *et al.* [30].

where the inequalities $\tilde{K} > 0$ and $G > 0$ hold as for the case of original Hencky material (firstly proposed in [39]). However, in contrast to Hencky's original material, $\tilde{K}$ is here defined as the *bulk parameter* (different from the tangent bulk modulus, described below). On the other hand, G is the shear modulus. Figures 4.5 and 4.6 provide a graphical illustration of the free energy function (Figures 4.5(a) and 4.6(a)), the effective pressure (Figures 4.5(b) and 4.6(b)) and the tangent elastic bulk modulus (Figures 4.5(c) and 4.6(c)) as a function of volumetric logarithmic strain, comparing the elastic (Figure 4.5) and the elasto-plastic (Figure 4.6) cases. Figure 4.5 also marks with a grey area the values of the volumetric logarithmic strain (elastic part) violated by the constraint (4.92a). As can be appreciated, the functions are not defined over that part of their domains. Some observations can be made of these figures and on Eq. (4.101) describing by the relative free energy function:

- as also suggested by keeping the same name for the shear modulus, there is no difference between the deviatoric behaviour of the original Hencky material and the improved one, which is constant in both models. However, in the literature, it has been

debated whether to keep the shear modulus constant rather than other elastic moduli constant while maintaining a variable bulk modulus. As pointed out by Zytinsky *et al.* [40] though, a material presenting a varying bulk modulus and a constant Poisson's ratio is non-conservative. Furthermore, in the case of the improved Hencky material, this material would fail to express the asymptotic behaviour when $\nu^{e\,tan} \to 0.5$, i.e., when approaching incompressible behaviour (as represented in Figure 4.5(d));

- since the improved Hencky material is a volumetric modification of the original one, it suffers from the same issues. In particular, both materials present a convex free energy function (which can be appreciated from Figures 4.5(a) and 4.6(a) for the improved case), and not poly-convex as required (see, for instance, Simo [41] for a discussion in this regard). However, as in the application for geo-materials, it is expected that plastic strains play a more dominant role with respect to elastic counterparts, so that issues in the elastic regime are mostly contained;

- the improved Hencky material described by Eq. (4.101) cannot be written in the form usually attributed to *decoupled* (or *uncoupled*, first described by Lubliner [42]) materials, these being $\Psi^{sk} = \Psi_1^{sk}(\boldsymbol{b}^e) + \Psi_2^{sk}(\boldsymbol{b}^p)$. In particular, the improved Hencky material exhibits what in the literature is usually referred to as *modulus coupling* (see Collins [43]). In the context of hyperplasticity, these materials require a non-associated flow rule. Further details about the compliance of the improved Hencky material with hyperplasticity are given in Section 4.5.1.2. Given Eq. (4.101) for the improved Hencky material, it is expected that the material, in the elasto-plastic case, exhibits a non-zero effective stress under purely plastic strains (see Figure 4.6). This can clearly understood considering that Eq. (4.101) includes the dependency on the Eulerian porosity n, this being, in turn, a function of the total Jacobian, as demonstrated by Eq. (4.89);

- the improved Hencky material presents an interesting analogy with free energy functions used in the context of damage mechanics (see, for instance, Houlsby and Puzrin for the thermodynamical background [27] and Murakami [44] for detailed explanation). For this category of materials, there exists a dependency of the elastic moduli from the so-called damage parameter, varying between 0 and 1, covering the range from totally undamaged to totally damaged material; and

- from a physical perspective, the improved Hencky material provides a stiffening behaviour when compressed, i.e., when water is progressively ejected from the decreasing pore volume. As such, the compressed material becomes gradually more similar to its solid constituent, this being incompressible by assumption. Most of the proposed geo-mechanics materials (see, for instance, Yamakawa [45] for an extension of previously-available models to the context of finite strain for solids) present stiffening behaviour when compressed. However, this is not sufficient to enforce (4.92a), which, as previously motivated, requires an asymptote, i.e., a very particular kind of stiffening.

(a) Volumetric part of the free energy function.

(b) Effective Kirchhoff pressure.

(c) Tangent elastic bulk modulus.
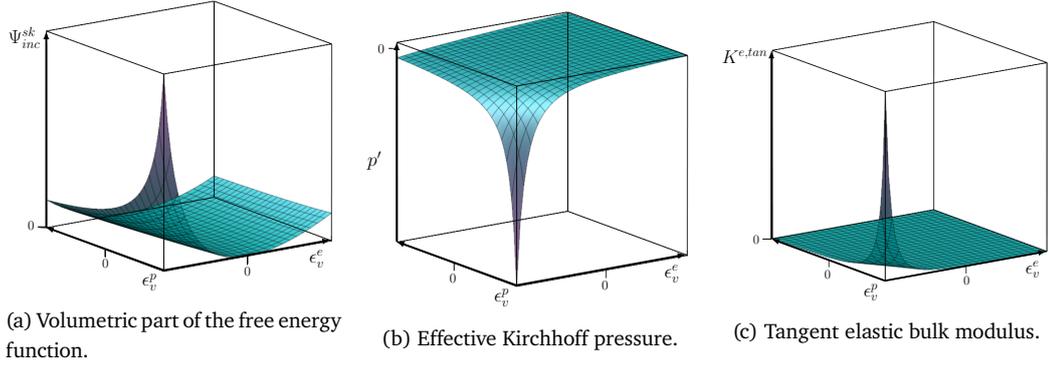
Figure 4.6: Three-dimensional plots of the volumetric parts described by Eq. (4.101), and its first and second derivatives with respect to ϵ_v^e. Figure reproduced from Pretti *et al.* [30].

Stress and elastic moduli computations are detailed in Appendix C.1 for the elastic case. Linearisation necessary to compute the stresses in correspondence with elasto-plastic strain as well as the tangent matrix $\mathcal{D}^{alg} := \frac{d\tau'}{d\epsilon^{e,tr}}$ for the Newton-Raphson (NR) procedure applied to the primary equations are given in Appendix C.2.

4.5.1.2 Dissipation function

If the part relative to the increment in fluid pressure is neglected, the Clausius-Planck inequality (4.98) can be concisely written as

$$\tau' : d = \dot{\hat{\Psi}}_{inc}^{(sk)} + \mathcal{D}^{(sk)}, \tag{4.102}$$

which constitutes the standard starting equation for hyperplastic formulations in the finite strain context. The reader can refer to Oliynyk and Tamagnini [46] or Pretti *et al.* [30] for further details on the subject.

In the case where no further dissipative mechanism is considered in the free energy function, i.e., $\dot{\tilde{\hat{\Psi}}}_{inc}^{(sk)}(\alpha) = 0$, the rate of the Helmholtz free energy function can be written as

$$\dot{\hat{\Psi}}_{inc}^{sk}(b^e, b^p) = \frac{\partial \hat{\Psi}_{inc}^{sk}}{\partial b_{ij}^e}\dot{b}_{ij}^e + \frac{\partial \hat{\Psi}_{inc}^{sk}}{\partial b_{ij}^p}\dot{b}_{ij}^p = 2\frac{\partial \hat{\Psi}_{inc}^{sk}}{\partial b_{ij}^e}\underbrace{b_{jk}^e d_{ki}^e}_{=\dot{b}_{ij}^e} + 2\frac{\partial \hat{\Psi}_{inc}^{sk}}{\partial b_{ij}^p}\underbrace{b_{jk}^p (F^e)_{il}^{-1} d_{lm}^p F_{mk}^e}_{=\dot{b}_{ij}^p}, \tag{4.103}$$

where the right-hand side of the above equation uses the kinematic relationships $\dot{b}^p = \bar{L}^p b^p + b^p (\bar{L}^p)^T$, with $\bar{L}^p := \dot{F}^p (F^p)^{-1}$ and $l^p := F^e \bar{L}^p (F^e)^{-1}$ (both have been previously defined in Eq. (3.43)). While Eq. (4.99a) gives the effective stress for a porous material, the same equation defines, in the context of hyperplasticity, what are referred to as *true stresses*. Two other measures of stress can be defined based on the quantities appearing in Eq. (4.102), these being the *shift stress* and the *dissipative stress*, defined respectively as

$$\chi_{lm}' := 2\frac{\partial \hat{\Psi}_{inc}^{sk}}{\partial b_{ij}^p}b_{jk}^p (F^e)_{il}^{-1} F_{mk}^e; \tag{4.104}$$

$$\varphi_{ij}' := \frac{\partial \left(\mathcal{D}^{sk}\right)}{\partial d_{ij}^p}. \tag{4.105}$$

Having excluded further dissipations, inequality (4.99c) suggests the dissipation mechanism being dependent only on the plastic stretching tensor $\boldsymbol{d}^p$. According to hypothesis **H.4.8** (rate effects are excluded), it can be postulated (see Oliynyk and Tamagnini [46]) that the dissipation is homogeneous of degree one in the plastic stretching tensor. Euler's theorem for this kind of functions leads to

$$\mathcal{D}^{sk} = \frac{\partial \left(\mathcal{D}^{sk} \right)}{\partial d_{ij}^p} d_{ij}^p = \varphi_{ij}' \, d_{ij}^p, \tag{4.106}$$

where definition (4.105) has been used on the RHS.

Eq. (4.102) can be re-written as (using Eqs. (4.103)-(4.106))

$$\tau_{ij}' \, d_{ij} = \varphi_{ij}' \, d_{ij}^p + \tau_{ij}' \, d_{ij}^e + \chi_{ij}' \, d_{ij}^p, \tag{4.107}$$

or, alternatively,

$$\tau_{ij}' = \varphi_{ij}' + \chi_{ij}'. \tag{4.108}$$

The above set of equations allows description of hyperplastic behaviour of a body, provided that the dissipation function $\mathcal{D}^{(sk)} \geq 0$ is given as a function of the plastic stretching tensor. In particular, if all of the terms of Eq. (4.106) are moved to one side, it gives the *dissipative yield condition*

$$\Phi^{\varphi'} := \mathcal{D}^{sk} - \varphi_{ij}' \, d_{ij}^p = 0. \tag{4.109}$$

The shift of the dissipative yield condition to the true stress space (via Eq. (4.108)) allows expression of the *yield function*

$$\Phi^{\varphi'} := \Phi^{\tau'} \leq 0, \tag{4.110}$$

as well as the direction of the plastic flow, this being

$$\frac{\partial \Phi^{\varphi'}}{\partial \varphi'} := \frac{\partial g^{\tau'}}{\partial \tau_{ij}'}. \tag{4.111}$$

From the above equations, it can be understood how prescribing a Helmholtz free energy function and a dissipation function is sufficient to establish the (non-associated) elasto-plastic behaviour of the skeleton.

4.5.1.3 Fluid free energy function

As previously mentioned in Section 4.5.1, the exclusion of thermal effects makes the fluid density $\rho^{(f)}$ exclusively a function of the pressure $p^{(f)}$. Even if (4.99b) does not exclude incompressible fluids, these are not considered in this work to keep the formulation as general as possible.

The fluid free enthalpy of a barotropic fluid $\breve{\mu}^{(f)}$ can be given by (see, for instance, Armero [47] or Gajo [48])

$$\breve{\mu}^{(f)} \left(\frac{1}{\rho^{(f)}} \right) = \frac{K^{(f)}}{\rho_{atm}^{(f)}} \left(1 - \frac{\rho_{atm}^{(f)}}{\rho^{(f)}} \right), \tag{4.112}$$

with $K^{(f)}$ being the fluid bulk modulus and $\rho_{atm}^{(f)}$ the density of the fluid at atmospheric pressure. In the case where water is considered (as for anchors embedding in the seabed), setting a particularly high value of $K^{(f)}$ makes the fluid slightly-compressible, thus achieving a good approximation of incompressible fluid. From the above equation and Eq. (4.99b), if follows that

$$\dot{\mu}^{(f)} = \frac{\dot{p}^{(f)}}{\rho^{(f)}} = \frac{K^{(f)}}{\left(\rho^{(f)}\right)^2} \dot{\rho}^{(f)}, \tag{4.113}$$

where a linear relationship between the ratio of density and the rate of pressure can be identified

$$\dot{\rho}^{(f)} = \frac{\rho^{(f)}}{K^{(f)}} \dot{p}^{(f)}. \tag{4.114}$$

If finite quantities are considered, the above equation becomes

$$\rho^{(f)} = \rho_{atm}^{(f)} \exp\left(\frac{p^{(f)}}{K^{(f)}}\right). \tag{4.115}$$

4.5.2 Fluid flux law

While the fluid constitutive equation has been introduced in the previous section, the fluid flow dissipation defined by (4.88b) still needs to be discussed. One simple way to satisfy inequality (4.88b) is to express the dissipation $\mathcal{D}^{(f)}$ as a quadratic in terms the flux $\boldsymbol{q}^{(f)}$, i.e.

$$\frac{1}{J}\mathcal{D}^{(f)} = \frac{1}{\rho^{(f)}\,k^{(f)}}\,\boldsymbol{q}^{(f)} \cdot \boldsymbol{q}^{(f)} \geq 0, \tag{4.116}$$

which implies the assumption of the so-called *Darcy-Weissbach law*, this being (for the case of isotropic permeability)

$$\boldsymbol{q}^{(f)} = -\rho^{(f)}\,k^{(f)}\left(\frac{\partial\,p^{(f)}}{\partial\boldsymbol{x}} - \rho^{(f)}\left(\boldsymbol{f} - \boldsymbol{a}^{(f)}\right)\right), \tag{4.117}$$

where $k^{(f)} \geq 0$ is the *mobility* of the sample.

From a physical perspective, the Darcy-Weissbach law expresses fluid flows with low *Reynolds numbers**, i.e., laminar flow (as assumed by **H.4.16**). As restrictive as this assumption may seem, it can be considered valid at least to a certain degree (depending on the permeability values and fluid velocity), especially distant from where the anchor interacts with the ground. Thus, it was the author's choice not to introduce a further non-linearity into this work (and, consequently, into the code inspired by it), nevertheless recognising its limitations.

As for the mobility $k^{(f)}$, this value is both a function of the fluid under consideration and the void network characterising the porosity. In the literature, there exist several other measures to account for the permeability of the sample. Among these, it is worth mentioning the *hydraulic conductivity* $\kappa^{(f)} := \rho^{(f)}\,g\,k^{(f)}$, with g being the gravity acceleration (used, for instance, by Bandara and Soga [49]), and the *absolute porosity* $\kappa := \eta^{(f)}\,k^{(f)}$ (with $\eta^{(f)}$ being

---

*For the sake of keeping a clear and compact discussion, basic subjects of the fluid dynamics such as the Reynolds number are not introduced and discussed in this work. The reader interested in this topic can refer to, for instance, Cengel and Cimbala [23].

the dynamic viscosity of the fluid phase) (see, for instance, Coussy [1]). The introduction of the former is justified in the text below, while the latter considers the permeability of the sample as dependent only the void network (i.e., excluding its dependency from the fluid under consideration). However, since data in the geotechnical literature (see, for instance, [50]) often make use of the hydraulic conductivity $\kappa^{(f)}$ and implicitly assume water filling the voids, this measure is considered throughout this work.

The above-mentioned dependency of the porosity network from the void network can be expressed via the *Kozeny–Carman* law. Following Bandara and Soga [49], the Kozeny–Carman law can be expressed by

$$\kappa^{(f)} = c_1 \frac{n^3}{\left(1 - n\right)^2}, \tag{4.118}$$

with c_1 being a constant parameter. As can be seen, this particular formula considers a relationship between the Eulerian value of the porosity n (and, in turn, of the Jacobian J via Eq. (4.89)) and the hydraulic conductivity $\kappa^{(f)}$.

The introduction of the fluid flux constitutive equation, such as the Darcy-Weissbach equation, allows one to compute the contribution of the force which one phase applies to the other, i.e., when equilibrium equations are written separately. If Eq. (4.44) is considered, via the use of Eq. (4.117), it follows that

$$f_i^{\rightarrow (f)} = p^{(f)} \frac{\partial n}{\partial x_i} - \frac{n}{k^{(f)} \rho^{(f)}} q_i^{(f)}. \tag{4.119}$$

While this equation is not particularly useful in the current work, it is here reported as its substitution into Eq. (4.44) constitutes the strong form of one of the primary equations for $\boldsymbol{v} - \boldsymbol{w}$ formulations, i.e., the fluid balance of rate of linear momentum.

4.6   The $\boldsymbol{u} - p^{(f)}$ formulation for the Material Point Method (MPM)

Having introduced so far all the necessary features (i.e., kinematics in Section 4.1, mass and momentum conservation in Sections 4.2 and 4.3, and the constitutive laws in Section 4.5 informed by the thermodynamics outlined in Section 4.4), it is now possible to model in all aspects a poromechanical problem. However, some further choices are available in the literature when considering MPM applied to this area.

Leaving aside the hybridisation of the MPM with other numerical methods, the literature offers possibilities regarding the primary equations to discretise and how this discretisation can be carried out in terms of set numbers of MPs. As for the former option, two possibilities are available, these being the $\boldsymbol{u} - p^{(f)}$ and the $\boldsymbol{v} - \boldsymbol{w}^*$ formulations, where the names of these methods are labelled after their primary variables. As expected, different primary values come with different discretised primary equations; both methods share the mixture balance

---

*While it is quite clear what $\boldsymbol{u} - p^{(f)}$ stands for, according to the notation used so far, $\boldsymbol{v} - \boldsymbol{w}$ refers to the velocities of the solid and fluid phase, respectively. In this work, the notation referring to these quantities (i.e., velocity of the solid and fluid phases) is different, these being $\boldsymbol{v}$ and $\boldsymbol{v}^{(f)}$. However, the above names are retained when referred to the methods to keep consistency with the published literature (see, for instance, Nøst [19]).

Table 4.1: Critical summary of the MPM formulations available in the literature.

	One set of MPs		Two sets of MPs	
$\boldsymbol{u} - p^{(f)}$ formulation	Conserving fluid mass; Consistent integration error between phases	$\boldsymbol{a} \approx \boldsymbol{a}^{(f)}$	Conserving fluid mass	$\boldsymbol{a} \approx \boldsymbol{a}^{(f)}$; Inconsistent integration between phases
$\boldsymbol{v} - \boldsymbol{w}$ formulation	No restricted assumption on $\boldsymbol{a}$; Consistent integration error between phases	Not conserving fluid mass	No restricted assumption on $\boldsymbol{a}$	Inconsistent integration between phases
	PROS	CONS	PROS	CONS

equation (whose integral form is given by Eq. (4.35)) as one of the primary equations. The second primary equation is given by the fluid mass balance (integral form given by Eq. (4.24)) for the $\boldsymbol{u} - p^{(f)}$ formulation and by the fluid balance equation (expressed by Eq. (4.44) in strong form) for the $\boldsymbol{v} - \boldsymbol{w}$ method.

With regards to the second option available in the literature (i.e., the set numbers of MPs), the spatial discretisation within the MPM can be performed either by one or two sets of MPs. In the former case, MPs are considered as carrying information relative to both the phases, while the latter case (namely two set of MPs) separates the phases and their relative data.

Table 4.1 reports an overview of the methods obtained combining the described options, considering strengths and weaknesses for each of them. In particular, $\boldsymbol{u} - p^{(f)}$ formulations are computationally less expensive compared to the $\boldsymbol{v} - \boldsymbol{w}$ counterpart, as there are less Degrees of Freedom (DoFs) per node (these are 4 and 6 for the three-dimensional case, respectively). However, this benefit for the $\boldsymbol{u} - p^{(f)}$ method comes with the price of assuming the acceleration of the solid phase to be very close to the fluid, i.e., $\boldsymbol{a} \approx \boldsymbol{a}^{(f)}$. In the literature (see Zhao and Choo [51] and references therein), this assumption is regarded as true when the porous material undergoes low frequency loading conditions, i.e., the applied loads do not induce a significant separation between the phases. Similarly, less expensive are methods using one set of MPs with respect to formulations using two, as these latter usually double the number of MPs to model consistently each phase.

Furthermore, when two sets of MPs are considered, the well-known integration error (see, for instance, Gan *et al.* [52] and the discussion is Section 3.8) afflicting the standard formulation of the MPM is exacerbated by the discrepancy of this error between the phases. This error inconsistency is due to the different position that the two sets of MPs may take within the mesh elements. As a hypothetical example, it may happen that the solid phase results are over-integrated due to the presence of multiple MPs representing this solid constituent, while the fluid part, in the same element, can be extremely poorly-integrated for the opposite reason.

Another feature in Table 4.1 concerns the fluid mass conservation for the $\boldsymbol{u} - p^{(f)}$ formu-

lations, satisfied via the considered primary equations. It is worth mentioning that this sentence differs from the conclusions drawn by Soga *et al.* [53], where it is stated that formulations discretised by one set of MPs do not conserve fluid mass. However, this sentence is not backed by further discussion on the topic. Owing to the above-mentioned presence of the fluid mass in the set of primary equations and given that there exists three-field mixed formulations $\boldsymbol{u} - \boldsymbol{w} - \boldsymbol{p}^f$ enforcing fluid mass conservation to methods in which $\boldsymbol{a} \neq \boldsymbol{a}^{(f)}$ (see, for instance, Noda and Toyoda [54] for the Finite Element Method (FEM) and Zheng *et al.* [55] for the MPM), the author reserves the right to disagree with Soga *et al.* [53] about this matter.

Thus, given all of the merits and drawbacks of each formulation, it was decided in this work to opt for a $\boldsymbol{u} - p^{(f)}$ formulation with one set of MPs. This formulation assumes that $\boldsymbol{a} \approx \boldsymbol{a}^{(f)}$, so that the mixture balance of rate of linear momentum Eq. (4.35) becomes

$$\frac{\partial}{\partial x_j} \sigma_{ij} + \rho \left(f_i - a_i \right) = \boldsymbol{0}, \tag{4.120}$$

where the density of the porous material ρ (introduced by Eq. (4.28)) has been used.

On the other hand, the fluid mass conservation expressed in the Lagrangian solid description Eq. (4.24) is given by

$$\dot{m}^{(f)} - J \frac{\partial}{\partial x_i} \left(\rho^{(f)} k^{(f)} \left(\frac{\partial p^{(f)}}{\partial x_i} - \rho^{(f)} \left(f_i - a_i \right) \right) \right) = 0, \tag{4.121}$$

where Darcy-Weissbach Eq. (4.117) has been substituted in the above equation. The definition of fluid mass content $m^{(f)}$ Eq. (4.23) and the introduction of the incremental linear relationship between pressure and density for barotropic fluids Eq. (4.114) leads to

$$\frac{\dot{m}^{(f)}}{\rho^{(f)}} = \dot{\phi} + \frac{\phi}{\rho^{(f)}} \dot{\rho^{(f)}} = \dot{\phi} + \frac{\phi}{K^{(f)}} \dot{p}^{(f)} = J \dot{\epsilon}_v + \frac{\phi}{K^{(f)}} \dot{p}^{(f)}, \tag{4.122}$$

which, plugged into Eq. (4.121), gives

$$\rho^{(f)} \dot{\epsilon}_v + \rho^{(f)} \frac{n}{K^{(f)}} \dot{p}^{(f)} - \frac{\partial}{\partial x_i} \left(\rho^{(f)} k^{(f)} \left(\frac{\partial p^{(f)}}{\partial x_i} - \rho^{(f)} \left(f_i - a_i \right) \right) \right) = 0. \tag{4.123}$$

From a physical perspective, the above equation reproduces the two mechanisms (both relative to the fluid part) to change volume of the porous material illustrated in Figure 4.4(c) and discursively described in Section 4.5.1. In particular, the volume change of the porous material (first term in Eq. (4.123)) is given by sum of the compressibility of the fluid phase (second term) with the relative fluid flow (third term).. This latter term is clearly allowed or not depending on the boundary conditions applied on the sample.

4.6.1 Weak form

Before moving to the weak form of the equations, it is necessary to fully introduce a proper description of the boundary γ of the porous material. Let us introduce a partition of the boundary (illustrated in Figure 4.7) given either by the Dirichlet (prescribed displacements
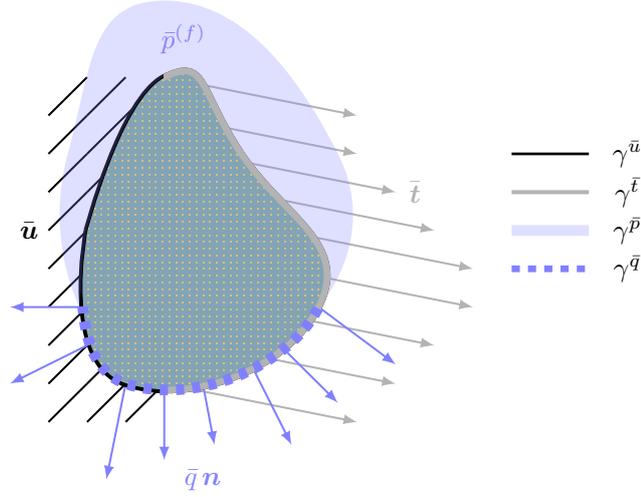
Figure 4.7: Partitioning of the boundary of the porous material according to Eqs. (4.124) and (4.125). Displacements and pressures are prescribed on the Dirichlet boundaries $\gamma^{\bar{u}}$ and $\gamma^{\bar{p}}$. Applied tractions force vectors and fluxes (these latter plotted with the help of the local normal direction vectors) are applied on the Neumann boundaries $\gamma^{\bar{t}}$ and $\gamma^{\bar{q}}$.

and pressure) and Neumann (applied tractions and fluxes) Boundary Conditions (BCs) as follows

$$\gamma = \gamma^{\bar{u}} \cup \gamma^{\bar{t}} = \gamma^{\bar{p}} \cup \gamma^{\bar{q}}; \tag{4.124}$$

$$\gamma^{\bar{u}} \cap \gamma^{\bar{t}} = \emptyset = \gamma^{\bar{p}} \cap \gamma^{\bar{q}}, \tag{4.125}$$

so that the applied BCs are

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \qquad \text{on } \gamma^{\bar{u}}; \tag{4.126}$$

$$\boldsymbol{\sigma}\,\boldsymbol{n} = \bar{\boldsymbol{t}} \qquad \text{on } \gamma^{\bar{t}}; \tag{4.127}$$

$$p^{(f)} = \bar{p} \qquad \text{on } \gamma^{\bar{p}}; \tag{4.128}$$

$$\boldsymbol{q}^{(f)} \cdot \boldsymbol{n} = \bar{q} \qquad \text{on } \gamma^{\bar{q}}, \tag{4.129}$$

with $\boldsymbol{n}$ being the outward normal to the surface in the current configuration.

In addition, the introduction of the spaces of admissible trial functions and weighting functions is necessary, these being

$$\mathscr{V}^{\boldsymbol{u}} = \left\{ \boldsymbol{u} \in \left[H^1\left(\omega \right) \right]^{n^{dim}} \mid J > 0 \land \boldsymbol{u} = \bar{\boldsymbol{u}} \text{ on } \gamma^{\bar{u}} \right\}; \tag{4.130}$$

$$\mathscr{V}^{p^{(f)}} = \left\{ p^{(f)} \in \left[H^1\left(\omega \right) \right] \mid p^{(f)} = \bar{p} \text{ on } \gamma^{\bar{p}} \right\}; \tag{4.131}$$

$$\mathscr{W}^{\boldsymbol{w}} = \left\{ \boldsymbol{w} \in \left[H^1\left(\omega \right) \right]^{n^{dim}} \mid \boldsymbol{w} = \boldsymbol{0} \text{ on } \gamma^{\bar{u}} \right\}; \tag{4.132}$$

$$\mathscr{W}^{\eta} = \left\{ \eta \in \left[H^1\left(\omega \right) \right] \mid \eta = 0 \text{ on } \gamma^{\bar{p}} \right\}, \tag{4.133}$$

with $H^1\left(\omega \right)$ denoting the *Sobolev* space of degree one over ω^*.

---

*On the limitations of the $H^1\left(\omega \right)$ space to compute fields related to the spatial gradients of the primary variables (e.g., stress or strain), the reader can refer to the discussion made in Section 3.6.1. In the case of porous material, the same discussion can be made for the derivatives of the pressure field.

While it is quite natural to integrate the mixture equilibrium equation on the whole current domain ω, the most natural choice about the fluid mass conservation would be to integrate over the fluid current domain $\omega^{(f)}$. However, given the relationship between the infinitesimal parts of these two volumes via the Eulerian porosity (i.e., $\omega^{(f)} = n\,\omega$), there is no difference between integrating Eq. (4.123) over these different volumes. As such, it follows that the integration of Eqs. (4.120) and (4.123) and their multiplication with relative weighting functions gives

$$\int_\omega w_i \left(\frac{\partial \sigma_{ij}}{\partial x_j} + \rho \left(f_i - a_i \right) \right) dv = \mathbf{0}, \qquad \forall \boldsymbol{w} \in \mathscr{W}^{\boldsymbol{w}}; \tag{4.134}$$

$$\int_\omega \eta \left(\rho^{(f)} \dot{\epsilon}_v + \rho^{(f)} \frac{n}{K^{(f)}} \dot{p}^{(f)} - \frac{\partial}{\partial x_i} \left(\rho^{(f)} k^{(f)} \left(\frac{\partial p^{(f)}}{\partial x_i} - \rho^{(f)} \left(f_i - a_i \right) \right) \right) \right) dv = 0,$$
$$\forall \eta \in \mathscr{W}^{\eta}. \tag{4.135}$$

Using the Fundamental Theorem of Calculus, the divergence theorem and applying Neumann BCs (4.127) and (4.129), the standard weak form for the above equations can be stated. This consists in finding $\left(\boldsymbol{u}, p^f \right) \in \left(\mathscr{V}^{\boldsymbol{u}}, \mathscr{V}^{p^{(f)}} \right)$ such that, for $t \in [0, T]$,

$$\int_\omega \frac{\partial w_i}{\partial x_j} \sigma_{ij} \, dv - \int_\omega \rho\, w_i \left(f_i - a_i \right) dv - \int_{\gamma^{\bar{t}}} w_i\, \bar{t}_i \, da = \mathbf{0}, \qquad \forall \boldsymbol{w} \in \mathscr{W}^{\boldsymbol{w}}; \tag{4.136}$$

$$\int_\omega \eta \left(\rho^{(f)} \dot{\epsilon}_v + \rho^{(f)} \frac{n}{K^{(f)}} \dot{p}^{(f)} \right) dv + \int_\omega \frac{\partial \eta}{\partial x_i} \left(\rho^{(f)} k^{(f)} \left(\frac{\partial p^{(f)}}{\partial x_i} - \rho^{(f)} \left(f_i - a_i \right) \right) \right) dv$$
$$+ \int_{\gamma^{\bar{q}}} \eta\, \bar{q} \, da = 0, \qquad \forall \eta \in \mathscr{W}^{\eta}, \tag{4.137}$$

given the initial conditions $\boldsymbol{u}\,(t=0) = \boldsymbol{u}_0$, $\dot{\boldsymbol{u}}\,(t=0) = \boldsymbol{v}_0$, and $p^{(f)}\,(t=0) = p_0^{(f)}$. It must also be noted that the above equations exploit the properties of the weighting functions to be zero on their Dirichlet boundary, as given by their spaces (4.132) and (4.133).

4.6.2 Discretisation processes

Section 4.6.2.1 introduces geometrical and primary field discretisations for the $\boldsymbol{u} - p^{(f)}$ formulation. Section 4.6.2.2 deals with temporal discretisation. Contrary to Chapter 3, the discussion of this latter section does not go into the details of compliance with the underlying continuum conservation laws. The reason for this choice is twofold: on the one hand, to the author's best knowledge, there are no studies in the literature that have discussed this subject; on the other hand, as evidenced by the discussion in Section 4.5.1, solid mass conservation is not automatically guaranteed for the current poromechanical problem, but must be explicitly enforced. Therefore, the author could not devote more time to the compliance of the discretised algorithm with its continuum formulation.

(a) Constant polynomial basis function defined by Eq. (4.138).

(b) Linear polynomial basis function defined by Eq. (4.139).

(c) Convolution of constant polynomial function with step characteristic function defined by Eq. (4.140).

(d) Convolution of linear polynomial function with step characteristic function defined by Eq. (4.141).
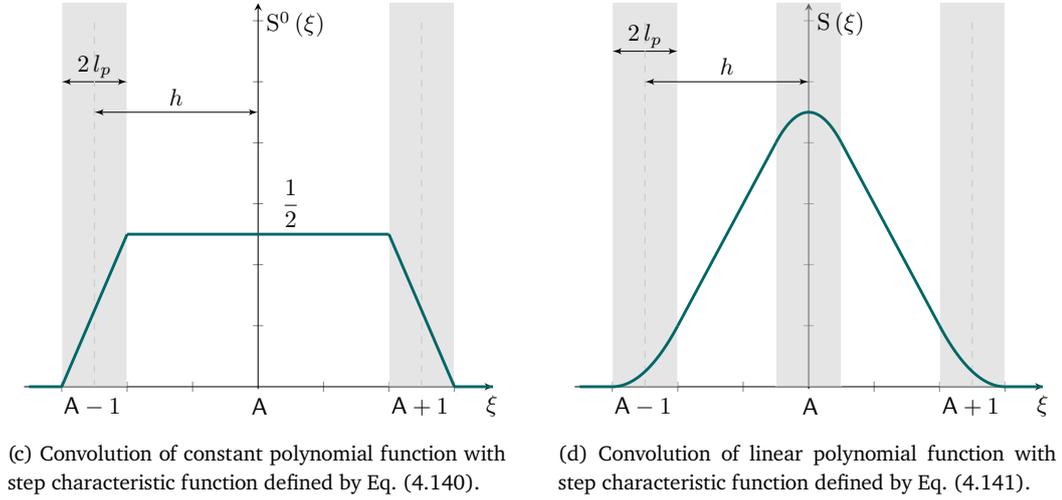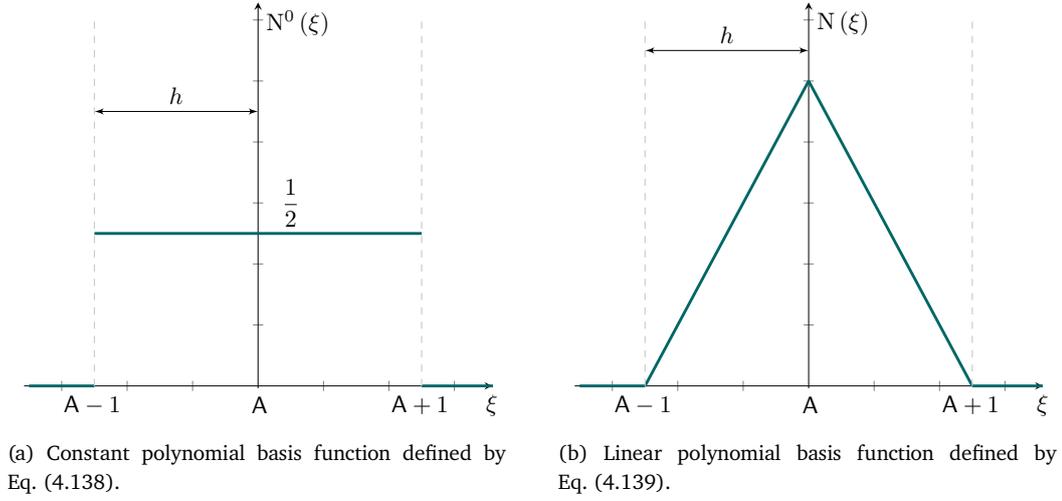
Figure 4.8: Qualitative illustration of the different basis functions used in $\boldsymbol{u} - p^{(f)}$ formulations. Shape functions are centred at the A-th node of the mesh.

4.6.2.1 Space and fields discretisation

As in Section 3.6.2.1 of Chapter 3, let us introduce a regular Cartesian grid, discretising the Euclidean space $\mathcal{E}$ and, consequently, the porous body $\mathcal{B}$ too in this fashion $^h\omega = \cup_{e=1}^{N^{els}} \omega^e$ (with $e = 1 \ldots N^{els}$ indicating each active element of the grid). Let the nodes of this mesh be denoted $\mathsf{A}, \mathsf{B}, \mathsf{C} = 1 \ldots N^{nds}$ and the vectorial DoFs be $I, J, K = 1 \ldots n^{dim} \times N^{nds}$. For purposes that will be clarified in Section 4.6.4.2, let the following constant and linear polynomial function be defined on the grid for the one-dimensional case, i.e.,

$$\mathrm{N}^0\left(\xi\right) = 1/2, \qquad \text{if} \quad -h < \xi \leq h; \tag{4.138}$$

$$\mathrm{N}\left(\xi\right) = \begin{cases} 1 + \xi/h, & \text{if} \quad -h < \xi \leq 0; \\ 1 - \xi/h, & \text{if} \quad 0 < \xi \leq h, \end{cases} \tag{4.139}$$

where ξ denotes the local coordinate and h is the mesh size.

Let the porous body $\mathcal{B}$ be discretised by one set of MPs, i.e., $^{MP}\omega = \cup_{mp=1}^{N^{mps}} v^{mp}$, where v^{mp} is the current total volume of a single MP. As for the solid case, let the characteristic functions (the functions defining the MP-based discretisation), to be the Dirac delta distribution for the MPM and the step functions (see details given in Section 3.6.2.1) for the Generalised Interpolation Material Point Method (GIMPM). The convolution of the grid functions with the characteristic ones gives basis and shape functions. In the case of MPM, the convolutions of Eqs. (4.138) and (4.139) with Dirac delta distribution leave these functions unaltered (as illustrated by Figures 4.8(a) and 4.8(b)). In the case of GIMPM, the convolutions of Eqs. (4.138) and (4.139) with the GIMPM characteristic function gives*

$$
S^0\left(\xi\right) = \begin{cases} \left(h + l_p + \xi\right) / \left(4l_p\right), & \text{if} \quad -h - l_p < \xi \leq -h + l_p; \\ 1/2, & \text{if} \quad -h + l_p < \xi \leq h - l_p; \\ \left(h + l_p - \xi\right) / \left(4l_p\right), & \text{if} \quad h - l_p < \xi \leq h + l_p; \end{cases} \tag{4.140}
$$

$$
S\left(\xi\right) = \begin{cases} \left(h + l_p + \xi\right)^2 / \left(4hl_p\right), & \text{if} \quad -h - l_p < \xi \leq -h + l_p; \\ 1 + \xi/h, & \text{if} \quad -h + l_p < \xi \leq -l_p; \\ 1 - \xi^2 / \left(2hl_p\right) + l_p^2 / \left(2hl_p\right), & \text{if} \quad -l_p < \xi \leq l_p; \\ 1 - \xi/h, & \text{if} \quad l_p < \xi \leq h - l_p, \\ \left(h + l_p - \xi\right)^2 / \left(4hl_p\right), & \text{if} \quad h - l_p < \xi \leq h + l_p. \end{cases} \tag{4.141}
$$

The above functions are represented in Figures 4.8(c) and 4.8(d).

Since the cloud of MPs is immersed in a new computational grid at every-time step, the two discretisations can be non-matching, i.e., $^h\omega \supseteq {}^{MP}\omega$. In Section 3.6.2.1 for the solid case, the choice of selecting the Dirichlet boundaries to be conforming (between the MP-based and the grid-based discretisations) and the use of boundary stabilisation for homogeneous Neumann boundaries was a good compromise between numerical effort and results. However, in the case of $\boldsymbol{u} - p^{(f)}$ formulations, the Dirichlet boundary (as can be seen from Eqs. (4.124) and (4.125) and Figure 4.7) is not unique as prescribed displacements and prescribed pressures boundaries exist. If these boundaries overlap, it is possible to return to assumptions similar to those made in Section 3.6.2.1 for a solid body. In the opposite scenario, it is necessary to choose which boundary should be conforming. The author's choice consisted of selecting the prescribed displacements boundary to be conforming, i.e., $^h\gamma^{\bar{u}} = {}^{MP}\gamma^{\bar{u}}$. Following this decision, the resulting reaction forces on this boundary can be correctly neglected, i.e.,

$$
\int_{\gamma^{\bar{u}}} w_i\, \sigma_{ij}\, n_j da = 0. \tag{4.142}
$$

However, as a result of the same decision, the contribution from fluxes arising on the prescribed pressure boundary cannot be set as zero. On this prescribed pressures boundary,

---

*The procedure to compute the convolution integrals resulting in Eqs. (4.140) was firstly given in Coombs *et al.* [56], while Eq. (4.141) in Bardenhagen *et al.* [57].

the zero value weighting function cannot be exploited, giving a non-null contribution

$$\int_{\gamma^{\bar{p}}} \eta \, \boldsymbol{q}^{(f)} \cdot \boldsymbol{n} \, da \neq 0. \tag{4.143}$$

For the same reasons described in Section 3.6.2.1 for tracking a Neumann boundary, this term is nonetheless neglected in what follows and pressure BCs, when prescribed, are enforced on the grid for simplicity. As will be clear from Example 4.7.2, the choice of prescribing displacement boundaries to be conforming and applying approximately the pressure on the grid is not always satisfactory and can be disadvantageous.

Following the introduced discretisations, let the finite-dimensional spaces of the trial and weighting functions be given by

$$^{h}\mathscr{V}^{\boldsymbol{u}} = \left\{ ^{h}u_i = \mathrm{N}^u_{iI} \, \mathrm{u}_I \; \middle| \; ^{h}u_i = \bar{u}_i \, (\mathrm{x}_I) \; \text{on} \; ^{h}\gamma^{\bar{u}} \right\}; \tag{4.144}$$

$$^{h}\mathscr{V}^{p^{(f)}} = \left\{ ^{h}p^{(f)} = \mathrm{n}^p_{\mathsf{A}} \, \mathrm{p}_{\mathsf{A}} \; \middle| \; ^{h}p^{(f)} = \bar{p} \, (\mathrm{x}_I) \; \text{on} \; ^{h}\gamma^{\bar{p}} \right\}; \tag{4.145}$$

$$^{h}\mathscr{W}^{\boldsymbol{w}} = \left\{ ^{h}w_i = \mathrm{N}^u_{iI} \, \mathrm{w}_I \; \middle| \; ^{h}w_i = \boldsymbol{0} \; \text{on} \; ^{h}\gamma^{\bar{u}} \right\}; \tag{4.146}$$

$$^{h}\mathscr{W}^{\eta} = \left\{ ^{h}\eta = \mathrm{n}^p_{\mathsf{A}} \, \eta_{\mathsf{A}} \; \middle| \; ^{h}\eta = 0 \; \text{on} \; ^{h}\gamma^{\bar{p}} \right\}. \tag{4.147}$$

De facto, unless otherwise specified (as in Section 4.6.4.2), the functions $\mathbf{N}^u$ and $\mathbf{n}^p$ are both given by Eq. (4.139) for the MPM, and by Eq. (4.141) for the GIMPM.

It must be highlighted that these spaces are not balanced, as expected by the so-called *inf-sup* or Ladyzhenskaya-Babuška-Brezzi (LBB) condition (see, for instance, White and Borja [58]). For this particular case of unbalanced spaces, some techniques are required (as addressed in Section 4.6.4.2). This work does not consider elements belonging to the *Taylor-Hood* family (see, for a detailed discussion of these, Brezzi and Fortin [59]), which are stable by design.

Owing to the above-introduced spaces, the finite-dimensional weighting and trial functions are given by

$$^{h}u_i \, (\boldsymbol{x}, t) = \mathrm{N}^u_{iI} \, (\boldsymbol{x}) \, \mathrm{u}_I \, (t); \tag{4.148}$$

$$^{h}p^{(f)} \, (\boldsymbol{x}, t) = \mathrm{n}^p_{\mathsf{A}} \, (\boldsymbol{x}) \, \mathrm{p}_{\mathsf{A}} \, (t); \tag{4.149}$$

$$^{h}w_i \, (\boldsymbol{x}) = \mathrm{N}^u_{iI} \, (\boldsymbol{x}) \, \mathrm{w}_I; \tag{4.150}$$

$$^{h}\eta \, (\boldsymbol{x}) = \mathrm{n}^p_{\mathsf{A}} \, (\boldsymbol{x}) \, \eta_{\mathsf{A}}. \tag{4.151}$$

The notation in the above equations highlights the different dependency from space and time of the trial functions, which are distributed between nodal time-dependent values and space-dependent shape functions. However, for the rest of this work, these dependencies are dropped to keep the notation as clear as possible.

It is now possible to write the discretised weak form for the $\boldsymbol{u} - p^{(f)}$ formulation, this consists in finding $\left(^{h}\boldsymbol{u}, {}^{h}p^f \right) \in \left(^{h}\mathscr{V}^{\boldsymbol{u}}, {}^{h}\mathscr{V}^{p^{(f)}} \right)$ such that, for $t \in [0, T]$,

$$\int_{^{h}\omega} \frac{\partial \mathrm{N}^u_{Ii}}{\partial x_j} \left(\sigma'_{ij} - \delta_{ij} \, \mathrm{n}^p_{\mathsf{A}} \, \mathrm{p}_{\mathsf{A}} \right) dv - \int_{^{h}\omega} \rho \, \mathrm{N}^u_{Ii} \left(f_i - \mathrm{N}^u_{iJ} \ddot{\mathrm{u}}_J \right) dv - \int_{^{h}\gamma^{\bar{t}}} \mathrm{N}^u_{Ii} \, \bar{t}_i \, da \approx \boldsymbol{0}; \tag{4.152}$$

$$\int_{^h\omega} \rho^{(f)} \mathrm{n}_\mathsf{A}^p \left(\delta_{ij} \frac{\partial \mathrm{N}_{iJ}^u}{\partial x_j} \dot{\mathrm{u}}_J + \frac{n}{K^{(f)}} \mathrm{n}_\mathsf{B}^p \dot{\mathrm{p}}_\mathsf{B} \right) dv$$

$$+ \int_{^h\omega} \rho^{(f)} k^{(f)} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_i} \left(\frac{\partial \mathrm{n}_\mathsf{B}^p}{\partial x_i} \mathrm{p}_\mathsf{B} - \rho^{(f)} \left(f_i - \mathrm{N}_{iI}^u \ddot{\mathrm{u}}_I \right) \right) dv + \int_{^h\gamma^{\bar{q}}} \mathrm{n}_\mathsf{A}^p \bar{q}\, da \approx \mathbf{0}, \quad (4.153)$$

given the initial conditions ${}^h\boldsymbol{u}\left(t=0 \right) = \boldsymbol{u}_0$, ${}^h\dot{\boldsymbol{u}}\left(t=0 \right) = \boldsymbol{v}_0$, and ${}^h p^{(f)}\left(t=0 \right) = p_0^{(f)}$. Comparing the above equations with their continuum counterparts given by Eqs. (4.134) and (4.135), it can be seen that Terzaghi effective stress decomposition (4.100) and the kinematic relationship $\dot{\epsilon}_v = \delta_{ij} \frac{\partial v_i}{\partial x_j}$ (compare with Eqs. (3.35) and (3.36)) have been applied. It can also be noticed that the independence of the Eqs. (4.152) and (4.153) from the weighting functions has been taken into account, simplifying their forms. These equations can be expressed in more concise ways by the matrix-vector system with entries

$$\mathrm{e}_I\left(\mathbf{u}, \mathbf{p} \right) := \mathrm{f}_I^{',int} + \left(\mathrm{Q}_{\nabla u,p}^{(1)} \right)_{IC} \mathrm{p}_\mathsf{C} + \mathrm{f}_I^{ext} + \mathrm{M}_{IK}\, \ddot{\mathrm{u}}_K \approx \mathbf{0}; \qquad (4.154)$$

$$\mathrm{d}_\mathsf{A}\left(\mathbf{u}, \mathbf{p} \right) := \left(\mathrm{C}_{p,p} \right)_{\mathsf{AC}} \dot{\mathrm{p}}_\mathsf{C} + \left(\mathrm{P}_{p,\nabla u} \right)_{\mathsf{A}K} \dot{\mathrm{u}}_K + \left(\mathrm{T}_{\nabla p,\nabla p} \right)_{\mathsf{AC}} \mathrm{p}_\mathsf{A}$$

$$+ \mathrm{f}_\mathsf{A}^{f,ext} + \left(\mathrm{Q}_{\nabla p,u}^{(2)} \right)_{\mathsf{A}K} \ddot{\mathrm{u}}_K \approx \mathbf{0}, \quad (4.155)$$

having set

$$\mathrm{f}_I^{',int}\left(\mathbf{u} \right) = \int_{^h\omega} \frac{\partial \mathrm{N}_{Ii}^u}{\partial x_j}\, \sigma_{ij}'\, dv \approx \sum_{mp}^{N^{mps}} v^{mp} \frac{\partial \mathrm{N}_{Ii}^u\left(\boldsymbol{x}^{mp} \right)}{\partial x_j}\, \sigma_{ij}'\left(\boldsymbol{x}^{mp} \right); \qquad (4.156)$$

$$\left(\mathrm{Q}_{\nabla u,p}^{(1)} \right)_{IC} = - \int_{^h\omega} \frac{\partial \mathrm{N}_{Ii}^u}{\partial x_j} \delta_{ij}\, \mathrm{n}_\mathsf{A}^p\, dv \approx - \sum_{mp}^{N^{mps}} v^{mp} \frac{\partial \mathrm{N}_{Ii}^u\left(\boldsymbol{x}^{mp} \right)}{\partial x_j} \delta_{ij}\, \mathrm{n}_\mathsf{A}^p\left(\boldsymbol{x}^{mp} \right); \qquad (4.157)$$

$$\mathrm{f}_I^{ext}\left(\mathbf{u}, \mathbf{p} \right) = - \int_{^h\omega} \rho\, \mathrm{N}_{Ii}^u f_i\, dv - \int_{^h\gamma^{\bar{t}}} \mathrm{N}_{Ii}^u\, \bar{t}_i\, da \approx - \sum_{mp}^{N^{mps}} \bar{m}^{mp}\, \mathrm{N}_{Ii}^u\left(\boldsymbol{x}^{mp} \right) f_i - \int_{^h\gamma^{\bar{t}}} \mathrm{N}_{Ii}^u\, \bar{t}_i\, da;$$

$$(4.158)$$

$$\mathrm{M}_{IK}\left(\mathbf{u}, \mathbf{p} \right) = \int_{^h\omega} \rho\, \mathrm{N}_{Ii}^u\, \mathrm{N}_{iK}^u\, dv \approx \sum_{mp}^{N^{mps}} \bar{m}^{mp}\, \mathrm{N}_{Ii}^u\left(\boldsymbol{x}^{mp} \right) \mathrm{N}_{iK}^u\left(\boldsymbol{x}^{mp} \right); \qquad (4.159)$$

$$\left(\mathrm{C}_{p,p} \right)_{\mathsf{AC}}\left(\mathbf{u}, \mathbf{p} \right) = \int_{^h\omega} \rho^{(f)} \frac{n}{K^{(f)}} \mathrm{n}_\mathsf{A}^p\, \mathrm{n}_\mathsf{C}^p\, dv \approx \sum_{mp}^{N^{mps}} \frac{\left(\bar{m}^{(f)} \right)^{mp}}{K^{(f)}} \mathrm{n}_\mathsf{A}^p\left(\boldsymbol{x}^{mp} \right) \mathrm{n}_\mathsf{C}^p\left(\boldsymbol{x}^{mp} \right); \quad (4.160)$$

$$\left(\mathrm{P}_{p,\nabla u} \right)_{\mathsf{A}K}\left(\mathbf{u}, \mathbf{p} \right) = \int_{^h\omega} \rho^{(f)} \mathrm{n}_\mathsf{A}^p\, \delta_{ij} \frac{\partial \mathrm{N}_{iK}^u}{\partial x_j}\, dv \approx \sum_{mp}^{N^{mps}} \left(\rho^{(f)} \right)^{mp} v^{mp}\, \mathrm{n}_\mathsf{A}^p\, \delta_{ij} \frac{\partial \mathrm{N}_{iK}^u\left(\boldsymbol{x}^{mp} \right)}{\partial x_j};$$

$$(4.161)$$

$$\left(\mathrm{T}_{\nabla p,\nabla p} \right)_{\mathsf{AC}}\left(\mathbf{u} \right) = \int_{^h\omega} \rho^{(f)} k^{(f)} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_i} \frac{\partial \mathrm{n}_\mathsf{C}^p}{\partial x_i}\, dv$$

$$\approx \sum_{mp}^{N^{mps}} \left(\rho^{(f)} \right)^{mp} \left(k^{(f)} \right)^{mp} v^{mp} \frac{\partial \mathrm{n}_\mathsf{A}^p\left(\boldsymbol{x}^{mp} \right)}{\partial x_i} \frac{\partial \mathrm{n}_\mathsf{C}^p\left(\boldsymbol{x}^{mp} \right)}{\partial x_i}; \quad (4.162)$$

$$\mathrm{f}_\mathsf{A}^{f,ext}\left(\mathbf{u}, \mathbf{p} \right) = - \int_{^h\omega} \left(\rho^{(f)} \right)^2 k^{(f)} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_i} f_i\, dv + \int_{^h\gamma^{\bar{q}}} \mathrm{n}_\mathsf{A}^p\, \bar{q}\, da$$

$$\approx -\sum_{mp}^{N^{mps}} \left(\left(\rho^{(f)}\right)^{mp}\right)^2 \left(k^{(f)}\right)^{mp} v^{mp} \frac{\partial \mathrm{n}_{\mathsf{A}}^p \left(\boldsymbol{x}^{mp}\right)}{\partial x_i} f_i + \int_{{}^h\gamma_{\bar{q}}} \mathrm{n}_{\mathsf{A}}^p \bar{q}\, da; \quad (4.163)$$

$$\left(\mathrm{Q}_{\nabla p,u}^{(2)}\right)_{\mathsf{A}K}(\mathbf{u},\mathbf{p}) = \int_{{}^h\omega} \left(\rho^{(f)}\right)^2 k^{(f)} \frac{\partial \mathrm{n}_{\mathsf{A}}^p}{\partial x_i} \mathrm{N}_{iK}^u\, dv$$

$$\approx \sum_{mp}^{N^{mps}} \left(\left(\rho^{(f)}\right)^{mp}\right)^2 \left(k^{(f)}\right)^{mp} v^{mp} \frac{\partial \mathrm{n}_{\mathsf{A}}^p \left(\boldsymbol{x}^{mp}\right)}{\partial x_i} \mathrm{N}_{iK}^u \left(\boldsymbol{x}^{mp}\right). \quad (4.164)$$

The RHSs of Eqs. (4.156)-(4.164) report the numerical approximations of the volume integrals, which, as standard in the MPM, are carried out via the MPs. Quantities denoted by the superscript $(\bullet)^{(mp)}$ are computed at the MP location and have all been previously defined, with the exclusion of $\bar{m}^{(f)} = \rho^{(f)} v^{(f)} = \rho^{(f)} n v$, and $\bar{m} = \bar{m}^{(sk)} + \bar{m}^{(f)}$, with $\bar{m}^{(sk)} = \rho^{(sk)} v^{(sk)} = \rho^{(sk)} (1-n)\, v$ for consistency.

As for the solid case, it can also be seen that the surface integrals on the RHSs of Eqs. (4.158) and (4.163) highlight again the above-discussed issue in the MPM, namely tracking and applying Neumann BCs.

4.6.2.2 Time discretisation

To fully discretise the weak form given by Eqs. (4.152) and (4.153), a temporal discretisation is required too. As in Section 3.6.2.2, the total time $[0,T]$ under consideration is divided into smaller fractions $\Delta t := t_{n+1} - t_n$. For the sake of clarity, let us also drop the subscript for the current time-step discretised quantities, i.e., $(\bullet) = (\bullet)_{n+1}$. Differently from Chapter 3 for the solid dynamic case*, let us relate the time-discretised variables according to the *Newmark* time-discretisation. If the primary variables are gathered in an unique vector as follows

$$\mathbf{x} := \begin{bmatrix} \mathbf{u} \\ \mathbf{p} \end{bmatrix}, \quad (4.165)$$

then the Newmark time-discretisation assumes these relationships among time derivatives

$$\mathbf{x} = \mathbf{x}_n + \Delta t\, \dot{\mathbf{x}}_n + \frac{\Delta t^2}{2} \left(1 - 2\,\tilde{\beta}\right) \ddot{\mathbf{x}}_n + \Delta t^2\, \tilde{\beta} \ddot{\mathbf{x}}; \quad (4.166)$$

$$\dot{\mathbf{x}} = \dot{\mathbf{x}}_n + \Delta t \left(1 - \tilde{\gamma}\right) \ddot{\mathbf{x}}_n + \Delta t\, \tilde{\gamma}\, \ddot{\mathbf{x}}, \quad (4.167)$$

with $\tilde{\gamma}$ and $\tilde{\beta}$ being fixed coefficients.

4.6.3 Linearisation

Introducing the above Newmark time-discretisation in Eqs. (4.154) and (4.155) provides the primary discrete equations for the $\boldsymbol{u} - p^{(f)}$ formulation considered in this work. As these

---

*While for the solid dynamic case the *mid-point* rule permits to inherit the conservation of angular momentum within the time-step from the continuum formulation (as proved by Simo and Tarnow [60]), no proof is available in the literature for the porous dynamics counterpart. The lack of such formulation, as well as the lack of techniques to map the fluid pressure from MPs to the grid and back minimising energy losses (see 4.6.5), motivates the use of the Newmark time-discretisation for this $\boldsymbol{u} - p^{(f)}$ formulation.

equations are non-linear, the NR algorithm is applied to find iteratively the solution for the primary unknowns. If the primary unknowns are denoted in a unique vector as follows

$$\mathbf{r} := \begin{bmatrix} \mathbf{e} \\ \mathbf{d} \end{bmatrix}, \tag{4.168}$$

the linearisation of these equations can be set to zero according to the NR iterative scheme

$$\mathbf{r}^{(k)} \approx \mathbf{r}^{(k-1)} + \left.\frac{\partial \mathbf{r}}{\partial \mathbf{x}}\right|^{(k-1)} \delta \mathbf{x} \approx \mathbf{0}, \tag{4.169}$$

with $(\bullet)^{(k)}$ being the current k-th iteration. The full expression of the Jacobian matrix appearing in the above equation is given by

$$\left.\frac{\partial \mathbf{r}}{\partial \mathbf{x}}\right|^{(k-1)} = \begin{bmatrix} \left.\frac{\partial \mathbf{e}}{\partial \mathbf{u}}\right|^{(k-1)} & \left.\frac{\partial \mathbf{e}}{\partial \mathbf{p}}\right|^{(k-1)} \\ \left.\frac{\partial \mathbf{d}}{\partial \mathbf{u}}\right|^{(k-1)} & \left.\frac{\partial \mathbf{d}}{\partial \mathbf{p}}\right|^{(k-1)} \end{bmatrix} := \begin{bmatrix} A_{IJ}^{(1)} & B_{I\mathrm{B}}^{(1)} \\ B_{\mathrm{A}J}^{(2)} & A_{\mathrm{AB}}^{(2)} \end{bmatrix}. \tag{4.170}$$

To improve the readability of the text, computations of the sub-matrices appearing in the Jacobian above are reported in Appendix B.2.

4.6.4 Stabilisation techniques

The primary Eqs. (4.154) and (4.155) are not stable for two different reasons. The first source of instability lies in the small-cut issue, which, as the MPM simulation evolves, is practically unavoidable. For a discussion on the small-cut issue and possible solutions, the reader is referred to Section 3.6.5. The application of the face penalty version of the ghost method to a $\boldsymbol{u} - p^{(f)}$ formulation has been proposed by Liu *et al.* [61] in the context of cut-FEM. This work adapts this stabilisation to the MPM, as detailed in Section 4.6.4.1.

On the other hand, the second instability is caused by the choice of the finite-dimensional space of functions for the displacement and fluid pressure fields. As well-recognised in the literature (see, for instance, White and Borja [58] for the FEM and Zhao and Choo [51] for the MPM), using polynomial linear shape function for both the primary unknowns violates the *inf-sup* conditions. An assessment of the different techniques available in the literature to stabilise inf-sup unstable methods has been carried out by Preisig and Prevost [62] for FEM. Among the proposed techniques, the Polynomial Pressure Projection (PPP) has been selected for this work, as its application to the MPM has already been proposed (see Zhao and Choo [51]). This technique is briefly discussed in Section 4.6.4.2.

4.6.4.1 Ghost stabilisation(s)

Following the proposal of Liu *et al.* [61], four different ghost stabilisations are here applied, two per each primary equation. In the case of linear polynomial functions or GIMPM functions (Eqs. (4.139) and (4.141), respectively), the matrix relative to the ghost stabilisations

of the mixture equilibrium equation is given by (similarly to Eq. (3.151))

$$(\mathrm{J}_G)_{IK} := \frac{h^3}{3} \int_{\gamma_{(G)}} \left(\frac{\partial \mathrm{N}_{Ii}^{u,+}}{\partial (x_n)_j} n_j^h - \frac{\partial \mathrm{N}_{Ii}^{u,-}}{\partial (x_n)_j} n_j^h \right) \left(\frac{\partial \mathrm{N}_{iK}^{u,+}}{\partial (x_n)_k} n_k^h - \frac{\partial \mathrm{N}_{iK}^{u,-}}{\partial (x_n)_k} n_k^h \right) da, \quad (4.171)$$

The meaning of the terms appearing in the above equation follows the same explanation given in Section 3.6.5. The introduction of the above matrix allows a stable mixture equilibrium equation via the stabilisation of the mass matrix

$$(\mathrm{M}_G)_{IK} := \gamma^M \ (\mathrm{J}_G)_{IK} \,, \tag{4.172}$$

and the stiffness matrix

$$(\mathrm{K}_G)_{IK} := \gamma^K \ (\mathrm{J}_G)_{IK} \,, \tag{4.173}$$

with γ^M and γ^K being user-selected parameters. When incorporating both the above stabilisations, Eq. (4.154) becomes

$$\tilde{\mathrm{e}}_I \left(\mathbf{u}, \mathbf{p} \right) := \mathrm{f}_I^{',int} + (\mathrm{K}_G)_{IK} \ \mathrm{u}_K + \left(\mathrm{Q}_{\nabla u,p}^{(1)} \right)_{I\mathsf{C}} \mathrm{p}_\mathsf{C} + \mathrm{f}_I^{ext} + (\mathrm{M}_{IK} + (\mathrm{M}_G)_{IK}) \ \ddot{\mathrm{u}}_K \approx \mathbf{0}. \quad (4.174)$$

It can be noted that, in *quasi-static* conditions, the stabilisation (4.172) is not required in Eq. (4.174).

Since this work uses the same basis functions interpolating the pressure field as those used for the displacement field, a term similar to (4.171) must be considered to stabilise the fluid mass conservation (4.155), this being

$$(\mathrm{i}_G)_{\mathsf{AC}} := \frac{h^3}{3} \int_{\gamma_{(G)}} \left(\frac{\partial \mathrm{n}_{\mathsf{A}}^{p,+}}{\partial (x_n)_j} n_j^h - \frac{\partial \mathrm{n}_{\mathsf{A}}^{p,-}}{\partial (x_n)_j} n_j^h \right) \left(\frac{\partial \mathrm{n}_{\mathsf{C}}^{p,+}}{\partial (x_n)_j} n_j^h - \frac{\partial \mathrm{n}_{\mathsf{C}}^{p,-}}{\partial (x_n)_j} n_j^h \right) da. \quad (4.175)$$

Specifically, the above term is used to stabilise a part relative to the pressure time derivative $\dot{\mathbf{p}}$

$$\left(\mathrm{M}_G^{(f)} \right)_{\mathsf{AC}} := \gamma^{M^{(f)}} \ (\mathrm{i}_G)_{\mathsf{AC}} \,, \tag{4.176}$$

and a part relative to pressure $\mathbf{p}$ itself

$$\left(\mathrm{K}_G^{(f)} \right)_{\mathsf{AC}} := \gamma^{K^{(f)}} \ (\mathrm{i}_G)_{\mathsf{AC}} \,, \tag{4.177}$$

with the parameters $\gamma^{M^{(f)}}$ and $\gamma^{K^{(f)}}$ being selected by the user. The stabilised version of Eq. (4.155) becomes

$$\tilde{\mathrm{d}}_\mathsf{A} := \mathrm{d}_\mathsf{A} + \left(\mathrm{M}_G^{(f)} \right)_{\mathsf{AC}} \dot{\mathrm{p}}_\mathsf{C} + \left(\mathrm{K}_G^{(f)} \right)_{\mathsf{AC}} \mathrm{p}_\mathsf{C} \approx \mathbf{0}. \tag{4.178}$$

4.6.4.2 Polynomial Pressure Projection stabilisation

When the same order of shape functions are considered for the displacement and pressure fields, the inf-sup condition is violated. As such, additional treatment of the primary equations is required. If the PPP (firstly proposed by Dohrmann and Bochev *et al.* [63] and adapted to $\boldsymbol{u} - p^{(f)}$ formulations by White and Borja [58]) is considered, the stabilising term must be added to Eq. (4.178), this being given by

$$\mathrm{R}_\mathsf{A}^{stab} = \int_{h_\omega} \rho^{(f)} \tau \left(\mathrm{n}_\mathsf{A}^p - \mathrm{n}_\mathsf{A}^{p,0} \right) \left(\mathrm{n}_\mathsf{C}^p - \mathrm{n}_\mathsf{C}^{p,0} \right) \dot{\mathrm{p}}_\mathsf{C} \, dv, \tag{4.179}$$

where the local expression for the shape functions $\mathrm{n}^{p,\,0}$ is given by Eq. (4.138) for the original MPM and Eq. (4.140) for the GIMPM. The parameter τ and its selection have been discussed widely in the literature (see, for instance, Zhao and Choo [51]).

4.6.5 Mapping processes

This section extends the discussion reported in Section 3.6.6, in the sense that a porous material requires considering which values should be mapped between the different load steps. For a solid dynamic problem, two forms of energies have been considered, namely the kinetic and the solid strain energy. In particular, only the assessment of the mappings involving the former energy has been carried out, as strain energy is stored entirely at the MPs. When it comes to a porous material, the fluid part can store energy too if this phase is compressible. This form of stored energy is expressed, for instance, by Eq. (4.99b), which particularises into (4.112) for a barotropic fluid. Hence, an assessment of this fluid stored energy for both the Point-to-grid (P2G) and Grid-to-point (G2P) mappings should be carried out. However, this evaluation is more complicated than the assessment of the kinetic energy. While this latter form of energy has the same form independently from the body under consideration (i.e., linearly proportional to the mass of the body and to the square of the velocity), the strain energy depends on the assumed constitutive relationship. From this perspective, it can be observed how fluid pressure mappings present all the typical difficulties of reconstructing the stress state. This is possibly the reason why, to the author's best knowledge, the few publications considering these mapping processes do not discuss them thoroughly in terms of energy loss. Zheng *et al.* [20, 64] move the nodal pressure value to central Gauss points and, in turn, use a Moving Least Square approximation to compute the pressure values at the MPs. However, no information is provided on the P2G mapping. Zhang *et al.* [65] and Zheng *et al.* [66] update the MP fluid pressure in a FLuid Implicit Particle (FLIP)-alike procedure for the G2P mapping, while the P2G mapping is not assessed. Zabala and Alonso [67] use the same FLIP-alike mapping from the G2P process, while mapping the pressure from MPs to grid nodes by averaging it over the volume.

From the lack of consistency across the literature, it can be seen how mapping processes in a $\boldsymbol{u}-p^f$ formulation are not trivial and require a rigorous assessment in terms of conservation laws, similar to that performed in Section 3.6.6 for the solid dynamic case. Acknowledging this intricacy, the mapping processes carried on below are not thoroughly evaluated, but are rooted in empirical considerations.

Nonetheless, some quite interesting observations can be made about the necessity of mapping in the first place when considering an incompressible fluid. From a physical perspective, such a fluid cannot store any strain energy. Mathematically, the fluid pressure acts as Lagrange Multiplier (LM) by shifting the dependency of the stresses on strains (i.e., the concept of effective stress). This shift depends on the drained/undrained/partially drained conditions. As such, it does not contribute to the internal energy of the body, and there is no need to map the fluid pressure under an energetic perspective.

If the case of incompressible fluid and quasi-static conditions, it can be seen that no need to perform the P2G mapping is required. However, this mapping can be performed to have better predictors (which influence the NR convergence, but do not affect the energy present in the body).

For a better explanation, the following parts regarding mapping techniques suspends the use of index notation and shows explicit sums of the saturated indices.

4.6.5.1 Point-to-grid (P2G) mapping

Velocity mapping   Following Section 3.6.6.1 for the solid dynamic case, the P2G mapping is performed as follows

$$\sum_{\mathsf{B}}^{N^{nds}} \left(\mathbf{M}_{\mathsf{AB}}\right)_n \left(\mathbf{v}_{\mathsf{B}}\right)_n = \sum_{mp}^{N^{mps}} \bar{m}_n^{mp} \, \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp}\right) \boldsymbol{v}_n^{mp}, \tag{4.180}$$

where the mass matrices are defined by Eq. (4.159) and (4.172). For the porous material under consideration, the evaluation of the use of effective mass matrices instead of the consistent mass matrix follows the explanation in Section 3.6.6.

Pressure mapping   As explained above, the P2G mapping of the fluid pressure for a barotropic fluid requires an assessment of the difference between the energy computed on the grid and the one stored at the MPs. As this is not available in the literature, this work considers to average the fluid pressure over the current volume as follows (following Zabala and Alonso's proposal [67])

$$\sum_{\mathsf{A}}^{N^{nds}} \left(\Omega_{\mathsf{AB}}\right)_n \left(p_{\mathsf{A}}\right)_n = \sum_{mp}^{N^{mps}} v_n^{mp} \, \mathrm{n}_{\mathsf{B}}^p(\boldsymbol{x}^{mp}) \left(p^{(f)}\right)_n^{mp}, \tag{4.181}$$

where the matrix Ω_{AB} is defined as

$$\left(\Omega_{\mathsf{AB}}\right)_n := \mathrm{diag}_A \left(\sum_{mp}^{N^{mps}} v_n^{mp} \, \mathrm{n}_{\mathsf{B}}^p(\boldsymbol{x}^{mp}) \right). \tag{4.182}$$

The choice of a lumped matrix for the P2G pressure mapping is a decision based on a lower computational cost, since the advantages of using a consistent one are not demonstrated.

4.6.5.2 Grid-to-point (G2P) mapping

Velocity mapping   This follows Section 3.6.6.2 for the solid dynamic part, favouring the FLIP procedure over Particle In Cell (PIC), i.e.,

$$\boldsymbol{v}^{mp} = \boldsymbol{v}_n^{mp} + \sum_{\mathsf{A}}^{N^{nds}} \mathrm{N}_{\mathsf{A}}^u \left(\boldsymbol{x}^{mp}\right) \Delta \mathbf{v}_{\mathsf{A}}. \tag{4.183}$$

It must be recognised that the energy conservation proof for the FLIP mapping explained in Section 3.6.6.2 for the solid dynamic case is entirely based on cancelling the error of

Table 4.2: Relationships between MP-stored primary variables $\left(\boldsymbol{u}, p^{(f)} \right)$ and secondary variables. Superscript $(\bullet)^{mp}$ is omitted.

	Solid $\boldsymbol{u}$	fluid $p^{(f)}$	mixture
Volumes and related variables	$\Omega^{(sk)} = \omega^{(sk)}$	$\tilde{\omega}^{(f)}$	ω J n
Densities	$\rho_0^{(sk)} = \rho^{(sk)}$	$\rho^{(f)}$	ρ
Masses	$\bar{m}_0^{(sk)} = \bar{m}^{(sk)}$	$\bar{m}^{(f)}$	$\bar{m}$

the initial mapping (4.180). This deletion was possible thanks to the mass conservation inherited by the discretisation process. However, in a $\boldsymbol{u} - p^{(f)}$ formulation with one set of MPs, a porous material can exchange mass based on the drainage at boundary conditions (Eq. (4.128) and (4.129)). Therefore, it can deduced that the proof of energy conservation for the initial mapping (4.180) together with FLIP (4.183) does not stand for the $\boldsymbol{u} - p^{(f)}$ case.

Pressure mapping The G2P mapping of the fluid pressure is not only necessary before resetting the grid at the end of the step, but also required to update the different (secondary) variables stored at the MPs and necessary to compute the integrals appearing in Eqs. (4.156)-(4.164) at every iteration.

Since in some components of the Jacobian matrix (as in Eq. (B.46), which is part of the submatrix (B.44)) the nodal updated fluid pressure is moved at the MP location directly via the shape functions, this same procedure is adopted to map the fluid pressure in general

$$\left(p^{(f)} \right)^{mp} = \sum_{A}^{N^{nds}} \mathrm{n}_A^p(\boldsymbol{x}^{mp}) \, \mathrm{p}_A. \tag{4.184}$$

In this fashion, fluid pressure is moved consistently from grid nodes to the MPs at every iteration.

Once this procedure is completed via Eqs. (4.183) and (4.184), other (secondary) variables at the same location follow. A graphical representation of the flow of relationships between the primary and the secondary variables at MP locations is shown in Table 4.2. This table shows the updating order (illustrated by the arrows) of the different considered variables stemming from the converged values of displacements and pressures, and based on the hypothesis of incompressibility of the solid skeleton.

4.6.6 Implementation details

As for Section 3.6.7, a few points on implementing the illustrated $\boldsymbol{u} - p^{(f)}$ formulation and their relative collocation in the MPM algorithm are provided in the following text. A Material
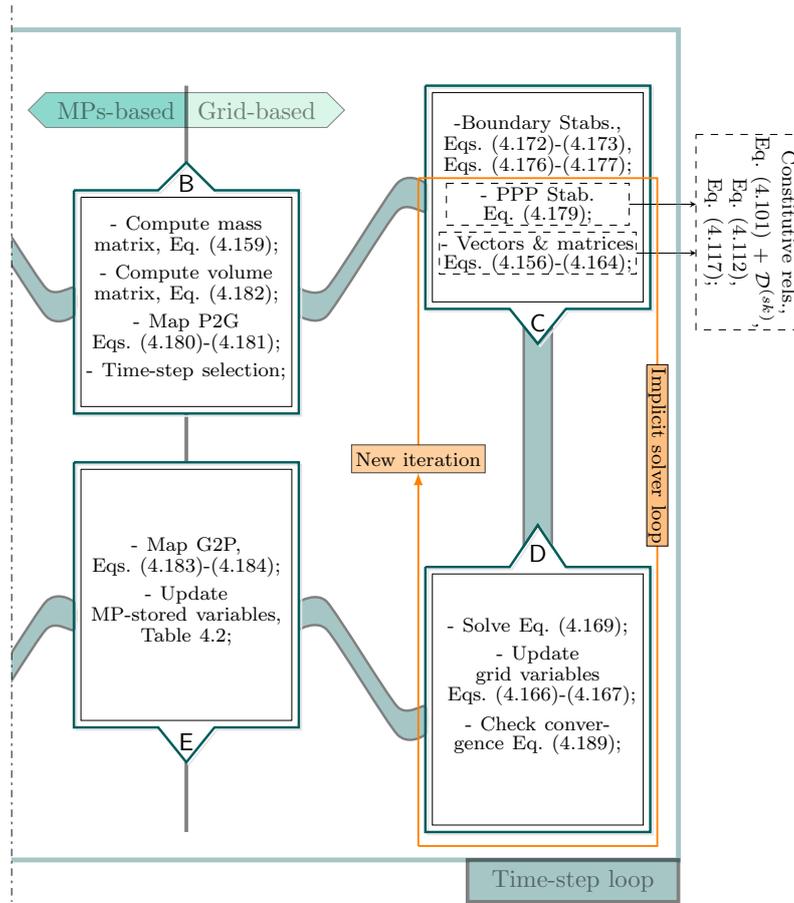
Figure 4.9: Main equations and relative use in the structure detailing the MPM algorithmic substeps. The considered equations are employed to design the $\boldsymbol{u} - p^{(f)}$ implicit dynamic code for porous materials and necessary stabilisations.

Point Learning Environment (AMPLE) (see Coombs and Augarde [68]) constitutes the basis code also for this formulation, modelling the dynamics of porous material. Again, only substeps from (B) to (E) are included in Figure 4.9 (compare with Figure 2.1 in Chapter 2).

Substep (B) starts with assembling the consistent mass matrix, Eq. (4.159) and the diagonal volume matrix, Eq. (4.182). These matrices are essential in performing the P2G mappings given by Eqs. (4.180) and (4.181) to reconstruct the velocity and fluid pressure at the grid nodes. In contrast to the solid dynamic case where the mass matrix can straightforwardly be used for the equilibrium equation, the presence of permeable boundaries allows drainage conditions. Hence, while the contribution of the solid constituent is constant during a simulation, the fluid phase can exchange mass. This implies that the mass matrix as a whole (or its fluid part, at least) must be recomputed at every iteration of the NR scheme. Moreover, the small cut issue can affect the P2G mappings, causing an ill-conditioned mass matrix or extremely low entries in the diagonal volume matrix. As both these situations are undesired, the stabilisations given by Eqs. (4.172) and (4.177) are added to the mass and volume matrices when required. The time-step selection, mostly non-adaptive during the simulation and based on a trial and error process, terminates substep (B).

Further stabilisations of the boundaries described by Eqs. (4.173) and (4.176), in addition to those that might already have been calculated (i.e., Eqs. (4.172) and (4.177)), are assembled out of the NR loop in substep (C). The NR iterations are then initiated using the following predictors:

$$\mathbf{a} = -\frac{1}{\tilde{\beta}\,\Delta t}\mathbf{v}_n, \qquad \mathbf{v} = \mathbf{v}_n\left(1 - \frac{\tilde{\gamma}}{\tilde{\beta}}\right); \tag{4.185}$$

$$\mathbf{p} = \mathbf{p}_n, \qquad\qquad\qquad \dot{\mathbf{p}} = \mathbf{0}, \tag{4.186}$$

which stem from the following assumptions relative to the displacement field

$$\begin{cases} \mathbf{v}_n \neq \mathbf{0}; \\ \mathbf{u}_n = \mathbf{a}_n = \Delta\mathbf{u} = \mathbf{0}, \end{cases} \tag{4.187}$$

and to the pressure field

$$\begin{cases} \mathbf{p}_n \neq \mathbf{0}; \\ \dot{\mathbf{p}}_n = \ddot{\mathbf{p}}_n = \Delta\mathbf{p} = \mathbf{0}. \end{cases} \tag{4.188}$$

It must be noted that the above assumptions are compatible with the mapping processes (4.180) and (4.181) and with the Newmark time discretisation Eqs. (4.166) and (4.167). These predictors are used to compute the Jacobian matrix in Eq. (4.170) (entries are given in detail in Appendix B.2) for the first iteration. The inversion of the Jacobian matrix in Eq. (4.169) (substep (D)), leads to a new grid solution for the displacement and pressure field. These nodal variables are updated according to the Newmark time scheme given in Eqs. (4.166) and (4.167). The updated values permit the computation of vectors and matrices from Eq. (4.156) to Eq. (4.164) appearing in the primary equations and the PPP stabilisation, Eq. (4.179), necessary to mitigate the LBB condition. To inform the primary equations, the constitutive relationships for the effective stress-strain relationship is provided by the free energy function of the improved Hencky material (Eq. (4.101)) for the elastic part and the dissipation function $\mathcal{D}^{(sk)}$ (see, for instance, Eq. (4.194)) for the plastic regime. The elastic and the elasto-plastic subroutines for this material are detailed in Appendix C. Other relationships, i.e., the fluid free energy function (4.112) and the Darcy-Weissbach law (4.117), complete the sets of the constitutive laws. NR convergence is checked (criterion in Eq. (4.189)) using the residuals (4.168) and the incremental solutions. If this criterion is not met, another NR iteration is taken.

The G2P mappings in substep (E) are performed via Eqs. (4.183) for the velocity and (4.184) for the pressure. These converged quantities, now stored at the MPs' location, are employed to update other variables, according to the flow illustrated in Table 4.2.

4.7   Validation tests and numerical analyses

This section reports four different examples, exploring different aspects of the formulation(s) explained above. Depending on the considered example, quasi-static or dynamic conditions are assumed.

The mono-dimensional consolidation Example 4.7.1 compares the numerical results with the analytical solution proposed by Terzaghi in the context of small strain theory. This example is straightforward but contains all the modelling difficulties for a porous material, the effects of different (spatial and time) discretisations and the validity of the PPP to address the inf-sup unstable spaces are explored.

Example 4.7.2 is intended to emphasise the invalidating effect of applying pressure Dirichlet BCs to the mesh even when the MPs' boundaries are non-conforming. While it can be expected that applying displacement Dirichlet BCs incorrectly would give some error which does not alter too significantly the simulations, this is not the case for the pressure BCs, as demonstrated in this case.

A variation of the Terzaghi mono-dimensional consolidation is explored in Example 4.7.3, where a column of elastic material is subjected to self-weight. It is shown that, even for moderate strains, the original Hencky material [39] (which is usually used to model finite strain elasto-plasticity, see Simo [69]) fails in considering the assumption of incompressibility of the solid phase*. The violation of this assumption leads to negative values of the Eulerian porosity and, simultaneously, to the violation of solid phase mass conservation. On the other hand, the improved Hencky material firstly proposed in Pretti *et al.* [30] (described by Eq. (4.101)) can reproduce physically meaningful results, retaining bounded values of the porosity and respecting the solid mass conservation.

The same comparison is carried out for the $3D$ flexible footing problem in the elasto-plastic regime, where it is shown the need to include the kinematic relationship (4.89) even for routine geotechnical analyses.

In all these analyses, the NR iterative process is considered as complete when the tolerance $tol = 10^{-11}$ at the $k-$th current iteration satisfies

$$\frac{|\mathbf{r}^{(k)}| \cdot |\delta\mathbf{x}^{(k)}|}{|\mathbf{r}^{(1)}| \cdot |\delta\mathbf{x}^{(1)}|} < tol, \tag{4.189}$$

with $|(\bullet)|$ being the absolute value of the vector $(\bullet)$. The vectors $\mathbf{r}$ and $\mathbf{x}$ have been defined by Eqs. (4.168) and (4.165), respectively. The above equation can be written in an expanded form as follows

$$\frac{|\mathbf{e}^{(k)}| \cdot |\delta\mathbf{u}^{(k)}| + |\mathbf{d}^{(k)}| \cdot |\delta\mathbf{p}^{(k)}|}{|\mathbf{e}^{(1)}| \cdot |\delta\mathbf{u}^{(1)}| + |\mathbf{d}^{(1)}| \cdot |\delta\mathbf{p}^{(1)}|} < tol. \tag{4.190}$$

It is expected that the contribution coming from $|\mathbf{e}^{(k)}| \cdot |\delta\mathbf{u}^{(k)}|$ weights usually less than the one provided by $|\mathbf{d}^{(k)}| \cdot |\delta\mathbf{p}^{(k)}|$ because of their different physical units. However, since both these contributions are expected to decrease with the iterations progressing, Eq. (4.189) can still operate as a global measure of the NR error.

4.7.1 Terzaghi mono-dimensional consolidation

Example scope The Terzaghi mono-dimensional problem [8] (illustrated in Figure 4.10) is considered to test results of the quasi-static $\boldsymbol{u} - p^{(f)}$ code against the analytical values

---

*The stress-strain relationship for an Hencky material can be obtained if, in Eq. (3.88), the stress on its LHS is the effective stress in lieu of the total one.

Table 4.3: Summary of the parameters considered in the analysis of the Terzaghi mono-dimensional consolidation.

Parameter Settings		
	$\bar{K}$, G	6, 3 MPa
Material Parameters	$K^{(f)}$	2.2 GPa
	$\kappa_0^{(f)}$	10^{-5} m s^{-1}
Geometry and	H	1 m
loading	w	100 kPa
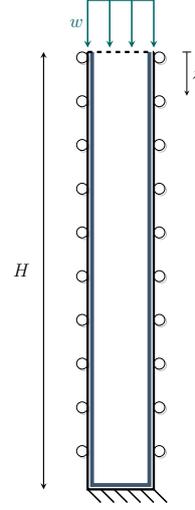
Figure 4.10: Illustration of the Terzaghi mono-dimensional problem. Dashed line means permeable surface.

of pressure. Furthermore, this example is designed to assess adapt time-step lengths and spatial discretisations necessary to approximate the analytical values. This solution for the pressure field is well-known (see, for instance, [8]) for the small strain regime and is given by the following equation:

$$P(Z, T) = \sum_{m=0}^{\infty} \frac{2}{M} \sin\left(MZ\right) \exp\left(-M^2\, T\right), \qquad \text{with} \quad M = \frac{\pi}{2}\left(2m + 1\right), \qquad (4.191)$$

where the non-dimensional quantities in the above equation are defined as

$$P \coloneqq \frac{p^{(f)}}{w}; \qquad Z \coloneqq \frac{z}{H}; \qquad T \coloneqq \frac{c_v}{H^2}\, t. \qquad (4.192)$$

Moreover, the relationship between the coefficient of consolidation and other hydro-mechanical parameters is given by

$$c_v = \frac{\left(\bar{K} + \frac{4}{3}G\right)}{\rho^{(f)}\, g}\, \kappa^{(f)}. \qquad (4.193)$$

Setup The example consists of a column of porous saturated material with a load applied at its top. The column is fixed at the bottom, free at the top, and roller displacement BCs are applied on both sides. All of the boundaries are impermeable, except for the top, where fluid flux is allowed, and zero (atmospheric) fluid pressures are applied as Dirichlet BCs directly on the grid. The column is loaded with the overburden w at the first time-step and hold constant for the rest of the analysis. The settlement of the column is allowed until the non-dimensional time T is equal to one.

The parameters considered in this simulation are listed in Table 4.3. Those mechanical elements are assigned so the column does not deform significantly (mimicking the small strain setting). At the same time, the hydraulic parameters are designed so that consolidation
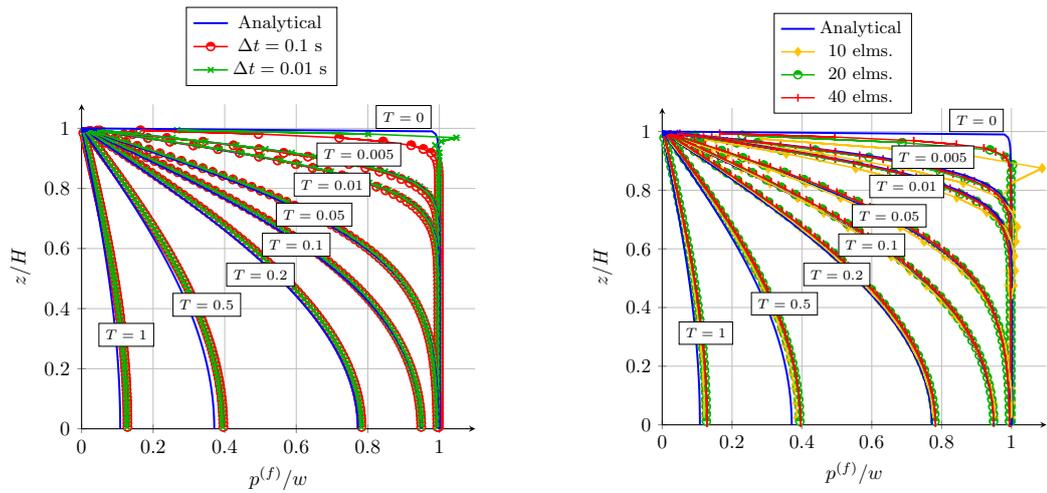
(a) Excess pore pressure isochrones for different time steps (40 vertical elements).

(b) Excess pore pressure isochrones for different vertical meshes ($\Delta t = \frac{T}{1000}$ s = 0.1 s).

Figure 4.11: Results for the Terzaghi mono-dimensional consolidation problem.

should occur in a reasonable (not instantaneous nor too long) time-scale. Further features of the simulation are as follows:

- GIMPM is considered;

- quasi-static code;

- gravity effects are neglected;

- PPP stabilisation is on, with $\tau = \frac{1}{2G}$;

- ghost stabilisations are off; and

- the material is the original Hencky;

Different time-step lenghts and spatial (vertical) discretisations are introduced and discussed below.

Results discussion   As it can be noticed from Figures 4.11(a) and 4.11(b), the numerical isochrones tend to slightly overestimate excess pore pressure over time if compared to the analytical solutions, especially from $T = 0.2$ to $T = 1$. This behaviour can be explained by two features influencing each other, which the code takes into account while Terzaghi's theory of consolidation does not. First, the code takes into account finite strain theory, and second, the hydraulic conductivity considered in the numerical analyses is dependent on the Eulerian porosity (via Eq. (4.118)), which, in turn, is a function of the finite volumetric strain (Eq. (4.89)). Hence, due to increasing strains allowed by the drainage at the top of the column, the hydraulic conductivity decreases, leading to a slower settlement modelled by numerical simulations.

As for different time-step lengths, numerical results in Figure 4.11(a) are very close to the analytical solution in both of the considered cases, with the time-steps being $\Delta t = 0.1$ in one analysis and $\Delta t = 0.01$ s in the other. The only exception lies in correspondence at zero non-dimensional time. The sharp gradient at the top of the column is more accurately followed along the $z-$coordinate by the lower time step discretisation. However, this advantage is balanced by the unnatural spike, which shows an increase in the fluid pressure, leading to a value bigger than the applied vertical stress. This peak is documented in the literature (see Preisig and Prevost [62]) and is due to the use of the PPP as stabilisation method which fails to completely remove pressure oscillation near the draining boundary. At the same time, however, PPP can smooth the gradient in pressure for the analysis with larger time-steps, even though it slightly underestimates the fluid pressure at the top MPs. This effectiveness of the PPP in the case of larger time-steps is not surprising as the PPP was originally (see Dohrmann and Bochev [63]) designed to stabilise Stokes problems, where the pressure is required to be only in $H^0(\omega)$ (conversely to the current problem, where ${}^h p \in H^1(\omega)$ is necessary, see Eq. (4.131)). As the numerical curves for the non-dimensional time $T = 0$ in Figure 4.11(a) are computed in correspondence of the first time-step (i.e., $T \approx 0$), the simulation with a smaller time-step is expected to show a sharper gradient of pressure. However, the PPP fails to stabilise the kind of components which include the gradients of pressure (such as $\mathrm{T}_{\nabla p, \nabla p}$). With larger time-steps and less sharp expected gradient, these components which incorporate the gradients of pressure play a lesser role, and the PPP stabilisation proves its effectiveness.

In regards to different spatial discretisations, it can be noticed from Figure 4.11(b) how the 10-element discretisation can be regarded as a poor choice: this behaviour is shown in particular for lower values of the non-dimensional times, between $T = 0$ and $T = 0.05$, where the gradient of the pressure is more pronounced. This trend becomes less and less significant over time. No relevant difference can be seen between the 20-element and 40-element discretisations, both of which can model the analytical behaviour. The difference between these two numerical analyses and the analytical solution at $T = 0$ can be attributed to the selected time-discretisation (equal to $\Delta t = 0.1$ s), as above detailed.

4.7.2 Strip flexible footing under harmonic loading conditions

Example scope   The scope of this example is to investigate how the hypothesis, introduced in Section 4.6.2.1 of conforming pressure BCs affects the results of the simulation.

Setup   A fully saturated soil undergoes a vertical loading whose magnitude varies with the time according to the equation shown in Figure 4.12, which is $w(t) = 3 - 3\cos(100\,t)$ MPa (with the time t expressed in seconds).

As illustrated by Figure 4.12, only half of the problem has been simulated. The bottom and lateral boundaries are subject to roller boundary conditions, and are considered as impermeable surfaces. On the other hand, the top boundary is subdivided into two parts: the former, whose length is 1 m, is impermeable and subject to the vertical loading, whereas

$$w(t) = 3 - 3\cos(100\,t) \text{ MPa}$$
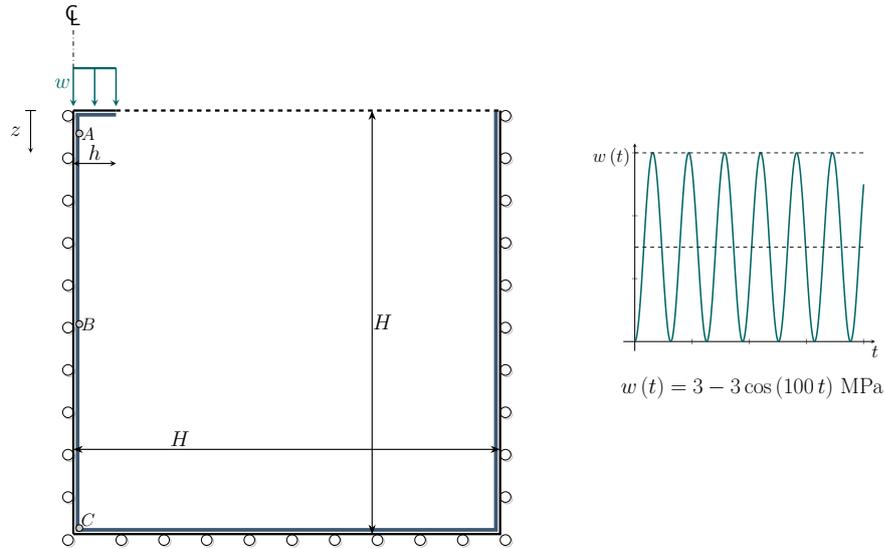
Figure 4.12: Planar section of half of the footing problem: geometry, boundary conditions and applied time-dependent load. Dashed line means permeable surface.

Table 4.4: Summary of the parameters considered in the analysis of the elastic flexible strip.

Parameter Settings		
	$\bar{K}$, G	12.13, 5.6 MPa
	$K^{(f)}$	2.2 GPa
Material Parameters	$\rho^{(sk)}$, $\rho_0^{(f)}$	2500, 1000 kg/m^3
	n_0	0.33
	$\kappa_0^{(f)}$	10^{-10} m s^{-1}
Geometry,	H	10 m
and Timings	h	1 m
	T	0.4 s
	h_x, h_y	≈ 0.33 m
Analysis Parameters	$\dagger mmp$	4
	Δt	0.004 s

$\dagger$ mmp is the number of MPs per direction per element.

the latter, which is the remaining part, is free to move and permeable. Moreover, zero fluid pressure is applied as a Dirichlet boundary condition over this latter top part. This atmospheric pressure has been applied on the fixed grid, which does not follow the moving MPs surface. The initial conditions considered (displacements, velocities and fluid pressure) are all null. For this analysis, parameters are given in Table 4.4. While these parameters are arbitrarily chosen (with the exclusion of the fluid bulk modulus, which represents water), the mechanical values are selected so that the column undergoes finite deformation and low values of hydraulic conductivity are considered to model nearly undrained conditions. These choices are designed to violate the hypothesis of conforming BCs so that the extent
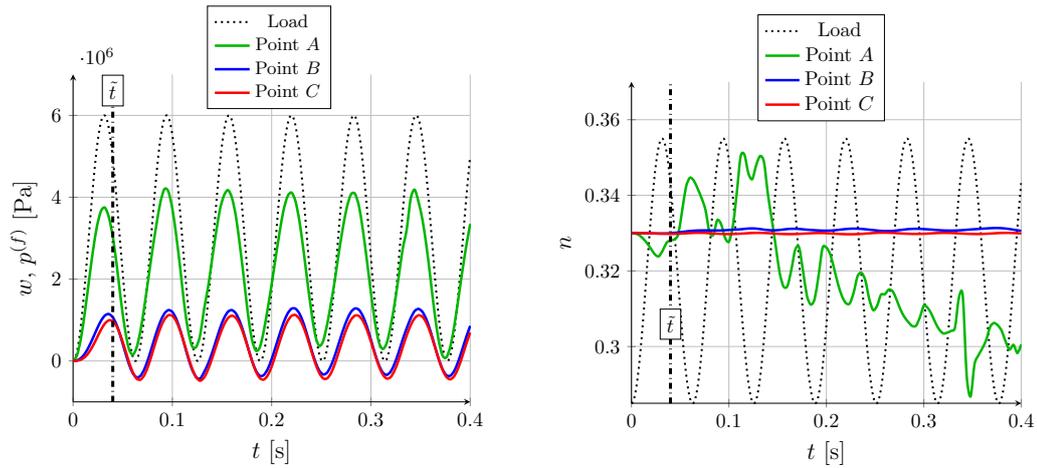
(a) Time evolution comparisons of the applied load w at the top surface and the excess pore pressure p^f evaluated at points A, B and C.

(b) Time evolution of the Eulerian porosity at points A, B and C. The qualitative behaviour of the applied load is shown for the sake of greater clarity.

Figure 4.13: Time behaviour of the MPs initially located at A ($z = 0.5417$ m), B ($z = 5.0417$ m) and C ($z = 9.9583$ m).

of this assumption can be tested. Other features of the simulation are as follows:

- GIMPM is used;

- the code includes inertia effects. In particular, the chosen coefficients for the Newmark time scheme (see Section 4.6.2.2) are for the so-called *trapezoidal rule*, these being $\tilde{\gamma} = \frac{1}{2}$, $\tilde{\beta} = \frac{1}{4}$;

- gravity is not considered;

- PPP stabilisation is on, with $\tau = \frac{1}{2G}$;

- ghost stabilisations are off; and

- the material is the original Hencky.

Results discussion   The graphs in Figure 4.13 illustrate the behaviour of the fluid pressure and the Eulerian porosity for the considered points in Figure 4.12 over time. In both of the graphs, the sinusoidal time evolution of the load has been reported to facilitate reading.

If Figure 4.13(a) is considered, it can be seen how the peaks of the fluid pressure do not coincide with those of the applied load, but they are slightly delayed with respect to the loading curve, in particular, for points B and C. Moreover, it can be seen how point A, which is the closest to a draining surface, does not experience negative (traction) fluid pressure, whereas points B and C exhibit negative fluid pressures. The reason for this non-physical values at the points B and C stems from having applied zero pressures as BCs on the fixed grid, without keeping track of the real surface of the body. This can be appreciated in Figure 4.14, where it can be seen how negative values of pressure on a few
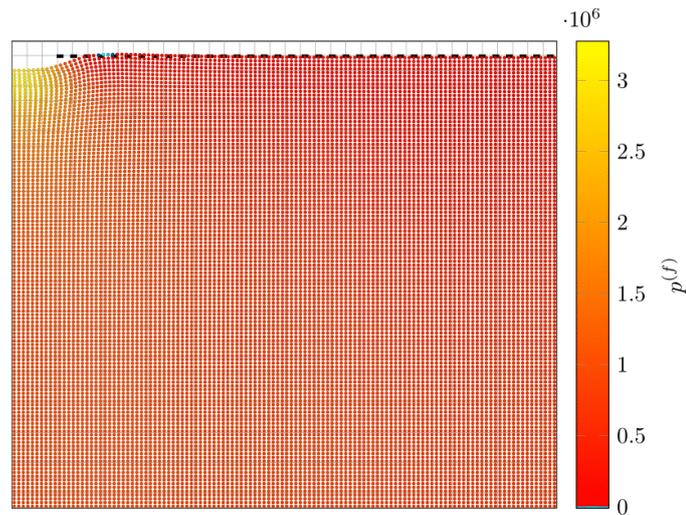
(a) Fluid pressure contours at the time $\tilde{t} = 0.04$ s.
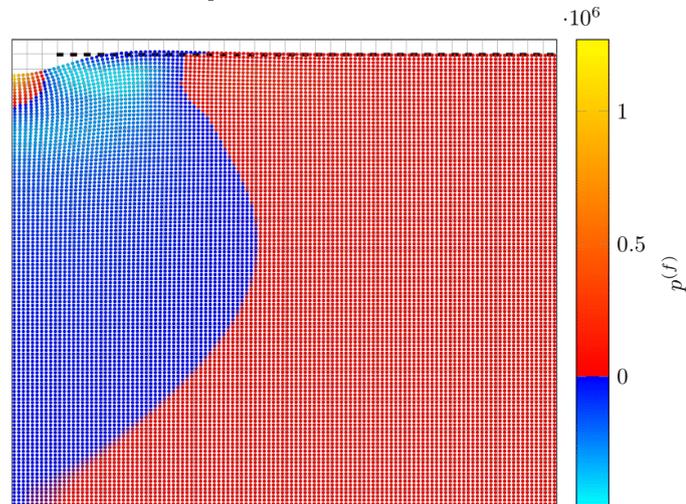
(b) Fluid pressure contours at the time $t = 0.052$ s.

Figure 4.14: Fluid pressure contours for the strip footing under harmonic loading. Dashed line means permeable surface.

MPs (illustrated in Figure 4.14(a) for the time $\tilde{t} = 0.04$ s) spread, within a few time-steps, to a very large portion of the material (Figure 4.14(b) for $t = 0.052$ s). The code is not equipped to reproduce this situation, which would physically imply the presence of air, i.e., a second fluid in the void volumes of the material. As such, results should not be considered after time $\tilde{t}$.

Considering only the time before $\tilde{t}$ in Figure 4.13(b), it can be seen that points B and C practically experience undrained conditions, while the change in porosity of A follows the periodic behaviour of the load. This different behaviour can be explained with the proximity (for the point A) or the distance (for points B and C) from the draining surface, together with low value of hydraulic conductivity $\kappa^{(f)}$.

4.7.3 Elastic column under self-weight

Table 4.5: Summary of the parameters considered in the analyses of the elastic column under self-weight.

Parameter Settings		
	$\ddagger\tilde{K}, \bar{K}$	500 kPa
	G	300 kPa
Material	K^f	2.2 GPa
Parameters	$\rho^{(sk)}$	2650 kg m^{-3}
	$\rho_0^{(f)}$	1000 kg m^{-3}
	$\kappa_0^{(f)}$	10^{-5} m s^{-1}
Geometry,	H	1 m
Loading	f_z	1000 m s^{-2}
and Timings	T	3000 s
Analysis	h_z, h_y	0.05 m
Parameters	$\S mmp$	4
	Δt	0.04 s

‡ $\tilde{K}$ refers to the parameter used in Eq. (4.101), while $\bar{K}$ is the bulk modulus of the Hencky material.
§ mmp is the number of MPs per direction per element.
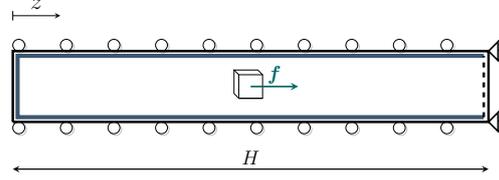
Figure 4.15: Illustration of the elastic column under self-weight. Dashed line means permeable surface.
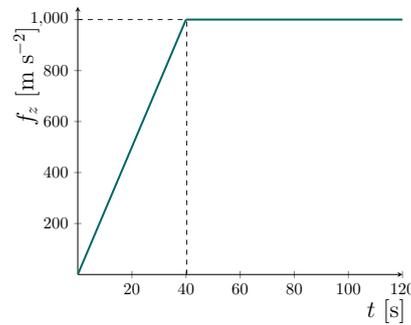
Figure 4.16: Applied load time variation.

Example scope This mono-dimensional numerical test (illustrated in Figure 4.15) is designed to emphasise the differences between the original and improved Hencky material in the case of elastic finite deformations.

Setup The column consists of an elastic material subjected to gravitational load whose time-dependency is given in Figure 4.16. Roller boundary conditions are applied to the upper and lower sides of the column, with the right surface being fixed and the left surface free. In addition, the entire external surface is impermeable, except for the right part, where zero atmospheric pressure is prescribed. The two simulations are run using a quasi-static formulation for $3,000$ pseudo time-steps, and the gravitational load is increased linearly for the first $1,000$ steps and kept constant for the remaining time. The selected gravitation* reaches $f = 1,000$ m s^{-2}. Other (arbitrarily chosen) parameters of the problems are given in Table 4.5†, and the following have been used for the analyses:

---

*The reason for selecting such a substantial gravitational load lies in the drastic reduction of simulation time, especially in correspondence with low hydraulic conductivity values such as the one considered in this example (see Table 4.5). The difference between the modified and the original Hencky model concerns the effective stress, whose increase occurs with the sample draining, i.e., with the time passing.

†The bulk modulus for the linear Hencky material $\bar{K}$ has been set so that it matches the initial elastic tangent bulk modulus $K_0^{e,tan}$ defined by Eq. (C.7) and has been computed in correspondence of initial porosity n_0 and zero elastic volumetric strain, i.e., $\tilde{K} = \bar{K} n_0 = 5 \cdot 10^5$ Pa. In this fashion, the stiffness of the two materials is initially the same.

(a) Porosity at the right-hand side of the column.

(b) Hydraulic conductivity at the right-hand side of the column.

(c) Horizontal displacement at the left-hand side of the column.

(d) Fluid pressure at the left-hand side of the column.
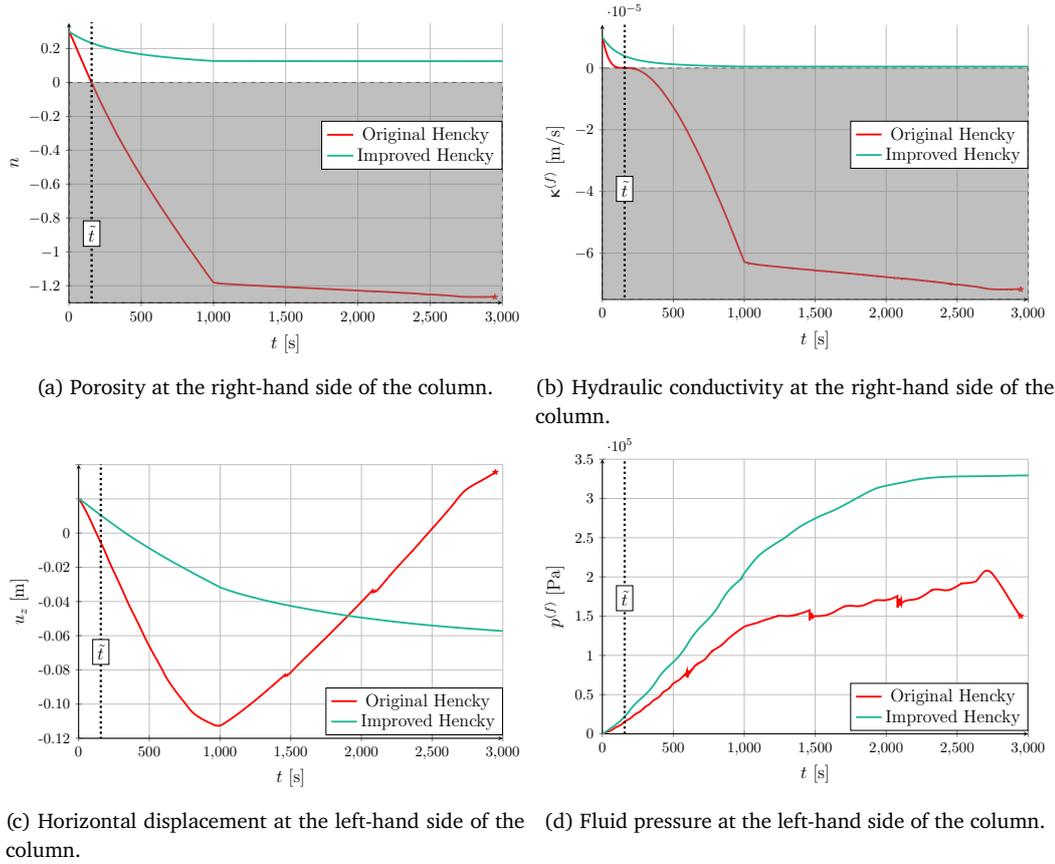
Figure 4.17: Time-plots of different quantities calculated at the right-hand side (top row) and left-hand side (bottom row) of the column under self-weight. Grey-shaded areas indicate the physically unfeasible regions. Figure reproduced from Pretti *et al.* [30].

- GIMPM is used;

- PPP stabilisation is on, with $\tau = \frac{1}{2G}$; and

- ghost stabilisations are off;

Results discussion   Figure 4.17 shows time-plots for different quantities comparing the two simulations. In particular, for the analysis where the original Hencky material is considered, time $\tilde{t}$ (corresponding to the 157th pseudo time-step) highlights the watershed between positive and negative values of the Eulerian porosity (see Figure 4.17(a), but also Figure 4.18(a)). From a physical perspective, it is clear that the simulation using the original Hencky material cannot be regarded as valid after $\tilde{t}$. Nonetheless, it can be noticed from Figures 4.17(c) and 4.17(d) how the values of primary variables, i.e., displacement $\boldsymbol{u}$ and fluid pressure p^f, differ consistently even for pseudo time-steps lower than $\tilde{t}$ and far from where more significant values of the deformation gradient are expected between the linear Hencky material and the current model. Thus, it can be seen how imposing constrained porosity values affects the entire simulation, leading to very different behaviour even for those parts of the body where such constraints are not expected to be violated (such as the

(a) Eulerian porosity for the linear Hencky material, $\tilde{t} = 157^{th}$ time-step.

(b) Eulerian porosity for the improved Hencky material, $\tilde{t} = 157^{th}$ time-step.

(c) Eulerian porosity for the linear Hencky material, final time-step.

(d) Eulerian porosity for the improved Hencky material, final time-step.
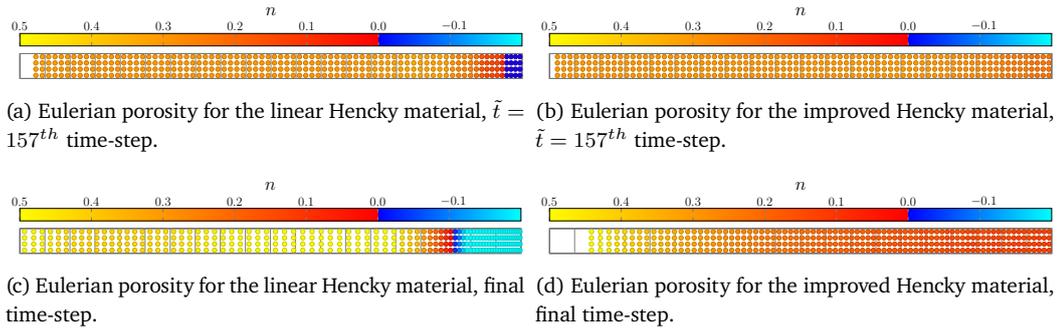
Figure 4.18: Deflection shapes of the columns at $\tilde{t} = 157^{th}$ time-step and the end of each simulation.

left-hand side of the column). This idea is further illustrated in Figures 4.17(c) and 4.18(c): when the simulation continues even for the original Hencky material, it begins to swell to the left (in a direction opposite to gravity). This leads the simulation itself to abort at the 2948th pseudo time-step. On the contrary, the new model behaves in accordance with the bounded porosity values and results in smoother transitions on both sides of the column, as highlighted by all of the time-plots of the different quantities in Figure 4.17.

A more uniform deformation state also emerges when comparing the Eulerian porosity values (and thus the Jacobian via Eq. (4.89)) between Figures 4.18(a)-4.18(b) at the 157th pseudo time-step, and Figures 4.18(c)-4.18(d) at the end of the simulations.

4.7.4 3D flexible footing

Example scope This example (illustrated in Figure 4.19) aims to assess the difference between the original and improved Hencky material in the case of a $3D$ routine geotechnical analysis (flexible footing) in the finite elasto-plastic regime for different material mechanical parameters.

Setup Since the problem presents two symmetry planes, only a quarter of the whole setup (as represented in Figure 4.19) is considered. All planes defining the soil boundaries have rollers and no-flux conditions applied, except for the top surface, which is free to move and with zero atmospheric pressure applied. The simulations are run considering quasi-static conditions, where the load w linearly ramps from zero to a value of $5 \cdot 10^5$ Pa during 10 pseudo time-steps (see Figure 4.20), while gravity effects are neglected. Eight MPs per cell initially populate each mesh element.

The chosen dissipation function is that proposed by Collins and Hilder [70] (see Coombs and Crouch [71] for derivation and implementation details) and it describes a *two-parameter*

Table 4.6: Summary of the shared parameters considered in the analyses of the elasto-plastic flexible footing under large deformations.

Parameter Setting		
Material Parameters	G	$28 \cdot 10^6$ Pa
	K^f	$2.2 \cdot 10^9$ Pa
	$\rho^{(sk)}, \rho_0^{(f)}$	$2650, 1000$ kg m^{-3}
	$\kappa_0^{(f)}$	$1 \cdot 10^{-2}$ m s^{-1}
	p_c	‡$250 \cdot 10^5$ Pa
	α, γ, M	$0.3, 0.9, 0.964$
Geometry, Loading and Timings	L_x, L_y, L_z	$10, 14, 10$ m
	a, l_y	$2.5, 10$ m
	w	$5 \cdot 10^5$ Pa
	T	10 s
Analysis Parameters	h_x, h_y, h_z	$0.5, 0.5, 0.5$ m
	†mmp	2
	Δt	1 s

‡ Geotechnical convention on signs is assumed.

† mmp is the number of MPs per direction per element.

Figure 4.19: Initial setup for the 3D flexible footing problem. Figure reproduced from Pretti *et al.* [30].

family of critical state [72]*

$$\mathcal{D}^{sk} = \sqrt{\left(d_v^p \, \tilde{A}\right)^2 + \left(d_\gamma^p \, \tilde{B}\right)^2} \geq 0, \qquad (4.194)$$

where the invariants are given by

$$d_v \coloneqq d_{ij}\delta_{ij} \qquad (4.195)$$

$$d_\gamma \coloneqq \sqrt{\left(d_{ij} - \frac{d_v}{3}\delta_{ij}\right) \mathbin{:} \left(d_{ij} - \frac{d_v}{3}\delta_{ij}\right)}, \qquad (4.196)$$

while the pressure-dependent parameters $\tilde{A}$ and $\tilde{B}$ are defined as

$$\tilde{A} \coloneqq (1 - \gamma)\,p' + \frac{\gamma}{2}p_c \qquad \tilde{B} \coloneqq M\left((1 - \alpha)\,p' + \alpha\frac{\gamma}{2}p_c\right). \qquad (4.197)$$

For the case under consideration, p_c is a constant prescribed value, while M indicates the slope of the critical state line. Following the procedure described in Section 4.5.1.2, it can be seen that the flow rule is non-associated for the improved Hencky material, due to the non-zero value of the shift stress (see definition (4.104)). On the other hand, the original Hencky material exhibits an associated flow rule.

Four simulations have been run: two analyses consider the elastic part described by the original linear Hencky material, while the other two use the improved Hencky material defined

---

*Since the dissipation function proposed by Collins and Hilder [70] comes from the field of geotechnics, it follows its convention, i.e., compressive state are denoted as positive. However, the current work employed the notation more commonly used in mechanics, i.e., compressive states are negative. As such, volumetric components appearing in Eq. (4.194) and related equations must change sign to be in compliance with the rest of this work.

Figure 4.20: Applied load time variation.

Table 4.7: Summary of the different parameters considered in the analyses of the elasto-plastic flexible footing under large deformations.

	Case (A)	Case (B)
$K_0^{e,tan} / \bar{K}$ [Pa]	$30 \cdot 10^6$	$50 \cdot 10^6$
n_0	0.2	0.1
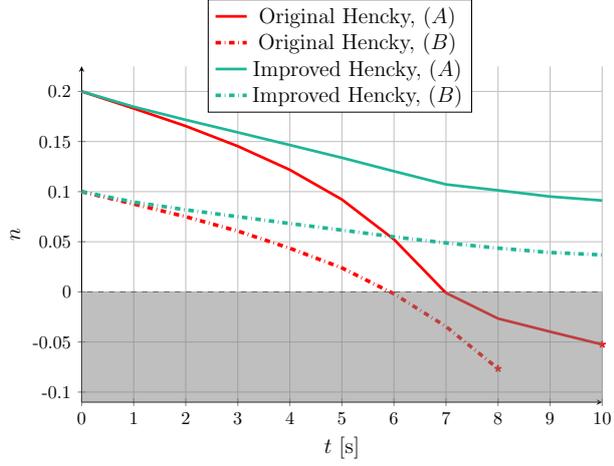
Figure 4.21: Time-plot of the porosity computed at the MP initially located at $(0.125, 9.625, 0.125)$ m for the different simulations. Figure reproduced from Pretti *et al.* [30].

by Eq. (4.101). All of the shared parameters of the simulations are listed in Table 4.6. These parameters are arbitrarily chosen, with the exception of those related to plastic behaviour, which are taken from Coombs and Crouch [71]. Different initial values of the Eulerian porosity and bulk modulus (or its tangent value) are enumerated in Table 4.7. Nonetheless, the reason why initial values of these parameters are considered lies in the direct implication of the Eulerian porosity and the bulk modulus on the volumetric behaviour. Moreover, as reported in Table 4.6, the hydraulic conductivity value is relatively low to quickly dissipate the excess pore pressure. Other choices used are:

- GIMPM is used;

- PPP stabilisation is on, with $\tau = \frac{1}{2G}$; and

- ghost stabilisations are on, with $\gamma^K = \bar{E} \cdot 10^{-6*}$, $\gamma^{M^{(f)}} = 0$, and $\gamma^{K^{(f)}} = \kappa_0^{(f)} \cdot 10^{-3}$.

Results discussion As can be seen from Figure 4.21, all analyses considering the original Hencky material fail within the 10th pseudo time-step, regardless of their initial value of Eulerian porosity or bulk modulus. Moreover, Figure 4.21 emphasises a strict correlation between negative values of the Eulerian porosity and failure of the analyses, with this latter phenomenon occurring a few time-steps after the former. The improved Hencky material permits all the considered simulations to be completed.

Contours of the Eulerian porosity plotted in Figure 4.22 show where the negative values of the original Hencky material occur, i.e., in the proximity of the applied load. Excluding a load disturbance zone (corresponding to the upper rows of MPs), the zone below the

---

*The constant Young's modulus used for stabilisation purposes is defined as $\bar{E} = \frac{9K\,G}{3K+G}$, with $K = \bar{K}$ for the original Hencky material, and $K = K_0^{e,tan}$ for its improved counterpart.

(a) Original Hencky, Analysis (A).

(b) Improved Hencky, Analysis (A).

(c) Original Hencky, Analysis (B).

(d) Improved Hencky, Analysis (B).
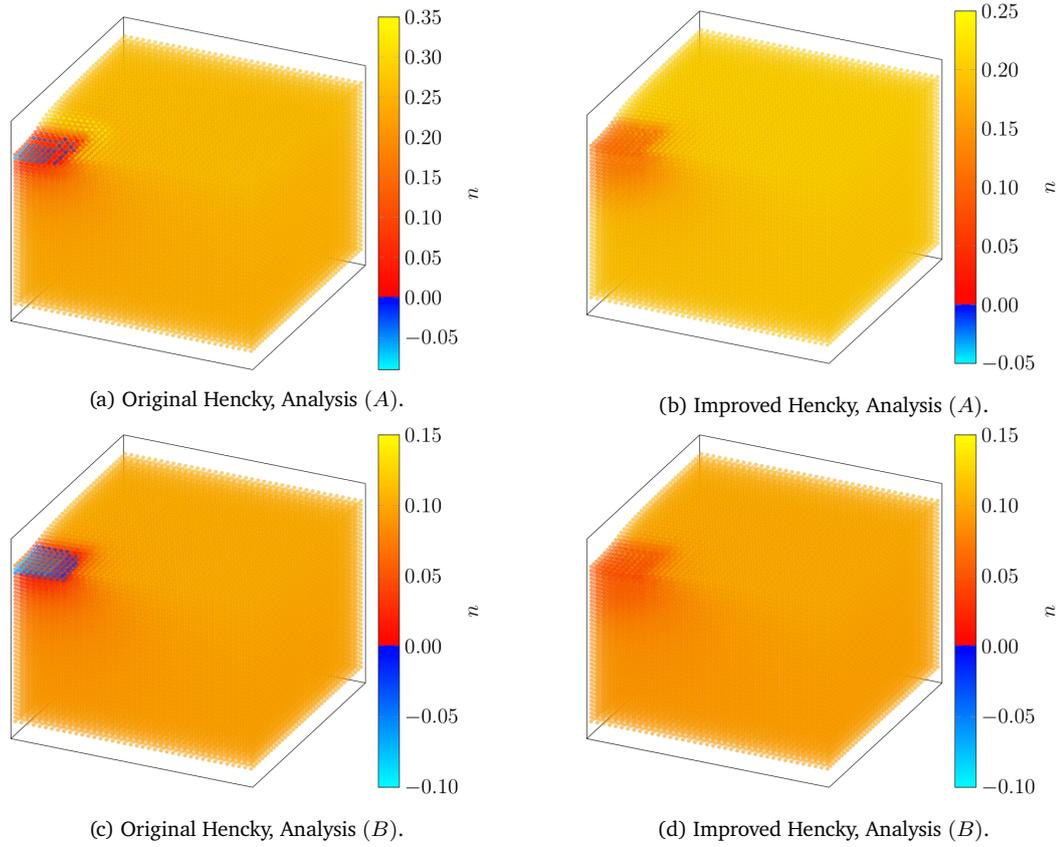
Figure 4.22: Contours of the Eulerian porosity values applied to the final deflection shapes of the flexible footing problem. Figure reproduced from Pretti *et al.* [30].

foundation is the area where, as expected, inequality (4.92a) is violated by the original Hencky material.

4.8 Observations

From the above examples and the narrative of this chapter describing porous materials and their application to the MPM, some conclusions can be drawn. In general, the main novelty of this chapter is given by the improved Hencky material, i.e., a material designed to comply with the conservation of the solid mass constituent when Terzaghi effective stress decomposition is considered. While the contribution of this novelty has few implications away from the limit dictated by (4.92a), it becomes essential when approaching such deformation states. Moreover, since this free energy function only affects constitutive modelling, its application is broader than the MPM.

Other novel contributions regard some aspects of the $\boldsymbol{u} - p^{(f)}$ formulations. The required boundary stabilisations outlined in Section 4.6.4.1 have been adapted from the cut-FEM community, but their role is fundamental to guarantee the stability of MPM numerical simulations. The adopted PPP stabilisation to account for the LBB conditions has shown some limitations when the pressure discretised field is $H^1(\omega)$. To the authors' best know-

ledge, this information is not available in the literature concerning the MPM, whereas it is for FEMs. Section 4.6.5 provides only some observations on the mappings required in a $\boldsymbol{u} - p^{(f)}$ formulation, which nonetheless constitute a small step forward compared to the existing literature. Some other discussions about applying non-conforming Neumann and Dirichlet BCs and the necessity to apply these consistently are made below.

As for the examples, Example 4.7.1 shows how the approach can match results in compliance with the literature for the case of small strains. However, when finite strains start to play a dominant role, the picture becomes more complicated, as seen in Example 4.7.2. Non-conforming atmospheric pressure can play an important role in the results, undermining the whole simulation when approximately applied on the grid. *Ergo*, techniques to apply non-conforming Dirichlet BCs are crucial in MPM simulations representing a porous material.

While this is not explored in this chapter, $\boldsymbol{u} - p^{(f)}$ formulations hide another issue. As illustrated by Figure 4.7 and reported in Eqs. (4.124) and (4.125), tractions can be applied on a portion of the body which might overlap with a portion where fluid pressures are prescribed, i.e., $\gamma^{\bar{t}} \cap \gamma^{\bar{p}} \neq \emptyset$ (this is the case, for instance, of top boundary of the Terzaghi column in Example 4.7.1). In this case, using the Cauchy stress theorem Eq. (4.32) and Terzaghi effective stress decomposition (4.100), it follows that

$$
\int_{\gamma^{\bar{t}} \cap \gamma^{\bar{p}}} \boldsymbol{w} \cdot \bar{\boldsymbol{t}} \, da = \int_{\gamma^{\bar{t}} \cap \gamma^{\bar{p}}} \boldsymbol{w} \cdot (\boldsymbol{\sigma} \cdot \boldsymbol{n}) \, da = \int_{\gamma^{\bar{t}} \cap \gamma^{\bar{p}}} \boldsymbol{w} \cdot \left(\boldsymbol{\sigma}' - \underbrace{p^{(f)} \, \boldsymbol{I}^{(2)}}_{= \bar{p}} \right) \cdot \boldsymbol{n} \, da, \qquad (4.198)
$$

i.e., fluid pressure (which is a Dirichlet BC) can appear as part of a natural BC for the mixture balance of rate of linear momentum Eq. (4.154), while being an essential BC for the fluid mass conservation Eq. (4.155)*. In the case where these BCs are applied to non-conforming boundaries, the contemporary presence of pressure as (part of) a natural and a essential BC requires that surface tracking and application of these BCs must be performed consistently. To the author's best knowledge, no technique satisfying this criterion is available in the literature for both kinds of BCs for the MPM†.

The use of different time-steps in Example 4.7.1 highlighted that, when sharp pressure gradients are expected, the PPP fails to stabilise the proposed formulation. Stabilisation

---

*At this stage, it is necessary to distinguish among the terminologies employed above. *Dirichlet* BCs specify the values of the solution field (primary unknown) along the domain's boundary. Conversely, *Neumann* BCs select the values of the spatial derivatives of the solution field applied at the domain's boundary. In irreducible formulations (as the displacement-based outlined in Section 3.6), Dirichlet BCs are satisfied in the variational statement by selecting appropriate space solutions and, as such, they are defined as *essential*. In these same irreducible formulations, Neumann BCs are satisfied by the weak form after applying the fundamental theorem of calculus, i.e., they appear *naturally* in the formulation. However, in mixed formulations (such as the $\boldsymbol{u} - p^{(f)}$), things can be more blurred. Therefore, the definitions of Dirichlet and Neumann BCs are kept as the ones above: the former specify the value of the solution and the latter the values of the derivatives of the solution. However, we define as essential BCs the ones satisfied by the variational statement via adequate space solutions, while natural BCs appear in the weak statement of equations. This nomenclature is consistent with the literature (see, for instance, Brezzi and Fortin [59]).

†Yamaguchi *et al.*'s [73] technique could resolve the above explained problem, even though using a different approach. This work adopts the Nitsche's method, requiring that also Dirichlet BCs are satisfied by the weak formulation.

alternatives, such as the fluid pressure Laplacian (see Brezzi and Pitkäranta [74]) or Variational multiscale methods (whose general idea was proposed by Hughes [75]), are available in the FEM literature, but, to the author's knowledge, have not been extended to the MPM. Same conclusions can be drawn for stable elements belonging to the Taylor-Hood family (see, for instance, Brezzi and Fortin [59]), whose application to the MPM has not been explored yet.

As indicated by Examples 4.7.3 and 4.7.4, the improved Hencky material Eq. 4.101 appears to be necessary for simulations whose strains can be close to the incompressible limit defined by (4.92a). As in the most general scenario it is not possible to establish *a priori* whether (4.92a) is satisfied or not, considering a material such as the improved Hencky or implementing strategies to satisfy the porosity constraints seems to be ineluctable. This aspect becomes quite urgent especially for the MPM, which is designed to deal with large deformation mechanics.

Moreover, to be sure that the MPM discretisation can mimic the conservation properties of its continuous underlying formulation, a rigorous assessment of the mappings (as the one proposed in Section 3.6.6 for the solid dynamic case) should be explored.

Chapter summary

Following a structure similar to Chapter 3, this chapter has discussed the underlying mechanics necessary to describe a porous body undergoing finite strain. The difference between Eulerian and Lagrangian descriptions has appeared necessary to describe the kinematics of such a material. The presence of two phases has consequences on the principles of mass conservation, balance of rate of momenta, and the principles of thermodynamics. The different phases have demanded different constitutive relationships. Among these, modifying the constitutive relationship describing the effective behaviour has been motivated by the mass conservation of the solid phase. The discretisation steps and the necessary stabilisations for a $\boldsymbol{u} - p^{(f)}$ MPM algorithm have been detailed, while a brief discussion has been reserved for the mapping processes. Numerical tests have shown the ability of the algorithm to reproduce analytical solutions and uncovered undesirable situations when pressure BCs are treated as conforming even when they are not. A discussion on using the PPP as a stabiliser for the poromechanical equations has been covered. The advancement of the newly introduced improved Hencky material over the standard one has been confirmed confirmed by two numerical tests.

References

[1] O. Coussy, *Poromechanics*. John Wiley & Sons, 2004.

[2] R. De Boer and W. Ehlers, 'A historical review of the formulation of porous media theories,' *Acta Mechanica*, vol. 74, no. 1-4, pp. 1–8, 1988.

[3] M. A. Biot, 'General theory of three-dimensional consolidation,' *Journal of Applied Physics*, vol. 12, no. 2, pp. 155–164, 1941.

[4] M. A. Biot, 'Theory of propagation of elastic waves in a fluid-saturated porous solid. II. higher frequency range,' *The Journal of the Acoustical Society of America*, vol. 28, no. 2, pp. 179–191, 1956.

[5] M. A. Biot, 'Mechanics of deformation and acoustic propagation in porous media,' *Journal of Applied Physics*, vol. 33, no. 4, pp. 1482–1498, 1962.

[6] K. von Terzaghi, 'Die berechnug der durchlassigkeit des tones aus dem verlauf der hydromechanischen spannungserscheinungen,' *Sitzungsber. Akad. Wiss.(Wien). Math.-Naturwiss. Kl., Abt. Iia*, vol. 132, pp. 125–138, 1923.

[7] K. Terzaghi, *Erdbaumechanik auf bodenphysikalischer Grundlage*. F. Deuticke, 1925.

[8] K. Terzaghi, *Theoretical soil mechanics*. John Wiley & sons, New York, 1943.

[9] N. Castelletto, G. Gambolati and P. Teatini, 'A coupled MFE poromechanical model of a large-scale load experiment at the coastland of venice,' *Computational Geosciences*, vol. 19, pp. 17–29, 2015.

[10] A. Karrech, M. Attar, M. Elchalakani, H. Basarir, F. Abbassi and A. Seibi, 'The poromechanics of massive fluid injection in natural environments,' in *ISRM 1st International Conference on Advances in Rock Mechanics-TuniRock 2018*, OnePetro, 2018, pp. 17–22.

[11] M. Hettema, P. Schutjens, B. Verboom and H. Gussinklo, 'Production-induced compaction of a sandstone reservoir: The strong influence of stress path,' *SPE Reservoir Evaluation & Engineering*, vol. 3, no. 04, pp. 342–347, 2000.

[12] A. Karrech, T. Poulet and K. Regenauer-Lieb, 'Poromechanics of saturated media based on the logarithmic finite strain,' *Mechanics of Materials*, vol. 51, pp. 118–136, 2012.

[13] S. C. Cowin, 'Bone poroelasticity,' *Journal of Biomechanics*, vol. 32, no. 3, pp. 217–238, 1999.

[14] A. Selvadurai and A. Suvorov, 'Coupled hydro-mechanical effects in a poro-hyperelastic material,' *Journal of the Mechanics and Physics of Solids*, vol. 91, pp. 311–333, 2016.

[15] Y. Zhang, 'Mechanics of adsorption–deformation coupling in porous media,' *Journal of the Mechanics and Physics of Solids*, vol. 114, pp. 31–54, 2018.

[16] R. M. Bowen, 'Compressible porous media models by use of the theory of mixtures,' *International Journal of Engineering Science*, vol. 20, no. 6, pp. 697–735, 1982.

[17] O. C. Zienkiewicz, A. Chan, M. Pastor, B. Schrefler and T. Shiomi, *Computational Geomechanics*. Citeseer, 1999, vol. 613.

[18] R. W. Lewis and B. Schrefler, *The finite element method in the static and dynamic deformation and consolidation of porous media*. John Wiley & Sons, 1998.

[19] H. A. Nøst, 'Iteratively coupled implicit dynamic MPM-FVM,' 2019.

[20] X. Zheng, F. Pisanò, P. J. Vardon and M. A. Hicks, 'Fully implicit, stabilised MPM simulation of large-deformation problems in two-phase elastoplastic geomaterials,' *Computers and Geotechnics*, vol. 147, p. 104 771, 2022.

[21] X. Zheng, S. Wang, F. Yang and J. Yang, 'Material point method simulation of hydro-mechanical behaviour in two-phase porous geomaterials: A state-of-the-art review,' *Journal of Rock Mechanics and Geotechnical Engineering*, 2023.

[22] O. Coussy, L. Dormieux and E. Detournay, 'From mixture theory to Biot's approach for porous media,' *International Journal of Solids and Structures*, vol. 35, no. 34-35, pp. 4619–4635, 1998.

[23] Y. Cengel and J. Cimbala, *Ebook: Fluid mechanics fundamentals and applications (si units)*. McGraw Hill, 2013.

[24] J. E. Marsden and T. J. Hughes, *Mathematical foundations of elasticity*. Courier Corporation, 1994.

[25] J. F. Thorpe, 'On the momentum theorem for a continuous system of variable mass,' *American Journal of Physics*, vol. 30, no. 9, pp. 637–640, 1962.

[26] C. K. Batchelor and G. K. Batchelor, *An introduction to fluid dynamics*. Cambridge University Press, 1967.

[27] G. T. Houlsby and A. M. Puzrin, *Principles of hyperplasticity: an approach to plasticity theory based on thermodynamic principles*. Springer Science & Business Media, 2007.

[28] B. Nedjar, 'Formulation of a nonlinear porosity law for fully saturated porous media at finite strains,' *Journal of the Mechanics and Physics of Solids*, vol. 61, no. 2, pp. 537–556, 2013.

[29] B. Nedjar, 'On finite strain poroplasticity with reversible and irreversible porosity laws. formulation and computational aspects,' *Mechanics of Materials*, vol. 68, pp. 237–252, 2014.

[30] G. Pretti, W. M. Coombs, C. E. Augarde, M. M. Puigvert and J. A. R. Gutiérrez, 'Imposing constrained porosity values in a bi-phase elasto-plastic material under Terzaghi's effective stress principle,' *Submitted to Mechanics of Materials*, 2023.

[31] E. Kröner, 'Allgemeine kontinuumstheorie der versetzungen und eigenspannungen,' *Archive for Rational Mechanics and Analysis*, vol. 4, no. 1, pp. 273–334, 1959.

[32] E. H. Lee, 'Elastic-plastic deformation at finite strains,' *Journal of Applied Mechanics*, vol. 36, no. 1, pp. 1–6, 1969.

[33] J. Mandel, *Plasticité Classique et Viscoplasticité* (CISM Courses and Lectures). Springer, 1971, vol. 97.

[34] B. Coleman and W. Noll, 'The thermodynamics of elastic materials with heat conduction,' *Archive for Rational Mechanics and Analysis*, vol. 13, p. 167, 1963.

[35] B. D. Coleman and M. E. Gurtin, 'Thermodynamics with internal state variables,' *The Journal of Chemical Physics*, vol. 47, no. 2, pp. 597–613, 1967.

[36] R. I. Borja and E. Alarcón, 'A mathematical framework for finite strain elastoplastic consolidation part 1: Balance laws, variational formulation, and linearization,' *Computer Methods in Applied Mechanics and Engineering*, vol. 122, no. 1-2, pp. 145–171, 1995.

[37] I. Collins and G. Houlsby, 'Application of thermomechanical principles to the modelling of geotechnical materials,' *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 453, no. 1964, pp. 1975–2001, 1997.

[38] P. Flory, 'Thermodynamic relations for high elastic materials,' *Transactions of the Faraday Society*, vol. 57, pp. 829–838, 1961.

[39] H. Hencky, 'The elastic behavior of vulcanized rubber,' *Journal of Applied Mechanics*, no. 1, pp. 45–48, 1933.

[40] M. Zytynski, M. Randolph, R. Nova and C. Wroth, 'On modelling the unloading-reloading behaviour of soils,' *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 2, no. 1, pp. 87–93, 1978.

[41] J. C. Simo, 'Algorithms for static and dynamic multiplicative plasticity that preserve the classical return mapping schemes of the infinitesimal theory,' *Computer Methods in Applied Mechanics and Engineering*, vol. 99, no. 1, pp. 61–112, 1992.

[42] J. Lubliner, 'On the thermodynamic foundations of non-linear solid mechanics,' *International Journal of Non-Linear Mechanics*, vol. 7, no. 3, pp. 237–254, 1972.

[43] I. Collins, 'Associated and non-associated aspects of the constitutive laws for coupled elastic/plastic materials,' *The International Journal Geomechanics*, vol. 2, no. 2, pp. 259–267, 2002.

[44] S. Murakami, *Continuum damage mechanics: a continuum mechanics approach to the analysis of damage and fracture*. Springer Science & Business Media, 2012, vol. 185.

[45] Y. Yamakawa, 'Hyperelastic constitutive models for geomaterials: Extension of existing models to include finite strains and their comparison,' *Computers and Geotechnics*, vol. 143, p. 104 600, 2022.

[46] K. Oliynyk and C. Tamagnini, 'Finite deformation hyperplasticity theory for crushable, cemented granular materials,' *Open Geomechanics*, vol. 2, pp. 1–33, 2020.

[47] F. Armero, 'Formulation and finite element implementation of a multiplicative model of coupled poro-plasticity at finite strains under fully saturated conditions,' *Computer Methods in Applied Mechanics and Engineering*, vol. 171, no. 3-4, pp. 205–241, 1999.

[48] A. Gajo, 'A general approach to isothermal hyperelastic modelling of saturated porous media at finite strains with compressible solid constituents,' *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 466, no. 2122, pp. 3061–3087, 2010.

[49]   S. Bandara and K. Soga, 'Coupling of soil deformation and pore fluid flow using material point method,' *Computers and Geotechnics*, vol. 63, pp. 199–214, 2015.

[50]   R. F. Craig, *Craig's soil mechanics*. CRC press, 2004.

[51]   Y. Zhao and J. Choo, 'Stabilized material point methods for coupled large deformation and fluid flow in porous materials,' *Computer Methods in Applied Mechanics and Engineering*, vol. 362, p. 112 742, 2020.

[52]   Y. Gan, Z. Sun, Z. Chen, X. Zhang and Y. Liu, 'Enhancement of the material point method using B-spline basis functions,' *International Journal for Numerical Methods in Engineering*, vol. 113, no. 3, pp. 411–431, 2018.

[53]   K. Soga, E. Alonso, A. Yerro, K. Kumar and S. Bandara, 'Trends in large-deformation analysis of landslide mass movements with particular emphasis on the material point method,' *Géotechnique*, vol. 66, no. 3, pp. 248–273, 2016.

[54]   T. Noda and T. Toyoda, 'Development and verification of a soil–water coupled finite deformation analysis based on u–w–p formulation with fluid convective nonlinearity,' *Soils and Foundations*, vol. 59, no. 4, pp. 888–904, 2019.

[55]   X. Zheng, F. Pisanò, P. J. Vardon and M. A. Hicks, 'Fully implicit, stabilised, three-field material point method for dynamic coupled problems,' *Engineering with Computers*, pp. 1–20, 2022.

[56]   W. M. Coombs, T. J. Charlton, M. Cortis and C. E. Augarde, 'Overcoming volumetric locking in material point methods,' *Computer Methods in Applied Mechanics and Engineering*, vol. 333, pp. 1–21, 2018.

[57]   S. G. Bardenhagen and E. M. Kober, 'The generalized interpolation material point method,' *Computer Modeling in Engineering and Sciences*, vol. 5, no. 6, pp. 477–496, 2004.

[58]   J. A. White and R. I. Borja, 'Stabilized low-order finite elements for coupled solid-deformation/fluid-diffusion and their application to fault zone transients,' *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 49-50, pp. 4353–4366, 2008.

[59]   F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*. Springer Science & Business Media, 2012, vol. 15.

[60]   J. Simo and N. Tarnow, 'The discrete energy-momentum method. Conserving algorithms for nonlinear elastodynamics,' *Zeitschrift für angewandte Mathematik und Physik ZAMP*, vol. 43, no. 5, pp. 757–792, 1992.

[61]   Z. Liu, Y. Zhang, Y. Jiang, H. Yang and Y. Yang, 'Unfitted finite element method for fully coupled poroelasticity with stabilization,' *Computer Methods in Applied Mechanics and Engineering*, vol. 397, p. 115 132, 2022.

[62]   M. Preisig and J. H. Prévost, 'Stabilization procedures in coupled poromechanics problems: A critical assessment,' *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 35, no. 11, pp. 1207–1225, 2011.

[63] C. R. Dohrmann and P. B. Bochev, 'A stabilized finite element method for the Stokes problem based on polynomial pressure projections,' *International Journal for Numerical Methods in Fluids*, vol. 46, no. 2, pp. 183–201, 2004.

[64] X. Zheng, F. Pisano, P. J. Vardon and M. A. Hicks, 'An explicit stabilised material point method for coupled hydromechanical problems in two-phase porous media,' *Computers and Geotechnics*, vol. 135, p. 104 112, 2021.

[65] H. Zhang, K. Wang and Z. Chen, 'Material point method for dynamic analysis of saturated porous media under external contact/impact of solid bodies,' *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 17-20, pp. 1456–1472, 2009.

[66] Y. Zheng, F. Gao, H. Zhang and M. Lu, 'Improved convected particle domain interpolation method for coupled dynamic analysis of fully saturated porous media involving large deformation,' *Computer Methods in Applied Mechanics and Engineering*, vol. 257, pp. 150–163, 2013.

[67] F. Zabala and E. Alonso, 'Progressive failure of aznalcóllar dam using the material point method,' *Géotechnique*, vol. 61, no. 9, pp. 795–808, 2011.

[68] W. M. Coombs and C. E. Augarde, 'AMPLE: A material point learning environment,' *Advances in Engineering Software*, vol. 139, p. 102 748, 2020.

[69] J. C. Simo, 'A framework for finite strain elastoplasticity based on maximum plastic dissipation and the multiplicative decomposition: Part I. continuum formulation,' *Computer Methods in Applied Mechanics and Engineering*, vol. 66, no. 2, pp. 199–219, 1988.

[70] I. F. Collins and T. Hilder, 'A theoretical framework for constructing elastic/plastic constitutive models of triaxial tests,' *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 26, no. 13, pp. 1313–1347, 2002.

[71] W. M. Coombs and R. S. Crouch, 'Algorithmic issues for three-invariant hyperplastic critical state models,' *Computer methods in applied mechanics and engineering*, vol. 200, no. 25-28, pp. 2297–2318, 2011.

[72] A. N. Schofield and P. Wroth, *Critical state soil mechanics*. McGraw-hill London, 1968, vol. 310.

[73] Y. Yamaguchi, S. Moriguchi and K. Terada, 'Extended B-spline-based implicit material point method,' *International Journal for Numerical Methods in Engineering*, vol. 122, no. 7, pp. 1746–1769, 2021.

[74] F. Brezzi and J. Pitkäranta, 'On the stabilization of finite element approximations of the Stokes equations,' in *Efficient Solutions of Elliptic Systems: Proceedings of a GAMM-Seminar Kiel, January 27 to 29, 1984*, Springer, 1984, pp. 11–19.

[75] T. J. Hughes, 'Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods,' *Computer methods in applied mechanics and engineering*, vol. 127, no. 1-4, pp. 387–401, 1995.

Chapter 5

Frictional contact

In the problem of focus of this thesis, i.e., anchor embedment, the role of the chain can be regarded as secondary in the inertial dissipation taking place during the embedment process. If this is the case, the anchor is expected to be slowed down principally by two mechanisms: the dissipation due to the plastic strains accumulating in the soil surrounding the anchor[†], and the frictional dissipation between the anchor surface and the soil. While the former mechanism can be represented via the usual plasticity theory (see Chapter 3), the latter needs its own treatment, which is what the current chapter deals with.

In particular, Section 5.1 considers the treatment of contact kinematics, statics, power-conjugation between them and necessary constitutive relationships of interest. These equations are primarily stated in a strong (point-wise) form, while Section 5.2 state them in an integral (weak) form and incorporate the contact power in a more general functional. Space and time discretisations follow in Section 5.3. Different ways of computing the contact forces are included in Section 5.4, while Section 5.5 is about numerical applications and relative discussions of the techniques introduced throughout Chapter 5.

Frictional contact is perhaps one of the most complicated challenges in computational mechanics, as demonstrated by the plethora of books and articles about it. Because of this extent, the current work does not attempt to present a complete account of all the previous results in this area[‡], but it provides the references that most significantly influenced this work. As further proof of this complexity, the following quote from a very recent publication by Andrews *et al.* [4] in the computer graphics community can be exemplifying

> ``... despite its extensive use in modern computer graphics, contact
> simulation remains of the most challenging problems in physics-based
> animation.''

[†]Since the soil is treated as elasto-plastic, a small part of the kinetic energy of the anchor is converted into elastic strain energy. However, the above-mentioned plastic dissipation is expected to play a more dominant role.

[‡]The reader interested in an overview and a reasonably recent state-of-the-art presentation of contact mechanics applied to the Finite Element Method (FEM) can refer to De Lorenzis *et al.* [1]. For a rigorous treatment of the same subject and its implementations within the FEM, Laursen [2] and Wriggers and Laursen [3] are considered to be excellent references.

The intricacies of treating frictional contact stem from multiple reasons, which can be grouped into three main categories:

1. the treatment of the finite contact kinematics;

2. the position where the contact conditions are imposed; and

3. the way contact forces are computed.

Regarding the kinematics (to which Sections 5.1.1 and 5.1.2 are dedicated), non-linearity of the finite strain theory must be added to the treatment of (usually curvilinear) surfaces, which, in turn, interact with each other. The description of the surfaces and their interaction usually rely on surface coordinates (see, for instance, Laursen [2]), whose mathematical treatment requires the use of curvilinear (non-orthonormal) local coordinates. In contrast with this approach, this work follows that proposed by Curnier *et al.* [5], Pietrzak and Curnier [6] and Poulios and Renard [7], where Cartesian coordinates are employed. In this fashion, frictional contact can be included in the finite strain mechanics of solid bodies previously described in Chapter 3 without further coordinate systems.

Nonetheless, the interaction between the surfaces cannot be entirely separated from the second challenge, i.e. where the contact conditions are applied, which is, in turn, a function of the surface discretisation process. Within the Material Point Method (MPM) community, the vast majority of publications (see [8–25]) apply contact conditions on the nodes of the background mesh. In this work, this approach is referred to as the *node-to-node* formulation, which is detailed in Section 5.3.1.1. Most publications adopt this simplified approach because of the lack of an explicit built-in representation of the body boundaries in the MPM (see, in this regard, the discussion developed in Section 3.6.2.1), in contrast to, for instance, the FEM. In the literature, techniques exist to reconstruct these boundaries (see Bing *et al.* [26], or Liu and Sun [27], or Chandra *et al.* [28], to name a few), but their general implementation is not straightforward. This work introduces another technique used for the FEM (see Poulios and Renard [7]) and adapted to the MPM with a compromise. This method, named *point-to-surface*, is outlined in Section 5.3.1.2.

The computation of the (frictional) contact forces, i.e. the forces taking place between the bodies which come into contact, represents the last (third) challenge. In this work, two main groups of methods were considered. In the first family of techniques, the interaction between bodies in contact is taken into account by adding the (real) power generated by contact forces (see Section 5.1.3) to the functionals describing the individual behaviour of each distinct body. As mentioned, it is clear how a treatment of the virtual power (see Section 5.1.4) generated by contact forces becomes necessary to develop this kind of methods. The same applies to constitutive relationships (see Section 5.1.5). Since any method belonging to this class of techniques can be recast in an Augmented Lagrangian framework, they are named the *Augmented Lagrangian method(s)*. The *penalty* (Penalty Method (PM)), the *Lagrange multiplier* (Lagrange Multiplier Method (LMM)), the (proper) *Augmented Lagrangian* (Augmented Lagrangian method (ALM)), and *Nitsche*'s method are all part of this group. This chapter considers all except Nitsche (see Sections from 5.4.1 to 5.4.3). To
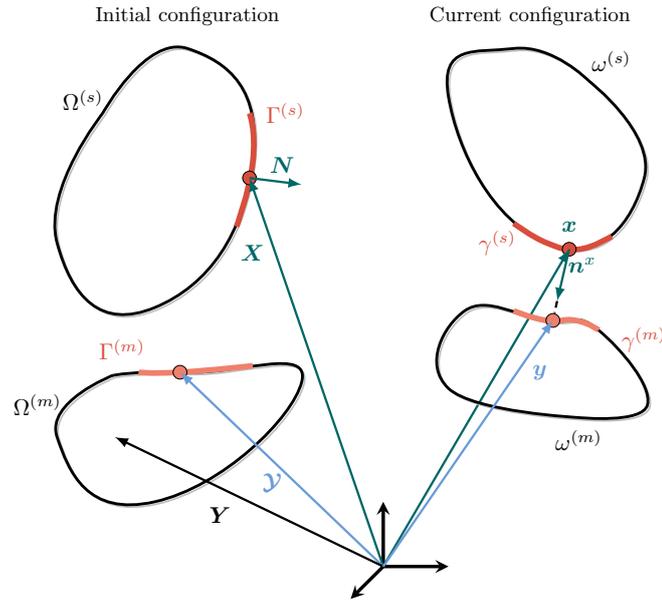
Figure 5.1: Main and secondary deformable bodies, represented in the initial and current configuration.

the author's knowledge, in the literature, the second class of techniques is restricted to MPM only (see [8–17, 19–24, 29]). The contact forces are calculated for these methods based on the (nodal) impulses exchanged between the contacting bodies. Owing to this behaviour, this group was named the *contact velocity-based method* (Contact Velocity-based Method (CVM)). In particular, this work considers the original version of this technique from Bardenhagen *et al.* [8], with some measures borrowed from González-Acosta *et al.* [23] to adapt this technique to the implicit MPM (see Section 5.4.4). As this kind of method does not rely on a weak formulation (contact forces are added to an already discretised version of the equilibrium equations), Sections 5.1-5.3 appear to be unnecessary to understand this method. From this perspective, Section 5.4.4 can be read almost entirely separately from the rest of the chapter.

The development of a single framework to describe the above-mentioned different methods for computing contact conditions (node-to-node and point-to-segment) and contact forces (ALM, LMM, PM and CVM) constitutes a novel contribution. Owing to this unique structure, the treatment and the consequential implementation of the different techniques are less onerous and error-prone. The single framework also allows the thorough assessment of these techniques regarding numerical dissipation. This examination complies with the bigger purpose designing MPM analyses more in compliance with their underlying continuum formulations.

5.1 Large deformation contact

Let first consider two deformable solids $\mathcal{B}^{(s)}$ and $\mathcal{B}^{(m)}$ (see Figure 5.1) occupying an initial volume $\Omega^{(s)}$ and $\Omega^{(m)}$ of the Euclidean space $\mathcal{E}$, where $(\bullet)^{(s)}$ stands for *secondary* and $(\bullet)^{(m)}$ for *main* body*. In order to keep the notation as simple as possible and without any loss of generality, the boundaries of main and secondary bodies are not partitioned into further parts to include Dirichlet or Neumann boundary conditions.

Let the *motion* mappings for each particle in the initial configuration be defined by the (sufficiently regular) bijection such that $\boldsymbol{\varphi}^{(s)} : \Omega^{(s)} \times [0, T] \to \mathcal{E}$ and $\boldsymbol{x} = \boldsymbol{\varphi}^{(s)}(\boldsymbol{X}, t) = \boldsymbol{x}(\boldsymbol{X}, t)$ for the secondary body, and the (sufficiently regular) bijection such that $\boldsymbol{\varphi}^{(m)} : \Omega^{(m)} \times [0, T] \to \mathcal{E}$ and $\boldsymbol{y} = \boldsymbol{\varphi}^{(m)}(\boldsymbol{Y}, t) = \boldsymbol{y}(\boldsymbol{Y}, t)$ for the main body, with $t \in [0, T]$ being the time. The above mappings allow describing the *displacement* of each particle, i.e., $\boldsymbol{u}(\boldsymbol{X}, t) = \boldsymbol{x}(\boldsymbol{X}, t) - \boldsymbol{X}$ and $\boldsymbol{u}(\boldsymbol{Y}, t) = \boldsymbol{y}(\boldsymbol{Y}, t) - \boldsymbol{Y}$. The deformations of the particles are expressed via the *deformation gradients* (similarly to Eq. (3.7)), i.e.,

$$\boldsymbol{F}(\boldsymbol{X}, t) := \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{X}}, \quad \text{with } \det(\boldsymbol{F}(\boldsymbol{X}, t)) > 0; \tag{5.1}$$

$$\boldsymbol{F}(\boldsymbol{Y}, t) := \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{Y}}, \quad \text{with } \det(\boldsymbol{F}(\boldsymbol{Y}, t)) > 0. \tag{5.2}$$

Let the current regions occupied by the bodies be given by $\omega^{(s)} = \boldsymbol{\varphi}^{(s)}(\Omega^{(s)}, t)$ and $\omega^{(m)} = \boldsymbol{\varphi}^{(m)}(\Omega^{(m)}, t)$. While the current surfaces not in contact follow trivially the logic dictated by the diverse motion mappings, the current surfaces in contact, which are denoted by $\gamma^{(m)}$ (originally $\Gamma^{(m)}$) and $\gamma^{(s)}$ (originally $\Gamma^{(s)}$), are identified via the selected projection procedure. The next subsection is devoted to the description of one of these projection techniques.

5.1.1 Solid surfaces interaction

To fully address the contact problem, some other quantities must be introduced. To this end, let the current normal direction to the secondary body be denoted by $\boldsymbol{n}(\boldsymbol{X}, t)$ (see Figure 5.1), which is related to its initial counterpart $\boldsymbol{N}(\boldsymbol{X})$ via Nanson's formula (compare with Eq. (3.11)), i.e.,

$$\tilde{\boldsymbol{n}}(\boldsymbol{X}, t)\, da = J\boldsymbol{F}^{-T}(\boldsymbol{X}, t)\, \boldsymbol{N}(\boldsymbol{X})\, dA, \quad \text{with } \|\boldsymbol{N}(\boldsymbol{X})\| = 1, \tag{5.3}$$

with da and dA indicating the current and reference infinitesimal surfaces. Since there is no guarantee that $\|\tilde{\boldsymbol{n}}\| = 1$, this vector can be normalised as follows

$$\boldsymbol{n}(\boldsymbol{X}, t) = \frac{\boldsymbol{F}^{-T}(\boldsymbol{X}, t)\, \boldsymbol{N}}{\left\|\boldsymbol{F}^{-T}(\boldsymbol{X}, t)\, \boldsymbol{N}\right\|} = \frac{\tilde{\boldsymbol{n}}}{\|\tilde{\boldsymbol{n}}\|} = \boldsymbol{n}^x(\boldsymbol{X}, t), \tag{5.4}$$

---

*Classically, the standard labels in the contact mechanics literature were designated by the words *slave* and *master*. However, this work adopts less controversial definitions, without changing the initial letters, so that the mathematical notation undergoes no changes.

where the superscript $(\bullet)^x$ is specified to highlight that this normal is computed on the current secondary surface. Having introduced the normal direction at the current secondary surface, the plane tangential to the same surface at the same point can be identified by the second-order tensor

$$\boldsymbol{T}^{n^x} := \boldsymbol{I}^{(2)} - \boldsymbol{n}^x\left(\boldsymbol{X},t\right) \otimes \boldsymbol{n}^x\left(\boldsymbol{X},t\right). \tag{5.5}$$

A similar procedure can be performed to compute the normal direction to the main body, which is indicated by $\boldsymbol{n}^y\left(\boldsymbol{Y},t\right)$.

As previously mentioned, it is necessary to establish a projection strategy, i.e., a bijection[*] that unambiguously defines the contact points on both current surfaces. In particular, in this work, the so-called *ray-tracing projection* is introduced[†], where the points $\boldsymbol{y}$ on the current main surface $\gamma^{(m)}$ are identified by the projection of points $\boldsymbol{x}$ along the normal to the current secondary surface $\boldsymbol{n}^x$. These points on the current main surface $\gamma^{(m)}$ are denoted by $\boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)$ (with a slight further abuse of notation). However, it can be appreciated how the difference between points obtained via this mapping and points that happen to belong to the current main body lies in the dependencies of the functions, the former being represented by $\boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)$ and the latter by $\boldsymbol{y}\left(\boldsymbol{Y},t\right)$.

Having defined the kind of projection between any contact point on the current secondary surface to its main counterpart, the *gap vector function*, which defines the vector distance (along $\boldsymbol{n}^x$) between the related points on the different surfaces, can be introduced as

$$\boldsymbol{g} := \boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right) - \boldsymbol{x}\left(\boldsymbol{X},t\right) = g\,\boldsymbol{n}^x\left(\boldsymbol{X},t\right), \tag{5.6}$$

where it can be seen that the magnitude g describes three different scenarios, that is

$$\begin{cases} g < 0, & \text{overlap;} \\ g = 0, & \text{contact;} \\ g > 0, & \text{gap.} \end{cases} \tag{5.7}$$

Since the above definition of the gap function requires that the distance between the two bodies has non-zero component only along the normal direction $\boldsymbol{n}^x$, it can be noticed that the body are in contact if

$$\boldsymbol{T}^{n^x}\boldsymbol{g} = \boldsymbol{0}. \tag{5.8}$$

If the above equation has a solution (see Curnier *et al.* [5] for existence of uniquess of this solution), this can also be found by the stationarity of the functional[‡]

$$\boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right) = \text{argmin}_{\boldsymbol{\mathcal{Y}} \in \Gamma^{(m)}}\left\{\frac{1}{2}\boldsymbol{y}\cdot\left(\boldsymbol{T}^{n^x}\boldsymbol{y}\right) - \boldsymbol{y}\cdot\left(\boldsymbol{T}^{n^x}\boldsymbol{x}\right)\right\}, \tag{5.9}$$

---

[*]Even if main and secondary surfaces are regular, a one-to-one correspondence between these surfaces is not always guaranteed (see, in this regard, Curnier *et al.* [5]). However, this issue is not further discussed in this work.

[†]In the literature, there exist other strategies to relate points on the secondary surface with points on the main one. Among all, it is worth mentioning the *closest-point projection* (see, for instance, Curnier *et al.* [5]). Differences between the ray-tracing and the closest-point projection are assessed by Poulios and Renard [7]. As pointed out by the same authors, the main advantage of the ray-tracing strategy over the closest-point projection lies in a simpler dependency of the normal $\boldsymbol{n}^x$ (whose position is known on the secondary surface) compared to $\boldsymbol{n}^y$ (whose position, being a function of the projection strategy, is unknown). The reader is referred to Poulios and Renard [7] for further details on the projection strategies.

[‡]If the solution exists and is unique, the solution of Eq. (5.8) and stationarity of the functional introduced in Eq. (5.9) are guaranteed to be the same if $\boldsymbol{T}^{n^x}$ is positive semi-definite.

which is given via the variational derivative of the above functional, i.e.,

$$\boldsymbol{T}^{n^x}\left(\boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)-\boldsymbol{x}\left(\boldsymbol{X},t\right)\right)\cdot\left(\underbrace{\frac{\partial\boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)}{\partial\boldsymbol{\mathcal{Y}}}}_{:=\boldsymbol{F}^{\mathcal{Y}}}\delta\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right)\right)=0,\qquad\forall\delta\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right).$$
(5.10)

As the variation $\delta\boldsymbol{\mathcal{Y}}$ belongs by definition to the plane tangent to the initial main surface $\Gamma^{(m)}$, it follows that the vector $\boldsymbol{F}^{\mathcal{Y}}\delta\boldsymbol{\mathcal{Y}}$ belongs to the plane tangent to the current main surface $\gamma^{(m)}$, i.e.,

$$\left(\boldsymbol{F}^{\mathcal{Y}}\delta\boldsymbol{\mathcal{Y}}\right)\cdot\boldsymbol{n}^{y}\left(\boldsymbol{y}\left(\boldsymbol{\mathcal{Y}},t\right),t\right)=0.$$
(5.11)

Even though at this stage the above orthogonality appears to have little practical applications, it will be useful in the development of the kinematics in the following.

5.1.2 Solid surfaces motions

To develop the complete kinematics defining how any point $\boldsymbol{x}\left(\boldsymbol{X},t\right)$ on the current secondary surface moves with respect to its projection $\boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)$ on the current main surface, the following notational simplifications for the rates are introduced

$$(\dot{\bullet})=\left.\frac{d\left(\bullet\right)}{dt}\right|_{X};$$
(5.12)

$$(\bullet)'=\left.\frac{d\left(\bullet\right)}{dt}\right|_{\mathcal{Y}}.$$
(5.13)

A first attempt to define the relative velocity between the related points on different surfaces can be the rate of the gap vector in the secondary reference frame, i.e.,

$$\dot{\boldsymbol{g}}=\dot{\boldsymbol{y}}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)-\dot{\boldsymbol{x}}\left(\boldsymbol{X},t\right)=\dot{g}\,\boldsymbol{n}^{x}+g\,\dot{\boldsymbol{n}}^{x},$$
(5.14)

where it can be seen that the above equation can be decomposed into a normal and a tangential part according to the *Frenet–Serret formulae* [30]

$$\dot{\boldsymbol{g}}_{n}:=\boldsymbol{n}^{x}\left(\dot{\boldsymbol{g}}\cdot\boldsymbol{n}^{x}\right)=\dot{g}\,\boldsymbol{n}^{x};$$
(5.15)

$$\dot{\boldsymbol{g}}_{t}:=\dot{\boldsymbol{g}}-\dot{\boldsymbol{g}}_{n}=g\,\dot{\boldsymbol{n}}^{x}.$$
(5.16)

Moreover, it is convenient to express the rate of the projection $\boldsymbol{y}$ also in the secondary relative frame, this being

$$\dot{\boldsymbol{y}}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)=\boldsymbol{y}'\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)+\boldsymbol{F}^{\mathcal{Y}}\dot{\boldsymbol{\mathcal{Y}}}\left(\boldsymbol{X},t\right).$$
(5.17)

From Eqs. (5.14) and (5.17), the rate of the projected position on the main surface can be isolated

$$\dot{\boldsymbol{y}}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)=\dot{\boldsymbol{x}}\left(\boldsymbol{X},t\right)+\dot{g}\,\boldsymbol{n}^{x}+g\,\dot{\boldsymbol{n}}^{x}=\boldsymbol{y}'\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right)+\boldsymbol{F}^{\mathcal{Y}}\dot{\boldsymbol{\mathcal{Y}}}\left(\boldsymbol{X},t\right).$$
(5.18)

Taking advantage of the orthogonality between $\boldsymbol{F}^{\mathcal{Y}}$ and $\boldsymbol{n}^{y}$ (see Eq. (5.11)), the projection of the above equation along $\boldsymbol{n}^{y}$ gives

$$\boldsymbol{n}^{y}\cdot\left(\dot{\boldsymbol{x}}\left(\boldsymbol{X},t\right)+\dot{g}\,\boldsymbol{n}^{x}+g\,\dot{\boldsymbol{n}}^{x}\right)=\boldsymbol{n}^{y}\cdot\boldsymbol{y}'\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X},t\right),t\right).$$
(5.19)
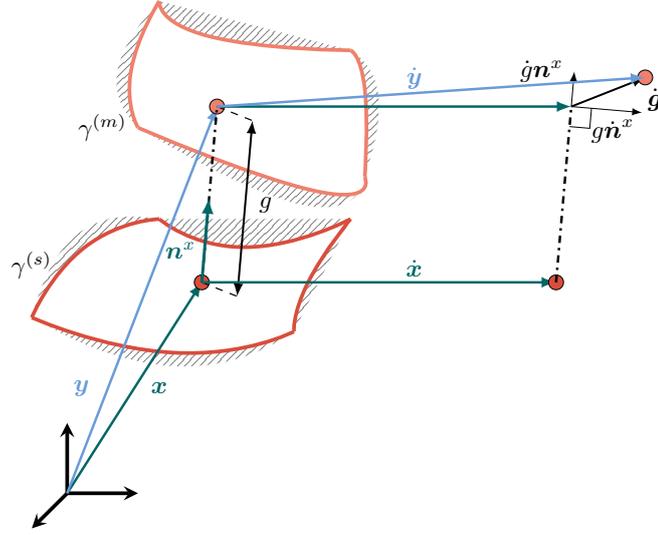
Figure 5.2: Illustration of current main and secondary surfaces (lying on their respective grey-shaded bodies). The considered material derivatives are computed with respect to the secondary body.

It hence follows that $\dot{g}$ can be computed as

$$\dot{g} = -\frac{1}{\boldsymbol{n}^y \cdot \boldsymbol{n}^x} \boldsymbol{n}^y \cdot \left(\dot{\boldsymbol{x}} \left(\boldsymbol{X}, t \right) - \boldsymbol{y}' \left(\boldsymbol{\mathcal{Y}} \left(\boldsymbol{X}, t \right), t \right) + g \, \dot{\boldsymbol{n}}^x \right). \tag{5.20}$$

Given the above result, Eq. (5.18) becomes

$$\dot{\boldsymbol{y}} \left(\boldsymbol{\mathcal{Y}} \left(\boldsymbol{X}, t \right), t \right) = \boldsymbol{y}' + \left(\boldsymbol{I}^{(2)} - \frac{\boldsymbol{n}^x \otimes \boldsymbol{n}^y}{\boldsymbol{n}^x \cdot \boldsymbol{n}^y} \right) \left(\dot{\boldsymbol{x}} \left(\boldsymbol{X}, t \right) - \boldsymbol{y}' \left(\boldsymbol{\mathcal{Y}} \left(\boldsymbol{X}, t \right), t \right) + g \, \dot{\boldsymbol{n}}^x \right), \tag{5.21}$$

and, from the RHS of Eq. (5.17), the rate of the reference position can be computed as

$$\dot{\boldsymbol{\mathcal{Y}}} \left(\boldsymbol{X}, t \right) = \left(\boldsymbol{F}^{\mathcal{Y}} \right)^{-1} \left(\boldsymbol{I}^{(2)} - \frac{\boldsymbol{n}^x \otimes \boldsymbol{n}^y}{\boldsymbol{n}^x \cdot \boldsymbol{n}^y} \right) \left(\dot{\boldsymbol{x}} \left(\boldsymbol{X}, t \right) - \boldsymbol{y}' \left(\boldsymbol{\mathcal{Y}} \left(\boldsymbol{X}, t \right), t \right) + g \, \dot{\boldsymbol{n}}^x \right). \tag{5.22}$$

Even though Eqs. (5.21) and (5.22) are not essential to develop the kinematics between the bodies, they will constitute the starting point in the linearisation (necessary for the Newton-Raphson (NR) solution scheme) with respect to the displacement field (see Appendix B.3).

If Eq. (5.16) is considered, it can be seen how the tangential component of the rate of the gap function cannot be a measure of slip rate, as this quantity entirely vanishes when the gap function is zero (i.e., when contact occurs).

Another quantity which could be seen as an adequate candidate to describe the slip rate is the *relative velocity*, i.e., the difference in velocity of the points $\boldsymbol{x}$ and $\boldsymbol{y}$ in their reference frame. This is given by the formula

$$\boldsymbol{v}^r := \boldsymbol{y}' \left(\boldsymbol{\mathcal{Y}} \left(\boldsymbol{X}, t \right), t \right) - \dot{\boldsymbol{x}} \left(\boldsymbol{X}, t \right). \tag{5.23}$$

Nonetheless, even this candidate measure of the slip rate must be ruled out. As pointed out by Curnier *et al.* [5], this quantity happens to be non-objective (see, for a discussion about objectivity, page 46). This is demonstrated by considering a rotation $\boldsymbol{Q}$ applied to the

quantities appearing in the above equations, where it can be observed that

$$\tilde{\boldsymbol{v}}^r = \boldsymbol{Q}\boldsymbol{v}^r \neq \underbrace{\boldsymbol{Q}\boldsymbol{y}' + \boldsymbol{Q}'\boldsymbol{y}}_{=\tilde{\boldsymbol{y}}'} - \underbrace{\boldsymbol{Q}\dot{\boldsymbol{x}} - \dot{\boldsymbol{Q}}\boldsymbol{x}}_{=\tilde{\dot{\boldsymbol{x}}}} = \dot{\boldsymbol{Q}}\boldsymbol{g} + \boldsymbol{Q}\boldsymbol{v}^r. \tag{5.24}$$

Thus, a modification of the relative velocity, named the *contact velocity*, is introduced, this being

$$\boldsymbol{v}^c := \dot{\boldsymbol{x}}\left(\boldsymbol{X}, t\right) - \boldsymbol{y}'\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X}, t\right), t\right) + g\,\dot{\boldsymbol{n}}^x = \boldsymbol{F}^{\mathcal{Y}}\dot{\boldsymbol{\mathcal{Y}}}\left(\boldsymbol{X}, t\right) - \dot{g}\,\boldsymbol{n}^x, \tag{5.25}$$

where Eqs. (5.14) and (5.17) are used to compute the RHS of the above equation. It can also be easily demonstrated that this quantity is objective, since

$$\tilde{\boldsymbol{v}}^c = \boldsymbol{Q}\boldsymbol{v}^c = -\dot{\boldsymbol{Q}}\boldsymbol{g} - \boldsymbol{Q}\boldsymbol{v}^r + g\left(\underbrace{\boldsymbol{Q}\,\dot{\boldsymbol{n}}^x + \dot{\boldsymbol{Q}}\boldsymbol{n}^x}_{=\tilde{\dot{\boldsymbol{n}}}^x}\right) =$$

$$= \dot{\boldsymbol{Q}}\left(g\,\boldsymbol{n}^x - \boldsymbol{g}\right) + \boldsymbol{Q}\left(g\,\dot{\boldsymbol{n}}^x - \boldsymbol{v}^r\right) = \boldsymbol{Q}\boldsymbol{v}^c. \tag{5.26}$$

From an objectivity perspective, the contact velocity can be seen as a possible measure of slip. However, some consideration on the physical meaning of this quantity can be made by projecting it along the normal and the tangential directions as follows

$$\boldsymbol{v}_n^c := \boldsymbol{n}^x\left(\boldsymbol{v}^c \cdot \boldsymbol{n}^x\right) = -\dot{g}\,\boldsymbol{n}^x = -\dot{\boldsymbol{g}}_n; \tag{5.27}$$

$$\boldsymbol{v}_t^c := \boldsymbol{v}^c - \boldsymbol{v}_n^c = \boldsymbol{F}^{\mathcal{Y}}\dot{\boldsymbol{\mathcal{Y}}} = v_t^c\,\boldsymbol{t}^y, \tag{5.28}$$

with $\boldsymbol{t}^y$ being a vector belonging to the plane orthogonal to $\boldsymbol{n}^y$. It can be observed that the contact velocity has a normal component along $-\boldsymbol{n}^x$, while its tangential part is along the plane tangent to the current main surface defined by the normal vector $\boldsymbol{n}^y$. This discrepancy is due to the chosen projection strategy (i.e., ray-tracing), which considers the gap vector function along the current secondary surface $\boldsymbol{n}^x$. Hence, in the case of ray-tracing projection, the contact velocity is quite an artificial measure of slip rate, which loses its physical meaning when contact does not occur, i.e., $g \neq 0$. Nonetheless, as will be shown in Section 5.1.5, since the contact velocity plays a role only in case of contact, i.e., $g = 0$, it can be overall seen as a physical meaningful measure of the slip rate. Indeed, in the case of contact, the two related points belonging to different body must have the same and opposite normal vector, i.e., if $g = 0$ then $\boldsymbol{n}^x = -\boldsymbol{n}^y$, so that the projection along uniquely-defined normal and tangential directions of the contact velocity is restored.

5.1.3 Contact statics

Since in the most general scenario the main and secondary contact surfaces can be in non-contact conditions ($g \neq 0$), it is not possible to use the action-reaction principle of stresses acting on the same areas (with opposite outward normal direction) as is conventionally done when cutting a continuous solid along an interface. Owing to this discontinuity in the involved areas, the action-reaction principle can nevertheless be formulated for distant particles in terms of both forces and angular momentum. If the first Piola-Kirchhoff stress

vector $\boldsymbol{p} = \boldsymbol{P}\boldsymbol{N}$ is introduced, this principle can be stated as balance of linear and angular forces in the following way

$$\int_{\Gamma^{(s)}} \boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) dA^{(s)} + \int_{\Gamma^{(m)}} \boldsymbol{p}\left(\boldsymbol{\mathcal{Y}}, t, \boldsymbol{N}\left(\boldsymbol{\mathcal{Y}}\right)\right) dA^{(m)} = \boldsymbol{0}; \tag{5.29}$$

$$\int_{\Gamma^{(s)}} \boldsymbol{x}\left(\boldsymbol{X}, t\right) \times \boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) dA^{(s)} + \int_{\Gamma^{(m)}} \boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X}, t\right), t\right) \times \boldsymbol{p}\left(\boldsymbol{\mathcal{Y}}, t, \boldsymbol{N}\left(\boldsymbol{\mathcal{Y}}\right)\right) dA^{(m)} = \boldsymbol{0}. \tag{5.30}$$

Because of the mapping between the points on the secondary surfaces $\boldsymbol{x}$ and its main counterparts $\boldsymbol{y}$, it is possible to express the infinitesimal area of the main surface as a function of the (infinitesimal) secondary one, i.e., $dA^{(m)} = J^{(m)} dA^{(s)}$. Owing to this relationship and assuming sufficient regularity (so that the above integral equations can be considered to be true point-wise), it can be found from Eq. (5.29) that

$$\boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) = -J^{(m)} \boldsymbol{p}\left(\boldsymbol{\mathcal{Y}}, t, \boldsymbol{N}\left(\boldsymbol{\mathcal{Y}}\right)\right). \tag{5.31}$$

If this equation is plugged into Eq. (5.30), it follows that

$$\left(\underbrace{\boldsymbol{x}\left(\boldsymbol{X}, t\right) - \boldsymbol{y}\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X}, t\right), t\right)}_{=-\boldsymbol{g}}\right) \times \boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) = \boldsymbol{0}. \tag{5.32}$$

It can be observed how the above equation can be zero either in a trivial way (being the magnitude of gap function zero) or in the case where the gap vector function is coaxial with the stress vector $\boldsymbol{p}$ (or in both of the cases), i.e.,

$$\boldsymbol{g} \times \boldsymbol{p} = \boldsymbol{0} \quad \Longleftrightarrow \quad \begin{cases} g = 0, & \forall \boldsymbol{p}; \\ \boldsymbol{p} = p\,\boldsymbol{n}^x, & \forall g. \end{cases} \tag{5.33}$$

It can hence be appreciated how, when contact occurs (i.e., $g = 0$), the stress vector can be decomposed into a normal and a tangential part along the normal to the current secondary surface

$$\boldsymbol{p}_n := \boldsymbol{n}^x \left(\boldsymbol{p}_n \cdot \boldsymbol{n}^x\right) = p\,\boldsymbol{n}^x; \tag{5.34}$$

$$\boldsymbol{p}_t := \boldsymbol{p} - \boldsymbol{p}_n = p_t\,\boldsymbol{t}^x, \tag{5.35}$$

with $\boldsymbol{t}^x$ being a vector belonging to the plane orthogonal to $\boldsymbol{n}^x$.

Since no limitations have been introduced so far, the magnitude of the stress vector defines these conditions

$$\begin{cases} p < 0, & \text{compression}; \\ p = 0, & \text{zero value}; \\ p > 0, & \text{tension}. \end{cases} \tag{5.36}$$

5.1.4 Contact power

Here it is investigated how the (virtual) power-conjugate of the first Piola-Kirchhoff stress vector can be related to the contact velocity given by Eq. (5.25). As will be shown in Sec-

tion 5.2, the computation of this quantity is pivotal to compute the weak form via a functional including the power of two bodies coming into contact. This quantity is

$$
\begin{aligned}
\mathcal{P}^c &= \int_{\Gamma^{(s)}} \boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) \cdot \dot{\boldsymbol{x}}\left(\boldsymbol{X}, t\right) dA^{(s)} + \int_{\Gamma^{(m)}} \boldsymbol{p}\left(\boldsymbol{\mathcal{Y}}, t, \boldsymbol{N}\left(\boldsymbol{\mathcal{Y}}\right)\right) \cdot \boldsymbol{y}'\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X}, t\right), t\right) dA^{(m)} \\
&= \int_{\Gamma^{(s)}} \boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) \cdot \left(\dot{\boldsymbol{x}}\left(\boldsymbol{X}, t\right) - \boldsymbol{y}'\left(\boldsymbol{\mathcal{Y}}\left(\boldsymbol{X}, t\right), t\right)\right) dA^{(s)} \\
&= \int_{\Gamma^{(s)}} \boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) \cdot \left(\boldsymbol{v}^c - g\dot{\boldsymbol{n}}^x\right) dA^{(s)} \\
&= \int_{\Gamma^{(s)}} \boldsymbol{p}\left(\boldsymbol{X}, t, \boldsymbol{N}\left(\boldsymbol{X}\right)\right) \cdot \left(\boldsymbol{v}^c\right) dA^{(s)},
\end{aligned}
\tag{5.37}
$$

where Eqs. (5.31), (5.25), and (5.32) have been used in the above chain of equations. Hence, any RHS of Eq. (5.37) defines a potential power-conjugate of the first Piola-Kirchhoff stress vector computed on the original secondary surface. As no constitutive relationship between the stress vector and their power-conjugates has been introduced, Eq. (5.37) expresses the *virtual* power.

In can be noticed how the quantities appearing in the RHS of Eq. (5.37) are all defined in the original configuration. These quantities can be mapped to the current configuration according to the usual rules, i.e., passing from a Total Lagrangian (TL) to an Updated Lagrangian (UL) description. However, as this step would not provide any additional information in terms of the knowledge of the described problem, this is not done in the current work.

5.1.5 Tribological laws

The introduction of all kinematic and static quantities of interest is, however, not a sufficient condition to describe the mechanical behaviour. In particular, it is necessary to relate the kinematic quantities to static quantities, i.e., it is necessary to introduce a constitutive relationship between these variables to describe the behaviour of materials at interfaces. When it comes to contact mechanics, these constitutive relationships are named *tribological laws*.

In particular, this work focuses on a normal contact law which does not allow penetration between the bodies in contact and which cannot sustain tension (adhesion behaviour is excluded) at the interface between the materials. This kind of behaviour is described by the *Signorini-Hertz-Moreau* conditions (see Section 5.1.5.1). Furthermore, it is of interest how the tangential behaviour governing two materials in contact can often be described by a slip/stick condition which can be triggered depending on the pressure acting at the interface between the bodies. *Coulomb's friction law* (sometimes referred to as *Amontons-Coulomb*, see Section 5.1.5.2) describes this kind of relationship between normal and tangential behaviour.

5.1.5.1 Normal contact law

As previously mentioned, the normal contact law which describes no-penetration and no-tension between the bodies in contact is described by the Signorini-Hertz-Moreau conditions

(see also Subfigure 5.3(a)), these being

$$g \begin{cases} = 0, & \text{contact;} \\ > 0, & \text{gap;} \end{cases} \qquad \underbrace{}_{\text{no-penetration}} \qquad p \begin{cases} < 0, & \text{compression;} \\ = 0, & \text{zero value;} \end{cases} \qquad \underbrace{}_{\text{no-tension}} \qquad \underbrace{g\,p = 0}_{\text{complementary condition}} \quad , \quad (5.38)$$

and represents the classical Karush-Kuhn-Tucker (KKT) conditions between the magnitude of the normal gap and the magnitude of the normal stress vector.

To have a more compact notation, it is also convenient to introduce the negative part operator on the generic quantity $(\bullet)$

$$[(\bullet)]_- = \begin{cases} -(\bullet), & \text{if } (\bullet) \le 0; \\ 0, & \text{if } (\bullet) > 0, \end{cases} \qquad (5.39)$$

so that the no-tension condition $p \le 0$ can be stated as $[p]_- \ge 0$.

5.1.5.2 Coulomb's friction law

Coulomb's friction law describes the tangential behaviour, presenting stick or slip conditions depending on the applied normal pressure. When it is also assumed that the contact velocity is collinear with the tangent stress vector, this is usually referred to as the isotropic case of Coulomb's law. In terms of KKT conditions (see also Subfigure 5.3(b)), this law can be expressed as

$$\underbrace{\boldsymbol{v}_t^c = -\|\boldsymbol{v}_t^c\| \frac{\boldsymbol{p}_t}{\|\boldsymbol{p}_t\|} \begin{cases} \|\boldsymbol{v}_t^c\| = \boldsymbol{0}, & \text{stick;} \\ \|\boldsymbol{v}_t^c\| \ge \boldsymbol{0}, & \text{slip;} \end{cases}}_{\text{slip rule}} \qquad \underbrace{\|\boldsymbol{p}_t\| + \mu\,p \begin{cases} < 0, & \text{shear;} \\ = 0, & \text{friction;} \end{cases}}_{\text{Coulomb's criterion}}$$

$$\underbrace{\|\boldsymbol{v}_t^c\| \, (\|\boldsymbol{p}_t\| + \mu\,p) = 0,}_{\text{complementarity}} \quad (5.40)$$

where μ is called the *friction coefficient*.

To easily describe the shear/friction inequality, the projection operator $\boldsymbol{P}_{B(\boldsymbol{n},s)}(\boldsymbol{x})$ can be defined by making use of the tangential projection operator in Eq. (5.5)

$$\boldsymbol{P}_{B(\boldsymbol{n},s)}(\boldsymbol{x}) = \begin{cases} \boldsymbol{T}^n \boldsymbol{x}, & \text{if } \|\boldsymbol{T}^n \boldsymbol{x}\| \le s; \\ s \dfrac{\boldsymbol{T}^n \boldsymbol{x}}{\|\boldsymbol{T}^n \boldsymbol{x}\|}, & \text{if } \|\boldsymbol{T}^n \boldsymbol{x}\| > s. \end{cases} \qquad (5.41)$$

In this way, the Coulomb shear/friction stress vector can be described by $\boldsymbol{P}_{B(\boldsymbol{n}^x,\mu p)}(\boldsymbol{p})$.

The minus sign in the description of the isotropic Coulomb's friction law is again due to the ray-projection strategy: the stress vector $\boldsymbol{p}$ is decomposed along a normal and a tangential component defined by the normal direction $\boldsymbol{n}^x$, while the contact velocity is projected along the coordinates defined by $\boldsymbol{n}^y$, as expressed by Eqs. (5.27) and (5.28). This implies that, when considering the power defined by Eq. (5.37) with the above introduced constitutive relationships (i.e., when that power is not virtual, but underlies the constitutive relationships), the sign of Eq. (5.37) must be changed to keep it positive.

(a) Signorini-Hertz-Moreau conditions.

(b) Coulomb's friction law.
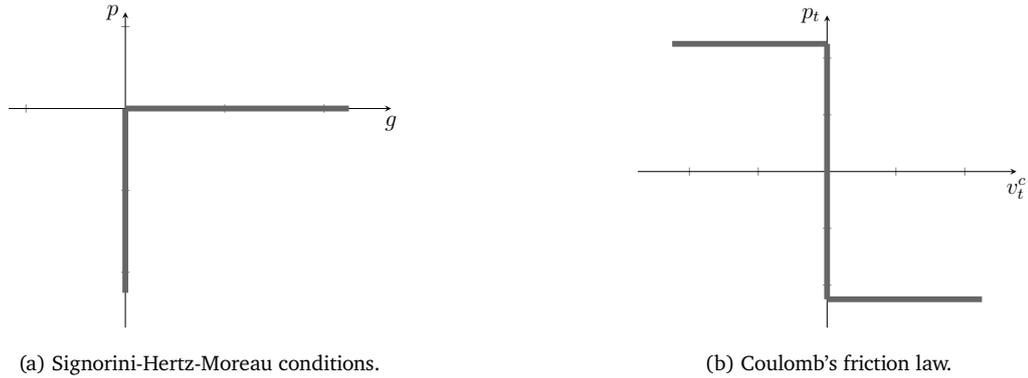
Figure 5.3: Graphical representation of the normal (a) and tangential (b) constitutive behaviour.

5.2 Contact functional

Now that the statics, kinematics and constitutive relationships have been stated, the weak form, which will constitute the basis for the MPM discretisations, is introduced and investigated. When dealing with such weak formulations, the introduction of the spaces of weighting and trial functions is also necessary. However, as the investigation of spaces depends on the chosen formulation and adds even more complexity to the problem under consideration, no hypotheses about the mentioned spaces are made at this stage. This will be considered later, in Section 5.3.1, where the discretised spaces of weighting and trial functions are introduced.

To proceed with the investigation of a weak form including contact and friction, let us first start by introducing the total functional of the union of the two separated bodies, which is given by the sum of their functionals, i.e.,

$$U^u \left(\dot{\boldsymbol{x}}, \boldsymbol{y}' \right) = U^{(s)} \left(\dot{\boldsymbol{x}} \right) + U^{(m)} \left(\boldsymbol{y}' \right). \tag{5.42}$$

The quantities constituting the functionals (such as the rate of internal energy or of kinetic energy, for instance) are not introduced for now, so that the treatment of the equations focuses only on the contact part and it can also be as general as possible.

When no interaction is happening between the two bodies, the stationarity is given by finding $(\dot{\boldsymbol{x}}, \boldsymbol{y}')$ such that*

$$\delta U^u \left(\dot{\boldsymbol{x}}, \boldsymbol{y}' \right) = \begin{cases} \delta U^{(s)} \left(\dot{\boldsymbol{x}} \right) = 0, & \forall \delta \boldsymbol{u}|_X; \\ \delta U^{(m)} \left(\boldsymbol{y}' \right) = 0, & \forall \delta \boldsymbol{u}|_Y. \end{cases} \tag{5.43}$$

As previously stated in Section 5.1.1, it must be highlighted that the points belonging to the secondary surface defined by the variations $\delta \boldsymbol{u}|_{\mathcal{Y}}$ are obtained via the mapping (5.6).

---

*For the sake of clarity, spaces of weighting and trial functions relative to the (real and virtual) displacements are not further characterised in this chapter. While, strictly speaking, they should satisfy some conditions on the Dirichlet boundaries (compare with spaces defined by Eqs. (3.103) and (3.104)), this is not explicitly stated in the above description that intends to keep the explanation as concise as possible.

Points belonging to the secondary body not subject to the considered mapping are instead defined by the variations $\delta \boldsymbol{u}|_Y$.

When interaction obeying Eqs. (5.38) and (5.40) and defined by the ray-tracing strategy (5.6) takes place, the stationarity of the functional (5.42) gives a constrained problem

$$\delta U^u \left(\dot{\boldsymbol{x}}, \boldsymbol{y}' \right) = 0, \quad \forall \delta \boldsymbol{u}|_X, \delta \boldsymbol{u}|_{\mathcal{Y}} : \left\{ \delta g \geq 0 \wedge \| \delta \left(\boldsymbol{v}_t^c \, t \right) \| \geq 0 \wedge \frac{\delta \left(\boldsymbol{v}_t^c \, t \right)}{\| \delta \left(\boldsymbol{v}_t^c \, t \right) \|} = -\frac{\boldsymbol{p}_t}{\| \boldsymbol{p}_t \|} \right\}, \quad (5.44)$$

where the variations of the gap function and the tangential contact velocity are computed via the directional derivatives (see Laursen [2] for this discussion). As such, these variations can be calculated in the same fashion as proposed by their rate in Eq. (5.25), i.e.,

$$\delta g = \boldsymbol{n}^x \cdot \left(\delta \boldsymbol{u}|_{\mathcal{Y}} - \delta \boldsymbol{u}|_X \right); \quad (5.45)$$

$$\delta \left(\boldsymbol{v}_t^c \, t \right) = \boldsymbol{T}^{n^y} \left(\delta \boldsymbol{u}|_X - \delta \boldsymbol{u}|_{\mathcal{Y}} \right). \quad (5.46)$$

It is emphasised that the second of the above equations is simplified to the case when contact occurs, i.e., $g = 0$. Moreover, the time $t \neq 0$ is added to the variation of the contact velocity to keep its dimensions consistent with the variation of the gap function.

As demonstrated by Pietrzak and Curnier [6], the constrained stationary problem (5.44) can be converted to an unconstrained one by considering the power generated during contact by the stress vector and its work-conjugate (Eq. (5.37)). In turn, if it is required that contact obeys the constitutive relationships (5.38) and (5.40), the power has to be modified as follows

$$-\int_{\Gamma^{(s)}} \boldsymbol{p} \cdot \boldsymbol{v}^c \, dA^{(s)} = -\int_{\Gamma^{(s)}} \left(p \, \boldsymbol{n}^x + \boldsymbol{P}_{B(\boldsymbol{n}^x, \mu p)} \left(\boldsymbol{p} \right) \right) \cdot \left(-\dot{g} \boldsymbol{n}^x + \boldsymbol{T}^{n^y} \boldsymbol{v}^c \right) dA^{(s)} =$$

$$= -\int_{\Gamma^{(s)}} \left([p]_- \, \dot{g} + \boldsymbol{P}_{B(\boldsymbol{n}^x, \mu p)} \left(\boldsymbol{p} \right) \cdot \boldsymbol{v}_t^c \right) dA^{(s)}, \quad (5.47)$$

where contact is considered to occur, which implies $\boldsymbol{T}^{n^x} = \left(\boldsymbol{I}^{(2)} - \boldsymbol{n}^x \otimes \boldsymbol{n}^x \right) = \left(\boldsymbol{I}^{(2)} - \left(-\boldsymbol{n}^y \right) \otimes \left(-\boldsymbol{n}^y \right) \right) = \boldsymbol{T}^{n^y}$ and $g = 0$. Hence, the unconstrained functional which considers the normal and tangential tribological laws introduced in Section 5.1.5 becomes

$$U^{sys} \left(\dot{\boldsymbol{x}}, \boldsymbol{y}', \boldsymbol{p} \right) = U^u \left(\dot{\boldsymbol{x}}, \boldsymbol{y}' \right) - \int_{\Gamma^{(s)}} \left([p]_- \, \dot{g} + \boldsymbol{P}_{B(\boldsymbol{n}^x, s)} \left(\boldsymbol{p} \right) \cdot \boldsymbol{v}_t^c \right) dA^{(s)}, \quad (5.48)$$

with its stationarity given by finding $\left(\dot{\boldsymbol{x}}, \boldsymbol{y}' \right)$ such that

$$\delta U^{sys} \left(\dot{\boldsymbol{x}}, \boldsymbol{y}', \boldsymbol{p} \right) = 0, \quad \forall \delta \boldsymbol{u}|_X, \delta \boldsymbol{u}|_{\mathcal{Y}}. \quad (5.49)$$

As pointed out by Poulios and Renard [7] and by Navarro-Jiménez *et al.* [31], it is not possible to obtain Coulomb's friction law via the stationarity expressed by Eq. (5.49). This is the reason why, in Eq. (5.48), the projector operator is written in terms of Tresca slip/stick conditions $\boldsymbol{P}_{B(\boldsymbol{n}^x, s)} \left(\boldsymbol{p} \right)$ (with s being a constant threshold, leading the tangential part of the function to be pressure-independent) in lieu of Coulomb's $\boldsymbol{P}_{B(\boldsymbol{n}^x, \mu p)} \left(\boldsymbol{p} \right)$. Nonetheless, to obtain a weak form considering Coulomb's friction law, the Coulomb projection operator is substituted directly in Eq. (5.49), while it is not used for the functional defined by Eq. (5.48).

Moreover, it can be observed that Eq. (5.49) can be directly used (providing an appropriate space for the Lagrange Multiplier (LM) and its weighting function is used) for the LMM or

for the PM (provided an appropriate *penalty parameter* r and penalty-type constitutive relationships). As the aim of the current work is to present a more general (and also smoother) scenario, the functional described by Eq. (5.48) is augmented to obtain an ALM, i.e.,

$$U^{aug}\left(\dot{\boldsymbol{x}}, \boldsymbol{y}', \boldsymbol{p}\right) = U^u\left(\dot{\boldsymbol{x}}, \boldsymbol{y}'\right)$$
$$+ \frac{1}{2\,r} \int_{\Gamma^{(s)}} \left(\left[\boldsymbol{\lambda} \cdot \boldsymbol{n}^x + r\,g\right]_-^2 - \|\boldsymbol{\lambda}\|^2 + \left\|\boldsymbol{P}_{B(\boldsymbol{n}^x, s)}\left(\boldsymbol{\lambda} - r\boldsymbol{v}^c\right)\right\|^2\right) dA^{(s)}, \quad (5.50)$$

where $\boldsymbol{\lambda}$ is the LM vector. It is worth noticing that, owing to the KKT conditions (5.38) and (5.40) for normal and Coulomb constitutive laws, functionals (5.48) and (5.50) are the same and the LM vector plays the role of the stress vector $\boldsymbol{p}$. Stationarity of (5.50) is given by finding $(\dot{\boldsymbol{x}}, \boldsymbol{y}', \boldsymbol{\lambda})$ such that[*]

$$\begin{cases} \delta U^u\left(\dot{\boldsymbol{x}}, \boldsymbol{y}'\right) - \int_{\Gamma^{(s)}} \left(\left[\boldsymbol{\lambda} \cdot \boldsymbol{n}^x + r\,g\right]_- \delta g + \boldsymbol{P}_{B(\boldsymbol{n}^x, \mu p)}\left(\boldsymbol{\lambda} - r\boldsymbol{v}^c\right) \cdot \delta\left(\boldsymbol{v}_t^c\,t\right)\right) dA^{(s)} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall \delta\boldsymbol{u}|_X, \delta\boldsymbol{u}|_{\mathcal{Y}}; \\ -\frac{1}{r} \int_{\Gamma^{(s)}} \left(\left[\boldsymbol{\lambda} \cdot \boldsymbol{n}^x + r\,g\right]_- \boldsymbol{n}^x + \boldsymbol{\lambda} - \boldsymbol{P}_{B(\boldsymbol{n}^x, \mu p)}\left(\boldsymbol{\lambda} - r\boldsymbol{v}^c\right)\right) \cdot \delta\boldsymbol{\lambda}\, dA^{(s)} = 0, \quad \forall \delta\boldsymbol{\lambda}. \end{cases}$$
$$(5.51)$$

Eq. (5.51) can also be recast in an unsymmetrical way by substituting the second equation into the first, leading to

$$\begin{cases} \delta U^u\left(\dot{\boldsymbol{x}}, \boldsymbol{y}'\right) - \int_{\Gamma^{(s)}} \boldsymbol{\lambda} \cdot \left(\delta\boldsymbol{u}|_X - \delta\boldsymbol{u}|_{\mathcal{Y}}\right) dA^{(s)} = 0, & \forall \delta\boldsymbol{u}|_X, \delta\boldsymbol{u}|_{\mathcal{Y}}; \\ -\frac{1}{r} \int_{\Gamma^{(s)}} \boldsymbol{C}\left(\boldsymbol{\lambda}, \boldsymbol{n}^x, \boldsymbol{v}^c, g\right) \cdot \delta\boldsymbol{\lambda}\, dA^{(s)} = 0, & \forall \delta\boldsymbol{\lambda}, \end{cases} \quad (5.52)$$

having defined the Coulomb constraint as

$$\boldsymbol{C}\left(\boldsymbol{\lambda}, \boldsymbol{n}^x, \boldsymbol{v}^c, g\right) := \left[\boldsymbol{\lambda} \cdot \boldsymbol{n}^x + r\,g\right]_- \boldsymbol{n}^x + \boldsymbol{\lambda} - \boldsymbol{P}_{B(\boldsymbol{n}^x, \mu p)}\left(\boldsymbol{\lambda} - r\boldsymbol{v}^c\right). \quad (5.53)$$

It can be noticed how Eq. (5.52) (or, equally, (5.51)) describes a *biased formulation*, i.e., one where the projection happens from the secondary surface onto the main. Indeed, it can be seen how contact forces are integrated over the secondary surface, while there is no information on the stress vector on the main surface. Thus, in this kind of formulation, the local action-reaction principle described by Eq. (5.31) is violated. This violation, in turn, leads to the loss of coaxiality between the first Piola-Kirchhoff stress vector and the gap vector in case of contact, which is used in the computation of the power generated by two surfaces in contact (see Eq. (5.37)). Following this rationale, it can be expected that some energy will be dissipated during the simulation for biased formulations. This drawback of biased formulations is well-known in the literature (see Curnier *et al.* [5] or Mlika *et al.* [32], among others). On the other hand, *unbiased formulations* do not only rely on the secondary surface, thus eliminating the asymmetry generated by the one-way secondary onto main bijection. Among others, one way to achieve this kind of formulation is presented by Mlika *et al.* [32], and is based on a double mapping procedure (from secondary to main and in the opposite way), so that the stress vectors are computed on both of the surfaces. Since the current formulation is not based on the actual representation of contact surfaces, a biased

---

[*]It can be noticed that the projector operator obeying to Coulomb's friction law has been already substituted in Eq. (5.51).

formulation is preferred in this work, although the limitations introduced by this further approximation are recognised.

The reason why two versions of the same equations are considered, i.e., Eq. (5.51) and Eq. (5.52), is that Eq. (5.51), which represents the most general case, is used as a starting point for the LMM or the PM, while Eq. (5.52) is used only for the ALM. Even though this use of Eqs. (5.51) and (5.52) seems counter-intuitive (the ALM is indeed the most general formulation), considering Eq. (5.52) for the ALM stems from a simplified (thus, easier to compute) Jacobian matrix in the NR iterative method (this will be detailed in Section 5.4.1). Since the LMM and the PM are particularisations of the ALM with less information (i.e., no augmentation for the LMM and no LM for the PM), the simplified Jacobian matrix cannot be used as it lacks some components. This justifies the use of Eq. (5.51) in lieu of Eq. (5.52) for the LMM and the PM.

As already mentioned in Section 5.1.4, the LM $\boldsymbol{\lambda}$ appearing in the above formulations is a first-Piola Kirchhoff measure of stress vector ($\boldsymbol{p}$ in the following equation), which can be related to a Cauchy measure of stress vector as follows

$$\int_{\Gamma^{(s)}} \boldsymbol{p}\, dA^{(s)} = \int_{\Gamma^{(s)}} \boldsymbol{P}\boldsymbol{N}\, dA^{(s)} = \int_{\gamma^{(s)}} \frac{1}{J} \boldsymbol{P}\boldsymbol{F}^{T} \tilde{\boldsymbol{n}}\, da^{s} = \int_{\gamma^{(s)}} \|\tilde{\boldsymbol{n}}\|\, \boldsymbol{\sigma}\boldsymbol{n}\, da^{s}, \qquad (5.54)$$

where Eq. (5.3) and the definition of the Cauchy stress vector have been used. The quantity appearing in the last integral is the Cauchy stress vector.

5.3   Discretisation processes

As already mentioned in Section 3.6.2.1 when discussing the tracking of Neumann boundaries, keeping track of and representing the surface is not a trivial challenge within the MPM, since this method retains no explicit boundary representation.

For the sake of simplicity, this work considers the integrals on the secondary surface appearing in Eqs. (5.51) and (5.52) as integrals on the grid faces, even for those cases where this assumption is not valid. The reason for such an assumption stems from the goal of the current research work, which is to assess different techniques to decide which can be considered the most suitable candidate to be later extended with a proper boundary representation for the MPM.

Given this assumption, both the methods evaluated in the following pages consider the contact on the mesh, in particular mesh nodes for the node-to-node technique and mesh faces for the point-to-surface procedure. Notwithstanding the extreme difference between the methods, they can be cast in a unique framework, as Figure 5.4 suggests. The node-to-node method relies on the Dirac delta distribution as shape and basis functions to cluster the information at the grid nodes so that the LM or the penalty-related quantities are computed only node-wise. On the other hand, the point-to-surface employs linear polynomial shape and basis functions to map information from the mesh faces to their relative nodes. As such, it can be understood how the point-to-surface formulation allows a proper mapping (such

as the ray-tracing technique, explained in Section 5.1.1) between the integration points on the secondary mesh and their relative projected positions on the main meshes. Dirac delta distributions as shape and basis functions for the node-to-node formulation makes any proper projection impracticable since information cannot be moved elsewhere from the node's position, as this is the only place where this distribution is non-zero. Therefore, given the impossibility of a proper mapping procedure for the node-to-node method, nodes of the main and secondary meshes occupying the same initial position at the beginning of the step are coupled *a priori*. This coupling leads to a considerable simplification of the kinematics explained in Section 5.1, which is based upon the assumption of the definition of the main surface via a mapping technique from the secondary one. Details are given in Section 5.3.1.1.

Given all of the differences in the contact discretisations, general space discretisation is introduced in Section 5.3.1, and particularised in Section 5.3.1.1 for the node-to-node and in Section 5.3.1.2 for the point-to-surface method. Time discretisation, which appears to be standard for the considered contact formulations, is introduced in Section 5.3.2.

5.3.1 Space and fields discretisations

Following the same procedure as the space discretisation performed for the dynamic formulation in Chapter 3, MP-based and (Cartesian) grid-based discretisations of the continuum bodies are introduced. To distinguish among the Degrees of Freedom (DoFs) relative to different bodies, let the indices $I, H, P = 1 \ldots n^{dim} \times N^{nds,\,(s)}$ express the grid DoFs for the secondary body (with n^{dim} being the considered Euclidean dimensions and $N^{nds,\,(s)}$ the number of grid nodes relative to the secondary body), while indices $J, K, Q = 1 \ldots n^{dim} \times N^{nds,\,(m)}$ are used for the main body grid DoFs (with $N^{nds,\,(m)}$ being the number of grid nodes relative to the main body). As in Section 3.6.2.1, let the linear polynomial piecewise functions be defined on the grid, while characteristic functions (different depending on whether original MPM or Generalised Interpolation Material Point Method (GIMPM) are considered) are established on the Material Points (MPs). Let convolution integrals of these two sets of functions give shape and basis functions for the two bodies. Finite-dimensional spaces for weighting and trial functions relative to the displacement fields are similar to those defined by Eqs. (3.110) and (3.111).

As previously mentioned, the considered contact surfaces are discretised by the mesh faces and a mapping from the secondary surface onto the main (i.e., a biased formulation) is assumed. Hence, on faces $^h\Gamma^{(s)}$ of the mesh defined by the DoFs $\mathcal{I}, \mathcal{J} = 1 \ldots n^{dim} \times \left(N^{nds,\,(s)} \cap N^{nds,\,(m)} \right)$ with $N^{nds,\,(s)} \cap N^{nds,\,(m)} \neq \emptyset$, let us introduce the finite-dimensional space for the LM functions and of their relative weighting function, this being

$$^h\mathscr{V}^{\boldsymbol{\lambda}} = {}^h\mathscr{W}^{\delta\boldsymbol{\lambda}} = \left\{ {}^h\boldsymbol{\lambda} : \mathrm{dom}\left({}^h\Gamma^{(s)} \right) \right\}, \qquad (5.55)$$

where $\mathrm{dom}\left({}^h\Gamma^{(s)} \right)$ highlights that the Lagrange multiplier is defined only on the contact faces of the mesh $^h\Gamma^{(s)}$. Owing to the introduction of the above spaces, the finite-

dimensional weighting and trial functions can be expressed as follows

$$\delta \boldsymbol{u}^h|_X = \mathbf{N}^u(\boldsymbol{X}) \, \delta \mathbf{u}|_X; \qquad \boldsymbol{u}^h|_X = \mathbf{N}^u(\boldsymbol{X}) \, \mathbf{u}|_X; \tag{5.56}$$

$$\delta \boldsymbol{u}^h|_Y = \mathbf{N}^u(\boldsymbol{Y}) \, \delta \mathbf{u}|_Y; \qquad \boldsymbol{u}^h|_Y = \mathbf{N}^u(\boldsymbol{Y}) \, \mathbf{u}|_Y; \tag{5.57}$$

$$\delta \boldsymbol{u}^h|_{\mathcal{Y}} = \mathbf{N}^u(\boldsymbol{\mathcal{Y}}) \, \delta \mathbf{u}|_Y; \qquad \boldsymbol{u}^h|_{\mathcal{Y}} = \mathbf{N}^u(\boldsymbol{\mathcal{Y}}) \, \mathbf{u}|_Y; \tag{5.58}$$

$$\delta \boldsymbol{\lambda}^h = \mathbf{N}^\lambda(\boldsymbol{X}) \, \delta \boldsymbol{\lambda}; \qquad \boldsymbol{\lambda}^h = \mathbf{N}^\lambda(\boldsymbol{X}) \, \boldsymbol{\lambda}, \tag{5.59}$$

where superscripts $(\bullet)^u$ and $(\bullet)^\lambda$ highlight the difference between the functions interpolating the displacement and the LM fields. Since different functions $\mathbf{N}^\lambda$ will be used depending on the selected projection strategy, the space $^h\mathscr{V}^\lambda$ given by the combinations of these functions is not further characterised at this stage. As in Sections 3.6.2.1 and 4.6.2.1, the dependency of the nodal displacement values with time is dropped for notation clarity (compare, for instance, with Eqs. (4.148)-(4.151)).

Given the introduced discretisations, the following notations are defined to make the equations more compact

$$\hat{\lambda}^n := \left[\mathrm{N}^\lambda_{j\mathcal{J}}(\boldsymbol{X}) \lambda_{\mathcal{J}} \, n^x_j + rg \right]_- ; \tag{5.60}$$

$$\hat{\lambda}^{t,\,tr}_i := T^{n_x}_{ij} \left(\mathrm{N}^\lambda_{j\mathcal{J}}(\boldsymbol{X}) \lambda_{\mathcal{J}} - rv^c_j \right); \tag{5.61}$$

$$\left(P_{B(\boldsymbol{n}^x, \mu\hat{\lambda}^n)} \right)_i (\boldsymbol{\lambda} - r\boldsymbol{v}^c) = \begin{cases} \hat{\lambda}^{t,\,tr}_i, & \text{if } ||\hat{\boldsymbol{\lambda}}^{t,\,tr}|| \le \mu\hat{\lambda}^n \quad \text{stick;} \\[2mm] \mu\hat{\lambda}^n \dfrac{\hat{\lambda}^{t,\,tr}_i}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||}, & \text{otherwise} \qquad \text{slip.} \end{cases} \tag{5.62}$$

Moreover, the discretised gap function and the contact velocity are given by

$$g = n^x_j \left(\mathrm{N}^u_{jK}(\boldsymbol{\mathcal{Y}}) \, \mathrm{y}_K - \mathrm{N}^u_{jH}(\boldsymbol{X}) \, \mathrm{x}_H \right); \tag{5.63}$$

$$v^c_i = \mathbf{N}^u(\boldsymbol{X})_{iH} \, \dot{x}_H - \mathbf{N}^u(\boldsymbol{\mathcal{Y}})_{iK} \, y'_K + g \, \dot{n}^x_i. \tag{5.64}$$

Owing to the discretisations and the use of the more compact notations introduced above, Eqs. (5.51) can be simplified into

$$\mathbf{R} := \begin{cases} \mathrm{r}^{(s)}_I := \mathrm{N}^u_{Ii}(\boldsymbol{X}) \dfrac{\partial U^{(s)}}{\partial \mathrm{u}_i|_X} \\[2mm] \qquad - \int_{^h\Gamma^{(s)}} \mathrm{N}^u_{Ii}(\boldsymbol{X}) \left(-\hat{\lambda}^n \, n^x_i + \left(P_{B(\boldsymbol{n}^x, \mu\hat{\lambda}^n)} \right)_i (\boldsymbol{\lambda} - r\boldsymbol{v}^c) \right) dA^{(s)} \approx \mathbf{0}; \\[3mm] \mathrm{r}^{(m)}_J := \mathrm{N}^u_{Ji}(\boldsymbol{Y}) \dfrac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} \\[2mm] \qquad + \int_{^h\Gamma^{(s)}} \mathrm{N}^u_{Ji}(\boldsymbol{\mathcal{Y}}) \left(-\hat{\lambda}^n \, n^x_i + \left(P_{B(\boldsymbol{n}^x, \mu\hat{\lambda}^n)} \right)_i (\boldsymbol{\lambda} - r\boldsymbol{v}^c) \right) dA^{(s)} \approx \mathbf{0}; \\[3mm] \chi_{\mathcal{I}} := -\dfrac{1}{r} \int_{^h\Gamma^{(s)}} \mathrm{N}^\lambda_{\mathcal{I}i}(\boldsymbol{X}) \, C_i(\boldsymbol{\lambda}, \boldsymbol{v}^c, \boldsymbol{n}^x, g) \, dA^{(s)} \approx \mathbf{0}. \end{cases} \tag{5.65}$$

with the constraint being

$$C_i := \mathrm{N}^\lambda_{i\mathcal{J}} \lambda_{\mathcal{J}} + \hat{\lambda}^n \, n^x_i - \left(P_{B(\boldsymbol{n}^x, \mu\hat{\lambda}^n)} \right)_i (\boldsymbol{\lambda} - r\boldsymbol{v}^c). \tag{5.66}$$

If, on the other hand, non-symmetric Eq. (5.52) is considered, it follows that

$$\tilde{\mathbf{R}} := \begin{cases} \tilde{\mathrm{r}}^{(s)}_I := \mathrm{N}^u_{Ii}(\boldsymbol{X}) \dfrac{\partial U^{(s)}}{\partial \mathrm{u}_i|_X} - \int_{^h\Gamma^{(s)}} \mathrm{N}^u_{Ii}(\boldsymbol{X}) \, \mathrm{N}^\lambda_{i\mathcal{I}}(\boldsymbol{X}) \lambda_{\mathcal{I}} \, dA^{(s)} \approx \mathbf{0}; \\[3mm] \tilde{\mathrm{r}}^{(m)}_J := \mathrm{N}^u_{Ji}(\boldsymbol{Y}) \dfrac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} + \int_{^h\Gamma^{(s)}} \mathrm{N}^u_{Ji}(\boldsymbol{\mathcal{Y}}) \, \mathrm{N}^\lambda_{i\mathcal{I}}(\boldsymbol{X}) \lambda_{\mathcal{I}} \, dA^{(s)} \approx \mathbf{0}; \\[3mm] \tilde{\chi}_{\mathcal{I}} := -\dfrac{1}{r} \int_{^h\Gamma^{(s)}} \mathrm{N}^\lambda_{\mathcal{I}i}(\boldsymbol{X}) \, C_i(\boldsymbol{\lambda}, \boldsymbol{v}^c, \boldsymbol{n}^x, g) \, dA^{(s)} \approx \mathbf{0}. \end{cases} \tag{5.67}$$

(a) Representation of bodies coming into contact, with proper contact boundary representation (dashdotted), active contact nodes and active elements (highlighted by different colours on the undeformed mesh). In this work, the mesh is considered to be fitting the body surfaces, even for cases as the one in the picture.
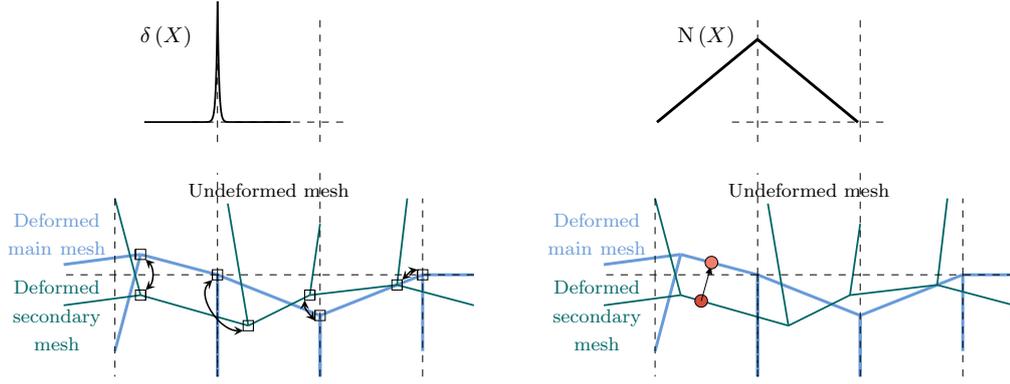
(b) Node-to-node representation: nodes from the secondary mesh are tied to their respective main mesh. The coupling is established by matching the nodes occupying the same position at the beginning of each timestep.

(c) Point-to-surface representation: integration points on the secondary deformed mesh project onto points on the deformed main mesh. Mappings to establish a relationship between these points are necessary and they affect the description of the kinematics involved.

Figure 5.4: Representation of considered quantities and different strategies adopted for the MPM formulations.

While the integrations over the volumes (contained in the variations of $U^{(s)}$ and $U^{(m)}$) follow the same procedure described in Chapter 3, the quadrature rule to approximate the surface integrals in Eqs. (5.65) and (5.67) must still be specified. Since these surface integrators are different in the different formulations, their explanation is given in Sections 5.3.1.1 and 5.3.1.2.

5.3.1.1 Node-to-node formulation

As previously mentioned, the node-to-node formulation considers the space of the Lagrange multiplier trial functions (and their relative weighting functions) to be given by the combination of Dirac delta distributions, i.e.,

$$
{}^{h}\mathcal{V}^{\boldsymbol{\lambda}} = {}^{h}\mathcal{W}^{\delta\boldsymbol{\lambda}} = \left\{ {}^{h}\boldsymbol{\lambda} : \mathrm{dom}\left({}^{h}\Gamma^{(s)}\right) \wedge {}^{h}\lambda_i = \sum_{\mathcal{J}} \delta^{\lambda}_{i\mathcal{J}}\left(\boldsymbol{X}\right)\lambda_{\mathcal{J}} \right\}, \tag{5.68}
$$

where the heuristic definition of the Dirac delta distribution is given by (assuming $\mathbf{X}^h$ is the initial position of the secondary grid node in contact)

$$\delta\left(\boldsymbol{X}\right) \simeq \begin{cases} +\infty & \text{if } \boldsymbol{X} = \mathbf{X}^h; \\ 0 & \text{otherwise,} \end{cases} \tag{5.69}$$

and completed by the following property (particularised to the discretised secondary surface)

$$\int_{{}_h\Gamma^{(s)}} \delta\left(\mathbf{X}\right) dA^{(s)} = 1. \tag{5.70}$$

Since the LM is defined only at the grid nodes where contact is detected according to the space defined by (5.68), the augmentations, i.e., the gap function and the contact velocity, must be computed at the same position for consistency, i.e.,

$$\hat{\lambda}^n = \left[\delta^\lambda_{i\mathcal{J}}\left(\mathbf{X}\right)\lambda_{\mathcal{J}}\, \mathrm{n}_i + r\mathrm{g}\right]_- ; \tag{5.71}$$

$$\hat{\lambda}^{t,\,tr}_i = \mathrm{T}^n_{ij}\left(\delta^\lambda_{j\mathcal{J}}\left(\mathbf{X}\right)\lambda_{\mathcal{J}} - r\mathrm{v}^c_j\right), \tag{5.72}$$

where the definitions (5.63) and (5.64) are particularised to node-to-node formulation as follows

$$\mathrm{g} = \mathrm{n}_j\left(\mathrm{N}^u_{jK}\left(\mathbf{Y}\right)\mathrm{y}_K - \mathrm{N}^u_{jH}\left(\mathbf{X}\right)\mathrm{x}_H\right) = \mathrm{n}_j\left(\mathrm{y}_j - \mathrm{x}_j\right); \tag{5.73}$$

$$\mathrm{v}^c_i = \mathrm{N}^u_{iH}\left(\mathbf{X}\right)\dot{\mathrm{x}}_H - \mathrm{N}^u_{iK}\left(\mathbf{Y}\right)\mathrm{y}'_K + \mathrm{g}\,\dot{\mathrm{n}}_i = \dot{\mathrm{x}}_i - \mathrm{y}'_i + \mathrm{g}\,\dot{\mathrm{n}}_i. \tag{5.74}$$

It is essential to highlight that the dependency of the quantities computed on the main surface and defined by the ray-tracing mapping (denoted by the dependency $\boldsymbol{Y}$) is lost in the above definitions. This applies not only to the gap function and the contact velocity but, more generally, to the whole node-to-node formulation. This is the reason why a simplification of the kinematics occurs for this method. In this case, owing to the lack of a proper mapping, the current position of the (nodal) points on the main surface depends only on its initial position and on time, i.e., $\mathbf{y} = \mathbf{y}\left(\mathbf{Y}, t\right)$. Therefore, there is no difference in the dependencies between the nodal position involved in the contact mechanics and any other position of the nodes belonging to the main body.

It can be noticed that the normal directions $\mathbf{n}$ computed at the grid nodes (and, subsequently, the tangential planes defined by such normal vectors $\mathbf{T}^n$) are not specified in Eqs. (5.73) and (5.74). Moreover, since the mapping between surfaces causes the biased formulations and since no proper mapping occurs for the node-to-node formulation, a sound choice of the normal direction can restore Newton's 3rd law. Following González-Acosta *et al.* [23], this can be done by defining the normal to each body at the A-th contact node as follows

$$\mathrm{n}^x_{\mathsf{A}i} := \frac{1}{\left\|\sum_{mp^{(s)}} m^{mp^{(s)}} \dfrac{\partial \mathrm{N}^u_{\mathsf{A}}(\boldsymbol{X}^{mp^{(s)}})}{\partial \boldsymbol{x}}\right\|} \sum_{mp^{(s)}} m^{mp^{(s)}} \frac{\partial \mathrm{N}^u_{\mathsf{A}}(\boldsymbol{X}^{mp^{(s)}})}{\partial x_i}; \tag{5.75}$$

$$\mathrm{n}^y_{\mathsf{A}i} := \frac{1}{\left\|\sum_{mp^{(m)}} m^{mp^{(m)}} \dfrac{\partial \mathrm{N}^u_{\mathsf{A}}(\boldsymbol{Y}^{mp^{(m)}})}{\partial \boldsymbol{y}}\right\|} \sum_{mp^{(m)}} m^{mp^{(m)}} \frac{\partial \mathrm{N}^u_{\mathsf{A}}(\boldsymbol{Y}^{mp^{(m)}})}{\partial y_i}, \tag{5.76}$$

where $m^{mp^{(m)}}$ and $m^{mp^{(s)}}$ are the MPs belonging to the grid element of the main and secondary bodies (which are in contact). The above-defined normals are then used to compute the unbiased normal to the secondary surface

$$\mathbf{n} := \frac{\mathbf{n}^x - \mathbf{n}^y}{\|\mathbf{n}^x - \mathbf{n}^y\|}. \tag{5.77}$$

The tangent plane to the above-defined direction is defined according to Eq. (5.5), where Eq. (5.77) is substituted to compute the normal to the secondary surface.

Having introduced all of the above quantities and considering that the I-th DoF is given by the A-th node in the i-th dimension (i.e., $I = A \times i$), Eq. (5.65) particularised to the node-to-node formulation reads

$$\mathbf{R} = \begin{cases} \mathrm{r}_I^{(s)} = \mathrm{N}_{Ii}^u(\boldsymbol{X}) \frac{\partial U^{(s)}}{\partial \mathrm{u}_i|_X} - \left(-\hat{\lambda}^n \, \mathrm{n}_I^x + \left(\mathrm{P}_{B(\mathbf{n}, \mu \hat{\lambda}^n)}\right)_I\right)(\boldsymbol{\lambda} - r\mathbf{v}^c)\right) \approx \mathbf{0}; \\ \mathrm{r}_J^{(m)} = \mathrm{N}_{Ji}^u(\boldsymbol{Y}) \frac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} + \left(-\hat{\lambda}^n \, \mathrm{n}_J^x + \left(\mathrm{P}_{B(\mathbf{n}, \mu \hat{\lambda}^n)}\right)_J\right)(\boldsymbol{\lambda} - r\mathbf{v}^c)\right) \approx \mathbf{0}; \\ \chi_{\mathcal{I}} = -\frac{1}{r} \, \mathrm{C}\left(\boldsymbol{\lambda}, \mathbf{v}^c, \mathbf{n}, \mathrm{g}\right) \approx \mathbf{0}, \end{cases} \tag{5.78}$$

or, equivalently, using Eq. (5.67),

$$\tilde{\mathbf{R}} := \begin{cases} \tilde{\mathrm{r}}_I^{(s)} = \mathrm{N}_{Ii}^u(\boldsymbol{X}) \frac{\partial U^{(s)}}{\partial \mathrm{u}_i|_X} - \lambda_I \approx \mathbf{0}; \\ \tilde{\mathrm{r}}_J^{(m)} = \mathrm{N}_{Ji}^u(\boldsymbol{Y}) \frac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} + \lambda_J \approx \mathbf{0}; \\ \chi_{\mathcal{I}} = -\frac{1}{r} \, \mathrm{C}_{\mathcal{I}}\left(\boldsymbol{\lambda}, \mathbf{v}^c, \mathbf{n}, \mathrm{g}\right) \approx \mathbf{0}, \end{cases} \tag{5.79}$$

where, in both of the above systems of equations, the constraint becomes nodal, i.e.,

$$\mathbf{C}\left(\boldsymbol{\lambda}, \mathbf{v}^c, \mathbf{n}, \mathrm{g}\right) = \boldsymbol{\lambda} + \hat{\lambda}^n \, \mathbf{n} - \left(\mathbf{P}_{B(\mathbf{n}, \mu \hat{\lambda}^n)}\right)(\boldsymbol{\lambda} - r\mathbf{v}^c). \tag{5.80}$$

As it appears from the equations, the integration rule is superfluous for the node-to-node formulation, as all of the quantities are nodal, in compliance with the use of the Dirac delta distribution.

The assembling of all the quantities introduced in the node-to-node formulation is given by each individual contact node contribution, except for the normals defined in Eqs. (5.75) and (5.76). These values are calculated by assembling the contribution of each element whose nodes (or at least one among them) are contact nodes.

5.3.1.2 Point-to-surface formulation

In the case of the point-to-surface formulation, the following discrete space for the LMs and their weighting functions is introduced

$${}^h\mathscr{V}^{\boldsymbol{\lambda}} = {}^h\mathscr{W}^{\delta\boldsymbol{\lambda}} = \left\{ {}^h\boldsymbol{\lambda} : \mathrm{dom}\left({}^h\Gamma^{(s)}\right) \wedge {}^h\lambda_i = \sum_{\mathcal{J}} \mathrm{N}_{i\mathcal{J}}^{\lambda,1}(\boldsymbol{X})\lambda_{\mathcal{J}} \right\}, \tag{5.81}$$

with $\mathrm{N}^{\lambda,1}(\boldsymbol{X})$ being the piecewise linear polynomial functions. For the mono-dimensional case, these are defined as

$$\mathrm{N}^{\lambda,1}(X) = \begin{cases} 1 + \frac{X - \mathrm{X}^h}{h} & \text{if } -h < X - \mathrm{X}^h \leq 0; \\ 1 - \frac{X - \mathrm{X}^h}{h} & \text{if } 0 < X - \mathrm{X}^h \leq h; \\ 0 & \text{otherwise,} \end{cases} \tag{5.82}$$
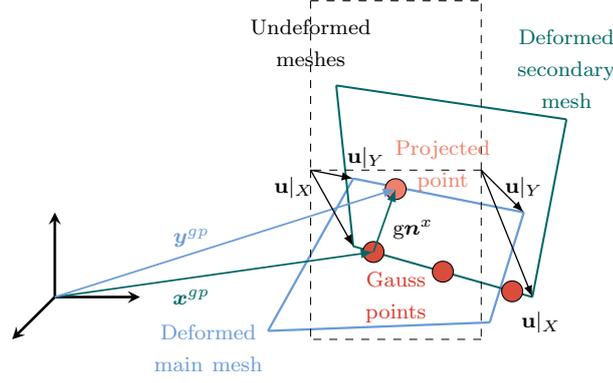
Figure 5.5: Illustration of the ray-tracing strategy ($n^{dim} = 2$) applied to the contact mesh faces. The position of the projected Gauss point on the main surface $\boldsymbol{y}^{gp}$ is computed via Eqs. (5.83) and (5.84). The original position $\boldsymbol{\mathcal{Y}}^{gp}$ is then computed via the interpolated main displacement field on the face, i.e., $\mathbf{N}^u \left(\boldsymbol{\mathcal{Y}}^{gp} \right) \mathbf{u}|_Y$.

with h being the mesh size.

As there is no appreciable simplification of Eqs. (5.65) and (5.67), these are not repeated here.

Since the point-to-surface formulation retains the integrals in Eqs. (5.65) and (5.67), a quadrature rule for these integrals on the secondary surface is required. In this work, the *Gauss-Legendre* quadrature rule is employed. For the integrands referring to the secondary surface, the computation of the weights and positions for the quadrature rule follows the classical procedure for the Gauss-Legendre method. For these quantities, the positions are selected as the appropriate positions of the Gauss points on the parent domain. These positions in the parent domain are straightforwardly mapped to the physical domain following the standard rules, due to the assumption that the mesh faces fit the physical surfaces. However, the other integrands appearing in Eqs. (5.65) and (5.67) refer to the main surface, and their positions cannot be selected according to the quadrature rule, as can be appreciated in Figure 5.5. To determine the position of these quantities, a proper bijection is required. As described in Section 5.1.1 for the ray-tracing strategy, the gap vector function definition (Eq. (5.6)) can be used to determine the projected position on the main surface $\boldsymbol{y}^{gp}$ of the integration points belonging to the secondary surface (defined by the global current position by $\boldsymbol{x}^{gp}$), i.e.,

$$g \, \boldsymbol{n}^x - \boldsymbol{y}^{gp} + \boldsymbol{x}^{gp} = \boldsymbol{0}, \tag{5.83}$$

where the gap distance g and the projected position of the Gauss point on the main surface $\boldsymbol{y}^{gp}$ are unknown. Hence, another equation is necessary to determine these unknowns, which is given by the requirement that the projected position of the Gauss point belongs to one of the main faces of the mesh. This additive requirement is represented by the equation of the plane passing through 3 points, i.e.,

$$a \, y_1^{gp} + b \, y_2^{gp} + c \, y_3^{gp} + d = 0, \tag{5.84}$$

where $\boldsymbol{y}^{gp} = [y_1^{gp}, y_2^{gp}, y_3^{gp}]^T$ and the coefficients a, b, c, d are given by 3 non-collinear mesh

nodes of the main surface[*]. Once $\boldsymbol{y}^{gp}$ is known, its initial position $\boldsymbol{\mathcal{Y}}^{gp}$ can be computed via the displacement field of the main surface $\boldsymbol{u}|_{\mathcal{Y}}$. The relationship between the initial position $\boldsymbol{\mathcal{Y}}^{gp}$ in the physical domain and its relative position in the parent domain follows the standard rules.

From the above descriptions of the positions of the secondary integration points and their projection onto the main surface, it can be appreciated how the assumption of mesh faces fitting the physical surfaces significantly simplifies the above calculations. If this assumption is neglected, it can be seen that an explicit parametrisation (such as the one proposed by Bing *et al.* [26], for instance) of the considered surfaces would be necessary.

Despite these complications in the case where proper boundary representations are performed, these difficulties concern only the treatment of the surfaces while adding no complexity to the basic equations described by the kinematics in Section 5.1.2, the statics in Section 5.1.3, and the tribological law in Section 5.1.5. Therefore, it can be deduced that the introduced simplification changes the overall level of algorithmic complexity for surfaces description, but it does not introduce any simplification in the underlying mathematics of the problem.

5.3.2 Time discretisation

From the definition of the tribological laws in Eqs. (5.38) and (5.40), it can be noted that these relationships are rate-independent. It follows that time discretisation does not affect the part of the functional dealing with friction. However, there are two exceptions to this, which are due to the presence of augmentations (especially the one which considers the tangential contact velocity), and the possible time-dependency of the rest of the functional. According to the hypothesis of finite strain, the configuration must be specified in which the equilibrium and the constraint equations are computed. Since the Conservation Law Consistent (CLC) dynamic formulation described in Chapter 3 provides a sound basis, the mid-point rule is also considered for this chapter. Plugging the frictional contact formulation into the CLC dynamic formulation allows the assessment of the numerical dissipation dealing only with the contact formulation, as no numerical dissipation takes place in the CLC formulation (see Chapter 3).

The mid-point rule follows these time relationships between the generic variable $(\bullet)$ and its time derivatives (compare with Eqs. (3.121)-(3.123))

$$(\bullet)_{\tilde{n}} = \frac{1}{2}\left((\bullet)_n + (\bullet)_{n+1}\right); \tag{5.85}$$

$$(\dot{\bullet})_{\tilde{n}} = \frac{(\bullet)_{n+1} - (\bullet)_n}{\Delta t}; \tag{5.86}$$

$$(\ddot{\bullet})_{\tilde{n}} = \frac{(\dot{\bullet})_{n+1} - (\dot{\bullet})_n}{\Delta t}. \tag{5.87}$$

---

[*]In the case of $n^{dim} = 2$, Eq. (5.84) is particularised to the straight line passing through 2 points $a\,y_1^{gp} + b\,y_2^{gp} + c = 0$, with the coefficients similarly given by the mesh nodes of the main surface.

Applying the above equation to the stationarity of the functional (5.52) leads to

$$
\begin{cases}
\delta U^u\left(\dot{\boldsymbol{x}}_{\tilde{n}}, \boldsymbol{y}'_{\tilde{n}}\right) - \int_{{}^h\Gamma^{(s)}} \boldsymbol{\lambda}_{\tilde{n}} \cdot \left(\delta \boldsymbol{u}|_X - \delta \boldsymbol{u}|_{\mathcal{Y}}\right) dA^{(s)} = 0, & \forall \delta \boldsymbol{u}|_X, \delta \boldsymbol{u}|_{\mathcal{Y}}; \\
-\frac{1}{r}\int_{{}^h\Gamma^{(s)}} \boldsymbol{C}_{\tilde{n}}\left(\boldsymbol{\lambda}_{\tilde{n}}, \boldsymbol{n}^x_{\tilde{n}}, \boldsymbol{v}^c_{\tilde{n}}, g_{\tilde{n}}\right) \cdot \delta\boldsymbol{\lambda}\, dA^{(s)} = 0, & \forall \delta\boldsymbol{\lambda},
\end{cases}
\tag{5.88}
$$

with the constraint being

$$
\boldsymbol{C}_{\tilde{n}}\left(\boldsymbol{\lambda}_{\tilde{n}}, \boldsymbol{n}^x_{\tilde{n}}, \boldsymbol{v}^c_{\tilde{n}}, g_{\tilde{n}}\right) = \left[\boldsymbol{\lambda}_{\tilde{n}} \cdot \boldsymbol{n}^x_{\tilde{n}} + r\, g_{\tilde{n}}\right]_- \boldsymbol{n}^x_{\tilde{n}} + \boldsymbol{\lambda}_{\tilde{n}} - \boldsymbol{P}_{B\left(\boldsymbol{n}^x_{\tilde{n}}, \mu p_{\tilde{n}}\right)}\left(\boldsymbol{\lambda}_{\tilde{n}} - r\boldsymbol{v}^c_{\tilde{n}}\right).
\tag{5.89}
$$

Moreover, the mid-point rule must be applied to the rate quantities appearing in the above equations, i.e.,

$$
\boldsymbol{v}^c_{\tilde{n}} = \dot{\boldsymbol{x}}_{\tilde{n}} - \boldsymbol{y}'_{\tilde{n}} + g_{\tilde{n}}\, \dot{\boldsymbol{n}}^x_{\tilde{n}};
\tag{5.90}
$$

$$
\dot{\boldsymbol{n}}^x_{\tilde{n}} = \frac{\boldsymbol{n}^x\left(\boldsymbol{F}^X_{n+1}\right) - \overbrace{\boldsymbol{n}^x\left(\boldsymbol{F}^X_n\right)}^{:=\boldsymbol{N}^X}}{\Delta t}.
\tag{5.91}
$$

Since specifying the notation relative to the mid-point configuration can make the expressions particularly cumbersome, the subscript $(\bullet)_{\tilde{n}}$ is dropped in the following equations, with few exceptions where time-related quantities are specified to avoid misunderstanding.

5.4 Contact force computation

This section focuses on explaining different methods to compute the contact forces, having introduced the discretisations explained in Section 5.3. Since spatial discretisation can be deployed in a node-to-node or point-to-surface formulation, in this section the considered technique is not specified (unless expressly stated otherwise) so that the discussion is general. Therefore, the interpolant functions for the LM will be generally denoted by $\mathbf{N}^\lambda$, the normal direction to the secondary surface by $\boldsymbol{n}$, and the plane orthogonal to such a direction by $\boldsymbol{T}^n$. In this fashion, node-to-node or point-to-surface formulations can be obtained following the explanations in Sections 5.3.1.1 and 5.3.1.2.

As detailed in the introduction to this chapter, there are various ways of calculating the contact forces, which can be grouped into two broad categories. In the first class of techniques, the interaction between bodies in contact is taken into account by adding the power generated by the contact forces to the functionals describing the individual behaviour of each distinct body. Hence, this family of methods relies on the weak formulation outlined in Section 5.2. The ALM constitutes the most general formulation available which, depending on the assumptions made, can be particularised to the LMM, as detailed in Section 5.4.2 and to the PM, as presented in Section 5.4.3. The PM is further diversified into a linear variant, explained in Section 5.4.3.1 and classically available in the literature (see, for instance, Ding and Schroeder [33]), and a non-linear one, outlined in Section 5.4.3.2.

In the second group of methods, named CVMs and discussed in Section 5.4.4, local forces are calculated based on the local (nodal) impulses exchanged between the bodies in contact. As presented in the literature, this method has no weak formulation. Therefore, this technique deviates from the outline in Sections 5.2 and 5.3, but is rather applied to an already

discretised form of the equilibrium equations. Moreover, since all the necessary quantities for this method are calculated at the mesh nodes, it does not introduce a mapping between the main and secondary surfaces. From this perspective, this method presents some similarities to the node-to-node approach, in the sense that the mesh nodes are coupled *a priori*. As with the node-to-node formulation, it follows that the kinematics is simplified, losing the proper projection between the points on the secondary surface and those on the main surface.

5.4.1 Augmented Lagrangian method (ALM)

As stated in Section 5.2, the ALM constitutes the most general form of the functional, Eq. (5.50), and, consequently, of its first variation, i.e., the weak form, Eq. (5.65) or (5.67). Therefore, the equations required for the ALM have already been expressed and are given by Eq. (5.67) (or (5.65)) for spatial discretisation and Eq. (5.88) for time discretisation.

Since Eq. (5.65) is solved using an implicit method, it is convenient to define only its parts dealing with contact as follows

$$\mathbf{R}^c := \begin{cases} \mathrm{r}_I^{(s),c} := -\int_{h\Gamma^{(s)}} \mathrm{N}_{Ii}^u\left(\boldsymbol{X}\right)\left(-\hat{\lambda}^n\,n_i + \left(P_{B\left(\boldsymbol{n},\mu\hat{\lambda}^n\right)}\right)_i \left(\boldsymbol{\lambda} - r\boldsymbol{v}^c\right)\right) dA^{(s)} \approx \mathbf{0}; \\[2mm] \mathrm{r}_J^{(m),c} := +\int_{h\Gamma^{(s)}} \mathrm{N}_{Ji}^u\left(\boldsymbol{\mathcal{Y}}\right)\left(-\hat{\lambda}^n\,n_i + \left(P_{B\left(\boldsymbol{n},\mu\hat{\lambda}^n\right)}\right)_i \left(\boldsymbol{\lambda} - r\boldsymbol{v}^c\right)\right) dA^{(s)} \approx \mathbf{0}; \\[2mm] \chi_{\mathcal{I}} := -\frac{1}{r}\int_{h\Gamma^{(s)}} \mathrm{N}_{\mathcal{I}i}^\lambda\left(\boldsymbol{X}\right) C_i\left(\boldsymbol{\lambda}, \boldsymbol{v}^c, \boldsymbol{n}, g\right) dA^{(s)} \approx \mathbf{0}, \end{cases}$$

(5.92)

or, equivalently, for Eq. (5.67)

$$\tilde{\mathbf{R}}^c := \begin{cases} \tilde{\mathrm{r}}_I^{(s),c} := -\int_{h\Gamma^{(s)}} \mathrm{N}_{Ii}^u\left(\boldsymbol{X}\right) \mathrm{N}_{i\mathcal{I}}^\lambda\left(\boldsymbol{X}\right) \lambda_{\mathcal{I}}\, dA^{(s)} \approx \mathbf{0}; \\[2mm] \tilde{\mathrm{r}}_J^{(m),c} := \mathrm{N}_{Ji}^u\left(\boldsymbol{Y}\right) \frac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} + \int_{h\Gamma^{(s)}} \mathrm{N}_{Ji}^u\left(\boldsymbol{\mathcal{Y}}\right) \mathrm{N}_{i\mathcal{I}}^\lambda\left(\boldsymbol{X}\right) \lambda_{\mathcal{I}}\, dA^{(s)} \approx \mathbf{0}; \\[2mm] \chi_{\mathcal{I}} := -\frac{1}{r}\int_{h\Gamma^{(s)}} \mathrm{N}_{\mathcal{I}i}^\lambda\left(\boldsymbol{X}\right) C_i\left(\boldsymbol{\lambda}, \boldsymbol{v}^c, \boldsymbol{n}, g\right) dA^{(s)} \approx \mathbf{0}. \end{cases}$$

(5.93)

If the NR procedure is considered, the linearisation of the equations requires that

$$\mathbf{R}^{(k+1)} \approx \mathbf{R}^{(k)} + \underbrace{\frac{\partial \mathbf{R}^{(k)}}{\partial \mathbf{x}}}_{:=\mathbf{J}} \underbrace{\left(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\right)}_{:=\delta\mathbf{x}} \approx \mathbf{0},$$

(5.94)

where $\mathbf{J}$ defines the Jacobian matrix of the system, and $\delta\mathbf{x}$ its incremental solution (having assumed $\mathbf{x} := [\mathbf{u}|_X, \mathbf{u}|_Y, \boldsymbol{\lambda}]^T$). As a result of the introduction of the parts that only involve contact, it is possible to consider the equilibrium equations $\mathbf{R}$ as the sum of a part given by the contact $\mathbf{R}^c$ and the remaining $(\mathbf{R} - \mathbf{R}^c)$. In this way, the Jacobian matrix can also be written as

$$\mathbf{J} = \mathbf{J}^c + (\mathbf{J} - \mathbf{J}^c),$$

(5.95)

where the entries of the Jacobian relative to frictional contact are given by

$$
\mathbf{J}^c =
\begin{bmatrix}
\dfrac{\partial \mathrm{r}_I^{(s),c}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \mathrm{r}_I^{(s),c}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \mathrm{r}_I^{(s),c}}{\partial \lambda_{\mathcal{J}}} \\[2em]
\dfrac{\partial \mathrm{r}_J^{(m),c}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \mathrm{r}_J^{(m),c}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \mathrm{r}_J^{(m),c}}{\partial \lambda_{\mathcal{J}}} \\[2em]
\dfrac{\partial \chi_{\mathcal{I}}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \chi_{\mathcal{I}}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \chi_{\mathcal{I}}}{\partial \lambda_{\mathcal{J}}}
\end{bmatrix} , \tag{5.96}
$$

The terms $(\mathbf{J} - \mathbf{J}^c)$ are left unknown for general purposes, as these parts of the Jacobian matrix deals with the unexpressed remaining part of the stationarity of the functional (i.e., the balances of rate of linear momentum on each bodies excluded the linearisation of contact forces). Equally, the same procedure can be applied when considering Eq. (5.93), resulting in

$$
\mathbf{J} = \tilde{\mathbf{J}}^c + \left(\mathbf{J} - \tilde{\mathbf{J}}^c \right) , \tag{5.97}
$$

with

$$
\tilde{\mathbf{J}}^c =
\begin{bmatrix}
\mathbf{0} & \mathbf{0} & \dfrac{\partial \tilde{\mathrm{r}}_I^{(s),c}}{\partial \lambda_{\mathcal{J}}} \\[2em]
\dfrac{\partial \tilde{\mathrm{r}}_J^{(m),c}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \tilde{\mathrm{r}}_J^{(m),c}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \tilde{\mathrm{r}}_J^{(m),c}}{\partial \lambda_{\mathcal{J}}} \\[2em]
\dfrac{\partial \chi_{\mathcal{I}}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \chi_{\mathcal{I}}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \chi_{\mathcal{I}}}{\partial \lambda_{\mathcal{J}}}
\end{bmatrix} . \tag{5.98}
$$

Entries of the Jacobian matrices (5.96) are listed in Appendix B.3. It can be now fully appreciated how the null entries in (5.98) justify why the reduced version defined by Eq. (5.67) cannot be used for the LMM and the PM. For the former method, there is no certainty that the derivatives of χ with respect to displacements are null (they are non-zero in the Jacobian matrix (5.96) often due only to the augmentations). For the latter, i.e., the PM, the Jacobian matrix is reduced to the first $I + J$-th rows and the first $H + K$-th columns, which makes the equilibrium equations for the secondary body uncoupled from those of the main body.

5.4.2 Lagrange Multiplier Method (LMM)

The functional for the LMM can be obtained either by vanishing the augmentations of the functional defined by (5.65) for the ALM or, equally, by introducing the LM as a stress vector, i.e., $\boldsymbol{p} = \boldsymbol{\lambda}$ into functional Eq. (5.48). Consequently, the space of functions defined by (5.55) for this quantity must be specified as well. In both of the described cases, by applying the discretisations described in Section 5.3, the stationarity of these potentials is

given by finding the displacement fields for each body, $\mathbf{u}|_X, \mathbf{u}|_Y \in {}^h\mathscr{V}^{\boldsymbol{u}}$, and the Lagrange multiplier $\boldsymbol{\lambda} \in {}^h\mathscr{V}^{\boldsymbol{\lambda}}$ such that

$$\mathbf{R} = \begin{cases} \mathrm{r}_I^{(s)} = \mathrm{N}_{Ii}^u\left(\boldsymbol{X}\right) \frac{\partial U^{(s)}}{\partial \mathrm{u}_i|_X} - \int_{h\Gamma^{(s)}} \mathrm{N}_{Ii}^u\left(\boldsymbol{X}\right)\left(-\lambda^n\, n_i + \left(P_{B(\boldsymbol{n},\mu\lambda^n)}\right)_i\left(\boldsymbol{\lambda}\right)\right)\, dA^{(s)} \approx \mathbf{0}; \\[2mm] \mathrm{r}_I^{(m)} = \mathrm{N}_{Ji}^u\left(\boldsymbol{Y}\right) \frac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} + \int_{h\Gamma^{(s)}} \mathrm{N}_{Ji}^u\left(\boldsymbol{\mathcal{Y}}\right)\left(-\lambda^n\, n_i + \left(P_{B(\boldsymbol{n},\mu\lambda^n)}\right)_i\left(\boldsymbol{\lambda}\right)\right)\, dA^{(s)} \approx \mathbf{0}; \\[2mm] \chi_{\mathcal{I}} = -\frac{1}{r}\int_{h\Gamma^{(s)}} \mathrm{N}_{\mathcal{I}i}^\lambda\left(\boldsymbol{X}\right) \mathcal{C}_i\left(\boldsymbol{\lambda},\boldsymbol{n}\right)\, dA^{(s)} \approx \mathbf{0}, \end{cases} \tag{5.99}$$

with

$$\lambda^n := \left[\mathrm{N}_{j\mathcal{J}}^\lambda\left(\boldsymbol{X}\right) \lambda_{\mathcal{J}}\, n_j\right]_- ; \tag{5.100}$$

$$\hat{\lambda}_i^{t,\,tr} := T_{ij}^n\left(\mathrm{N}_{j\mathcal{J}}^\lambda\left(\boldsymbol{X}\right)\lambda_{\mathcal{J}}\right); \tag{5.101}$$

$$\boldsymbol{P}_{B(\boldsymbol{n}^x,\mu\lambda^n)}\left(\boldsymbol{\lambda}\right) = \begin{cases} \boldsymbol{\lambda}^{t,\,tr}, & \text{if } ||\boldsymbol{\lambda}^{t,\,tr}|| \leq \mu\lambda^n \quad \text{stick}; \\[3mm] \mu\lambda^n \dfrac{\boldsymbol{\lambda}^{t,\,tr}}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||}, & \text{otherwise} \qquad\qquad \text{slip}; \end{cases} \tag{5.102}$$

$$\mathcal{C}_i = \mathrm{N}_{i\mathcal{J}}^\lambda \lambda_{\mathcal{J}} + \lambda^n\, n_i - \left(P_{B(\boldsymbol{n},\mu\lambda^n)}\right)_i\left(\boldsymbol{\lambda}\right). \tag{5.103}$$

It is well-known that, while satisfying the KKT conditions for the normal and the tangential behaviour, the LMM can struggle to converge, especially in correspondence of the loss of strong convexity of the functional (see Chen *et al.* [18]). As underlined by Burman and Hansbo [34], the augmentations present in Eqs. (5.60) and (5.61) but absent in Eqs. (5.100) and (5.101) strengthen the contact conditions, even though this does not imply a stabilisation of the method.

Since the Jacobian matrix for the LMM can be deduced by vanishing the augmentations in the entries of matrix (5.96), this is not explicitly reported here.

5.4.3 Penalty Method (PM)

As with the LMM, the functional for the PM can be obtained by not considering the LM in Eq. (5.65) or by considering the penalty constitutive relations for the stress vector in the functional (5.48). Introducing the discretisations described in Section 5.3, the stationarity of these potentials is given by finding the displacement fields for each body, $\mathbf{u}|_X, \mathbf{u}|_Y \in {}^h\mathscr{V}^{\boldsymbol{u}}$, such that

$$\mathbf{R} = \begin{cases} \mathrm{r}_I^{(s)} = \mathrm{N}_{Ii}^u\left(\boldsymbol{X}\right) \frac{\partial U^{(s)}}{\partial \mathrm{u}_i|_X} \\[2mm] \qquad\quad - \int_{h\Gamma^{(s)}} \mathrm{N}_{Ii}^u\left(\boldsymbol{X}\right)\left(-[p\left(g\right)]_-\, n_i + \left(P_{B(\boldsymbol{n},\mu[p(g)]_-)}\right)_i\left(\boldsymbol{p}\left(g,\boldsymbol{v}^c\right)\right)\right)\, dA^{(s)} \approx \mathbf{0}; \\[3mm] \mathrm{r}_J^{(m)} = \mathrm{N}_{Ji}^u\left(\boldsymbol{Y}\right) \frac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} \\[2mm] \qquad\quad + \int_{h\Gamma^{(s)}} \mathrm{N}_{Ji}^u\left(\boldsymbol{X}\right)\left(-[p\left(g\right)]_-\, n_i + \left(P_{B(\boldsymbol{n},\mu[p(g)]_-)}\right)_i\left(\boldsymbol{p}\left(g,\boldsymbol{v}^c\right)\right)\right)\, dA^{(s)} \approx \mathbf{0}. \end{cases} \tag{5.104}$$

As can be noticed from the above equations, the PM is a an irreducible approach, where only the displacement fields are primary variables. This feature can be particularly attractive not only because it reduces the number of unknowns, but also because it removes the non-trivial discussion about the *inf-sup* condition to satisfy for mixed formulations.

Since the constitutive relationships of the penalty kind considered in this work are different, they are not reported here, but are specified for the linear case in Section 5.4.3.1 and the non-linear case in Section 5.4.3.2.

5.4.3.1   Linear Penalty Method (PM)

If the constitutive equations between the stress vector and the associated kinematic variables imitating the behaviour of Eqs. (5.38) and (5.40) are of the linear spring type (see Figure 5.6), they are expressed by the following equations

$$[p(g)]_- = -r\,g; \tag{5.105}$$

$$p_i^{t,tr} = -r\,T_{ij}^n\,v_j^c = -r\,(v_t^c)_i\,; \tag{5.106}$$

$$\left(P_{B(\boldsymbol{n}^x, -\mu\,r\,g)}\right)_i (\boldsymbol{p}(g,\boldsymbol{v}^c)) = \begin{cases} p_i^{t,tr}, & \text{if } ||\boldsymbol{p}^{t,tr}|| \le -\mu\,r\,g \quad \text{stick;} \\ -\mu\,r\,g\,\dfrac{p_i^{t,tr}}{||\boldsymbol{p}^{t,tr}||}, & \text{otherwise} \qquad\qquad \text{slip.} \end{cases} \tag{5.107}$$

(a) Linear PM: normal behaviour.   (b) Linear PM: tangential behaviour.

Figure 5.6: Linear PM approximating the no-penetration no-tension normal behaviour, and the Coulomb's friction law.

From a visual comparison between Figures 5.3 and 5.6, it can also be seen that the compatibility conditions cannot be fulfilled for the introduced penalty-type equations. Mathematically, this can be seen by substituting Eqs. (5.105) and (5.106) into (5.38) and (5.40). By doing so, it is clear how the compatibility conditions can only be approximately close to zero, but not strictly zero (see, for further discussion, Laursen and Chawla [35]).

From a physical point of view, the violation of the compatibility conditions allows bodies to overlap (violation in the normal direction) and slip for values of tangential forces lower than those allowed by a Coulomb-type condition (violation in the tangential direction).

For both cases, the violation is proportional to the penalty parameter r: the higher r, the lower the violation of the compatibility conditions. It would therefore be quite tempting to set high values for the penalty parameter. However, as extensively documented in the literature (see Chen *et al.* [18], or Laursen [2], for instance), excessively high values for

the penalty parameter can lead to excessively high values in some entries of the Jacobian and, thus, to an ill-conditioned Jacobian. In conclusion, it is necessary to take both of these opposite tendencies into account; on the one hand, setting high values to violate the compatibility conditions as little as possible, and, on the other hand, moderating the same value to achieve convergence.

The Jacobian matrix for the linear PM can again be obtained via (5.96), neglecting the influence of the LM (third column) and the rows occupied by χ (third row).

5.4.3.2 Non-linear Penalty Method (PM)

The introduction of a PM more complex than the linear one presented in Section 5.4.3.2 can be justified by multiple reasons. Firstly, it is intended to make the transition of on/off conditions gentler, as is the case with contact/separation or slip/stick (see Figure 5.7). This modification should allow, in turn, a smoother convergence of the NR algorithm. In addition, having a non-linear PM allows further adaptation with regard to the normal behaviour. As pointed out by *Zavarise et al.* [36], selecting high penalty parameters to violate KKT conditions as little as possible can lead not only to ill-conditioned Jacobian matrices, as previously reported in Section 5.4.3.1. Indeed, it may also happen that the predictor forces in the normal direction are overestimated, especially in the first iterations, where the overlap between bodies may be pronounced. This overestimation can cause the NR not to converge, as these predictors are too far from the exact solution. A non-linear formulation, on the other hand, allows for a harder material in correspondence with small overlaps, and a softer material for larger values, i.e., $r_n^{in} > r_n^{fin}$. In this fashion, predictors (which can end up in the less stiff region, since the initial overlap is usually considerable) are not overestimated. At the same time, KKT conditions are enforced more strictly in the successive iterations (with the overlap decreasing) in the part where stiffer behaviour is considered.

(a) Non-linear PM: normal behaviour. (b) Non-linear PM: tangential behaviour.
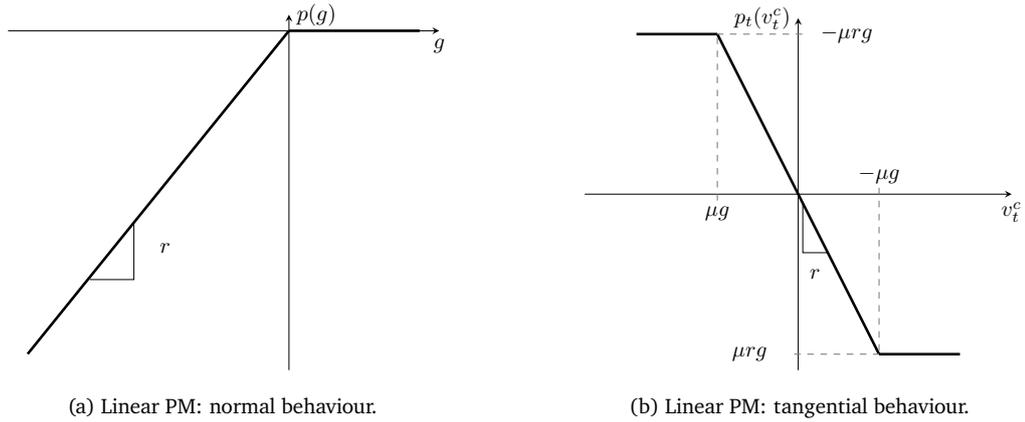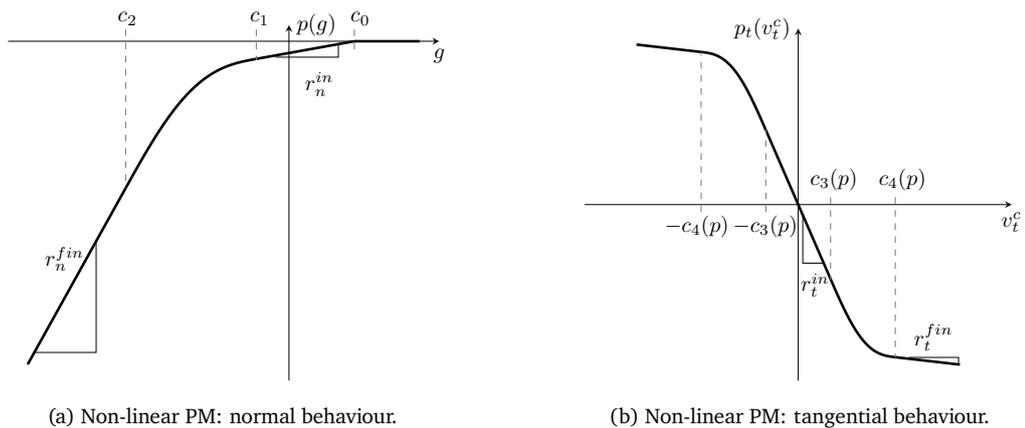
Figure 5.7: Non-linear PM approximating the no-penetration no-tension normal behaviour, and the Coulomb's friction law.

In this work, it is assumed that the normal constitutive behaviour for the non-linear PM is

given by

$$
p(g) = \begin{cases}
0, & \text{if } g > c_0 \\
r_n^{in}(-c_0 + g), & \text{if } c_0 \geq g > c_1; \\
\alpha_0 + \alpha_1 g + \alpha_2 g^2 + \alpha_3 g^3 + \alpha_4 g^4, & \text{if } c_1 \geq g > c_2; \\
r_n^{in} c_0 - \frac{1}{2}\left(r_n^{fin} - r_n^{in}\right)(c_1 + c_2) + r_n^{fin} g, & \text{otherwise,}
\end{cases}
\tag{5.108}
$$

where values of the gap functions c_0, c_1 and c_2 are user-selected, while the coefficients for the 4th order polynomial are computed to allow up to second-order derivatives to be continuous with the piecewise linear parts, i.e.,

$$
\alpha_0 = \frac{(r_n^{fin} - r_n^{in})c_1^4 + 2c_1^3\left(c_2\left(-r_n^{fin} + r_n^{in}\right) + c_0 r_n^{in}\right) - 2 c_0 c_2 r_n^{in}\left(3 c_1^2 + 3 c_1 c_2 - c_2^2\right)}{2(c_2 - c_1)^3};
\tag{5.109}
$$

$$
\alpha_1 = \frac{c_2^3 r_n^{in} - 3 c_1 c_2^2 r_n^{in} + 3 c_1^2 c_2 r_n^{fin} - c_1^3 r_n^{fin}}{(c_2 - c_1)^3};
\tag{5.110}
$$

$$
\alpha_2 = -\frac{3 c_1 c_2 (r_n^{fin} - r_n^{in})}{(c_2 - c_1)^3};
\tag{5.111}
$$

$$
\alpha_3 = \frac{(r_n^{fin} - r_n^{in})(c_1 + c_2)}{(c_2 - c_1)^3};
\tag{5.112}
$$

$$
\alpha_4 = -\frac{r_n^{fin} - r_n^{in}}{2(c_2 - c_1)^3}.
\tag{5.113}
$$

It can also be noticed that setting $c_0 > 0$ allows for some interaction at a distance between bodies.

For the tangential behaviour of the non-linear PM, it is assumed that

$$
p_t = \begin{cases}
-\frac{1}{2}\left(r_t^{fin} - r_t^{in}\right)(c_3(p) + c_4(p)) - r_t^{fin} v_t^c, & \text{if } v_t^c \leq -c_4(p); \\
-\beta_0(p) + \beta_1(p) v_t^c - \beta_2(p)\left(v_t^c\right)^2 + \beta_3(p)\left(v_t^c\right)^3 - \beta_4(p)\left(v_t^c\right)^4, & \text{if } -c_4(p) < v_t^c \leq -c_3(p); \\
-r_t^{in} v_t^c, & \text{if } -c_3(p) < v_t^c \leq c_3(p); \\
\beta_0(p) + \beta_1(p) v_t^c + \beta_2(p)\left(v_t^c\right)^2 + \beta_3(p)\left(v_t^c\right)^3 + \beta_4(p)\left(v_t^c\right)^4, & \text{if } c_3(p) < v_t^c \leq -c_4(p); \\
\frac{1}{2}\left(r_t^{fin} - r_t^{in}\right)(c_3(p) + c_4(p)) - r_t^{fin} v_t^c, & \text{if } v_t^c > c_4(p),
\end{cases}
\tag{5.114}
$$

where

$$
\beta_0(p) = \frac{c_3(p)^3\left(r_t^{fin} - r_t^{in}\right)(c_3 - 2 c_4)}{2\left(c_3(p) - c_4(p)\right)^3};
\tag{5.115}
$$

$$
\beta_1(p) = \frac{-c_3(p)^3 r_t^{fin} + 3c_3(p) c_4(p)\left(c_3(p) r_t^{fin} - c_4(p) r_t^{in}\right) + c_4(p)^3 r_t^{in}}{\left(c_3(p) - c_4(p)\right)^3};
\tag{5.116}
$$

$$
\beta_2(p) = -\frac{-c_3(p) c_4(p)\left(r_t^{fin} - r_t^{in}\right)}{\left(c_3(p) - c_4(p)\right)^3};
\tag{5.117}
$$

$$
\beta_3(p) = \frac{\left(r_t^{fin} - r_t^{in}\right)(c_3(p) + c_4(p))}{\left(c_3(p) - c_4(p)\right)^3};
\tag{5.118}
$$

$$\beta_4(p) = -\frac{\left(r_t^{fin} - r_t^{in}\right)}{2\left(c_3(p) - c_4(p)\right)^3}.$$ (5.119)

The coefficients of the above polynomials follow the same rationale explained for their normal counterparts. As this formulation is a PM and, as such, it constitutes an approximation of the KKT conditions, it is debatable whether the slip conditions begin at $c_3(p)$ or $c_4(p)$. In this paper, the former option is chosen, which implies that $c_3(p) = -\mu\,p(g)/r_t^{in}$. The value of $c_4(p)$ is then chosen as a multiple of $c_3(p)$.

5.4.4 Contact Velocity-based Method (CVM)

As this method belongs to a family of techniques entirely different form the Augmented Lagrangian group, the CVM is not based on the stationarity of any functional. Thus, it differs from the presentations made so far in Sections 5.2 and 5.3, as the contact forces are added to a previously discretised version of the stationarity of the functionals of the two separate bodies. Furthermore, to the author's best knowledge, this method has only been applied in the MPM (see [8–17, 19–24, 29]), referring to the grid for applying both contact and friction conditions. As it is based on nodal values, it has similarities with node-to-node formulations in the sense that no mapping between contact surfaces takes place properly, but the mesh nodes are coupled *a priori*. This method can also become unbiased by following the same steps as for the node-to-node formulation (uniquely defined normal and tangent plane). The CVM is a formulation in which the primary variables are only given by the displacement field, thus differing in this respect from mixed formulations such as ALM and LMM.

This kind of method is based on the calculation of the following quantity at the A-th node

$$\tilde{\mathsf{v}}_{\mathsf{A}i}^c := \frac{1}{\left(\bar{\mathsf{M}}^{(s)} + \bar{\mathsf{M}}^{(m)}\right)|^{\mathsf{AA}}} \left(\bar{\mathsf{M}}^{(s)}|^{\mathsf{AA}} \left(\dot{\mathsf{u}}|_X\right)|_i^{\mathsf{A}} + \bar{\mathsf{M}}^{(m)}|^{\mathsf{AA}} \left(\mathsf{u}'|_Y\right)|_i^{\mathsf{A}} \right) =$$

$$= \begin{cases} \left(\dot{\mathsf{u}}|_X\right)|_i^{\mathsf{A}}, & \text{if } \mathsf{A} \in {}^h\omega^{(s)}; \\ \left(\mathsf{u}'|_Y\right)|_i^{\mathsf{A}}, & \text{if } \mathsf{A} \in {}^h\omega^{(m)}; \qquad\qquad \text{(no sum over A)} \\ \neq \left(\dot{\mathsf{u}}|_X\right)|_i^{\mathsf{A}} \ \wedge \ \neq \left(\mathsf{u}'|_Y\right)|_i^{\mathsf{A}}, & \text{if } \mathsf{A} \in {}^h\omega^{(s)} \cap {}^h\omega^{(m)}, \end{cases}$$ (5.120)

with $\bar{\mathsf{M}}^{(m)}$ and $\bar{\mathsf{M}}^{(s)}$ being the main and secondary lumped matrices (see Eq. (3.162) for the definition of the lumped mass matrix). In the literature, Eq. (5.120) is often called contact velocity (or, similarly, average velocity) but, as noticeable, it is defined very differently from that introduced in Eq. (5.25). It can be seen that this velocity is a proxy for identifying the nodes at which contact dynamics may occur, since it differs from the nodal velocities of individual bodies only at the nodes where both bodies are mapped.

Nevertheless, a further step is necessary to establish whether this velocity is properly a contact velocity or a separation one. For instance, if the secondary body is considered, this

is performed by checking the following inequality

$$
\left((\dot{\mathrm{u}}|_X) \,|_i^{\mathsf{A}} - (\tilde{\mathrm{v}}^c) \,|_i^{\mathsf{A}} \right) \mathrm{n}|_i^{\mathsf{A}}
\begin{cases}
> 0, & \text{contact;} \\
\text{otherwise,} & \text{separation.}
\end{cases}
\qquad \text{(no sum over A)} \qquad (5.121)
$$

The contact forces added directly to the separate bodies equilibrium equations are computed proportionally to the difference between the velocity of the body and the contact velocity defined by (5.120). If, again, the secondary body is taken into account, it gives

$$
\left(\mathrm{f}_c^{(s)} \right)_I := \left(\mathrm{f}_c^{(s),n} \right)_I + \left(\mathrm{P}_{B\left(\mathbf{n}, \left\| \left(\mathbf{f}_c^{(s),n} \right) \right\| \right)} \right)_I \left(\mathbf{f}_c^{(s)} \right); \tag{5.122}
$$

$$
\left(\mathrm{f}_c^{(s),n} \right)_J := \frac{1}{\Delta t} \underbrace{\tilde{\mathrm{M}}_{JI}}_{\tilde{\mathrm{M}}_{\mathsf{B}j\mathsf{A}i}} \underbrace{\mathrm{n}_I}_{=\mathrm{n}_{\mathsf{A}i}} \left(\mathrm{n}|_h^{\mathsf{A}} \left((\dot{\mathrm{u}}|_X) \,|_h^{\mathsf{A}} - (\tilde{\mathrm{v}}^c) \,|_h^{\mathsf{A}} \right) \right); \tag{5.123}
$$

$$
\left(\mathrm{f}_c^{(s),tr} \right)_I := \underbrace{\mathrm{T}_{IJ}^{\mathrm{n}}}_{=\mathrm{T}_{\mathsf{A}i\mathsf{B}j}^{\mathrm{n}}} \underbrace{\left(\mathrm{f}_c^{(s)} \right)_J}_{\left(\mathrm{f}_c^{(s)} \right)_{\mathsf{B}j}}; \tag{5.124}
$$

$$
\left(\mathrm{P}_{B\left(\mathbf{n}, \left\| \left(\mathbf{f}_c^{(s),n} \right) \right\| \right)} \right)_I \left(\mathbf{f}_c^{(s)} \right) :=
\begin{cases}
\left(\mathrm{f}_c^{(s),tr} \right)_I, & \text{if } \left\| \left(\mathrm{f}_c^{(s),tr} \right) \,|_i^{\mathsf{A}} \right\| \leq \mu \left\| \left(\mathrm{f}_c^{(s),n} \right) \,|_i^{\mathsf{A}} \right\| \quad \text{stick;} \\[3mm]
\mu \left\| \left(\mathrm{f}_c^{(s),n} \right) \,|_i^{\mathsf{A}} \right\| \dfrac{\left(\mathrm{f}_c^{(s),tr} \right) |_i^{\mathsf{A}}}{\left\| \left(\mathbf{f}_c^{(s),tr} \right) |^{\mathsf{A}} \right\|}, & \text{otherwise} \quad \text{slip,}
\end{cases}
$$
$$\tag{5.125}$$

where, in particular, equilibrium equations take the form

$$
\begin{cases}
\mathrm{r}_I^{(s)} = \mathrm{N}_{Ii}^u \left(\boldsymbol{X} \right) \dfrac{\partial U^{(s)}}{\partial \mathrm{u}_i|_X} - \left(\mathrm{f}_c^{(s)} \right)_I; \\[3mm]
\mathrm{r}_J^{(m)} = \mathrm{N}_{Ji}^u \left(\boldsymbol{Y} \right) \dfrac{\partial U^{(m)}}{\partial \mathrm{u}_i|_Y} + \left(\mathrm{f}_c^{(s)} \right)_J.
\end{cases}
\tag{5.126}
$$

In the equations above, it can be seen that there are global indices relative to the DoFs of the mesh (I, J), indices relative to the nodes of the mesh (A, B), and indices relative to the dimension of the Euclidean space (i, j). This non-standard use of indices occurs because this work made an attempt to generalise the contact forces to the case where the consistent mass matrix is also considered. This generalisation was attempted because of the motivations given in Chapter 3 regarding the violation of conservation laws when the lumped mass matrix is taken into account. While there is no way to generalise the contact velocity (5.120) using the consistent mass matrix (the inversion of the matrix $\mathbf{M}^{(s)} + \mathbf{M}^{(m)}$ would lead to contact velocities different from the velocities of the individual bodies everywhere on the mesh), this generalisation is instead feasible in the calculation of the contact forces (which is why they make use of the effective mass matrix $\tilde{\mathbf{M}}$ in Eq. (5.123)). The mixed use of indices is therefore due to the necessity of relating global indices (as those relative to the effective mass matrix) to nodal ones (as the ones used in the definition of contact velcity (5.120)).

These two variations of the same method (i.e., with a consistent mass matrix or lumped for calculating contact forces) will be critically assessed in Section 5.5.

Even though it is difficult to notice via the equations, this method violates the KKT conditions, making it inconsistent with the underlying thermodynamics. Owing to the lack

of definition of the gap function and the contact velocity defined in Eq. (5.74)*, it is not possible to plug this method into the normal and tangential compatibility equations. Nonetheless, it can be observed how the contact forces are computed proportionally to a constraint violation, i.e., Eq. (5.121). Due to this proportionality, it can be understood how the compatibility equation cannot be exactly satisfied.

5.5 Numerical simulations

This section includes numerical examples of the previously detailed formulations, distinguishing among those to which a node-to-node or a point-to-surface discretisation have been applied. In particular, since the implementations of node-to-node formulations are more straightforward, it was first considered in Section 5.5.1 to evaluate all of the different formulations to compute the contact forces reported in Section 5.4. Some details on the implementations for the node-to-node techniques are illustrated in Section 5.5.1.1. The numerical analyses were evaluated from a thermodynamic point of view (conservation of energy), from the point of view of the convergence (of the NR scheme) and, lastly, from the compliance with some analytical examples. Point-to-surface type formulation is considered in Section 5.5.2. Instead of implementing all the methods previously explained, it was considered to implement only the most numerically robust and physically consistent among those compared in Section 5.5.1, which appeared to the ALM. An algorithmic overview of the point-to-surface implementation for the ALM is given in Section 5.5.2.1. Observations and comments on all the above methods and their numerical results are finally given in Section 5.6.

The tolerance $tol = 10^{-11}$ is compared to the $(k+1)$-th NR error computed as follows

$$
\begin{aligned}
&\frac{|\mathbf{R}^{(k+1)}| \cdot |\delta\mathbf{x}^{(k+1)}|}{|\mathbf{R}^{(1)}| \cdot |\delta\mathbf{x}^1|} < tol \quad \vee \\
&\left((k+1) > 5 \wedge \left| \left[\left(\mathrm{r}^{(s)}\right)^{(k+1)}, \left(\mathrm{r}^{(m)}\right)^{(k+1)} \right]^T \right| \cdot \left| \left[\delta\mathbf{u}|_X^{(k+1)}, \delta\mathbf{u}|_Y^{(k+1)} \right] \right| < tol^{2/3} \right),
\end{aligned}
$$

(5.127)

with $|(\bullet)|$ being the absolute value of $(\bullet)$. While the criterion on the Left-Hand Side (LHS) is standard to compare how both the residual and the iterative variable decrease, the one on the Right-Hand Side (RHS) is mainly due to the stagnation discussed in Example 5.5.2.3.

---

*The contact velocity projected along the normal direction Eq. (5.121) acts as a constraint equation, but its definition and role are quite different from the gap function (5.73). The same rationale can be applied to the projection of such quantity along the tangential direction and its differences with the contact velocity defined in Eq. (5.74) for the node-to-node case.
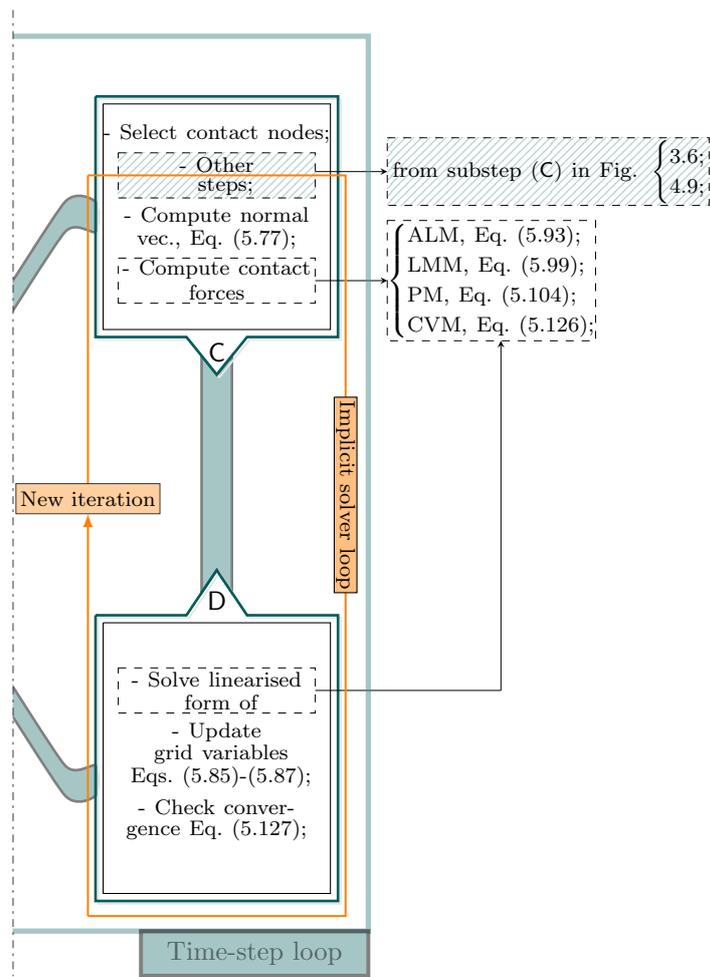
Figure 5.8: Main equations and relative use in the structure detailing the MPM algorithmic substeps. The considered equations are employed to design the node-to-node formulation combined with all the explained techniques to compute contact forces.

5.5.1 Node-to-node implementations

5.5.1.1 Implementation details

This section illustrates an overview of the node-to-node implementation integrated with the different techniques to compute contact forces, i.e., the ALM, LMM, PM and CVMs. From a technical point of view, the contact functional described by Eq. (5.50) can be coupled with a functional describing a general material, i.e., a solid or a porous one, for this work. This explains why Figure 5.8 focuses only on substeps (C) and (D), i.e., the ones dealing with the assemblage of the grid primary equations and their solution. Substeps (B) and (E) (as some parts of substep (C), in the box with diagonal line pattern) follow the ones provided by Figures 3.6 for a solid material and 4.9 for a porous one, and will not be treated below[*]. The first action in substep (C) determines the (plausible) contact nodes

---

[*]Nonetheless, the examples provided below consider only the coupling with a solid material. This was motivated by the intention to obtain preliminary results for the friction contact algorithm.

shared between the bodies under consideration, i.e., those active for both bodies. This selection is performed outside of the NR loop, and it is followed by the box mentioned above describing other steps relative to the selected material. This box is voluntarily positioned to partially overlap with the NR loop since the included steps may be in or out of this cycle. The node-to-node formulation considers the grid's normal directions on both bodies, i.e., Eqs. (5.75) and (5.76), to have an unbiased normal direction, Eq. (5.77). This vector allows the calculation of the contact forces, distinguishing between the normal (contact or separation) and the tangential directions (slip or stick). As the curly bracket indicates, these forces are computed differently according to the selected method.

The linearised form of the chosen contact forces is added to the linearised form of the equilibrium equations (and to other sets of primary equations, if these are present) to assemble the Jacobian matrix in the first iteration of substep (D). The inversion of this matrix permits the calculation of the primary unknowns. In the case of the ALM and LMM, this set of unknowns includes the LM too. The incremental solution of the unknowns permits the computation of the residuals, whose values, in turn, are checked using the criterion (5.127) to establish whether or not the NR solution has converged.

If the condition (5.127) is met, solutions are converged, and substep (E) follows.

5.5.1.2 Collision of elastic rectangular prisms
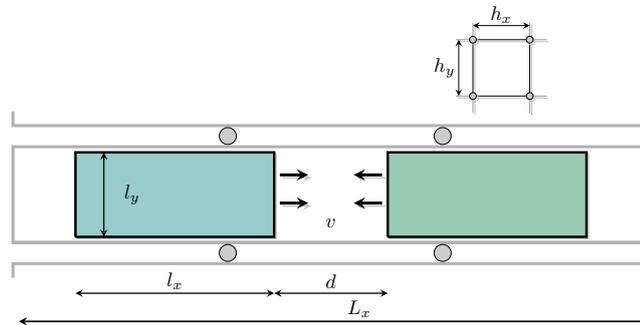
Figure 5.9: Illustration of the collision problem between elastic rectangular prisms.

Example scope This low-frequency elastic collision test is designed to assess the energetical behaviour of the different formulations discussed in the chapter, these being:

- the ALM;

- the LMM;

- the PM (with its linear and non-linear variations); and

- the CVM (used with the consistent mass matrix $\mathbf{M}$ and the lumped mass matrix $\bar{\mathbf{M}}$).

Table 5.1: Summary of the parameters considered in the analysis of the elastic prism collision, where the subscripts $(\bullet)_1$ and $(\bullet)_2$ denote the two bodies.

Parameter Settings		
Material Parameters	$\bar{E}^{(1)}$, $\bar{E}^{(2)}$	1000 Pa
	$\nu^{(1)}$, $\nu^{(2)}$	0.3
	$\rho_0^{(1)}$, $\rho_0^{(2)}$	50 kg/m³
Geometry, Velocities and Timings	L_x	10 m
	$l_x \times l_y$	2.5 m ×1 m
	d	2 m
	$\|v\|$	1 m/s
	T	2.5 s
Analysis Parameters	h_x, h_y	1 m
	mmp	16

† mmp is the number of MPs per direction per element.

Setup Two sections of rectangular prisms (see Figure 5.9) are initially separated by a distance d and with opposite velocity v applied at the beginning of the simulation. The bodies can only move horizontally since roller boundary conditions are prescribed both on the top and bottom of the domain. The stress-strain relationship for the considered materials is taken as the elastic CLC Hencky model (see Section 3.6.3 or Pretti *et al.* [37]). Most of the features concerning the analysis and the discretisations are listed in Table 5.1. Further features useful to reproduce the results are:

- the standard MPM is considered;

- the time-step length is five time more the approximated Courant-Friedrichs-Lewy (CFL) condition (see Eq. (3.127) in Section 3.6.2.2 or Pretti *et al.* [37]);

- the Ghost stabilisation method is employed, with $\gamma^K = 0$ and $\gamma^M = \rho_0^{(1)} \cdot 10^{-6}$ (see Eqs. (3.152) and (3.153));

- gravity effects are neglected; and

- the parameters used are

  - ALM: $r = 5\,(\bar{E}^{(1)} + \bar{E}^{(2)})$;

  - linear PM: $r = 5\,(\bar{E}^{(1)} + \bar{E}^{(2)})$; and

  - non-linear PM*: $r_n^{in} = 5\,(\bar{E}^{(1)} + \bar{E}^{(2)})$ with $c_0 = 0, c_1 = -0.01d$;
    $r_n^{fin} = 100\,r_n^{in\dagger}$ with $c_2 = -0.03d$ where $d \in \sqrt{h_x^2 + h_y^2}$.

---

*The tangent parameter is not stated as no tangential behaviour is expected for these analyses.

†It is worth mentioning that even the case where $r_n^{fin} < r_n^{in}$ was considered, in compliance to the explanation given in Section 5.4.4 about predictors and correctors strategy. However, this option is not included here as the method stagnates at the first time-step where contact is detected.
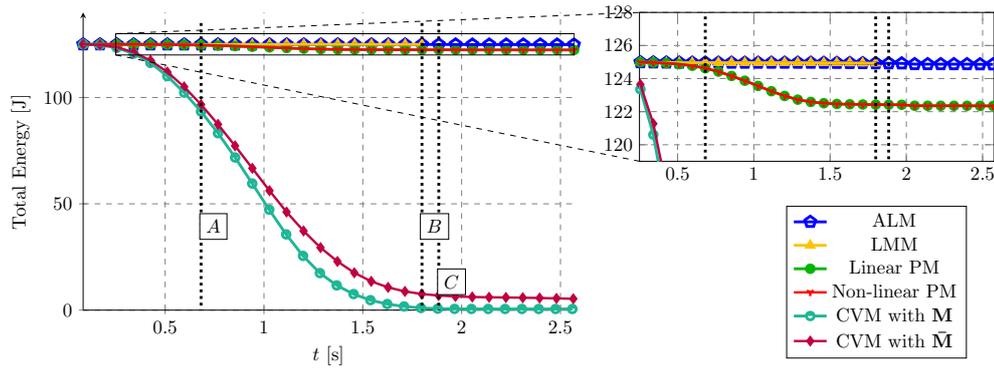
Figure 5.10: Time plot of energies obtained using different techniques.

While the elastic parameters and the velocities of the bodies are chosen to represent a low-frequency problem, the estimation of the penalty parameters is informed by the elastic moduli. However, this choice is ultimately based on some arbitrariness.

Results discussion - Energy comparison The choice of the constitutive model described in Section 3.6.3.1 and in Pretti *et al.* [37] allows the assessment of the validity of the different methods in terms of energy conservation when simulating the contact between bodies. As a matter of fact, the normal contact expected in this problem should not dissipate any energy since dissipation can only take place in the tangential direction according to the *Coulomb* friction law. Therefore, whenever it occurs, energy loss has to be seen as algorithmic dissipation.

Figure 5.10 shows the total energy trend (given by the sum of strain and kinetic energies) over time. Three particular time-steps are highlighted: A 9th step, B 21st step, and C 22nd step. As observed in the same figure, all the calculation methods complete the simulation in 30 time-steps, except for the LMM, which stops at the 21st step. As is well-known, the LMM struggles to cope with contact release, as also highlighted by Figure 5.11, representing the kinetic energies for the considered methods. Back to Figure 5.10, as far as energy evolution is concerned, it can be noticed that the CVM, independently of the mass matrix considered, dissipates substantial amounts of energy from the first contact moments. When the lumped matrix method is used a 95.73% of the total initial energy is observed, compared to a dissipation of 99.57% in the consistent mass matrix case. As mentioned in Section 5.4.4, the most dissipative method between the two variations is expected to be the one with the lumped mass matrix, as the numerical dissipation due to contact forces is added to the numerical dissipation due to the use of the lumped mass matrix. On the other hand, only the former dissipation is expected when the consistent mass matrix is used, i.e., the one relative to the contact algorithm. However, the nodal masses participating in the calculation of the contact velocity (and, in turn, of the contact forces) are bigger for the consistent mass matrix case. Thus, it appears that the sole dissipation for the consistent mass matrix overcomes the sums of the two mentioned above for the lumped mass matrix case. Owing to the considerable amount of energy dissipated by the extreme scenarios (lumped and
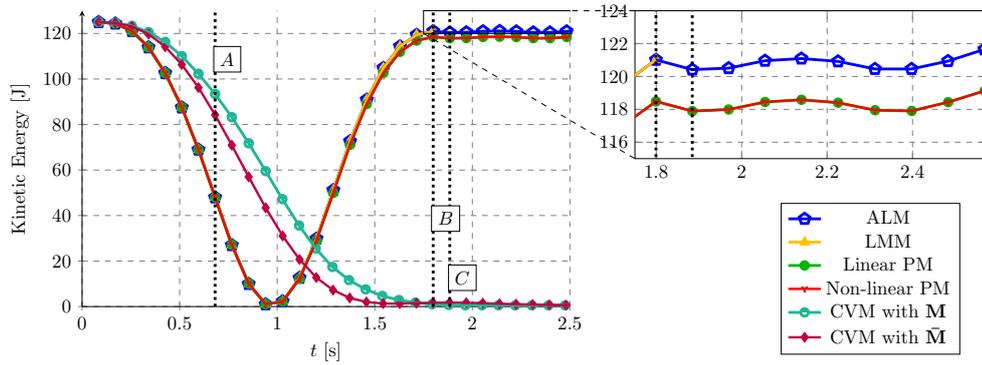
Figure 5.11: Time plot of kinetic energies obtained using different techniques.

consistent mass matrices), intermediate values for the effective mass matrix have not been studied.

To see the amount of dissipation for the other methods considered, it is necessary to observe the magnification on the right-hand side of Figure 5.10. As predicted, it can be seen that both the ALM and LMM do not dissipate energy. In contrast, the PM (both linear and non-linear) dissipates around 2.12% of the initial total energy. The reason why both these variations dissipate the same amount of energy lies in the choice of the penalty parameter. The parameter for the linear case has been chosen equal to the initial value for the non-linear case, i.e., $r = r_n^{in}$. As the converged values of the overlaps are small, these overlaps lie in the range where the slope is r_n^{in}, leading to the same results in terms of violation of the normal KKT conditions. As far as the ALM is concerned, although this method has a force part proportional to the overlapping bodies, it satisfies the KKT conditions as successive iterations proceed and the normal augmentation vanishes (see, for a discussion, Laursen and Chawla [35]).

However, it is legitimate to ponder why the results obtained here in terms of energy conservation are drastically different from those obtained by González-Acosta *et al.* [23] for a similar example using the CVM. As those authors underline, the energy dissipated also depends on the length of the load step (the longer the time-step, the greater the violation of the KKT conditions). Therefore, a shorter time-step length could have led to less marked dissipations for the CVM. Moreover, the method proposed by González-Acosta *et al.* [23] introduces a proximity rule (not included in this work) which activates contact only when the distance between the MPs from different bodies is lower than a chosen distance (typically less than the mesh size). This procedure reduces the interaction distance and, consequently, the time when the two bodies are considered to be in contact. Hence, if this proximity rule is applied, the time of the simulation when KKT conditions are violated is shorter, leading to a more energy-conservative algorithm compared to the one adopted in this work and reported in Figure 5.10. Although this arrangement is useful, it does not change the underlying strongly dissipative nature of the CVM.

Figure 5.11 compares the kinetic energies relative to the selected methods over time. Two interesting trends can be seen from this figure. First, the decrease in kinetic energy when
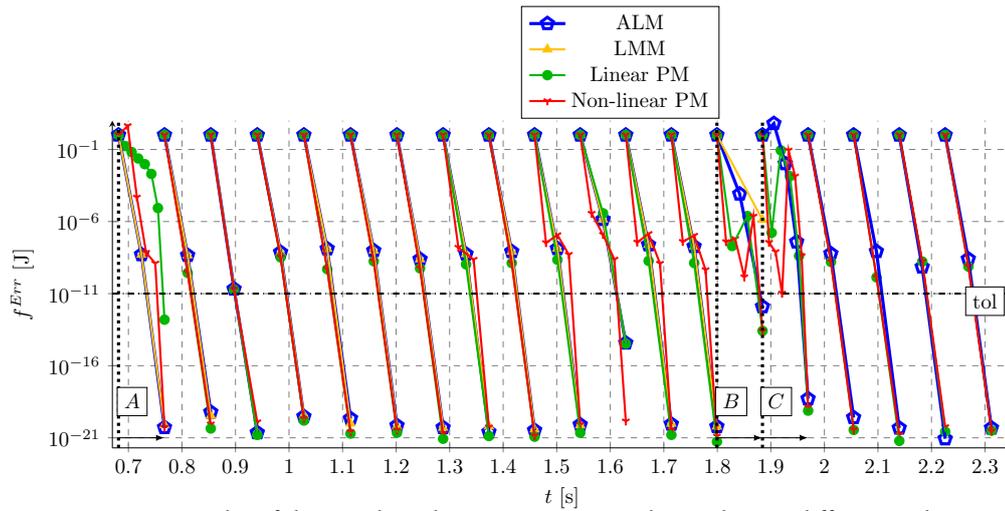
Figure 5.12: Time plot of the NR algorithmic convergences obtained using different techniques.

contact occurs in the CVMs presents gentler slopes than the steeper ones of the other methods. However, this softer decrease represents entirely numerical dissipation, as CVMs cannot recover these kinetic energies. A second pattern is highlighted if the kinetic energies in Figure 5.11 are compared with their total in Figure 5.10. The CVM using the consistent mass matrix exhibits higher values of the kinetic energies for most of the simulation compared to the CVM using the lumped mass matrix, thus showing a trend reversal from Figure 5.10. This difference implies that the CVM using the lumped mass matrix can convert kinetic energy more efficiently into strain energy during the initial phases of contact. Nonetheless, when times higher than ≈ 1.5 s are considered, low values of kinetic energies can be seen for both the CVMs.

Results discussion - Convergence comparison   In order to compare the NR algorithmic convergence of the primary equations among the different methods, time-steps from the 9th to the 26th (i.e., the most salient ones in terms of contact and subsequent release) are considered in Figure 5.12. In the same figure, it was decided to exclude the CVM graphs since the substantial energy dissipation made the convergence smooth. In addition, Figure 5.13(a) shows the energy evolution (calculated using the ALM). Figures 5.13(b)-5.13(g) present the deformed shapes of steps A, B, and C, with their respective strain and kinetic energies distributions. In step A, both the deformation and kinetic energy have an opposite gradient, with a smooth horizontal transition (see Figures 5.13(b) and 5.13(c)). This gradient is not as smooth for B and C when considering the strain energy (see Figures 5.13(d) and 5.13(f)). This behaviour is predictable since B marks the end of the contact between the bodies, leaving them to oscillate according to the propagation of elastic waves inside the bodies themselves. This separation and the consequent wave propagation are the cause of the difficult convergence that can be found for all methods for steps B and C.

In Figure 5.12, the ALM and PM are the approaches with smoother convergence. The only time steps that allow these techniques to be distinguished are B and C. In these steps, the ALM converges, albeit with a gentler slope in B, and requires 5 iterations in C compared
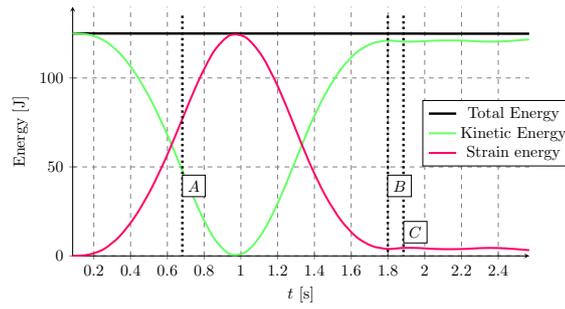
(a) Time plot of the different energies.
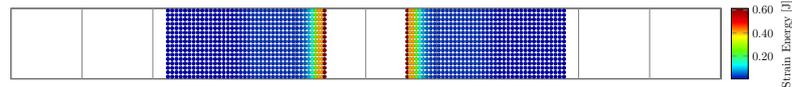
(b) Strain Energy, time-step A.
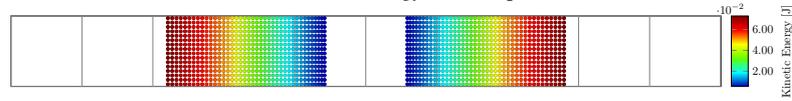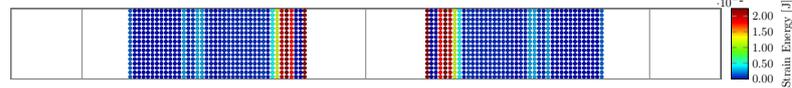
(c) Kinetic Energy, time-step A.

(d) Strain Energy, time-step B.
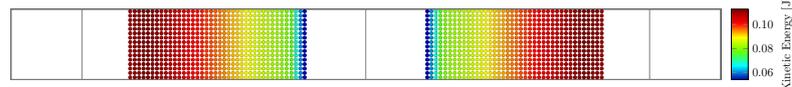
(e) Kinetic Energy, time-step B.
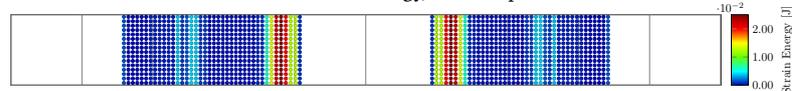
(f) Strain Energy, time-step C.
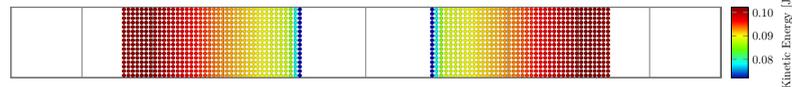
(g) Kinetic Energy, time-step C.

Figure 5.13: Time plot of the energies during the simulation (top row) and deflection shapes representing the strain and kinetic energies for time-steps A, B, and C. Plots are obtained using the ALM.

to the 3 iterations required in the rest of the simulation. On the other hand the LMM cannot converge in B, where the simulation ends due to an ill-conditioned Jacobian. These results are in agreement with the theory (see, e.g., Chen *et al.* [18]), which, as mentioned in Section 5.4.2, requires a strictly convex functional for the LM. When this is not the case, convergence is expected to be more difficult or even not achievable. It is also known that the ALM does not require the functional to be convex, and this allows for greater robustness in convergence. Both linear and non-linear PMs struggle converging at step A, using 8 and 6 iterations, respectively. Going forward, the non-linear PM presents some sharp behaviours from approximatively 1.3 s to time-step C. This response is illustrative of the additional non-linearity introduced by the method, which does not provide the hoped-for results in

terms of convergence[*]. In contrast, linear PM struggles only at time-steps B and C.

5.5.1.3 Cylinder rolling on an inclined surface
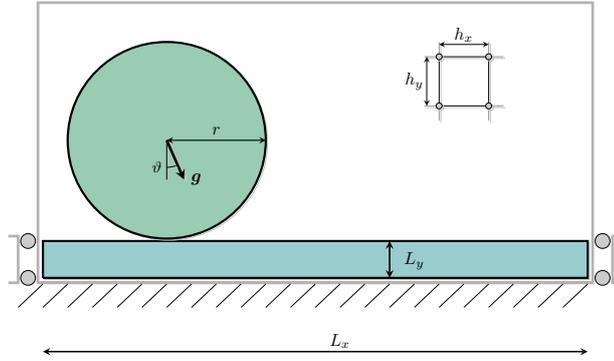
Figure 5.14: Illustration of the cylinder rolling over an inclined surface problem.

Example scope The scope of the example illustrated in Figure 5.14 consists in testing the displacements resulting from the node-to-node ALM with the analytical solution (obtained in the case of rigid bodies). Two scenarios are considered in the simulations since different behaviours are expected when varying the ratio between the tangent to the inclined plane, $\tan(\vartheta)$, and the frictional coefficient, μ. These are expressed by the following equations (see, for instance, Bardenhagen *et al.* [8])

$$u_x(t) = \begin{cases} \frac{1}{3}g\,t^2\sin(\vartheta) & \text{if } \dfrac{\tan(\vartheta)}{\mu} \leq 3 \quad \text{(rolling without slipping)}; \\ \frac{1}{2}g\,t^2\left(\sin(\vartheta) - \mu\,cos(\vartheta)\right) & \text{otherwise} \quad \text{(rolling and slipping)}. \end{cases} \tag{5.128}$$

Setup This example considers an elastic cylinder rolling by gravity on an inclined elastic surface, which is simulated by the application of an inclined gravity to a level surface, as illustrated in Figure 5.14. The cylinder is free to move, and its initial velocity and displacement values are null. The surface has rollers on its left and right edges, while it is fixed on the bottom. The first group of simulations present a value of $\mu = 0.4$, while in the second the value is changed to $\mu = 0.2$. Hence, given an inclination of the surface of $\pi/4$, it is expected that the cylinder will roll without slip in the first group of simulations, while it will roll and slip at the same time in the second ones, according to the analytical solution for the rigid bodies The other parameters of the simulations are outlined in Table 5.2 and the features are:

- standard MPM is considered;

---

[*]From these analyses, it appears that the non-linear PM does not give the results for which it was conceived, especially a smoother NR convergence. Possibly, this is due to the introduction of a further non-linearity in the system of equations, which, instead of smoothening the algorithm, introduces an additional degree of complication. Besides, this method was intended as a simplification of the work of Zavarise *et al.* [36], who proposed a much more sophisticated algorithm.

Table 5.2: Summary of the parameters considered in the analysis of the elastic cylinder, $(\bullet)^{(c)}$, rolling over an elastic inclined surface, $(\bullet)^{(s)}$.

Parameter Settings		
Material Parameters and Gravity	$\bar{E}^{(c)}$, $\bar{E}^{(s)}$	10^7 Pa
	$\nu^{(c)}$, $\nu^{(s)}$	0.3
	$\rho_0^{(c)}$, $\rho_0^{(s)}$	10^3, 10^4 kg/m^3
	$g^\dagger$	0.141 m/s^2
Geometry and Timings	$L_x \times L_y$	10 m $\times$0.2 m
	r	2 m
	ϑ	$\frac{\pi}{4}$
	T	2.5 s
Analyses Parameters	mmp	2

† The gravity effects have been applied only on the cylinder.

Table 5.3: Relative errors of the centre of cylinder displacements with respect to the analytical solution given by Eq. (5.128) for different mesh resolutions.

Case	time [s]	40×40	50×50	60×60	70×70	80×80
Rolling without	1.25	$\approx 45.03\%$	$\approx 33.33\%$	$\approx 22.66\%$	$\approx 38.92\%$	$\approx 31.95\%$
slipping	2.5	$\approx 49.84\%$	$\approx 38.39\%$	$\approx 26.32\%$	$\approx 43.35\%$	$\approx 36.14\%$
Rolling &	1.25	$\approx 26.38\%$	$\approx 28.55\%$	$\approx 26.96\%$	$\approx 36.47\%$	$\approx 28.13\%$
slipping	2.5	$\approx 30.99\%$	$\approx 32.67\%$	$\approx 33.70\%$	$\approx 37.92\%$	$\approx 26.89\%$

- the time-step length is five time more the approximated CFL condition, and an elastic constitutive model is considered (for details of both, see Section 3.6.3.1 or Pretti *et al.* [37]);

- the Ghost stabilisation method is employed, with $\gamma^K = 0$ and $\gamma^M = \rho_0^{(c)} \cdot 10^{-6}$ (see Eqs. (3.152) and (3.153)); and

- ALM is used with the penalty parameters being $r = 5\,(\bar{E}^{(c)} + \bar{E}^{(s)})$.

While elastic parameters were selected to approximate a rigid body without incurring technical difficulties, penalty parameters are (as in the previous example) approximatively based on the elastic moduli values.

As the mesh dependency of this implementation has been highlighted in the literature by several authors (see Bardenhagen *et al.* [8] and Chandra *et al.* [28]), five different meshes have been considered. Moreover, it must be emphasised that the analytical solution is given for undeformable bodies, while some compliance has been considered for the two bodies here. Hence, some (minor) deviation from the analytical solution must be expected.
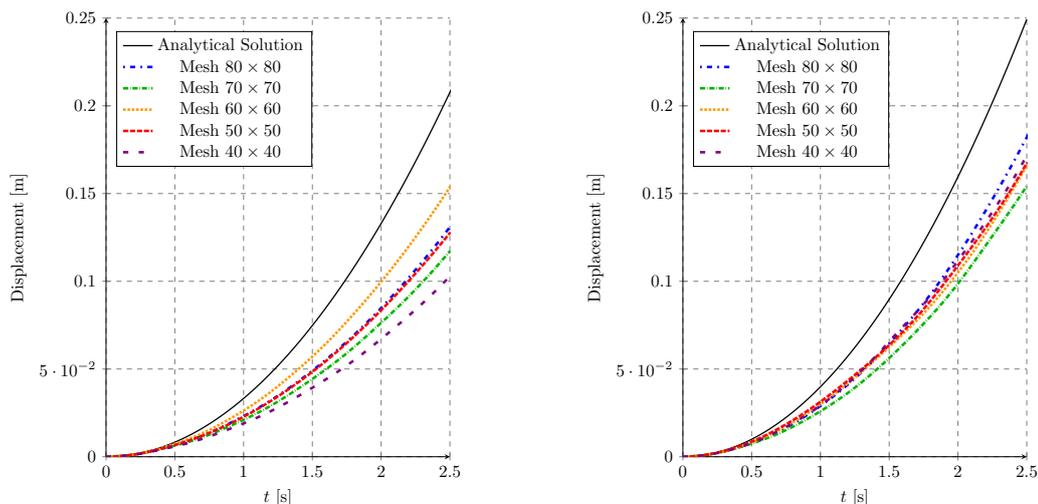
(a) Centre of the cylinder displacement over time for the rolling without slipping case.

(b) Centre of the cylinder displacement over time for the rolling and slipping case.

Figure 5.15: Comparisons of the nodal representation results between different contact conditions for different mesh resolutions. Analytical solution is given by Eq. (5.128).

Results discussion   Time trends of the centre of cylinder displacements are reported in Figure 5.15 for both rolling without slipping and rolling with slipping (see Figures 5.15(a) and 5.15(b), respectively). Moreover, to quantify the difference between the numerical results and the analytical solution, relative errors for all of the meshes computed at half way through the simulations (1.25 s) and at the end (2.5 s) are given in Table 5.3. Figure 5.15 shows that widespread results are obtained for the case of rolling without slipping (Figure 5.15(a)). In this case, particularly unexpected is the worsening of the relative errors with finer meshes for the former case, as happens in the case of the 70×70 and 80×80 meshes, with relative errors (for 2.5 s) of $\approx 43.35\%$ and $\approx 36.14\%$, respectively. A plausible explanation can be found when comparing the results of the current work with those from Chandra *et al.* [28] for a similar example. To achieve better results for the case of the rolling without slipping case, these authors considered a mesh 6 times finer than the one considered here. Since the computational effort available to the author of this work is not able to cope with such accuracy, a different approach has been used to investigate the spread of results for the rolling without slipping case. In particular, the goal of this further investigation is to understand if the initial set-up of the MPs with respect to the grid position plays a role. This would indeed explain why coarser meshes can sometimes produce results closer to the analytical solution. The results of this investigation are reported in Figure 5.16, where different initial positioning (quantified by the shift with respect to the mesh size) has been tried for the 40×40 grid with shorter simulations ($T = 0.5$ s). In particular, Figure 5.16(a) reports the displacement of the centre of the cylinder for the 10 case studies of shift at time $T = 0.5$ s. Figure 5.16(a) shows that no symmetry with respect to the normalised shift 0.5 is observed, as the highest and lowest values of displacements are given by 0.5 and 0.7 normalised shifts, respectively. Nonetheless, it should be considered that the relative error difference in displacements between these simulations is not particularly considerable, i.e.,
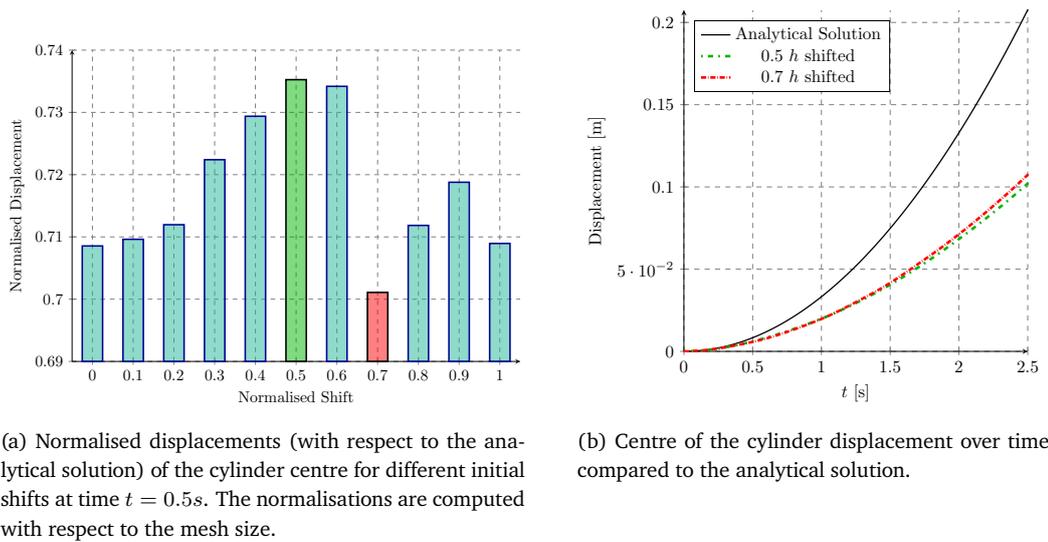
(a) Normalised displacements (with respect to the analytical solution) of the cylinder centre for different initial shifts at time $t = 0.5s$. The normalisations are computed with respect to the mesh size.

(b) Centre of the cylinder displacement over time compared to the analytical solution.

Figure 5.16: Graphs refer to the roll without slip case for a mesh with 40×40 elements.

slightly more than $\approx 3\%$. These two simulations (being the maximum and minimum values of displacements at $t = 0.5$ s) are then allowed to run for the whole analysis time (2.5 s), thus checking if these differences increase over time. As reported in Figure 5.16(b), displacements at the end of the simulation are inverted compared to 0.5 s, with the 0.7 shift displacement being larger than the 0.5. The relative errors of these two simulations compared to the analytical solution are $\approx 50.68\%$ for the 0.5 shift and $\approx 48.12\%$ for the 0.7 one. It is noteworthy how these errors are considerable, even though they are close to each other. Due to these results, the dependency of the initial set-up of MPs with respect to the mesh position can be considered as a secondary effect, even though not entirely negligible in explaining the results of Figure 5.15(a) and Table 5.3. Hence, in the absence of further explanations (also lacking in Chawla *et al.* [28]), it is concluded that the mesh lengths of the considered simulations need to be finer to capture analytical behaviour in the case of the rolling without slipping.

In the case of rolling and slipping, results are more clustered together, and closer to the analytical solution, with the best- and worst-case scenario represented by 80×80 and 70×70 grid, with relative errors of $\approx 26.89\%$ and $\approx 37.92\%$, respectively. The fact that these results are close to the analytical solution (compared to the rolling without slipping case) for coarser meshes is also reported in Chandra *et al.* [28]. Nonetheless, it is noteworthy that the mesh lengths are not a proxy for closeness to the analytical solution, as the errors are not sorted according to the mesh scale. Again, a plausible explanation of this phenomenon could be found in the set-ups of the different simulations, as proved for the rolling without slipping case.
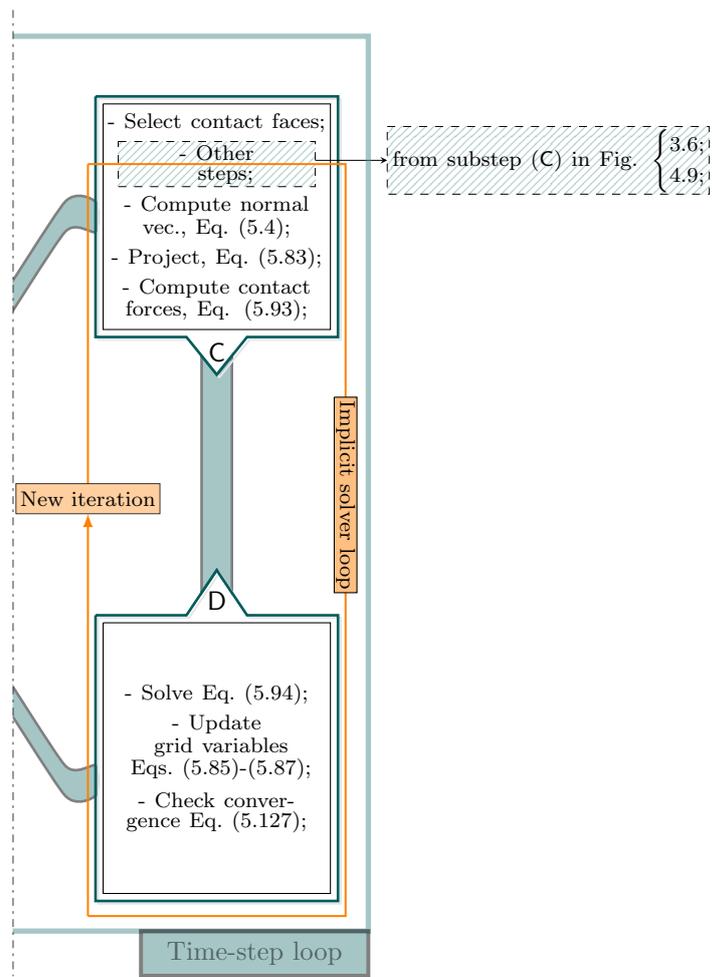
Figure 5.17: Main equations and relative use in the structure detailing the MPM algorithmic substeps. The considered equations are employed to design the point-to-surface formulation using the ALM to compute contact forces.

5.5.2 Point-to-surface implementation

5.5.2.1 Implementation details

Similarly to Section 5.5.1.1 for the node-to-node method, the point-to-surface formulation combined with the ALM to compute the contact forces is detailed below. The reason for selecting only the ALM lies in the demonstrated robustness in Example 5.5.1.2 over the other methods. Figure 5.17 focuses on substeps (C) and (D) for the same reasons of generality explained in Section 5.5.1.1.

The selection of the contact faces (and relative nodes) is employed in substep (C) to avoid tracking the boundaries*. The faces on the secondary surface are then seeded by Gauss Points (as explained in Section 5.3.1.2 and illustrated in Figure 5.5) for integration purposes. Other actions between in and out of the NR loop from substep (C) in Figures 3.6 or 4.9

---

*If a technique to reconstruct the boundaries were to be operated (see, for instance, Bing *et al.* [26]), proximity detection between bodies should be performed at this stage.

follow. The biased normal vector is computed inside the NR cycle. The first computation of this quantity at every Gauss Point is performed using the face direction, while its update follows the normalised Nanson's formula (5.4). Ray-tracing projection (5.83) is performed for every Gauss Point, which allows the computation of the gap function (5.6), the current position of the projection on the main surface (5.9), and the contact velocity (5.25). Based (also) on these quantities, contact forces are assembled according to the ALM described by Eq. (5.93).

The solution to Eq. (5.94) is given by the inversion of the Jacobian matrix (linearisation of the ALM method for the point-to-surface formulation is given in Appendix B.3) and it provides updated incremental values at the grid nodes in substep (D). The considered time-discretisation (Eqs. (5.85)-(5.87), if the mid-point rule is considered) informs the calculation of the vector and matrices appearing in the primary grid equations. The residual and the incremental solution are then tested against criterion (5.127). Failing this condition implies the assembly of a new Jacobian, looping again between substeps (C) and (D).

On the other hand, the satisfaction of criterion (5.127) defines the grid converged values, and substep (E) follows to map these solutions to the MPs.

5.5.2.2 Collision of elastic rectangular prisms

Example scope   The same problem considered in Section 5.5.1.2 (and illustrated in Figure 5.9) is here analysed to assess the point-to-surface method (in lieu of the node-to-node technique) with the ALM. The selection of the ALM is justified by the results given in Section 5.5.1.2, where it seems that this technique provided the most robust results.

Setup   To compare the results obtained for the node-to-node implementation, this simulation is run with the same parameters given in Section 5.5.1.2 for the ALM and in Table 5.1, with the additional use of 4 Gauss points per element face to integrate the contact surface. Despite 4 Gauss points per element face can be seen as a high number of Gauss points to integrate values described by linear polynomial functions, this number is still considerably lower than the mmp, i.e., the number of MPs per direction per elements, which is used to integrate over the mesh volume. Thus, 4 Gauss points per element face can be regarded as a compromise between these two opposite considerations. Even though it is not particularly significant for this example, the secondary body is the one on the left-hand side of Figure 5.9.

Results discussion   As already anticipated in 5.2, biased formulations lead to numerical dissipation, due to the violation of Newton's 3rd action-reaction principle, both in terms of linear and angular momentum and, consequently, of energy conservation. This is confirmed by the current analysis, as it can be observed in Figure 5.18(a), where the total dissipation at the end of the simulation is $\approx 3.93\%$.

However, as it can be appreciated from Figure 5.18(b), convergence of the NR algorithm is particularly smooth. No issues are encountered even at difficult time-steps, which makes the

(a) Time plot of energies.

(b) Time plot of the NR algorithmic convergence.

Figure 5.18: Graphs refer to the point-to-surface ALM.

current method stabler than the node-to-node technique (see Figure 5.12 for comparison). In particular, the current method works better in proximity of the release case, where the discontinuity in contact does not affect the algorithmic convergence. In the author's opinion, this smoother convergence is mostly due to the point-to-surface formulation more than the slight dissipation of energy occurring for biased formulations. Indeed, the dissipation is not too different ($\approx 96.07\%$ of the total energy is retained) from the value reported in Example 5.5.1.2 for the PM ($\approx 97,88\%$), whose convergence is nonetheless more difficult (see Figure 5.12, particularly time-steps B and C).

5.5.2.3 Frictionless sliding prism over inclined surface
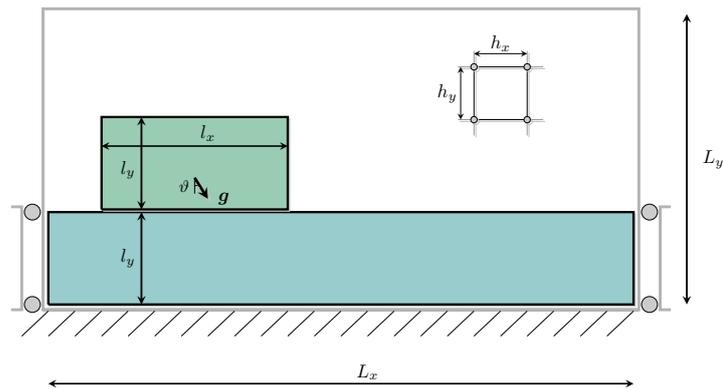
Figure 5.19: Illustration of the frictionless prism sliding over an inclined surface problem.

Example scope Despite retaining all of the complexities necessary to describe a large deformation contact formulation, the considered point-to-surface method still relies on the (Cartesian) mesh, as assumed for this methodology in Section 5.5.2. This is the reason why the example presented here, represented in Figure 5.19, takes a prism (in lieu of a

Table 5.4: Summary of the parameters considered in the analysis of the elastic prism, $(\bullet)^{(p)}$, sliding over an elastic inclined surface, $(\bullet)^{(s)}$.

Parameter Settings		
Material Parameters & Gravity	$\bar{E}^{(p)}, \bar{E}^{(s)}$	$5 \cdot 10^5$ Pa, 10^5 Pa
	$\nu^{(p)}, \nu^{(s)}$	0.35
	$\rho_0^{(p)}, \rho_0^{(s)}$	2000 kg/m^3
	$g^\dagger$	10 m/s^2
Geometry and Timings	$L_x \times L_y$	6×6 m
	l_x, l_y	3 m, 1.5 m
	ϑ	$\frac{15}{180}\pi$
	T	1 s
Analysis Parameters	mmp	3
	$fGp^\ddagger$	3
	h_x, h_y	0.5 m
	c	$100/h \left(E^{(p)} + E^{(s)} \right)$

† The gravity effects have been applied only on the sliding prism.
‡ In this case, fGp are the number of Gauss points per face per dimension.

cylinder, as in Section 5.5.1.3) sliding over an inclined surface into account. In this fashion, the representation of the contact surfaces via the meshes should provide enough accuracy (even in the presence of coarse meshes) to describe the real surface of the bodies coming into contact. The results of the analysis can be then compared with the analytical solution for the rigid body cases, this being (see, for instance, Jiang and Yang [38])

$$u_x = \frac{1}{2} \left(g \sin \vartheta \right).$$

Setup   For this analysis, the inclined surface is fixed on the bottom (dimensions $L_x \times l_y$). The dimensions of the prism rolling over this surface are $l_x \times l_y$. These and other arbitrarily selected parameters are listed in Table 5.4. Other features necessary to reproduce these data are:

- standard MPM is considered;

- the time-step length is five time more the approximated CFL condition, and elastic constitutive model is considered (for details of both, see Section 3.6.3.1 or Pretti *et al.* [37]); and

- the Ghost stabilisation method is employed, with $\gamma^K = 0$ and $\gamma^M = \rho_0^{(p)} \cdot 10^{-6}$ (see Eqs. (3.152) and (3.153)); and

- ALM is used with the penalty parameters being $r = 5 \left(\bar{E}^{(p)} + \bar{E}^{(s)} \right)$.
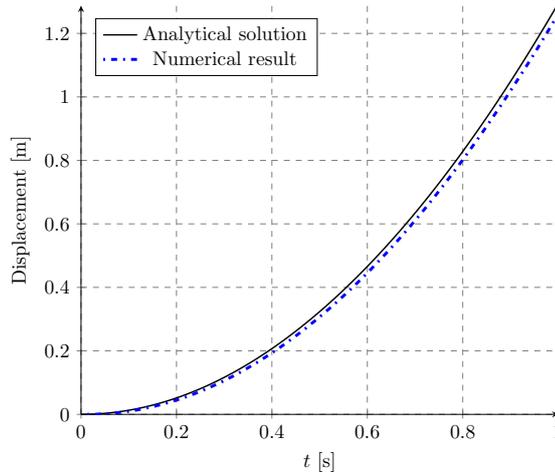
Figure 5.20: Centre of the prism displacement over time compared to the analytical solution.

Results discussion   Figure 5.20 shows that the current algorithm can reproduce the analytical solution with good accuracy. The relative error of the horizontal displacement computed at the end of the simulation ($t = 1$ s) is $\approx 2.53\%$. The underestimation of the displacement field for the numerical example could be due to the loss of the energy for biased formulations, as reported for the Example 5.5.2.2. On the other hand, the smoothness of the NR convergence, could possibly be improved with a more continuous representation of the contact surfaces. While slipping, new contact mesh faces (and their Gauss points) are activated. In this progressive movement, the analysis may undergo several scenarios, the extremes of which are represented in Figures 5.21(a) and 5.21(b). While the case in Figure 5.21(a) is not problematic, as the mesh fits the actual body boundaries, this is not true for case in Figure 5.21(b). This situation introduces strong discontinuities in the analysis (contact nodes are either on/off, regardless of how MPs partially fill the element associated with them), leading to a stagnation in the correspondence of the newly activated nodes. The contact constraint equations at these nodes cannot be satisfied with the necessary precision, due to a discrepancy between the physical progressive phenomenon and the discontinuous representation relying only on the mesh. Nonetheless, it is observed that the convergence of the NR for the equilibrium equations is achieved with no issues even for these nodes. As predictable, the above-described mismatch would be even more substantial for a simulation such as that in Section 5.5.1.3, where the difference between the meshes and the real body surfaces is more emphasised.

5.6   Observations

Some remarks and conclusions can be drawn from the examples given in Section 5.5 and the discussion carried on throughout this chapter.

The introduction of a unique novel framework for dealing with different ways of enforcing contact conditions (node-to-node and point-to surfaces) and computing contact forces (via
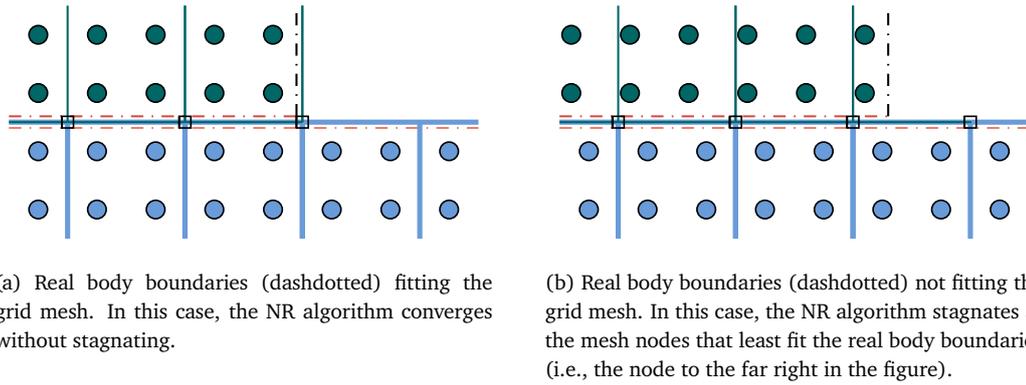
(a) Real body boundaries (dashdotted) fitting the grid mesh. In this case, the NR algorithm converges without stagnating.

(b) Real body boundaries (dashdotted) not fitting the grid mesh. In this case, the NR algorithm stagnates at the mesh nodes that least fit the real body boundaries (i.e., the node to the far right in the figure).

Figure 5.21: While sliding over the inclined surface, the top prism can experience both scenarios represented in the above figures. In the case of Figure (*b*), one can appreciate how the contact nodes are activated even when the elements associated with them are partially filled with MPs.

ALM, LMM, linear and non-linear PM and CVMs) has allowed a straightforward comparison among these options. Exploring the theoretical setting in a unified manner (and how these techniques are located within it) has also facilitated an understanding of the sources for numerical dissipation.

It was shown how the CVM and the PM (and each of their variations) violate the KKT conditions. In particular, it was found that, under the same conditions, the CVM clearly violates these contact conditions by the degree of artificial energy dissipation, while dissipation caused by the PM is much more superficial. Hence, without any particular treatment of this issue (as suggested in González-Acosta *et al.* [23]), the CVM is not feasible for the time-step lengths used in implicit discretisations. On the other hand, the compliance of the LMM and ALM with the KKT conditions comes with further additional costs, i.e., the additive equations necessary to compute the tractions on the contact surfaces. A way to avoid these two problems (i.e., KKT conditions and additional computational costs) could be using Nitsche's method (see, for instance, Mlika *et al.* [32]), whose implementation, nonetheless, is not trivial, especially in a dynamic energy-consistent formulation.

As proved by Example 5.5.1.2, the robustness of these algorithms can be problematic even for very simple benchmarks. The assessment of the CVM from this perspective was neglected because of its extremely damped results. However, it was noticed how the linear PM converged in the whole analysis, even though it displayed some non-smooth behaviour at difficult time-steps. The non-linear PM performed in a poorer manner compared to its linear counterpart, so it can be concluded that the results for the non-linear PM cannot be regarded as particularly promising. While the LMM failed at the contact-release step, the ALM did not, proving that the augmentation plays a primary role in the strength of the algorithm.

From a comparison of Section 5.5.1 with Section 5.5.2, several differences can be highlighted between the node-to-node and the point-to-surface implementations. While the kinematics of the former technique is easier to understand and implement, the latter algorithm, even though still relying on mesh representation in this work, can provide a solid

basis for contact in the context of large deformations. The reason is that there are no simplifications in the kinematics involved. Hence, the proposed point-to-surface approach can be regarded as a first step towards a solid representation of the frictional contact conditions within the MPM. Some techniques for capitalising on this point-to-surface algorithm could be based on B-splines (as proposed in Bing *et al.* [26], for instance) to render the boundaries, or on local enrichment of the element (as suggested in Liu and Borja [39] for the case of eXtended-FEM), or even on local mesh refinement techniques (see, for instance, Mao *et al.* [40]). Another issue which should be explored is why the point-to-surface algorithm performs well in the frictionless case (as in Example 5.5.2.3), while it becomes unstable when friction is considered. As a matter of fact, all the proofs of stability for the ALM (see El-Abbasi and Bathe [41] from a numerical perspective and Bathe and Brezzi [42] from an analytical one) consider only contact, without taking friction into account. Even though this difference could seem negligible in terms of stability, it should be considered that the Jacobian matrix changes between frictionless and frictional cases, especially for the sticking case (see Appendix B.3). Under this condition, the derivative of the constraint with respect to the LM is zero, making the Jacobian matrix similar to a saddle-point problem, which has the potential to cause numerical instability. In this regard, an adaptation of the Polynomial Pressure Projection (PPP) from Liu and Borja [43] to the frictional case was tried. However, this attempt was not successful in improving the convergence of the NR. Ghost stabilisation approaches (suggested in Burman and Hansbo [44] for constant interpolant function representing the space of the LM and adapted from Sticko *et al.* [45] for higher order spaces) cannot work in a biased formulation, as they need to stabilise the tractions (i.e., the LM) computed on two different bodies. Another option could be represented by the stabilised LM described by Heintz and Hansbo [46] and adapted to large deformation mechanics by Navarro-Jiménez *et al.* [31]. Here the LM space is stabilised by the addition of a Nitsche-like traction term. However, this approach was not considered. As previoulsy mentioned, the linearisation of a Nitsche's term is not straightforward (as it can be seen in Mlika *et al.* [32]) and its adaptation to the used energy conserving algorithm requires more investigation. Non-monolithic solvers for mixed problems, such as the *Uzawa algorithm* (see, for instance, Arrow *et al.* [47] or Bertsekas [48]) or the stabilised fixed-point iteration presented in Tur *et al.* [49], were not considered in this work as they are not second-order accurate with regard to convergence of the iterative method.

It is worth mentioning that there are other promising approaches in the computer graphics community, including that proposed by Hu *et al.* [50]. This approach is innovative, not in terms of how contact forces are calculated (it is indeed a PM), but because it does not require a boundary representation. Coupling a rigid body with a deformable one, this method does not require the surface of the deformable body to be represented, thus relying on the rigid body to represent the contact surface. It thus differs not only from the approaches considered in this work, which must be based on the mesh (nodes or faces), but also from the ones where parametrisation of the surfaces is involved (such as Liu and Sun [27]).

Chapter summary

This chapter has shown the continuum principles necessary to describe frictional contact in the context of finite strain mechanics. The kinematics, the statics, and the power dissipated by two surfaces coming into contact have been detailed. Tribological laws for non-adhesive frictional contact have been detailed. Different functionals capable of including this interaction between two surfaces and two kinds of plausible discretisations were defined in a unique framework to favour a comparison among the different methods. Several numerical examples have supported a thorough discussion on the sources of numerical dissipation and the capability of the different discretisations to reproduce analytical results.

References

[1]   L. De Lorenzis, P. Wriggers and C. Weißenfels, 'Computational contact mechanics with the finite element method,' *Encyclopedia of Computational Mechanics second edition*, pp. 1–45, 2017.

[2]   T. A. Laursen, *Computational contact and impact mechanics: fundamentals of modeling interfacial phenomena in nonlinear finite element analysis*. Springer Science & Business Media, 2003.

[3]   P. Wriggers and T. A. Laursen, *Computational contact mechanics*. Springer, 2006, vol. 2.

[4]   S. Andrews, K. Erleben and Z. Ferguson, 'Contact and friction simulation for computer graphics,' in *ACM SIGGRAPH 2022 Courses*, 2022, pp. 1–172.

[5]   A. Curnier, Q.-C. He and A. Klarbring, 'Continuum mechanics modelling of large deformation contact with friction,' in *Contact Mechanics*, Springer, 1995, pp. 145–158.

[6]   G. Pietrzak and A. Curnier, 'Large deformation frictional contact mechanics: Continuum formulation and augmented Lagrangian treatment,' *Computer Methods in Applied Mechanics and Engineering*, vol. 177, no. 3-4, pp. 351–381, 1999.

[7]   K. Poulios and Y. Renard, 'An unconstrained integral approximation of large sliding frictional contact between deformable solids,' *Computers & Structures*, vol. 153, pp. 75–90, 2015.

[8]   S. Bardenhagen, J. Brackbill and D. Sulsky, 'The material-point method for granular materials,' *Computer Methods in Applied Mechanics and Engineering*, vol. 187, no. 3-4, pp. 529–541, 2000.

[9]   S. Bardenhagen, 'An improved contact algorithm for the material point method and application to stress propagation in granular material,' *Computer Modeling in Engineering & Sciences*, vol. 2, no. 4, p. 509, 2001.

[10] W. Hu and Z. Chen, 'A multi-mesh MPM for simulating the meshing process of spur gears,' *Computers & Structures*, vol. 81, no. 20, pp. 1991–2002, 2003.

[11] X.-F. Pan, A.-G. Xu, G.-C. Zhang, P. Zhang, J.-S. Zhu, S. Ma and X. Zhang, 'Three-dimensional multi-mesh material point method for solving collision problems,' *Communications in Theoretical Physics*, vol. 49, no. 5, p. 1129, 2008.

[12] V. Lemiale, J. Nairn and A. Hurmane, 'Material point method simulation of equal channel angular pressing involving large plastic strain and contact through sharp corners,' *Computer Modeling in Engineering and Sciences*, vol. 70, no. 1, p. 41, 2010.

[13] Z. Ma, X. Zhang and P. Huang, 'An object-oriented MPM framework for simulation of large deformation and contact of numerous grains,' *Computer Modeling in Engineering and Sciences*, vol. 55, no. 1, p. 61, 2010.

[14] P. Huang, X. Zhang, S. Ma and X. Huang, 'Contact algorithms for the material point method in impact and penetration simulation,' *International Journal for Numerical Methods in Engineering*, vol. 85, no. 4, pp. 498–517, 2011.

[15] J. Nairn, 'Modeling imperfect interfaces in the material point method using multimaterial methods,' *Comput. Model. Eng. Sci*, vol. 1, no. 1, pp. 1–15, 2013.

[16] J. Ma, D. Wang and M. Randolph, 'A new contact algorithm in the material point method for geotechnical simulations,' *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 38, no. 11, pp. 1197–1210, 2014.

[17] S. Nezamabadi, F. Radjai, J. Averseng and J.-Y. Delenne, 'Implicit frictional-contact model for soft particle systems,' *Journal of the Mechanics and Physics of Solids*, vol. 83, pp. 72–87, 2015.

[18] Z.-P. Chen, X. Zhang, X.-M. Qiu and Y. Liu, 'A frictional contact algorithm for implicit material point method,' *Computer Methods in Applied Mechanics and Engineering*, vol. 321, pp. 124–144, 2017.

[19] M. A. Homel and E. B. Herbold, 'Field-gradient partitioning for fracture and frictional contact in the material point method,' *International Journal for Numerical Methods in Engineering*, vol. 109, no. 7, pp. 1013–1044, 2017.

[20] L. G. Acosta, I. Pantev, P. Vardon and M. Hicks, 'The use of MPM to estimate the behaviour of rigid structures during landslides,' in *Numerical Methods in Geotechnical Engineering IX*, CRC Press, 2018, pp. 651–656.

[21] C. Liu, Q. Sun and G. G. Zhou, 'Coupling of material point method and discrete element method for granular flows impacting simulations,' *International Journal for Numerical Methods in Engineering*, vol. 115, no. 2, pp. 172–188, 2018.

[22] X. Han, T. F. Gast, Q. Guo, S. Wang, C. Jiang and J. Teran, 'A hybrid material point method for frictional contact with diverse materials,' *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, vol. 2, no. 2, pp. 1–24, 2019.

[23] J. L. G. Acosta, P. J. Vardon and M. A. Hicks, 'Development of an implicit contact technique for the material point method,' *Computers and Geotechnics*, vol. 130, p. 103 859, 2021.

[24] L. Gao, N. Guo, Z. Yang and R. Jardine, 'MPM modeling of pile installation in sand: Contact improvement and quantitative analysis,' *Computers and Geotechnics*, vol. 151, p. 104 943, 2022.

[25] V. P. Nguyen, A. de Vaucorbeil and S. Bordas, *The Material Point Method: Theory, Implementations and Applications*. Springer Nature, 2023.

[26] Y. Bing, M. Cortis, T. Charlton, W. Coombs and C. Augarde, 'B-spline based boundary conditions in the material point method,' *Computers & Structures*, vol. 212, pp. 257–274, 2019.

[27] C. Liu and W. Sun, 'ILS-MPM: An implicit level-set-based material point method for frictional particulate contact mechanics of deformable particles,' *Computer Methods in Applied Mechanics and Engineering*, vol. 369, p. 113 168, 2020.

[28] B. Chandra, V. Singer, T. Teschemacher, R. Wüchner and A. Larese, 'Nonconforming Dirichlet boundary conditions in implicit material point method by means of penalty augmentation,' *Acta Geotechnica*, vol. 16, no. 8, pp. 2315–2335, 2021.

[29] K. Nakamura, S. Matsumura and T. Mizutani, 'Particle-to-surface frictional contact algorithm for material point method using weighted least squares,' *Computers and Geotechnics*, vol. 134, p. 104 069, 2021.

[30] F. Frenet, 'Sur les courbes à double courbure,' *Journal de mathématiques pures et appliquées*, vol. 17, pp. 437–447, 1852.

[31] J. M. Navarro-Jiménez, M. Tur, J. Albelda and J. J. Ródenas, 'Large deformation frictional contact analysis with immersed boundary method,' *Computational Mechanics*, vol. 62, no. 4, pp. 853–870, 2018.

[32] R. Mlika, Y. Renard and F. Chouly, 'An unbiased Nitsche's formulation of large deformation frictional contact and self-contact,' *Computer Methods in Applied Mechanics and Engineering*, vol. 325, pp. 265–288, 2017.

[33] O. Ding and C. Schroeder, 'Penalty force for coupling materials with coulomb friction,' *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 7, pp. 2443–2455, 2019.

[34] E. Burman and P. Hansbo, 'Deriving robust unfitted finite element methods from augmented Lagrangian formulations,' in *Geometrically unfitted finite element methods and applications*, Springer, 2017, pp. 1–24.

[35] T. Laursen and V. Chawla, 'Design of energy conserving algorithms for frictionless dynamic contact problems,' *International Journal for Numerical Methods in Engineering*, vol. 40, no. 5, pp. 863–886, 1997.

[36] G. Zavarise, L. De Lorenzis and R. L. Taylor, 'A non-consistent start-up procedure for contact problems with large load-steps,' *Computer Methods in Applied Mechanics and Engineering*, vol. 205, pp. 91–109, 2012.

[37] G. Pretti, W. M. Coombs, C. E. Augarde, B. Sims, M. M. Puigvert and J. A. R. Gutiérrez, 'A conservation law consistent updated Lagrangian material point method for dynamic analysis,' *Journal of Computational Physics*, vol. 485, p. 112 075, 2023.

[38] Q. Jiang and M. Yeung, 'A model of point-to-face contact for three-dimensional discontinuous deformation analysis,' *Rock Mechanics and Rock Engineering*, vol. 37, no. 2, pp. 95–116, 2004.

[39] F. Liu and R. I. Borja, 'A contact algorithm for frictional crack propagation with the extended finite element method,' *International Journal for Numerical Methods in Engineering*, vol. 76, no. 10, pp. 1489–1512, 2008.

[40] S. Mao, Q. Chen, D. Li and Z. Feng, 'Modeling of free surface flows using improved material point method and dynamic adaptive mesh refinement,' *Journal of Engineering Mechanics*, vol. 142, no. 2, p. 04 015 069, 2016.

[41] N. El-Abbasi and K.-J. Bathe, 'Stability and patch test performance of contact discretizations and a new solution algorithm,' *Computers & Structures*, vol. 79, no. 16, pp. 1473–1486, 2001.

[42] K. J. Bathe and F. Brezzi, 'Stability of finite element mixed interpolations for contact problems,' *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni*, vol. 12, no. 3, pp. 167–183, 2001.

[43] F. Liu and R. I. Borja, 'Stabilized low-order finite elements for frictional contact with the extended finite element method,' *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 37-40, pp. 2456–2471, 2010.

[44] E. Burman and P. Hansbo, 'Fictitious domain finite element methods using cut elements: I. a stabilized lagrange multiplier method,' *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 41-44, pp. 2680–2686, 2010.

[45] S. Sticko, G. Ludvigsson and G. Kreiss, 'High-order cut finite elements for the elastic wave equation,' *Advances in Computational Mathematics*, vol. 46, no. 3, pp. 1–28, 2020.

[46] P. Heintz and P. Hansbo, 'Stabilized lagrange multiplier methods for bilateral elastic contact with friction,' *Computer Methods in Applied Mechanics and Engineering*, vol. 195, no. 33-36, pp. 4323–4333, 2006.

[47] K. J. Arrow, H. Azawa, L. Hurwicz, H. Uzawa, H. B. Chenery, S. M. Johnson and S. Karlin, *Studies in linear and non-linear programming*. Stanford University Press, 1958, vol. 2.

[48] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[49] M. Tur, J. Albelda, J. M. Navarro-Jimenez and J. J. Rodenas, 'A modified perturbed Lagrangian formulation for contact problems,' *Computational Mechanics*, vol. 55, no. 4, pp. 737–754, 2015.

[50] Y. Hu, Y. Fang, Z. Ge, Z. Qu, Y. Zhu, A. Pradhana and C. Jiang, 'A moving least squares material point method with displacement discontinuity and two-way rigid body coupling,' *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 1–14, 2018.

Chapter 6

Conclusions

The scope of this Chapter is to summarise the main findings (underlined in the text) of this work and discuss the open issues encountered throughout this thesis. To begin with, comments and outcomes of a more specific nature are developed first in the relevant paragraphs, while some general conclusions are given in the last section.

Dynamics of solid materials

Sections 3.1-3.5 of Chapter 3 has reviewed the continuum formulation (kinematics, mass conservation, conservation of rate of linear and angular momentum, first and second principles of thermodynamics and constitutive elasto-plastic relationships) necessary to describe a solid body in dynamic conditions in the finite strain regime. Section 3.6 has presented a discrete Material Point Method (MPM) algorithm inheriting the conservation properties of its continuum counterpart. Table 6.1 summarises and compares which properties the discrete formulations satisfy, restricting to the grid-based parts (i.e., mappings are not considered in this table).

Novelties have been established in Sections 3.6.2 and 3.6.3, where an Updated Lagrangian (UL) Conservation Law Consistent (CLC) formulation is presented. Section 3.6.6 quantifies and clarifies the dissipation sources for the Particle In Cell (PIC) or FLuid Implicit Particle (FLIP) mappings are quantified and clarified. However, in regard to Sections 3.6.2 and 3.6.3, the CLC formulation came with the cost of modifying the considered stress-strain relationships. As can be noticed from Appendix B.1, the linearisation process is particularly intricate, making its implementation more error-prone. The numerical simulations presented in Section 3.7 serve a three-fold purpose: the implementation of the algorithm is verified, a comparison of the CLC method with other techniques is performed, and the capabilities of the same formulation are demonstrated. Section 3.8 has provided an overview of the chapter and indicated possible investigation directions for future work.

Table 6.1: Summary of the compliance of the MPM discrete grid-based equations with underlying continuum basic principles. Mapping processes are not considered in this table.

Principle		Dynamics of solid materials	Dynamics of porous materials	
			solid phase	fluid phase
Mass conservation	✓ / ✗	✓	✓	✓
	How	MPs store constant mass, Eq. (3.118).	Via improved Hencky mat., Eq. (4.101).	Via primary grid-based Eq. (4.155).
Balance of rate of linear momentum	✓ / ✗	✓	✓	
	How	Via primary grid-based Eq. (3.114).	Via primary grid-based Eq. (4.154).	
Balance of rate of angular momentum	✓ / ✗	✓	✗	
	How	Via the mid-point rule, Eqs. (3.121)-(3.123).	Not assessed.	
First principle of thermodynamics	✓ / ✗	✓	✗	
	How	Via CLC stress-strain rel., Eq. (3.142).	Not assessed.	
Second principle of thermodynamics	✓ / ✗	✓	✗	
	How	Via CLC stress-strain rel, Eq. (3.142).	Not assessed.	

Dynamics of porous materials

As with the previous chapter, Sections 4.1-4.5 have been dedicated to the bases of poromechanics, revisiting kinematics, mass conservation, balance of rate of momenta, laws of thermodynamics and constitutive equations. A significant novelty is explained in Section 4.5.1.1, where the introduction of the underlined improved Hencky material permits the conservation of the solid mass constituent (see Table 6.1) in the case of Terzaghi effective stress decomposition. The improved Hencky material can equally be extended to the Finite Element Method (FEM) simulations. However, its use as a standard method is more urgent for MPM analyses where larger deformations are expected (see examples in Sections 4.7.3 and 4.7.4).

Among the possible formulations and discretisation available in the literature, Section 4.6 considers the $u - p^{(f)}$ formulation with one set of Material Points (MPs). For this method, the situation is more ambiguous with respect to the conservation properties which the algorithm should inherit from its continuum formulation, especially in the dynamic case. Further studies are required to check the balance of rate of angular momentum for the grid-based primary discretised equations (as in the procedure explained by Simo and Tarnow [1] for solid materials). As for the compliance with the laws of thermodynamics, the modification of the stress-strain relationship outlined in Section 3.6.3.1 for solid materials could be extended to the effective stress-strain relationship for porous materials. Even so, this is insufficient to guarantee that grid-based equations comply with Clausius-Duhem inequality,

as the constitutive law of the fluid part (given in Section 4.5.1.3) should be tested too. A similar assessment should be conducted for the mapping processes outlined in Section 4.6.5 relative to a $\boldsymbol{u} - p^{(f)}$ formulation. When the pressure field is mapped from the MPs to the grid and back (as in the case of a barotropic fluid), these processes should be performed with the lowest possible loss of energy (possibly zero) relative to the fluid part. Another novelty of this chapter is also the recognition of the detrimental role played by the wrong application of non-conforming pressure Boundary Conditions (BCs). This is particularly emphasised by the example in Section 4.7.2. Chapter 4 devotes attention to other aspects not notably recognised (to the best of the author's knowledge) in the MPM community. Among these, it is worth mentioning the need for boundary stabilisations to run analyses (regardless of the small-cut issue) and the limits of the Polynomial Pressure Projection (PPP) stabilisation for the considered equations. Section 4.8 terminates this chapter providing an overview of the explored techniques and indicating plausible directions of future work.

Frictional contact

Chapter 5 has presented the necessary development of kinematics and statics for frictional contact, considering Coulomb's friction law as a tribological constitutive relationship (Section 5.1). Section 5.2 has documented how the considered frictional contact can be included in a variational statement via different potential functionals. In Section 5.3, some contact algorithms available in the FEM and MPM literature are framed in a new single methodological format.

This framework has permitted a systematic comparison among different ways of computing contact forces (Section 5.4) and presents two methods to implement contact conditions (Section 5.3.1). Some conclusions can be drawn about the diverse fashions of calculating contact forces, informed by the examples provided in Section 5.5. While the Augmented Lagrangian method (ALM) represents a robust and precise technique (Karush-Kuhn-Tucker (KKT) conditions are strictly satisfied), it comes with additional equations and relative computational costs. In addition to suffering from the same problem, the Lagrange Multiplier Method (LMM) proved to be less robust, especially when the bodies separate (see the results in Section 5.5.1.2). The Penalty Method (PM) appears to be a good compromise between satisfaction of the KKT conditions (which are partially violated, though) and computational costs. However, tailoring the penalty parameters for non-linear problems is highly challenging. As appealing as the Contact Velocity-based Method (CVM) methods can appear (mainly due to their simplicity), it must be said that they cannot be regarded as representative of the physical phenomena they should represent.

As for where contact conditions are applied, the node-to-node formulation and the point-to-surface formulation have been considered. While the former method emerges as more accessible in terms of theory and implementation (no projection occurs), its mesh dependency cannot be ruled out. On the other hand, point-to-surface formulation seems to possess all the properties to remove mesh dependency, if boundaries should be reconstructed (which was beyond the scope of this work). In terms of underlying physics and implementation,

the point-to-surface method exhibits all the intricacies of the case (see, in particular, the linearisation in Appendix B.3), yet maintains a substantial theoretical background to deliver physically meaningful results.

Section 4.8 finally develops a discussion on all of these considered methodologies while providing insights into the ambiguities which require additional work.

General conclusions

From a broad perspective, the primary goal of this work has consisted of designing MPM algorithms inheriting the conservation properties of their underlying continuum formulation. The associated numerical modelling aims to represent physical phenomena involved in the embedment process of drag anchors, with the hope of paving the way to future parametric studies in this area. However, the application of these algorithms could be extended to more general geotechnical problems dealing with fully saturated soils undergoing considerable deformations. As highlighted in Chapter 5, the designed algorithms, if equipped with a proper boundary representation, could also be used for soil-structure interaction problems with the inclusion of frictional contact between surfaces. It is in this aspect in particular (i.e., the absence of boundary representation) that the illustrated MPM formulations have presented some limitations. This is particularly true for those boundaries where displacements BCs are not prescribed and their evolution cannot be easily tracked. As highlighted at different points in this work, the literature presents some techniques (see Bing *et al.* [2], Liu and Sun [3], Chandra *et al.* [4], Yamaguchi *et al.* [5]) to reconstruct the boundaries. However, at a deeper level, it should be questioned how a numerical technique based on a cloud of points could keep track of its boundaries (this issue has already pointed out by Augarde *et al.* [6], for instance). This is more easily comprehensible if considering a material that is first divided into two distinct halves that are subsequently recomposed (as might happen before and after the passage of an anchor in a soft non-cohesive soil, for example). It can be understood how crafting an algorithm that keeps track of the creation of a new surface from scratch and its subsequent closure from a cloud of MPs is not a trivial task. At the same time, relying on the mesh for this purpose introduces a strong mesh dependency and other issues as the ones described in Section 5.6.

This work cannot be regarded even remotely as a culmination for the MPM. In all fairness, such a scope is beyond any manuscript. Yet, these problems, once solved, might constitute future developments allowing the MPM to become a cornerstone technique in the Olympus of numerical methods. Among these unsolved issues (see, in this regards, Chapter 2), it is worth mentioning the integration error (see, for instance, Steffen *et al.* [7]), and the lack of an error estimate (see, e.g., Sołowski *et al.* [8]).

References

[1] J. Simo and N. Tarnow, 'The discrete energy-momentum method. Conserving algorithms for nonlinear elastodynamics,' *Zeitschrift für angewandte Mathematik und Physik ZAMP*, vol. 43, no. 5, pp. 757–792, 1992.

[2] Y. Bing, M. Cortis, T. Charlton, W. Coombs and C. Augarde, 'B-spline based boundary conditions in the material point method,' *Computers & Structures*, vol. 212, pp. 257–274, 2019.

[3] C. Liu and W. Sun, 'ILS-MPM: An implicit level-set-based material point method for frictional particulate contact mechanics of deformable particles,' *Computer Methods in Applied Mechanics and Engineering*, vol. 369, p. 113 168, 2020.

[4] B. Chandra, V. Singer, T. Teschemacher, R. Wüchner and A. Larese, 'Nonconforming Dirichlet boundary conditions in implicit material point method by means of penalty augmentation,' *Acta Geotechnica*, vol. 16, no. 8, pp. 2315–2335, 2021.

[5] Y. Yamaguchi, S. Moriguchi and K. Terada, 'Extended B-spline-based implicit material point method,' *International Journal for Numerical Methods in Engineering*, vol. 122, no. 7, pp. 1746–1769, 2021.

[6] C. E. Augarde, S. J. Lee and D. Loukidis, 'Numerical modelling of large deformation problems in geotechnical engineering: A state-of-the-art review,' *Soils and Foundations*, vol. 61, no. 6, pp. 1718–1735, 2021.

[7] M. Steffen, R. M. Kirby and M. Berzins, 'Analysis and reduction of quadrature errors in the material point method (MPM),' *International Journal for Numerical Methods in Engineering*, vol. 76, no. 6, pp. 922–948, 2008.

[8] W. T. Sołowski, M. Berzins, W. M. Coombs, J. E. Guilkey, M. Möller, Q. A. Tran, T. Adibaskoro, S. Seyedan, R. Tielen and K. Soga, 'Material point method: Overview and challenges ahead,' *Advances in Applied Mechanics*, vol. 54, pp. 113–204, 2021.

Appendix A

A few hints on tensor algebra

For the sake of completeness, a few equations necessary to manipulate tensors and their algebra are briefly discussed below. For this purpose, let us consider a *symmetric* second-order (real-valued) tensor, i.e., $\boldsymbol{A} = \boldsymbol{A}^T$. Let us also consider for the current section $n^{dim} = 3$.

If the *eigenvalues* λ_i and *eigenvectors* $\boldsymbol{v}$ of $\boldsymbol{A}$ are of interest, they satisfy the equation

$$\begin{cases} \boldsymbol{A}\,\boldsymbol{v} = \lambda\,\boldsymbol{v}; \\ \sqrt{\boldsymbol{v}\cdot\boldsymbol{v}} = 1, \end{cases} \tag{A.1}$$

where the second equation is added to ensure that the eigenvectors have unit magnitude and to avoid trivial solutions (i.e., $\boldsymbol{v} \neq \boldsymbol{0}$). Solutions to the above problem are given by the roots of the *characteristic equation*, which is defined by the following polynomial function

$$\det\left(\boldsymbol{A} - \lambda\,\boldsymbol{I}^{(2)}\right) = 0 = -\lambda^3 + I_1\left(\boldsymbol{A}\right)\lambda^2 - I_2\left(\boldsymbol{A}\right)\lambda + I_3\left(\boldsymbol{A}\right), \tag{A.2}$$

where $\boldsymbol{I}^{(2)}$ is the second-order identity tensor and I_1, I_2 and I_3 are the *principal invariants* of $\boldsymbol{A}$, given by

$$\begin{cases} I_1\left(\boldsymbol{A}\right) = A_{ii} = \mathrm{tr}\boldsymbol{A}; \\ I_2\left(\boldsymbol{A}\right) = \frac{1}{2}\left(A_{ii}^2 - A_{ij}\,A_{ij}\right) = \frac{1}{2}\left(\left(\mathrm{tr}\boldsymbol{A}\right)^2 - \mathrm{tr}\left(\boldsymbol{A}^2\right)\right); \\ I_3\left(\boldsymbol{A}\right) = \frac{1}{6}\epsilon_{ijk}\,\epsilon_{pqr}\,A_{ip}\,A_{jq}\,A_{kr} = \det\boldsymbol{A}. \end{cases} \tag{A.3}$$

Alternatively, if $(\lambda_1, \lambda_2, \lambda_3)$ are the roots of the characteristic functions, the above invariants can be written as

$$\begin{cases} I_1\left(\boldsymbol{A}\right) = \lambda_1 + \lambda_2 + \lambda_3; \\ I_2\left(\boldsymbol{A}\right) = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3; \\ I_3\left(\boldsymbol{A}\right) = \lambda_1\lambda_2\lambda_3. \end{cases} \tag{A.4}$$

It can also be shown that to each eigenvalue λ_i corresponds an eigenvector $\boldsymbol{v}_i$. In particular, if $\lambda_1 \neq \lambda_2$, the above equations gives

$$\boldsymbol{A}\,\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1; \tag{A.5}$$

$$\boldsymbol{A}\,\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2. \tag{A.6}$$

If Eq. (A.6) is subtracted to Eq. (A.5), it follows that

$$0 = (\lambda_1 - \lambda_2)\,\boldsymbol{v}_1 \cdot \boldsymbol{v}_2, \tag{A.7}$$

i.e., if two eigenvalues are different, their corresponding eigenvectors must be orthogonal.

Eigenvalues and eigenvectors also permit the spectral decomposition of the symmetric tensor $\boldsymbol{A}$, this being

$$\boldsymbol{A} = \sum_{i=1}^{n^{dim}} \lambda_i\,\boldsymbol{v}_i \otimes \boldsymbol{v}_i. \tag{A.8}$$

Appendix B

Linearisations

B.1 Linearisation of the modified internal force vector for solid dynamics

This section details the linearisation of the Conservation Law Consistent (CLC) Updated Lagrangian (UL) internal force vector with respect to the current nodal displacements. The definition of this quantity is here recalled from Eq. (3.133)

$$\left(\tilde{\mathrm{f}}_{\tilde{n}}^{int}\right)_I = \int_\Omega \frac{\partial \mathrm{N}_{Ip}^u}{\partial x_r} \left(\Delta F_{1-\tilde{n}}^{-1}\right)_{pq} \tau_{qr}^{mp} \, dV. \tag{B.1}$$

The quantity of interest is the stiffness matrix (see Eq. (3.150)), which can be obtained by making use of the chain rule as follows

$$
\begin{aligned}
\mathrm{K}_{IJ} = \frac{\partial \left(\tilde{\mathrm{f}}_{\tilde{n}}^{int}\right)_I}{\partial \left(\mathrm{u}_{n+1}\right)_J} &= \int_\Omega \left(\frac{\partial \left(\bullet\right)_I}{\partial \left(F\right)_{mn}} \frac{\partial \left(F\right)_{mn}}{\partial \left(\mathrm{u}_{n+1}\right)_J} \right) dV \\
&= \int_\Omega \left(\frac{\partial \left(\bullet\right)_I}{\partial \left(F\right)_{mn}} \frac{\partial}{\partial \left(\mathrm{u}_{n+1}\right)_J} \underbrace{\left(\frac{\partial x_m}{\partial X_n}\right)}_{=\frac{\partial \left(X + u_{n+1}\right)_m}{\partial X_n}} \right) dV = \\
&= \int_\Omega \left(\frac{\partial \left(\bullet\right)_I}{\partial \left(F\right)_{mn}} \frac{\partial}{\partial \left(\mathrm{u}_{n+1}\right)_J} \left(\delta_{mn} + \frac{\partial \mathrm{N}_{mH}^u \left(\mathrm{u}_{n+1}\right)_H}{\partial X_n} \right) \right) dV \\
&= \int_\Omega \left(\frac{\partial \left(\bullet\right)_I}{\partial \left(F\right)_{mn}} \frac{\partial \mathrm{N}_{mJ}^u}{\partial X_n} \right) dV = \\
&= \int_\Omega \left(\frac{\partial \left(\bullet\right)_I}{\partial \left(F\right)_{mn}} \frac{\partial \mathrm{N}_{mJ}^u}{\partial x_s} F_{sn} \right) dV.
\end{aligned}
\tag{B.2}
$$

In this fashion, the linearisation of the quantities appearing in the internal force vector can be performed with respect to the deformation gradient relative to the current configuration. Applying this rationale, it follows that

$$
\begin{aligned}
\mathrm{K}_{IJ} = \int_\omega \Bigg(&\frac{1}{J} \frac{\partial}{\partial F_{mn}} \left(\frac{\partial \mathrm{N}_{Ip}^u}{\partial x_r} \right) \left(\Delta F_{1-\tilde{n}}^{-1}\right)_{pq} \tau_{qr}^{mp} \frac{\partial \mathrm{N}_{mJ}^u}{\partial x_s} F_{sn} \\
&+ \frac{1}{J} \frac{\partial \mathrm{N}_{Ip}^u}{\partial x_r} \frac{\partial \left(\Delta F_{1-\tilde{n}}^{-1}\right)_{pq}}{\partial F_{mn}} \tau_{qr}^{mp} \frac{\partial \mathrm{N}_{mJ}^u}{\partial x_s} F_{sn} + \frac{1}{J} \frac{\partial \mathrm{N}_{Ip}^u}{\partial x_r} \left(\Delta F_{1-\tilde{n}}^{-1}\right)_{pq} \frac{\partial \tau_{qr}^{mp}}{\partial F_{mn}} \frac{\partial \mathrm{N}_{mJ}^u}{\partial x_s} \Bigg) dv
\end{aligned}
$$

$$= \int_\omega \left(\frac{1}{J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial X_u} \frac{\partial \left(F^{-1} \right)_{ur}}{\partial F_{mn}} \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq} \tau^{mp}_{qr} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} F_{sn} \right.$$

$$\left. + \frac{1}{J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \frac{\partial \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq}}{\partial F_{mn}} \tau^{mp}_{qr} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} F_{sn} + \frac{1}{J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq} \frac{\partial \tau^{mp}_{qr}}{\partial F_{mn}} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} F_{sn} \right) dv. \tag{B.3}$$

If focusing on the linearisation of $\Delta \boldsymbol{F}^{-1}_{1-\tilde{n}}$, it can be noticed how this quantity can be written as follows

$$\Delta \boldsymbol{F}^{-1}_{1-\tilde{n}} = \frac{\partial \boldsymbol{x}_{\tilde{n}}}{\partial \boldsymbol{X}} \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{x}} = \frac{\partial}{\partial \boldsymbol{X}} \left(\boldsymbol{x}_n + \vartheta \Delta \boldsymbol{u} \right) \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{x}} = \frac{\partial}{\partial \boldsymbol{X}} \left((1 - \vartheta) \, \boldsymbol{x}_n + \vartheta \boldsymbol{x} \right) \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{x}}$$

$$= \left((1 - \vartheta) \, \boldsymbol{F}_n + \vartheta \boldsymbol{F} \right) \boldsymbol{F}^{-1} = (1 - \vartheta) \, \Delta \boldsymbol{F}^{-1} + \vartheta \boldsymbol{I}^{(2)}, \tag{B.4}$$

where $\vartheta = \frac{1}{2}$ in the case of the mid-point rule, which is the case of interest (see discussion in Section 3.6.2.2). Thus, it follows that

$$\frac{\partial \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq}}{\partial F_{mn}} = \frac{1}{2} \frac{\partial}{\partial F_{mn}} \left((F_n)_{p\alpha} \left(F^{-1} \right)_{\alpha q} + \delta_{pq} \right) = -\frac{1}{2} (F_n)_{p\alpha} \left(F^{-1} \right)_{\alpha m} \left(F^{-1} \right)_{nq}. \tag{B.5}$$

In light of the above derivative, Eq. (B.3) becomes

$$\mathrm{K}_{IJ} = \int_\omega \left(-\frac{1}{J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_m} \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq} \tau^{mp}_{qr} \delta_{rs} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} - \frac{1}{2 J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \tau^{mp}_{qr} \left(\Delta F^{-1} \right)_{pm} \delta_{qs} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} \right.$$

$$\left. + \frac{1}{J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq} \frac{\partial \tau^{mp}_{qr}}{\partial F_{mn}} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} F_{sn} \right) dv. \tag{B.6}$$

The non-symmetric stress-tensor appearing in the above equation can be gathered in the following way (see, in this regard, the discussion in Section 3.6.3.1)

$$\check{P}_{pr} := \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq} \tau^{mp}_{qr}. \tag{B.7}$$

To collect all of the terms pre- and post-multiplying the derivatives of the shape functions, the following index saturation and desaturation must be performed

$$\frac{\partial \mathrm{N}^u_{Ip}}{\partial x_m} \check{P}_{pr} \delta_{rs} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} = \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_m} \check{P}_{ps} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} = \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \check{P}_{ps} \delta_{rm} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s}, \tag{B.8}$$

which allow to express the stiffness matrix as follows

$$\mathrm{K}_{IJ} = \int_\omega \left(-\frac{1}{J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \check{P}_{ps} \delta_{rm} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} - \frac{1}{2 J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \tau^{mp}_{sr} \left(\Delta F^{-1} \right)_{pm} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} \right.$$

$$\left. + \frac{1}{J} \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq} \frac{\partial \tau^{mp}_{qr}}{\partial F_{mn}} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} F_{sn} \right) dv. \tag{B.9}$$

The above equation can be expressed in a more compact form if the following definition is introduced

$$\check{a}_{prms} := \underbrace{-\frac{1}{J} \check{P}_{ps} \delta_{rm}}_{:=\mathcal{P}_{prms}} \underbrace{-\frac{1}{2 J} \tau^{mp}_{sr} \left(\Delta F^{-1} \right)_{pm} + \frac{1}{J} \left(\Delta F^{-1}_{1-\tilde{n}} \right)_{pq} \frac{\partial \tau^{mp}_{qr}}{\partial F_{mn}} F_{sn}}_{:=\Theta_{prms}}, \tag{B.10}$$

so that the stiffness matrix is simply given by

$$\mathrm{K}_{IJ} = \int_\omega \frac{\partial \mathrm{N}^u_{Ip}}{\partial x_r} \check{a}_{prms} \frac{\partial \mathrm{N}^u_{mJ}}{\partial x_s} \, dv. \tag{B.11}$$

B.1.1   Linearisation of the algorithmic Kirchhoff stress

As it can be noticed from Eq. (B.10), the linearisation of the modified stress with respect to the deformation gradient in also required. Recalling its definition (given by Eq. (3.142)), this computation can be performed as follows

$$
\begin{aligned}
\frac{\partial \tau_{qr}^{mp}}{\partial F_{mn}} &= \frac{\partial}{\partial F_{mn}} \left(\bar{\tau}_{qr} + \Delta^{PF} b_{qr} \left(\frac{\Delta \mathscr{H}^{int}}{\|\Delta \boldsymbol{C}\|^2} \right) \right) \\
&= \frac{\partial \bar{\tau}_{qr}}{\partial F_{mn}} + \frac{\partial \Delta^{PF} b_{qr}}{\partial F_{mn}} \left(\frac{\Delta \mathscr{H}^{int}}{\|\Delta \boldsymbol{C}\|^2} \right) \\
&\quad + \frac{\Delta^{PF} b_{qr}}{\|\Delta \boldsymbol{C}\|^4} \left(\frac{\partial \Delta \mathscr{H}^{int}}{\partial F_{mn}} \|\Delta \boldsymbol{C}\|^2 - \Delta \mathscr{H}^{int} \frac{\partial \|\Delta \boldsymbol{C}\|^2}{\partial F_{mn}} \right),
\end{aligned} \tag{B.12}
$$

where the derivative appearing in the above equation are detailed below, these being

$$
\begin{aligned}
\frac{\partial \bar{\tau}_{qr}}{\partial F_{mn}} &= \frac{1}{2} \frac{\partial}{\partial F_{mn}} \left(\Delta F_{q\alpha} \, (\tau_n)_{\alpha\beta} \, \Delta F_{r\beta} + (\tau_{n+1})_{qr} \right) \\
&= \frac{1}{2} \left(\delta_{qm} \left(F_n^{-1} \right)_{n\alpha} (\tau_n)_{\alpha\beta} \, \Delta F_{r\beta} + \Delta F_{q\alpha} \, (\tau_n)_{\alpha\beta} \, \delta_{rm} \left(F_n^{-1} \right)_{n\beta} + \frac{\partial (\tau_{n+1})_{qr}}{\partial F_{mn}} \right);
\end{aligned} \tag{B.13}
$$

$$
\begin{aligned}
\frac{\partial \Delta^{PF} b_{qr}}{\partial F_{mn}} &= \frac{\partial}{\partial F_{mn}} \left(F_{q\alpha} \, F_{\gamma\alpha} \, F_{\gamma\beta} \, F_{r\beta} - F_{q\alpha} \, (F_n)_{\gamma\alpha} \, (F_n)_{\gamma\beta} \, F_{r\beta} \right) \\
&= \delta_{qm} \, (\Delta C)_{n\beta} \, F_{r\beta} + F_{q\alpha} \, (\Delta C)_{\alpha n} \, \delta_{rm} + F_{qn} \, F_{m\beta} \, F_{r\beta} + F_{q\alpha} \, F_{m\alpha} \, F_{rn};
\end{aligned} \tag{B.14}
$$

$$
\frac{\partial \|\Delta \boldsymbol{C}\|^2}{\partial F_{mn}} = \frac{\partial}{\partial F_{mn}} \left((\Delta C)_{\alpha\beta} \, (\Delta C)_{\alpha\beta} \right) = 2 \left(F_{m\beta} \, \Delta C_{n\beta} + F_{m\alpha} \, \Delta C_{\alpha n} \right); \tag{B.15}
$$

$$
\begin{aligned}
\frac{\partial \Delta \mathscr{H}^{int}}{\partial F_{mn}} &= \frac{\partial}{\partial F_{mn}} \left(2 \left(\Delta \Psi + \Delta \mathcal{D} \right) - 2 \, \bar{\tau}_{\delta\varepsilon} \, \bar{d}_{\delta\varepsilon} \right) \\
&= 2 \frac{\partial \Delta \Psi}{\partial F_{mn}} + 2 \frac{\partial \Delta \mathcal{D}}{\partial F_{mn}} - 2 \left(\frac{\partial \bar{\tau}_{\delta\varepsilon}}{\partial F_{mn}} \bar{d}_{\delta\varepsilon} + \frac{\partial \bar{d}_{\delta\varepsilon}}{\partial F_{mn}} \bar{\tau}_{\delta\varepsilon} \right).
\end{aligned} \tag{B.16}
$$

As the most general case for this work consisted in an Hencky material (described by Eq. (3.87)) with linear hardening (see Example 3.7.6), this free energy function is considered $\Psi = \underbrace{\frac{1}{2} \boldsymbol{\epsilon}^e \, \boldsymbol{\mathcal{D}}^e \, \boldsymbol{\epsilon}^e}_{\hat{\Psi}_1(\boldsymbol{\epsilon}^e)} + \underbrace{\frac{1}{2} H \, \alpha^2}_{\hat{\Psi}_2(\alpha)}$ (with $H \geq 0$ being the hardening parameter). Hence, it follows that

$$
\begin{aligned}
\frac{\partial \Delta \Psi}{\partial F_{mn}} &= \frac{\partial \Psi \left(\boldsymbol{\epsilon}^e_{n+1}, \alpha_{n+1} \right)}{\partial F_{mn}} \\
&= \frac{\partial \hat{\Psi}_1 \left(\boldsymbol{\epsilon}^e_{n+1} \right)}{\partial \left(\epsilon^e_{n+1} \right)_{cd}} \frac{\partial \left(\epsilon^e_{n+1} \right)_{cd}}{\partial \left(\tau_{n+1} \right)_{ab}} \frac{\partial \left(\tau_{n+1} \right)_{ab}}{\partial F_{mn}} + \frac{\partial \hat{\Psi}_2 \left(\alpha_{n+1} \right)}{\partial \alpha_{n+1}} \frac{\partial \alpha_{n+1}}{\partial q_{n+1}} \frac{\partial q_{n+1}}{\partial F_{mn}} = \\
&= \left(\epsilon^e_{n+1} \right)_{ab} \frac{\partial \left(\tau_{n+1} \right)_{ab}}{\partial F_{mn}} - \alpha_{n+1} \frac{\partial q_{n+1}}{\partial F_{mn}}.
\end{aligned} \tag{B.17}
$$

The derivative of the time-difference of dissipation $\Delta \mathcal{D}$ appearing in Eq. (B.16) can be computed as follows

$$
\frac{\partial \Delta \mathcal{D}}{\partial F_{mn}} = \frac{\partial \Delta \mathcal{D}}{\partial (\epsilon^{e,trial}_{n+1})_{hk}} \frac{\partial (\epsilon^{e,trial}_{n+1})_{hk}}{\partial F_{mn}}
$$

$$= \frac{\partial}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \left(\Delta\gamma^p \, (\tau_{n+1})_{\alpha\beta} \frac{\partial\Phi}{\partial(\tau_{n+1})_{\alpha\beta}} + \Delta\gamma^p \, q_{n+1} \frac{\partial\Phi}{\partial q_{n+1}} \right) \frac{\partial(\epsilon_{n+1}^{e,trial})_{hk}}{\partial F_{mn}}$$

$$= \frac{\partial(\epsilon_{n+1}^{e,trial})_{hk}}{\partial F_{mn}} \left(\frac{\partial\Delta\gamma^p}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \left((\tau_{n+1})_{\alpha\beta} \frac{\partial\Phi}{\partial(\tau_{n+1})_{\alpha\beta}} + q_{n+1} \frac{\partial\Phi}{\partial q_{n+1}} \right) \right.$$

$$+ \Delta\gamma^p \, I_{\alpha\beta ab}^{4,sym} \frac{\partial(\tau_{n+1})_{ab}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \frac{\partial\Phi}{\partial(\tau_{n+1})_{\alpha\beta}}$$

$$+ \Delta\gamma^p \, (\tau_{n+1})_{\alpha\beta} \left(\frac{\partial}{\partial(\tau_{n+1})_{ab}} \left(\frac{\partial\Phi}{\partial(\tau_{n+1})_{\alpha\beta}} \right) \frac{\partial(\tau_{n+1})_{ab}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \right.$$

$$+ \frac{\partial}{\partial q_{n+1}} \left(\frac{\partial\Phi}{\partial(\tau_{n+1})_{\alpha\beta}} \right) \frac{\partial q_{n+1}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \right)$$

$$+ \Delta\gamma^p \left(\frac{\partial q_{n+1}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \frac{\partial\Phi}{\partial q_{n+1}} \right.$$

$$+ q_{n+1} \left(\frac{\partial}{\partial(\tau_{n+1})_{ab}} \left(\frac{\partial\Phi}{\partial q_{n+1}} \right) \frac{\partial(\tau_{n+1})_{ab}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} + \right.$$

$$\left.\left.\left.\frac{\partial^2\Phi}{(\partial q_{n+1})^2} \frac{\partial q_{n+1}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \right) \right) \right)$$

$$= \left(\left(\partial\Delta\mathcal{D}_{,(1)} \right)_{hk} + \left(\partial\Delta\mathcal{D}_{,(2)} \right)_{ab} \frac{\partial(\tau_{n+1})_{ab}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \right) \frac{\partial(\epsilon_{n+1}^{e,trial})_{hk}}{\partial F_{mn}}, \tag{B.18}$$

where the quantities $\left(\partial\Delta\mathcal{D}_{,(1)}\right)_{hk}$ and $\left(\partial\Delta\mathcal{D}_{,(2)}\right)_{ab}$ have been defined as

$$\left(\partial\Delta\mathcal{D}_{,(1)}\right)_{hk} := \frac{\partial\Delta\gamma^p}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \left((\tau_{n+1})_{\alpha\beta} \frac{\partial\Phi}{\partial(\tau_{n+1})_{\alpha\beta}} + q_{n+1} \frac{\partial\Phi}{\partial q_{n+1}} \right)$$

$$+ \Delta\gamma^p \frac{\partial q_{n+1}}{\partial(\epsilon_{n+1}^{e,trial})_{hk}} \left((\tau_{n+1})_{\alpha\beta} \frac{\partial}{\partial q_{n+1}} \left(\frac{\partial\Phi}{\partial(\tau_{n+1})_{\alpha\beta}} \right) \right.$$

$$\left. + \frac{\partial\Phi}{\partial q_{n+1}} + q_{n+1} \frac{\partial^2\Phi}{(\partial q_{n+1})^2} \right);$$

$$\left(\partial\Delta\mathcal{D}_{,(2)}\right)_{ab} := \Delta\gamma^p \left(\frac{\partial\Phi}{\partial(\tau_{n+1})_{ab}} + (\tau_{n+1})_{\alpha\beta} \frac{\partial^2\Phi}{\partial(\tau_{n+1})_{ab}\,\partial(\tau_{n+1})_{\alpha\beta}} \right.$$

$$\left. + q_{n+1} \frac{\partial}{\partial(\tau_{n+1})_{ab}} \left(\frac{\partial\Phi}{\partial q_{n+1}} \right) \right), \tag{B.19}$$

In the case of von Mises yield function (as it is the case of Examples 3.7.5 and (3.7.6)), it can be seen that the following simplifications can be done $\frac{\partial^2\Phi}{(\partial q_{n+1})^2} = 0$ and $\frac{\partial}{\partial\boldsymbol{\tau}_{n+1}} \left(\frac{\partial\Phi}{\partial\boldsymbol{\tau}_{n+1}} \right)$: $\boldsymbol{\tau}_{n+1} = \frac{\partial}{\partial\boldsymbol{\tau}_{n+1}} \left(\frac{\partial\Phi}{\partial q_{n+1}} \right) = \mathbf{0}$.

The other derivatives in Eq. (B.16) are given by

$$\frac{\partial\bar{\tau}_{\delta\varepsilon}}{\partial F_{mn}} \bar{d}_{\delta\varepsilon} = \frac{1}{2} \left(\delta_{\delta m} \left(F_n^{-1} \right)_{n\alpha} (\tau_n)_{\alpha\beta} \Delta F_{\varepsilon\beta} \right.$$

$$\left. + \Delta F_{\delta\alpha} (\tau_n)_{\alpha\beta} \delta_{\varepsilon m} \left(F_n^{-1} \right)_{n\beta} + \frac{\partial(\tau_{n+1})_{\delta\varepsilon}}{\partial F_{mn}} \right) \bar{d}_{\delta\varepsilon}$$

$$
\begin{aligned}
=\frac{1}{2}\Bigg(& \left(F_n^{-1}\right)_{n\alpha}\ (\tau_n)_{\alpha\beta}\ \Delta F_{\varepsilon\beta}\ \bar{d}_{m\varepsilon} \\
& + \Delta F_{\delta\alpha}\ (\tau_n)_{\alpha\beta}\ \left(F_n^{-1}\right)_{n\beta}\ \bar{d}_{\delta m} + \frac{\partial\left(\tau_{n+1}\right)_{\delta\varepsilon}}{\partial F_{mn}}\ \bar{d}_{\delta\varepsilon}\Bigg);
\end{aligned}
\tag{B.20}
$$

$$
\begin{aligned}
\frac{\partial\bar{d}_{\delta\varepsilon}}{\partial F_{mn}}\ \bar{\tau}_{\delta\varepsilon} &= \bar{\tau}_{\delta\varepsilon}\ \frac{\partial}{\partial F_{mn}}\left(\frac{1}{2}\left(\delta_{\delta\varepsilon} - \Delta F_{\alpha\delta}^{-1}\ \Delta F_{\alpha\varepsilon}^{-1}\right)\right) \\
&= \frac{1}{2}\ \bar{\tau}_{\delta\varepsilon}\left(\left(\Delta F^{-1}\right)_{\alpha m}\left(F^{-1}\right)_{n\delta}\left(\Delta F^{-1}\right)_{\alpha\varepsilon} + \left(\Delta F^{-1}\right)_{\alpha\delta}\left(\Delta F^{-1}\right)_{\alpha m}\left(F^{-1}\right)_{n\varepsilon}\right).
\end{aligned}
\tag{B.21}
$$

The linearisation of the Kirchhoff stress relative to a Hencky material is usually (see, for instance, de Souza Neto *et al.* [1]) performed as

$$
\frac{\partial\left(\tau_{n+1}\right)_{ab}}{\partial F_{mn}} = \frac{\partial\left(\tau_{n+1}\right)_{ab}}{\partial(\epsilon_{n+1}^{e,\,trial})_{hk}}\ \frac{\partial(\epsilon_{n+1}^{e,\,trial})_{hk}}{\partial(b_{n+1}^{e,\,trial})_{lu}}\ \frac{\partial(b_{n+1}^{e,\,trial})_{lu}}{\partial F_{mn}}.
\tag{B.22}
$$

Using Eqs. (B.13)-(B.22), Eq. (B.12) can be expressed as

$$
\frac{\partial\tau_{qr}^{mp}}{\partial F_{mn}} = \mathcal{G}_{qrmn} + \left(\mathcal{F}_{qrab}\ \frac{\partial\left(\tau_{n+1}\right)_{ab}}{\partial(\epsilon_{n+1}^{e,\,trial})_{hk}} + \mathcal{H}_{qrhk}\right)\frac{\partial(\epsilon_{n+1}^{e,\,trial})_{hk}}{\partial F_{mn}},
\tag{B.23}
$$

where the (components of the) fourth-order tensors appearing in the above equation are defined as

$$
\begin{aligned}
\mathcal{G}_{qrmn} :=&\ \frac{1}{2}\left(\delta_{qm}\left(F_n^{-1}\right)_{n\alpha}(\tau_n)_{\alpha\beta}\Delta F_{r\beta} + \Delta F_{q\alpha}(\tau_n)_{\alpha\beta}\delta_{rm}\left(F_n^{-1}\right)_{n\beta}\right) \\
&+ \frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\left(\delta_{qm}(\Delta C)_{n\beta}F_{r\beta} + F_{q\alpha}(\Delta C)_{\alpha n}\delta_{rm} + F_{qn}F_{m\beta}F_{r\beta} + F_{q\alpha}F_{m\alpha}F_{rn}\right) \\
&- \frac{1}{\|\Delta\boldsymbol{C}\|^2}\Delta^{PF}b_{qr}\Bigg(\left(F_n^{-1}\right)_{n\alpha}(\tau_n)_{\alpha\beta}\Delta F_{\varepsilon\beta}\bar{d}_{m\varepsilon} + \Delta F_{\delta\alpha}(\tau_n)_{\alpha\beta}\left(F_n^{-1}\right)_{n\beta}\bar{d}_{\delta m} \\
&+ \bar{\tau}_{\delta\varepsilon}\left(\Delta F^{-1}\right)_{\alpha m}\left(F^{-1}\right)_{n\delta}\left(\Delta F^{-1}\right)_{\alpha\varepsilon} + \bar{\tau}_{\delta\varepsilon}\left(\Delta F^{-1}\right)_{\alpha\delta}\left(\Delta F^{-1}\right)_{\alpha m}\left(F^{-1}\right)_{n\varepsilon} \\
&\hspace{4cm} + 2\frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2}\left(F_{m\beta}\Delta C_{n\beta} + F_{m\alpha}\Delta C_{\alpha n}\right)\Bigg);
\end{aligned}
\tag{B.24}
$$

$$
\mathcal{F}_{qrab} := \frac{1}{2}I_{qrab}^{4,sym} + \frac{1}{\|\Delta\boldsymbol{C}\|^2}\Delta^{PF}b_{qr}\left(2\left(\epsilon_{n+1}^e\right)_{ab} + 2\left(\partial\Delta\mathcal{D}_{,(2)}\right)_{ab} - \bar{d}_{ab}\right);
\tag{B.25}
$$

$$
\mathcal{H}_{qrhk} := \frac{2}{\|\Delta\boldsymbol{C}\|^2}\Delta^{PF}b_{qr}\left(\left(\partial\Delta\mathcal{D}_{,(1)}\right)_{hk} - \alpha_{n+1}\frac{\partial q_{n+1}}{\partial(\epsilon_{n+1}^{e,\,trial})_{hk}};\right).
\tag{B.26}
$$

As it can be seen from Eq. (B.10), the derivative Eq. (B.23) is post-multiplied by $\boldsymbol{F}$, which gives

$$
\frac{\partial\tau_{qr}^{mp}}{\partial F_{mn}}F_{sn} = \mathcal{G}_{qrmn}F_{sn} + \left(\mathcal{F}_{qrab}\frac{\partial\left(\tau_{n+1}\right)_{ab}}{\partial(\epsilon_{n+1}^{e,\,trial})_{hk}} + \mathcal{H}_{qrhk}\right)\frac{\partial(\epsilon_{n+1}^{e,\,trial})_{hk}}{\partial F_{mn}}F_{sn},
\tag{B.27}
$$

where, in particular,

$$\mathcal{G}_{qrmn} F_{sn} = \delta_{qm} \left(\frac{1}{2} \left(\tau_n^{PF} \right)_{rs} + \frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2} \Delta^{PF} b_{rs} \right) + \delta_{rm} \left(\frac{1}{2} \left(\tau_n^{PF} \right)_{qs} + \frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2} \Delta^{PF} b_{qs} \right)$$
$$+ \frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2} \left(b_{mp} b_{qs} + b_{qm} b_{rs} \right) - \frac{2}{\|\Delta\boldsymbol{C}\|^2} \Delta^{PF} b_{qr} \left(\bar{\tau}_{sm} - (\tau_{n+1})_{s\varepsilon} \, \bar{d}_{m\epsilon} + 2 \frac{\Delta\mathscr{H}^{int}}{\|\Delta\boldsymbol{C}\|^2} \Delta^{PF} b_{sm} \right);$$

$$\tag{B.28}$$

$$\mathcal{F}_{qrab} \left(\frac{\partial \left(\tau_{n+1} \right)_{ab}}{\partial F_{mn}} \right) F_{sn} = \frac{1}{2} \mathcal{F}_{qrab} \, \mathcal{D}_{abhk}^{alg} \, \mathcal{L}_{hklu} \, \mathcal{B}_{lums}; \tag{B.29}$$

$$\mathcal{H}_{qrhk} \frac{\partial (\epsilon_{n+1}^{e,trial})_{hk}}{\partial F_{mn}} F_{sn} = \frac{1}{2} \mathcal{H}_{qrhk} \, \mathcal{L}_{hklu} \, \mathcal{B}_{lums}, \tag{B.30}$$

having defined

$$\mathcal{D}_{abhk}^{alg} := \frac{\partial (\tau_{n+1})_{ab}}{\partial (\epsilon_{n+1}^{e,\,trial})_{hk}}; \tag{B.31}$$

$$\mathcal{L}_{hklu} := 2 \frac{\partial (\epsilon_{n+1}^{e,trial})_{hk}}{\partial (b_{n+1}^{e,\,trial})_{lu}} = \frac{\partial \ln(b_{n+1}^{e,\,trial})_{hk}}{\partial (b_{n+1}^{e,\,trial})_{lu}}; \tag{B.32}$$

$$\mathcal{B}_{lums} := \frac{\partial (b_{n+1}^{e,\,trial})_{lu}}{\partial F_{mn}} F_{sn} = (b_{n+1}^{e,\,trial})_{us} \delta_{lm} + (b_{n+1}^{e,\,trial})_{ls} \delta_{um}. \tag{B.33}$$

All of the above formulae allow the following expression for Eq. (B.10)

$$\check{a}_{prms} = \mathcal{P}_{prms} + \Theta_{prms} + \frac{1}{J} \left(\Delta F_{1-\tilde{n}}^{-1} \right)_{pq} \left(\mathcal{G}_{qrmn} F_{sn} + \frac{1}{2} \left(\mathcal{F}_{qrab} \, \mathcal{D}_{abhk}^{alg} + \mathcal{H}_{qrhk} \right) \mathcal{L}_{hklu} \, \mathcal{B}_{lums} \right).$$

$$\tag{B.34}$$

B.2 Linearisation of the $\boldsymbol{u} - p^{(f)}$ formulation

This section provides the details of the entries appearing in the linearised Eq. (4.170), which we recall below

$$
\begin{bmatrix}
\frac{\partial \mathbf{e}}{\partial \mathbf{u}}\Big|^{(k-1)} & \frac{\partial \mathbf{e}}{\partial \mathbf{p}}\Big|^{(k-1)} \\[2mm]
\frac{\partial \mathbf{d}}{\partial \mathbf{u}}\Big|^{(k-1)} & \frac{\partial \mathbf{d}}{\partial \mathbf{p}}\Big|^{(k-1)}
\end{bmatrix}
:=
\begin{bmatrix}
\mathrm{A}^{(1)}_{IJ} & \mathrm{B}^{(1)}_{I\mathrm{B}} \\[2mm]
\mathrm{B}^{(2)}_{\mathrm{A}J} & \mathrm{A}^{(2)}_{\mathrm{AB}}
\end{bmatrix}.
\tag{B.35}
$$

To facilitate consultation, addends of the the sub-matrices in the above Jacobian matrix are given in Sections B.2.1-B.2.4.

Hence, before detailing each component of these sub-matrices, some intermediate steps are introduced, so that the final derivatives can appear as clear as possible. For this purpose, similarly to what done in 3.6.4 for Eq. (B.2), the chain rule is applied to the derivative of a generic term with respect to the nodal displacements

$$
\begin{aligned}
\frac{\partial\,(\bullet)}{\partial \mathrm{u}_J} &= \frac{\partial\,(\bullet)}{\partial F_{mn}}\frac{\partial F_{mn}}{\partial \mathrm{u}_J} \\
&= \frac{\partial\,(\bullet)}{\partial F_{mn}}\frac{\partial}{\partial \mathrm{u}_J}\left(\frac{\partial}{\partial X_n}\left(X_m + \mathrm{N}^u_{mK}\,\mathrm{u}_K\right)\right) \\
&= \frac{\partial\,(\bullet)}{\partial F_{mn}}\frac{\partial}{\partial \mathrm{u}_J}\left(\delta_{nm} + \frac{\partial \mathrm{N}^u_{mK}}{\partial X_n}\mathrm{u}_K\right) \\
&= \frac{\partial\,(\bullet)}{\partial F_{mn}}\frac{\partial \mathrm{N}^u_{mJ}}{\partial X_n} = \frac{\partial\,(\bullet)}{\partial F_{mn}}\frac{\partial \mathrm{N}^u_{mJ}}{\partial x_\alpha}F_{\alpha n}.
\end{aligned}
\tag{B.36}
$$

Owing to the above chain of formulae, the linearisation can be performed with respect to the deformation gradient $\boldsymbol{F}$. This is the case, for instance, of the derivative of the Jacobian $J = \det \boldsymbol{F}$, which gives

$$
\frac{\partial\,(J)}{\partial F_{mn}} = J\,F^{-1}_{nm}.
\tag{B.37}
$$

The current gradient of the shape function recurs multiple times in the linearisation process. This is given by

$$
\begin{aligned}
\frac{\partial}{\partial F_{mn}}\left(\frac{\partial \mathrm{N}^u_{Ip}}{\partial x_q}\right) &= \frac{\partial}{\partial F_{mn}}\left(\frac{\partial}{\partial X_s}\left(\mathrm{N}^u_{Ip}\left(F^{-1}\right)_{sq}\right)\right) \\
&= \frac{\partial \mathrm{N}^u_{Ip}}{\partial X_s}\frac{\partial\left(F^{-1}\right)_{sq}}{\partial F_{mn}} \\
&= -\frac{\partial \mathrm{N}^u_{Ip}}{\partial X_s}F^{-1}_{sm}\,F^{-1}_{nq} \\
&= -\frac{\partial \mathrm{N}^u_{Ip}}{\partial x_m}F^{-1}_{nq}.
\end{aligned}
\tag{B.38}
$$

Owing to relationship Eq. (4.89), the Eulerian porosity n can be linearised using its relationship with the Jacobian, i.e.,

$$
\begin{aligned}
\frac{\partial n}{\partial F_{mn}} &= \frac{\partial n}{\partial J}\frac{\partial J}{\partial F_{mn}} \\
&= \frac{\partial}{\partial J}\left(1 - \frac{1}{J}\left(1 - n_0\right)\right)J\,F^{-1}_{nm}
\end{aligned}
$$

$$= \frac{1}{J} \left(1 - n_0\right) F_{nm}^{-1} = \left(1 - n\right) F_{nm}^{-1}. \tag{B.39}$$

The derivatives of the hydraulic conductivity, via the Kozeny-Carman formula (4.118), is given by

$$\begin{aligned}
\frac{\partial \kappa^{(f)}}{\partial F_{mn}} &= \frac{\partial \kappa^{(f)}}{\partial n} \frac{\partial n}{\partial F_{mn}} \\
&= \frac{\partial}{\partial n} \left(c_1 \frac{n^3}{\left(1 - n\right)^2} \right) \frac{1}{J} \left(1 - n_0\right) F_{nm}^{-1} \\
&= \left(c_1 \frac{n^2 \left(3 - n\right)}{\left(1 - n\right)^3} \right) \left(1 - n\right) F_{nm}^{-1} \\
&= \left(c_1 \frac{n^2 \left(3 - n\right)}{\left(1 - n\right)^2} \right) F_{nm}^{-1}. \tag{B.40}
\end{aligned}$$

The density of the porous material, defined by Eq. (4.28), exploits a similar relationship, resulting in

$$\begin{aligned}
\frac{\partial \rho}{\partial F_{mn}} &= \frac{\partial \rho}{\partial n} \frac{\partial n}{\partial F_{mn}} \\
&= \frac{\partial}{\partial n} \left(\rho^{(f)} n + \rho^{(sk)} \left(1 - n\right) \right) \frac{\partial n}{\partial F_{mn}} \\
&= \left(\rho^{(f)} - \rho^{(sk)} \right) \left(1 - n\right) F_{nm}^{-1}. \tag{B.41}
\end{aligned}$$

The linearisation introduced so far are useful when equations are derived with respect to the displacement field, as Eq. (B.36) highlights. However, similar linearisation must be carried on with respect to the other primary variable, namely the nodal pressure. This can be done by noticing that, having considered a barotropic fluid (see Eq. (4.115)), the linearisation of the current fluid density with respect to the nodal pressure gives

$$\begin{aligned}
\frac{\partial \rho^{(f)}}{\partial \mathrm{p_B}} &= \frac{\partial \rho^{(f)}}{\partial p^{(f)}} \frac{\partial p^{(f)}}{\partial \mathrm{p_B}} \\
&= \frac{\partial}{\partial p^{(f)}} \left(\rho_0^{(f)} \exp\left(\frac{p^{(f)}}{K^{(f)}} \right) \right) \mathrm{n_B^p} \\
&= \frac{\rho_0^{(f)}}{K^{(f)}} \exp\left(\frac{p^{(f)}}{K^{(f)}} \right) \mathrm{n_B^p} \\
&= \frac{\rho^{(f)}}{K^{(f)}} \mathrm{n_B^p}. \tag{B.42}
\end{aligned}$$

This results in a dependency of the whole density from its fluid constituent part, which gives the following derivative

$$\begin{aligned}
\frac{\partial \rho}{\partial \mathrm{p_B}} &= n \frac{\partial \rho^{(f)}}{\partial \mathrm{p_B}} \\
&= \frac{n \, \rho^{(f)}}{K^{(f)}} \mathrm{n_B^p}. \tag{B.43}
\end{aligned}$$

B.2.1 Components of $\mathbf{A}^{(1)}$

Sub-matrix $\mathbf{A}^{(1)}$ is obtained by linearising the balance of rate of linear momentum of a porous material Eq. (4.154) with respect to the nodal displacements, which results in

$$A_{IJ}^{(1)} = \left.\frac{\partial(e)_I}{\partial(u)_J}\right|^{(k-1)} = \frac{\partial\left(f^{',int}\right)_I}{\partial u_J} + \frac{\partial\left(Q_{\nabla u,p}^{(1)}\right)_{IC}}{\partial u_J}\,p_C + \frac{\partial\left(f^{ext}\right)_I}{\partial u_J}$$
$$+ \frac{\partial\,(M)_{IK}}{\partial u_J}\,\ddot{u}_K + \frac{1}{\Delta t^2\,\bar{\beta}}\,M_{IJ}. \quad \text{(B.44)}$$

Derivatives in the above equation are listed below and they make use of intermediate derivatives (B.36)-(B.41)

$$\frac{\partial f_I^{',int}}{\partial u_J} = \int_{h\Omega} \frac{\partial}{\partial u_J}\left(\frac{\partial N_{Ip}^u}{\partial x_q}\tau'_{pq}\right)dV$$
$$= \int_{h\Omega}\left(-\frac{\partial N_{Ip}^u}{\partial x_m}F_{nq}^{-1}\,\tau'_{pq} + \frac{\partial N_{Ip}^u}{\partial x_q}\frac{\partial\tau'_{pq}}{\partial F_{mn}}\right)\frac{\partial N_{mJ}^u}{\partial x_\alpha}F_{\alpha n}\,dV =$$
$$= \int_{h\Omega}\frac{\partial N_{Ip}^u}{\partial x_q}\left(-\tau'_{p\alpha}\,\delta_{qm} + \frac{\partial\tau'_{pq}}{\partial F_{mn}}F_{\alpha n}\right)\frac{\partial N_{mJ}^u}{\partial x_\alpha}\,dV; \quad \text{(B.45)}$$

$$\frac{\partial\left(Q_{\nabla u,p}^{(1)}\right)_{IC}}{\partial u_J} = -\int_{h\omega}\frac{\partial}{\partial F_{mn}}\left(\frac{\partial N_{Ip}^u}{\partial x_q}\delta_{pq}\,n_C^p\,J\right)\frac{\partial N_{mJ}^u}{\partial x_\alpha}F_{\alpha n}\,dV$$
$$= -\int_{h\omega}\left(\frac{\partial N_{Ip}^u}{\partial X_s}\frac{\partial\,(F)_{sq}^{-1}}{\partial F_{mn}} + \frac{\partial N_{Ip}^u}{\partial x_q}F_{nm}^{-1}\right)\delta_{pq}\frac{\partial N_{mJ}^u}{\partial x_\alpha}F_{\alpha n}\,n_C^p\,dv =$$
$$= -\int_{h\omega}\frac{\partial N_{Ip}^u}{\partial x_q}\left(-\delta_{p\alpha}\,\delta_{qm} + \delta_{\alpha m}\,\delta_{qp}\right)\frac{\partial N_{mJ}^u}{\partial x_\alpha}\,n_C^p\,dv; \quad \text{(B.46)}$$

$$\frac{\partial f_I^{ext}}{\partial u_J} = -\int_{h\omega}N_{Ip}^u f_p\left(\frac{\partial}{\partial F_{mn}}\,(\rho\,J)\right)\frac{\partial N_{mJ}^u}{\partial x_\alpha}F_{\alpha n}\,dV =$$
$$= -\int_{h\omega}N_{Ip}^u f_p\left(J\frac{\partial\rho}{\partial F_{mn}} + \rho\frac{\partial J}{\partial F_{mn}}\right)\frac{\partial N_{mJ}^u}{\partial x_\alpha}F_{\alpha n}\,dV$$
$$= -\int_{h\omega}N_{Ip}^u f_p\underbrace{\left(\left(\rho^{(f)} - \rho^{(sk)}\right)\frac{1}{J}\,(1 - n_0) + \rho\right)}_{=\rho^{(f)}}\delta_{\alpha m}\frac{\partial N_{mJ}^u}{\partial x_\alpha}\,dv; \quad \text{(B.47)}$$

$$\frac{\partial M_{IK}}{\partial u_J} = \int_{h\Omega}N_{Ip}^u N_{pK}^u\left(\frac{\partial}{\partial F_{mn}}\,(\rho\,J)\right)\frac{\partial N_{mJ}^u}{\partial x_\alpha}F_{\alpha n}\,dV$$
$$= \int_{h\omega}\rho^{(f)}N_{Ip}^u N_{pK}^u\,\delta_{\alpha m}\frac{\partial N_{mJ}^u}{\partial x_\alpha}\,dv. \quad \text{(B.48)}$$

B.2.2 Components of $\mathbf{B}^{(1)}$

The linearisation of the balance of rate of linear momentum for a porous material Eq. (4.154) with respect to the nodal pressure gives the sub-matrix $\mathbf{B}^{(1)}$, this being given by

$$B_{IB}^{(1)} = \left.\frac{\partial(e)_I}{\partial(p)_B}\right|^{(k-1)} = \left(Q_{\nabla u,p}^{(1)}\right)_{IB} + \frac{\partial\left(f^{ext}\right)_I}{\partial p_B} + \frac{\partial\,(M)_{IK}}{\partial p_B}\,\ddot{u}_K. \quad \text{(B.49)}$$

In this case, intermediate derivatives in Eqs. (B.42) and (B.43) are helpful to details the addends in the above equation, i.e.,

$$\frac{\partial f_I^{ext}}{\partial p_B} = - \int_{h_\omega} N_{Ip}^u \, f_p \left(\frac{\partial \rho}{\partial p_B} \right) dv = - \int_{h_\omega} N_{Ip}^u \, f_p \, \frac{n \, \rho^{(f)}}{K^{(f)}} n_B^p \, dv; \qquad (B.50)$$

$$\frac{\partial M_{IK}}{\partial p_B} = \int_{h_\omega} N_{Ip}^u \, N_{pK}^u \left(\frac{\partial \rho}{\partial p_B} \right) dv = \int_{h_\omega} N_{Ip}^u \, N_{pK}^u \left(\frac{n \, \rho^{(f)}}{K^{(f)}} n_B^p \right) dv. \qquad (B.51)$$

B.2.3 Components of $\mathbf{B}^{(2)}$

Sub-matrix $\mathbf{B}^{(2)}$ is given by linearising the fluid mass conservation Eq. (4.155) with respect to the nodal displacements, i.e.,

$$B_{AJ}^{(2)} = \frac{\partial (d)_A}{\partial (u)_J} \bigg|^{(k-1)} = \frac{\partial (C_{p,p})_{AC}}{\partial u_J} \dot{p}_C + \frac{\partial (P_{p,\nabla u})_{AK}}{\partial u_J} \dot{u}_K + \frac{\tilde{\gamma}}{\Delta t \, \tilde{\beta}} (P_{p,\nabla u})_{AJ} + \frac{\partial \left(f^{f,ext} \right)_A}{\partial u_J}$$

$$+ \frac{\partial (T_{\nabla u,\nabla u})_{AC}}{\partial u_J} p_C + \frac{\partial \left(Q_{\nabla p,u}^{(2)} \right)_{AK}}{\partial u_J} \ddot{u}_K + \frac{1}{\Delta t^2 \, \tilde{\beta}} \left(Q_{\nabla p,u}^{(2)} \right)_{AJ}. \qquad (B.52)$$

Using Eqs. (B.36)-(B.41), the above derivatives are given by

$$\frac{\partial (C_{p,p})_{AC}}{\partial u_J} = \int_{h_\omega} \frac{\rho^{(f)}}{K^{(f)}} n_A^p \, n_C^p \left(\frac{\partial}{\partial F_{mn}} (n \, J) \right) \frac{\partial N_{mJ}^u}{\partial x_\alpha} F_{\alpha n} \, dV$$

$$= \int_{h_\omega} \frac{\rho^{(f)}}{K^{(f)}} n_A^p \, n_C^p \left(J \frac{\partial n}{\partial J} + n \right) \frac{\partial J}{\partial F_{mn}} \frac{\partial N_{mJ}^u}{\partial x_\alpha} F_{\alpha n} \, dV$$

$$= \int_{h_\omega} \frac{\rho^{(f)}}{K^{(f)}} \, n_A^p \, n_C^p \left(\underbrace{\frac{1}{J} (1 - n_0) + n}_{=1} \right) \delta_{\alpha m} \frac{\partial N_{mJ}^u}{\partial x_\alpha} \, dv; \qquad (B.53)$$

$$\frac{\partial (P_{p,\nabla u})_{AK}}{\partial u_J} = \int_{h_\omega} \rho^{(f)} \, n_A^p \, \delta_{pq} \frac{\partial}{\partial F_{mn}} \left(\frac{\partial N_{pK}^u}{\partial x_q} J \right) \frac{\partial N_{mJ}^u}{\partial x_\alpha} F_{\alpha n} \, dV$$

$$= \int_{h_\omega} \rho^{(f)} \, n_A^p \, \delta_{pq} \left(-\frac{\partial N_{pK}^u}{\partial x_m} \delta_{q\alpha} + \frac{\partial N_{pK}^u}{\partial x_q} \delta_{m\alpha} \right) \frac{\partial N_{mJ}^u}{\partial x_\alpha} \, dv$$

$$= \int_{h_\omega} \rho^{(f)} \, n_A^p \, \frac{\partial N_{pK}^u}{\partial x_q} (-\delta_{mq} \delta_{p\alpha} + \delta_{pq} \delta_{m\alpha}) \frac{\partial N_{mJ}^u}{\partial x_\alpha} \, dv; \qquad (B.54)$$

$$\frac{\partial (T_{\nabla u,\nabla u})_{AC}}{\partial u_J} = \int_{h_\omega} \frac{1}{g} \frac{\partial}{\partial F_{mn}} \left(\kappa^{(f)} \frac{\partial n_A^p}{\partial x_p} \frac{\partial n_C^p}{\partial x_p} J \right) \frac{\partial N_{mJ}^u}{\partial x_\alpha} F_{\alpha n} \, dV$$

$$= \int_{h_\omega} \frac{1}{g} \left(\left(c_1 \frac{n^2 (3 - n)}{(1 - n)^2} + \kappa^{(f)} \right) \frac{\partial n_A^p}{\partial x_p} \frac{\partial n_C^p}{\partial x_p} \delta_{m\alpha} \right.$$

$$\left. - \kappa^{(f)} \left(\frac{\partial n_A^p}{\partial x_m} \frac{\partial n_C^p}{\partial x_\alpha} + \frac{\partial n_A^p}{\partial x_\alpha} \frac{\partial n_C^p}{\partial x_m} \right) \right) \frac{\partial N_{mJ}^u}{\partial x_\alpha} \, dv; \qquad (B.55)$$

$$\frac{\partial f_A^{f,ext}}{\partial u_J} = - \int_{h_\omega} \frac{\rho^{(f)}}{g} \frac{\partial}{\partial F_{mn}} \left(\kappa^{(f)} \frac{\partial n_A^p}{\partial x_p} J \right) f_p \frac{\partial N_{mJ}^p}{\partial x_\alpha} F_{\alpha n} \, dV$$

$$= -\int_{{}^h\omega} \frac{\rho^{(f)}}{g} \left(\left(c_1 \frac{n^2 (3-n)}{(1-n)^2} + \kappa^{(f)} \right) \delta_{m\alpha} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_p} - \kappa^{(f)} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_m} \delta_{p\alpha} \right) f_p \frac{\partial \mathrm{N}_{mJ}^p}{\partial x_\alpha} \, dv;$$

$$(\text{B.56})$$

$$
\frac{\partial \left(\mathrm{Q}_{\nabla p, u}^{(2)} \right)_{\mathsf{A}K}}{\partial \mathrm{u}_J} = \int_{{}^h\omega} \frac{\rho^{(f)}}{g} \frac{\partial}{\partial F_{mn}} \left(\kappa^{(f)} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_p} J \right) \mathrm{N}_{pK}^u \frac{\partial \mathrm{N}_{mJ}^u}{\partial x_\alpha} F_{\alpha n} \, dV
$$

$$
= \int_{{}^h\omega} \frac{\rho^{(f)}}{g} \left(\left(c_1 \frac{n^2 (3-n)}{(1-n)^2} + \kappa^{(f)} \right) \delta_{m\alpha} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_p} - \kappa^{(f)} \frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_m} \delta_{p\alpha} \right) \mathrm{N}_{pK}^u \frac{\partial \mathrm{N}_{mJ}^u}{\partial x_\alpha} \, dv.
$$

$$(\text{B.57})$$

B.2.4 Components of $\mathbf{A}^{(2)}$

The bottom-right sub-matrix $\mathbf{A}^{(2)}$ is the result of the linearisation of fluid mass conservation Eq. (4.155) with respect to nodal pressure. This is given by

$$
A_{\mathsf{AB}}^{(2)} = \frac{\partial (\mathrm{d})_\mathsf{A}}{\partial (\mathrm{p})_\mathsf{B}} \bigg|^{(k-1)} = \frac{\partial (\mathrm{C}_{p,p})_{\mathsf{AC}}}{\partial \mathrm{p}_\mathsf{B}} \dot{\mathrm{p}}_\mathsf{C} + \frac{\tilde{\gamma}}{\Delta t \, \tilde{\beta}} (\mathrm{C}_{p,p})_{\mathsf{AB}} + \frac{\partial (\mathrm{P}_{p,\nabla u})_{\mathsf{A}K}}{\partial \mathrm{p}_\mathsf{B}} \dot{\mathrm{u}}_K + \frac{\partial \left(\mathrm{f}^{f,ext} \right)_\mathsf{A}}{\partial \mathrm{p}_\mathsf{B}}
$$

$$
+ (\mathrm{T}_{\nabla p, \nabla p})_{\mathsf{AB}} + \frac{\partial \left(\mathrm{Q}_{\nabla p, u}^{(2)} \right)_{\mathsf{A}K}}{\partial \mathrm{p}_\mathsf{B}} \ddot{\mathrm{u}}_K. \quad (\text{B.58})
$$

where, with the help provided by Eqs. (B.42) and (B.43), the derivatives in the above formula can be computed as follows

$$
\frac{\partial (\mathrm{C}_{p,p})_{\mathsf{AC}}}{\partial \mathrm{p}_\mathsf{B}} = \int_{{}^h\omega} \frac{n}{K^{(f)}} \mathrm{n}_\mathsf{A}^p \, \mathrm{n}_\mathsf{C}^p \left(\frac{\partial \rho^{(f)}}{\partial \mathrm{p}_\mathsf{B}} \right) dv = \int_{{}^h\omega} \frac{n}{K^{(f)}} \mathrm{n}_\mathsf{A}^p \, \mathrm{n}_\mathsf{C}^p \left(\frac{\rho^{(f)}}{K^{(f)}} \mathrm{n}_\mathsf{B}^p \right) dv; \quad (\text{B.59})
$$

$$
\frac{\partial (\mathrm{P}_{p,\nabla u})_{\mathsf{A}K}}{\partial \mathrm{p}_\mathsf{B}} = \int_{{}^h\omega} \mathrm{n}_\mathsf{A}^p \, \delta_{pq} \left(\frac{\partial \mathrm{N}_{pK}^u}{\partial x_q} \right) \left(\frac{\partial \rho^{(f)}}{\partial \mathrm{p}_\mathsf{B}} \right) dv = \int_{{}^h\omega} \mathrm{n}_\mathsf{A}^p \, \delta_{pq} \left(\frac{\partial \mathrm{N}_{pK}^u}{\partial x_q} \right) \left(\frac{\rho^{(f)}}{K^{(f)}} \mathrm{n}_\mathsf{B}^p \right) dv;
$$

$$(\text{B.60})$$

$$
\frac{\partial \mathrm{f}_\mathsf{A}^{f,ext}}{\partial \mathrm{p}_\mathsf{B}} = -\int_{{}^h\omega} \frac{\kappa^{(f)}}{g} \left(\frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_p} \right) f_p \left(\frac{\partial \rho^{(f)}}{\partial \mathrm{p}_\mathsf{B}} \right) dv = -\int_{{}^h\omega} \frac{\kappa^{(f)}}{g} \left(\frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_p} \right) f_p \left(\frac{\rho^{(f)}}{K^{(f)}} \mathrm{n}_\mathsf{B}^p \right) dv;
$$

$$(\text{B.61})$$

$$
\frac{\partial \left(\mathrm{Q}_{\nabla p, u}^{(2)} \right)_{\mathsf{A}K}}{\partial \mathrm{p}_\mathsf{B}} = \int_{{}^h\omega} \frac{\kappa^{(f)}}{g} \left(\frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_p} \right) \mathrm{N}_{pK}^u \left(\frac{\partial \rho^{(f)}}{\partial \mathrm{p}_\mathsf{B}} \right) dv = \int_{{}^h\omega} \frac{\kappa^{(f)}}{g} \left(\frac{\partial \mathrm{n}_\mathsf{A}^p}{\partial x_p} \right) \mathrm{N}_{pK}^u \left(\frac{\rho^{(f)}}{K^{(f)}} \mathrm{n}_\mathsf{B}^p \right) dv.
$$

$$(\text{B.62})$$

B.3 Linearisation of frictional contact formulations

This section details the linearisation process for the frictional contact formulations provided in Chapter 5 in the most general case. The Jacobian matrix for most[*] of the formulations of Chapter 5 can be obtained by vanishing wisely some terms or slightly modifying some others of the below equations. The most general method is the symmetric Augmented Lagrangian method (ALM) formulation using the point-to-surface technique, described by Eq. (5.65). Since this section focuses only on the terms expressing frictional contact, Eq. (5.92) and its Jacobian matrix (5.96) are of interest here. For readability, the Jacobian matrix (5.96) is here re-written

$$
\mathbf{J}^c = \begin{bmatrix}
\dfrac{\partial \mathrm{r}_I^{(s),c}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \mathrm{r}_I^{(s),c}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \mathrm{r}_I^{(s),c}}{\partial \lambda_{\mathcal{J}}} \\[2.2em]
\dfrac{\partial \mathrm{r}_J^{(m),c}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \mathrm{r}_J^{(m),c}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \mathrm{r}_J^{(m),c}}{\partial \lambda_{\mathcal{J}}} \\[2.2em]
\dfrac{\partial \chi_{\mathcal{I}}}{\partial \mathrm{u}_H|_X} & \dfrac{\partial \chi_{\mathcal{I}}}{\partial \mathrm{u}_K|_Y} & \dfrac{\partial \chi_{\mathcal{I}}}{\partial \lambda_{\mathcal{J}}}
\end{bmatrix}. \tag{B.63}
$$

Each of the following sections is dedicated to report the different columns of the above matrix.

B.3.1 First column

The entries of the first column of matrix (B.63) can be extensively written as follows

$$
\frac{\partial \mathrm{r}_I^{(s),c}}{\partial \mathrm{u}_H|_X} = - \int_{{}^h \Gamma^{(s)}} \mathrm{N}_{Ii}^u (\boldsymbol{X}) \left(-\frac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_H|_X} n_i^x - \hat{\lambda}^n \frac{\partial n_i^x}{\partial \mathrm{u}_H|_X} \right.
$$
$$
\left. + \frac{\partial \left(P_{B(\boldsymbol{n},\mu \hat{\lambda}^n)} \right)_i}{\partial \mathrm{u}_H|_X} (\boldsymbol{\lambda} - r \boldsymbol{v}^c) \right) dA^{(s)}; \tag{B.64}
$$

$$
\frac{\partial \mathrm{r}_J^{(m),c}}{\partial \mathrm{u}_H|_X} = \int_{{}^h \Gamma^{(s)}} \left(\frac{\partial \mathrm{N}_{Ji}^u (\boldsymbol{Y})}{\partial \mathrm{u}_H|_X} \left(-\hat{\lambda}^n n_i^x + \left(P_{B(\boldsymbol{n},\mu \hat{\lambda}^n)} \right)_i (\boldsymbol{\lambda} - r \boldsymbol{v}^c) \right) \right.
$$
$$
\left. + \mathrm{N}_{Ji}^u (\boldsymbol{Y}) \left(-\frac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_H|_X} n_i^x - \hat{\lambda}^n \frac{\partial n_i^x}{\partial \mathrm{u}_H|_X} + \frac{\partial \left(P_{B(\boldsymbol{n},\mu \hat{\lambda}^n)} \right)_i}{\partial \mathrm{u}_H|_X} (\boldsymbol{\lambda} - r \boldsymbol{v}^c) \right) \right) dA^{(s)}; \tag{B.65}
$$

$$
\frac{\partial \chi_{\mathcal{I}}}{\partial \mathrm{u}_H|_X} = -\frac{1}{r} \int_{{}^h \Gamma^{(s)}} \mathrm{N}_{\mathcal{I}i}^{\lambda} (\boldsymbol{X}) \left(\frac{\partial \mathcal{C}_i}{\partial v_{\alpha}^c} \frac{\partial v_{\alpha}^c}{\partial \mathrm{u}_H|_X} + \frac{\partial \mathcal{C}_i}{\partial n_{\beta}^x} \frac{\partial n_{\beta}^x}{\partial \mathrm{u}_H|_X} + \frac{\partial \mathcal{C}_i}{\partial g} \frac{\partial g}{\partial \mathrm{u}_H|_X} \right) dA^{(s)}. \tag{B.66}
$$

The above equations, in turn, rely on the derivatives with respect to the nodal displacements of the secondary body $\mathbf{u}|_X$. To compute these, the different definitions must be recalled.

---

[*]The exception to the formulations which can be linearised using the formulae available in this section is the Contact Velocity-based Method (CVM). The decision to exclude this method rather than others is because, overall, the linearisation of this one involves fewer challenges.

Using Eq. (5.60), the following derivative can be computed

$$\frac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_H|_X} = \begin{cases} -\left(\mathrm{N}_{j\mathcal{J}}^{\lambda}\left(\boldsymbol{X}\right)\lambda_{\mathcal{J}}\frac{\partial n_j^x}{\partial \mathrm{u}_H|_X} + r\frac{\partial g}{\partial \mathrm{u}_H|_X} \right), & \text{if } \hat{\lambda}^n \geq 0 \quad \text{contact;} \\ 0, & \text{otherwise} \quad \text{no contact,} \end{cases} \tag{B.67}$$

where the derivatives appearing above are given via the definitions of the normal direction (5.4) and the gap function (5.63), resulting in

$$\frac{\partial n_i^x}{\partial \mathrm{u}_H|_X} = -T_{i\alpha}^{n^x}\frac{\partial \mathrm{N}_{H\beta}^u\left(\boldsymbol{X}\right)}{\partial x_\alpha}n_\beta^x; \tag{B.68}$$

$$\frac{\partial g}{\partial \mathrm{u}_H|_X} = -\frac{1}{\boldsymbol{n}^y\cdot\boldsymbol{n}^x}n_\alpha^y\left(\mathrm{N}_{\alpha H}^u\left(\boldsymbol{X}\right) + g\frac{\partial n_\alpha^x}{\partial \mathrm{u}_H|_X}\right). \tag{B.69}$$

The projector operator (5.62) gives the following derivative

$$\frac{\partial \left(P_{B(\boldsymbol{n},\mu\hat{\lambda}^n)}\right)_i}{\partial \mathrm{u}_H|_X}\left(\boldsymbol{\lambda} - r\boldsymbol{v}^c\right) =$$

$$= \begin{cases} \frac{\partial \hat{\lambda}_i^{t,tr}}{\partial \mathrm{u}_H|_X} = \frac{\partial T_{ij}^{n^x}}{\partial \mathrm{u}_H|_X}\left(\mathrm{N}_{j\mathcal{J}}^{\lambda}\left(\boldsymbol{X}\right)\lambda_{\mathcal{J}} - rv_j^c\right) - rT_{ij}^{n^x}\frac{\partial v_j^c}{\partial \mathrm{u}_H|_X}, & \text{if } ||\hat{\boldsymbol{\lambda}}^{t,tr}|| \leq \mu\hat{\lambda}^n \quad \text{stick;} \\ \mu\left(\frac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_H|_X}\frac{\hat{\lambda}_i^{t,tr}}{||\hat{\boldsymbol{\lambda}}^{t,tr}||} + \frac{\hat{\lambda}^n}{||\hat{\boldsymbol{\lambda}}^{t,tr}||}\left(\frac{\partial \hat{\lambda}_i^{t,tr}}{\partial \mathrm{u}_H|_X} - \hat{\lambda}_i^{t,tr}\frac{\hat{\lambda}_\alpha^{t,tr}}{||\hat{\boldsymbol{\lambda}}^{t,tr}||^2}\frac{\partial \hat{\lambda}_\alpha^{t,tr}}{\partial \mathrm{u}_H|_X}\right)\right), & \text{otherwise} \quad \text{slip,} \end{cases} \tag{B.70}$$

where the derivatives of the plane tangential to the normal secondary surface (5.5), the trial stick state along the tangential direction (5.61), and the contact velocity (5.90) are

$$\frac{\partial T_{ij}^{n^x}}{\partial \mathrm{u}_H|_X} = -\left(\frac{\partial n_i^x}{\partial \mathrm{u}_H|_X}n_j^x + n_i^x\frac{\partial n_j^x}{\partial \mathrm{u}_H|_X}\right); \tag{B.71}$$

$$\frac{\partial \hat{\lambda}_i^{t,tr}}{\partial \mathrm{u}_H|_X} = \frac{\partial T_{ij}^{n^x}}{\partial \mathrm{u}_H|_X}\left(\mathrm{N}_{j\mathcal{J}}^{\lambda}\left(\boldsymbol{X}\right)\lambda_{\mathcal{J}} - rv_j^c\right) - rT_{ij}^{n^x}\frac{\partial v_j^c}{\partial \mathrm{u}_H|_X}; \tag{B.72}$$

$$\frac{\partial v_\alpha^c}{\partial \mathrm{u}_H|_X} = \frac{2}{\Delta t}\left(\mathrm{N}_{\alpha H}^u\left(\boldsymbol{X}\right) - \left(\delta_{\alpha\beta} - \frac{n_\alpha^x n_\beta^y}{\boldsymbol{n}^x\cdot\boldsymbol{n}^y}\right)\left(\mathrm{N}_{\beta H}^u\left(\boldsymbol{X}\right) - g\frac{\partial \mathrm{N}_{H\gamma}^u\left(\boldsymbol{X}\right)}{\partial x_\beta}n_\gamma^x\right)\right)$$
$$+ \frac{\partial g}{\partial \mathrm{u}_H|_X}\left(\dot{n}_{\tilde{n}}^x\right)_\alpha + g\frac{\partial \left(\dot{n}_{\tilde{n}}^x\right)_\alpha}{\partial \mathrm{u}_H|_X}. \tag{B.73}$$

In particular, the second term appearing on the Right-Hand Side (RHS) of the last equation is computed with the help of Eq. (5.21). The derivative of the rate of the normal direction to the secondary surface appearing in Eq. (B.66) is computed by recalling its mid-point time discretisation given by Eq. (5.91), and it gives

$$\frac{\partial \left(\dot{n}_{\tilde{n}}^x\right)_\alpha}{\partial \mathrm{u}_H|_X} = \frac{1}{\Delta t}\left(\frac{\partial \left(n^x\right)_\alpha}{\partial \mathrm{u}_H|_X}\right). \tag{B.74}$$

Eqs. (B.67)-(B.74) listed above are necessary to compute only the first entry, i.e., Eq. (B.64). If the second entry described by Eq. (B.65) is considered, the following equation must also be computed

$$\frac{\partial \mathrm{N}_{\mathcal{J}i}^u\left(\boldsymbol{\mathcal{Y}}\right)}{\partial \mathrm{u}_H|_X} = \frac{\partial \mathrm{N}_{\mathcal{J}i}^u\left(\boldsymbol{\mathcal{Y}}\right)}{\partial y_\beta}\left(\delta_{\beta\gamma} - \frac{n_\beta^x n_\gamma^y}{\boldsymbol{n}^x\cdot\boldsymbol{n}^y}\right)\left(\mathrm{N}_{\gamma H}^u\left(\boldsymbol{X}\right) - g\frac{\partial \mathrm{N}_{mH}^u\left(\boldsymbol{X}\right)}{\partial x_\gamma}n_m^x\right), \tag{B.75}$$

where Eq. (5.22) is used to computed the RHS of the above equation.

As for the third entry given by Eq. (B.66), the constraint (5.89) must be recalled to compute its derivatives, which are as follow

$$
\frac{\partial \mathcal{C}_i}{\partial v_\alpha^c} = -
\begin{cases}
\dfrac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial v_\alpha^c} = -r\,T_{i\alpha}^{n^x}, & \text{if } ||\hat{\boldsymbol{\lambda}}^{t,\,tr}|| \leq \mu \hat{\lambda}^n \quad \text{stick;} \\[2ex]
\mu \dfrac{\hat{\lambda}^n}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||} \left(\dfrac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial v_\alpha^c} - \dfrac{\hat{\lambda}_\beta^{t,\,tr}}{||\hat{\boldsymbol{\lambda}}^{t,tr}||^2} \dfrac{\partial \hat{\lambda}_\beta^{t,\,tr}}{\partial v_\alpha^c} \hat{\lambda}_i^{t,\,tr} \right), & \text{otherwise} \quad\quad\quad \text{slip;}
\end{cases}
$$

(B.76)

$$
\frac{\partial \mathcal{C}_i}{\partial n_\beta^x} = \frac{\partial \hat{\lambda}^n}{\partial n_\beta^x} n_i^x + \hat{\lambda}^n\, \delta_{i\beta} -
\begin{cases}
\dfrac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial n_\beta^x}, & \text{if } \left\| \hat{\boldsymbol{\lambda}}^{t,\,tr} \right\| \leq \mu \hat{\lambda}^n \quad \text{stick;} \\[2ex]
\mu \left(\dfrac{\partial \hat{\lambda}^n}{\partial n_\beta^x} \dfrac{\hat{\lambda}_i^{t,\,tr}}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||} + \dfrac{\hat{\lambda}^n}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||} \left(\dfrac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial n_\beta^x} - \dfrac{\hat{\lambda}_i^{t,\,tr}}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||^2} \dfrac{\partial \hat{\lambda}_\gamma^{t,\,tr}}{\partial n_\beta^x} \hat{\lambda}_\gamma^{t,\,tr} \right) \right), & \\[1ex]
& \text{otherwise} \quad\quad \text{slip;}
\end{cases}
$$

(B.77)

$$
\frac{\partial \mathcal{C}_i}{\partial g} = n_i^x \frac{\partial \hat{\lambda}^n}{\partial g} -
\begin{cases}
\mathbf{0}, & \text{if } ||\hat{\boldsymbol{\lambda}}^{t,\,tr}|| \leq \mu \hat{\lambda}^n \quad \text{stick;} \\[2ex]
\mu \dfrac{\hat{\lambda}_i^{t,\,tr}}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||} \dfrac{\partial \hat{\lambda}^n}{\partial g}, & \text{otherwise} \quad\quad\quad \text{slip,}
\end{cases}
$$

(B.78)

with

$$
\frac{\partial \hat{\lambda}^n}{\partial n_\beta^x} =
\begin{cases}
-\mathrm{N}_{\beta\mathcal{J}}^\lambda (\boldsymbol{X})\,\lambda_\mathcal{J}, & \text{if } \hat{\lambda}^n \geq 0 \quad \text{contact;} \\[1ex]
\mathbf{0}, & \text{otherwise} \quad \text{no contact,}
\end{cases}
$$

(B.79)

$$
\frac{\partial \hat{\lambda}^n}{\partial g} =
\begin{cases}
-r, & \text{if } \hat{\lambda}^n \geq 0 \quad \text{contact;} \\[1ex]
0, & \text{otherwise} \quad \text{no contact,}
\end{cases}
$$

(B.80)

$$
\frac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial n_\beta^x} = - \left(\delta_{i\beta} n_\epsilon^x + n_i^x \delta_{\epsilon\beta} \right) \left(\lambda_\epsilon - r^t v_\epsilon^c \right).
$$

(B.81)

B.3.2 Second column

The second row of entries of the Jacobian matrix (5.96) is given by the following equations

$$
\frac{\partial \mathrm{r}_I^{(s),c}}{\partial \mathrm{u}_K|_Y} = - \int_{{}_h\Gamma^{(s)}} \mathrm{N}_{Ii}^u(\boldsymbol{X}) \left(-\frac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_K|_Y} n_i^x + \frac{\partial \left(P_{B(\boldsymbol{n},\mu\hat{\lambda}^n)} \right)_i}{\partial \mathrm{u}_K|_Y} (\boldsymbol{\lambda} - r\boldsymbol{v}^c) \right) dA^{(s)};
$$

(B.82)

$$
\frac{\partial \mathrm{r}_J^{(m),c}}{\partial \mathrm{u}_K|_Y} = \int_{{}_h\Gamma^{(s)}} \left(\frac{\partial \mathrm{N}_{Ji}^u(\boldsymbol{\mathcal{Y}})}{\partial \mathrm{u}_K|_Y} \left(-\hat{\lambda}^n\, n_i^x + \left(P_{B(\boldsymbol{n},\mu\hat{\lambda}^n)} \right)_i (\boldsymbol{\lambda} - r\boldsymbol{v}^c) \right) \right.
$$

$$
\left. + \mathrm{N}_{Ji}^u(\boldsymbol{\mathcal{Y}}) \left(-\frac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_K|_Y} n_i^x + \frac{\partial \left(P_{B(\boldsymbol{n},\mu\hat{\lambda}^n)} \right)_i}{\partial \mathrm{u}_K|_Y} (\boldsymbol{\lambda} - r\boldsymbol{v}^c) \right) \right) dA^{(s)};
$$

(B.83)

$$
\frac{\partial \chi_\mathcal{I}}{\partial \mathrm{u}_K|_Y} = -\frac{1}{r} \int_{{}_h\Gamma^{(s)}} \mathrm{N}_{\mathcal{I}i}^\lambda(\boldsymbol{X}) \left(\frac{\partial \mathcal{C}_i}{\partial v_\alpha^c} \frac{\partial v_\alpha^c}{\partial \mathrm{u}_K|_Y} + \frac{\partial \mathcal{C}_i}{\partial g} \frac{\partial g}{\partial \mathrm{u}_K|_Y} \right) dA^{(s)}.
$$

(B.84)

Similarly to Section B.3.1, the partial derivatives appearing in the above equations with respect to the main surface displacements $\mathbf{u}|_Y$ must be calculated by recalling their respective definitions. Using Eq. (5.60), its partial derivative is

$$\frac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_K|_Y} = \begin{cases} -r \dfrac{\partial g}{\partial \mathrm{u}_K|_Y}, & \text{if } \hat{\lambda}^n \geq 0 \quad \text{contact;} \\ \mathbf{0}, & \text{otherwise} \quad \text{no contact,} \end{cases} \tag{B.85}$$

with

$$\frac{\partial g}{\partial \mathrm{u}_K|_Y} = \frac{1}{\boldsymbol{n}^y \cdot \boldsymbol{n}^x} n_\alpha^y \mathrm{N}_{\alpha K}^u (\boldsymbol{\mathcal{Y}}). \tag{B.86}$$

The derivative of the projector operator (5.62) is as follows

$$\frac{\partial \left(P_{B(\boldsymbol{n},\mu\hat{\lambda}^n)} \right)_i}{\partial \mathrm{u}_K|_Y} (\boldsymbol{\lambda} - r\boldsymbol{v}^c) =$$

$$= \begin{cases} \dfrac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial \mathrm{u}_K|_Y} = -r\, T_{ij}^{n^x} \dfrac{\partial v_j^c}{\partial \mathrm{u}_K|_Y}, & \text{if } ||\hat{\boldsymbol{\lambda}}^{t,\,tr}|| \leq \mu\hat{\lambda}^n \quad \text{stick;} \\ \mu \left(\dfrac{\partial \hat{\lambda}^n}{\partial \mathrm{u}_K|_Y} \dfrac{\hat{\lambda}_i^{t,\,tr}}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||} + \dfrac{\hat{\lambda}^n}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||} \left(\dfrac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial \mathrm{u}_K|_Y} - \hat{\lambda}_i^{t,\,tr} \dfrac{\hat{\lambda}_\alpha^{t,\,tr}}{||\hat{\boldsymbol{\lambda}}^{t,\,tr}||^2} \dfrac{\partial \hat{\lambda}_\alpha^{t,\,tr}}{\partial \mathrm{u}_K|_Y} \right) \right), & \text{otherwise} \quad \text{slip,} \end{cases} \tag{B.87}$$

where, in particular,

$$\frac{\partial \hat{\lambda}_i^{t,\,tr}}{\partial \mathrm{u}_K|_Y} = -r\, T_{ij}^{n^x} \frac{\partial v_j^c}{\partial \mathrm{u}_K|_Y}; \tag{B.88}$$

$$\frac{\partial v_\gamma^c}{\partial \mathrm{u}_K|_Y} = -\frac{2}{\Delta t} \left(\mathrm{N}_{\gamma K}^u (\boldsymbol{\mathcal{Y}}) - \left(\delta_{\gamma\beta} - \frac{n_\gamma^x n_\beta^y}{\boldsymbol{n}^x \cdot \boldsymbol{n}^y} \right) \mathrm{N}_{\beta K}^u (\boldsymbol{\mathcal{Y}}) \right) + \frac{\partial g}{\partial \mathrm{u}_K|_Y} (\dot{n}_{\tilde{n}}^x)_\gamma. \tag{B.89}$$

The derivative of the shape functions computed on the main surface with respect to the nodal displacement of the same body (appearing in Eq. (B.83)) is computed with the help of Eq. (5.22) as follows

$$\frac{\partial \mathrm{N}_{Ji}^u (\boldsymbol{\mathcal{Y}})}{\partial \mathrm{u}_K|_Y} = \frac{\partial \mathrm{N}_{Ji}^u (\boldsymbol{\mathcal{Y}})}{\partial y_\beta} \left(\delta_{\beta\gamma} - \frac{n_\beta^x n_\gamma^y}{\boldsymbol{n}^x \cdot \boldsymbol{n}^y} \right) \mathrm{N}_{\gamma K}^u (\boldsymbol{\mathcal{Y}}). \tag{B.90}$$

B.3.3 Third column

The last column of the Jacobian matrix (5.96) presents the following entries

$$\frac{\partial \mathrm{r}_I^{(s),c}}{\partial \lambda_{\mathcal{J}}} = -\int_{h\Gamma^{(s)}} \mathrm{N}_{Ii}^u (\boldsymbol{X}) \left(-\frac{\partial \hat{\lambda}^n}{\partial \lambda_{\mathcal{J}}} n_i^x + \frac{\partial \left(P_{B(\boldsymbol{n},\mu\hat{\lambda}^n)} \right)_i}{\partial \lambda_{\mathcal{J}}} (\boldsymbol{\lambda} - r\boldsymbol{v}^c) \right) dA^{(s)}; \tag{B.91}$$

$$\frac{\partial \mathrm{r}_J^{(m),c}}{\partial \lambda_{\mathcal{J}}} = \int_{h\Gamma^{(s)}} \mathrm{N}_{Ji}^u (\boldsymbol{\mathcal{Y}}) \left(-\frac{\partial \hat{\lambda}^n}{\partial \lambda_{\mathcal{J}}} n_i^x + \frac{\partial \left(P_{B(\boldsymbol{n},\mu\hat{\lambda}^n)} \right)_i}{\partial \lambda_{\mathcal{J}}} (\boldsymbol{\lambda} - r\boldsymbol{v}^c) \right) dA^{(s)}; \tag{B.92}$$

$$\frac{\partial \chi_{\mathcal{I}}}{\partial \lambda_{\mathcal{J}}} = -\frac{1}{r} \int_{h\Gamma^{(s)}} \mathrm{N}_{\mathcal{I}i}^\lambda (\boldsymbol{X}) \frac{\partial \mathcal{C}_i}{\partial \lambda_{\mathcal{J}}} dA^{(s)}. \tag{B.93}$$

Recalling Eq. (5.60), its derivative with respect to the nodal Lagrange Multiplier (LM) λ is as follows

$$\frac{\partial \hat{\lambda}^n}{\partial \lambda_{\mathcal{J}}} = \begin{cases} n_j^x \mathrm{N}_{j\mathcal{J}}^\lambda (\boldsymbol{X}), & \text{if } \hat{\lambda}^n \geq 0 \quad \text{contact;} \\ \mathbf{0}, & \text{otherwise} \quad \text{no contact.} \end{cases} \tag{B.94}$$

The partial derivative with respect to the same argument of the projection operation (5.62) is given by

$$
\frac{\partial \left(P_{B\left(\boldsymbol{n}^x, \mu \hat{\lambda}^n\right)} \right)_i}{\partial \lambda_{\mathcal{J}}} \left(\boldsymbol{\lambda} - r \boldsymbol{v}^c \right) =
$$

$$
= \begin{cases}
\dfrac{\partial \hat{\lambda}_i^{t,tr}}{\partial \lambda_{\mathcal{J}}} = T_{ij}^{n^x} \mathrm{N}_{j\mathcal{J}}^{\lambda} \left(\boldsymbol{X} \right), & \text{if } \left\| \hat{\boldsymbol{\lambda}}^{t,tr} \right\| \leq \mu \hat{\lambda}^n \quad \text{stick;} \\[2ex]
\mu \left(\dfrac{\partial \hat{\lambda}^n}{\partial \lambda_{\mathcal{J}}} \dfrac{\hat{\lambda}_i^{t,tr}}{\|\hat{\boldsymbol{\lambda}}^{t,tr}\|} + \dfrac{\hat{\lambda}^n}{\|\hat{\boldsymbol{\lambda}}^{t,tr}\|} \left(\dfrac{\partial \hat{\lambda}_i^{t,tr}}{\partial \lambda_{\mathcal{J}}} - \hat{\lambda}_i^{t,tr} \dfrac{\hat{\lambda}_\alpha^{t,tr}}{\|\hat{\boldsymbol{\lambda}}^{t,tr}\|^2} \dfrac{\partial \hat{\lambda}_\alpha^{t,tr}}{\partial \lambda_{\mathcal{J}}} \right) \right), & \text{otherwise} \quad \text{slip.}
\end{cases}
$$

$$\text{(B.95)}$$

Last, the derivative of the constraint (5.89) gives

$$
\frac{\partial \mathcal{C}_i}{\partial \lambda_{\mathcal{J}}} = \mathrm{N}_{i\mathcal{J}}^{\lambda} + n_i^x \frac{\partial \hat{\lambda}^n}{\partial \lambda_{\mathcal{J}}} - \begin{cases}
\dfrac{\partial \hat{\lambda}_i^{t,tr}}{\partial \lambda_{\mathcal{J}}} = T_{i\epsilon}^{n_x} \left(\mathrm{N}_{\epsilon\mathcal{J}}^{\lambda} \left(\boldsymbol{X} \right) \right), & \text{if } \left\| \hat{\boldsymbol{\lambda}}^{t,tr} \right\| \leq \mu \hat{\lambda}^n \quad \text{stick;} \\[2ex]
\mu \left(\dfrac{\partial \hat{\lambda}^n}{\partial \lambda_{\mathcal{J}}} \dfrac{\hat{\lambda}_i^{t,tr}}{\left\| \hat{\boldsymbol{\lambda}}^{t,tr} \right\|} + \dfrac{\hat{\lambda}^n}{\left\| \hat{\boldsymbol{\lambda}}^{t,tr} \right\|} \left(\dfrac{\partial \hat{\lambda}_i^{t,tr}}{\partial \lambda_{\mathcal{J}}} - \dfrac{\hat{\lambda}_i^{t,tr}}{\left\| \hat{\boldsymbol{\lambda}}^{t,tr} \right\|^2} \dfrac{\partial \hat{\lambda}_\gamma^{t,tr}}{\partial \lambda_{\mathcal{J}}} \hat{\lambda}_\gamma^{t,tr} \right) \right), & \\
& \text{otherwise} \qquad \text{slip.}
\end{cases}
$$

$$\text{(B.96)}$$

Appendix C

Notes on the improved Hencky material

Since this appendix is focused on the (effective) stress-strain relationship and being both the considered tensors (namely, Kirchhoff stress and logarithmic strain) symmetric, *Voigt* notation is considered.

Based on the definition of the invariants used in Eq. (4.101) and the kinematic relationship Eq. (4.89), it is convenient to noticed that

$$\frac{\partial \epsilon_v}{\partial \boldsymbol{\epsilon}} = \boldsymbol{I}^{(2)}; \tag{C.1}$$

$$\frac{\partial \epsilon_q}{\partial \boldsymbol{\epsilon}} = \frac{2}{3\,\epsilon_q} \bar{\boldsymbol{e}}; \tag{C.2}$$

$$\frac{\partial n}{\partial \epsilon_v^e} = \frac{\partial n}{\partial \epsilon_v^p} = \frac{1 - n_0}{\exp\left(\epsilon_v\right)} = 1 - n, \tag{C.3}$$

with the quantity $\bar{e}$ being

$$\bar{\boldsymbol{e}} = \left[e_{xx}, \ e_{yy}, \ e_{zz}, \ 2\,e_{xy}, \ 2\,e_{xz}, \ 2\,e_{yz} \right]^T. \tag{C.4}$$

Generally speaking, when a symmetric second-order tensor $(\bullet)$ is denoted by $(\bar{\bullet})$, it implies that its components $[\bullet_{xy}, \ \bullet_{xz}, \ \bullet_{yz}]$ are doubled, as in the above case.

C.1 Stress computation for the elastic case

This section considers $\boldsymbol{F} = \boldsymbol{F}^e$. While, for this case, inequality 4.99c is trivially satisfied, Eq. (4.99a) gives

$$
\begin{aligned}
\boldsymbol{\tau}' &= 2 \frac{\partial \hat{\Psi}_{inc}^{sk}}{\partial \boldsymbol{b}^e} \boldsymbol{b}^e = \frac{\partial \hat{\Psi}_{inc}^{sk}}{\partial \boldsymbol{\epsilon}^e} \\
&= \frac{\tilde{K}}{2} \frac{\partial}{\partial \epsilon_v^e} \left(\frac{(\epsilon_v^e)^2}{n} \right) \boldsymbol{I}^{(2)} + \frac{3}{2} G \frac{\partial}{\partial \epsilon_q^e} \left((\epsilon_q^e)^2 \right) \frac{2}{3\,\epsilon_q^e} \bar{\boldsymbol{e}} \\
&= \underbrace{\frac{\tilde{K}\,\epsilon_v^e}{n} \left(1 + \frac{\epsilon_v^e}{2\,n}\,(n-1) \right)}_{:=p'} \boldsymbol{I}^{(2)} + \underbrace{2G\,\bar{\boldsymbol{e}}^e}_{:=\bar{\boldsymbol{s}}'}.
\end{aligned} \tag{C.5}
$$

Being the model non-linear in terms of strain, the incremental form of Eq. (C.5) is necessary to study the tangent elastic moduli, i.e.,

$$
d\boldsymbol{\tau}' = \underbrace{\frac{\partial^2 \hat{\Psi}_{inc}^{sk}}{\partial \boldsymbol{\epsilon}^e \otimes \partial \boldsymbol{\epsilon}^e}}_{:=\mathcal{D}^e} : d\boldsymbol{\epsilon}^e = \frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\epsilon}^e}
$$

$$
= \frac{\partial}{\partial \epsilon_v^e} \left(\frac{\tilde{K}\,\epsilon_v^e}{n} \left(1 + \frac{\epsilon_v^e}{2\,n}\,(n-1) \right) \right) \boldsymbol{I}^{(2)} \otimes \boldsymbol{I}^{(2)} + 2\,G\,\frac{\partial \bar{\boldsymbol{e}}^e}{\partial \boldsymbol{\epsilon}^e}
$$

$$
= \left(K^{e,\,tan}\,\boldsymbol{I}^{(2)} \otimes \boldsymbol{I}^{(2)} + 2\,G\,\boldsymbol{I}^{(4),\,dev} \right) : d\boldsymbol{\epsilon}^e, \tag{C.6}
$$

with

$$
K^{e,\,tan} := \frac{\tilde{K}}{2\,n^3} \left(n^2 \left((\epsilon_v^e)^2 + 4\,\epsilon_v^e + 2 \right) - \epsilon_v^e\,n\,(3\,\epsilon_v^e + 4) + 2\,(\epsilon_v^e)^2 \right), \tag{C.7}
$$

and

$$
\bar{\boldsymbol{I}}^{(4),\,dev} := \begin{bmatrix} 2/3 & -1/3 & -1/3 & 0 & 0 & 0 \\ -1/3 & 2/3 & -1/3 & 0 & 0 & 0 \\ -1/3 & -1/3 & 2/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}. \tag{C.8}
$$

From Eq. (C.7), the limits for the extremes values of the Eulerian porosity can be computed (see also Figure 4.5(c))

$$
\lim_{n \to 0^+} K^{e,\,tan} = +\infty; \tag{C.9}
$$

$$
\lim_{n \to 1^-} K^{e,\,tan} = \tilde{K}. \tag{C.10}
$$

The tangent Poisson's ratio can be computed via the tangent bulk modulus Eq. (C.7) and the constant shear modulus, giving

$$
\nu^{e,\,tan} := \frac{3K^{e,\,tan} - 2G}{2\,(3K^{e,\,tan} + G)}, \tag{C.11}
$$

whose limits are (in compliance with Figure 4.5(d))

$$
\lim_{n \to 0^+} \nu^{e,\,tan} = \frac{1}{2}; \tag{C.12}
$$

$$
\lim_{n \to 1^-} \nu^{e,\,tan} = \frac{3\tilde{K} - 2G}{2\,(3\tilde{K} + G)}. \tag{C.13}
$$

C.2   Stress computation for the elasto-plastic case

The more general scenario $\boldsymbol{F} = \boldsymbol{F}^e\,\boldsymbol{F}^p$ is assumed in this section. While the relationship between the yield function $\Phi^{\tau'}$ and the direction of the flow rule $\frac{\partial g^{\tau'}}{\partial \tau'}$ via $\mathcal{D}^{(sk)}$ according

to the principles of hyperplasticity is detailed in 4.5.1.2, this appendix focuses on the linearisation to compute the elasto-plastic subroutine when no further dissipation mechanism is taken into account, i.e., $\tilde{\Psi}_{inc}^{(sk)}(\boldsymbol{\alpha}) = 0$ in Eq. (4.101). The standard equation to implement this are given by the (time-discretised) decomposition in the logarithmic strain space retaining the small strain setting, and by the satisfaction of the yield function for the stress tensor, i.e.,

$$\begin{cases} \boldsymbol{\epsilon}^e - \boldsymbol{\epsilon}^{e,\,tr} + \Delta\gamma^p \dfrac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'} = \boldsymbol{0}; & \text{(C.14a)} \\[3mm] \Phi^{\tau'}(\boldsymbol{\tau}') = 0, & \text{(C.14b)} \end{cases}$$

where $\Delta\gamma^p \geq 0$ is the increment in the plastic multiplier, which, along with $\Phi^{\tau'}(\boldsymbol{\tau}') \leq 0$ and the compatibility condition $\Phi^{\tau'}\Delta\gamma^p = 0$, gives the classical Karush-Kuhn-Tucker (KKT) conditions. Linearisation of the above system of equations with respect to the unknown quantities (with the elastic trial strain an unknown in this case) gives

$$\begin{cases} d\boldsymbol{\epsilon}^e - d\boldsymbol{\epsilon}^{e,\,tr} + d\Delta\gamma^p \dfrac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'} + \Delta\gamma^p \dfrac{\partial^2 g^{\tau'}}{\partial \boldsymbol{\tau}' \otimes \partial \boldsymbol{\tau}'} : d\boldsymbol{\tau}' = \boldsymbol{0}; & \text{(C.15a)} \\[3mm] \dfrac{\partial \Phi^{\tau'}}{\partial \boldsymbol{\tau}'} : d\boldsymbol{\tau}' = 0. & \text{(C.15b)} \end{cases}$$

The increment of stress for the elasto-plastic case is given by

$$d\boldsymbol{\tau}' = \underbrace{\frac{\partial^2 \hat{\Psi}_{inc}^{sk}}{\partial \boldsymbol{\epsilon}^e \otimes \partial \boldsymbol{\epsilon}^e}}_{=\boldsymbol{\mathcal{D}}^e} : d\boldsymbol{\epsilon}^e + \underbrace{\frac{\partial^2 \hat{\Psi}_{inc}^{sk}}{\partial \epsilon_v^p \, \partial \boldsymbol{\epsilon}^e} \otimes \boldsymbol{I}^{(2)}}_{:=\boldsymbol{\mathcal{D}}^p} : \left(\underbrace{d\Delta\gamma^p \frac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'} + \Delta\gamma^p \frac{\partial^2 g^{\tau'}}{\partial \boldsymbol{\tau}' \otimes \partial \boldsymbol{\tau}'} : d\boldsymbol{\tau}'}_{=d\boldsymbol{\epsilon}^{e,\,tr} - d\boldsymbol{\epsilon}^e} \right), \quad \text{(C.16)}$$

where $\boldsymbol{\mathcal{D}}^e$ is given by Eq. (C.6), while $\boldsymbol{\mathcal{D}}^p$ is defined as

$$\boldsymbol{\mathcal{D}}^p := \frac{\partial^2 \hat{\Psi}_{inc}^{sk}}{\partial \epsilon_v^p \, \partial \boldsymbol{\epsilon}^e} \otimes \boldsymbol{I}^{(2)} = \underbrace{\frac{\tilde{K}}{2} \left(\frac{\epsilon_v^e}{n} \right)^2 (1-n) \left(1 + \frac{2}{n}(1-n) \right)}_{:=K^{p,\,tan}} \boldsymbol{I}^{(2)} \otimes \boldsymbol{I}^{(2)}. \quad \text{(C.17)}$$

Eq. (C.16) can be rearranged as

$$\left(\boldsymbol{I}^{4,\,sym} - \Delta\gamma^p \boldsymbol{\mathcal{D}}^p : \frac{\partial^2 g^{\tau'}}{\partial \boldsymbol{\tau}' \otimes \partial \boldsymbol{\tau}'} \right) : d\boldsymbol{\tau}' = \boldsymbol{\mathcal{D}}^e : d\boldsymbol{\epsilon}^e + d\Delta\gamma^p \boldsymbol{\mathcal{D}}^p : \frac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'}, \quad \text{(C.18)}$$

where it is convenient to define the above-appearing fourth-order tensor as

$$\text{Ϙ} := \left(\boldsymbol{I}^{4,\,sym} - \Delta\gamma^p \boldsymbol{\mathcal{D}}^p : \frac{\partial^2 g^{\tau'}}{\partial \boldsymbol{\tau}' \otimes \partial \boldsymbol{\tau}'} \right)^{-1}. \quad \text{(C.19)}$$

Hence, Eq. (C.18) can be concisely written as

$$d\boldsymbol{\tau}' = \text{Ϙ} : \boldsymbol{\mathcal{D}}^e : d\boldsymbol{\epsilon}^e + d\Delta\gamma^p \, \text{Ϙ} : \boldsymbol{\mathcal{D}}^p : \frac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'}. \quad \text{(C.20)}$$

Substitution of the above equation into Eqs. (C.14) results in

$$\begin{cases} d\boldsymbol{\epsilon}^e - d\boldsymbol{\epsilon}^{e,\,tr} + d\Delta\gamma^p \dfrac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'} \\[3mm] \qquad + \Delta\gamma^p \dfrac{\partial^2 g^{\tau'}}{\partial \boldsymbol{\tau}' \otimes \partial \boldsymbol{\tau}'} : \left(\text{Ϙ} : \boldsymbol{\mathcal{D}}^e : d\boldsymbol{\epsilon}^e + d\Delta\gamma^p \, \text{Ϙ} : \boldsymbol{\mathcal{D}}^p : \dfrac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'} \right) = \boldsymbol{0}; & \text{(C.21a)} \\[4mm] \dfrac{\partial \Phi^{\tau'}}{\partial \boldsymbol{\tau}'} : \left(\text{Ϙ} : \boldsymbol{\mathcal{D}}^e : d\boldsymbol{\epsilon}^e + d\Delta\gamma^p \, \text{Ϙ} : \boldsymbol{\mathcal{D}}^p : \dfrac{\partial g^{\tau'}}{\partial \boldsymbol{\tau}'} \right) = 0, & \text{(C.21b)} \end{cases}$$

which can be rearranged in a linear system

$$
\begin{bmatrix}
\boldsymbol{I}^{4,\,sym} + \Delta\gamma^p \frac{\partial^2 g^{\tau'}}{\partial\boldsymbol{\tau}'\otimes\partial\boldsymbol{\tau}'} : \Upsilon : \boldsymbol{\mathcal{D}}^e & \left(\boldsymbol{I}^{4,\,sym} + \Delta\gamma^p \frac{\partial^2 g^{\tau'}}{\partial\boldsymbol{\tau}'\otimes\partial\boldsymbol{\tau}'} : \Upsilon : \boldsymbol{\mathcal{D}}^p\right) : \frac{\partial g^{\tau'}}{\partial\boldsymbol{\tau}'} \\[2ex]
\frac{\partial\Phi^{\tau'}}{\partial\boldsymbol{\tau}'} : \Upsilon : \boldsymbol{\mathcal{D}}^e & \frac{\partial\Phi^{\tau'}}{\partial\boldsymbol{\tau}'} : \Upsilon : \boldsymbol{\mathcal{D}}^p : \frac{\partial g^{\tau'}}{\partial\boldsymbol{\tau}'}
\end{bmatrix}
\begin{bmatrix} d\boldsymbol{\epsilon}^e \\[2ex] d\Delta\gamma^p \end{bmatrix}
$$

$$
= \begin{bmatrix} d\boldsymbol{\epsilon}^{e,tr} \\[2ex] 0 \end{bmatrix}. \quad \text{(C.22)}
$$

To compute the tangent $\boldsymbol{\mathcal{D}}^{alg} := \frac{\partial\boldsymbol{\tau}'}{\partial\boldsymbol{\epsilon}^{e,\,tr}}$, it is necessary to re-write Eq. (C.22) explicitly including $d\boldsymbol{\tau}'$ in lieu of $d\boldsymbol{\epsilon}^e$ (i.e., using again Eq. (C.20)), which gives

$$
\begin{bmatrix}
(\boldsymbol{\mathcal{D}}^e)^{-1} : \Upsilon^{-1} + \Delta\gamma^p \frac{\partial^2 g^{\tau'}}{\partial\boldsymbol{\tau}'\otimes\partial\boldsymbol{\tau}'} & \left(\boldsymbol{I}^{4,\,sym} - (\boldsymbol{\mathcal{D}}^e)^{-1} : \boldsymbol{\mathcal{D}}^p\right) : \frac{\partial g^{\tau'}}{\partial\boldsymbol{\tau}'} \\[2ex]
\frac{\partial\Phi^{\tau'}}{\partial\boldsymbol{\tau}'} & 0
\end{bmatrix}
\begin{bmatrix} d\boldsymbol{\tau}' \\[2ex] d\Delta\gamma^p \end{bmatrix}
= \begin{bmatrix} d\boldsymbol{\epsilon}^{e,tr} \\[2ex] 0 \end{bmatrix}.
$$

$$\text{(C.23)}$$

The inversion of (C.23) leads to

$$
\begin{bmatrix} d\boldsymbol{\tau}' \\[2ex] d\Delta\gamma^p \end{bmatrix}
= \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\[2ex] \mathbf{D}_{21} & \mathbf{D}_{22} \end{bmatrix}
\begin{bmatrix} d\boldsymbol{\epsilon}^{e,tr} \\[2ex] 0 \end{bmatrix}, \quad\quad \text{(C.24)}
$$

where $\boldsymbol{\mathcal{D}}^{alg} = \frac{d\boldsymbol{\tau}'}{d\boldsymbol{\epsilon}^{e,tr}} = \mathbf{D}_{11}$.