Topology optimization in idealized-bolt contact, plasticity, and fracture mechanics using a level-set method

Jeet Samir Desai

Université de Paris

École doctorale de Sciences Mathématiques de Paris Centre, 386

Laboratoire Jacques-Louis Lions

Topology optimization in contact, plasticity, and fracture mechanics using a level-set method

Par Jeet DESAI

Thèse de doctorat de Mathématiques Appliquées

Dirigée par François Jouve
Et Grégoire Allaire

Présentée et soutenue publiquement le 28 septembre 2021

Devant un jury composé de :

Marc Dambrine, Professeur à l'Université de Pau et des Pays de l'Adour, Rapporteur
Francisco Periago, Professeur associé à Technical University of Cartagena, Rapporteur
Grégoire Allaire, Professeur à l'Ecole Polytechnique, Co-directeur de thèse
Samuel Amstutz, Maître de conférences-HDR à l'Ecole Polytechnique, Examinateur
Maya de Buhan, CR-CNRS, Université Paris Saclay, Examinateur
Yves Capdeboscq, Professeur à l'Université de Paris, Président du Jury
Andrei Constantinescu, DR-CNRS à l'Ecole Polytechnique, Examinateur
François Jouve, Professeur à l'Université de Paris, Directeur de thèse
Chiara Nardoni, Ingénieure chercheuse à l'IRT systemX, Invitée

To my parents Samir and Rupa, my sister Pooja and *FreeFEM*

Acknowledgements

णमो अरिहंताणं ।	Namo Arihantanam
णमो सिध्दाणं ।	Namo Siddhanam
णमो आयरियाणं ।	Namo Ayriyanam
णमो उवज्झयाणं ।	Namo Uvajjhayanam
णमो लोए सव्वसाहुणं ।	Namo Loye Savvasahunam

"I bow down to all souls, who have either conquered their passions, or are in the process of doing so."

I have come a long way, due to the support of a lot of people, to whom I am indebted. I thank Prof. Grégoire Allaire for giving me the opportunity to do this PhD. His constant guidance and motivation inspired me to do my best. Over the past three years, he has scrutinized each and every idea that I could think of. His ability to think critically of a concept is something that I look up to. I attribute my success, where I am and where I will be to him. Likewise, I thank Prof. François Jouve for guiding and supporting me all along the three years of this PhD. His deep insights into the numerical aspects were crucial in shaping my implementational skills. I am glad that I could absorb some of his years of consolidated expertise.

I sincerely thank the reviewers, Prof. Periago and Prof. Dambrine for sparing time and giving their valuable report on this manuscript. Their report not only improved the quality of this thesis but also opened perspectives for future works. I am grateful to Prof. Yves Capdeboscq for presiding over my thesis defense, along with jury examiners Prof. Andrei Constantinescu, Prof. Samuel Amstutz and Dr. Maya de Buhan.

My heartfelt thanks to my ex-supervisor and a friend, Georgios Michailidis, who collaborated with me and ignited my love for topology optimization. Special thanks to Prof. Frédéric Hecht, for having developed FreeFEM, which is the basis of all of the simulations of this thesis. I also thank him for organizing FreeFEM days and increasing the size of the FreeFEM community. All the high-performance computing of this thesis were possible due to the assistance of Pierre Jolivet and Pierre-Henri Tournier. I thank Prof. Blaise Bourdin and Prof. Antonin Chambolle for sharing their valuable insights into the fracture problem.

I find myself fortunate, to be surrounded by benevolent people at IRT systemX who helped me with every aspect of my PhD. I thank Chiara, who is like an elder sister now, for ensuring that every deadline is met. Given my nonexistent managerial skills, her presence ensured that I never end up in trouble. I am extremely grateful to Chetra, who taught me plasticity, fracture and helped me with code troubleshooting. Without his help, it would have taken me a few additional months to finish this PhD. I thank David and Ahmadali for their technical discussions on optimization, science and philosophical discussions on everything that possibly exists. I thank my boss Xavier, for constantly ensuring that this work finds relevance for the industrial problems. I thank Rim and Jakob, who arranged regular meetings, to ensure the smooth unwinding of this PhD and for helping me with my startup ventures. Thanks to Yannick, I could do useful research on the potential applications of this thesis and find applications in civil engineering. I sincerely thank Bruno for having connected me to several companies, which could potentially use this work. I thank the members of AMC team, Adrien, Romain, Mathieu, Henri for the infinite discussions and the weekly pot de croissant. I thank Pascal for having taken the initiative to ensure that every doctoral student at IRT is doing well by creating the DOCSX association.

The industrial partners of project TOP helped me in several ways. I thank Luca for being a mentor and advising me on a range of subjects, from coding to business and entrepreneurship. I thank Felipe for sharing his valuable knowledge on plasticity and fracture, and Patrick for sharing his understanding of plasticity and for finding relevant industrial test cases.

I thank Norah Jones for being such an exceptional singer. Her magical voice kept me motivated to write the entire thesis. I would also like to thank Jasper Newton Daniel, for making such amazing Whiskey. These two make a deadly combo and can help in calming the mind down in the most difficult situations.

The list of friends that I wish to thank is exhaustive. In particular, I wish to express my gratitude to thank all of the women in my life, who played a role, in some way or the other in shaping my beliefs and making my life, a better place. I thank Maria for helping me with everything, ranging from my startup ventures to presentations. I thank Yuliya for being there each and every day of the COVID crisis and cheering me up. I thank Tora for having to give me a different perspective on life and for having maintained contact for two years. I thank Diala for counteracting and analyzing my decisions, eventually saving me from detours. I am glad to have met Raha, whose positivism inspired me. I thank Luci, Iris and Heather for being such nice friends. I thank Aarzoo for sharing her deep insights on love. I thank Hanen, Laurence, Jyoti and Arushi with whom I spent memorable time. I thank the director of the Canadian residence, France for organizing events and for maintaining an incredible joyful environment during the COVID crisis. Thanks to her, I could restart sketching after six years. I thank Lalaina for the collaboration, that turned out fruitful.

During the three years spent at the cité universitaire, I met possibly hundreds of people. The friendship I could make would, by nature of this place, last a year, after which people leave for different places. Despite this flux, I am grateful that a few of them stayed: Vincent, Ajitesh, Salim, Zakari, Léo and Naji. There is something or the other, that I learnt from all of them. I thank Stefan, Jean François, Nathan and Sofiane for the memorable time spent at the Canadian residence. I thank Guillermo for having hosted me at his place and taking very good care of me.

It has been 4 years that I have lost physical contact with my Indian friends, but I am glad that our friendship is still as ripe as before. We have all scattered in different parts of the world, but by the virtue of quantum entanglement, we remain connected! I thank Vivek, Abhishek, Pratyaksh, Pragyan, Sourabh and Rajeev for the support on all fronts, from emotional to motivational. Our friendship is now 8 years old, and I am pretty confident that it will last at least a lifetime. I thank Sagar S., Aditya P., Srajan and Nimit for having maintained contact for so many years.

I thank my colleagues from CMAP, Alex, Abdelhak, Florian, Perle, Mathilde, Matias and Martin for the discussions and the annual get-together. In particular, I thank Alex for teaching me skiing and saving me from a potential accident. I thank Alexandra for being so quick in responding and helping with the administrative stuff.

I thank my guru and now a friend, Amol sir, for constantly motivating me to do something useful for the planet. Eventually, sensitizing me into finding a path that helps our nature. The decisions that I would take on my future ventures are attributed to him.

And last but not the least, I thank my family, for supporting my choices. I am grateful to my mother and sister, who continually showered me with their unconditional love and to my father, who guided me all along my life and shaped my current self.

Topology optimization in contact, plasticity, and fracture mechanics using a level-set method

Abstract

The main contribution of this thesis is the theoretical and the numerical study shape and topology optimization for nonlinear phenomena, like contact, plasticity, and fracture using the level-set method. One application of the contact boundary condition for an idealized-bolt model is also proposed. The governing equations of the three physics dealt with in this thesis: contact, plasticity and damage, are theoretically not shape-differentiable. In each case, we construct an approximation by penalization, regularization or a combination of the two. The approximations for contact and plasticity are shown to be well-posed and to admit solutions that converge to the exact solution. For each physics, the shape sensitivity analysis is performed on the approximate model and the resultant adjoint problem is shown to be well-posed under technical assumptions. The shape optimization is implemented numerically using a level-set method with body-fitted remeshing, which captures the boundary of the shapes while allowing for topology changes. Numerical results are presented in 2D and 3D. We also discuss high-performance computing for linear elasticity and for fracture model, and present a few 3D results.

Keywords

Shape and topology optimization, level-set method, damage model, contact model, simplified-bolt model, fracture mechanics, plasticity with isotropic and kinematic hardening

Optimisation topologique en mécanique du contact, de la plasticité, et de l'endommagement par une méthode de lignes de niveaux

Abstract

La contribution principale de cette thèse est l'étude théorique et numérique de l'optimisation topologique pour des phénomènes non-linéaires, comme le contact, la plasticité, et l'endommagement en utilisant la méthode des lignes de niveaux. Une application de la condition de contact dans un modèle de vis idéalisée est aussi proposée. Les trois modèles physiques traités – le contact, la plasticité et l'endommagement – ne sont pas différentiables théoriquement par rapport la forme. Pour chaque modèle nous construisons une approximation par pénalisation, régularisation, ou une combinaison des deux. Les problèmes approchés sont alors bien posés (sauf dans le cas de l'endommagement) et convergent vers les solutions des problèmes initiaux. Pour chaque physique, nous faisons une analyse de sensibilité sur le modèle approché et nous démontrons que le problème d'adjoint est bien posé sous quelques hypothèses. L'optimisation de forme est implémentée numériquement avec une méthode de ligne de niveaux, qui capte le bord des formes, en permettant les changements topologiques. Les résultats numériques sont présentés en 2D et en 3D. Nous abordons le calcul haute performance (HPC) pour l'élasticité linéaire et le modèle de fracture et nous présentons quelques résultats numériques.

Mots clés

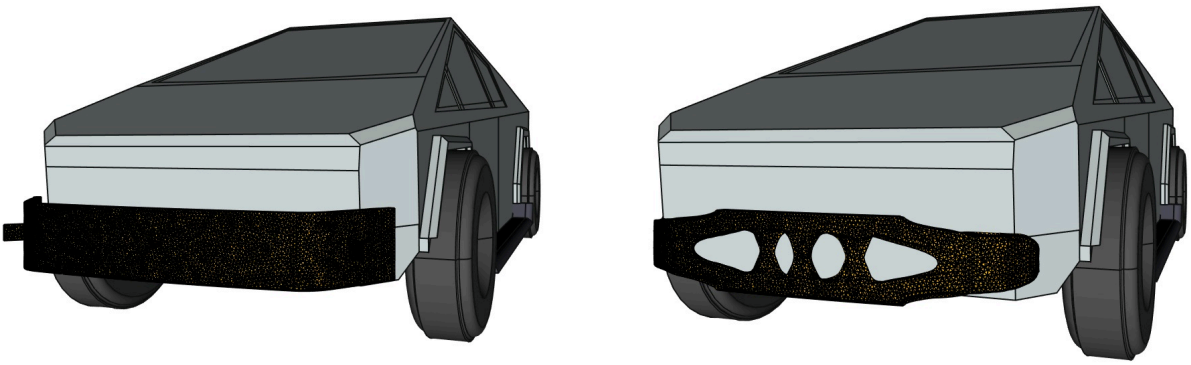
Optimisation géométrique et topologique, méthode des lignes de niveaux, modèle d'endommagement, modèle de contact, modèle de vis simplifié, mécanique de la rupture, plasticité avec l'écrouissage cinématique et isotrope

Résumé de thèse (en français)

L'objectif de cette thèse est d'étudier l'optimisation topologique pour des phénomènes non linéaires, comme le contact, la plasticité et la rupture en utilisant la méthode des level-sets.

L'optimisation topologique de forme est l'art de trouver une structure optimale. Cet art nécessite un moyen de décrire la forme et la topologie mathématiquement. Il existe deux méthodes connues pour représenter une forme sur un domaine fixe : la première, où la forme est approchée par une fonction de densité; la deuxième, où le bord d'une forme est défini implicitement par l'isovaleur zéro d'une fonction scalaire, appelée méthode des lignes de niveaux. D'une manière générale, la plupart des algorithmes d'optimisation topologique peuvent être classifiés en fonction de la densité ou de la ligne de niveau.

Il existe plusieurs logiciels commerciaux capables d'optimiser la topologie sous de nombreuses contraintes, cependant ils sont pour la plupart basés sur des méthodes de densité comme SIMP (Solid Isotropic Material with Penalization). Une forme optimale obtenue par une telle méthode est en effet une fonction de densité, qui a potentiellement des densités intermédiaires. Celles-ci sont difficiles à interpréter physiquement et une interprétation incorrecte peut conduire à une forme qui n'est plus optimale. Une méthode de ligne de niveau utilise également des densités intermédiaires si on utilise un maillage fixe, mais la zone où ces densités se produisent est beaucoup plus petite que dans le cas de SIMP, restreinte aux mailles qui contiennent la courbe de niveau 0.



(a) Maillage de pare-choc non optimisé (b) Pare-choc optimisé pour minimiser la compliance

Figure 0.1: Maillage d'un pare-choc optimisé par l'auteur pour Cybertruck (CAO pris de GradCAD)[1]

La méthode des lignes de niveau suscite un intérêt croissant et quelques logiciels commerciaux ont commencé à intégrer cette méthode. Un remède connu pour éviter les densités intermédiaires est l'approche de remaillage [7], qui capture la forme avec un maillage d'éléments finis entièrement recalculé. Un exemple d'une forme optimale obtenue par remaillage est montré dans la Fig.0.1. Comme on le voit sur cette figure, la forme finale obtenue a une forme lisse, une topologie assez simple et est apparemment facile à fabriquer

En 2018, un projet intitulé TOP (Topology Optimization Platform)[2]a été lancé à l'IRT systemX, en région parisienne, en collaboration avec Safran, Airbus, Renault et ESI Group (éditeur de logiciels). Ce projet d'une durée de quatre ans a fourni des bureaux d'études aux partenaires industriels susmentionnés. L'objectif de ce projet était de développer des outils d'optimisation topologique basés sur la méthode des lignes de niveaux, et de démontrer leur applicabilité sur des cas-test proposés par les partenaires industriels. Une plateforme de recherche et développement robuste, appelée PISCO, a été développée. La nouveauté de cette plate-forme étant l'utilisation de remaillage et la possibilité de brancher facilement n'importe quel logiciel commercial pour effectuer la simulation mécanique pour l'optimisation topologique.

Cette thèse a été initiée dans le cadre du projet TOP afin d'étendre les capacités de la plateforme, d'effectuer l'optimisation topologique de phénomènes non linéaires en utilisant la méthode des lignes de niveau avec remaillage. Ceci est pertinent car les applications industrielles réalistes exigent des modèles sophistiqués,

[1] https://www.irt-systemx.fr/en/projets/top/

qui sont généralement non linéaires. Par exemple, le contact entre deux pièces jointes mécaniquement, la déformation plastique ou permanente dans les structures metalliques, et l'endommagement et la rupture dans les structures en béton.

La recherche de Jeet Desai a été réalisée dans le cadre de l'IRT SystemX, Paris-Saclay, France. Elle est subventionnée par des fonds publics dans le cadre du "Programme d'Investissements d'Avenir".

Quelle est la nouveauté de cette thèse sur l'optimisation de la topologie en mécanique non linéaire ?

Les trois phénomènes non linéaires traités dans cette thèse – le contact, la plasticité et la rupture – sont, par nature, non réguliers. L'optimisation topologique pour ces phénomènes non réguliers existait dans la littérature scientifique, au moins depuis les deux dernières décennies.

La méthode de densité ainsi que la méthode de ligne de niveau a été appliqué à ces phénomènes. Les méthodes de densité étant plus anciennes et plus répandues, ont avancé plus rapidement. Il y a des résultats pour le contact [174, 111], la plasticité avec l'écrouissage [182, 103, 33, 116] et la mécanique de la rupture [96, 101, 96]. La méthode des lignes de niveau a été appliquée au contact avec remaillage [113, 114] et sans remaillage [129], à la plasticité parfaite [130] et à la fracture [186].

Nous listons les nouveautés de cette thèse par rapport aux travaux antérieurs dans les deux sous-sections suivantes.

Analyse de sensibilité pour le problème continue

La plupart des travaux précédents (mentionnés ci-dessus) discrétisent le problème non lisse en espace et en temps. La sensibilité de forme et le problème adjoint sont alors calculés pour une stratégie de discrétisation particulière. Les auteurs ignorent généralement les termes non dérivables dans le calcul de la sensibilité de forme et du problème adjoint.

Un tel problème non régulier ne peut admettre qu'une dérivée dite conique, dont l'analyse mathématique est très technique [171]. De plus, une telle dérivée n'a jamais été implémentée numériquement à notre connaissance. Dans cette thèse, nous contournons les problèmes de non différentiabilité en construisant un problème approché qui est différentiable par rapport à la forme et qui est facile à mettre en œuvre numériquement. On montre que l'approximation pour chaque phénomène est bien posée (sauf pour la fracture, qui est mal posée) et qu'elle admet une solution qui converge vers la solution exacte. De plus, l'équation adjointe (utilisée pour calculer la sensibilité à la forme) est bien posée pour chaque problème, sous certaines hypothèses. Aussi, une étude formelle de la dérivée de forme ainsi calculée est présentée lorsque les paramètres d'approximation tendent vers zéro. Connaître le problème de l'adjoint analytique présente plusieurs avantages.

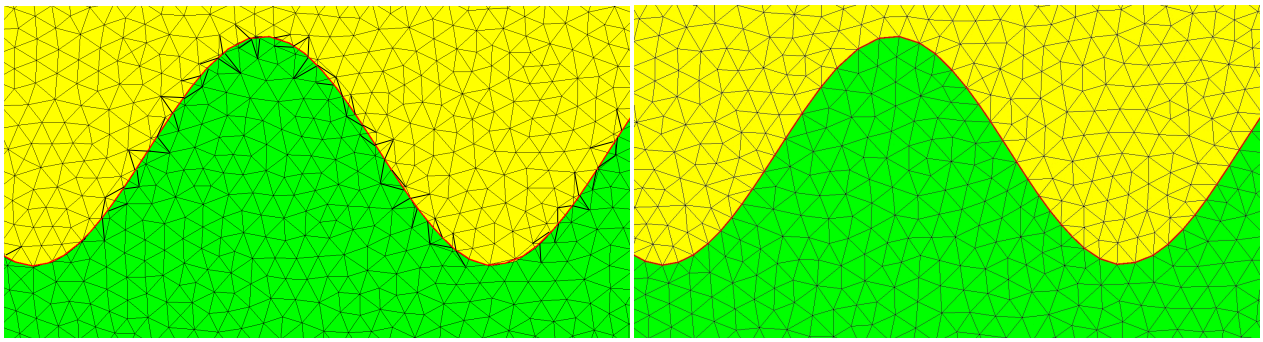
1. Si le problème adjoint est connu pour être bien posé, on est assuré de l'absence totale d'instabilités numériques lors de la résolution du problème adjoint
2. La connaissance des propriétés mathématiques du problème adjoint aide à choisir le solveur itératif pour la résolution numérique
3. On peut maintenant effectuer une analyse de sensibilité numérique pour une forme et une solution obtenues à partir d'un logiciel (boîte noire) commercial, dont la stratégie de discrétisation est a priori inconnue

Le dernier point permet même de créer une plateforme logicielle d'optimisation topologique, qui peut être branchée sur n'importe quel logiciel commercial.

Remaillage adapté au corps

D'après la littérature mentionnée ci-dessus, il existe très peu de travaux l'approche par remaillage global est utilisée. Les deux seules approches utilisées dans les travaux précédents : XFEM [180] et CutFEM [41] impliquent le remaillage uniquement de la frontière de la forme. Tandis que XFEM souffre d'un mauvais conditionnement de matrice de la rigidité et une solution polluée autour du bord, CutFEM y remédie en construisant des termes appelés "ghost penalty" afin de contrôler le conditionnement de la matrice. Au

[2] https://grabcad.com/library



(a) Maillage XFEM (b) Remaillage adapté au corps

Figure 0.2: Forme Ω (en verte) avec le bord $\partial\Omega$ (en rouge)

contraire, le remaillage adapté que nous utilisons implique de remailler l'espace de design entier (voir Fig.0.2 pour une comparaison), sans avoir besoin de construire et d'analyser des termes "ghost penalty" [61]. On obtient donc une meilleure qualité de maillage, une meilleure solution physique, une meilleure dérivée de forme et – on l'espère – une meilleure forme optimale.

Nous présentons à la fin de ce document des cas-test 3D qui utilisent le calcul haute performance ce qui nous permet de traiter de maillages très raffinés.

Résumé de tous les chapitres

Le chapitre 1 est une courte introduction.

Le chapitre 2 présente un résumé des méthodes d'optimisation topologique de formes et met l'accent sur la combinaison de l'analyse de sensibilité et de la représentation par lignes de niveaux pour capturer le bord d'une forme. La méthode avec suivi de frontière et remaillage, qui est utilisée pour l'ensemble des simulations numériques de cette thèse, est aussi présentée.

Dans le chapitre 3, nous étudions l'optimisation topologique en élasticité linéarisée avec condition de contact. Nous introduisons un modèle modèle de vis simplifié avec contact. La connexion mécanique entre deux solides par une vis est modélisée à l'aide d'un ressort. Le modèle de contact est incorporé par pénalisation, ce qui aboutit à un problème approché qui est différentiable par rapport à la forme. La dérivée de forme est calculée pour le problème approché en utilisant la méthode adjointe. Le problème adjoint est bien posée grâce à la régularisation-pénalisation proposée. Enfin, l'efficacité de la méthode de level-set pour le modèle de ressort-contact est illustrée numériquement par des exemples 2D et 3D.

Dans le chapitre 4, nous étudions l'optimisation topologique en plasticité quasi-statique avec écrouissage cinématique linéaire et isotrope linéaire. Nous considérons la formulation variationnelle primale du problème de plasticité. Cette formulation est régularisée et pénalisée afin d'obtenir un problème approché dérivable par rapport la forme. La dérivée de forme pour le problème approché est calculée en utilisant la méthode adjointe. Grâce à la pénalisation et à la régularisation proposées, la discrétisation temporelle du problème adjoint est bien posée. Nous comparons cette dérivée de forme obtenue pour le problème régularisé avec celle du problème d'origine calculée de manière formelle. Ces résultats sont implémentés et de nombreux résultats numériques 2D et 3D obtenus par la méthode des courbes de niveaux sont présentés.

Puis, au chapitre 5, nous étudions le problème de plasticité discrétisé en temps en utilisant une approche incrémentale. Cette approche facilite la conversion de l'inégalité de plasticité en égalité et se substitue à la pénalisation. La dérivée de formes est obtenue pour cette approche incrémentale mais elle n'est pas implémentée.

Dans le chapitre 6, nous proposons un algorithme d'optimisation topologique pour un modèle de fracture. Dans l'esprit de la théorie de Griffith, la fracture fragile est modélisée par le modèle énergétique de Francfort-Marigo, avec sa régularisation à la Ambrosio-Tortorelli, qui peut également être considérée comme un modèle d'endommagement de gradient. Ce modèle d'endommagement de gradient quasi-statique et irréversible est

approché en utilisant la pénalisation et la régularisation pour autoriser la différenciation de forme. La dérivée de forme est déterminée en utilisant la méthode adjointe. Le modèle est implémenté numériquement et de nombreux cas-test 2D et 3D sont proposés. La méthode s'avère très efficace pour concevoir des structures sans fissures.

Enfin, au chapitre 7, nous abordons le calcul haute performance (HPC) pour l'élasticité linéaire et le modèle de fracture. Les détails algorithmiques sont présentés et quelques cas de tests esthétiques 3D sont présentés. Pour le problème de fracture, une colonne 3D réaliste est optimisée.

Contents

Chapter 1

Introduction

Contents

1.1 Motivation

The objective of this thesis is to study shape and topology optimization for nonlinear phenomena, like contact, plasticity, and fracture using the level-set method.

By topology of a structure, we mean the number of holes in the structure in two-dimension and the Euler characteristic in three-dimension. Shape and topology optimization is the art of finding an optimal structure. This art necessitates a way to describe the shape or topology mathematically. There are two well-known methods to represent a shape over a fixed domain. First, where the shape is approximated using a scalar density function. Second, where the shape boundary is defined (explicitly or implicitly) using the zero isovalue of a scalar function, also known as a level-set function. Broadly speaking, all optimization algorithms can be classified into density based or boundary based.

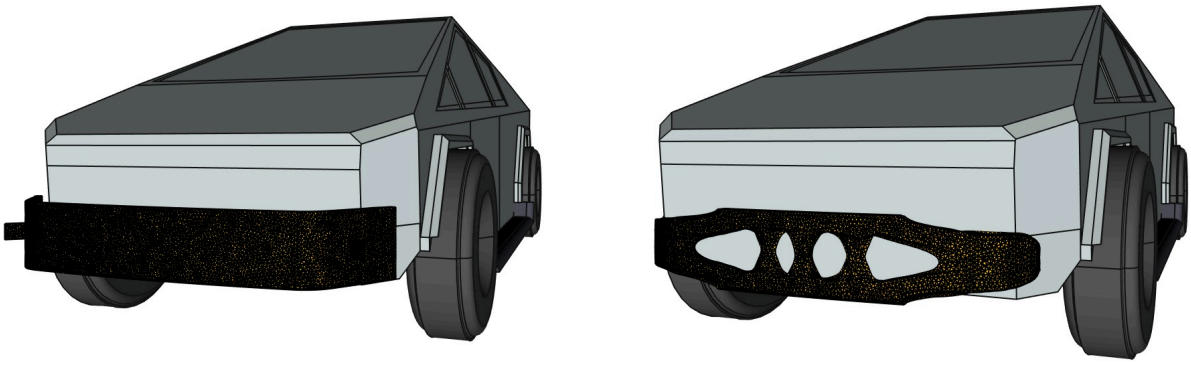
While there exists several commercial software, capable of performing topology optimization under many constraints, they are all typically based out of density-based methods, like SIMP (Solid Isotropic material with Penalization). An optimal shape obtained by such a method is indeed a density function, which potentially has intermediate densities. These are hard to interpret physically and incorrect interpretation can lead to a shape which is no longer optimal. A level-set method as well suffers from intermediate densities in case of a fixed mesh, although the zone where these densities occur is far smaller than in the case of SIMP.

The level-set method has been receiving growing interest and a few commercial software have started integrating this method. One time tested remedy to avoid the intermediate densities is body-fitted remeshing approach [7], which captures the shape with a fully recomputed finite element mesh. An example of an optimal shape obtained via body-fitted remeshing is shown in Fig.1.1. As seen in this figure, the final shape obtained has smooth boundary, fairly simple topology and is seemingly easy to fabricate.

In 2018, a project called TOP (Topology Optimization Platform)[2]was launched at IRT systemX, in Paris area, in collaboration with Safran, Airbus, Renault and ESI Group (software editor). This four-year long project provided industrial research departments to the aforementioned industrial partners. The goal of this project was to develop level-set based topology optimization tools, and demonstrate their industrial applicability on test cases proposed by the industrial partners. A robust research and development software platform, called PISCO, was developed. The novelty of this platform being the use of body-fitted remeshing and capability to plug any commercial software to perform the mechanical simulation for optimization.

This thesis is a consequence of a PhD initiated in the framework of project TOP. This PhD so-launched envisaged to extend the capabilities of the software platform, to perform topology optimization for nonlinear

[0]https://www.irt-systemx.fr/en/projets/top/



(a) Unoptimized bumper mesh

(b) Bumper optimized for minimal compliance

Figure 1.1: Mesh of bumper optimized by the author for Cybertruck (CAD taken from GradCAD)[1]

phenomena using the level-set method and remeshing. This is relevant as realistic industrial applications demand sophisticated models, that are typically nonlinear. For example, the contact between two mechanically joint parts, the plastic or permanent deformation in shafts, and the damage and fracture in concrete structures.

The research of J. Desai has been carried out in the framework of IRT SystemX, Paris-Saclay, France, and therefore granted with public funds within the scope of "Programme d'Investissements d'Avenir".

1.2 What is new about this PhD on topology optimization in nonlinear mechanics?

The three nonlinear phenomena treated in this thesis: contact, plasticity and fracture are, by nature, non-smooth. Topology optimization for these non-smooth phenomena has existed in the scientific literature, at least for the past two decades.

Both density and level-set methods have been applied to tackle with these phenomena. Density based methods being the older method, have taken the lead. They have been applied to contact [174, 111], plasticity with hardening [182, 103, 33, 116] and fracture mechanics [96, 101, 96]. Likewise, the level-set based methods have been applied to contact with remeshing [113, 114] and without remeshing [129], perfect plasticity [130] and fracture [186].

Then the question that arises is, what is new about this work? We address this question in the following two subsections.

1.2.1 Shape sensitivity analysis for the continuous problem

Most of the previous works (mentioned above) discretize the non-smooth problem in space and time. The shape sensitivity and the adjoint problem are computed for the specific discretization strategy. Authors typically ignore the non-differentiable terms in the computation of the shape sensitivity and the adjoint problem. While this approach might seem to give results in the literature, it is not rigorously justified as the non-differentiable terms in the theoretical problem were ignored.

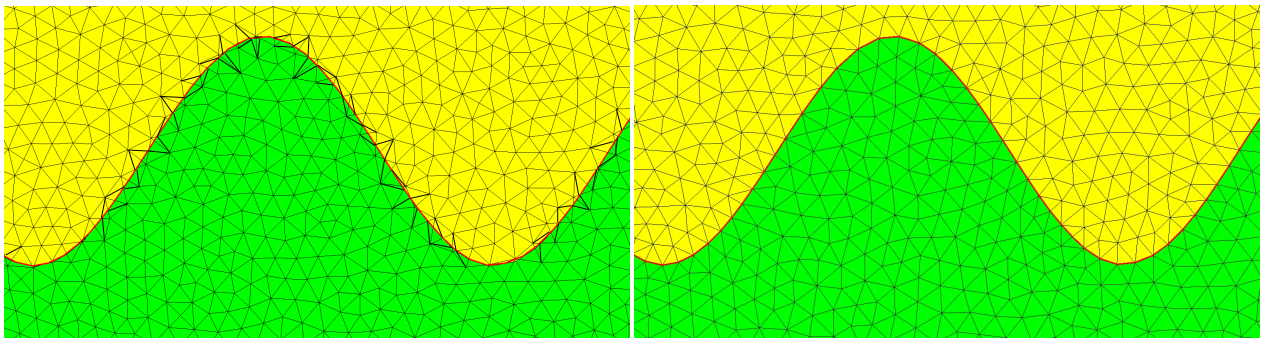
Such a non-smooth problem can admit only a so-called conical derivative, the mathematical analysis of which is very technical [171]. Moreover, such a derivative has not been implemented numerically to the best of our knowledge. In this thesis, we circumvent the non-differentiability issues by constructing an approximate problem that is shape-differentiable and is easy to implement, numerically. The approximation for each phenomenon is shown to be well-posed (except for fracture, which is ill-posed) and to admit a solution

[2] https://grabcad.com/library

that converges to the exact solution. Moreover, the adjoint equation (used to compute shape-sensitivity) is shown to be well-posed for each problem, under suitable technical assumptions. Also, a formal study of the shape derivative so-calculated is presented when the approximation parameters tend to zero. Knowing the analytical adjoint problem has several advantages.

1. If the adjoint problem is known to be well-posed, one is assured of complete absence of numerical instabilities while solving the adjoint problem
2. The knowledge of the mathematical properties of the adjoint problem helps in choosing the iterative solver for the numerical resolution
3. One can now perform a sensitivity analysis numerically for a shape and a solution obtained from a commercial blackbox software, whose discretization strategy is a priori unknown

The last point even permits us to create a software platform for topology optimization, which can be plugged to any commercial software.



(a) XFEM mesh (b) Body-fitted remesh

Figure 1.2: Shape Ω (in green) with boundary $\partial\Omega$ (in red)

1.2.2 Body-fitted remeshing

From the above mentioned literature, there are very few works where some kind of remeshing approach is used. The only two approaches used in previous works: XFEM [180] and CutFEM [41] involve remeshing only the boundary of the shape. While XFEM suffers from ill-conditioning of stiffness matrix and a polluted solution around the shape boundaries, CutFEM remedies it by constructing additional ghost penalty terms so as to control the conditioning of the stiffness matrix. On the contrary, the body-fitted remeshing approach (which we resort to) involves remeshing the entire design space (see Fig.1.2 for comparison), with no need to construct and analyze additional ghost penalty terms [61]. Hence, better mesh quality, better physical solution, better shape derivative and eventually a better shape.

Moreover, we present 3D test cases and use high performance computing for a few test cases, permitting us to deal with large meshes. In all, the above two points (in the two subsections) are addressed in the context of the three non-smooth phenomena (contact, plasticity and fracture) in this thesis. The details are summarized in the following section.

1.3 Summary of the chapters

We now present a summary of all chapters. A few aspects remain common to all chapters: Céa's technique [53] (based on the adjoint method) is used to compute the shape derivative, the adjoint problem is analyzed and is shown to be well-posed, the numerical simulations are performed using an open source software `FreeFEM` [92] and the remeshing is performed using an open source tool, `MMG` [7].

Chapter 2: State of the art

This chapter aims at providing a literature review on the existing methods of shape optimization, while delineating the level-set based method. The density based methods are discussed and compared to the level-

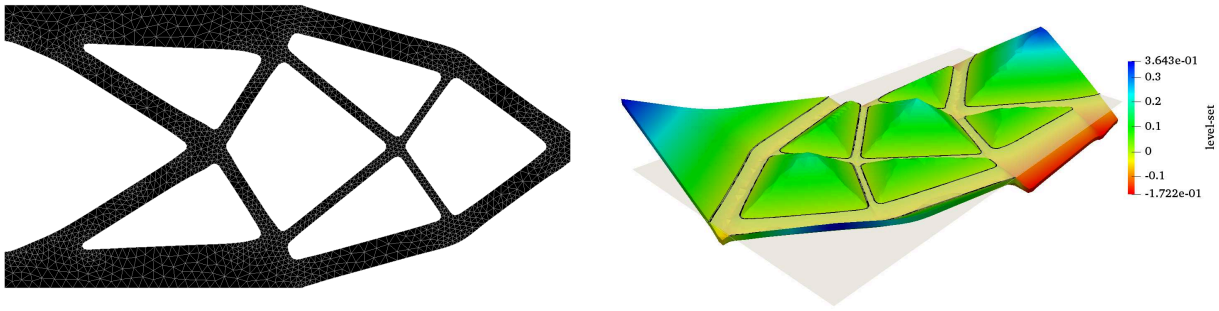


Figure 1.3: 2D cantilever beam, optimized for linear elasticity (left), corresponding level-set (right)

set based methods. The Hadamard's method for defining a shape derivative is recalled. The Céa's method is presented for the sensitivity analysis for linear elasticity and the shape derivative is computed.

The level-set method is then presented. As mentioned earlier, the level-set method is a technique to capture shape using the zero isovalue of a scalar function, also known as a level-set function. For example, for a shape Ω conceived in a design space D, the level-set function ϕ is given by

$$\begin{cases} \phi(\boldsymbol{x}) < 0 & \text{if } \boldsymbol{x} \in \Omega, \\ \phi(\boldsymbol{x}) = 0 & \text{if } \boldsymbol{x} \in \partial\Omega, \\ \phi(\boldsymbol{x}) > 0 & \text{if } \boldsymbol{x} \in D \backslash \Omega \end{cases}$$

A graphical example is given in Fig.1.3, where we show the level-set function of an optimized 2D cantilever beam (along with its body-fitted mesh). We then discuss the level-set method in the context of shape optimization following the approach in [12]. The ersatz material method is briefly recalled and compared to remeshing approaches like XFEM [180] and CutFEM approaches [41]. The body-fitted remeshing approach is then presented, along with its implementation using `MMG` [61].

We then present the topology optimization algorithm used in this thesis, which involves a simple gradient descent algorithm and a Lagrange-mulitplier method. Finally a few 3D numerical results are presented. In the results section, we compare a cantilever obtained via ersatz material method, which contains bar-like features versus the one obtained via remeshing, which has plate-like features. Here, pointing out the striking differences in the optimal shapes obtained, we advocate for the remeshing approach.

Chapter 3: Contact mechanics

This chapter is devoted to the study of topology optimization for contact mechanics, along with its application in an idealized-bolt setting. It is an advancement of an earlier work [128], where the penalized contact problem was treated for the level-set method. The contribution of this chapter is the extension of the same work in an idealized bolt setting and the usage of remeshing for problems involving rigid contact boundaries.

The contact model considered here involves linear elasticity and nonlinear contact boundary conditions. The contact boundary conditions include the non-penetration equations of Signorini and the Coulomb's friction law. The contact problem put in the variational form renders itself as an inequality, that is well-posed if the friction coefficient is small [71]. In order to render the problem shape-differentiable, the no penetration conditions are penalized, and the friction boundary condition is regularized [129, 71]. This furnishes a variational equation, amenable to shape differentiation.

The shape derivative computation is then presented, which involves an adjoint equation. Since the adjoint equation corresponds to the approximate problem, a formal limit analysis of the adjoint equation is also presented. In the numerical section, we discuss the implementation of the contact boundary conditions on `FreeFEM` [92]. We explain how the contact boundary conditions are numerically integrated, along with a few test cases in 2D and 3D. For the topology optimization using level-set method, the ersatz material approach (see Subsection 2.2.1) as well as the remeshing approach is used. Then, in the results section, a few interesting 2D and 3D test cases are presented. The contact zone is prescribed to be included in a fixed and given surface. We optimize a 3D cube, a chair (see Fig.1.4) and a gripper.

In the last section of this chapter, we explain how this contact model can be extended in an idealized-bolt setting and subjected to topology optimization. The mechanical connections in a structure using bolts are hard to model physically. One possibility is to simplify the connection by assuming that it acts like a spring

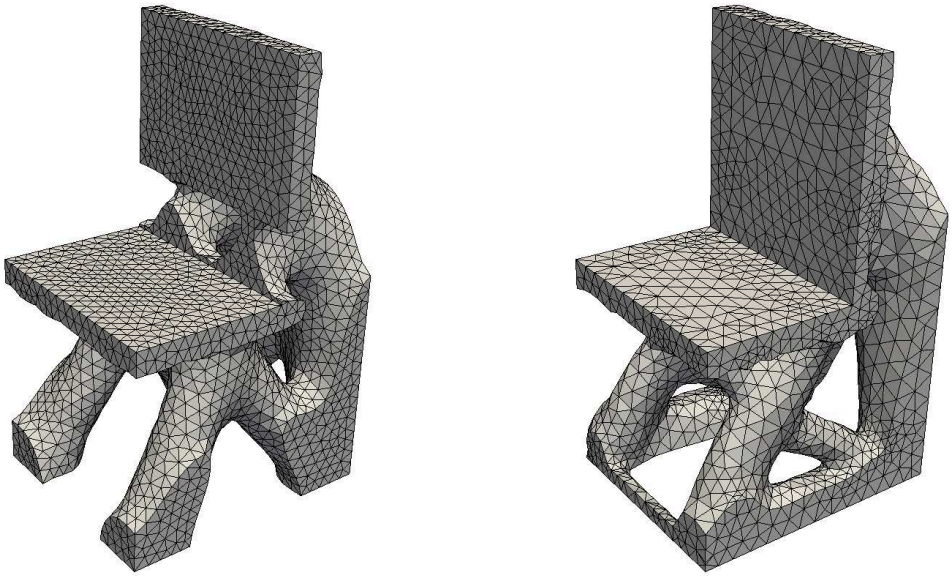


Figure 1.4: Chair optimized for linear elasticity, without contact boundary conditions (left), with contact boundary conditions (right)

[151]. The two structures are assumed to satisfy the linear elasticity model and the contact between the two structures is taken into account via the contact boundary conditions. In addition, the bolt is assumed to be given a specified pretension. The model is written in the variational inequality form and is shown to be well-posed. The same inequality is treated as before and the shape derivative is computed.

Finally, a cube with a crack at the center is optimized for the idealized bolt-contact model. Different topologies obtained by altering the bolt pretension are presented. Coupled optimization, where the topology and bolt position optimized are also presented.

Chapter 4: Plasticity

Here, quasi-static plasticity with linear kinematic and isotropic hardening is studied for topology optimization using the level-set method.

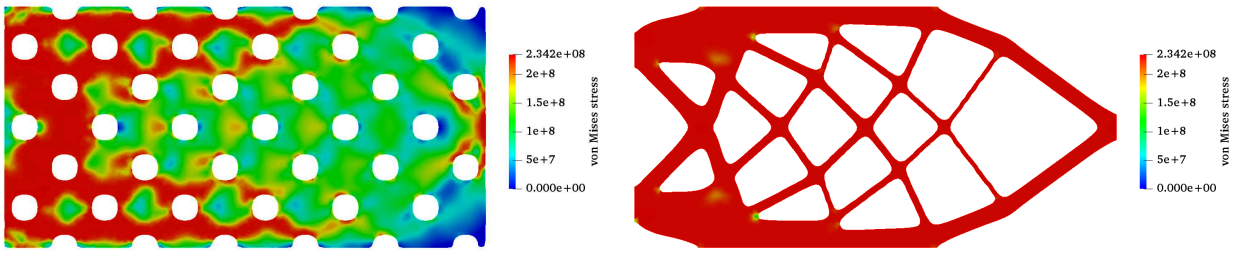
The plasticity model governs the permanent deformation in structures, which occur when they undergo a stress exceeding their yield limit. This happens often in mechanical parts subjected to large-magnitude forces, for instance, shafts or piston in engines. The quasi-static assumption is made when the evolution of the external forces is slow.

The yield limit is a surface in the stress plane for the three-dimensional problem. The von Mises yield limit is considered in this chapter. An associative flow rule, which which assumes that the plastic flow variables are proportional to the normal to the yield surface, is considered [162]. The plasticity model is presented in the primal and in the dual forms for the associative flow rule. The primal form of plasticity is treated in this chapter. The same written using a variational formulation, renders itself as a variational inequality [86].

The variational inequality is not shape-differentiable and we approximate it using regularization and penalization, to render it shape differentiable. The approximation is shown to be well-posed and to admit a unique solution, which converges to the exact solution.

Then the shape derivative is determined for the approximate model using the Céa's method and the adjoint problem is shown to be well-posed. To understand what does the adjoint problem correspond to, as the regularization and penalization parameters tend to zero, a formal limit analysis is presented. Then, the implementation details are outlined for the numerical resolution of the approximate model and the adjoint equation. Since typical commercial packages use radial return algorithm to solve the primal formulation, we do the same here. Though we compute the adjoint problem for the approximate plasticity model, we also state how this adjoint problem can be used in the exact case (see Remark 12). We also compare the solution to the approximate model to the one obtained via the radial return algorithm.

Here, we point out at the novelty of our work:



(a) Initial shape (b) Shape optimized for compliance

Figure 1.5: 2D cantilever beam optimization under plasticity; red zones undergo plastic deformation

- In existing literature, the plasticity problem is discretized using a certain implicit scheme (and subjected to the radial-return algorithm). Then exactly for this discretization, the shape derivative and the adjoint problem are computed. If one changes the discretization scheme, one must recompute the sensitivities and the adjoint problem analytically. In our work, the shape sensitivity and the adjoint problem are independent of the discretization scheme.

- Most of the existing works use density methods like SIMP, where the material properties used in the plasticity model are approximated using the density function. There are no error estimates for such approximations. We on the other hand, give a precise estimate of the error arising from penalization. Since we use body-fitted remeshing, there are no intermediate densities and mechanical simulation is precise.

Finally, we present a few numerical results in 2D and 3D. The important conclusion of this chapter is that the shapes obtained for plasticity are indeed very similar to ones obtained for linear elasticity. When one minimizes the compliance of the structure, one naturally expects the optimal structure to undergo less plastic deformation. But in reality, the optimal shapes are the ones in which the plastic flow occurs everywhere (for example, see Fig.1.5). This helps the structure as plastic flow is followed by hardening, which causes an increase in the yield limit. This increase renders the structure more resistant to an increasing external loading.

Chapter 5: Plasticity, Incremental approach

The previous chapter deals with plasticity model which deals only with associative flow law and the von Mises yield limit. In order to determine a shape derivative for a more general model, we consider the time-discretized version of the plasticity problem and apply the incremental approach [154]. The presentation in this chapter is given only for the von Mises criteria with a associated flow-rule, but it can easily be extended to a general criteria yield criterion with a non-associated flow rule. In the incremental approach, a projection operator is used and one ends up with a variational equation, instead of a variational inequality (like in the case of primal formulation). Such a variational equation is easily amenable to shape-differentiation using Céa's technique. Only in this chapter, we do not present any numerical results.

Chapter 6: Fracture

In this chapter, we treat quasi-static brittle fracture for topology optimization.

Brittle fracture implies fracture without change in material properties, very often seen in glass and polymers. The brittle fracture model is governed by the three laws of Griffith and can be written in a variational form using the Francfort-Marigo energy (the total mechanical energy)[77]. The Francfort-Marigo energy is approximated using Ambrosio-Tortorelli regularization. The regularization problem is known to be ill-posed and to admit an infinity of solutions. The same regularization can also be seen as a damage gradient model, which governs progressive damage in structures [148]. If this progressive damage leads to a zero stiffness, fracture is assumed to occur. Such a damage model governs fracture in quasi-brittle materials, such as concrete.

Figure 1.6: Intermediate shapes encountered while finding a crack-free optimal column (black zones represent cracks)

The Ambrosio-Tortorelli regularization can also be viewed as a phase field approach [16]. We consider this regularization model in this thesis. This model demands the computation of two variables: the displacement vector and the damage variable. The damage variable measures the extent of degradation of the material stiffness and is assumed to satisfy the irreversibility constraint, i.e., damage cannot be reversed. The fracture model is derived by applying the optimality conditions on the approximate Francfort-Marigo energy, which results in a variational inequality. The computation of the shape derivative for this inequality, while taking the irreversibility condition into account is difficult.

As usual, we overcome this difficulty by approximating the inequality using penalization. We can show that the penalized model admits at least one solution and no uniqueness can be demonstrated. The shape derivative is then computed using Céa's technique. The adjoint equation is determined. For the computation of the adjoint equation, we have to assume that the damage variable evolves smoothly in time, which is a very strong hypothesis. The adjoint equation is hard to analyze. However, for a time-discretized version, we can show the well-posedness of the adjoint equation (under certain hypothesis). We then discuss the numerical implementation aspect of the damage problem, resolved using a backtracking algorithm [34]. The algorithm for shape optimization is also outlined. Finally, we show several numerical results in 2D and in 3D.

From the numerical perspective, this is one of the hardest non-smooth phenomenon to deal with for shape

optimization (compared to contact and plasticity). This is because slight changes in shape or loading can cause the appearance or disappearance of cracks, and hence violent oscillations in the objective function being minimized. Despite this, we show that the our proposed optimization algorithm is effective in conceiving crack-free optimal structures (see Fig. 1.6 for example).

Other than use of body-fitted remeshing, the novelty of this work is attributed to the following:

- Analysis of the adjoint equation, which despite being based out of an indefinite operator, is shown to be well-posed (under some technical assumptions)
- Unlike the previous works, which focus on damage minimization or crack direction control, we focus on crack elimination
- Numerical resolution of the fracture problem using backtracking [34], which potentially escapes local minima and results in cracks more often. This is results in the higher stability of the optimal shape with respect to shape perturbations (see Remark 26).

Chapter 7: High performance computing

In this chapter, we discuss the implementation of the topology optimization using parallel computing on a multi-core CPU. We consider the linear elasticity and the fracture model. The use of HPC (High performance computing) allows for a simulation of shapes meshed with up to $400,000$ vertices. We talk briefly about the need of parallel computing, and compare multi-core CPU to GPU.

We then give the implementation details of HPC using an open source package `PETSc` [1], which has been integrated to `FreeFEM` [99]. We then present a few interesting 3D test cases for linear elasticity, a cantilever, a wedge, a bridge and a column. While remeshing results in slight oscillations in the objective function (being minimized for the topology optimization), resolution on a highly refined mesh (using HPC) results a convergence that is much smoother. Finally, we present a 3D column optimized for the fracture model.

Scientific Communications

The work presented in this thesis has been documented in three articles, submitted for publication:

1. L.Rakotondrainibe, J.Desai, P.Orval, G.Allaire. *Coupled topology optimization of structure and connections for bolted mechanical systems* [152]
2. J.Desai, G.Allaire, F.Jouve, C.Mang. *Topology optimization in quasi-static plasticity with hardening using a levet-set method* [67]
3. J.Desai, G.Allaire, F.Jouve. *Topology optimization of structures undergoing fracture* [66]

Chapter 2

State of the art

Contents

Optimization is the minimization of a mathematical criterion. Shape optimization is finding a shape that minimizes the criterion.

A mechanical structure is an assemblage of materials, intended to sustain or to resist loads [82]. Structural optimization is a type of shape optimization where the shape is a mechanical structure, governed by physical laws. Structural optimization as a tool, is inevitable if one seeks to conceive a structure with desired mechanical properties.

Traditional structural design techniques involve a lot of heuristics and empirical methods. These techniques would be iterative, where the engineer would make a design, analyze it using empirical formulas, prototype, test and repeat, until a structure with desired mechanical properties was determined. This results in a high design cost and an improvement of the structure, which is totally limited to the experience and the expertise of the engineer. Moreover, for complex phenomena like plastic deformations and crack formations, the intuition of an engineer is limited. Here, the engineer has to additionally perform several tests on the design specimen to check for plastic deformations and crack formations, increasing the costs of design conception, further.

With the advent of computers with powerful processors, and the progress of the finite element method, structural design process underwent a revolution. The structural analysis could now be performed for very complex shapes on a computer. The prototyping and testing of every design was no longer needed. Coupling structural analysis to the structural optimization algorithms, the process of design, analysis and testing could be automated. This lead to a dramatic change in the structural design process in the past few decades. Not only did the design process become faster and cheaper, newer shapes beyond the imagination of the engineer were conceived.

The ultimate industrial goal of structural optimization is to determine structures that are resistant and light-weight, hence less energy-consuming. From an industrial perspective, a few relevant criteria in structural

optimization are compliance and maximum stress. Today, structural optimization finds application not only in the design of mechanical components used in automobiles, aeroplanes, turbines, etc., but also in the design of civil structures like bridges, buildings, etc.

In this chapter, we first recall the state of the art on topology optimization in Subsection 2.1.1. We then present the mathematical formulation of the shape optimization problem and briefly discuss methods to determine the shape derivative in Subsection 2.1.2. In Subsection 2.1.3, we present the linear elasticity model and derive the shape derivative using Céa's method. In Section 2.2, we present the level-set method and explain its implementation using body-fitted remeshing. The detailed version of the content in the Section 2.2 can be found in [12, 7, 39]. Finally, in section 2.3, a few 3D numerical results are presented.

2.1 Shape optimization

2.1.1 Existing methods

Depending on the number of parameters involved in the structural optimization, it can be broadly divided into three categories,

1. **Parametric optimization**: The shape is parametrized using a limited number of variables, for instance, the "thickness" of a structure [147], or the truss configuration [81]. Given the simplicity of parametric optimization, it has a wide variety of industrial applications

2. **Geometric shape optimization**: The shape is parametrized using the boundary of the structure. This parametrization can be brought about using the mesh boundary [14], splines [156], NURBS (non-uniform rational basis splines) [52], etc. This method is very robust, except, it cannot introduce topological changes.

3. **Topology optimization**: By topology of a structure, we mean the number of holes in the structure. Two structures have the same topology, if one structure can be warped smoothly into another without changing its number of holes. In topology optimization, the topology is defined by a density function [32], a level-set [8] or a characteristic function. The function is then evolved using a heuristic approach or a shape derivative, and the topology of the structure is optimized.

Evidently, there is gradual rise in the number of optimization parameters as we move from parametric optimization to topology optimization. Topology optimization is what we deal with in this thesis.

A holistic review on the several topology optimization methods can be found in [161]. Any topology optimization method must facilitate a "way" to define the topology. There are several ways, which exist in literature. The oldest known method is using the homogenization method [4, 30, 90, 91]. This method results in an optimal shape that is a density function taking values between zero and one. For instance, a 2D cantilever beam (optimized for compliance) obtained via this method is plotted in Fig.2.1. As can be seen from the image, there is no clear description of a boundary and there are large gray zones (the region where the optimized density function takes a value between 0 and 1).

In order to overcome this drawback of homogenization, one resorts to a penalization approach, where the density function (representing the shape) is forced to takes binary values, 0 or 1. One of the most successful penalization approach is the so-called SIMP (solid isotropic material with penalization) [31, 32]. In this approach, the density of the structure is raised to an exponent p. The value of this exponent is chosen heuristically, typically 3. The previous 2D cantilever beam obtained using SIMP is plotted in Fig.2.1. As seen in the figure, the density function seems to have lesser gray zones. This marks the success of SIMP, making it one of the most widely used methods in literature and in commercial packages like Optistruct and ANSYS. However, one very well-known drawback of SIMP is the fact that there are still a few gray zones in the optimized density function. These gray zones are hard to interpret from the fabrication point of view. Incorrect interpretation can result in loss of optimality of the optimized density function. There are several ways to avoid these gray zones, the most famous being the density-filter developed by O. Sigmund [160]. Another possibility is to project the density function and eliminate the gray zones [183].

ESO (evolutionary structural optimization) is a heuristic method, where one removes material from low stressed regions in the structure [188]. BESO (bi-evolutionary structural optimization) is a modification of ESO, where one can in addition add material wherever needed [94] in a fixed design space. The boundary of the structure is clearly defined as there are no density functions. These two methods are very easy to implement and known to work well for compliance minimization. These methods can be combined to XFEM

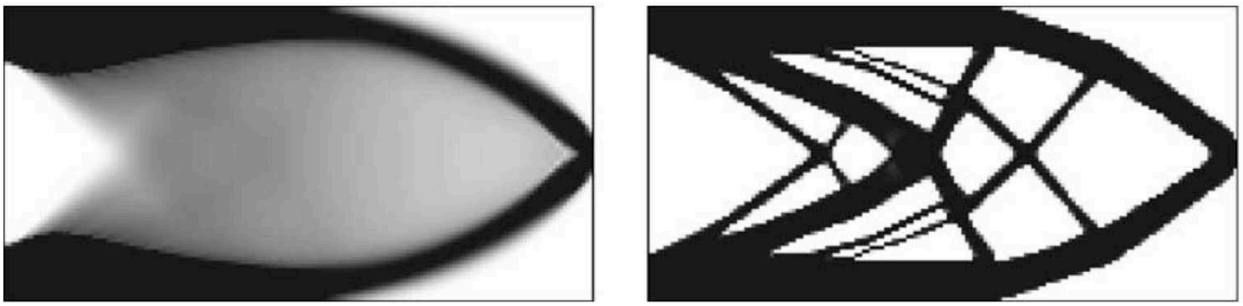


Figure 2.1: Cantilever computed with the homogenization method. Composite (left) and penalized solution (right), taken from [12]

(extended finite element method) approach to remesh the structure at every shape optimization iteration [126, 118].

One strength of ESO/BESO is the fact that it can introduce holes in the structure and hence bring about topological changes quite easily. But the two methods are heuristic. A mathematically rigorous way of introducing holes in the structure is using the so-called topological derivative [169]. In this approach, one computes the topological derivative everywhere in the structure and determines the most feasible location of adding an infinitesimal hole. The topological derivative can then be considered in conjunction with the shape derivative of the structure to optimize the topology of the structure [6].

Another method to deal with topological changes is using the level-set function. The level-set function is a continuous function, whose zero isovalue defines the shape (see (2.16) for a precise definition). One can then work with this level-set function, instead of dealing only with the boundary of the shape like in typical shape optimization approaches. This facilitates changes in topology. The level-set method was integrated into the shape optimization framework in [157, 145, 12, 184], and developed over the years, finding several applications such as robust design [65, 127], manufacturing constraints [11], composite design [9], etc. In the method proposed by G. Allaire et al [12] and M. Wang [184], the evolution of the level-set function is brought about by solving the transport equation (see Eqn.(2.17)). By doing so, the zero isovalue of the level-set function (representing the boundary of the structure) moves and the creation of holes is facilitated. Another way to bring about the evolution of the level-set function is by solving the convection-diffusion equation, proposed by T.Yamada [189]. Using this approach, one can create holes inside the structure without waiting for the boundary to move and collapse, unlike in the case of transport equation. The level-set method can also be combined with a topological derivative to deal with internal porosity constraints [125].

Most of the above mentioned works deploy a gradient descent method for the shape optimization algorithm. Gradient descent is a first order method and the rate of convergence is linear, hence slow. One possible method to converge at a faster rate if the second derivative, or the Hessian is computed. The Hessian in the context of shape optimization has been shown to be stable [58] and has been implemented using the level-set method [179].

We now present the mathematical formulation of the shape optimization problem. The structure or the shape is represented by an open set $\Omega \subset \mathbb{R}^d$ ($d = 2$ or 3 being the dimension of the problem). The structure has a boundary $\partial\Omega = \Gamma_N \cup \Gamma_D \cup \Gamma$, such that it is fixed on Γ_D and loaded on Γ_N as shown in Fig.2.2.

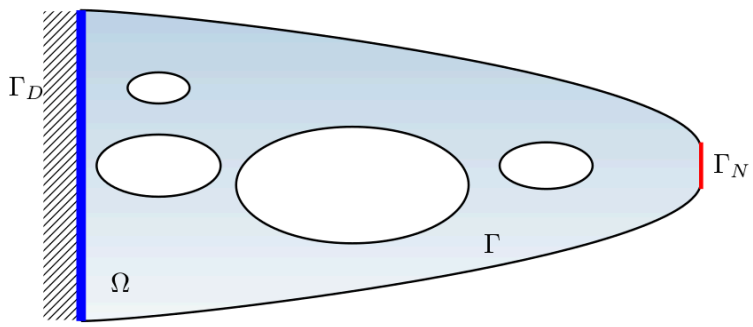


Figure 2.2: Shape Ω and boundary conditions

The open set $\Omega \in \mathcal{U}_{ad}$, which is a set of admissible shapes (typically defined by the user, see Def.(2.4) for an example). As stated earlier, via structural optimization, we wish to determine structures with desired mechanical properties. These mechanical properties can be viewed as an objective function denoted by $J(\Omega)$, that ought to be minimized. A general mathematical formulation of a shape optimization problem then reads

$$\inf_{\Omega \in \mathcal{U}_{ad}} J(\Omega). \tag{2.1}$$

One particular form of $J(\Omega)$ that one typically considers is

$$J(\Omega) = \int_{\Omega} m(\boldsymbol{u}(\Omega))\ dx + \int_{\partial\Omega} p(\boldsymbol{u}(\Omega))ds, \tag{2.2}$$

where $\boldsymbol{u}(\Omega)$ is solution to a partial differential equation (for example, solution to equation (2.7)), and $m(\cdot)$, $p(\cdot)$ are functions that are assumed to be at least $\mathcal{C}^1$ smooth such that m', p' have a linear growth rate. These regularity assumptions are made on $m(\cdot)$ and $p(\cdot)$ so that the objective function is well-defined.

The above expression of the objective $J(\Omega)$ can represent several mechanical properties such as the total compliance, elastic energy, stress, etc. as well as geometric properties such as volume and perimeter. One objective function that almost every researcher working on shape optimization deals with is compliance, given by

$$J(\Omega) = \int_{\Omega} \boldsymbol{f} \cdot \boldsymbol{u}(\Omega) dx + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{u}(\Omega)\ ds, \tag{2.3}$$

where $\boldsymbol{f}$ is the body force and $\boldsymbol{g}$, the surface force applied on the structure. Compliance is the work done by the external force. Lesser compliance implies a lesser deformation at the zone of application of the force (Γ_N) and in the shape Ω, and hence results in a lesser energy transferred to the structure from the external force. In the framework of linear elasticity, compliance is directly proportional to the inverse of rigidity of the structure. Thus, minimizing compliance implies maximizing rigidity.

Typically, the structure or the shape Ω is conceived inside a fixed design space $D \subset \mathbb{R}^d$ (for example, see Fig.2.4). The set of admissible shapes $\mathcal{U}_{ad}$ is user-defined. In this thesis, we define it as

$$\mathcal{U}_{ad} = \left\{ \Omega \subset D : \int_{\Omega} dx = V_t \right\}, \tag{2.4}$$

where V_t is a target volume. The existence of optimal shapes is subject to the choice of $\mathcal{U}_{ad}$. For instance, if one puts a perimeter constraint on Ω in (2.4), one can show the existence of a global minimum to the problem (2.1). In general, the question of existence of optimal shapes Ω is theoretically involved and we shall not dwell into it (see [93, 48] for more details). Rather, we content ourselves with computing numerical minimizers, using a gradient-descent method. A priori, this results in the determination of a local minimum.

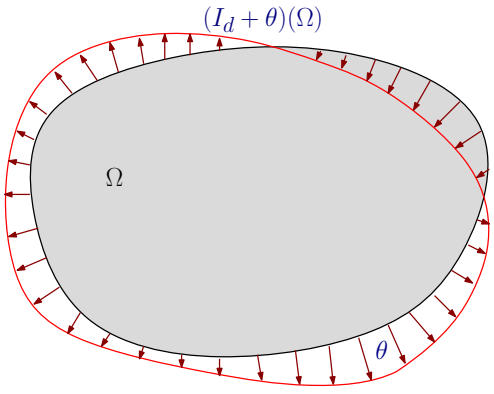


Figure 2.3: Perturbation of Ω using Hadamard's method

2.1.2 Hadamard's approach

In all the chapters, we shall resort to a gradient-descent optimization algorithm. Gradient-based optimization methods demand an explicit expression of the derivative of the objective function $J(\Omega)$. The gradient in the

context of shape optimization is based on the notion of the Hadamard shape derivative [2, 8, 93, 171]. As shown in Fig.2.1.2, starting from a smooth domain Ω, the perturbation of the domain is expressed as

$$\Omega_{\boldsymbol{\theta}} = (I_d + \boldsymbol{\theta})(\Omega),$$

where $\boldsymbol{\theta} \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ and I_d is the identity map. It is well-known that when the $W^{1,\infty}$ norm of $\boldsymbol{\theta}$ is sufficiently small, the map $I_d + \boldsymbol{\theta}$ is a diffeomorphism in $\mathbb{R}^d$. With this perturbation of the domain, one can define the notion of a Fréchet derivative for the function $J(\Omega)$.

Definition 1. *The shape derivative of $J(\Omega)$ at Ω is defined as the Fréchet derivative in $W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ evaluated at 0 for the mapping $\boldsymbol{\theta} \mapsto J((I_d + \boldsymbol{\theta})(\Omega))$ i.e.,*

$$J((I_d + \boldsymbol{\theta})(\Omega)) = J(\Omega) + J'(\Omega)(\boldsymbol{\theta}) + o(\boldsymbol{\theta}) \qquad \text{with} \qquad \lim_{\boldsymbol{\theta} \to 0} \frac{o(\boldsymbol{\theta})}{\|\boldsymbol{\theta}\|_{W^{1,\infty}}} = 0,$$

where $J'(\Omega)$ is a continuous linear form on $W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$.

For the objective function (2.2), the shape derivative $J'(\Omega)$ can be written using a function $j(\Omega)$ such that

$$J'(\Omega)(\boldsymbol{\theta}) = \int_{\partial\Omega} j(\Omega)\boldsymbol{\theta} \cdot \boldsymbol{n}\, dx$$

($\boldsymbol{n}$ being the outward normal to the Ω). We can then set the descent direction to $\boldsymbol{\theta} = -t\, j(\Omega)\boldsymbol{n}$, where $t \in \mathbb{R}^+$ is the descent step. For the new shape $\Omega_t = (I_d + \boldsymbol{\theta})\Omega$, we can formally write

$$J(\Omega_t) = J(\Omega) - t \int_{\partial\Omega} j(\Omega)^2 ds + \mathcal{O}(t^2).$$

Choosing a descent step t sufficiently small, one can show that the left hand side of the above is smaller than $J(\Omega)$, which guarantees a descent. Given an initial shape Ω, one can then apply the above gradient $J'(\Omega)$, and move the shape iteratively, minimizing the objective. In general, nothing ensures that our iterations would converge. Moreover, even in the case of convergence, one ends up in a final shape, which is often a local minimum, depending on the choice of the initial shape.

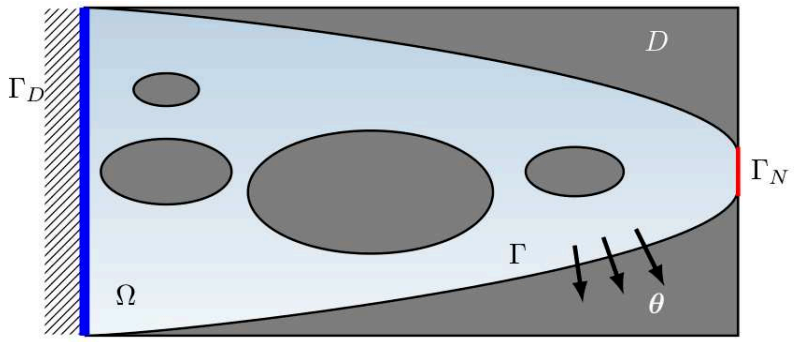


Figure 2.4: Design domain D and the shape Ω

As shown in Fig.2.4, the blue region represents the shape Ω, and the blue and grey area represent the design space D. Typically when a structure is designed, the clamped and the forced boundaries are assumed to be non-optimizable. Hence in all our optimizations, we allow only Γ to move along $\boldsymbol{\theta}$ as shown in Fig.2.4. To incorporate this constraint, we introduce the space

$$W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d) = \{\boldsymbol{\theta} \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d) : \ \boldsymbol{\theta} = \mathbf{0} \text{ on } \Gamma_N \cup \Gamma_D\}$$

and state a classical lemma we shall use for the derivation of shape derivative in every chapter.

Lemma 1. *Let Ω be a smooth bounded open set and $\varphi, \psi, \xi \in W^{1,1}(\mathbb{R}^d, \mathbb{R})$. Define $J_V(\Omega)$, $J_S(\Omega)$ and $J_N(\Omega)$ by*

$$J_V(\Omega) = \int_\Omega \varphi(x)dx \qquad J_S(\Omega) = \int_\Gamma \xi(x)\, ds, \qquad J_N(\Omega) = \int_{\Gamma_N} \psi(x)\, ds$$

then the functions $J_V(\Omega)$, $J_S(\Omega)$ and $J_N(\Omega)$ are differentiable at Ω in the direction $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ with the derivative

$$J'_V(\Omega)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} \varphi \, ds, \quad J'_S(\Omega)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} \left(\frac{\partial \xi}{\partial n} + H\xi \right) ds, \quad J'_N(\Omega)(\boldsymbol{\theta}) = 0,$$

where H is the mean curvature of $\partial\Omega$ (assumed to be smooth).

In Lemma 1, one differentiates a functional whose integrand functions are independent of Ω. To determine a shape derivative, one ought to differentiate (2.2), whose integrand functions ($m(\boldsymbol{u}(\Omega))$ and $p(\boldsymbol{u}(\Omega))$) are dependent on Ω. This dependency arises from the fact that $\boldsymbol{u}(\Omega)$ is solution to the partial differential equation defined on Ω. We thus need to determine the derivative of $\boldsymbol{u}(\Omega)$ with respect to Ω. In this regard, there exist two notions of derivatives: Eulerian and Lagrangian. We do not expound on the Eulerian and Lagrangian derivatives (see [2] for more details). Instead, we use the "fast derivation" method, proposed by Céa [53]. This method is presented for linear elasticity in next subsection.

2.1.3 Shape derivative computation for linear elasticity

Structural optimization is often performed assuming the structure to be linear with respect to external forcing. Linear elasticity is a mathematical model that assumes all mechanical deformations to be linear with respect to the external force. This assumption is a simplification of a physical phenomena that is by nature, nonlinear. The beauty of this assumption being, a lot of mathematical analysis tools are now rendered accessible.

Throughout this chapter and this thesis, we use the convention of representing all scalar mathematical entities by lowercase italic alphabets. The structure is represented by Ω as shown in Fig.2.2. Let

- $\mathcal{M}_s^d$ represent the set of symmetric $d \times d$ matrices
- $\mathbb{I}$ represent the fourth-order identity tensor of dimension d
- $\boldsymbol{u} : \Omega \to \mathbb{R}^d$ denote the displacement field
- $\boldsymbol{\sigma}$ denote the second order stress tensor
- $\boldsymbol{\varepsilon}(\boldsymbol{u})$ denote the second order strain tensor
- $\boldsymbol{n}$ denote the outward normal to $\partial\Omega$.

$\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}(\boldsymbol{u})$ are second order tensors of dimension d. The structure when subjected to a body force $\boldsymbol{f} : \Omega \to \mathbb{R}^d$ and a surface force $\boldsymbol{g} : \Gamma_N \to \mathbb{R}^d$ respects the momentum balance equation:

$$\begin{aligned}
\operatorname{div}(\boldsymbol{\sigma}) + \boldsymbol{f} &= \boldsymbol{0} && \text{in } \Omega, && (2.5a)\\
\boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N, && (2.5b)\\
\boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma, && (2.5c)\\
\boldsymbol{u} &= \boldsymbol{0} && \text{on } \Gamma_D. && (2.5d)
\end{aligned}$$

In the framework of linear elasticity, the mechanical displacement is assumed to be small. In this case, the strain tensor $\boldsymbol{\varepsilon}(\boldsymbol{u})$ is given by

$$\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2} \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T \right),$$

and the constitutive law is given by

$$\boldsymbol{\sigma} = \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}), \tag{2.6}$$

where $\mathbb{C}$ is the Hooke's tensor. The Hooke's tensor $\mathbb{C}$ is a fourth order tensor of dimension d and for a structure composed of an isotropic material, given by

$$\mathbb{C} = \lambda \mathbf{1} \otimes \mathbf{1} + 2\mu\mathbb{I},$$

where λ, μ are the Lamé coefficients and $\mathbf{1}$ is second order identity tensor of dimension d. For most materials, the tensor $\mathbb{C}$ is coercive, i.e., there exists $c_0 > 0$ such that,

$$\mathbb{C}\boldsymbol{\xi} : \boldsymbol{\xi} \geq c_0 |\boldsymbol{\xi}|^2 \qquad \forall \boldsymbol{\xi} \in \mathcal{M}_s^d.$$

Then using Eqns.(2.5) and the constitutive equation (2.6), the linear elasticity model is given by

$$\operatorname{div}(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})) + \boldsymbol{f} = \boldsymbol{0} \qquad \text{in } \Omega, \tag{2.7a}$$
$$\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n} = \boldsymbol{g} \qquad \text{on } \Gamma_N, \tag{2.7b}$$
$$\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n} = \boldsymbol{0} \qquad \text{on } \Gamma, \tag{2.7c}$$
$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } \Gamma_D. \tag{2.7d}$$

The above set of equations is also known as the strong form of linear elasticity, and the solution $\boldsymbol{u}$ to it, the strong solution.

The aim is now to find a solution $\boldsymbol{u}$ such that it satisfies the set of partial differential equations (2.7). In general, the solution to the strong form is known to exist under very strong conditions [3]. What one instead does is, convert the set of equations (2.7) into a "weak" form, that admits solutions under weaker conditions. This weak form is simply obtained multiplying the first equation in (2.7) by a (smooth) test function $\boldsymbol{v}$, integrating it by parts, and using the boundary conditions prescribed in (2.7), resulting in

$$\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v})\, dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, dx + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\, ds \tag{2.8}$$

The above is known as the weak form of linear elasticity. In the context of shape optimization, it is called as the state equation and the solution $\boldsymbol{u}$ to it, the state solution. In order to find the state solution, we define the displacement space

$$V = \{\boldsymbol{u} \in H^1(\Omega)^d \text{ such that } \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_D\}. \tag{2.9}$$

This space is complete with a norm $\|\cdot\|_V$ defined as

$$\|\boldsymbol{u}\|_V^2 = \int_\Omega \boldsymbol{u} \cdot \boldsymbol{u}\, dx + \int_\Omega \nabla \boldsymbol{u} : \nabla \boldsymbol{u}\, dx.$$

Let $a : V \times V \to \mathbb{R}$ be a bilinear form, defined as

$$a(\boldsymbol{u}, \boldsymbol{\varphi}) = \int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{\varphi}) dx. \tag{2.10}$$

and $l : V \to \mathbb{R}$ be a linear form given by

$$l(\boldsymbol{\varphi}) = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{\varphi}\, dx + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\varphi}\, ds.$$

Then using the bilinear form $a(\cdot, \cdot)$ and the linear form $l(\cdot)$, and choosing the test function $\boldsymbol{v} \in V$, the weak form (2.8) can be written in a compact manner as: find $\boldsymbol{u} \in V$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) = l(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V. \tag{2.11}$$

Given the coercivity of the Hooke's tensor and the Korn's inequality, one can show that

$$\exists \alpha > 0 \quad \text{such that} \quad \alpha \|\boldsymbol{v}\|_{H^1(\Omega)}^2 \leq a(\boldsymbol{v}, \boldsymbol{v}).$$

Thus the bilinear form $a(\cdot, \cdot)$ is coercive, or stable. If $\boldsymbol{f}, \boldsymbol{g} \in V'$ (the dual space of V), one can show that the linear form $l(\cdot)$ is continuous. The coercivity of $a(\cdot, \cdot)$ and the continuity of $l(\cdot)$ allows us to apply Lax-Milgram's Lemma to (2.11) leading to the existence of a unique solution $\boldsymbol{u} \in V$, which satisfies

$$\|\boldsymbol{u}\|_V \leq \frac{C}{\alpha} \left(\|f\|_{V'} + \|g\|_{V'} \right),$$

where C is some constant.

We now consider the objective function (2.2) (where $\boldsymbol{u}(\Omega)$ is the solution to (2.11)). The shape derivative for this objective is determined using Céa's method. Explained in one line, this method is a Lagrange-multiplier method where the objective (2.2) is minimized with the differential equation (2.7) viewed as a constraint. In this method, one constructs a Lagrangian by assuming the variable $\boldsymbol{u}$ to be independent of Ω, and by introducing a Lagrange-multiplier $\boldsymbol{v}$ (a variable defined in $\mathbb{R}^d$). One then proceeds to apply the optimality condition on the Lagrangian with respect to its variables, $\boldsymbol{u}, \boldsymbol{v}, \Omega$. The optimality condition

on the resulting Lagrangian results in an adjoint equation (a partial differential equation to determine $\boldsymbol{v}$). For this optimality condition to hold, the adjoint equation must be well-posed. Further, to determine the shape derivative, one must differentiate the Lagrangian with respect to Ω at $\boldsymbol{u}(\Omega)$. This differentiation is justified if $\boldsymbol{u}(\Omega)$ is differentiable with respect to the shape Ω. This is a strong requirement and often taken for granted. Hence, Céa's method can be used to determine the shape derivative if the adjoint equation is well-posed and $\boldsymbol{u}(\Omega)$ is differentiable respect to Ω.

In the case of linear elasticity, under additional hypothesis on the regularity of the boundary $\partial\Omega$, one can show that $\boldsymbol{u}(\Omega)$ is indeed differentiable with respect to Ω [10, 2]. However, in the case of nonlinear phenomena (like contact, plasticity and fracture), this far from evident. Nevertheless, in every chapter, we assume that $\boldsymbol{u}(\Omega)$ is differentiable with respect to Ω and the adjoint equation is well-posed. After deriving the shape derivative, we show that adjoint equation is indeed well-posed. However, we do not show that the solution is not differentiable with respect to Ω (which is out of the scope of this thesis).

Theorem 1. *Let $\Omega \subset \mathbb{R}^d$ be a smooth bounded open set. Let $\boldsymbol{f} \in H^1(\mathbb{R}^d)^d$, $\boldsymbol{g} \in H^2(\mathbb{R}^d)^d$ and $\boldsymbol{u}(\Omega) \in V$ the solution to (2.11). Then the shape derivative of $J(\Omega)$ along $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, $J'(\Omega)(\boldsymbol{\theta})$ is given by*

$$J'(\Omega)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} \Big(m(\boldsymbol{u}) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{f} \cdot \boldsymbol{v} \Big) ds \tag{2.12}$$

where $\boldsymbol{v}(\Omega) \in V$ is the solution to the adjoint problem,

$$a(\boldsymbol{v}, \boldsymbol{\varphi}) = -\int_\Omega m'(\boldsymbol{u}) \cdot \boldsymbol{\varphi}\, dx - \int_{\Gamma_N} p'(\boldsymbol{u}) \cdot \boldsymbol{\varphi}\, ds \quad \forall \boldsymbol{\varphi} \in V, \tag{2.13}$$

which can be easily shown to be well-posed.

Proof. Define space $\tilde{V}$ (analog of the space V given in (2.9))

$$\tilde{V} = \{\boldsymbol{u} \in H^1(\mathbb{R}^d)^d \text{ such that } \boldsymbol{u} = 0 \text{ on } \Gamma_D\}. \tag{2.14}$$

As per Céa's technique [53], for $\tilde{\boldsymbol{u}}$, $\tilde{\boldsymbol{v}} \in \tilde{V}$, we construct a Lagrangian

$$\mathcal{L}(\Omega, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) = \int_\Omega m(\tilde{\boldsymbol{u}})\, dx + \int_{\Gamma_N} p(\tilde{\boldsymbol{u}})\, ds + a(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) - \int_\Omega \boldsymbol{f} \cdot \tilde{\boldsymbol{v}}\, dx - \int_{\Gamma_N} \boldsymbol{g} \cdot \tilde{\boldsymbol{v}}\, ds. \tag{2.15}$$

The variables $\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}$ are defined on the full space $\mathbb{R}^d$ and are thus independent of Ω. Although $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ are chosen such that they vanish on Γ_D, they do not depend on Ω since Γ_D is a fixed boundary.

Given the construction of $\mathcal{L}$ (2.15), applying the optimality condition on $\mathcal{L}$ with respect to the independent variables $\boldsymbol{u}, \boldsymbol{v}$ and Ω yields the state equation, the adjoint equation and the shape derivative respectively.

The optimality condition on $\mathcal{L}$ with respect to $\tilde{\boldsymbol{v}}$ amounts to differentiating it with respect to the adjoint variable $\tilde{\boldsymbol{v}}$ along $\boldsymbol{\varphi} \in \tilde{V}$, equating it to zero, followed by the substitution $\tilde{\boldsymbol{u}} = \boldsymbol{u}$, resulting in

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{v}}(\boldsymbol{\varphi}) = a(\boldsymbol{u}, \boldsymbol{\varphi}) - \int_\Omega \boldsymbol{f} \cdot \boldsymbol{\varphi}\, dx - \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{\varphi}\, ds = 0 \qquad \forall \boldsymbol{\varphi} \in \tilde{V}.$$

The bilinear form $a(\cdot, \cdot)$ in the above being defined only on Ω, we can replace $\tilde{V}$ by V. We thus recover the state equation (2.11). In a similar way, we apply the optimality condition on $\mathcal{L}$ with respect to $\tilde{\boldsymbol{u}}$ at $\tilde{\boldsymbol{u}} = \boldsymbol{u}$, $\tilde{\boldsymbol{v}} = \boldsymbol{v}$,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{u}}(\boldsymbol{\varphi}) = \int_\Omega m'(\boldsymbol{u}) \cdot \boldsymbol{\varphi}\, dx + \int_{\Gamma_N} p'(\boldsymbol{u}) \cdot \boldsymbol{\varphi}\, ds + a(\boldsymbol{\varphi}, \boldsymbol{v}) = 0 \qquad \forall \boldsymbol{\varphi} \in \tilde{V}.$$

Since all integrals in the above are defined only on Ω, we can replace $\tilde{V}$ by V. We thus obtain the following adjoint equation:

$$a(\boldsymbol{v}, \boldsymbol{\varphi}) = -\int_\Omega m'(\boldsymbol{u}) \cdot \boldsymbol{\varphi}\, dx - \int_{\Gamma_N} p'(\boldsymbol{u}) \cdot \boldsymbol{\varphi}\, ds \quad \forall \boldsymbol{\varphi} \in V.$$

Finally, using the relation $J(\Omega) = \mathcal{L}(\Omega, \boldsymbol{u}, \tilde{\boldsymbol{v}})$, we determine the shape derivative $J'(\Omega)(\boldsymbol{\theta})$ for any $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ by chain differentiation rule

$$J'(\Omega)(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{\theta}) + \frac{\partial \mathcal{L}}{\partial \boldsymbol{u}} \left(\frac{\partial \boldsymbol{u}}{\partial \Omega}(\boldsymbol{\theta}) \right).$$

Now, using the adjoint equation, the last term annihilates in the above and we get

$$J'(\Omega)(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{\theta}).$$

Formula (2.12) is deduced by straight forward application of Lemma 1. □

As mentioned earlier, the Céa's method in the above was applied assuming the adjoint equation to be well-posed. Fortunately, the adjoint equation (2.13) turned out to be well-posed. This is because the constraint in the method was the elasticity equation (2.7), which is linear and well-posed. However, if instead we had a nonlinear equation or an inequation as constraint (which is the case in the Chapters 3, 4, 5 and 6), nothing would guarantee that the adjoint equation shall be well-posed. In such a case, we apply the Céa's method assuming the adjoint equation to be well-posed. And after deriving the shape derivative, we check if it the adjoint equation is well-posed or not. This shall be our approach in all the chapters in this thesis.

2.2 Level-set method

Any topology optimization method necessitates a way to define and alter a topology of the shape. The method must ensure that topology changes are captured. One obvious possibility is to parametrize the boundary Γ of the shape Ω and move the boundary along its normal $\boldsymbol{n}$ by the magnitude of the shape derivative. This approach cannot capture topology changes.

The level-set method was introduced by Osher and Sethian [144] to capture the flame front in combustion. They proposed to use the zero isovalue of a continuous function to define the boundary $\partial\Omega$ of a shape Ω. The function defining a shape is such a manner is called a level-set function as each level or value represents a boundary and thus a shape. The mathematical formulation being, for $\Omega \in \mathbb{R}^d$, the level-set function $\phi : D \to \mathbb{R}$ is defined as

$$\begin{cases} \phi(\boldsymbol{x}) < 0 & \text{if } \boldsymbol{x} \in \Omega, \\ \phi(\boldsymbol{x}) = 0 & \text{if } \boldsymbol{x} \in \Gamma, \\ \phi(\boldsymbol{x}) > 0 & \text{if } \boldsymbol{x} \in \overline{\Omega}^c \end{cases} \tag{2.16}$$

where Γ is the movable part of the boundary $\partial\Omega$ and D is the design space as shown in Fig.2.4. This level-set method was adapted to the shape optimization framework in [12, 157, 184]. The crux of the method lies in letting the shape deform along a velocity field $\boldsymbol{\theta} : D \to \mathbb{R}^d$. The evolution of the shape is governed by the transport equation

$$\frac{\partial \phi}{\partial t} + \boldsymbol{\theta} \cdot \nabla \phi = 0. \tag{2.17}$$

Very often, the velocity field is oriented along the normal, namely $\boldsymbol{\theta} = \theta \boldsymbol{n}$ where $\boldsymbol{n} = \nabla\phi/|\nabla\phi|$ and the scalar function θ is the normal velocity. In such a case, (2.17) can be re-written as a Hamilton-Jacobi equation

$$\frac{\partial \phi}{\partial t} + \theta |\nabla \phi| = 0. \tag{2.18}$$

In the framework of shape optimization, the normal velocity θ is chosen to be $j(\Omega)$ (the integrand in (2.12)). In the Hamilton Jacobi equation above, if the initial level-set ϕ is such that $|\nabla\phi(\boldsymbol{x})| = 1$, every point $\boldsymbol{x} \in \Gamma$ shall move along the normal $\boldsymbol{n}$ by a value that is exactly equal to the magnitude of the shape derivative. Thus it is preferable to construct a level-set such that $|\nabla\phi(\boldsymbol{x})| = 1$ on Γ. One possible construction is the signed distance function [143], d_Ω is defined such that

$$\forall \boldsymbol{x} \in D,\ d_\Omega(\boldsymbol{x}) = \begin{cases} \min|P_\Gamma(\boldsymbol{x}) - \boldsymbol{x}| & \text{if } \boldsymbol{x} \in \Omega, \\ 0 & \text{if } \boldsymbol{x} \in \Gamma, \\ -\min|P_\Gamma(\boldsymbol{x}) - \boldsymbol{x}| & \text{if } \boldsymbol{x} \in \overline{\Omega}^c, \end{cases} \tag{2.19}$$

where P_Γ is the orthogonal projection of $\boldsymbol{x}$ on Γ. Since Ω is not necessarily convex, the orthogonal projection $P_\Gamma(\boldsymbol{x})$ is not necessarily unique. In order to define $d_\Omega(\boldsymbol{x})$ uniquely, we take the min of all possible projections $P_\Gamma(\boldsymbol{x})$ in the definition (2.19).

As an example, we plot the signed distance level-set function d_Ω in Fig.2.5, where Ω is the statue of liberty. When one solves the transport equation (2.17), the level-set function ϕ evolves and its corresponding shape. Since the shape evolves, the shape derivative evolves as well. The descent direction $\boldsymbol{\theta}$, that depends

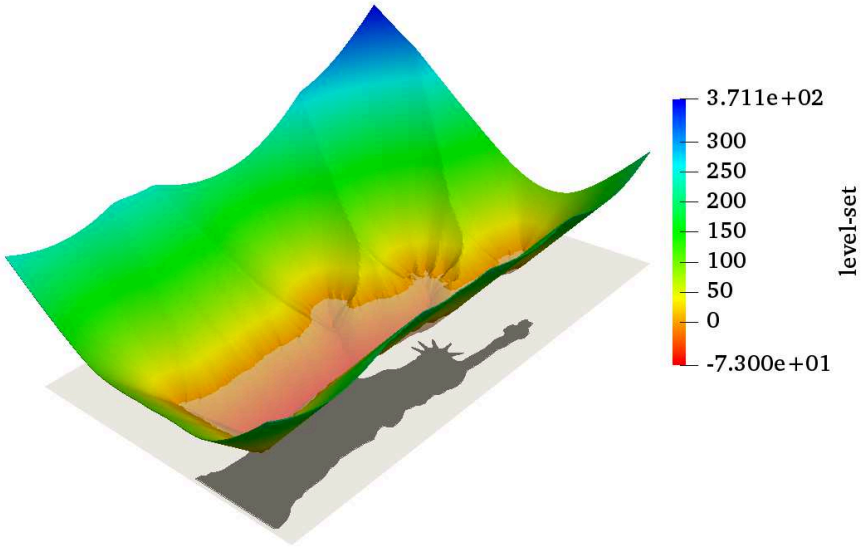


Figure 2.5: Signed distance level-set function ϕ representing statue of liberty (in black)

on the shape derivative must evolve, but is assumed to be constant. Since the transport equation (2.17) is solved for a very small time step (2.28), this assumption on $\boldsymbol{\theta}$ turns out to be fairly robust.

Moreover, descent direction $\boldsymbol{\theta}$ is defined using the shape derivative and not the topological derivative. Yet resolution of the transport equation results in topological changes in the shape. When the topological changes occur, there is no guarantee that the objective function shall diminish. Nevertheless, we use the shape derivative to perform topology optimization.

In our numerical setting, we work with the linear transport equation (2.17) because we use non cartesian meshes. The level-set function is a $\mathbb{P}^1$ function on a simplicial mesh. The initial level-set function that is given as input to the Eqn.(2.17) is chosen as $\phi = d_\Omega$ (2.19). We rely on the library `advect` [39] which solves (2.17) by the method of characteristics, known to be unconditionally stable. The equation (2.17) is solved for a small time interval δt (specified in algorithm 1). The result is a new level-set function representing the updated shape. However, the new level-set function is no longer a signed-distance function. Hence, the level-set function ought to be re-initialized. This is brought about by using a library `mshdist` [62]. The library `mshdist` re-initialize by solving the Eikonal equation.

2.2.1 Ersatz material approach

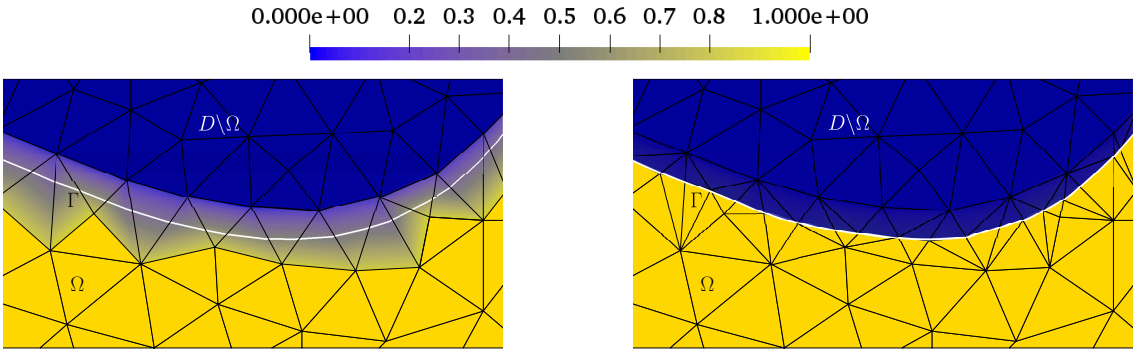
In the original level-set method ([12, 157, 184]), the mesh for the design space D (see Fig.2.4) is kept fixed for all the shape optimization iterations. The computation of the state equation (2.11) and the adjoint equation (2.13), is brought about in the entire design space D (instead of Ω). This is done by assuming that the "void" space $D\backslash\Omega$ has an ersatz material (having a very small stiffness). The Hooke's tensor $\mathbb{C}$ is modified to $\mathbb{C}_\epsilon(\Omega)$, using a density function $\rho_\epsilon(\Omega)$

$$\rho_\epsilon(\Omega) = (\chi_\Omega + (1 - \chi_\Omega)\epsilon)\,, \qquad \mathbb{C}_\epsilon(\Omega) = \rho_\epsilon(\Omega)\mathbb{C}, \tag{2.20}$$

where χ_Ω is the characteristic function of the domain Ω and ϵ is the residual stiffness. The displacement solution now becomes a function of the ersatz material stiffness, denoted by $\boldsymbol{u}_\epsilon(\Omega)$ and is slightly different from the solution obtained without ersatz, $\boldsymbol{u}(\Omega)$. This is difference is proportional to ϵ, and is given by the estimate [57]

$$\|\boldsymbol{u} - \boldsymbol{u}_\epsilon\|_{H^1(\Omega)} \leq C(\Omega, D, \boldsymbol{f}, \boldsymbol{g})\epsilon$$

where C is a constant. In practice, ϵ is typically taken to be 10^{-3}. The characteristic function χ_Ω is $\mathbb{P}^1$ on a simplicial mesh. Using this stiffness the variational formulation (2.11) is defined on the entire design space D. The ersatz approach shall be considered for the shape optimization only in Chapter 3. The advantage



(a) Fixed mesh (b) XFEM mesh

Figure 2.6: Ersatz material density $\rho_\epsilon(\Omega)$

of the ersatz approach is the fact that no remeshing is needed and all the calculations can be performed on the same mesh. The topological changes are easily captured on the same mesh. However, there are two drawbacks:

1. The computation time of state and adjoint equations on D is higher compared to the same computation on Ω
2. There are intermediate densities in the mesh elements containing $\partial\Omega$ as seen in Fig.2.6a

In order to deal with the first drawback, one may consider eliminating the entries in the stiffness matrix corresponding to the ersatz material, to end up with a stiffness matrix that is relatively smaller. However, in this apprach, the intermediate densities still remain. To deal with the second drawback, there are two solutions:

1. XFEM [180]
   This involves eliminating mesh elements consisting of ersatz and cutting the mesh elements on boundary Γ, so as to capture the shape Ω as shown in Fig.2.6b. The issue with this approach being, one might end up with heavily distorted boundary elements. This leads to an ill-conditioned stiffness matrix having a very small coercivity constant.
2. CutFEM [41]
   This involves eliminating mesh elements consisting of ersatz and enriching the finite-element basis function of the interface zone (the elements containing the boundary Γ). This along with a ghost penalty term on the interface zone can ensure that the coercivity constant is controlled and the stiffness matrix remains well-conditioned.

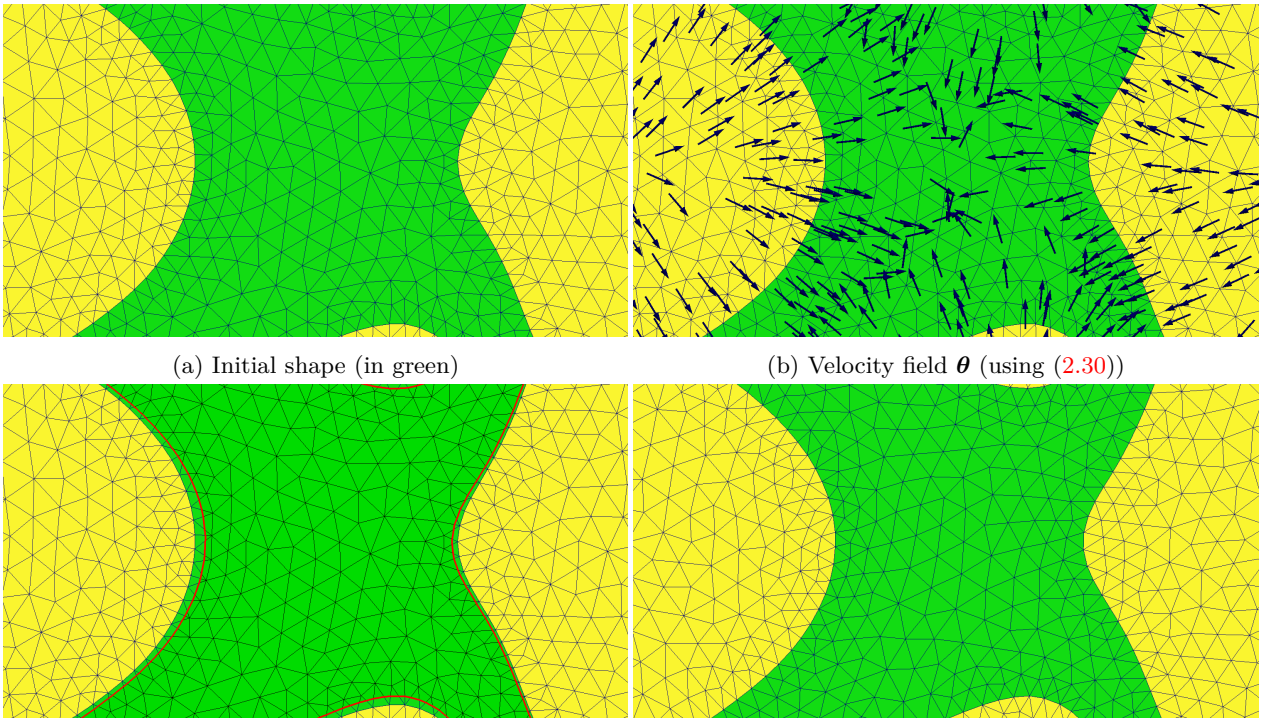
An alternative to the XFEM and the CutFEM approach is remeshing the whole domain D, explained in the next subsection.

2.2.2 Remeshing

One method to avoid intermediate densities on the boundary Γ (as seen in Fig.2.6a) is remeshing. By remeshing, we mean construction of a body-fitted mesh that captures the boundary of the shape Ω. In this remeshing approach, the mesh quality can be controlled, resulting in stiffness matrix, that is well-conditioned compared to the XFEM approach. Moreover, no additional penalty terms are needed as in CutFEM approach. The remeshing approach (using the level-set method) was first adapted to topology optimization by C.Dapogny [60, 61, 39, 7]. Later, several applications were developed, for instance for fluid-structure interaction [73], heat exchangers [74], etc. The only downside of this approach is the remeshing computing time (see Chapter 7 for more details).

The application of remeshing in the context of shape optimization is explained with the following example. As shown in Fig.2.7a, consider an initial shape and its corresponding mesh (in green). For this shape, the velocity for advection is computed using the formula (2.30), resulting in $\boldsymbol{\theta}$ as shown in Fig.2.7b. Using this velocity field, the transport equation (2.17) is solved for a time step τ (given in (2.28)) using `advect` [39],

resulting in the new shape in Fig.2.7c (marked with red lines). This new shape is then remeshed (using a tool `MMG`, explained in the next subsection) to obtain the body-fitted mesh (2.7d). Finally, for the new shape, the signed distance level-set function is computed using `mshdist` [62].



(a) Initial shape (in green)

(b) Velocity field $\boldsymbol{\theta}$ (using (2.30))

(c) New shape after solving transport Eqn.2.17

(d) New shape captured using body-fitted mesh

Figure 2.7

2.2.3 Remeshing using `MMG`

We perform remeshing using an open source library `MMG 5.5.2` [61]. `MMG` is a remeshing tool that can not only remesh to capture a shape (defined by a level-set function) but also to improve the quality of the mesh. The documentation of `MMG` can be found online [1]. The program has to be launched via command line. In order to launch it, the user has to give several parameters (a few optional ones) as input. The parameters that ought to be defined for our application are

1. `hmax`: the prescribed maximal mesh size
2. `hmin`: the prescribed minimal mesh size
3. `hausd`: the Hausdorff metric (to refine mesh according to the curvature of $\partial\Omega$)
4. `rmc`: the volume fraction of small floating island inside or outside the domain Ω to be removed; as islands outside the domain Ω may make the stiffness matrix non-invertible
5. `nr`: a flag to avert `MMG` about detecting edges and corners of Ω
6. `sol`: a flag to give the level-set function as input

`hmin` is the prescribed lower bound on the minimal mesh size and not the actual minimal mesh size. Hence, the minimal mesh size of the mesh generated by `MMG` can be very well bigger than `hmin`. In the same way `hmax` is the prescribed upper bound on the maximal mesh size and not the actual maximal mesh size. The parameter that takes precedence over `hmin` and `hmax` is `hausd`. This parameter is a measure of mesh refinement, lesser the value, more the refinement. The mesh refinement is by default according to the

[1] https://www.mmgtools.org/mmg-remesher-try-mmg/mmg-remesher-options

curvature of the boundary $\partial\Omega$, but it can also be according to a predefined metric function. If `hausd` is too small, the minimal mesh size may be lesser than `hmin` as `hausd` is given preference. The parameters `hmin` and `hmax` are thus not always respected in order to create a good quality mesh.

For any test case in this entire thesis, we first choose $h_{\min}$ in a heuristic manner. Using this, we define the above parameters as for 2D geometries as

$$\texttt{hmax} = 2\,\texttt{hmin},\ \texttt{hausd} = 0.1\,\texttt{hmin},\ \texttt{rmc} = 20\,\texttt{hmin}^2, \tag{2.21}$$

and for 3D geometries as

$$\texttt{hmax} = 2\,\texttt{hmin},\ \texttt{hausd} = 0.5\,\texttt{hmin},\ \texttt{rmc} = 10\,\texttt{hmin}^3. \tag{2.22}$$

This choice of these parameters is very much based out of heuristics.

The mesh file given as input to `MMG` to construct a body-fitted mesh must contain the list of corners and edge vertices of the design space D. If the mesh file does not contain this list, while `MMG` remeshes, the corners and edges might be smoothened, leading to a fillet. In order to avert `MMG` about the corners and edges, one needs to remesh the first mesh of the shape optimization using `MMG` without the flag `nr`. An example of the command that ought to be given in the command-line is

```
mmg2d_03 -rmc 0.0005 -hmax 0.01 -hmin 0.005  -hausd 0.0005 levelsetmesh.mesh -sol
    levelsetmesh.sol
```

This creates a list of corner and edge vertices in the output file. For the succeeding iterations, the flag `nr` need to be given and the command to remesh reads

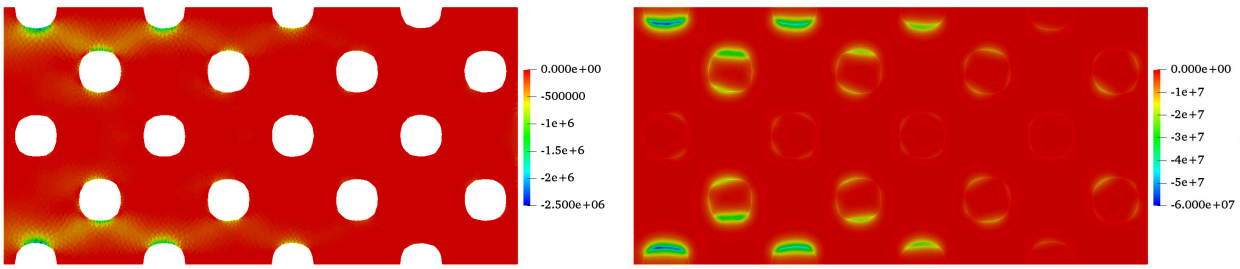
```
mmg2d_03 -rmc 0.0005 -hmax 0.01 -hmin 0.005 -nr -hausd 0.0001  levelsetmesh.mesh
   -sol levelsetmesh.sol
```

for the remaining iterations.

2.2.4 Regularization and extension of the shape derivative (2.12)

In a typical topology optimization algorithm, the shape derivative (also known as sensitivity) must be smoothed. This is typically done by solving a partial differential equation. In the density-based methods, such a smoothing algorithm is known as a sensitivity filter [115, 32], and plays a pivotal role in avoiding numerical instabilities. In level-set based methods, smoothing (or regularization) plays a crucial role in ensuring the smooth evolution of the topology [12, 64, 40]. Such a smoothing is very classical and well-known in geometric optimization since a long time [135].

The shape derivative computed numerically using the expression (2.12) is typically not very smooth. There are two factors that contribute to this lack of smoothness. First, during the shape optimization process, the intermediate shapes may have sharp boundaries. So the shape derivative may not be rigorously defined on the optimizible boundary Γ. Second, when $\boldsymbol{u}$ is $\mathbb{P}^1$ smooth, the shape derivative (2.12) is $\mathbb{P}^0$ smooth (an example plotted in Fig.2.8a). Given the lack of smoothness, it is imperative to regularize the



(a) Shape derivative $j(\Omega)$(2.12) extended to Ω (b) Regularized shape derivative $dj_\alpha(\Omega)$ (using (2.23))

shape derivative [40, 64] in such a way that it is still a descent direction. One possibility is to consider the H^1 scalar product instead of the L^2 scalar product by finding a function $dj_\alpha(\Omega) \in H^1(D)$. We introduce a regularization parameter $0 < \alpha \ll 1$ and solve the variational problem: find $dj_\alpha(\Omega) \in H^1(\Omega)$ such that

$$\int_D (\alpha^2 \nabla dj_\alpha(\Omega) \cdot \nabla\varphi + dj_\alpha(\Omega)\varphi) dx = \int_\Gamma j(\Omega)\varphi\ ds \qquad \forall \varphi \in H^1(\Omega). \tag{2.23}$$

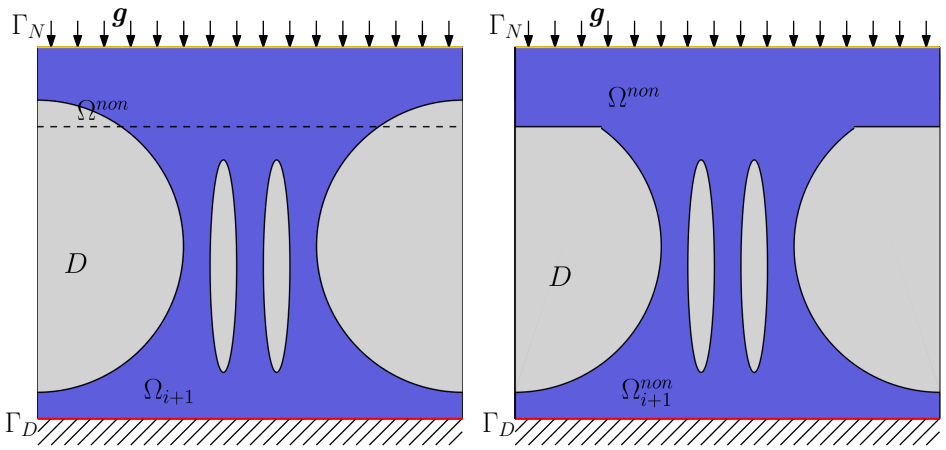


Figure 2.9: The shape Ω_{i+1} after the application of the non-optimizable domain using (2.24)

Typically, α is taken to be $h_{\min}$ or $h_{\max}$ (the minimal or maximal mesh size, respectively). Larger the parameter, greater the regularization. One can quite easily demonstrate that the above problem (2.23) is well-posed and admits a unique solution $dj_\alpha(\Omega) \in H^1(\Omega)$.

2.2.5 Imposing a non-optimizable domain

Having regularized the shape derivative in the previous subsection, we can easily perform the advection by solving (2.17) (using `advect` library [39]). After advection, we might wish to impose a non-optimizable domain Ω^{non}. Typical non optimizable domains are around the zone of application of force Γ_N, or around the Dirichlet boundary, an example is given in Fig.2.9. To do this, assume that after the advection and the volume correction at optimization iteration i (the 8-th step of the algorithm 1), we determine the level-set ϕ_{i+1} and a shape Ω_{i+1}, which must contain a non-optimizable domain. Let ϕ^{non} be the signed distance function corresponding to Ω^{non}. Then in order to impose this domain, we simply modify ϕ_{i+1} using

$$\phi^{non}_{i+1} = \min(\phi_{i+1}, \phi^{non}), \tag{2.24}$$

just before remeshing at the end of optimization iteration i. The levelset function ϕ^{non}_{i+1} is later reinitialized using `mshdist`. An an example, we show Fig.2.9, where a non optimizable domain has been imposed. Another possibility of imposing a non-optimizable domain is via solving (2.23) with a homogeneous Dirichlet boundary condition imposed on $dj(\Omega)$ in Ω^{non}. This possibility is not explored here as it might result in artificial changes in the shape derivative around Ω^{non}.

2.2.6 Shape optimization algorithm

Assume that the design space D contains a non-optimizable domain Ω^{non} and its complementary set, the optimizable domain, $D\backslash\Omega^{non}$. By volume of the domain, we mean the volume of the optimizable domain. The space of admissible spaces is thus redefined as

$$\mathcal{U}_{ad} = \left\{ \Omega \subset D : \; V_\Omega = \int_{\Omega\backslash\Omega^{non}} dx = V_t \right\}, \tag{2.25}$$

We consider the shape optimization problem

$$\min_{\Omega \in \mathcal{U}_{ad}} J(\Omega).$$

A simple strategy to perform the above shape optimization is given in [12].

2.2.6.1 Strategy in [12]

A Lagrangian (cost function) is constructed as

$$\mathcal{L}(\Omega, \lambda) = J(\Omega) + \ell V_\Omega,$$

where ℓ is the Lagrange multiplier that ensures that the volume is minimized and the target volume is attained.

The Hamilton–Jacobi equation (2.18) is solved using an explicit scheme. A time step τ is updated at each optimization iteration such that the CFL condition is respected. The decrease of the cost function $\mathcal{L}(\Omega, \lambda)$ is monitored and a time stepping scheme is performed in case of an increase of the cost function.

2.2.6.2 Strategy used in this thesis

In this thesis, in order to ensure that the optimized shape satisfies the volume constraint in (2.25), we follow a slightly different strategy by constructing the following Lagrangian

$$\mathcal{L}(\Omega, \lambda) = J(\Omega) + \frac{\lambda}{C_V}\left(V_\Omega - V_t\right), \tag{2.26}$$

where λ is the Lagrange multiplier for the volume constraint, V_{Ω^h} is the volume of the optimizable domain, and C_V is a normalization constant. The optimizable domain is D unless otherwise mentioned. Denoting by Ω_0 the initial shape, the constant C_V is defined by

$$C_V = \left|V_{\Omega_0} - V_f\right|. \tag{2.27}$$

We apply a standard gradient-based Uzawa-type algorithm to the Lagrangian (4.111). Let $I_{\max} = 200$ be the maximal number of shape optimization iterations. The iteration number is denoted by i with $1 \leq i \leq I_{\max}$. At each iteration i, once the shape derivative $dj_\alpha(\Omega_i)$ is evaluated by (2.23), a pseudo-time step (or descent step) τ is defined by

$$\tau = \frac{h_{\min}}{2C_i}, \tag{2.28}$$

where $h_{\min}$ is the minimal mesh size of the first iteration and C_i is a normalization constant, given by

$$C_i = \int_{\partial\Omega_i} \left|dj_\alpha(\Omega_i)\right| dx.$$

Updating the constant C_i at every iteration of the optimization process ensures a control of the descent step τ.

For every optimization iteration, the simplest choice of the multiplier λ is the one that ensures that the volume constraint is satisfied. If the initial volume is much larger than the target volume, we might deflate the holes too much and break the structure Ω. We thus prefer not to satisfy the volume constraint at all iterations, instead update the multiplier λ at each iteration by

$$\lambda_{i+1} = \lambda_i + \frac{C_i \tau}{C_V}\left(V_{\Omega_i} - V_f\right), \tag{2.29}$$

ensuring that the volume will converge (slowly) to the target volume. Then, for the descent step τ, the transport equation (2.17) is solved with a velocity θ_i, given by

$$\boldsymbol{\theta}_i = \left(dj_\alpha(\Omega_i) + \frac{\lambda_{i+1}}{C_V}\right)\boldsymbol{n}, \tag{2.30}$$

where $\boldsymbol{n} = \nabla\phi_i$ is the normal to the level-set function associated to the shape Ω_i.

To improve the satisfaction of the volume constraint, we apply the following trick. As soon as the volume is close to the volume target, namely $|V_{i+1} - V_f| \leq 10^{-1} V_f$, we apply a projection algorithm to satisfy the target volume exactly. More precisely, the level-set ϕ_{i+1} is iteratively updated by

$$\phi_{i+1} = \phi_{i+1} + \frac{V_{\Omega_{i+1}} - V_f}{\int_{\partial\Omega_{i+1}} ds}, \tag{2.31}$$

until $|V_{\Omega_{i+1}} - V_f| \leq 10^{-4} V_f$. The newly obtained shape $V_{\Omega_{i+1}}$ is remeshed with `MMG` [61]. Eventually, the objective function $J(\Omega_{i+1})$ is evaluated but is not compared to the previous value $J(\Omega_i)$. Summing up this subsection, we basically implement Algorithm 1. The shape optimization algorithm used in the succeeding chapters is very similar to the algorithm 1, modulo a few differences, which shall be specified.

Algorithm 1 Repeat over $i = 1, \cdots, I_{max}$

1. Solve (2.11) for the state solution $\boldsymbol{u}$ on the mesh of Ω_i
2. Solve (2.13) for the adjoint solution $\boldsymbol{v}$ on the mesh of Ω_i
3. Compute the shape derivative $J'(\Omega_i)$ using (2.12) and regularize it by solving (2.23) for $dj_\alpha(\Omega_i)$
4. Update the multiplier λ_{i+1} using (2.29)
5. Solve the transport equation (2.17) with the velocity given by (2.30) for the pseudo-time step τ given by (2.28) to obtain the new level-set function $\tilde{\phi}_{i+1}$
6. Re-initialize $\tilde{\phi}_{i+1}$ to the signed distance function ϕ_{i+1} (defining a new shape Ω_{i+1})
7. Compute the volume $V_{\Omega_{i+1}}$. If it is close to the volume target, apply the projection algorithm (2.31) to satisfy exactly the volume constraint.
8. Impose non-optimizable domain using (2.24)
9. Remesh the box D using `MMG` [61] to obtain the body-fitted mesh of the new shape Ω_{i+1}

2.3 3D Numerical Results

In this section, we shall study briefly, a few well-known test cases (cantilever, L-beam, etc.). The material properties are taken to be $E = 210$GPa and $\nu = 0.3$. The finite element computations are brought about on an open source software `FreeFEM` [92]. In all the test cases in this section, only surface forces are applied and body forces are not applied. The finite element mesh consists of tetrahedral elements and the displacement solution $\boldsymbol{u} \in \mathbb{P}^1(\Omega)^3$.

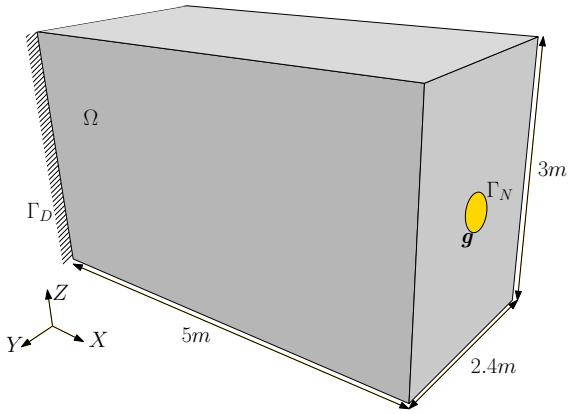
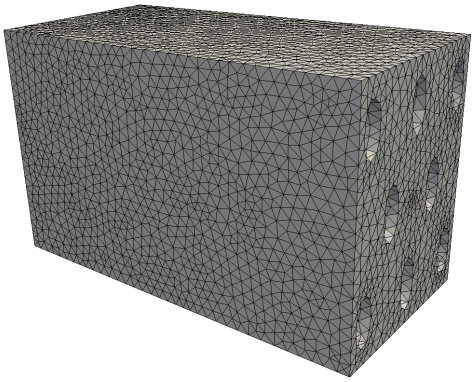


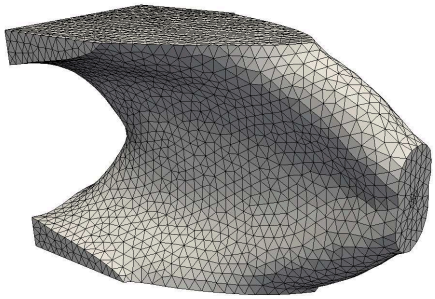
Figure 2.10: Boundary condition applied on the 3D cantilever beam

2.3.1 Cantilever beam

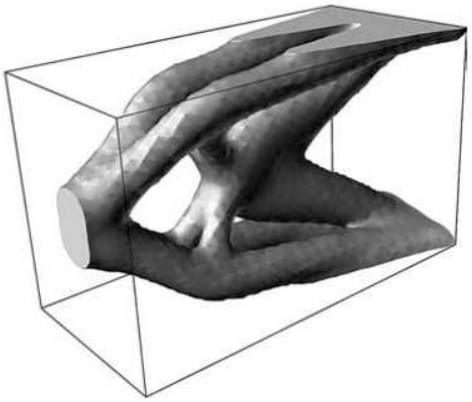
We consider a cantilever beam of dimensions $5m \times 2.4m \times 3m$, as shown in the Fig.2.10. The cantilever beam is fixed on its leftmost plane, forced downwards on a circular region of radius $0.1m$ on its rightmost plane (in chrome) with $\boldsymbol{g} = (0, 0, -1000)$MN/m. The compliance (2.3) is minimized. A target volume $V_t = 6m^3$ is imposed. We show the initial shape and the final shape in Fig.2.11. The corresponding convergence histories are plotted in Fig.2.12 The optimal shape we have obtained in Fig.2.11 seems to have a fairly simple topology, with a plate-like structure at the center. The shape resembles an I-beam with varying length. This is quite expected from an engineering perspective as I-beams are known to be extremely stiff structures. The same 3D cantilever beam optimized using the ersatz material approach might result in an optimal shape with several bars (see Fig.2.11c) instead of plate like-structures. Moreover, in the ersatz material approach, if the target volume fraction is decreased, the chances of seeing bar-like structures is higher rather than seeing



(a) Initial shape



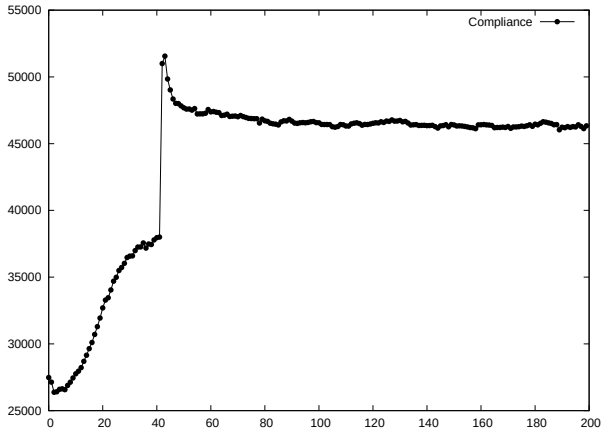
(b) Final shape



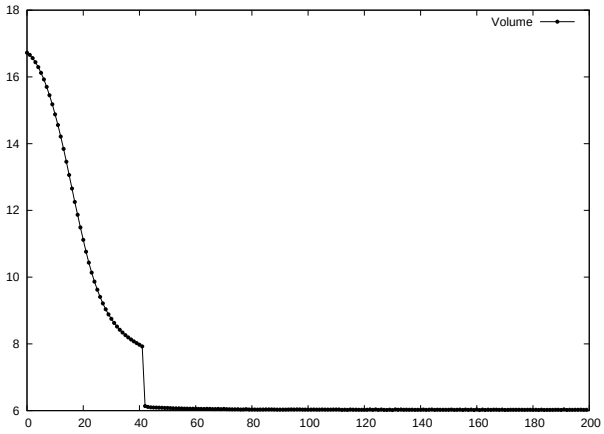
(c) Optimal shape using ersatz approach (taken from [12])

Figure 2.11: Cantilever beam, minimization of compliance 2.3

plate-like structures, like in the remeshing approach. This is because remeshing can capture thin structures precisely, leading to a more accurate shape derivative and a higher tendency of creating plates rather than bars.



(a) Compliance (2.3) vs iterations



(b) Volume vs iterations

Figure 2.12: Convergence history for the shape in Fig.2.11

In Fig.2.12, we observe that the compliance grows initially. This is because the optimization algorithm 1 seems to favor the minimization of volume at the start. Then at iteration 44, the levelset is projected (as explained in 1) so as to satisfy the target volume V_t exactly. This results in a spike in the compliance. After this iteration, the volume remains constant (because of the projection of the levelset) whereas the compliance

effectively decreases. Also, we note that there are small oscillations in the compliance in Fig.2.12. This is because of the fact that we compute the continuous shape derivative and use it for the discretized problem; rather than finding the shape derivative for the discretized problem. While remeshing helps to capture the shape Ω very well, resulting in a better calculation of the solution $\boldsymbol{u}$ and the shape derivative (compared to ersatz material approach in 2.2.1), it adds to a slight fluctuation in the objective function. This is because the change in mesh is not taken into account in the continuous shape derivative.

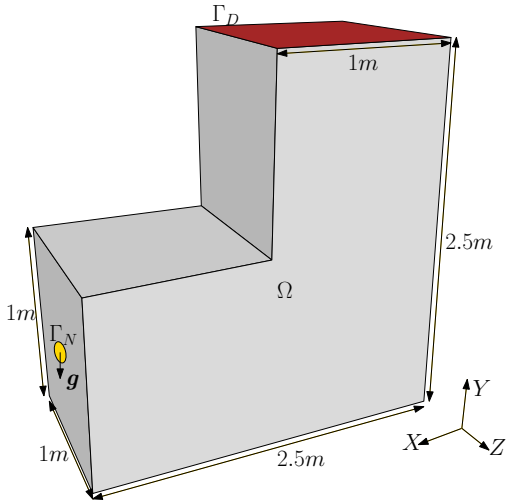
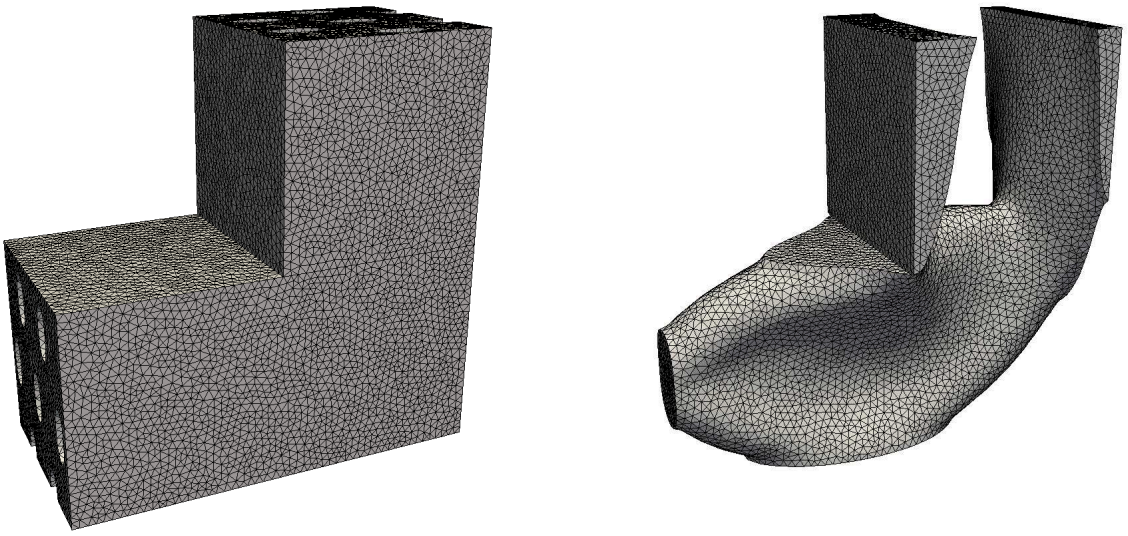


Figure 2.13: 3D L-beam boundary conditions

2.3.2 L-beam

We study an L-beam as shown in Fig.2.13. The L-beam is fixed on the topmost surface (in red) and forced on a small circular part on the left (in gold). The circular region has a radius of 0.1m. The surface force reads $\boldsymbol{g} = (0, -1, 0)$MN/m. Here again the compliance (2.3) is minimized. We show the initial meshed shape and the final meshed shape for linear elasticity in Fig.2.14 and the corresponding convergence histories are plotted in Fig.2.15. Like in the cantilever test case, the optimal shape consists of a plate-like structure at

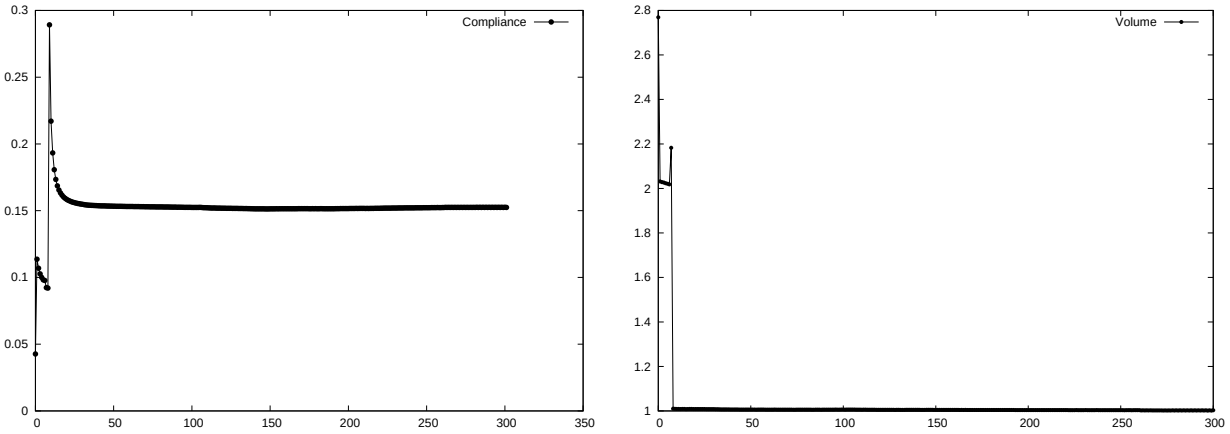


(a) Initial shape (b) Final shape

Figure 2.14: Meshed L-beam, minimization of compliance 2.3

the center. As seen in Fig.2.15, the compliance decreases well after the volume constraint is satisfied at the 15-*th* iteration. In the convergence for compliance seems to be smoother than the same convergence for the

cantilever beam in the previous section. This is because the mesh considered for the L-beam optimization is more refined than the one used for cantilever beam. A refined mesh implies the computation of a better shape derivative. This shall be discussed at length in the Section 7.2 of high performance computing Chapter 7.



(a) Compliance (2.3) vs iterations

(b) Volume vs iterations

Figure 2.15: Convergence history for the shape in Fig.2.14

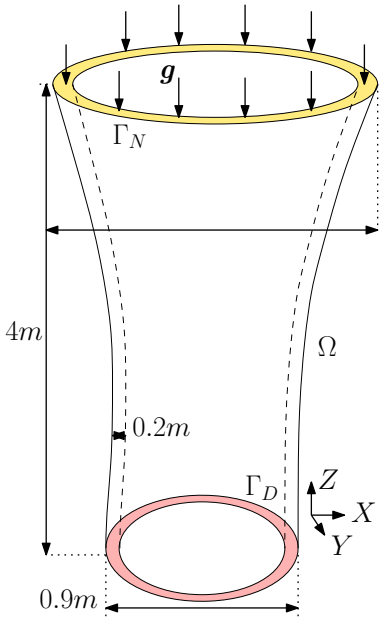


Figure 2.16: 3D column boundary conditions

2.3.3 Column

We now consider a 3D column of height $4m$ as shown in the Fig.2.16. The column is fixed on its lowermost plane (in red), forced downwards on the upper plane (in gold) with $\boldsymbol{g} = (0, 0, 1)MN/m$. A target volume $V_t = 2.5\text{m}^3$ is imposed. We show the initialized meshed shape and the final meshed shape for linear elasticity Fig.2.17. The mesh is taken sufficiently fine such that there are at least 3 mesh elements in the column thickness. As seen in the final shape in Fig.2.17, the optimal shape has curved bars, instead of strictly vertical bars. This is because of the slight curvature we have given to the design space of the column.

(a) Initial shape (b) Final shape

Figure 2.17: Meshed column, minimization of compliance 2.3

Chapter 3

Contact Mechanics

Contents

Contact mechanics is the study of the deformation of two bodies subjected to contact. Contact of sufficiently rough surfaces ensures locomotion of vehicles and plays a crucial role in mechanical engineering. Typical applications where contact mechanics plays crucial role include design of tires, braking systems, wear and tear of machines, mechanical linkages, etc.

Several attempts have been made to model contact over the last few centuries. Earlier attempts date back to the Greek civilization when the cause and mitigation of friction was studied. The Greeks were

aware of differences between static and kinetic friction with Themistius stating in 350 A.D. that "it is easier to further the motion of a moving body than to move a body at rest". Further developments date back to the 16^{th} century when Leonardo Da Vinci studied the rotational resistance of axles and the mechanics of screw threads and derived the classic laws of sliding friction. Friction modeling was later developed by Charles-Augustin de Coulomb in the 18^{th} century. Most of the attempts were directed towards characterizing the friction force applied on an object sliding on a rigid body. Since then, various friction laws have been proposed, like Coulomb law, Stokes law, Norton-Hoff law, normal compliance model, etc [129].

One big difficulty with modeling contact is that, before the loads are applied, the contact surface is not known. In 1933 Signorini wrote the contact problem in linearized elasticity for small strains which he called "a problem with ambiguous boundary conditions". The existence and uniqueness of the solution to this problem was given by one of his students Gaetano Fichera and is now known as the Signorini problem [76], which models a sliding contact (without friction). The equations governing the mechanical system with contact take the form of a variational inequality which highlights its non linearity and the resulting computational difficulties. Adding friction to Signorini problem potentially leads to mathematically ill-posed problems.

Ever since, there has been significant progress in the contact mechanics, both from a theoretical and a numerical perspective. The contact mechanics problem in its variational inequality form has been studied widely [104, 72, 167, 168]. The contact problem without friction is known to be a well-posed problem [168]. The contact problem with Coulomb friction can be shown to be well-posed if the friction coefficient is small [72]. For the numerical resolution of the contact problem, several attempts have been made: using penalization [105], Lagrange multipliers [178], Augmented Lagrangian method [163], Nitsche's method [49], primal-dual active set method [95], etc.

Shape and topology optimization in contact mechanics has received significant attention over the past two decades. Most attempts in topology optimization are based out of density methods, where the shape is approximated by a density function. Topology optimization for linear elasticity with unilateral contact was studied using SIMP [174]. The same was considered with local failure constraints [111]. In order to improve the numerical implementation of contact boundary conditions using SIMP, a mortar finite element approach was considered [173]. Nonlinearities arising from large-deformation were incorporated using SIMP [37]. One common feature in all the previous works is the representation of the mechanical properties (like the Young's modulus) using the density function, raised to a certain exponent. If the optimal shape has intermediate densities at the contact boundary, the numerical solution might feature artificial contact.

Several approaches that do not rely on density methods were also studied. The shape optimization for contact was carried out by parametrizing the shape boundary using B-splines [75]. Compliant mechanisms (where the shape is represented by a truss), designed to gain rigidity using contact were studied and designed [120]. Level-set approaches to optimize topology using X-FEM approach have also been studied for unilateral sliding contact with linear elasticity [113], bilateral contact for nonlinear material [114]. In all the above mentioned works on shape and topology optimization, the contact problem was discretized in space and the corresponding shape derivative was computed. However, it is well-known that the continuous contact problem is formulated using a variational inequality, which is not differentiable in the classical sense. In such a case, one may resort to a conical derivative and determine the shape derivative [171]. Unfortunately, the numerical implementation of a conical derivative is very difficult. Another possibility to find a continuous shape derivative is by penalizing the contact problem [88]. The numerical implementation of the shape derivative for the penalized contact problem is feasible and can be brought about using the level-set method [129, 128, 137].

In this chapter, we study a contact model governed by linear elasticity system for shape and topology optimization using the level-set method. The application of topology optimization using the level-set method in contact mechanics (via the penalization approach) was studied at length in the PhD thesis of A.Maury [128]. While some of the content from this thesis can be found in this chapter, the contribution of this work is two fold. First, the topology optimization with rigid contact boundary conditions is treated using remeshing. Second, the contact problem for a bolted connection is considered for topology optimization. The bolted connection between two solids shall be simplified and modeled as a spring., resulting in a "Idealized contact bolt model".

We first present the governing laws for contact model without friction (using Signorini laws) and then contact model with friction (using Coulomb's law) in Section 3.1. In Section 3.2, we put the contact laws under a variational inequality. We approximate the variational formulation using penalization and regularization rendering it shape-differentiable in Section 3.3. We demonstrate the well-posedness of the approximate model and show its convergence towards the actual model towards the end of this section. We then present

the shape derivative for the approximate formulation in Section 3.4. A formal limit analysis of the adjoint equation is also presented. We then talk about the numerical implementation, highlighting the issues involved resolution of the contact problem in Section 3.5. Finally, we show some 2D and 3D results for the contact model only in the Section 3.6. Finally, in Section 3.7 we present the idealized contact bolt model, the variational formulation, its approximation, the shape derivative and a few numerical results.

The content of the section 3.7 has been written in an article, submitted for publication:
L.Rakotondrainibe, J.Desai, P.Orval, G.Allaire. *Coupled topology optimization of structure and connections for bolted mechanical systems*
The section 3.7 contains some additional numerical results, not documented in the above article. The work in this section was done in a joint collaboration with Lalaina Rakotondrainibe during her PhD at Renault, in the framework of TOP Project. The modeling of the idealized bolt, the optimization of the bolt position and the computation of the topological derivative to add bolts was studied in her thesis [151].

3.1 Governing Laws

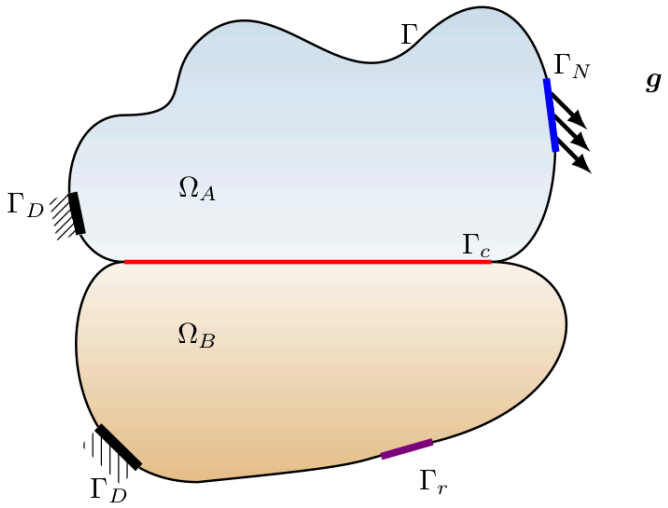


Figure 3.1: Boundary conditions of structure represented by $\Omega_A \cup \Omega_B$

The contact may be classified as unilateral and bilateral (depending on the number of directions there can be contact). In this thesis, we consider unilateral contact only, which is of two types:

- **Rigid body contact**
  Contact between a deformable body and a rigid body. In practice, rigid body is an assumption that simplifies the analysis of various complex industrial contact problems.
- **Self-contact**
  Contact between two deformable bodies. This typically means the contact of a deformable body with its self but it could also mean the contact between two disconnected bodies.

Consider two structures that shall come in unilateral contact on being subjected to an external forcing as shown in Fig.3.1. Let

- $d = 2, 3$ represent the dimension of the problem
- Ω_A, $\Omega_B \subset \mathbb{R}^d$ be two smooth bounded domains representing the two structures in Fig.3.1
- $\Omega = \Omega_A \cup \Omega_B$
- $\partial\Omega = \Gamma \cup \Gamma_c \cup \Gamma_N \cup \Gamma_D \cup \Gamma_r$, union of disjoint boundaries in Fig.3.1
- $\boldsymbol{n}$ denote the outward normal to $\partial\Omega$

- Γ_D denote the Dirichlet boundary
- Γ_c denote the zone of unilateral contact, composed of two coincident surfaces Γ_c^+ and Γ_c^-, with outward normals $\boldsymbol{n}^+$ and $\boldsymbol{n}^-$ respectively
- Γ denote the traction-free boundary
- Γ_r denote the rigid contact boundary
- $\boldsymbol{g} : \Gamma_N \to \mathbb{R}^d$ represent a surface force applied on Γ_N
- $\boldsymbol{f} : \Omega \to \mathbb{R}^d$ represent a body force applied on Ω
- Ω_A and Ω_B be composed of linear isotropic elastic materials with Lamé coefficients μ_A, λ_A and $\mu_B,\ \lambda_B$ respectively
- $\boldsymbol{u} : \Omega \to \mathbb{R}^d$ represent the displacement vector field in the two structures
- $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}(\nabla \boldsymbol{u} + \nabla^T \boldsymbol{u})$ denote the strain tensor
- $\mathcal{M}_s^d$ represent the set of symmetric $d \times d$ matrices
- $\mathbb{I}$ represent the fourth-order identity tensor of dimension d
- $\mathbb{C}$ denote the Hooke's fourth-order tensor of dimension d given by

$$\mathbb{C} = \begin{cases} 2\lambda_A \mathbf{1} \otimes \mathbf{1} + 2\mu_A \mathbb{I} & \text{in } \Omega_A \\ 2\lambda_B \mathbf{1} \otimes \mathbf{1} + 2\mu_B \mathbb{I} & \text{in } \Omega_B \end{cases}. \tag{3.1}$$

- $\boldsymbol{\sigma} : \Omega \to \mathcal{M}_s^d$ denote the stress tensor, given by the constitutive relation

$$\boldsymbol{\sigma} = \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \tag{3.2}$$

- $\mu \in \mathbb{R}^+$ be the coefficient of friction
- subscript t denote the tangential of a vector $\boldsymbol{\tau} \in \mathbb{R}^d$ on the boundary $\partial\Omega$, given by

$$\boldsymbol{\tau}_t = \boldsymbol{\tau} - (\boldsymbol{\tau} \cdot \boldsymbol{n})\boldsymbol{n}$$

- For a given displacement $\boldsymbol{u}$, and a boundary $S \subset \partial\Omega$, the normal force magnitude $\mathcal{F}_S(\boldsymbol{u})$ and the tangential force vector $\mathcal{T}_S(\boldsymbol{u})$ are given by

$$\begin{aligned} \mathcal{F}_S(\boldsymbol{u}) &= (\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n}) \cdot \boldsymbol{n} \\ \mathcal{T}_S(\boldsymbol{u}) &= (\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n})_t. \end{aligned} \tag{3.3}$$

  where $\boldsymbol{n}$ is the normal to S.

The assembly Ω satisfies the following momentum balance equations

$$\begin{aligned} \operatorname{div}(\boldsymbol{\sigma}) + \boldsymbol{f} &= \mathbf{0} && \text{in } \Omega && \text{(3.4a)} \\ \boldsymbol{u} &= \mathbf{0} && \text{on } \Gamma_D && \text{(3.4b)} \\ \boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N && \text{(3.4c)} \\ \boldsymbol{\sigma} \cdot \boldsymbol{n} &= \mathbf{0} && \text{on } \Gamma && \text{(3.4d)} \end{aligned}$$

where $\boldsymbol{\sigma}$ satisfies (3.2). We present the contact boundary conditions without friction in Subsection 3.1.1 and with friction in Subsection 3.1.2.

3.1.1 Contact model without friction

The value of $\boldsymbol{u}$ on Γ_c^+ and Γ_c^- is denoted by $\boldsymbol{u}^+$ and $\boldsymbol{u}^-$, respectively. Let the jump in displacement $[\boldsymbol{u}] = \boldsymbol{u}^- - \boldsymbol{u}^+$. For the structures Ω_A and Ω_B, the Signorini contact boundary conditions are formulated assuming

- Absence of penetration:

$$\begin{aligned} \boldsymbol{u}\cdot\boldsymbol{n} &\leq 0 \qquad \text{on } \Gamma_r, \\ \text{and } [\boldsymbol{u}]\cdot\boldsymbol{n}^- &\leq 0 \qquad \text{on } \Gamma_c \end{aligned} \tag{3.5}$$

- Contact reaction force can be compressive only:

$$\begin{aligned} &\mathcal{F}_{\Gamma_r}(\boldsymbol{u}) \leq 0, \\ &\text{and } \mathcal{F}_{\Gamma_c^+}(\boldsymbol{u}) = \mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}) \leq 0 \end{aligned} \tag{3.6}$$

- Complementary relation of contact and contact-force:

$$\begin{aligned} \boldsymbol{u}\cdot\boldsymbol{n}\mathcal{F}_{\Gamma_r}(\boldsymbol{u}) &= 0 \qquad \text{on } \Gamma_r, \\ \text{and } [\boldsymbol{u}]\cdot\boldsymbol{n}^-\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}) &= 0 \qquad \text{on } \Gamma_c. \end{aligned} \tag{3.7}$$

The complementary relation implies that either there is no contact ($\boldsymbol{u}\cdot\boldsymbol{n} < 0$, $[\boldsymbol{u}]\cdot\boldsymbol{n}^- < 0$) and a zero contact force ($\mathcal{F}_{\Gamma_r}(\boldsymbol{u}) = 0$, $\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}) = 0$) or there is contact ($\boldsymbol{u}\cdot\boldsymbol{n} = 0$, $[\boldsymbol{u}]\cdot\boldsymbol{n}^- = 0$) and the contact force is non-zero. These contact boundary conditions (3.5)-(3.7) ought to be satisfied with the momentum balance equation 3.4, resulting in the frictionless contact model

$$\begin{aligned} \operatorname{div}(\boldsymbol{\sigma}) + \boldsymbol{f} &= \boldsymbol{0} && \text{in } \Omega \\ \boldsymbol{u} &= \boldsymbol{0} && \text{on } \Gamma_D \\ \boldsymbol{\sigma}\cdot\boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N \\ \boldsymbol{\sigma}\cdot\boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma \\ \boldsymbol{u}\cdot\boldsymbol{n} &\leq 0 && \text{on } \Gamma_r, \\ [\boldsymbol{u}]\cdot\boldsymbol{n}^- &\leq 0 && \text{on } \Gamma_c \\ \mathcal{F}_{\Gamma_r}(\boldsymbol{u}) &\leq 0 && \text{on } \Gamma_r, \\ \mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}) = \mathcal{F}_{\Gamma_c^+}(\boldsymbol{u}) &\leq 0 && \text{on } \Gamma_c. \end{aligned} \tag{3.8}$$

3.1.2 Contact model with friction

To incorporate the friction between the surfaces Γ_c^+ and Γ_c^-, there are several models [104, 55]: Coulomb, Tresca, Normal Compliance and Norton-Hoff. In this thesis, we consider the Coulomb's friction model only, which makes the following hypotheses on Γ_r:

$$\begin{aligned} &|\mathcal{T}_{\Gamma_r}(\boldsymbol{u})| \leq \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u})|, \\ &|\mathcal{T}_{\Gamma_r}(\boldsymbol{u})| < \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u})| \implies \boldsymbol{u}_t = 0, \\ &|\mathcal{T}_{\Gamma_r}(\boldsymbol{u})| = \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u})| \implies \exists\zeta \geq 0,\ \boldsymbol{u}_t = -\zeta\mathcal{T}_{\Gamma_r}(\boldsymbol{u}), \end{aligned} \tag{3.9}$$

where $|\cdot|$ denotes the Euclidean norm. Similarly, the same model hypothesizes that on the surface Γ_c,

$$\begin{aligned} &\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}) = \mathcal{F}_{\Gamma_c^+}(\boldsymbol{u}) \\ &\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u}) = -\mathcal{T}_{\Gamma_c^+}(\boldsymbol{u}) \\ &|\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u})| \leq \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u})|, \\ &|\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u})| < \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u})| \implies [\boldsymbol{u}]_t = 0, \\ &|\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u})| = \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u})| \implies \exists\zeta \geq 0,\ [\boldsymbol{u}]_t = -\zeta\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u}). \end{aligned} \tag{3.10}$$

Again, these friction laws ought to be satisfied along with the contact boundary conditions (3.5)-(3.7) and momentum balance equation 3.4, resulting in the friction contact model:

$$\begin{aligned}
\operatorname{div}(\boldsymbol{\sigma}) + \boldsymbol{f} &= \boldsymbol{0} && \text{in } \Omega \\
\boldsymbol{u} &= \boldsymbol{0} && \text{on } \Gamma_D \\
\boldsymbol{\sigma}\cdot\boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N \\
\boldsymbol{\sigma}\cdot\boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma \\
\boldsymbol{u}\cdot\boldsymbol{n} &\leq 0 && \text{on } \Gamma_r \\
[\boldsymbol{u}]\cdot\boldsymbol{n}^- &\leq 0 && \text{on } \Gamma_c \\
\mathcal{F}_{\Gamma_r}(\boldsymbol{u}) &\leq 0 && \text{on } \Gamma_r \\
\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u}) = -\mathcal{T}_{\Gamma_c^+}(\boldsymbol{u}),\ \mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}) &= \mathcal{F}_{\Gamma_c^+}(\boldsymbol{u}) \leq 0 && \text{on } \Gamma_c \\
|\mathcal{T}_{\Gamma_r}(\boldsymbol{u})| \leq \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u})|,\ (|\mathcal{T}_{\Gamma_r}(\boldsymbol{u})| - \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u})|)\boldsymbol{u}_t &= 0 && \text{on } \Gamma_r \\
|\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u})| \leq \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u})|,\ (|\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u})| - \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u})|[\boldsymbol{u}_t] &= 0 && \text{on } \Gamma_c
\end{aligned} \tag{3.11}$$

3.2 Variational formulation

We now derive the variational formulation for linear elasticity with contact boundary conditions, discussed in the previous chapter. It is well-known that the incorporation of the contact boundary conditions results in a variational inequality, rather than a variational equation.

The Hooke's tensor $\mathbb{C}$ is assumed to be coercive, i.e., $\exists\ c_0 > 0$ such that,

$$\mathbb{C}\boldsymbol{\xi}\colon\boldsymbol{\xi} \geq c_0|\boldsymbol{\xi}|^2 \qquad \forall \boldsymbol{\xi} \in \mathcal{M}_s^d. \tag{3.12}$$

Let V be the space of displacement $\boldsymbol{u}$

$$V = \{\boldsymbol{u} \in H^1(\Omega)^d,\ \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_D\}$$

Let $K \subset V$ be the set, where the non penetration contact condition on Γ_r and Γ_c is satisfied,

$$K = \{\boldsymbol{u} \in V : \quad \boldsymbol{u}\cdot\boldsymbol{n} \leq 0, \quad \text{on } \Gamma_r \quad \text{and} \qquad [\boldsymbol{u}]\cdot\boldsymbol{n}^- \leq 0 \text{ on } \Gamma_c\}.$$

Quite easily, one can verify that K is a convex subset of V. We introduce a bilinear form $a : V \times V \to \mathbb{R}$ defined as

$$a(\boldsymbol{u},\boldsymbol{v}) = \int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\colon\boldsymbol{\varepsilon}(\boldsymbol{v})\ dx, \tag{3.13}$$

We can show that $a(\cdot,\cdot)$ is a symmetric continuous bilinear form. Using Korn's inequality, $a(\cdot,\cdot)$ can be shown to be coercive under two assumptions:

V.1 the positive definiteness of the Hooke's tensor (3.12)

V.2 the Dirichlet boundary is assumed to be present on both the subdomains Ω_A, Ω_B, i.e., $\partial\Omega_A \cap \Gamma_D \neq \emptyset$, $\partial\Omega_B \cap \Gamma_D \neq \emptyset$.

We then introduce a linear form: $l : V^* \to \mathbb{R}$

$$l(\boldsymbol{u}) = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{u}\ dx + \int_{\Gamma_N} \boldsymbol{g}\cdot\boldsymbol{u}\ ds.$$

In the next subsection, we present variational formulations for the contact model without friction and for the contact model with friction. Before moving to the variational formulations, we state a theorem from [107].

Theorem 2. *Let X be a Hilbert space, $a : X \times X \to \mathbb{R}$ be a bilinear coercive form, $f \in X^*$ (the dual of X) and K be a non-empty convex set in X. Then the problem: find $\boldsymbol{u} \in K$ such that*

$$a(\boldsymbol{u}, \boldsymbol{v}-\boldsymbol{u}) \geq \langle f, \boldsymbol{v}-\boldsymbol{u}\rangle_{X^*,X} \quad \forall \boldsymbol{v} \in K$$

admits a unique solution. Moreover, if $a(\cdot,\cdot)$ is symmetric, $\boldsymbol{u}$ is also the solution to the minimization problem: find $u \in K$ such that

$$\boldsymbol{u} = \arg\min_{\boldsymbol{v}\in K} \left(\frac{1}{2}a(\boldsymbol{v},\boldsymbol{v}) - \boldsymbol{l}(\boldsymbol{v})\right).$$

3.2.1 Contact model without friction

Lemma 2. *The solution to the frictionless contact model (3.8), $\boldsymbol{u}$ satisfies the variational inequality: $\boldsymbol{u} \in K$ and*

$$a(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) \geq l(\boldsymbol{v} - \boldsymbol{u}) \qquad \forall \boldsymbol{v} \in K. \tag{3.14}$$

Proof. For simplicity, consider the case $\Gamma_c = \emptyset$. By definition, the solution to (3.8) belongs to the space K. We multiply both sides of the Eqn.(3.4a) by $(\boldsymbol{v} - \boldsymbol{u}) \in K$, where $\boldsymbol{v} \in K$ and obtain

$$-\int_\Omega \operatorname{div}(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})) \cdot (\boldsymbol{v} - \boldsymbol{u})\, dx = \int_\Omega \boldsymbol{f} \cdot (\boldsymbol{v} - \boldsymbol{u})\, dx \quad \forall \boldsymbol{v} \in K.$$

Integrating the first term in the above by parts and applying the boundary conditions on Γ, Γ_N, Γ_r, and Γ_D, we get

$$\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \colon \boldsymbol{\varepsilon}(\boldsymbol{v} - \boldsymbol{u})\, dx - \int_{\Gamma_r} \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\boldsymbol{n} \cdot (\boldsymbol{v} - \boldsymbol{u})\, ds = \int_\Omega \boldsymbol{f} \cdot (\boldsymbol{v} - \boldsymbol{u})\, dx + \int_{\Gamma_N} \boldsymbol{g} \cdot (\boldsymbol{v} - \boldsymbol{u})\, dx \quad \forall \boldsymbol{v} \in K. \tag{3.15}$$

For the integral on Γ_r, we can simplify

$$\begin{aligned}
(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n}) \cdot (\boldsymbol{v} - \boldsymbol{u}) &= ((\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n}) \cdot \boldsymbol{n})\boldsymbol{n} \cdot (\boldsymbol{v} - \boldsymbol{u}) + (\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n})_t \cdot (\boldsymbol{v} - \boldsymbol{u}) \\
&= ((\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n}) \cdot \boldsymbol{n})\boldsymbol{n} \cdot (\boldsymbol{v} - \boldsymbol{u}) \\
&= ((\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \cdot \boldsymbol{n}) \cdot \boldsymbol{n})(\boldsymbol{v} \cdot \boldsymbol{n}) \geq 0, \quad \forall \boldsymbol{v} \in K.
\end{aligned}$$

Thus, Eqn.(3.15) simplifies to the following variational inequality

$$\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) \colon \boldsymbol{\varepsilon}(\boldsymbol{v} - \boldsymbol{u})\, dx \geq \int_\Omega \boldsymbol{f} \cdot (\boldsymbol{v} - \boldsymbol{u})\, dx + \int_{\Gamma_N} \boldsymbol{g} \cdot (\boldsymbol{v} - \boldsymbol{u})\, dx \quad \forall \boldsymbol{v} \in K. \tag{3.16}$$

Using the definition of the bilinear form $a(\cdot, \cdot)$ and the linear form l, we get (3.14) from the above. Note that in the derivation of the above (3.16), we only used two Signorini conditions ((3.6) and (3.7)) while the condition (3.5) was not used, since the solution $\boldsymbol{u}$ was assumed to satisfy it by definition. The above derivation can be generalized to the case where $\Gamma_c \neq \emptyset$. The inequality obtained in this case is the same as (3.14). □

The existence and uniqueness of solution to the inequality (3.14) can be shown easily using Theorem 2.

Theorem 3. *The inequality (3.14) admits a unique solution $\boldsymbol{u} \in K$. Moreover, the solution $\boldsymbol{u}$ is the unique minimizer of the problem*

$$\min_{\boldsymbol{v} \in K} \left(\frac{1}{2} a(\boldsymbol{v}, \boldsymbol{v}) - l(\boldsymbol{v}) \right).$$

Proof. It suffices to take $X = V$ and $f = l$ in Theorem 2 to show the existence of unique $\boldsymbol{u} \in K$. □

3.2.2 Contact model with friction

Lemma 3. *The solution $\boldsymbol{u}$ to the friction contact model (3.11) satisfies: $\boldsymbol{u} \in K$ such that*

$$a(\boldsymbol{u}, \boldsymbol{v} - \boldsymbol{u}) + j_c(\boldsymbol{u}, \boldsymbol{v}) - j_c(\boldsymbol{u}, \boldsymbol{u}) \geq l(\boldsymbol{v} - \boldsymbol{u}) \quad \forall \boldsymbol{v} \in K, \tag{3.17}$$

where

$$j_c(\boldsymbol{u}, \boldsymbol{v}) = \mu \left(\int_{\Gamma_r} |\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\boldsymbol{n} \cdot \boldsymbol{n}|\ |\boldsymbol{v}_t| ds + \int_{\Gamma_c} |\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\boldsymbol{n} \cdot \boldsymbol{n}|\ |[\boldsymbol{v}]_t| ds \right) \tag{3.18}$$

Proof. Performing a similar computation in proof to Lemma 2 and using the friction boundary conditions, we get the inequality (3.17). □

The inequality (3.17) cannot be recast as a minimization problem (because of term $j_c(\boldsymbol{u}, \boldsymbol{v})$). If only the term $j_c(\cdot, \cdot)$ was linear in its second argument, (3.17) would have been equivalent to an optimization problem. One can however show existence and uniqueness of a solution $\boldsymbol{u} \in K$ [89, 71], by approximating the inequality (3.17) by penalizing the constraint $\boldsymbol{u} \in K$. The sequence of penalized solutions is shown to

converge weakly. The limit of the sequence satisfies the equation (3.17) if the friction coefficient μ is assumed to be small and bounded by [71]

$$\mu \leq C = \begin{cases} \dfrac{\sqrt{3-4\nu}}{2-2\nu}, & d = 2, \\ \sqrt{\dfrac{3-4\nu}{4-4\nu}}, & d = 3, \end{cases} \tag{3.19}$$

where ν is the Poisson's ratio and the material is assumed to be isotropic.

3.3 Approximation via penalization and regularization

We need to determine the shape derivative for the contact problem (3.14), (3.17) and the idealized bolt model (3.59)-(3.60) by differentiating them with respect to their respective solution. Since these problems are inequalities, defined on a convex set K and their respective solutions belong to the convex set K, the differentiation must be brought about only along the directions $\boldsymbol{\varphi} \in K$. Here, the classical notion of a derivative is lost and one ought to resort to the conical derivative. The conical derivative has been derived in the context of the static contact problem with and without friction [171]. Determining a conical derivative is mathematically involved and numerically implementing a conical derivative is difficult.

One possibility to circumvent a conical derivative is to approximate the inequalities (3.14) and (3.17) using penalization. The penalization method in the context of contact mechanics has been elaborated in [72]. The idea of the penalization method is to search for a solution in V by penalizing the constraints (3.6) and (3.7), thereby forcing the solution to belong to K. The penalization of the contact boundary condition converts the inequality to equality, making the contact problem easily differentiable. We introduce a penalization constant $0 < \epsilon \ll 1$ and define a function $j_\epsilon : V \to \mathbb{R}$

$$j_\epsilon(\boldsymbol{u}) = \frac{1}{\epsilon}\left(\int_{\Gamma_r}\int_0^{\boldsymbol{u}\cdot\boldsymbol{n}} \max(t,0)dt\ ds + \int_{\Gamma_c}\int_0^{[\boldsymbol{u}]\cdot\boldsymbol{n}^-} \max(t,0)dt\ ds\right). \tag{3.20}$$

In the above, one might as well consider a regularized form of the max function. But we use the genuine max function for all of our numerical simulations. The term $j_\epsilon(\boldsymbol{u})$ penalizes the constraints (3.5)-(3.7). We present the penalization approach for the contact model with and without friction, followed by the same for the idealized bolt model.

3.3.1 Contact model without friction

The penalized form of the variational inequality (3.14) reads: find $\boldsymbol{u}_\epsilon \in V$ such that

$$\boldsymbol{u}_\epsilon = \arg\min_{\boldsymbol{v}\in V}\left(\frac{1}{2}a(\boldsymbol{v},\boldsymbol{v}) + j_\epsilon(\boldsymbol{v}) - l(\boldsymbol{v})\right). \tag{3.21}$$

Theorem 4. *There exists a unique solution $\boldsymbol{u}_\epsilon \in V$ to the minimization problem (3.21). Moreover, as ϵ tends to zero, the sequence of solutions $\boldsymbol{u}_\epsilon$ converges strongly to the solution of the problem (3.14) $\boldsymbol{u}$ in V.*

Proof. Given the convexity of $j_\epsilon(\cdot)$, the functional on the right in (3.21) is clearly convex. Using the convexity and the fact that it is a minimization problem on the entire space V, one can state that there exists a unique solution $\boldsymbol{u}_\epsilon \in V$. The convergence $\boldsymbol{u}_\epsilon \longrightarrow \boldsymbol{u}$ in V as $\epsilon \longrightarrow 0$ is really classical (see [80], Chapter 1, Theorem 7.1). □

The biggest advantage of the penalization is the fact the problem can now be written in an equation form: find $\boldsymbol{u}_\epsilon \in V$

$$a(\boldsymbol{u}_\epsilon,\boldsymbol{v}) + j'_\epsilon(\boldsymbol{u}_\epsilon)\cdot\boldsymbol{v} = l(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V, \tag{3.22}$$

where the derivative j'_ϵ along the direction $\boldsymbol{v}$ is defined as

$$j'_\epsilon(\boldsymbol{u}_\epsilon)\cdot\boldsymbol{v} = \frac{1}{\epsilon}\left(\int_{\Gamma_r}\max(\boldsymbol{u}_\epsilon\cdot\boldsymbol{n},0)\boldsymbol{v}\cdot\boldsymbol{n}\,ds + \int_{\Gamma_c}\max([\boldsymbol{u}_\epsilon]\cdot\boldsymbol{n}^-,0)[\boldsymbol{v}]\cdot\boldsymbol{n}\,ds\right)$$

This is the desired variational equation, the approximation of the inequality (3.14). We note that the contact boundary conditions are now reformulated as the following Neumann boundary conditions

$$\begin{aligned}(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\cdot\boldsymbol{n})\cdot\boldsymbol{n} &= -\frac{1}{\epsilon}\max(\boldsymbol{u}_\epsilon\cdot\boldsymbol{n},0) \qquad \text{on } \Gamma_r\\ (\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\cdot\boldsymbol{n})\cdot\boldsymbol{n} &= -\frac{1}{\epsilon}\max([\boldsymbol{u}_\epsilon]\cdot\boldsymbol{n}^-,0) \qquad \text{on } \Gamma_c.\end{aligned} \tag{3.23}$$

The Theorem 4 shows convergence of the solution $\boldsymbol{u}_\epsilon$ but says nothing about the rate of convergence with respect to ϵ. This is given in [42], Theorem 4.4, where the authors show that (using our notations),

$$a(\boldsymbol{u}-\boldsymbol{u}_\epsilon,\boldsymbol{u}-\boldsymbol{u}_\epsilon)\leq\frac{\epsilon}{2}\left\|(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\cdot\boldsymbol{n})\cdot\boldsymbol{n}\right\|^2_{L^2(\Omega_c^+\cup\Omega_c^-\cup\Omega_r)}.$$

This implies that

$$\|\boldsymbol{u}-\boldsymbol{u}_\epsilon\|_V\leq C_{\boldsymbol{u}}\sqrt{\epsilon}, \tag{3.24}$$

where $C_{\boldsymbol{u}}$ is constant that depends on $\boldsymbol{u}$. This roughly translates to an error estimation of $\mathcal{O}(\sqrt{\epsilon})$.

3.3.2 Contact model with friction

Unlike the inequality (3.14), the inequality (3.17) cannot be cast to an optimization problem. Because of this, one cannot apply penalization via a minimization problem like in the previous subsection, using (3.21). As seen in the previous subsection, adding the Neumann boundary conditions (3.23) (or the term j'_ϵ) forces the solution $\boldsymbol{u}_\epsilon$ to satisfy the contact boundary conditions (3.5)-(3.7). Hence, we do the same here for (3.17) and obtain: find $\boldsymbol{u}_\epsilon\in K$

$$a(\boldsymbol{u}_\epsilon,\boldsymbol{v}-\boldsymbol{u}_\epsilon)+j'_\epsilon(\boldsymbol{u}_\epsilon)\cdot(\boldsymbol{v}-\boldsymbol{u}_\epsilon)+j_{c,\epsilon}(\boldsymbol{u}_\epsilon,\boldsymbol{v})-j_{c,\epsilon}(\boldsymbol{u}_\epsilon,\boldsymbol{u}_\epsilon)\geq l(\boldsymbol{v}-\boldsymbol{u}_\epsilon)\quad\forall\boldsymbol{v}\in V, \tag{3.25a}$$

$$\text{where } j_{c,\epsilon}=\frac{\mu}{\epsilon}\left(\int_{\Gamma_r}\max(\boldsymbol{u}_\epsilon\cdot\boldsymbol{n},0)|\boldsymbol{v}_t|ds+\int_{\Gamma_c}\max([\boldsymbol{u}_\epsilon]\cdot\boldsymbol{n}^-,0)|[\boldsymbol{v}]_t|ds\right) \tag{3.25b}$$

is obtained by injecting (3.23) in (3.18). In order to convert the above to an equation, we replace $\boldsymbol{v}$ by $\boldsymbol{u}_\epsilon\pm\eta\boldsymbol{v}$ and pass η to 0. This demands the Gâteaux differentiability of the integrand in (3.25b). In order to ensure so, we must regularize the Euclidean norm, one possible regularization being [72], $0<\eta\ll 1$, $\mathcal{N}_\eta:\mathbb{R}\to\mathbb{R}$ given by

$$\mathcal{N}_\eta(x):\ x\to\begin{cases}|x| & |x|\geq\eta,\\ -\dfrac{|x|^4}{8\eta^3}+\dfrac{3|x|^2}{4\eta}+\dfrac{3\eta}{8} & |x|<\eta,\end{cases}$$

with the gradient

$$\nabla\mathcal{N}_\eta(x)=\beta(x)\frac{x}{|x|},\ \beta(x)=\begin{cases}1 & |x|\geq\eta,\\ -\dfrac{|x|^3}{2\eta^3}+\dfrac{3|x|}{2\eta} & |x|<\eta.\end{cases}$$

Replacing the Euclidean norm by $\mathcal{N}_\eta$ in (3.25), $\boldsymbol{v}$ by $\boldsymbol{u}_\epsilon\pm\eta\boldsymbol{v}$ and passing η to 0, we get the penalized contact problem: find $\boldsymbol{u}_{\epsilon,\eta}\in V$ such that

$$a(\boldsymbol{u}_{\epsilon,\eta},\boldsymbol{v})+j'_\epsilon(\boldsymbol{u}_{\epsilon,\eta})\cdot\boldsymbol{v}+j_{\epsilon,\eta}(\boldsymbol{u}_{\epsilon,\eta},\boldsymbol{v})=l(\boldsymbol{v})\quad\forall v\in V,\quad\text{where} \tag{3.26}$$

$$j_{\epsilon,\eta}(\boldsymbol{u},\boldsymbol{v})=\frac{\mu}{\epsilon}\left(\int_{\Gamma_r}\max(\boldsymbol{u}\cdot\boldsymbol{n},0)\mathcal{N}'_\eta(\boldsymbol{u}_t)\cdot\boldsymbol{v}_t ds+\int_{\Gamma_c}\max([\boldsymbol{u}]\cdot\boldsymbol{n}^-,0)\mathcal{N}'_\eta([\boldsymbol{u}]_t)\cdot[\boldsymbol{v}]_t ds\right). \tag{3.27}$$

The above is to be shown to be well-posed in Theorem 1 [71]. The approximate frictionless contact formulation (3.22) can evidently be retrieved from the above by substituting $\mu=0$. Hence, we treat only the formulation (3.26) for the shape derivative computation in the next section. The equation (3.26) shall be called the state equation and the solution $\boldsymbol{u}_{\epsilon,\eta}$ to it, the state solution.

3.4 Shape derivative

In this chapter we determine the shape derivative for the penalized regularized friction formulation (3.26). The shape derivative computation in Subsection (3.4.1) involves an adjoint problem that depends on the penalization regularization parameters. In Subsection 3.36, we formally study what happens when these parameters tend to zero.

We consider an objective function of the following form

$$J(\Omega) = \int_{\Omega} m(\boldsymbol{u}_{\epsilon,\eta}(\Omega))dx + \int_{\Gamma_N} p(\boldsymbol{u}_{\epsilon,\eta}(\Omega))ds, \tag{3.28}$$

where $\boldsymbol{u}_{\epsilon,\eta}(\Omega)$ is the solution to the problem (3.26), $m(\cdot)$, $p(\cdot)$ are assumed to be $\mathcal{C}^1$ smooth and m', p' are assumed to have a linear growth rate.

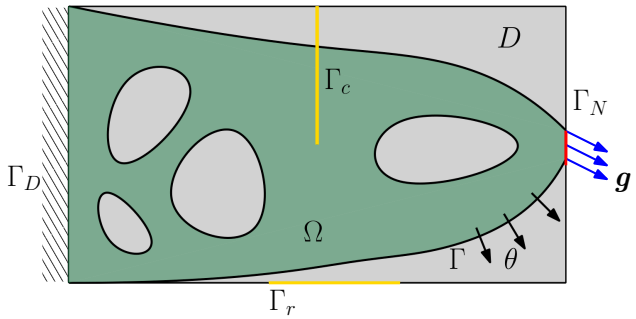


Figure 3.2: Design domain D and the shape Ω

As mentioned in Subsection 2.1.2, for all the chapters in this thesis, only Γ is allowed to move and all other boundaries (Γ_D, Γ_N) are assumed non-optimizable. As shown in Fig.3.2, even Γ_c and Γ_r as fixed. In [128], the case where the contact boundaries Γ_c and Γ_r are assumed optimizable has been studied. As seen in Fig.3.2, only Γ is allowed to move.

3.4.1 Shape derivative computation

Only in this subsection, we simplify notations by dropping the indices ϵ, η in $\boldsymbol{u}_{\epsilon,\eta}$ and denoting it by $\boldsymbol{u}$. The regularized nonlinearity $j_{\epsilon,\eta}$ is Gâteaux differentiable. The following theorem gives an expression of the shape derivative for minimizing the objective function 3.28. The theorem assumes the adjoint equation to be well-posed, which is later shown to be well-posed for the frictionless case ($\mu = 0$). We assume that $l(\cdot)$, $m(\cdot)$ are $\mathcal{C}^1$ and coercive, Γ_m is the optimizable subset of boundary $\partial\Omega$.

Theorem 5. *Let Ω_A and Ω_B be two smooth bounded domains sharing a boundary Γ_c and $\Omega = \Omega_A \cup \Omega_B$. Let $\boldsymbol{f} \in H^1(\mathbb{R}^d)^d$, $\boldsymbol{g} \in H^2(\mathbb{R}^d)^d$ and $\boldsymbol{u}(\Omega) \in V$ be the solution to (3.26). Then the shape derivative of $J(\Omega)$ (3.28) along $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, is given by*

$$J'(\Omega)\boldsymbol{\theta} = \int_{\Gamma} \boldsymbol{\theta}\cdot\boldsymbol{n}\,(m(\boldsymbol{u}) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{f}\cdot\boldsymbol{v})\,ds, \tag{3.29}$$

where $\boldsymbol{v} \in V$ is the solution to the adjoint problem:

$$a(\boldsymbol{\varphi}, \boldsymbol{v}) + (j_\epsilon''(\boldsymbol{u})\cdot\boldsymbol{\varphi})\cdot\boldsymbol{v} + \partial_{\boldsymbol{u}} j_{\epsilon,\eta}(\boldsymbol{u}, \boldsymbol{v})\cdot\boldsymbol{\varphi} = -\int_{\Omega} \partial_{\boldsymbol{u}} m(\boldsymbol{u})\cdot\boldsymbol{\varphi}\,dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}} p(\boldsymbol{u})\cdot\boldsymbol{\varphi}\,ds \qquad \forall \boldsymbol{\varphi} \in V, \tag{3.30}$$

where the derivatives of $j'_\epsilon, j_{\epsilon,\eta}$ are defined in (3.33).

Proof. Introduce a space

$$\tilde{V} = \{\boldsymbol{v} \in H^1(\mathbb{R}^d)^d : \boldsymbol{v} = \boldsymbol{0} \qquad \text{on } \Gamma_D\}. \tag{3.31}$$

We use the fast derivation technique of Céa [53] by constructing the Lagrangian:

$$\mathcal{L}(\Omega, \tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) = \int_{\Omega} m(\tilde{\boldsymbol{u}})\ dx + \int_{\Gamma_N} p(\tilde{\boldsymbol{u}})ds + a(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) + j'_\epsilon(\tilde{\boldsymbol{u}})\cdot\boldsymbol{v} + j_{\epsilon,\eta}(\tilde{\boldsymbol{u}}, \boldsymbol{v}) - l(\tilde{\boldsymbol{v}}), \tag{3.32}$$

where $\tilde{\boldsymbol{u}} \in \tilde{V}$ is the state variable and $\tilde{\boldsymbol{v}} \in \tilde{V}$ is the adjoint variables, both of which are defined on the full space $\mathbb{R}^d$. Since Γ_D is non-optimizable, the variables $\boldsymbol{u}, \boldsymbol{v}$ are independent of Ω.

The optimality conditions applied to the Lagrangian in (3.32) results in the determination of the state equation (3.26), adjoint equation and the shape derivative. First we apply optimality with respect to the adjoint variables $\boldsymbol{v}$ along $\boldsymbol{\varphi} \in \tilde{V}$ at $\tilde{\boldsymbol{u}} = \boldsymbol{u}$ and obtain

$$a(\boldsymbol{u}, \boldsymbol{\varphi}) + j'_\epsilon(\boldsymbol{u}) \cdot \boldsymbol{\varphi} + j_{\epsilon,\eta}(\boldsymbol{u}, \boldsymbol{\varphi}) = l(\boldsymbol{\varphi}) \quad \forall \boldsymbol{\psi} \in \tilde{V}.$$

Since all integrals in the above are defined only on Ω, we can replace $\tilde{V}$ by V and obtain (3.26). Second, we apply the optimality condition on (3.32) with respect to the state variables $\tilde{\boldsymbol{u}}$ at $(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) = (\boldsymbol{u}, \boldsymbol{v})$ and obtain

$$a(\boldsymbol{\psi}, \boldsymbol{v}) + (j''_\epsilon(\boldsymbol{u}) \cdot \boldsymbol{\psi}) \cdot \boldsymbol{v} + \partial_{\boldsymbol{u}} j_{\epsilon,\eta}(\boldsymbol{u}, \boldsymbol{v}) \cdot \boldsymbol{\psi} = -\int_\Omega \partial_{\boldsymbol{u}} m(\boldsymbol{u}) \cdot \boldsymbol{\psi}\, dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}} p(\boldsymbol{u}) \cdot \boldsymbol{\psi}\, ds \qquad \forall \boldsymbol{\varphi} \in \tilde{V}$$

$$\text{where } (j''_\epsilon(\boldsymbol{u}) \cdot \boldsymbol{\psi}) \cdot \boldsymbol{v} = \frac{1}{\epsilon}\left(\int_{\Gamma_r} \mathcal{H}(\boldsymbol{u}\cdot\boldsymbol{n})(\boldsymbol{\psi}\cdot\boldsymbol{n})(\boldsymbol{v}\cdot\boldsymbol{n})\, ds + \int_{\Gamma_c} \mathcal{H}([\boldsymbol{u}]\cdot\boldsymbol{n})([\boldsymbol{\psi}]\cdot\boldsymbol{n}^-)([\boldsymbol{v}]\cdot\boldsymbol{n}^-)\, ds\right), \tag{3.33a}$$

$$\begin{aligned}\partial_{\boldsymbol{u}} j_{\epsilon,\eta}(\boldsymbol{u}, \boldsymbol{v}) \cdot \boldsymbol{\psi} = {} & \frac{\mu}{\epsilon}\left(\int_{\Gamma_r} \mathcal{H}(\boldsymbol{u}\cdot\boldsymbol{n})(\boldsymbol{\psi}\cdot\boldsymbol{n})\mathcal{N}'_\eta(\boldsymbol{u}_t)\cdot\boldsymbol{v}_t ds + \int_{\Gamma_c} \mathcal{H}([\boldsymbol{u}]\cdot\boldsymbol{n})([\boldsymbol{\psi}]\cdot\boldsymbol{n}^-)\mathcal{N}'_\eta([\boldsymbol{u}]_t)\cdot[\boldsymbol{v}]_t ds\right) \\ & + \frac{\mu}{\epsilon}\left(\int_{\Gamma_r} \max(\boldsymbol{u}\cdot\boldsymbol{n}, 0)(\mathcal{N}''_\eta(\boldsymbol{u}_t)\cdot\boldsymbol{\psi}_t)\cdot\boldsymbol{v}_t ds + \int_{\Gamma_c} \max([\boldsymbol{u}]\cdot\boldsymbol{n}, 0)(\mathcal{N}''_\eta([\boldsymbol{u}]_t)\cdot[\boldsymbol{\psi}]_t)\cdot[\boldsymbol{v}]_t ds\right)\end{aligned} \tag{3.33b}$$

and $\mathcal{H}$ denotes the Heaviside function. Replacing $\tilde{V}$ by V in the above, we get the adjoint equation (3.30). Then, we observe that (using the state equation (3.26))

$$J(\Omega) = \mathcal{L}(\Omega, \boldsymbol{u}(\Omega), \tilde{\boldsymbol{v}}).$$

Using the chain differentiation rule on the above, we get

$$J'(\Omega)(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}(\Omega, \boldsymbol{u}(\Omega), \tilde{\boldsymbol{v}})(\boldsymbol{\theta}) + \frac{\partial \mathcal{L}}{\partial \boldsymbol{u}}(\Omega, \boldsymbol{u}(\Omega), \tilde{\boldsymbol{v}})\frac{\boldsymbol{u}(\Omega)}{\partial \Omega}(\boldsymbol{\theta}).$$

Substituting $\tilde{\boldsymbol{v}} = \boldsymbol{v}(\Omega)$, using the adjoint equation (3.30) and applying the Lemma 1, we obtain the expression for shape derivative in (3.29). □

In the previous theorem, we assumed the adjoint equation (3.30) to be well-posed (which allowed us to use Céa's method). The adjoint solution $\boldsymbol{v}$ does depend on the parameters ϵ and η, the indices being dropped here to simplify notations. Showing the adjoint equation to be well-posed is not straight-forward as it contains the adjoint of linearization of the state equation (3.26), which cannot be shown to be the minimization of a convex function. However, for the frictionless case ($\mu = 0$), the state equation (3.26) is effectively the minimization of a convex function (see (3.21)), and the corresponding adjoint equation can be shown to be well-posed. We show this well-posedness in the the following theorem.

Theorem 6. *Under the assumptions* ***V.1***-***V.2*** *and for* $\mu = 0$, *the adjoint equation* (3.30) *admits a unique solution* $\boldsymbol{v} \in V$.

Proof. The proof is rather straightforward. Since the functions $m(\cdot), p(\cdot)$ are assumed to be $\mathcal{C}^1$ smooth with linear growth, the right hand side of the equation (3.30) is continuous. The bilinear operator on the left hand side can be written for all $\boldsymbol{\phi} \in V$ as

$$a(\boldsymbol{\phi}, \boldsymbol{\phi}) \leq a(\boldsymbol{\phi}, \boldsymbol{\phi}) + \frac{1}{\epsilon}\left(\int_{\Gamma_r} \mathcal{H}(\boldsymbol{u}\cdot\boldsymbol{n})(\boldsymbol{\phi}\cdot\boldsymbol{n})^2 ds + \int_{\Gamma_c} \mathcal{H}([\boldsymbol{u}]\cdot\boldsymbol{n})([\boldsymbol{\phi}]\cdot\boldsymbol{n}^-)^2 ds\right)$$

The form $a(\cdot,\cdot)$ is coercive under the assumptions **V.1**-**V.2**. Hence the bilinear operator on the left hand side can be shown to be coercive and we can apply Lax-Milgram theorem to (3.30) to state that there exists a unique $\boldsymbol{v} \in V$. □

3.4.2 Formal limit analysis of the adjoint equation (3.69) for $\mu = 0$

We derived the adjoint equation for the penalized regularized problem (3.26) in the previous subsection. However, it remains unclear as to what does the adjoint so derived correspond to as the limit $\epsilon, \eta \longrightarrow 0$. In fact, at the limit $\epsilon = 0, \eta = 0$ we obtain the problem (3.17), that is non-smooth and indeed not differentiable in the classical sense. One instead needs to compute the conical derivative. The conical derivative has been well studied for contact mechanics [171]. This is very technical and involves several theorems. Here, we wish to present a shape-derivative by a formal computation, bypassing very technical analysis. In order to do so, we must reformulate the contact problem (3.26). For this reformulation, there are two possibilities:

- Add more constraints (to be specified later) to the space $V \times V$ and convert (3.26) to an equation
- Express the contact problem (3.26) using Lagrange multipliers.

We only consider the first possibility in this subsection. A similar approach of converting the variational inequality of frictionless contact (3.16) to variational equation has been studied [108].

To simplify the presentation, we assume that $\mu = 0$ (without this assumption, the analysis is very complicated). We split the contact boundary $\Gamma_c \cup \Gamma_r = \Gamma_{\boldsymbol{u}}^1 \cup \Gamma_{\boldsymbol{u}}^2$ such that

$$\Gamma_{\boldsymbol{u}}^1 = \{\boldsymbol{x} \in \Gamma_r : \ \boldsymbol{u} \cdot \boldsymbol{n}(\boldsymbol{x}) < 0 \quad \text{a.e. on } \Gamma_r\} \cup \{\boldsymbol{x} \in \Gamma_c : \ [\boldsymbol{u}] \cdot \boldsymbol{n}^-(\boldsymbol{x}) < 0 \quad \text{a.e. on } \Gamma_c\},$$
$$\Gamma_{\boldsymbol{u}}^2 = \{\boldsymbol{x} \in \Gamma_r : \ \boldsymbol{u} \cdot \boldsymbol{n}(\boldsymbol{x}) = 0 \quad \text{a.e. on } \Gamma_r\} \cup \{\boldsymbol{x} \in \Gamma_c : \ [\boldsymbol{u}] \cdot \boldsymbol{n}^-(\boldsymbol{x}) = 0 \quad \text{a.e. on } \Gamma_c\}.$$

where $\boldsymbol{u}$ is the solution to inequality (3.16). For this solution $\boldsymbol{u}$, we minimize the same functional as in the previous subsection

$$J(\Omega) = \int_\Omega m(\boldsymbol{u}(\Omega))dx + \int_{\Gamma_N} p(\boldsymbol{u}(\Omega))ds. \tag{3.34}$$

Introduce a space $K_{\boldsymbol{u}}$ defined as

$$K_{\boldsymbol{u}} = \{\boldsymbol{v} \in V : \boldsymbol{v} \cdot \boldsymbol{n}(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Gamma_{\boldsymbol{u}}^2 \cap \Gamma_r, \quad [\boldsymbol{v}] \cdot \boldsymbol{n}^-(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Gamma_{\boldsymbol{u}}^2 \cap \Gamma_c\}.$$

Multiplying the strong form (3.4a) by $\boldsymbol{v} \in K_{\boldsymbol{u}}$ and performing the same the step as in Lemma 2, we arrive at: find $\hat{\boldsymbol{u}} \in K_{\boldsymbol{u}}$ such that

$$a(\hat{\boldsymbol{u}}, \boldsymbol{v}) = l(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in K_{\boldsymbol{u}}. \tag{3.35}$$

Assuming Ω to be smooth, it can be shown that the solution $\hat{\boldsymbol{u}}$ to the above satisfies the inequality (3.16) as well. The following lemma demonstrates it.

Lemma 4. *If the domain Ω is smooth, the solution to the problem (3.35) $\hat{\boldsymbol{u}}$ satisfies the inequality (3.16). Moreover, $\hat{\boldsymbol{u}} = \boldsymbol{u}$.*

Proof. For this proof, assume $\Gamma_c = \emptyset$ (the following argument can be easily generalized for $\Gamma_c \neq \emptyset$). We perform a simple integration by parts in (3.35):

$$\begin{aligned} a(\hat{\boldsymbol{u}}, \boldsymbol{v}) - l(\boldsymbol{v}) &= \int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) : \boldsymbol{\varepsilon}(\boldsymbol{v})\, dx - \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} dx - \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\, ds \\ &= -\int_\Omega (\operatorname{div}(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})) + \boldsymbol{f}) \cdot \boldsymbol{v}\, dx + \int_{\Gamma_N} (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n} - \boldsymbol{g}) \cdot \boldsymbol{v}\, ds \\ &\quad + \int_{\Gamma_r} (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{v}\, dx + \int_{\Gamma} (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{v}\, ds \\ &= 0 \end{aligned}$$

Choosing test functions $\boldsymbol{v} = \boldsymbol{0}$ on $\partial\Omega$, we recover the force balance equation

$$\operatorname{div}(\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}})) + \boldsymbol{f} = \boldsymbol{0} \quad \text{on } \Omega.$$

Varying the test function $\boldsymbol{v}$ over Γ_N and Γ, we recover the Neumann and the traction-free boundary conditions

$$\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n} = \boldsymbol{g} \quad \text{on } \Gamma_N, \quad \text{and } \mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n} = \boldsymbol{0} \quad \text{on } \Gamma.$$

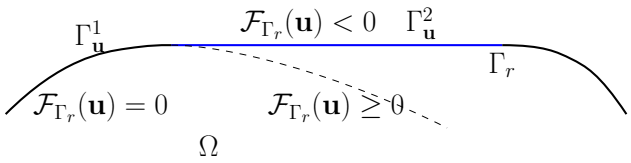


Figure 3.3: $\Gamma_r = \Gamma_{\boldsymbol{u}}^1 \cup \Gamma_{\boldsymbol{u}}^2$; dotted line shows the deformation in Ω

Since $\Gamma_c = \emptyset$, $\Gamma_r = \Gamma_{\boldsymbol{u}}^1 \cup \Gamma_{\boldsymbol{u}}^2$ and we can write

$$\begin{aligned} a(\hat{\boldsymbol{u}}, \boldsymbol{v}) - l(\boldsymbol{v}) &= \int_{\Gamma_{\boldsymbol{u}}^1} (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{v}\, dx + \int_{\Gamma_{\boldsymbol{u}}^2} (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{v}\, dx \\ &= \int_{\Gamma_{\boldsymbol{u}}^1} (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{v}\, dx + \int_{\Gamma_{\boldsymbol{u}}^2} ((\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{n})(\boldsymbol{v} \cdot \boldsymbol{n})\, dx + \int_{\Gamma_{\boldsymbol{u}}^2} ((\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{t})(\boldsymbol{v} \cdot \boldsymbol{t})\, dx \\ &= \int_{\Gamma_{\boldsymbol{u}}^1} (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{v}\, dx + \int_{\Gamma_{\boldsymbol{u}}^2} ((\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{t})(\boldsymbol{v} \cdot \boldsymbol{t})\, dx \quad (\text{since } \boldsymbol{v} \in K_{\boldsymbol{u}}) \\ &= 0. \end{aligned}$$

Hence, varying test function $\boldsymbol{v}$ over $\partial\Omega$, we find that the following holds

$$\begin{aligned} \mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n} &= \boldsymbol{0} \quad \text{on } \Gamma_{\boldsymbol{u}}^1 \\ (\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{t} &= 0 \quad \text{on } \Gamma_{\boldsymbol{u}}^1 \cup \Gamma_{\boldsymbol{u}}^2 \quad (= \Gamma_r) \\ (\hat{\boldsymbol{u}} \cdot \boldsymbol{n})((\mathbb{C}\boldsymbol{\varepsilon}(\hat{\boldsymbol{u}}) \cdot \boldsymbol{n}) \cdot \boldsymbol{n}) &= 0 \quad \text{since } \hat{\boldsymbol{u}} \in K_{\boldsymbol{u}}. \end{aligned}$$

Given the continuity of $\hat{\boldsymbol{u}}$, we claim that the normal force $\mathcal{F}_{\Gamma_r}(\hat{\boldsymbol{u}})$ on $\Gamma_{\boldsymbol{u}}^2$ is compressive only ($\mathcal{F}_{\Gamma_r}(\hat{\boldsymbol{u}}) \leq 0$). Indeed, if is not compressive and tensile instead, with ($\mathcal{F}_{\Gamma_r}(\hat{\boldsymbol{u}}) > 0$) the body would not stick ($\hat{\boldsymbol{u}} \cdot \boldsymbol{n} = 0$) on $\Gamma_{\boldsymbol{u}}^2$, instead detach from $\Gamma_{\boldsymbol{u}}^2$ (resulting in $\hat{\boldsymbol{u}} \cdot \boldsymbol{n} < 0$) as shown in Fig.3.3, contradicting the definition of $\Gamma_{\boldsymbol{u}}^2$. One can then follow the steps in Lemma 2, use the above stated boundary conditions and find that $\hat{\boldsymbol{u}}$ satisfies the weak form (3.16). Moreover, since $\boldsymbol{u}$ is unique, we can state that $\boldsymbol{u} = \hat{\boldsymbol{u}}$. □

We can then identify $\hat{\boldsymbol{u}}$ with $\boldsymbol{u}$. We now determine the shape derivative for minimizing the functional (3.34). In this regard, we propose the following theorem

Theorem 7. *Assume that when Ω perturbs to $(I + \boldsymbol{\theta})\Omega$, the contact zone $\Gamma_{\boldsymbol{u}}^2$ does not move. Let $\boldsymbol{f} \in H^1(\mathbb{R}^d)^d$, $\boldsymbol{g} \in H^2(\mathbb{R}^d)^d$ and $\boldsymbol{u} \in V$ solution to Eqn.(3.35). Then the Gâteaux derivative of $J(\Omega)$ along $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, $J'(\Omega)(\boldsymbol{\theta})$ is given by*

$$J'(\Omega)(\boldsymbol{\theta}) = \int_{\Gamma} \boldsymbol{\theta} \cdot \boldsymbol{n}\, (m(\boldsymbol{u}) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{f} \cdot \boldsymbol{v})\, ds, \tag{3.36}$$

where $\boldsymbol{v} \in K_{\boldsymbol{u}}$ is the adjoint variable satisfying the adjoint equation:

$$a(\boldsymbol{v}, \boldsymbol{\varphi}) = -\int_{\Omega} \partial_{\boldsymbol{u}} m(\boldsymbol{u}) \boldsymbol{\varphi}\, dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}} p(\boldsymbol{u}) \boldsymbol{\varphi}\, ds \qquad \forall \varphi \in K_{\boldsymbol{u}}. \tag{3.37}$$

Proof. The proof follows from a quick application of the Céa's method. Let the Lagrangian be

$$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}, \tilde{\lambda}, \Omega) = \int_{\Omega} m(\tilde{\boldsymbol{u}}) dx + \int_{\Gamma_N} p(\tilde{\boldsymbol{u}}) ds + a(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}) - l(\tilde{\boldsymbol{v}}) + \int_{\Gamma_{\boldsymbol{u}}^2 \cap \Gamma_r} \tilde{\lambda} \tilde{\boldsymbol{u}} \cdot \boldsymbol{n}\, ds + \int_{\Gamma_{\boldsymbol{u}}^1 \cap \Gamma_c} \tilde{\lambda}[\tilde{\boldsymbol{u}}] \cdot \boldsymbol{n}\, ds, \tag{3.38}$$

where $\tilde{\boldsymbol{u}} \in \tilde{V}$ is the state variable, $\tilde{\boldsymbol{v}} \in V$ is the adjoint variable and $\tilde{\lambda}$ is the multiplier that penalizes the condition $\tilde{\boldsymbol{u}} \in K_{\boldsymbol{u}}$. Applying the optimality condition of the Lagrangian (3.38) with respect to $\tilde{\boldsymbol{v}}$ and $\tilde{\lambda}$ at $\tilde{\boldsymbol{u}} = \boldsymbol{u}$ results in the state equation (3.35).

In order to determine the adjoint equation, we differentiate (3.38) with respect to $\tilde{\boldsymbol{u}}$ along $\boldsymbol{\varphi} \in V$ and equate it to zero at $(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}, \tilde{\lambda}) = (\boldsymbol{u}, \boldsymbol{v}, \lambda)$ and obtain

$$a(\boldsymbol{v}, \boldsymbol{\varphi}) + \int_{\Omega} \partial_{\boldsymbol{u}} m(\boldsymbol{u}) \boldsymbol{\varphi}\, dx + \int_{\Gamma_{\boldsymbol{u}}^2 \cap \Gamma_r} \lambda \boldsymbol{\varphi} \cdot \boldsymbol{n}\, ds + \int_{\Gamma_{\boldsymbol{u}}^1 \cap \Gamma_c} \lambda[\boldsymbol{\varphi}] \cdot \boldsymbol{n}\, ds, = 0 \qquad \forall \boldsymbol{\varphi} \in V.$$

Choosing $\boldsymbol{\varphi} \in K_{\boldsymbol{u}}$, we obtain the weak form of the adjoint equation: find $\boldsymbol{v} \in K_{\boldsymbol{u}}$ such that

$$a(\boldsymbol{v}, \boldsymbol{\varphi}) = -\int_\Omega \partial_{\boldsymbol{u}} m(\boldsymbol{u})\boldsymbol{\varphi}\, dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}} p(\boldsymbol{u})\boldsymbol{\varphi}\, ds \qquad \forall \boldsymbol{\varphi} \in K_{\boldsymbol{u}}.$$

To determine the shape derivative, we need to differentiate $J(\Omega) = \mathcal{L}(\boldsymbol{u}, \tilde{\boldsymbol{v}}, \tilde{\lambda}, \Omega)$ with respect to Ω. Given the assumption that $\Gamma^2_{\boldsymbol{u}}$ does not move, this differentiation is straightforward (using Lemma 1) and yields the shape derivative (3.36). □

One may now ask, if the adjoint solution to (3.37) comparable to the adjoint problem corresponding to the penalized case (3.30). We rewrite (3.30) here for the case $\mu = 0$: find $\boldsymbol{v} \in V$ such that

$$a(\boldsymbol{v}, \boldsymbol{\varphi}) + \frac{1}{\epsilon}\left(\int_{\Gamma_r} \mathcal{H}(\boldsymbol{u}\cdot\boldsymbol{n})(\boldsymbol{v}\cdot\boldsymbol{n})(\boldsymbol{\varphi}\cdot\boldsymbol{n})ds + \int_{\Gamma_c} \mathcal{H}([\boldsymbol{u}]\cdot\boldsymbol{n}^-)([\boldsymbol{v}]\cdot\boldsymbol{n}^-)([\boldsymbol{\varphi}]\cdot\boldsymbol{n}^-)ds\right)$$
$$= -\int_\Omega \partial_{\boldsymbol{u}} m(\boldsymbol{u})\boldsymbol{\varphi}\, dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}} p(\boldsymbol{u})\boldsymbol{\varphi}\, ds \qquad \forall \boldsymbol{\varphi} \in V.$$

Indeed, the above is nothing but an approximation of (3.37) where the constraint $\boldsymbol{v} \in K_{\boldsymbol{u}}$ has been penalized using the second term on the left in the above.

3.5 Numerical implementation

We first present the aspects of numerical resolution of the state equations (3.26), then of the adjoint (3.30) and finally of the shape optimization algorithm.

We discretize the domain Ω to Ω^h using a simplicial unstructured mesh. The contact boundary $\Gamma_c = \Gamma_c^+ \cup \Gamma_c^-$ (the left and the right interface) is discretized to Γ_c^h, consisting of left and right interface, both of which contain coincident mesh vertices as shown in Fig.3.4 (in red). While using coincident vertices incredibly simplifies the resolution of the contact problem (see Subsection 3.5.1), it restrains us from using body-fitted remeshing (explained in 2.2.2). This is because it is very difficult to remesh, while ensuring that the mesh vertices on Γ_c remain coincident. Hence, for this chapter only, we shun remeshing and instead apply the traditional ersatz material approach (explained in Subsection 2.2.1) for test cases with self-contact boundaries Γ_c. However, for test cases with rigid contact boundary only (Γ_r), there are no vertices to preserve and we can remesh easily. We drop the superscript h from Γ^h, Γ^h_D and Γ^h_N hereafter.

The space V is discretized to V^h using the finite element framework. We let

$$V^h = \mathbb{P}^1(\Omega^h)^d \tag{3.39}$$

Let the maximal mesh size be $h_{\max}$, minimal mesh size be $h_{\min}$. We assume the mesh to be regular, or $h_{\max}$ and $h_{\min}$ to be of the same order. For the numerical examples concerning contact only, we shall treat test cases with and without friction (however, for the numerical examples involving contact and bolt model, we shall treat frictionless test cases only).

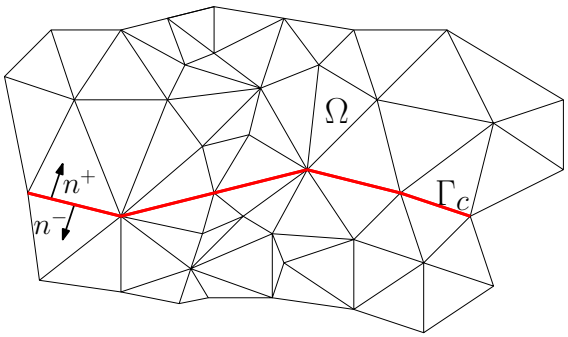


Figure 3.4: Coincident boundary mesh of Γ^h, representing the contact surface Γ_c

Since we have used $\mathbb{P}^1$ elements for the displacement $\boldsymbol{u}^h$, we shall commit an error of $\mathcal{O}(h_{\max})$ in the displacement solution. For the contact problem (3.22), as per the estimate (3.24), penalization alone induces

additional error in displacement, of $\mathcal{O}(\sqrt{\epsilon})$. In order to ensure that this additional error is of the same order as $\mathcal{O}(h_{\max})$, we take

$$\epsilon < E\left(\frac{h_{\max}}{L}\right)^2 \tag{3.40}$$

where L is the characteristic length of Ω. The regularization parameter is chosen as $\eta = \epsilon^2$. All of our numerical experiments have been performed on the open source software `FreeFEM` [92].

3.5.1 Resolution of the contact problem (3.26)

The discretized state solution (to problem (3.26)), $\boldsymbol{u}^h_{\epsilon,\eta} \in V^h$ is denoted by $\boldsymbol{u}^h$ here to simplify notations. Alongside we drop the superscript h for discretized contact interface Γ_c, Γ_c^+ and Γ_c^-. The discretized version of the contact problem (3.26) reads: find $\boldsymbol{u}^h \in V^h$ such that

$$a(\boldsymbol{u}^h, \boldsymbol{v}^h) + j'_\epsilon(\boldsymbol{u}^h)\cdot \boldsymbol{v}^h + j'_{c,\epsilon}(\boldsymbol{u}, \boldsymbol{v}^h) = l(\boldsymbol{v}^h) \quad \forall v^h \in V^h, \tag{3.41}$$

where it is understood that the bilinear form and linear forms are now defined on Ω^h. For the numerical resolution of the above, one can apply a simple fixed point or a Newton algorithm. We apply a fixed point algorithm to (3.41) using the following construction: find $\boldsymbol{u}^h_n \in V^h, 1 \le n < \infty$ such that

$$\begin{aligned} a(\boldsymbol{u}^h_n, \boldsymbol{v}^h) + \frac{1}{\epsilon}\Bigg(\int_{\Gamma_r^h} \mathcal{H}(\boldsymbol{u}^h_{n-1}\cdot\boldsymbol{n})(\boldsymbol{u}^h_n\cdot\boldsymbol{n})(\boldsymbol{v}^h\cdot\boldsymbol{n})ds + \int_{\Gamma_c} \mathcal{H}([\boldsymbol{u}^h_{n-1}]\cdot\boldsymbol{n}^-)\ ([\boldsymbol{u}^h_{n+1}]\cdot\boldsymbol{n}^-)([\boldsymbol{v}^h]\cdot\boldsymbol{n}^-)ds\Bigg) \\ + \frac{\mu}{\epsilon}\int_{\Gamma_c} \phi([\boldsymbol{u}^h_{n-1}]\cdot\boldsymbol{n}^-)\kappa([\boldsymbol{u}^h_{u-1}]_t)\ [\boldsymbol{u}^h_n]_t\ [\boldsymbol{v}^h]\cdot\boldsymbol{n}^- ds = l(\boldsymbol{v}^h) \qquad \forall \boldsymbol{v}^h \in V^h. \end{aligned} \tag{3.42}$$

The convergence $\boldsymbol{u}^h_n \longrightarrow \boldsymbol{u}_n$ as $n \longrightarrow \infty$ is expected. This convergence is observed numerically always when $\mu = 0$ (since j_ϵ is convex). For $\mu \neq 0$, the convergence is not guaranteed.

We discuss the numerical implementation of (3.42) on `FreeFEM` in 2D for the frictionless case $\mu = 0$. The extension to the 3D and friction case is straight-forward. Let φ be the $\mathbb{P}^1$ Lagrange finite element basis and $\mathcal{N}$ be the number of mesh vertices, then $\boldsymbol{u}^h_n = \sum_{i=1}^{\mathcal{N}} [c_i\varphi_i, d_i\varphi_i]$. Let $\boldsymbol{U}_n = [c_1, c_2, \cdots, d_1, d_2, \cdots]$. Then (3.42) can be written in the matricial form as

$$(\boldsymbol{K} + \boldsymbol{M}_c(\boldsymbol{U}_{n-1}) + \boldsymbol{M}_r(\boldsymbol{U}_{n-1}))\boldsymbol{U}_n = \boldsymbol{F} \tag{3.43a}$$

$$\text{where} \quad \boldsymbol{K}\boldsymbol{U}_n \equiv \int_{\Omega^h} \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^h_n) : \boldsymbol{\varepsilon}(\boldsymbol{v}^h)\ dx, \tag{3.43b}$$

$$\boldsymbol{M}_r(\boldsymbol{U}_{n-1})\boldsymbol{U}_n \equiv \frac{1}{\epsilon}\int_{\Gamma_r} \mathcal{H}(\boldsymbol{u}^h_{n-1}\cdot\boldsymbol{n})(\boldsymbol{u}^h_n\cdot\boldsymbol{n})(\boldsymbol{v}^h\cdot\boldsymbol{n})ds,$$

$$\boldsymbol{M}_c(\boldsymbol{U}_{n-1})\boldsymbol{U}_n \equiv \frac{1}{\epsilon}\int_{\Gamma_c} \mathcal{H}([\boldsymbol{u}^h_{n-1}]\cdot\boldsymbol{n}^-)\ [\boldsymbol{u}^h_n]\cdot\boldsymbol{n}^-[\boldsymbol{v}^h]\cdot\boldsymbol{n}^- ds.$$

A simple initialization for a fixed point algorithm (3.43a) is the solution to the elasticity system without the contact boundary conditions. `FreeFEM` does not have an inbuilt contact mechanics solver. Rather it has all the functions and libraries to create meshes, finite element basis and the stiffness matrix. One can simply give the expression of variational formulation as input and let `FreeFEM` construct the corresponding stiffness matrix. Thus the construction of the stiffness matrix K and of $M_r(\boldsymbol{U}_{n-1})$ on `FreeFEM` is very easy. However, the construction of $M_c(\boldsymbol{U}_n)$ is not straight-forward as one must determine the jump $[\boldsymbol{v}]$ on Γ_c. The complexity of determining this jump is dependent on the mesh at the boundary Γ_c. If the vertices on the two surfaces Γ_c^+ and Γ_c^- are coincident, the determination of $[\boldsymbol{u}^h_{n-1}]\cdot\boldsymbol{n}$ is quick. If the vertices on the two surfaces are not coincident, then one has to resort to several complicated projection strategies such as the mortar finite elements [27], slave-master approach [146], etc.

Let $F = \boldsymbol{M}_c(\boldsymbol{U}_{n-1})$. In 2D, the matrix F has 4 blocks of size $\mathcal{N}\times\mathcal{N}$, schematically given

$$F = \left[\begin{array}{c|c} F^{11} & F^{12} \\ \hline F^{21} & F^{22} \end{array}\right],$$

Letting $\boldsymbol{n}^- = [n_x^-, n_y^-]$, $\boldsymbol{n}^+ = [n_x^+, n_y^+]$, the ij-th entry of the first block F_{11} is constructed as

$$\begin{aligned}
F^{11} &= \frac{1}{\epsilon}\int_{\Gamma_c} (\varphi_i^- - \varphi_i^+)n_x^-\mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)(\varphi_j^- - \varphi_j^+)n_x^- \; ds \\
&= \int_{\Gamma_c} \varphi_i^- n_x^- \varphi_j^- n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds + \int_{\Gamma_c} \varphi_i^+ n_x^- \varphi_j^+ n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds \\
&\quad - \int_{\Gamma_c} \varphi_i^- n_x^- \varphi_j^+ n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds - \int_{\Gamma_c} \varphi_i^+ n_x^- \varphi_j^- n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds \\
&= \int_{\Gamma_c^-} \varphi_i n_x^- \varphi_j n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds + \int_{\Gamma_c^+} \varphi_i n_x^+ \varphi_j n_x^+ \mathcal{H}(-[\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^+)\; ds \\
&\quad - \int_{\Gamma_c} \varphi_i^- n_x^- \varphi_j^+ n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds - \int_{\Gamma_c} \varphi_i^+ n_x^- \varphi_j^- n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds,
\end{aligned}$$

where we have used the fact that $\boldsymbol{n}^+ = \boldsymbol{n}^-$ because the mesh vertices are coincident. Note that the term $[\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-$ is the same for Γ_c^+ and Γ_c^-. Finally in the matrix form, we have

$$F^{11} = \begin{array}{c} \\ \Gamma_c^- \\ \Gamma_c^+ \end{array} \begin{array}{c} \begin{array}{cc} \cdots & \Gamma_c^- \end{array} \quad \Gamma_c^+ \\ \left[\begin{array}{ccc|c} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \hline 0 & \cdots & M_{11} & -M_{12} \\ 0 & \cdots & -M_{21} & M_{22} \end{array}\right] \end{array},$$

$$\text{where} \quad M_{11} \equiv \int_{\Gamma_c^-} \varphi_i n_x^- \varphi_j n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds, \tag{3.44}$$

$$M_{12} \equiv \int_{\Gamma_c} \varphi_i^- n_x^- \varphi_j^+ n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds, \tag{3.45}$$

$$M_{21} \equiv \int_{\Gamma_c} \varphi_i^+ n_x^- \varphi_j^- n_x^- \mathcal{H}([\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^-)\; ds, \tag{3.46}$$

$$M_{22} \equiv \int_{\Gamma_c^+} \varphi_i n_x^+ \varphi_j n_x^+ \mathcal{H}(-[\boldsymbol{u}_{n-1}^h]\cdot\boldsymbol{n}^+)\; ds. \tag{3.47}$$

Before integrating the terms in F^{11}, we need to determine the jump $[\boldsymbol{u}_{n-1}^h]$ on Γ_c^+ (and likewise on Γ_c^-). The value $\boldsymbol{u}_{n-1}^h|_{\Gamma_c^-}$ can be projected on Γ_c^+ with ease. Let this projection be $\bar{\boldsymbol{u}}_{n-1}^h$. It suffices to find a vertex $x_j \in \Gamma_c^-$ which infinitesimally close to the vertex $x_k \in \Gamma_c^+$ (existence guaranteed by construction), equating the value $\bar{\boldsymbol{u}}_{n-1}^h(x_k) = \boldsymbol{u}_{n-1}^h(x_j)$ and then repeating this over all points $x_k \in \Gamma_c^+$. The jump on Γ_c^+ is thus $\boldsymbol{u}_{n-1}^h|_{\Gamma_c^+} - \bar{\boldsymbol{u}}_{n-1}^h$.

After determining the jump, the matrices M_{11} and M_{22} can then be constructed easily. Our numerical simulations show that if the jump of $[\boldsymbol{u}_{n-1}^h]$ it is not well calculated, the fixed point construction (3.43) may not converge.

Now we talk about the construction of M_{12} and M_{21}. As seen in (3.45) above, we need to integrate the product of the test function φ_i^+ (defined on the left Γ_c^+) and the test function φ_j^- (defined on the right Γ_c^-) on the boundary Γ_c (or Γ_c^+ since the boundaries are coincident). In order to find this product for M_{12} (say), it suffices to consider the projection of φ_j^- on Γ_c^+ using the same technique as before (finding the nearest neighbor $x_k \in \Gamma_c^+$ to $x_j \in \Gamma_c^-$), and computing the product $\varphi_i^+\varphi_k^+$ on Γ_c^+. However, we note that the product $\varphi_i^+\varphi_k^+$ was already computed while constructing M_{11}. Hence in order to construct M_{12} it suffices to equate the indices $M_{12}(i,j) = M_{11}(i,k)$. In the same way, one can construct M_{21} by equating $M_{12}(i,j) = M_{22}(i,k)$.

Additionally, with the fixed point method it is easier to converge for penalization parameter ϵ "relatively small". Consider the material parameter E of the order unity for a unit square geometry and ϵ satisfies (3.40). In this case, the small magnitude of ϵ is of the order 10^{-4}. However, with ϵ "very small", say 10^{-8}, we may have slower convergence. Here, we start the iteration with the solution obtained at $\epsilon = 10^{-4}$, using it as initialization for the case $\epsilon = 10^{-5}$, converge and reduce ϵ further. With this strategy, we find rapid convergence. The same holds for Newton-Raphson as well. For Newton-Raphson, it is more legitimate to start the iteration with solution obtained from fixed point iteration. Since Newton-Raphson converges at a quadratic rate, we can set tighter tolerance for convergence and attain a better solution quickly.

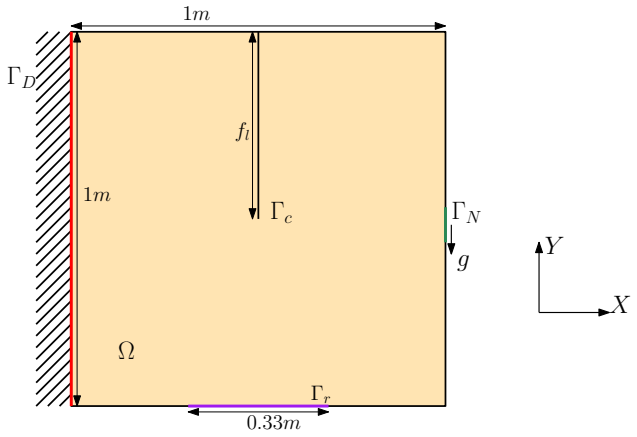
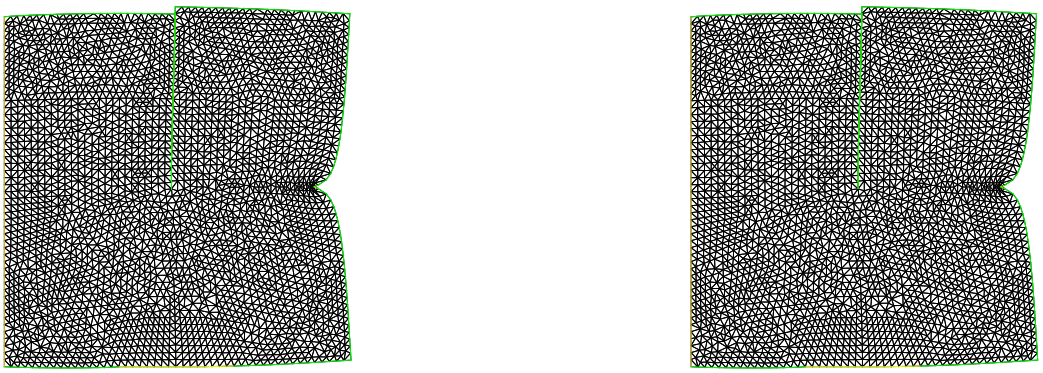


Figure 3.5: Boundary conditions for a $1m \times 1m$ square with a 0.5 long fissure at the center

3.5.2 Numerical verification of the fixed point scheme (3.43a)

Let $E = 20, \nu = 0.3$. To verify the above fixed point scheme (3.43a), we consider two test cases: a $2D$ square and a $3D$ cube.

First, we take up a $1m \times 1m$ square shape. As shown in Fig.3.5, the shape is fixed on the leftmost edge, and forced on a small part of the rightmost edge and has a fissure at the center. The fissure starts from the topmost edge and has a length $f_l = 0.5m$. We plot the deformed shape, obtained for several values of $\epsilon,\ \mu,\ \eta,\ \boldsymbol{g}$ and compare them.



(a) $\epsilon = 10^{-4},\ \mu = 0$ (b) $\epsilon = 10^{-10},\ \mu = 0$

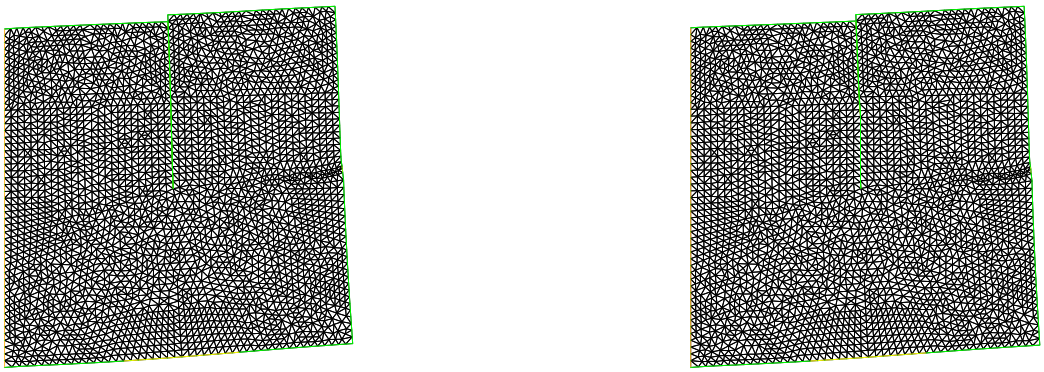
Figure 3.6: Mesh deformed by u for $\boldsymbol{g} = (-50,\ 0)\text{N}/m,\ \mu = 0$

To begin with, we take a compressive force $\boldsymbol{g} = [-50,\ 0],\ \mu = 0$ and plot the deformed shape obtained for $\epsilon = 10^{-4}$ and $\epsilon = 10^{-10}$ in Fig.3.6. As seen in this figure, the difference between the two shapes is very negligible.

We then consider a shear force, by taking $\boldsymbol{g} = (0,\ 10)\text{N}/m$ for $\mu = 0$. Again, we plot the solution obtained at $\epsilon = 10^{-4}$ and $\epsilon = 10^{-10}$ in Fig.3.7. We repeat the same experiment ($\mu = 0$) by changing the direction of the force, $\boldsymbol{g} = (0,\ -10)\text{N}/m$, and plot deformed shape in Fig.3.8. We see that in each case, the difference between the deformed shape for $\epsilon = 10^{-4}$ and $\epsilon = 10^{-10}$ is negligible. Implying it suffices to take $\epsilon = 10^{-4}$ for our numerical experiments in Section 3.6.

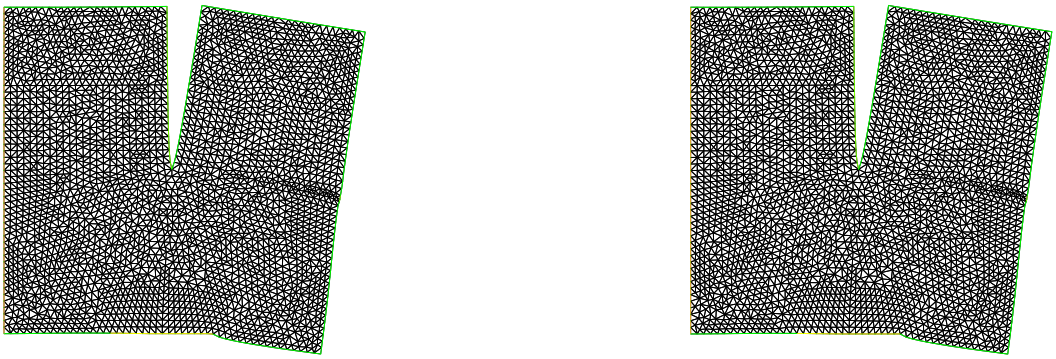
Now we consider non-zero friction, $\mu = 0.2$ with a shear force $\boldsymbol{g} = (0,\ 10)\text{N}/m$. As discussed earlier, we apply the fixed point scheme (3.43a) for a lower value of (ϵ, η) (for instance, $\eta = 0.1,\ \epsilon = 10^{-3}$), converge to the solution, reduce (ϵ, η) by a factor of 10 and repeat. We plot the solution obtained at $\epsilon = 10^{-4}, \eta = 10^{-2}$ and at $\epsilon = 10^{-10}, \eta = 10^{-8}$ in Fig.3.9.

Remark 1. *A possible way to bypass the construction of Γ_c^- and Γ_c^+ and instead have a single boundary Γ_c meshed is by considering the discontinuous finite element space. Computation using discontinuous Galerkin method can be performed easily used on* `FreeFEM`*, using its* P1dc *elements.*



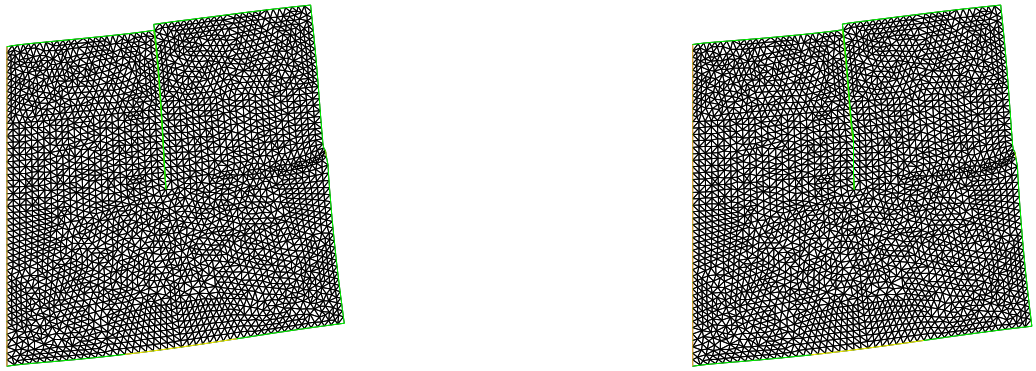
(a) $\epsilon = 10^{-4},\ \mu = 0$ (b) $\epsilon = 10^{-10},\ \mu = 0$

Figure 3.7: Mesh deformed by u for $\boldsymbol{g} = (0,\ 10)\mathrm{N}/m,\ \mu = 0$



(a) $\epsilon = 10^{-4},\ \mu = 0$ (b) $\epsilon = 10^{-10},\ \mu = 0$

Figure 3.8: Mesh deformed by u for $\boldsymbol{g} = (0,\ -10)\mathrm{N}/m$

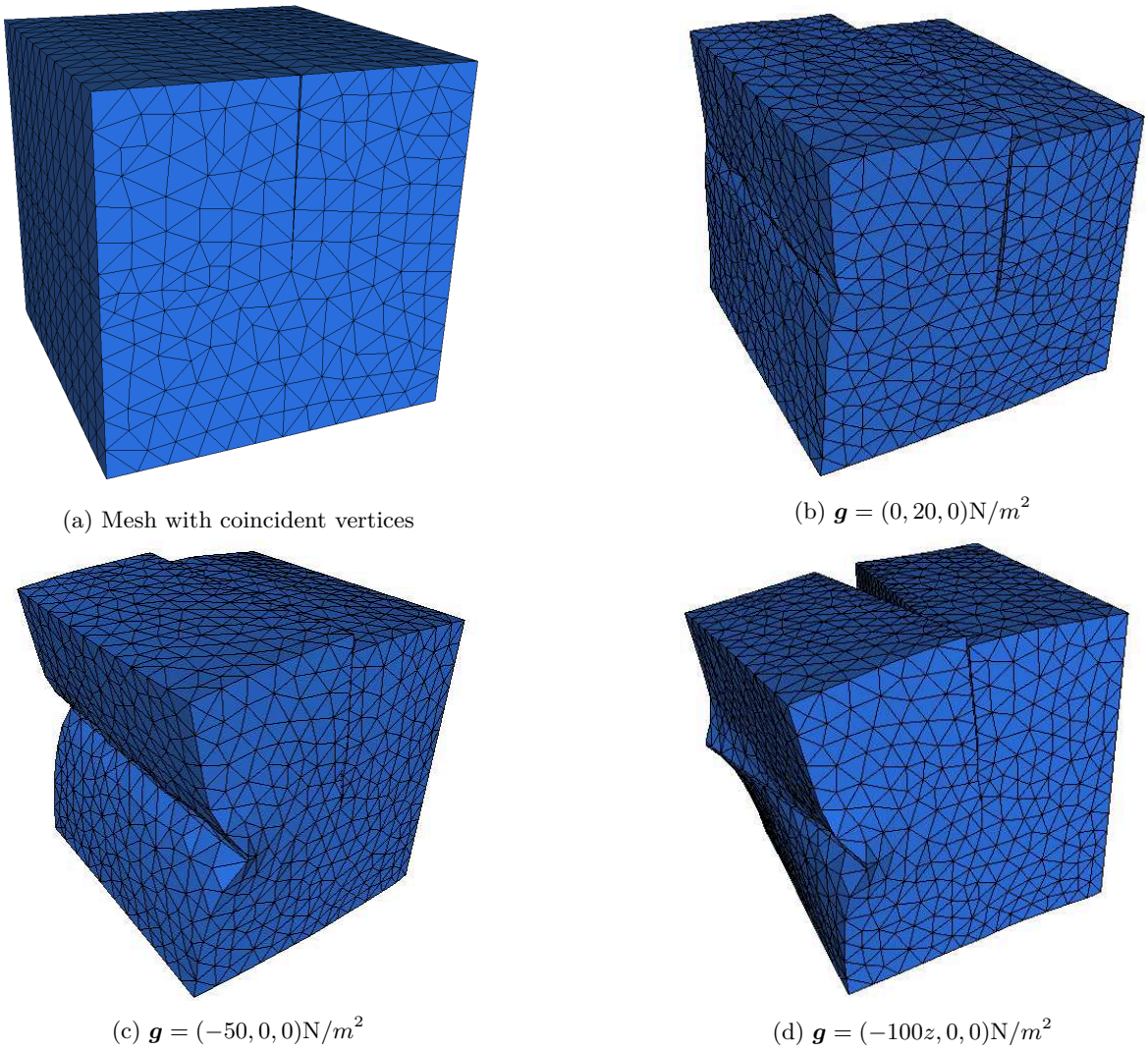


(a) $\epsilon = 10^{-4},\ \eta = 10^{-2},\ \mu = 0.2$ (b) $\epsilon = 10^{-10},\ \eta = 10^{-8},\ \mu = 0.2$

Figure 3.9: Mesh deformed by u for $\boldsymbol{g} = (0,\ 10)\mathrm{N}/m$

However, compared to the classical Galerkin method (with continuous P1 *elements) the number of degrees of freedom is* 6 *times higher in 2D and* 8 *times in 3D for* P1dc*. In order to reduce the computation time, we prefer not to use* P1dc*.*

Finally, we take up a $3D$ cube of unit length, fixed on one face, forced on the opposite face and with a fissure at the center of dimension $1 \times 1/2$. As shown in Fig.3.10a, the geometry is meshed with co-incident vertices. The solutions obtained for $\epsilon = 10^{-5},\ \mu = 0,\ E = 20,\ \nu = 0.3$ and three different forcing functions: $\boldsymbol{g} = (0, 20, 0)\mathrm{N}/m^2, (-50, 0, 0)\mathrm{N}/m^2, (-100z, 0, 0)\mathrm{N}/m^2$ are plotted in Fig.3.10.



(a) Mesh with coincident vertices

(b) $\boldsymbol{g} = (0, 20, 0)\mathrm{N}/m^2$

(c) $\boldsymbol{g} = (-50, 0, 0)\mathrm{N}/m^2$

(d) $\boldsymbol{g} = (-100z, 0, 0)\mathrm{N}/m^2$

Figure 3.10: Solution for $\epsilon = 10^{-6}, \eta = 0,\ \mu = 0$

3.5.3 Resolution of the adjoint problem (3.30)

The space discretized version of the contact problem (3.30) reads: find $\boldsymbol{v}^h \in V^h$ such that

$$a(\boldsymbol{\varphi}, \boldsymbol{v}^h) + (j_\epsilon''(\boldsymbol{u}^h) \cdot \boldsymbol{\varphi}) \cdot \boldsymbol{v}^h + \partial_{\boldsymbol{u}^h} j_{\epsilon,\eta}(\boldsymbol{u}^h, \boldsymbol{v}^h) \cdot \boldsymbol{\varphi} = -\int_\Omega \partial_{\boldsymbol{u}^h} m(\boldsymbol{u}^h) \cdot \boldsymbol{\varphi}\, dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}^h} p(\boldsymbol{u}^h) \cdot \boldsymbol{\varphi}\, ds \qquad \forall \boldsymbol{\varphi} \in V^h \tag{3.48}$$

where $\boldsymbol{u}^h$ is the solution to the problem (3.41). The above equation is evidently linear with respect to $\boldsymbol{v}^h$. The construction of the stiffness matrix for the problem (3.48) involves calculating the jump $[\boldsymbol{u}^h]$ as well as the jump in the test functions φ on the contact boundary Γ_c. The construction of this jump is explained in Subsection 3.5.1.

Alternatively, if one uses a Newton-Raphson algorithm to solve the equation (3.41), it suffices to use the tangent matrix of the last iteration (of Newton-Raphson) instead of construction the stiffness matrix for (3.48) from scratch.

Using the adjoint solution $\boldsymbol{v}^h$, the space discretized shape derivative can be expressed as

$$\begin{aligned} J'(\Omega^h)\boldsymbol{\theta} &= \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} j(\Omega^h) ds \\ &= \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} \left(m(\boldsymbol{u}^h) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^h) : \boldsymbol{\varepsilon}(\boldsymbol{v}^h) - \boldsymbol{f} \cdot \boldsymbol{v}^h \right) ds, \end{aligned} \tag{3.49}$$

where $\boldsymbol{u}^h$ is the solution to (3.41).

3.5.4 Shape optimization algorithm

For all the test cases in this chapter, we minimize the compliance, defined as

$$J(\Omega^h) = \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{u} \, ds. \tag{3.50}$$

We stick to the ersatz approach (see Subsection 2.2.1) instead of using remeshing for reasons explained in the beginning of Section 3.5 for test cases involving self-contact boundary Γ_c. For these test cases, the design space D has a fixed mesh and all shapes Ω^h are captured on it using a level-set function $\psi : D \to \mathbb{R}$. For the test cases involving rigid boundary Γ_r only, we apply the remeshing algorithm (discussed in Subsection 2.2.2).

The space of admissible spaces $\mathcal{U}_{ad}$ is given by (2.25) (shapes that satisfy a target volume). We then consider the shape optimization problem

$$\min_{\Omega \in \mathcal{U}_{ad}} J(\Omega^h).$$

As we did in the previous chapter in Subsection 2.2.6, we construct a Lagrangian corresponding for the above problem as

$$\mathcal{L}(\boldsymbol{u}, \boldsymbol{v}, \Omega^h) = J(\Omega^h) + \frac{\lambda}{C_V} \left(V_{\Omega^h} - V_f \right). \tag{3.51}$$

where λ is the Lagrange multiplier for the volume constraint, V_{Ω^h} is the volume of the optimizable domain, and C_V is a normalization constant, given by (2.27). The optimizable domain is D unless otherwise mentioned. As usual, we begin with an initial shape Ω_0^h containing periodic holes and a standard gradient-based Uzawa-type algorithm is applied to the Lagrangian (3.51). The initial volume is typically taken to be much greater than the target volume V_f. Let $I_{\max} = 100$ be the maximal number of shape optimization iterations.

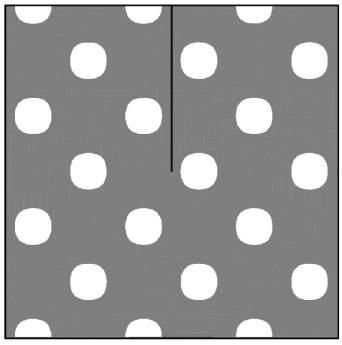
For the shape at each iteration i, the shape derivative $j(\Omega_i^h)$ in (3.49) is regularized to $dj_\alpha(\Omega_i^h)$ by solving (2.23). The pseudo-time step (or descent step) τ is defined by (2.28). The volume constraint is satisfied gradually by updating the multiplier λ at each iteration i using (2.29). Then, the transport equation the transport equation (2.17) is solved with a velocity θ_i (2.30) for the descent step τ. If the volume $V_{\Omega_i^h}$ is close to the target volume, we apply the volume projection scheme (2.31). We thus perform the algorithm 2.

Algorithm 2 Initialize with a shape $\Omega_0^h, \boldsymbol{u}_0 = \boldsymbol{0}$ and repeat over $i = 1, \cdots, I_{\max}$

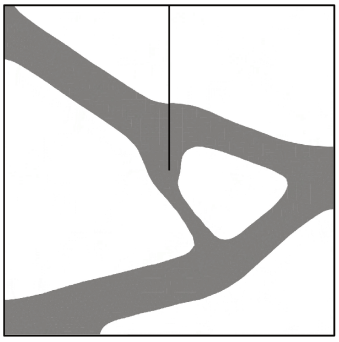
1. Solve state equation (3.41) for $\boldsymbol{u}$ on the mesh of Ω_i^h using the fixed point (3.43a) and initial guess $\boldsymbol{u}_{i-1}$
2. Solve adjoint equation (3.48) for $\boldsymbol{v}$ on the mesh of Ω_i^h
3. Compute the shape derivative (3.49) and regularize it using (2.23) to find $dj_\alpha(\Omega_i^h)$
4. Update the Lagrange multiplier using (2.29), using the time step (2.28)
5. Solve the transport equation (2.17) with the velocity given by (2.30) for the pseudo-time step τ given by (2.28) to obtain the new level-set function $\tilde{\phi}_{i+1}$
6. Re-initialize $\tilde{\phi}_{i+1}$ to the signed distance function ϕ_{i+1} (defining a new shape Ω_{i+1}^h)
7. Compute the volume $V_{\Omega_{i+1}^h}$. If it is close to the volume target, apply the projection algorithm (2.31) to satisfy exactly the volume constraint
8. Impose non-optimizable domain using (2.24)
9. If $\Gamma_c = \emptyset$, remesh the box D using `MMG` [61] to obtain the body-fitted mesh of the new shape Ω_{i+1}

3.6 Results: Contact

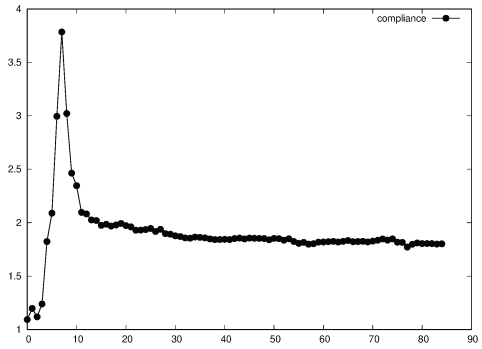
For all the test cases, we minimize compliance objective function (2.3) using the algorithm 2. Only for this section, the elasticity coefficients are chosen to be $E = 20$, $\nu = 0.3$. The body force $\boldsymbol{f}$ is assumed to be zero. For all 3D test cases, the coefficient of friction $\mu = 0$. The only test cases in this chapter which involve remeshing are studied in Subsection 3.6.3 (a chair) and 3.6.4 (a 3D gripper).



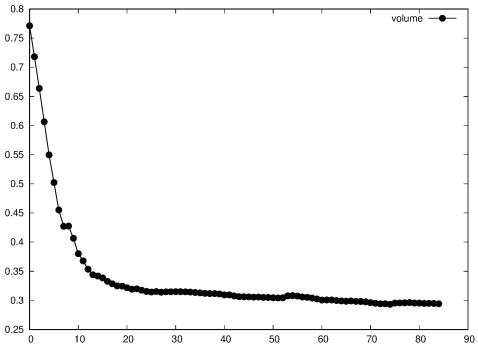
(a) Initial shape



(b) Final shape



(c) Compliance vs iterations



(d) Volume vs iterations

Figure 3.11: $f_l = 0.5m,\ \mu = 0$

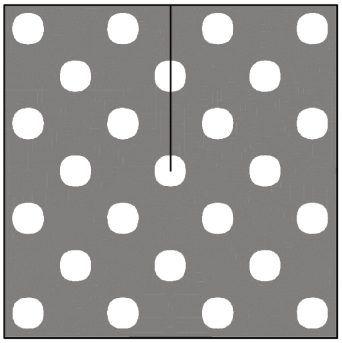
3.6.1 Square

In this chapter, we consider only one 2D test case, a $1m \times 1m$ square shape. As show in Fig.3.5, the square shape is fixed on its leftmost edge, forced on a small part of the rightmost edge and has a fissure at the center. The fissure starts from the topmost edge and has a length f_l. In addition, the square is given a rigid support Γ_r on the bottom most edge in the middle with a length of $0.33m$. The surface force is chosen as $g = [0,\ 10]\text{N}/m$. As the frictionless contact problem is scalable, any magnitude of force gives the same optimized shape. First we investigate the influence of the initial shape for $f_l = 0.5m$ and absence of friction $\mu = 0$. As shown in Fig.3.11 and in Fig.3.12, for two different initial shapes, we get to different optimized shapes.

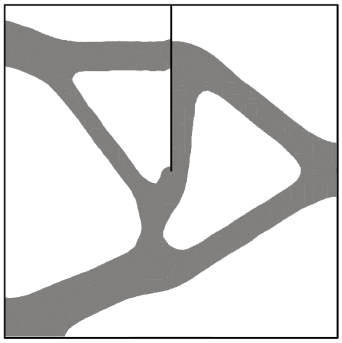
Now to see the influence of the crack length f_l, we reduce it to $f_l = 0.25m$ and obtain the shape in Fig.3.13. As seen in this figure, the influence is very significant and the optimization algorithm tries to avoid putting material around the fissure. To see the impact of friction, we optimize for $\mu = 0.2, f_l = 0.5m$ and obtain the shape in Fig.3.14. This shape very much resembles the final shape obtain for the frictionless case in Fig.3.12. We note that for all optimized topologies obtained, the boundary of the geometry remains almost continuous near the contact boundary Γ_c. This is expected in the presence of a discontinuous surface, the structure would have a vertical surface that neither acts as a support, nor is supported.

3.6.2 Cube

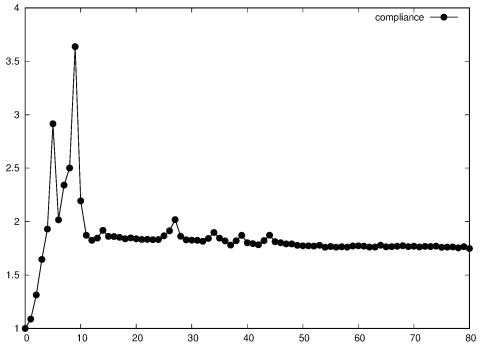
We consider a cube of unit side length. As seen in Fig.3.15a, the cube is fixed on red face and forced on a small square region with $\boldsymbol{g}$. The cube has a fissure at its center of dimension $1m \times 0.5m$ (the plane in purple). The initial shape is taken to be (3.15b). To see the influence of the direction of the force, we consider two cases (same target volume $V_f = 0.25m^3$): upward force $\boldsymbol{g} = (0, 10, 0)\text{N}/m^2$ and downward force $\boldsymbol{g}(0, -10, 0)\text{N}/m^2$. The two forces result in two very different topologies as seen in Fig.3.15. Here as well, the optimized topologies so obtained in Fig.3.15 have no surface discontinuity around the interface Γ_c.



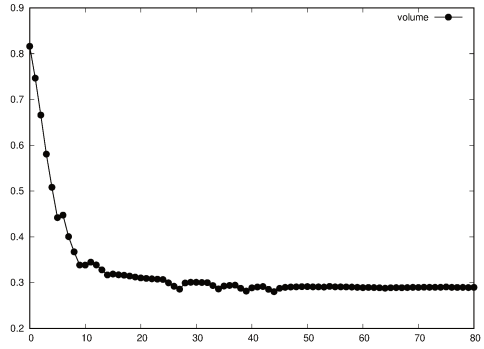
(a) Initial shape



(b) Final shape

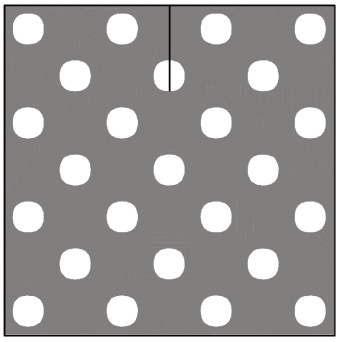


(c) Compliance vs iterations

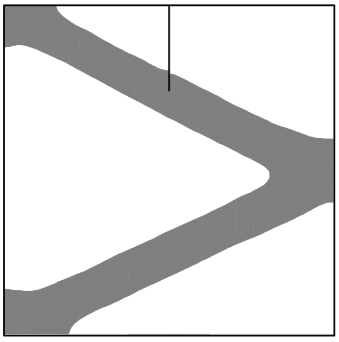


(d) Volume vs iterations

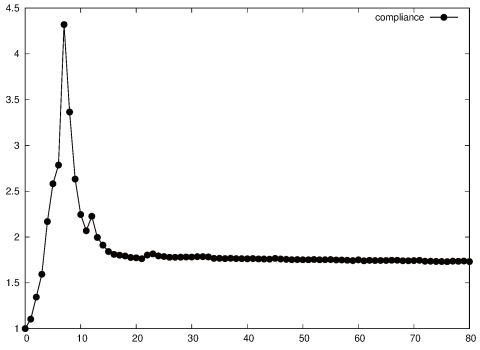
Figure 3.12: $f_l = 0.5m,\ \mu = 0$



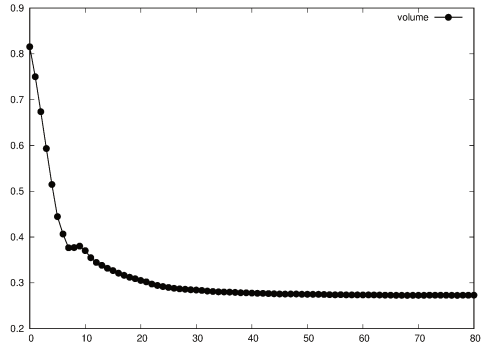
(a) Initial shape



(b) Final shape

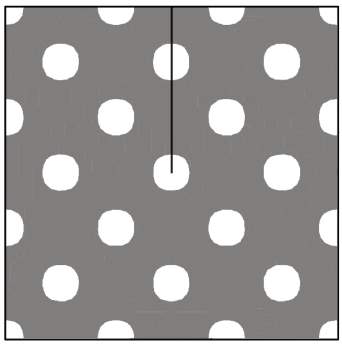


(c) Compliance vs iterations

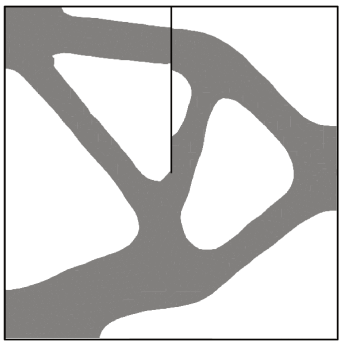


(d) Volume vs iterations

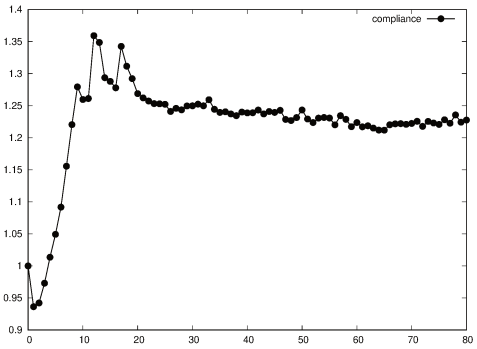
Figure 3.13: $f_l = 0.25m,\ \mu = 0$



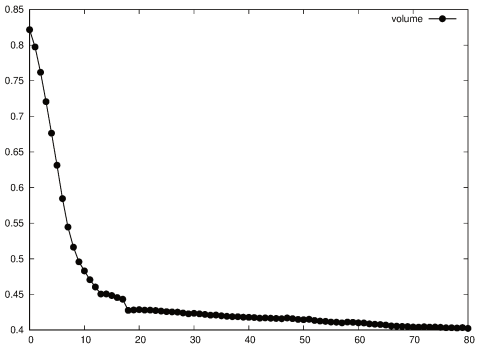
(a) Initial shape



(b) Final shape



(c) Compliance vs iterations



(d) Volume vs iterations

Figure 3.14: $f_l = 0.5m,\ \mu = 0.2$

3.6.3 Chair

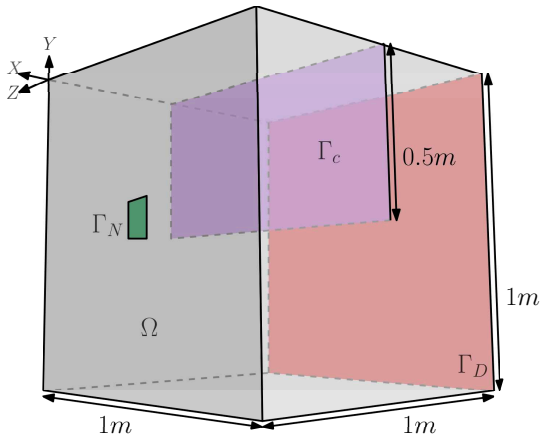
As a realistic test case, we now consider a chair. As shown in Fig.3.16a, the chair is supported on four circular regions at the bottom and subjected to two forces $\boldsymbol{g}_1 = (-10, 0, 0)\text{N}/m^2$ and $\boldsymbol{g}_2 = (0, -10, 0)\text{N}/m^2$, equal in magnitude on Γ_N^1 and Γ_N^2, respectively. The compliance (2.3) computed for both the forces $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$, is minimized. Since there are rigid contact supports only, we perform remeshing.

On the seat Γ_N^1 and on the backrest Γ_N^2, there are thin plate-like regions, assumed to be non-optimizable. A target volume $V_f = 0.65m^3$ is imposed. As can be seen in Fig.3.16a, two supports are clamped (in red) and two supports act as rigid contact (in purple). The initial shape in Fig.3.16b converges to the shape in Fig.3.16d. Here we compare the optimized chair to the one that would be obtained for linear elasticity. For linear elasticity, we assume that all the four supports are clamped (hence act as Dirichlet boundary Γ_D). Initialized from (3.16b), the shape obtained for this case is plotted in Fig.3.16c. There is a striking difference between the chair (3.16c) and the chair (3.16d). The four supports legs of the chair (3.16d) are connected by horizontal bars, unlike the chair (3.16c). This is because of the frictionless boundary condition which permit the support legs of (3.16d) to slide along the horizontal surface, which is prevented by the horizontal bars.

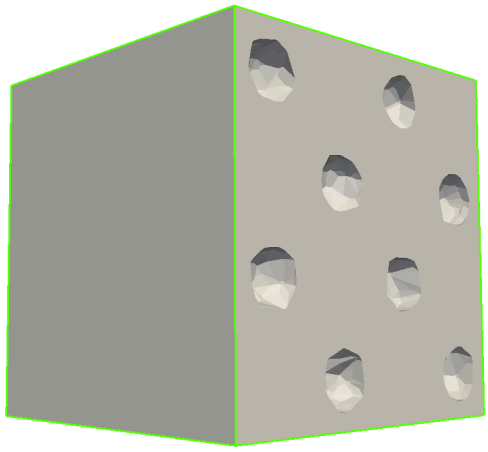
3.6.4 3D Gripper

We consider a 3D gripper as shown in Fig.3.17a. The gripper is supported on two circular hole-like regions, denoted by Γ_r. The friction is assumed to be absent and these regions act like pin joints. Here as well, there are rigid contact supports only, hence we apply remeshing.

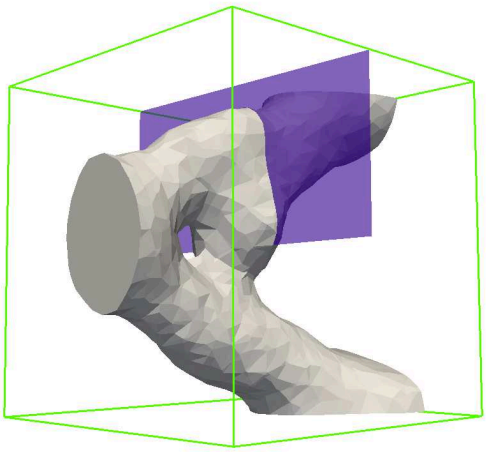
The gripper is forced on Γ_N^1 (the teeth in green) and on Γ_N^2 (the blue plane) with $\boldsymbol{g}_1 = (0.6\text{sign}(y), 0, 0)\text{N}/m^2$ and $\boldsymbol{g}_2 = (-1, 0, 0)\text{N}/m^2$, respectively. The term $\text{sign}(y)$ in $\boldsymbol{g}_1$ simply ensures that the direction of the force is against the teeth of the gripper. A target volume $V_f = 0.175m^3$ is imposed. The initial shape in Fig.(3.17b) converges to the shape in Fig.3.17d. A similar test case for linear elasticity is considered where the frictionless boundary condition is replaced by a fixed boundary Γ_D. The results in the gripper (3.17c), which is very similar to the gripper (3.17d).



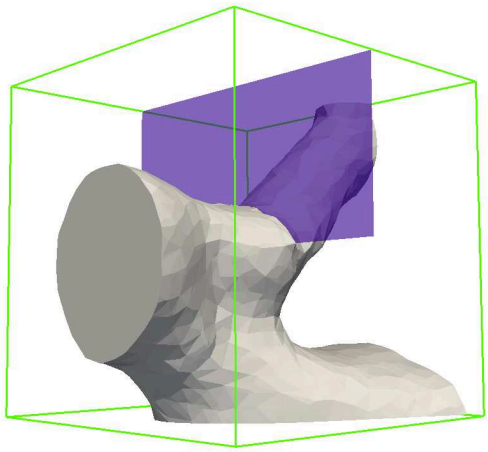
(a) Boundary conditions



(b) Initial shape



(c) Final shape, $\boldsymbol{g} = (0, 10, 0)\mathrm{N}/m^2$



(d) Final shape, $\boldsymbol{g} = (0, -10, 0)\mathrm{N}/m^2$

Figure 3.15: Compliance (2.3) minimization, $V_f = 0.25m^3, \mu = 0$

3.7 Application to idealized bolt setting

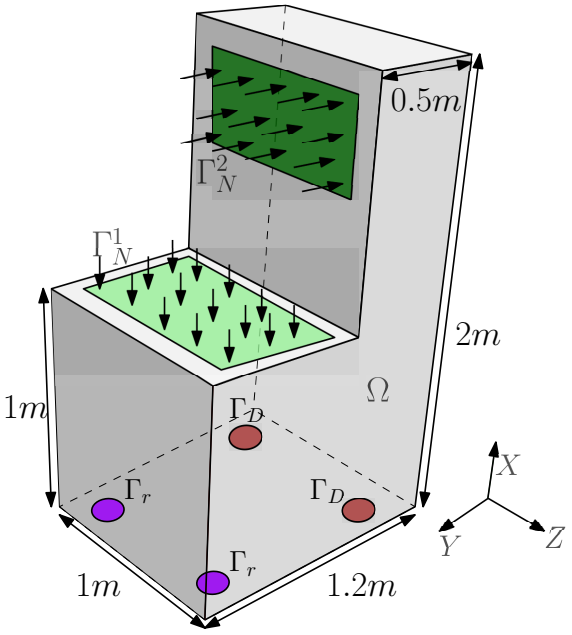
The work in the previous sections was done by the author all alone. On the contrary, the work in this section was done by the author in collaboration with Lalaina Rakontodrainibe, who worked on the idealized bolt model and its numerical implementation on `FreeFEM`. For more details on the idealized bolt model, the reader can refer to [151, 150].

Condensed in one line, the idealized bolt model aims at modeling mechanical connections (like bolts) in structures using an ideal spring. More details on the model shall be given in the Subsection 3.7.1. The model then shall be cast as a variational inequality in Subsection 3.7.2. This inequality shall be approximated using penalization and regularization (like in the case of contact problem above) in Subsection 3.7.3. The shape derivative shall be derived in Subsection 3.7.4 and a strategy to optimize the location of the bolts shall be proposed in Subsection 3.7.5. The numerical details shall be briefly discussed in Subsection 3.7.6. Finally, the a few 3D results shall be presented in Subsection 3.7.7.

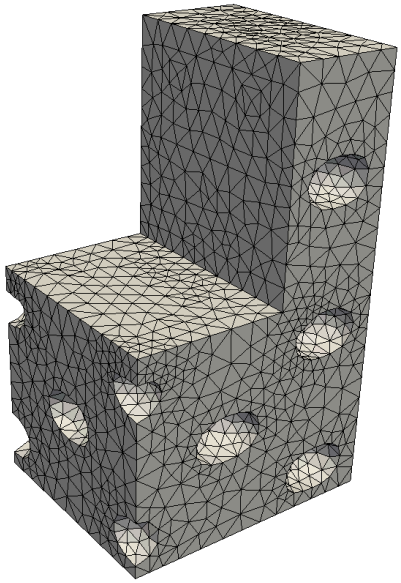
3.7.1 Governing laws

The idealized bolt-contact model consists of linear elasticity (momentum balance equations 3.4) coupled to

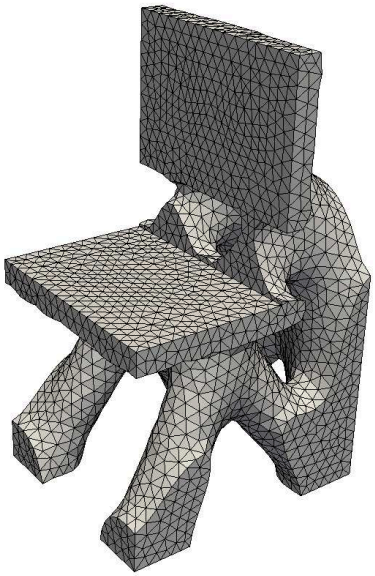
- contact boundary conditions (see subsections 3.1.1-(3.1.2))



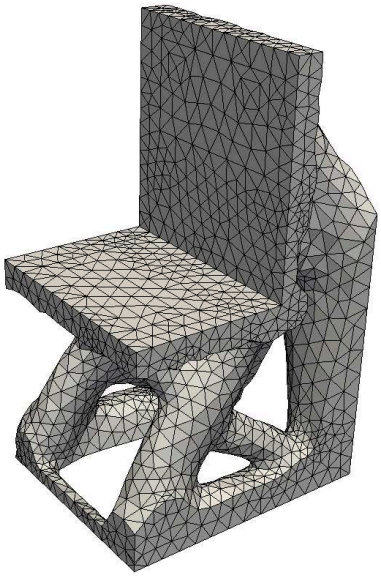
(a) Boundary conditions



(b) Initial shape



(c) Linear elasticity (assuming Γ_r to be Γ_D)



(d) Frictionless contact (3.22)

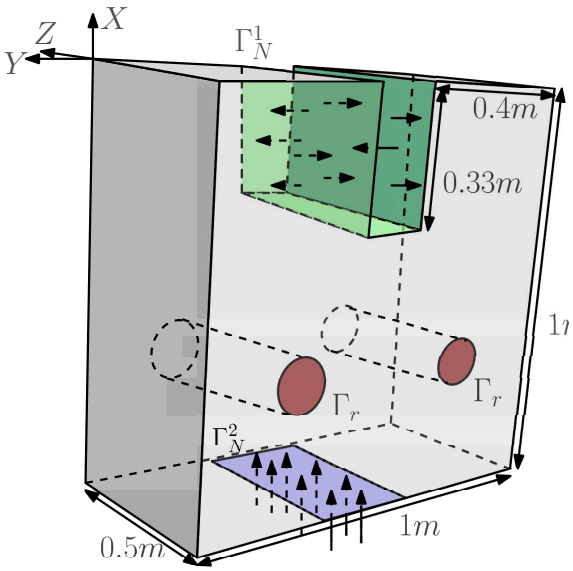
Figure 3.16: Meshes obtained for compliance (3.50) minimization

- idealized bolt model (see below).

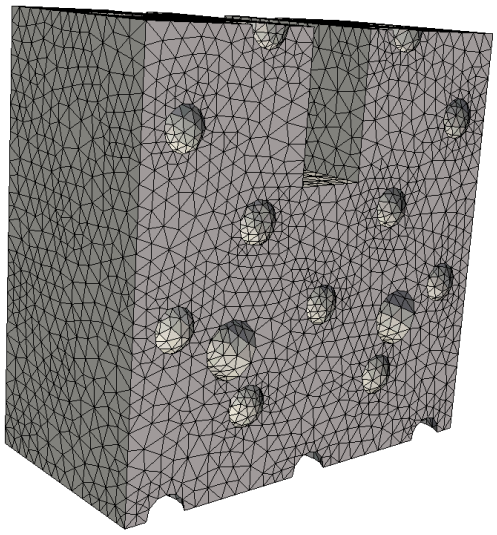
We now connect the two structures Ω_A and Ω_B in Fig.3.1 using a bolt. This bolt can be simplified as shown in Fig.3.18 and assumed to act like a spring. Hence the name, idealized bolt model. Before the application of the forces $\boldsymbol{f}$ and $\boldsymbol{g}$, the bolt is tightened. This tightening induces a pre-stress and an initial displacement $\boldsymbol{u}_0$ (step 1). We then apply external loads $\boldsymbol{f}$ and $\boldsymbol{g}$ to these pre-stressed structures and determine the final displacement $\boldsymbol{u}$ (step 2), as shown in Fig. 3.19b. Hence there are two steps, formulated in (3.55) and (3.56). At both the steps, the Signorini contact boundary conditions (3.6)-(3.7) and the friction laws (3.9)-(3.10) must be satisfied.

We first introduce the notations needed for the idealized bolt model, and then present the governing equations. Let

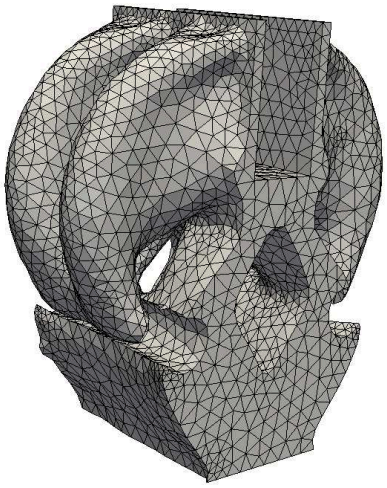
- $\rho > 0$ be a dimensional-less factor



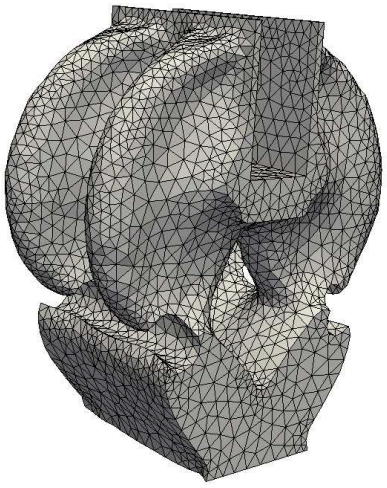
(a) Boundary conditions



(b) Initial shape



(c) Linear elasticity (replacing Γ_r by Γ_D)



(d) Frictionless contact (3.22)

Figure 3.17: Compliance (3.50) minimization, target volume $V_f = 0.175m^3$

- ω be a shape representing one of the ends of the bolt
- $\boldsymbol{e}$ be a unit vector along the orientation of the bolt
- $\ell > 0$ be the length of the bolt
- ω_A, ω_B be small inclusions each of shape ω, centered at $\boldsymbol{x}_0$ and $\boldsymbol{x}_0 - \ell\boldsymbol{e}$ resp. and given by

$$\omega_A = \left\{ \boldsymbol{x} \in \mathbb{R}^d, \frac{\boldsymbol{x} - \boldsymbol{x}_0}{\rho} \in \omega \right\} \text{ and } \omega_B = \left\{ \boldsymbol{x} \in \mathbb{R}^d, \frac{\boldsymbol{x} - \boldsymbol{x}_0 - \ell\boldsymbol{e}}{\rho} \in \omega \right\}$$

- $\boldsymbol{\Phi}$ be the pretension force on the pair of elastic spheres as shown in Fig.3.19a
- $\boldsymbol{u}_0 : \Omega \to \mathbb{R}^d$ be the displacement field of the two structures at step 1
- $\boldsymbol{u} : \Omega \to \mathbb{R}^d$ be the displacement field of the two structures at step 2

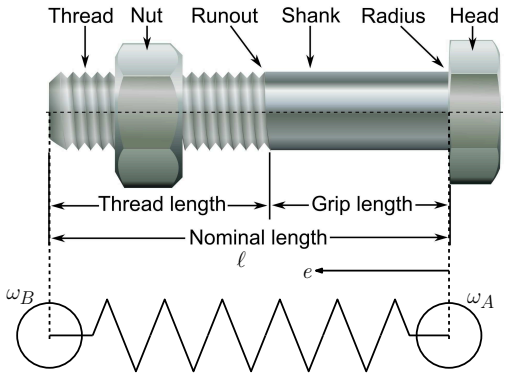
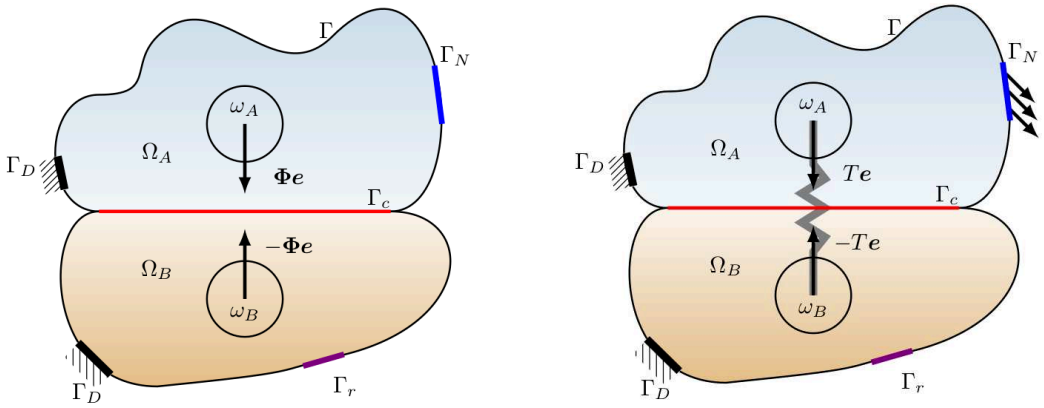


Figure 3.18: A 1/2-13UNC-2A-3 (ANSI (American National Standards Institute) nomenclature [83]) bolt [1](on top); Idealized bolt (on bottom)



(a) Step 1

(b) Step 2

- κ be the stiffness of the bolt

The elongation of the bolt/spring is the difference between the average displacements in ω_A and ω_B along the axis of the spring.

We define the average of a function $\varphi \in L^1(\Omega)$ on a domain ω as

$$\fint_\omega \varphi \, dx = \frac{1}{|\omega|} \int_\omega \varphi \, dx. \tag{3.52}$$

3.7.1.1 Step 1

Figure 3.18 shows the different parts of the bolt. In practice, the bolt connecting the structures is tightened by applying a torque to its head. This creates a traction and torsion in the shank of the bolt along with local compressive stresses in the structure. The tightening torsion is ignored here. A prescribed tensile force is applied between both the extremities of the shank to induces pre-stress, and the axial, torsional elongations of the shank are ignored. This approximation is justified due to Euler-Bernoulli beam theory, exactly as proposed in commercial software, such as ABAQUS [175] and NX NASTRAN [141].

By background domain, we mean the zone in Ω where one wishes to add a small idealized bolt connection. This domain may already contain n pre-stressed idealized bolts. The n bolts, idealized as springs are modeled by the function $\mathcal{S}$ given by

$$\mathcal{S}(\boldsymbol{w}) = \sum_{i=1}^{n} \kappa_i \left(\fint_{\omega_{B_i}} \boldsymbol{w} \cdot \boldsymbol{e}_i dx - \fint_{\omega_{A_i}} \boldsymbol{w} \cdot \boldsymbol{e}_i dx \right) \left(\frac{\mathbf{1}_{\omega_{B_i}}}{|\omega_{B_i}|} - \frac{\mathbf{1}_{\omega_{A_i}}}{|\omega_{A_i}|} \right) \boldsymbol{e}_i, \tag{3.53}$$

[1] http://www.gizmology.net/nutsbolts.html

where $\mathbf{1}_{\omega_{A_i}}, \mathbf{1}_{\omega_{B_i}}$ denote the characteristic function of the *i-th* bolt extremities (assumed to have a shape ω_{A_i} and ω_{B_i}). The n tension forces imposed on the n springs are then modeled by

$$\boldsymbol{t}(\boldsymbol{u}) = \sum_{i=1}^{n} \left(\kappa_i \left(\fint_{\omega_{B_i}} \boldsymbol{u}\cdot\boldsymbol{e}_i dx - \fint_{\omega_{A_i}} \boldsymbol{u}\cdot\boldsymbol{e}_i dx \right) - \boldsymbol{\Phi}_i \right) \left(\frac{\mathbf{1}_{\omega_{B_i}}}{|\omega_{B_i}|} - \frac{\mathbf{1}_{\omega_{A_i}}}{|\omega_{A_i}|} \right) \boldsymbol{e}_i, \tag{3.54}$$

where $\boldsymbol{\Phi}_i$ is the pretension force, which results from the first step of the pre-stressed state of the spring i of axis $\boldsymbol{e}_i$. The functions $\mathcal{S}$ and $\boldsymbol{t}$ are zero if there is no finite size idealized bolt in the background domain.

Using the above notations, the first step can then be given by the following equations

$$\begin{cases} -\mathrm{div}(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}_0)) + \mathcal{S}(\boldsymbol{u}_0) = \boldsymbol{t}(\mathbf{0}) & \text{in } \Omega \\ \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}_0)\boldsymbol{n} = \mathbf{0} & \text{on } \Gamma_N \cup \Gamma \\ \boldsymbol{u}_0 = \mathbf{0} & \text{on } \Gamma_D \\ \boldsymbol{u}_0\cdot\boldsymbol{n} \le 0 & \text{on } \Gamma_r \\ [\boldsymbol{u}_0]\cdot\boldsymbol{n}^- \le 0 & \text{on } \Gamma_c \\ \mathcal{F}_{\Gamma_r}(\boldsymbol{u}_0) \le 0 & \text{on } \Gamma_r \\ \mathcal{T}_{\Gamma_c^-}(\boldsymbol{u}_0) = -\mathcal{T}_{\Gamma_c^+}(\boldsymbol{u}_0),\ \mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}_0) = \mathcal{F}_{\Gamma_c^+}(\boldsymbol{u}_0) \le 0 & \text{on } \Gamma_c \\ |\mathcal{T}_{\Gamma_r}(\boldsymbol{u}_0)| \le \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u}_0)|,\ (|\mathcal{T}_{\Gamma_r}(\boldsymbol{u}_0)| - \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u}_0)|)\boldsymbol{u}_{0t} = 0 & \text{on } \Gamma_r \\ |\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u}_0)| \le \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}_0)|,\ (|\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u}_0)| - \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}_0)|[\boldsymbol{u}_{0t}] = 0 & \text{on } \Gamma_c \end{cases}. \tag{3.55}$$

In the absence of any tightened bolt, the displacement field $\boldsymbol{u}_0$ is zero.

3.7.1.2 Step 2

External loads are applied in the structure after applying a pre-stress in step 1. The solution $\boldsymbol{u}$ satisfies the following system of equations

$$\begin{cases} -\mathrm{div}(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})) + \mathcal{S}(\boldsymbol{u}) = \boldsymbol{f} + \boldsymbol{t}(\boldsymbol{u}_0) & \text{in } \Omega \\ \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\boldsymbol{n} = \boldsymbol{g} & \text{on } \Gamma_N \\ \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\boldsymbol{n} = \mathbf{0} & \text{on } \Gamma \\ \boldsymbol{u} = \mathbf{0} & \text{on } \Gamma_D \\ \boldsymbol{u}\cdot\boldsymbol{n} \le 0 & \text{on } \Gamma_r, \\ [\boldsymbol{u}]\cdot\boldsymbol{n}^- \le 0 & \text{on } \Gamma_c \\ \mathcal{F}_{\Gamma_r}(\boldsymbol{u}) \le 0 & \text{on } \Gamma_r \\ \mathcal{T}_{\Gamma_c^-}(\boldsymbol{u}) = -\mathcal{T}_{\Gamma_c^+}(\boldsymbol{u}),\ \mathcal{F}_{\Gamma_c^-}(\boldsymbol{u}) = \mathcal{F}_{\Gamma_c^+}(\boldsymbol{u}) \le 0 & \text{on } \Gamma_c \\ |\mathcal{T}_{\Gamma_r}(\boldsymbol{u})| \le \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u})|,\ (|\mathcal{T}_{\Gamma_r}(\boldsymbol{u})| - \mu|\mathcal{F}_{\Gamma_r}(\boldsymbol{u})|)\boldsymbol{u}_t = 0 & \text{on } \Gamma_r \\ |\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u})| \le \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u})|,\ (|\mathcal{T}_{\Gamma_c^-}(\boldsymbol{u})| - \mu|\mathcal{F}_{\Gamma_c^-}(\boldsymbol{u})|[\boldsymbol{u}_t] = 0 & \text{on } \Gamma_c \end{cases}. \tag{3.56}$$

3.7.2 Variational formulation

Introduce two bilinear forms $a_r : V \times V \to \mathbb{R}$ and $h : V \times V^* \to \mathbb{R}$ defined respectively by

$$a_r(\boldsymbol{u},\boldsymbol{v}) = a(\boldsymbol{u},\boldsymbol{v}) + \int_\Omega \mathcal{S}(\boldsymbol{u})\boldsymbol{v}\, dx, \tag{3.57}$$

$$h(\boldsymbol{u},\boldsymbol{v}) = \int_\Omega t(\boldsymbol{u})\boldsymbol{v}\, dx, \tag{3.58}$$

$\kappa(\rho)$ is the stiffness of the bolt and is positive and a is defined in (3.13). Quite easily, one can show that $a_r(\cdot,\cdot)$ is a symmetric bilinear forms. Moreover, $a_r(\cdot,\cdot)$ and $h(\cdot,\cdot)$ are continuous on $V \times V$. Under the assumptions **V.1** and **V.2**, the bilinear form $a_r(\cdot,\cdot)$ can be shown to be coercive.

Lemma 5. *The solution to Step 1 (3.55), $\boldsymbol{u}_0$ and the solution to Step 2 (3.56) satisfy: $(\boldsymbol{u}_0, \boldsymbol{u}) \in K \times K$ such that*

$$a_r(\boldsymbol{u}_0, \boldsymbol{v}-\boldsymbol{u}_0) + j_c(\boldsymbol{u}_0,\boldsymbol{v}) - j_c(\boldsymbol{u}_0,\boldsymbol{u}_0) \ge h(\mathbf{0}, \boldsymbol{v}-\boldsymbol{u}_0) \qquad \forall \boldsymbol{v} \in K \tag{3.59}$$

$$a_r(\boldsymbol{u}, \boldsymbol{v}-\boldsymbol{u}) + j_c(\boldsymbol{u},\boldsymbol{v}) - j_c(\boldsymbol{u},\boldsymbol{u}) \ge l(\boldsymbol{v}-\boldsymbol{u}) + h(\boldsymbol{u}_0, \boldsymbol{v}-\boldsymbol{u}) \qquad \forall \boldsymbol{v} \in K. \tag{3.60}$$

Proof. Applying the same steps as in proof of Lemma 2, we can obtain the formulations (3.59)-(3.60). □

The inequalities (3.59)-(3.60) are very similar to the (3.17) as they possess the same nonlinear function j_c. The next theorem talks about the well-posedness of the (3.59)-(3.60).

Theorem 8. *Let $\Omega \subset \mathbb{R}^d$ be smooth and Lipschitz. If the friction coefficient μ satifies the bound (3.19), the problem (3.59)-(3.60) admits a unique solution $\boldsymbol{u}_0, \boldsymbol{u} \in K$.*

Proof. Fortunately, the equations (3.59)-(3.60) are decoupled in $(\boldsymbol{u}_0, \boldsymbol{u})$. The Theorem 3 in [71] shows existence and uniqueness of $\boldsymbol{u} \in K$ to the problem (3.17). The proof of this theorem uses the coercivity of the bilinear form $a(\cdot,\cdot)$ and the continuity of the linear form $l(\cdot)$. Given that the bilinear form $a_r(\cdot,\cdot)$ is coercive and the linear form, $h(\mathbf{0},\cdot)$ continuous, the Theorem 3 of [71] can be easily extended to show the existence of a solution to $\boldsymbol{u}_0 \in K$ to problem (3.59). For the solution $\boldsymbol{u}_0 \in K$, the linear form $l(\cdot) + h(\boldsymbol{u}_0,\cdot)$ can be shown to be continuous on V and the Theorem 3 can be again applied to show the existence and uniqueness of $\boldsymbol{u} \in K$. □

Remark 2. *Consider the frictionless case, $\mu = 0$. The solution $(\boldsymbol{u}_0, \boldsymbol{u})$ is scalable with respect to the forces $\boldsymbol{f}, \boldsymbol{g}$ and the pretension $\boldsymbol{\Phi}_i$. If the pretension is kept constant and the forces are increased, then the solution is no longer scalable. We shall show the influence of increasing the external forces at a fixed pretension on the optimized shapes in the Section 3.7.7.*

3.7.3 Approximation via penalization and regularization

As discussed in Section 3.3, the contact problems (3.16) and (3.17) are not shape-differentiable in the classical sense and ought to be approximated using penalization and regularization approach. Likewise, (3.59)-(3.60) a variational inequalities, which ought to be approximated using penalization and regularization technique.

Following the idea in Subsection 3.3.1, we enforce the Signorini contact boundary conditions using the term j_ϵ (see (3.20)) and the Coulomb friction boundary conditions using $j_{\epsilon,\eta}$ (see Subsection 3.3.2) on the idealized bolt model (3.59)-(3.60). This results in the following regularized-penalized idealized bolt model: find $(\boldsymbol{u}_{0,\epsilon,\eta}, \boldsymbol{u}_{\epsilon,\eta}) \in V \times V$ such that

$$a_r(\boldsymbol{u}_{0,\epsilon,\eta}, \boldsymbol{v}) + j'_\epsilon(\boldsymbol{u}_{0,\epsilon,\eta}) \cdot \boldsymbol{v} + j_{\epsilon,\eta}(\boldsymbol{u}_{0,\epsilon,\eta}, \boldsymbol{v}) = h(\mathbf{0}, \boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V \tag{3.61}$$

$$a_r(\boldsymbol{u}_{\epsilon,\eta}, \boldsymbol{v}) + j'_\epsilon(\boldsymbol{u}_{\epsilon,\eta}) \cdot \boldsymbol{v} + j_{\epsilon,\eta}(\boldsymbol{u}_{\epsilon,\eta}, \boldsymbol{v}) = l(\boldsymbol{v}) + h(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V. \tag{3.62}$$

This is the set of equations we shall be dealing with for the determining the shape derivative. Henceforth, they shall be called state equations and the solution corresponding to them shall be called the state solutions.

Remark 3. *Note that the penalized frictionless contact model (3.22) can be viewed as a special case of (3.62) with $\boldsymbol{u}_{0,\epsilon,\eta} = \mathbf{0}$ (absence of pretension) $\mu = 0$. Similarly, the penalized regularized friction contact model (3.26) can be recovered from (3.62) with $\boldsymbol{u}_{0,\epsilon,\eta} = \mathbf{0}$.*

Showing the strong convergence of the sequence of solutions $(\boldsymbol{u}_{0,\epsilon,\eta}, \boldsymbol{u}_{\epsilon,\eta}) \longrightarrow (\boldsymbol{u}_0, \boldsymbol{u})$ (solution to (3.59)-(3.60)) as $\epsilon \longrightarrow 0, \eta \longrightarrow 0$ is beyond the scope of this chapter. We can however show the convergence for the frictionless case ($\mu = 0$). The problem (3.61)-(3.62) in the absence of friction can be written as: find $(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{u}_\epsilon) \in V \times V$ such that

$$\boldsymbol{u}_{0,\epsilon} = \min_{\boldsymbol{v} \in V} \left(\frac{1}{2} a_r(\boldsymbol{v}, \boldsymbol{v}) - h(\mathbf{0}, \boldsymbol{v}) + j_\epsilon(\boldsymbol{v}) \right), \tag{3.63}$$

$$\boldsymbol{u}_\epsilon = \min_{\boldsymbol{v} \in V} \left(\frac{1}{2} a_r(\boldsymbol{v}, \boldsymbol{v}) - l(\boldsymbol{v}) - h(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v}) + j_\epsilon(\boldsymbol{v}) \right) \tag{3.64}$$

where we have dropped the index η as the term (3.27) is absent. To simplify the presentation, we consider the case of a single bolt (instead of multiple bolts), having a pretension $\boldsymbol{\Phi}$. The following theorem show the strong convergence of the solution to (3.63)-(3.64).

Theorem 9. *The minimization problem (3.63)-(3.64) admits a unique solution $(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{u}_\epsilon) \in V \times V$, satisfying the bound*

$$\|\boldsymbol{u}_{0,\epsilon}\|_V + \|\boldsymbol{u}_\epsilon\|_V \leq C,$$

where C is a constant independent of ϵ. Moreover, the sequence of solutions $(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{u}_\epsilon)$ converges strongly to

$$\boldsymbol{u}_{0,\epsilon} \longrightarrow \boldsymbol{u}_0, \qquad \text{and} \qquad \boldsymbol{u}_\epsilon \longrightarrow \boldsymbol{u} \qquad \text{in } V.$$

where $(\boldsymbol{u}_0, \boldsymbol{u})$ is the solution to the problem (3.59)-(3.60).

Proof. The minimization problem (3.63) can be equivalently written in the form of an inequality of the second kind: find $\boldsymbol{u}_{0,\epsilon} \in V$ such that

$$a_r(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v} - \boldsymbol{u}_{0,\epsilon}) + j_\epsilon(\boldsymbol{v}) - j_\epsilon(\boldsymbol{u}_{0,\epsilon}) \geq h(\boldsymbol{0}, \boldsymbol{v} - \boldsymbol{u}_{0,\epsilon}) \quad \forall \boldsymbol{v} \in V. \tag{3.65}$$

where the form $a_r(\cdot,\cdot)$ is coercive, the nonlinear form j_ϵ is convex lower semi continuous and proper. Thus, it admits a unique solution $\boldsymbol{u}_{0,\epsilon} \in V$ [80]. Similarly, (3.63) written under the form

$$a_r(\boldsymbol{u}_\epsilon, \boldsymbol{v} - \boldsymbol{u}_\epsilon) + j_\epsilon(\boldsymbol{v}) - j_\epsilon(\boldsymbol{u}_\epsilon) \geq l(\boldsymbol{u}_\epsilon, \boldsymbol{v} - \boldsymbol{u}_\epsilon) + h(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v} - \boldsymbol{u}_\epsilon) \quad \forall \boldsymbol{v} \in V \tag{3.66}$$

admits a unique solution $\boldsymbol{u}_\epsilon \in V$.

Now, we obtain a bound on $\boldsymbol{u}_{0,\epsilon}, \boldsymbol{u}_\epsilon$. To simply the presentation, assume that there is only one bolt, having a shape ω_A, ω_B, stiffness κ and a pretension $\boldsymbol{\Phi}$. Let α and M be the coercivity and continuity constants of $a_r(\cdot,\cdot)$ In order to obtain the bound on $\boldsymbol{u}_{0,\epsilon}$, we take $\boldsymbol{v} = \boldsymbol{v}_0 \in K$ (an arbitrary element) in (3.64) and obtain

$$\begin{aligned}
\alpha\|\boldsymbol{u}_{0,\epsilon}\|_V^2 \leq a_r(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{u}_{0,\epsilon}) &\leq a_r(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v}_0) + j_\epsilon(\boldsymbol{v}_0) - j_\epsilon(\boldsymbol{u}_{0,\epsilon}) + h(\boldsymbol{0}, \boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0) \\
&= a_r(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v}_0) - j_\epsilon(\boldsymbol{u}_{0,\epsilon}) + h(\boldsymbol{0}, \boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0) \qquad (\text{since } \boldsymbol{v}_0 \in K) \\
&\leq a_r(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v}_0) + h(\boldsymbol{0}, \boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0) \\
&\leq M \|\boldsymbol{u}_{0,\epsilon}\|_V \|\boldsymbol{v}_0\|_V + \boldsymbol{\Phi} \left(\fint_{\omega_B} |(\boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0) \cdot \boldsymbol{e}_i| \, dx + \fint_{\omega_A} |(\boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0) \cdot \boldsymbol{e}_i| \, dx \right) \\
&\leq M \|\boldsymbol{u}_{0,\epsilon}\|_V \|\boldsymbol{v}_0\|_V + \frac{\boldsymbol{\Phi}}{\min(|\omega_B|, |\omega_A|)} \left(\int_\Omega |\boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0| \, dx \right) \\
&\leq M \|\boldsymbol{u}_{0,\epsilon}\|_V \|\boldsymbol{v}_0\|_V + \frac{\boldsymbol{\Phi}}{\min(|\omega_B|, |\omega_A|)} \left(\int_\Omega |\boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0| \, dx \right) \\
&\leq M \|\boldsymbol{u}_{0,\epsilon}\|_V \|\boldsymbol{v}_0\|_V + \frac{\boldsymbol{\Phi}}{\min(|\omega_B|, |\omega_A|)} \sqrt{|\Omega|} \, \|\boldsymbol{u}_{0,\epsilon} - \boldsymbol{v}_0\|_V \\
&\leq (C_1 + M \|\boldsymbol{v}_0\|_V) \|\boldsymbol{u}_{0,\epsilon}\|_V + C_1 \|\boldsymbol{v}_0\|_V \\
&\leq \frac{\alpha}{2} \|\boldsymbol{u}_{0,\epsilon}\|_V^2 + \frac{1}{2\alpha} (M \|\boldsymbol{v}_0\|_V + C_1)^2 + C_1 \|\boldsymbol{v}_0\|_V \\
&\leq \frac{1}{\alpha} (M \|\boldsymbol{v}_0\|_V + C_1)^2 + 2C_1 \|\boldsymbol{v}_0\|_V = C_2
\end{aligned}$$

where C_1, C_2 depend on the forces $\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{v}_0$, κ but not ϵ. Similarly, we take $\boldsymbol{v} = \boldsymbol{v}_0 \in K$ in (3.64)

$$\begin{aligned}
\alpha \|\boldsymbol{u}_\epsilon\|_V^2 \leq a_r(\boldsymbol{u}_\epsilon, \boldsymbol{u}_\epsilon) &\leq M \|\boldsymbol{u}_\epsilon\|_V \|\boldsymbol{v}_0\|_V \\
&\quad + \frac{1}{\min(|\omega_B|, |\omega_A|)} \left(\boldsymbol{\Phi} + \frac{\kappa}{\min(|\omega_B|, |\omega_A|)} \int_\Omega |\boldsymbol{u}_{0,\epsilon}| \, dx \right) \left(\int_\Omega |\boldsymbol{u}_\epsilon - \boldsymbol{v}_0| \, dx \right) \\
&\leq M \|\boldsymbol{u}_\epsilon\|_V \|\boldsymbol{v}_0\|_V + \frac{1}{\min(|\omega_B|, |\omega_A|)} \left(\boldsymbol{\Phi} + \frac{\kappa\sqrt{|\Omega|}}{\min(|\omega_B|, |\omega_A|)} \|\boldsymbol{u}_{0,\epsilon}\|_V \right) \|\boldsymbol{u}_\epsilon - \boldsymbol{v}_0\|_V \\
&\leq M \|\boldsymbol{u}_\epsilon\|_V \|\boldsymbol{v}_0\|_V + \left(\frac{1}{\min(|\omega_B|, |\omega_A|)} \right) \left(\boldsymbol{\Phi} + \frac{\kappa\sqrt{|\Omega|}}{\min(|\omega_B|, |\omega_A|)} \|\boldsymbol{u}_{0,\epsilon}\|_V \right) \|\boldsymbol{u}_\epsilon - \boldsymbol{v}_0\|_V \\
&\leq M \|\boldsymbol{u}_\epsilon\|_V \|\boldsymbol{v}_0\|_V + C_3 \|\boldsymbol{u}_\epsilon - \boldsymbol{v}_0\|_V + C_4 \|\boldsymbol{u}_{0,\epsilon}\|_V \|\boldsymbol{u}_\epsilon - \boldsymbol{v}_0\|_V \\
&\leq \left(C_3 + M \|\boldsymbol{v}_0\|_V + C_4 \|\boldsymbol{u}_{0,\epsilon}\|_V\right) \|\boldsymbol{u}_\epsilon\|_V + (C_3 + C_4 \|\boldsymbol{u}_{0,\epsilon}\|_V) \|\boldsymbol{v}_0\|_V \\
&\leq \frac{1}{\alpha} \left(C_3 + M \|\boldsymbol{v}_0\|_V + C_4 \sqrt{\frac{C_2}{\alpha}} \right)^2 + \left(C_3 + C_4 \sqrt{\frac{C_2}{\alpha}} \right) \|\boldsymbol{v}_0\|_V = C_5
\end{aligned}$$

where again, C_3, C_4, C_5 are constants independent of ϵ. Thus we have the bounds

$$\|\boldsymbol{u}_{0,\epsilon}\|_V \leq \sqrt{\frac{C_2}{\alpha}} \qquad \text{and} \qquad \|\boldsymbol{u}_\epsilon\|_V \leq \sqrt{\frac{C_5}{\alpha}}.$$

Given the above bounds in V, we deduce that there exists a weakly convergent subsequence $\boldsymbol{u}_{0,\epsilon^k}, \boldsymbol{u}_{\epsilon^k} \in V$ such that

$$\boldsymbol{u}_{0,\epsilon^k} \rightharpoonup \boldsymbol{u}_0^*, \qquad \text{and} \qquad \boldsymbol{u}_{\epsilon^k} \rightharpoonup \boldsymbol{u}^* \qquad \text{in } V.$$

Using the fact that $a_r(\cdot,\cdot)$ is inf-sup stable, we can easily pass to the limit $\epsilon \longrightarrow 0$ in (3.65) and establish $\boldsymbol{u}_0^* = \boldsymbol{u}_0$. The inequality (3.65) is an inequality of the second kind. For such an inequality, using Theorem 7.1 [80], we state that there exists a subsequence $\boldsymbol{u}_{0,\epsilon^k}$

$$\boldsymbol{u}_{0,\epsilon^k} \longrightarrow \boldsymbol{u}_0 \qquad \text{in } V.$$

Since the limit $\boldsymbol{u}_0$ is unique, we state that all the subsequences converge to $\boldsymbol{u}_0$. Given the strong convergence $\boldsymbol{u}_{0,\epsilon} \longrightarrow \boldsymbol{u}_0$ and the weak convergence $\boldsymbol{u}_{0,\epsilon^k} \rightharpoonup \boldsymbol{u}_0$ we can show that as $\epsilon \longrightarrow 0$, the functional

$$h(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v} - \boldsymbol{u}_\epsilon) \longrightarrow h(\boldsymbol{u}_{0,\epsilon}, \boldsymbol{v} - \boldsymbol{u}_\epsilon).$$

Using this and the inf-sup stability of $a_r(\cdot,\cdot)$, we can pass to the limit $\epsilon \longrightarrow 0$ in (3.66) and establish $\boldsymbol{u}^* = \boldsymbol{u}$. Finally, we use the use Theorem 7.1 [80] again and state that the subsequence $\boldsymbol{u}_{\epsilon^k}$

$$\boldsymbol{u}_{\epsilon^k} \longrightarrow \boldsymbol{u} \qquad \text{in } V.$$

□

3.7.4 Shape derivative computation

To simplify notations, we denote $(\boldsymbol{u}_{0,\epsilon,\eta}, \boldsymbol{u}_{\epsilon,\eta})$ by $(\boldsymbol{u}_0, \boldsymbol{u})$. We consider an objective function (similar to (3.28)) of the following form

$$J(\Omega) = \int_\Omega m(\boldsymbol{u}_0(\Omega), \boldsymbol{u}(\Omega))dx + \int_{\Gamma_N} p(\boldsymbol{u}_0(\Omega), \boldsymbol{u}(\Omega))ds, \tag{3.67}$$

where $(\boldsymbol{u}_0(\Omega), \boldsymbol{u}(\Omega))$ are solution to the problem (3.61)-(3.62), and $m(\cdot)$, $p(\cdot)$ are assumed to be Lipschitz and $\mathcal{C}^1$ smooth. For the position of the bolt, we either let it fixed, or move its location along a derivative computed in Subsection 3.7.5. The shape derivative can be computed using the Céa's technique as in Subsection 3.4.1. Here again, we must assume the adjoint equation to be well-posed and perform the computation. We state the following theorem, which gives the expression of the shape derivative.

Theorem 10. *Let Ω_A and Ω_B be two smooth bounded domains sharing a boundary Γ_c and $\Omega = \Omega_A \cup \Omega_B$. Let $\boldsymbol{f} \in H^1(\mathbb{R}^d)^d$, $\boldsymbol{g} \in H^2(\mathbb{R}^d)^d$ and $(\boldsymbol{u}_0(\Omega), \boldsymbol{u}(\Omega)) \in V \times V$ be the solution to (3.61)-(3.62). Then the shape derivative of $J(\Omega)$ (3.67) along $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, is given by*

$$\begin{aligned} J'(\Omega)\boldsymbol{\theta} = \int_\Gamma \boldsymbol{\theta}\cdot\boldsymbol{n}\,(m(\boldsymbol{u}_0,\boldsymbol{u}) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}_0)\colon\boldsymbol{\varepsilon}(\boldsymbol{v}_0) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u})\colon\boldsymbol{\varepsilon}(\boldsymbol{v}) + \mathcal{S}(\boldsymbol{u}_0)\boldsymbol{v}_0 + \mathcal{S}(\boldsymbol{u})\boldsymbol{v} + M_\epsilon(\boldsymbol{u}_0)\boldsymbol{v}_0 \\ + M_\epsilon(\boldsymbol{u})\boldsymbol{v} - h(\boldsymbol{0},\boldsymbol{v}_0) - h(\boldsymbol{u}_0,\boldsymbol{v}) + \boldsymbol{f}\cdot\boldsymbol{v})\,ds, \end{aligned} \tag{3.68}$$

where $(\boldsymbol{v}_0, \boldsymbol{v})$ is the solution to the adjoint problem:

$$\begin{aligned} a_r(\boldsymbol{\varphi},\boldsymbol{v}_0) + (j_\epsilon''(\boldsymbol{u}_0)\cdot\boldsymbol{\varphi})\cdot\boldsymbol{v}_0 + \partial_{\boldsymbol{u}_0} j_{\epsilon,\eta}(\boldsymbol{u}_0,\boldsymbol{v}_0)\cdot\boldsymbol{\varphi} \\ = h(\boldsymbol{\varphi},\boldsymbol{v}) - \int_\Omega \partial_{\boldsymbol{u}_0} m(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\varphi}\,dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}_0} p(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\varphi}\,ds \qquad \forall\boldsymbol{\varphi}\in V, \end{aligned} \tag{3.69a}$$

$$a_r(\boldsymbol{\psi},\boldsymbol{v}) + (j_\epsilon''(\boldsymbol{u})\cdot\boldsymbol{\psi})\cdot\boldsymbol{v} + \partial_{\boldsymbol{u}} j_{\epsilon,\eta}(\boldsymbol{u},\boldsymbol{v})\cdot\boldsymbol{\psi} = -\int_\Omega \partial_{\boldsymbol{u}} m(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\psi}\,dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}} p(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\psi}\,ds \qquad \forall\boldsymbol{\varphi}\in V, \tag{3.69b}$$

where $\mathcal{H}$ denotes the Heaviside function and the derivatives of $j'_\epsilon, j_{\epsilon,\eta}$ are defined in (3.33).

Proof. Introduce a space

$$\tilde{V} = \{\boldsymbol{v} \in H^1(\mathbb{R}^d) : \boldsymbol{v} = \boldsymbol{0} \qquad \text{on } \Gamma_D\}. \tag{3.70}$$

We use the fast derivation technique of Céa [53] by constructing the Lagrangian:

$$\begin{aligned}\mathcal{L}(\Omega,\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}_0,\tilde{\boldsymbol{v}}) = &\int_\Omega m(\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{u}})\,dx + \int_{\Gamma_N} p(\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{u}})ds + a_r(\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{v}}) + j'_\epsilon(\tilde{\boldsymbol{u}}_0)\cdot\boldsymbol{v} + j_{\epsilon,\eta}(\tilde{\boldsymbol{u}}_0,\boldsymbol{v}) \\ &- h(\mathbf{0},\tilde{\boldsymbol{v}}) + a_r(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}) + j'_\epsilon(\tilde{\boldsymbol{u}})\cdot\tilde{\boldsymbol{v}} + j_{\epsilon,\eta}(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}) - l(\tilde{\boldsymbol{v}}) - h(\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{v}}),\end{aligned} \tag{3.71}$$

where $\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{u}} \in \tilde{V}$ and $\tilde{\boldsymbol{v}}_0,\tilde{\boldsymbol{v}}$ are the adjoint variables. We mention that here the variables $\boldsymbol{u}_0,\boldsymbol{u},\boldsymbol{v}_0,\boldsymbol{v}$ are defined on the full space $\mathbb{R}^d$ and zero on Γ_0. Since Γ_D is non-optimizable, the variables $\boldsymbol{u}_0,\boldsymbol{u},\boldsymbol{v}_0,\boldsymbol{v}$ are independent of Ω. The optimality conditions applied to the Lagrangian in (3.71) results in the determination of the state equations (3.61)-(3.62), adjoint equation and the shape derivative. First we apply optimality with respect to the adjoint variables $(\boldsymbol{v}_0,\boldsymbol{v})$ along $(\boldsymbol{\varphi},\boldsymbol{\psi}) \in \tilde{V}\times\tilde{V}$ at $(\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{u}}) = (\boldsymbol{u}_0,\boldsymbol{u})$ and obtain

$$\begin{aligned} a_r(\boldsymbol{u}_0,\boldsymbol{\varphi}) + j'_\epsilon(\boldsymbol{u}_0)\cdot\boldsymbol{\varphi} + j_{\epsilon,\eta}(\boldsymbol{u}_0,\boldsymbol{\varphi}) &= h(\mathbf{0},\boldsymbol{\varphi}) && \forall\boldsymbol{\varphi}\in\tilde{V} \\ a_r(\boldsymbol{u},\boldsymbol{\psi}) + j'_\epsilon(\boldsymbol{u})\cdot\boldsymbol{\psi} + j_{\epsilon,\eta}(\boldsymbol{u},\boldsymbol{\psi}) &= l(\boldsymbol{\psi}) + h(\boldsymbol{u}_0,\boldsymbol{\psi}) && \forall\boldsymbol{\psi}\in\tilde{V}.\end{aligned}$$

Since all integrals in the above are defined only on Ω, we can replace $\tilde{V}$ by V and obtain (3.61)-(3.62). We then apply the optimality condition with respect to the state variables $(\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{u}})$ at $(\tilde{\boldsymbol{u}}_0,\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}_0,\tilde{\boldsymbol{v}}) = (\boldsymbol{u}_0,\boldsymbol{u},\boldsymbol{v}_0,\boldsymbol{v})$ and obtain

$$\begin{aligned} a_r(\boldsymbol{\varphi},\boldsymbol{v}_0) + (j''_\epsilon(\boldsymbol{u}_0)\cdot\boldsymbol{\varphi})\cdot\boldsymbol{v}_0 + \partial_{\boldsymbol{u}_0} j_{\epsilon,\eta}(\boldsymbol{u}_0,\boldsymbol{v}_0)\cdot\boldsymbol{\varphi} & \\ = h(\boldsymbol{\varphi},\boldsymbol{v}) - \int_\Omega \partial_{\boldsymbol{u}_0} m(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\varphi}\,dx - \int_{\Gamma_N}\partial_{\boldsymbol{u}_0} p(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\varphi}\,ds & \qquad \forall\boldsymbol{\varphi}\in V,\end{aligned}$$

$$a_r(\boldsymbol{\psi},\boldsymbol{v}) + (j''_\epsilon(\boldsymbol{u})\cdot\boldsymbol{\psi})\cdot\boldsymbol{v} + \partial_{\boldsymbol{u}} j_{\epsilon,\eta}(\boldsymbol{u},\boldsymbol{v})\cdot\boldsymbol{\psi} = -\int_\Omega \partial_{\boldsymbol{u}} m(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\psi}\,dx - \int_{\Gamma_N}\partial_{\boldsymbol{u}} p(\boldsymbol{u}_0,\boldsymbol{u})\cdot\boldsymbol{\psi}\,ds \qquad \forall\boldsymbol{\varphi}\in V.$$

$$\text{where } (j''_\epsilon(\boldsymbol{u})\cdot\boldsymbol{\psi})\cdot\boldsymbol{v} = \frac{1}{\epsilon}\left(\int_{\Gamma_r}\mathcal{H}(\boldsymbol{u}\cdot\boldsymbol{n})(\boldsymbol{\psi}\cdot\boldsymbol{n})(\boldsymbol{v}\cdot\boldsymbol{n})\,ds + \int_{\Gamma_c}\mathcal{H}([\boldsymbol{u}]\cdot\boldsymbol{n})([\boldsymbol{\psi}]\cdot\boldsymbol{n}^-)([\boldsymbol{v}]\cdot\boldsymbol{n}^-)\,ds\right), \tag{3.72a}$$

$$\begin{aligned}\partial_{\boldsymbol{u}} j_{\epsilon,\eta}(\boldsymbol{u},\boldsymbol{v})\cdot\boldsymbol{\psi} = &\frac{\mu}{\epsilon}\left(\int_{\Gamma_r}\mathcal{H}(\boldsymbol{u}\cdot\boldsymbol{n})(\boldsymbol{\psi}\cdot\boldsymbol{n})\mathcal{N}'(\boldsymbol{u}_t)\cdot\boldsymbol{v}_t ds + \int_{\Gamma_c}\mathcal{H}([\boldsymbol{u}]\cdot\boldsymbol{n})([\boldsymbol{\psi}]\cdot\boldsymbol{n}^-)\mathcal{N}'([\boldsymbol{u}]_t)\cdot[\boldsymbol{v}]_t ds\right) \\ &+ \frac{\mu}{\epsilon}\left(\int_{\Gamma_r}\max(\boldsymbol{u}\cdot\boldsymbol{n},0)(\mathcal{N}''(\boldsymbol{u}_t)\cdot\boldsymbol{\psi}_t)\cdot\boldsymbol{v}_t ds + \int_{\Gamma_c}\max([\boldsymbol{u}]\cdot\boldsymbol{n},0)(\mathcal{N}''([\boldsymbol{u}]_t)\cdot[\boldsymbol{\psi}]_t)\cdot[\boldsymbol{v}]_t ds\right)\end{aligned} \tag{3.72b}$$

Thus we get the adjoint equation (3.69). To determine the shape derivative (3.68), it suffices to differentiate

$$J(\Omega) = \mathcal{L}(\Omega,\boldsymbol{u}_0(\Omega),\boldsymbol{u}(\Omega),\tilde{\boldsymbol{v}}_0,\tilde{\boldsymbol{v}})$$

with respect to Ω, use the adjoint equation (3.69) and apply the Lemma 1. □

Remark 4. *For the state equations (3.61)-(3.62), one first solves for $\boldsymbol{u}_0$ and then for $\boldsymbol{u}$. For the adjoint equation (3.69) being retrograde, one first solves for $\boldsymbol{v}$ and then for $\boldsymbol{v}_0$.*

Theorem 11. *Under the assumptions **V.1-V.2** and for $\mu = 0$, the system of equations (3.69a)-(3.69b) admit a unique solution $(\boldsymbol{v}_0,\boldsymbol{v}) \in V\times V$*

Proof. Given the $\mathcal{C}^1$ smoothness assumptions on $m(\cdot)$ and $p(\cdot)$, the linear form on the right in (3.69b) is continuous. Moreover, the assumptions **V.1-V.2** and the positivity of $\mathcal{S}$, the bilinear form a_r is coercive. Then, following the arguments in the proof of Theorem 6, one can show the existence of a unique solution $\boldsymbol{v} \in V$. Using the continuity of the h, and repeating the same procedure for (3.69a), one can show the existence of a unique solution $\boldsymbol{v}_0 \in V$. □

3.7.5 Optimization of the idealized bolt location

In the previous computation of the shape derivative, the location of the bolt was assumed to be constant. In this subsection, we briefly present a gradient-descent approach to optimize the location of the bolt in the absence of pretension (i.e., $\boldsymbol{u}_0 = \mathbf{0}$) to minimize the compliance (3.67). The (parametric) gradient reads

$$\boldsymbol{x}_0^{i+1} = \boldsymbol{x}_0^i - \delta^i \frac{\partial J}{\partial \boldsymbol{x}_0}(\Omega), \tag{3.73}$$

where $\boldsymbol{x}_0^i$ is the center of ω_A at the iteration i, δ^i is the descent step and $\dfrac{\partial J}{\partial \boldsymbol{x}_0}(\Omega)$ is the partial derivative of the compliance (3.67) (computed at , $\boldsymbol{x}_0^i$), given by equation (7) in [150].

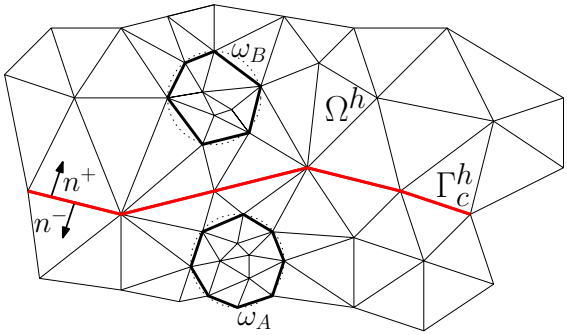


Figure 3.20: Bolt ends ω_A and ω_B, coincident boundary mesh of Γ^h, representing the contact surface Γ_c

3.7.6 Resolution of the simplified-bolt contact problem (3.61)-(3.61)

In order to ensure precise mechanical computation of the structure, the ends of the idealized bolts must be accurately meshed. The mesh of Ω^h is created that the bolt ends are captured using a body-fitted mesh using the tool `MMG`. A 2D example is shown in Fig.3.20. Evidently, the bolt ends ω_A and ω_B, are circles in 2D and spheres in 3D.

The space V is dicretized to $V^h = \mathbb{P}^1(\Omega^h)^d$. The space discretized state solution $(\boldsymbol{u}_{0,\epsilon}^h, \boldsymbol{u}_\epsilon^h) \in V^h \times V^h$ and adjoint solution $(\boldsymbol{v}_{0,\epsilon}^h, \boldsymbol{v}_\epsilon^h) \in V^h \times V^h$. In order to simplify notations in this section, we denote $(\boldsymbol{u}_{0,\epsilon}^h, \boldsymbol{u}_\epsilon^h)$ by $(\boldsymbol{u}_0^h, \boldsymbol{u}^h)$. All computations shall be performed on `FreeFEM`. The material and the bolt parameters considered are shown in Table 3.1. For all the numerical examples in this section, we shall treat frictionless test cases

E	ν	ϵ	ρ	$\kappa(\rho)$	ℓ
210GPa	0.3	$10^{-4}E$	$0.04m$	$4E$	8ρ

Table 3.1: Mild steel properties

only. We write the space discretized version of (3.61)-(3.62) for $\mu = 0$: find $(\boldsymbol{u}_0^h, \boldsymbol{u}^h) \in V^h \times V^h$ such that

$$a_r(\boldsymbol{u}_0^h, \boldsymbol{v}^h) + j'_\epsilon(\boldsymbol{u}_0^h)\boldsymbol{v}^h = h(\mathbf{0}, \boldsymbol{v}^h) \qquad \forall \boldsymbol{v}^h \in V^h, \tag{3.74a}$$

$$a_r(\boldsymbol{u}^h, \boldsymbol{v}^h) + j'_\epsilon(\boldsymbol{u}^h)\boldsymbol{v}^h = l(\boldsymbol{v}^h) + h(\boldsymbol{u}_0^h, \boldsymbol{v}^h) \qquad \forall \boldsymbol{v}^h \in V^h, \tag{3.74b}$$

$$\text{where} \quad j'_\epsilon(\boldsymbol{u}^h)\boldsymbol{v}^h = \frac{1}{\epsilon}\left(\int_{\Gamma_r} \max(\boldsymbol{u}\cdot\boldsymbol{n}, 0)\boldsymbol{v}\cdot\boldsymbol{n} ds + \int_{\Gamma_c} \max([\boldsymbol{u}^h]\cdot\boldsymbol{n}^-, 0)[\boldsymbol{v}^h]\cdot\boldsymbol{n}^- \, ds\right). \tag{3.74c}$$

We solve the variational Eqn.(3.74a) first followed by (3.74b) using a fixed point strategy, similar to (3.43). For the numerical implementation of the term j'_ϵ in the above, we follow the steps mentioned in Subsection 3.5.1. For the above frictionless case, the space discretized adjoint equation reads: find $(\boldsymbol{v}_0^h, \boldsymbol{v}^h) \in V^h \times V^h$ such that

$$a_r(\boldsymbol{\varphi}, \boldsymbol{v}_0^h) + (j''_\epsilon(\boldsymbol{u}_0^h)\cdot\boldsymbol{\varphi})\cdot\boldsymbol{v}_0^h = h(\boldsymbol{\varphi}, \boldsymbol{v}^h) - \int_\Omega \partial_{\boldsymbol{u}_0^h} m(\boldsymbol{u}_0^h, \boldsymbol{u}^h)\cdot\boldsymbol{\varphi}\, dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}_0^h} p(\boldsymbol{u}_0^h, \boldsymbol{u}^h)\cdot\boldsymbol{\varphi}\, ds \qquad \forall \boldsymbol{\varphi} \in V^h, \tag{3.75a}$$

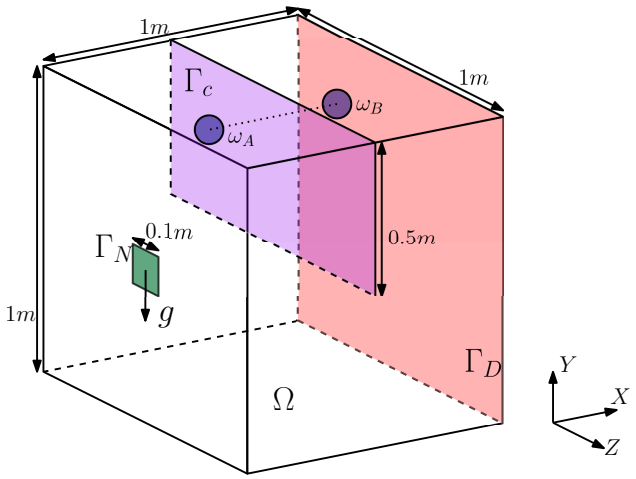


Figure 3.21: Boundary conditions

$$a_r(\boldsymbol{\psi}, \boldsymbol{v}^h) + (j_\epsilon''(\boldsymbol{u}^h) \cdot \boldsymbol{\psi}) \cdot \boldsymbol{v}^h = -\int_\Omega \partial_{\boldsymbol{u}^h} m(\boldsymbol{u}_0^h, \boldsymbol{u}^h) \cdot \boldsymbol{\psi}\, dx - \int_{\Gamma_N} \partial_{\boldsymbol{u}^h} p(\boldsymbol{u}_0^h, \boldsymbol{u}^h) \cdot \boldsymbol{\psi}\, ds \qquad \forall \boldsymbol{\varphi} \in V^h, \quad (3.75b)$$

The discretized shape derivative for this section reads

$$\begin{aligned} J'(\Omega^h)\boldsymbol{\theta} = \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} \big(& m(\boldsymbol{u}_0^h, \boldsymbol{u}^h) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}_0^h) : \boldsymbol{\varepsilon}(\boldsymbol{v}_0^h) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}^h) : \boldsymbol{\varepsilon}(\boldsymbol{v}^h) + \mathcal{S}(\boldsymbol{u}_0^h)\boldsymbol{v}_0^h + \mathcal{S}(\boldsymbol{u}^h)\boldsymbol{v}^h + M_\epsilon(\boldsymbol{u}_0^h)\boldsymbol{v}_0^h \\ & + M_\epsilon(\boldsymbol{u}^h)\boldsymbol{v}^h - h(\boldsymbol{0}, \boldsymbol{v}_0^h) - h(\boldsymbol{u}_0^h, \boldsymbol{v}^h) + \boldsymbol{f} \cdot \boldsymbol{v}^h \big)\, ds, \end{aligned} \quad (3.76)$$

where $(\boldsymbol{u}_0^h, \boldsymbol{u}^h)$ is the solution to (3.74) and $(\boldsymbol{v}_0^h, \boldsymbol{v}^h)$ is the solution to (3.75).

The resulting shape optimization algorithm is similar as the algorithm 2, given below.

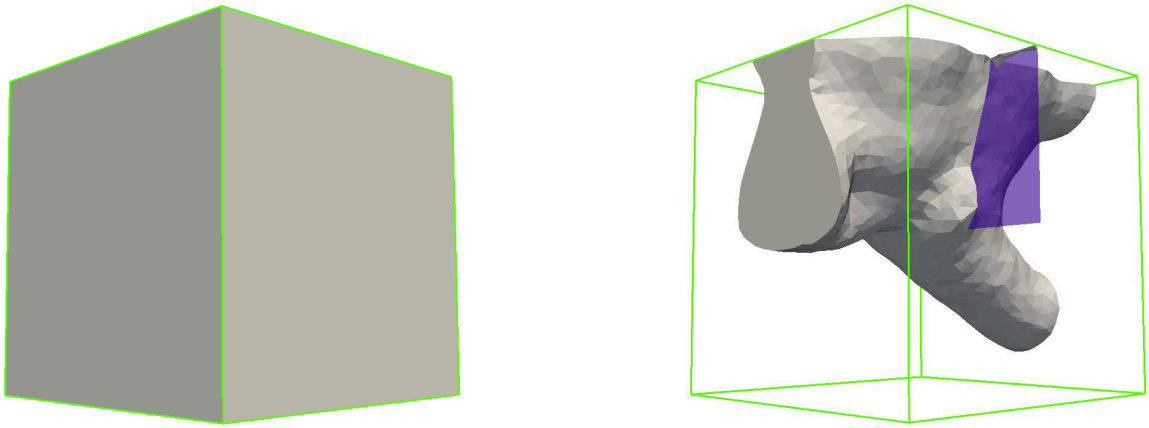
Algorithm 3 Initialize with a shape Ω_0^h and repeat over $i = 1, \cdots, I_{\max}$

1. Solve state equation (3.74) for $(\boldsymbol{u}_0, \boldsymbol{u})$ on the mesh of Ω_i^h
2. Solve adjoint equation (3.75) for $(\boldsymbol{v}_0, \boldsymbol{v})$ on the mesh of Ω_i^h
3. Compute the shape derivative (3.76) and regularize it using (2.23) to find $dj_\alpha(\Omega_i^h)$
4. Update the Lagrange multiplier using (2.29), using the time step (2.28)
5. Solve the transport equation (2.17) with the velocity given by (2.30) for the pseudo-time step τ given by (2.28) to obtain the new level-set function $\tilde{\phi}_{i+1}$
6. Re-initialize $\tilde{\phi}_{i+1}$ to the signed distance function ϕ_{i+1} (defining a new shape Ω_{i+1}^h)
7. Compute the volume $V_{\Omega_{i+1}^h}$. If it is close to the volume target, apply the projection algorithm (2.31) to satisfy exactly the volume constraint.

3.7.7 Numerical Results: Pretension

For all the test cases, we again minimize compliance objective function (3.67) using the algorithm 3. The material and bolt properties chosen here are given in Table 3.1. The body force $\boldsymbol{f}$ is assumed to be zero. As mentioned in Section 3.5, for all the test cases here, the coefficient of friction $\mu = 0$.

We consider a cube of unit side length with a vertical crack plane at the center as shown in Fig.3.21. The cube is subjected to same boundary conditions as the cube in Fig.3.15a. The only difference is that now, there is bolted connection across the fissure Γ_c. The bolt heads are represented by ω_A and ω_B. The bolt lies at the middle (along Z) and is $0.2m$ downwards from the topmost surface. The bolts are given a pre tension of 25MN. Here, we study the impact of



(a) Initial shape

(b) Final shape, $\boldsymbol{g} = (0, -50, 0)$ N/m^2

(c) Final shape, $\boldsymbol{g} = (0, -100, 0)$N/$m^2$

(d) Final shape, $\boldsymbol{g} = (0, -200, 0)$ N/m^2

Figure 3.22: Initial and final shapes for compliance (3.50) minimization, target volume $V_f = 0.25m^3$

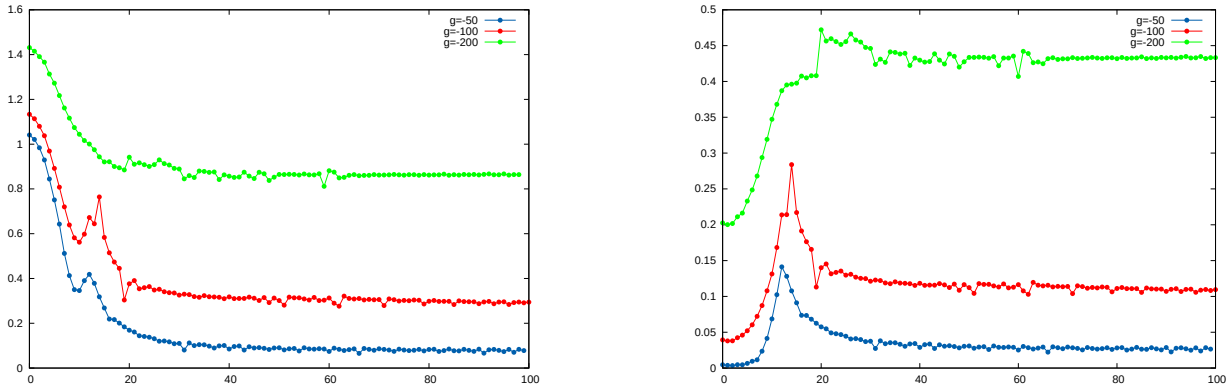


Figure 3.23: Lagrangian (3.51) (left) and compliance (3.50)(right) for the 3 cases in (3.26)

1. Increasing the magnitude of external force $\boldsymbol{g}$
2. Increasing the number of bolts

on the optimal shape. For the first case, We plot the initial shape and the final shape for several forces in Fig.3.26. The initial shape has periodic holes inside the cube. We observe that as an when the magnitude

of the force rises, the structure becomes more and more stiff at the lower half. The pretension plays lesser role in stiffening and decreasing the compliance of the structure. In addition, we compare the convergence history for the three cases in Fig.3.23. Here we observe that for the three cases, the compliance increases initially in favor of decreasing the volume. And once the target volume is attained, the compliance decreases consistently and then remains constant towards convergence. Now for the second case, we consider the same cube as before but with two bolts as shown in the Fig.3.24. The external force is taken to be much larger than before, $\boldsymbol{g} = (0, -400, 0)$ N$/m^2$ since there are two bolts. The initial shape in Fig.3.25a (same as (3.22a)) converges to the shape in Fig.3.25b. We observe in Fig.3.25b that the optimal shape has two bars on the top, each connecting the bolt to Γ_D and Γ_N, which is completely different from the shapes obtained for a single bolt in (3.26).

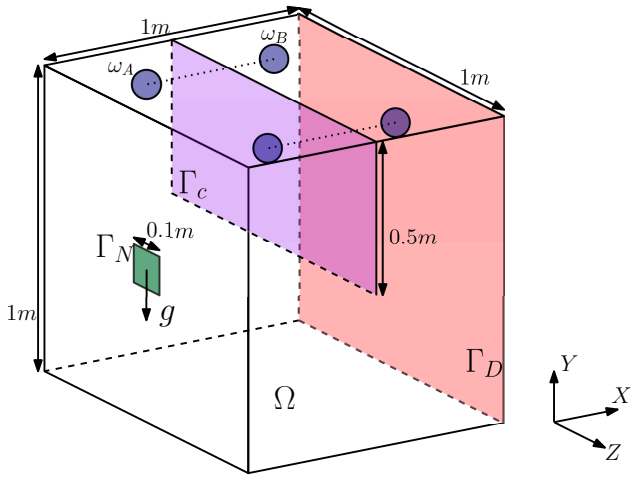
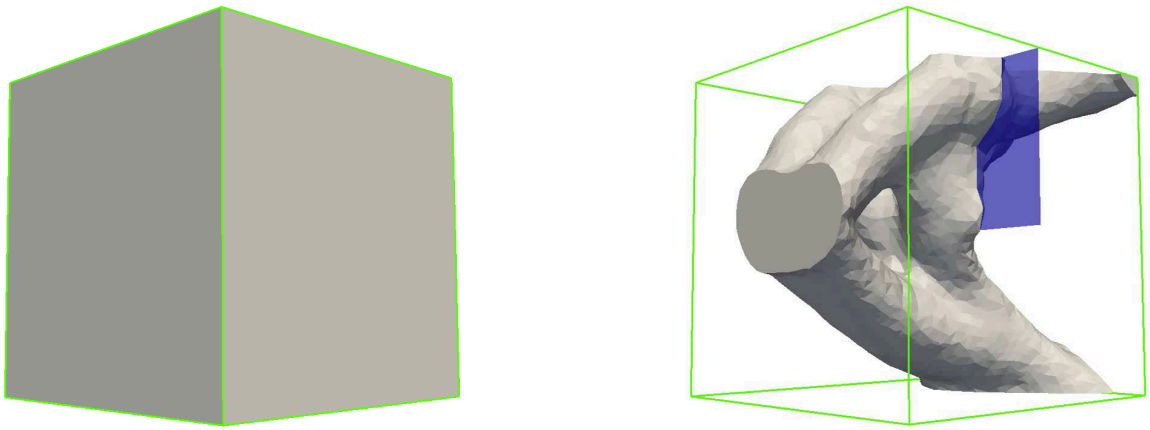


Figure 3.24: Boundary conditions

3.7.8 Numerical Results: Bolt position and topology optimization without pretension

In this subsection, the position of the bolt along with the topology of the structure is optimized. We assume absence of pretension ($\boldsymbol{\phi} = 0$) and of friction ($\mu = 0$). The shape optimization algorithm 3 is applied along with an additional step (3.73) at the end of each iteration. The compliance (3.67) is minimized. For all examples, a target volume $V_f = 0.25m^3$ is imposed. Here as well, we consider a cube of unit side length with

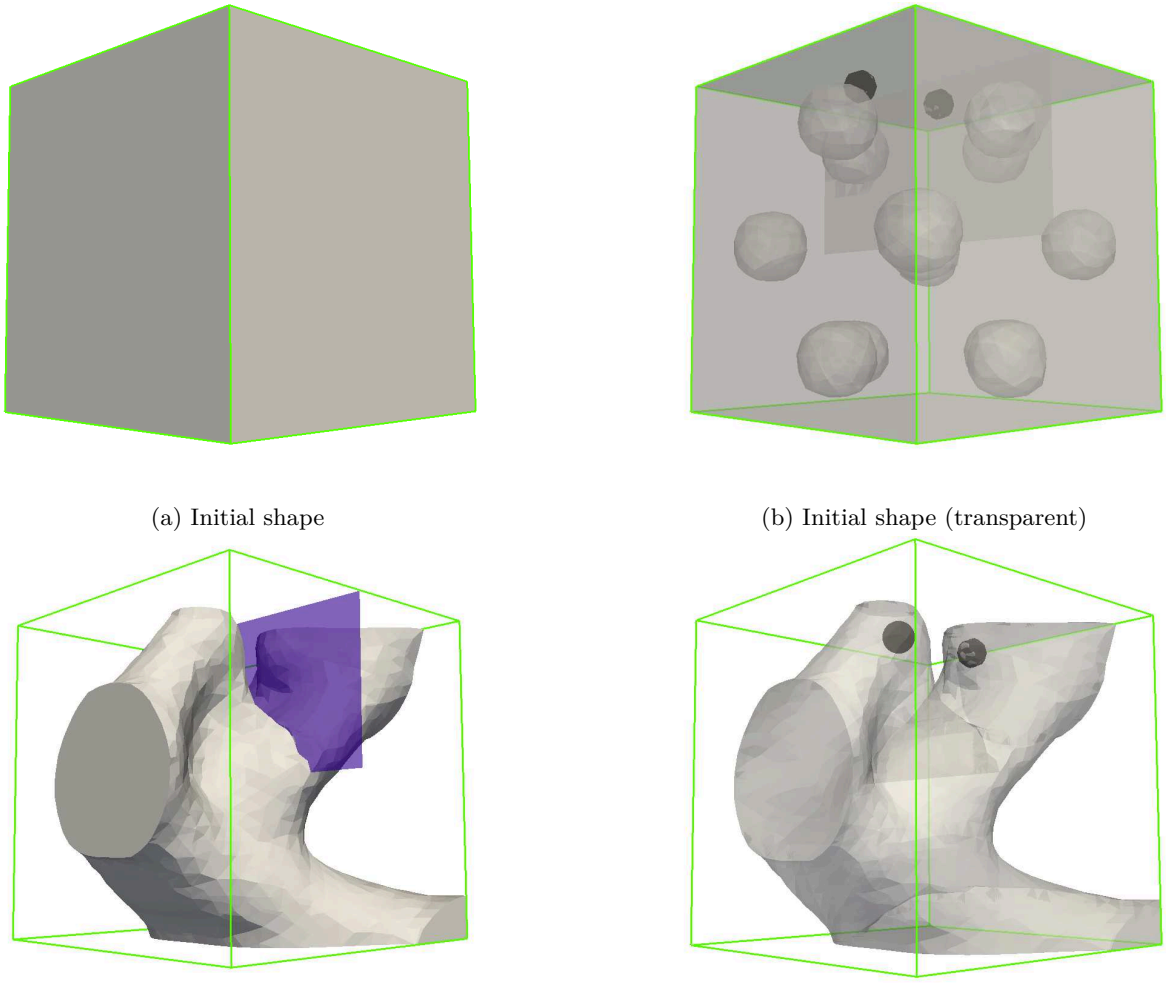


(a) Initial shape

(b) Final shape, $\boldsymbol{g} = (0, -400, 0)$ N$/m^2$

Figure 3.25: Initial and final shapes for compliance (3.50) minimization

a vertical crack plane at the center as shown in Fig.3.21. The material and bolt properties chosen here are given in Table 3.1 and no body force $\boldsymbol{f}$ is applied.



(a) Initial shape

(b) Initial shape (transparent)

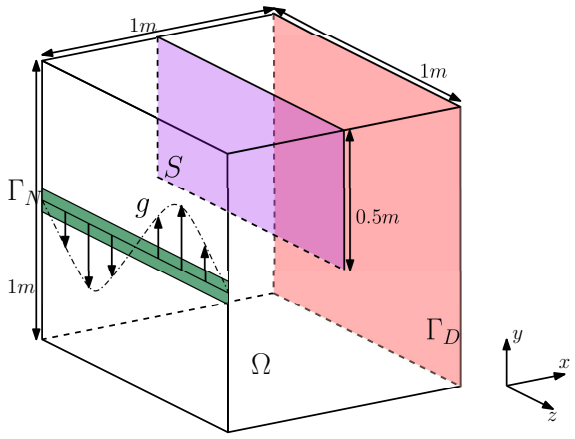
(c) Final shape

(d) Final shape (transparent)

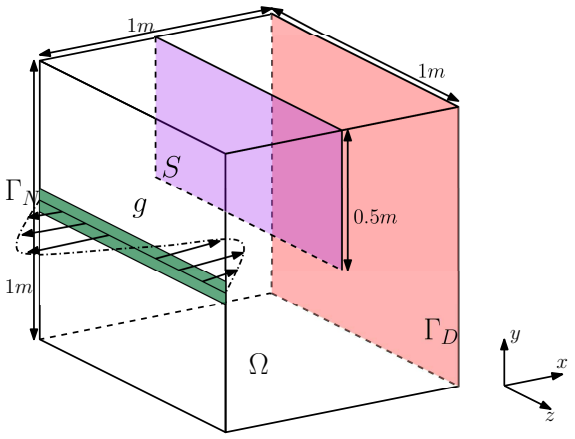
Figure 3.26: Initial and final shapes, for compliance (3.67) minimization with bolt location, $\boldsymbol{g} = (0, -100, 0)\ N/m^2$

We first consider a (vertically) downward force $\boldsymbol{g} = (0, -100, 0)\ N/m^2$. The shape with periodic holes is initialized as shown in Fig.3.26b. This initial shape contains an idealized bolt on the central plane in the upper half as seen in Fig.3.26b. On applying the bolt position-topology coupled optimization, we end up with the shape in Fig.3.26c and the bolt position, as seen in Fig.3.26d. As one can see from the optimized shape, since the force is applied vertically downward, the optimization algorithm tries to place the bolts on the top of the cube, as much as possible. At the same time, since there is absence of contact (due to the force direction), there is almost no mass near the contact plane (in purple).

We then apply two sinusoidal forces. As shown in the Fig.3.27 two forces $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ are considered (specified in the caption). We apply such a force to see if the optimizing the location bolt results in an asymmetric shape. Indeed it does. The shape is initialized as shown in Fig.3.26a and the initial bolt location can be seen in Fig.3.26b. The final shape obtained for the two forces $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ are plotted in Fig.3.28. For the force $\boldsymbol{g}_1$, the bolt ends up being outside the optimal shape Ω. This implies that the bolt is not needed to minimize the compliance; instead, the tube like structure suffices. For the force $\boldsymbol{g}_2$ the bolt does end up being used. It's location is inside the small protrusion coming out as seen in Fig.3.28. The algorithm deals with "two" zones with smart strategies: the zone the external force is compressive, it adds mass near the contact surface so as to support the structure and rigidify it and in the zone, where the external force is tensile, it removes mass and keeps the bolt, so as to ensure that the deformation is minimized and no mass

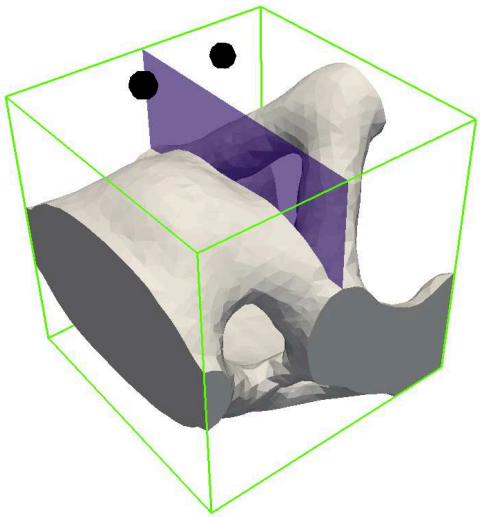


(a) $\boldsymbol{g}_1 = \left[0,\ \sin\left(\frac{\pi z}{2}\right),\ 0\right]$

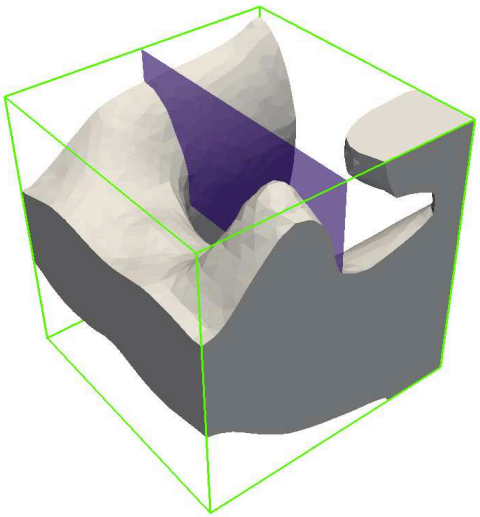


(b) $\boldsymbol{g}_2 = \left[\sin\left(\frac{\pi z}{2}\right),\ 0,\ 0\right]$

Figure 3.27: Domain Ω fixed on the leftmost edge, forced on the right, supported at the bottom.



(a) $\boldsymbol{g}_1 = \left[0,\ \sin\left(\frac{\pi z}{2}\right),\ 0\right]$



(b) $\boldsymbol{g}_2 = \left[\sin\left(\frac{\pi z}{2}\right),\ 0,\ 0\right]$

Figure 3.28: Domain Ω fixed on the leftmost edge, forced on the right, supported at the bottom.

is wasted.

Chapter 4

Plasticity

Contents

In structural design involving steel or soil, a plasticity model is often used to determine the plastic strain, or the permanent deformation. The structure on being subjected to an external force, exhibits plastic strain when it undergoes a stress, that exceeds a value known as the yield strength [86]. As the time-dependent force evolves, if the yield strength remains constant everywhere in the structure, the resulting phenomenon is called perfect plasticity, otherwise it is called plasticity with hardening. Using a hardening law, one can determine the shift in the yield strength and measure how ductile the material is.

Plasticity modeling has been developed significantly since its inception in the 1960's. At the heart of the model lies the Hill's principle and its equivalent Drucker Illyushin principle [122]. The plasticity model is often simplified by assuming that the evolution of the force is slow. This assumption is valid for mechanical processes that are slow, for instance metal-drawing; and this assumption is absolutely false for instantaneous processes like vehicle crash-test or cold-forging. This assumption results in a quasi-static plasticity model, which has been largely studied theoretically. The model, when written in its variational formulation, is an inequality which can be expressed either in a dual form or in a primal form. Under certain hypothesis on the material properties, both the forms be shown to be a well-posed problem [87]. While the primal formulation illuminates the theoretical properties of the solution to the plasticity problem, it is not easily amenable to numerical resolution. Therefore, one instead resorts to the radial-return algorithm [165, 162, 185] which

discretizes the governing equations of hardening-based plasticity using an implicit Euler scheme. The radial return algorithm is used in virtually every commercial software.

Most of the shape and topology optimization algorithms are developed for linearized elasticity problems while less focus is given to nonlinear ones. Nonlinearities can arise due to material properties (plasticity, damage), contact boundary conditions, hyperelasticity [109], large displacement (large strain [181, 38], finite strain [182]) and structural buckling [117]. Topology optimization using density approaches or SIMP (Solid Isotropic material with Penalization) was applied to elasto-plastic problems [131, 182, 103, 33, 116], visco-elastic problems [97] and visco-elasto-plastic problems [140]. A common feature in all the previous works is the determination of a design gradient by differentiating the space and time-discretized schemes of the plasticity models, which are approximated using a fictitious material density. The material properties like the Young's modulus and the hardening coefficients are modified using this material density raised to a certain exponent. This exponent is different for every mechanical property and ought to be chosen in an ad-hoc manner, ensuring numerical stability. The optimized shape obtained have intermediate densities undergoing a plastic flow, which might actually be artificial. In the level-set framework with body-fitted remeshing, since the material properties are not approximated using material densities, such artificial plastic zones are avoided.

The level-set method for topology optimization was applied to a simplified version of perfect plasticity in [130]. There, the first time-step of time-discretized perfect plasticity, also known as the Hencky's model, was approximated using Perzyna penalization and the resulting approximation was shown to be well-posed. The model did not take into account hardening laws, the time dependence or the irreversibility of the plasticity problem. As soon as one incorporates the irreversibility of the plastic flow and hardening laws, one ends up with a variational inequation with a complex theoretical and numerical treatment.

Another approach to topology optimization involves a scalar phase-field function, featuring a small parameter δ. This approach is similar to density approach, where the shape is represented using the phase field function. The plasticity problem is approximated using the phase field function and one can show that as the parameter δ tends to zero, the solution to the approximate problem converges to the solution to the exact problem. This analysis was performed for plasticity with linear kinematic hardening only (no isotropic hardening) in [15].

In this chapter, we apply the level-set method to quasi-static plasticity with linear kinematic and linear isotropic hardening. Unlike most of the other previous works, the quasi-static plasticity is considered in its primal formulation and in the next chapter, the same is considered in its incremental formulation. For the two cases (primal and incremental formulation), the shape derivative is determined for the continuous problem. The primal form being non-smooth is not differentiable. Nevertheless, we construct an approximate problem that is differentiable using a penalization and regularization technique. We show that the approximate problem is well-posed and the corresponding solution converges to the exact solution. Then, we compute the shape derivative for minimizing an objective function with the approximate problem as a constraint. As usual, this shape derivative involves an adjoint problem. Thanks to the proposed penalization and regularization, the time discretized version of this adjoint problem is proved to be well-posed. It is well known that the original primal problem is not differentiable in the usual sense but admits only a so-called conical derivative (see [133, 171, 130]). Similarly, there is no rigorous notion of adjoint for this primal problem. Nevertheless, to make a comparison with our regularized adjoint, we present a formal approach, relying on strong assumptions (not always realistic), which allows us to give a shape derivative and an adjoint problem for the primal formulation. In some sense, this "formal" or "naive" shape derivative and adjoint equation of the original problem should be the limits of our regularized shape derivative and adjoint equations when the penalization and regularization coefficients go to zero. However, we do not perform such a limit analysis, which of course would require strong assumptions, and rather we content ourselves with pursuing a pedagogical goal in Subsection 4.3.2.

The efficiency of the shape derivative (obtained with our penalization and regularization process) is assessed by numerically optimizing some geometries in 2D and 3D. The plasticity problem is numerically solved using the radial return algorithm. One salient feature of the approach used here is that the geometry is captured at each iteration of the optimization process by a body-fitted mesh (using `MMG` [61]), which allows for topology changes (see Chapter 2, Subsection 2.2.2 for details). Unlike XFEM [70], where the mesh elements can become heavily distorted, capturing of the geometry using `MMG` ensures a much better mesh and thus an accurate calculation of the shape derivative. A CutFEM [41] type approach may also be considered; however, constructing a ghost penalty term for the primal plasticity formulation is theoretically involved. We present numerical case studies of a cantilever and a wedge in 2D and in 3D. The two geometries are loaded with a uniaxial force that increases monotonically in time. We compare the shapes optimized

for plasticity with the ones optimized for linear elasticity. For the 2D wedge, we also consider a force that changes its direction with time. While optimizing the shape for such a force, we show that the dependence of the optimized topology on the forcing history is significant.

From an industrial perspective, mechanical parts are never conceived to undergo plastic deformation when put to use. It is assumed that maximum stress, that a mechanical part shall undergo would always remain below the yieldstrength. The ratio between the yieldstrength and the maximum stress that the mechanical part is designed to undergo, also called as the factor of safety is thus more than one. Hence, plasticity modeling is not quite useful from the perspective of conceiving shapes. Optimizing shapes to minimize plastic deformation then make no sense for typical mechanical parts. Nevertheless, one can permit slight damage or plastic deformations to design "fail-safe" structures [98, 190]. There are a few applications of "fail-safe" structures in the aerospace industry or crash tests for automobile industry [106]. Another possibility to circumvent optimization for plasticity is to design structures for worst-case scenario. Here, one optimizes topology, assuming uncertainties in Lamé parameters [59], in the external forcing [5, 56] or both [127].

Instead of using plasticity for conceiving mechanical parts, one can use it optimize fabrication purposes. Typically, metal components for automobile and aerospace industry are manufactured by the process of cold-working. In cold-working, the metal is kept at a room temperature and subjected to a variety of external forces of large magnitudes. These forces are applied via drawing, punching, cold-forging, stamping, sheet metal forming, etc. For an automobile, around 45% of the weight comes from the chassis and the suspension [177]. This weight is assumed to be "shape optimizable" (for linear elasticity). Most of the parts making up the 45% weight are manufactured using sheet metal forming [177]. Shape optimization to maximize plasticity in sheet metal forming seems to be a relevant industrial subject. While the cold-working involves several nonlinear phenomena like contact, plasticity, and large deformation, our work here is limited to quasi-static plasticity only.

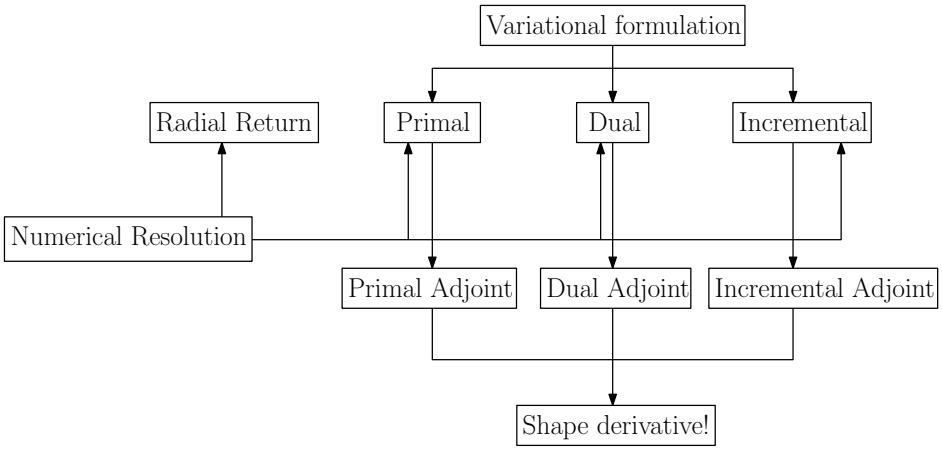


Figure 4.1: Various recipes for shape derivative

We now discuss the layout of this chapter as well as the next chapter. The governing equations for plasticity are presented in Section 4.1. These governing equations are condensed in two forms

1. Primal form (see Subsection 4.1.2)
2. Dual form (see Subsection 4.1.3),

both of which, are derived rigorously in Section 4.1. Using the governing equations, one can determine a variational formulation in three possible ways (see Fig.4.1),

1. Primal formulation:
   This formulation is derived from the Primal form, treated analytically in Section 4.2 and numerically in Section 4.4. The shape derivative shall be determined for this formulation in Section 4.3.
2. Dual formulation
   This formulation is derived from the Dual form. This formulation shall not be treated analytically in or numerically in this thesis. The reason being the fact that numerical treatment involves stress formulation and requires very exotic polynomial spaces [155].

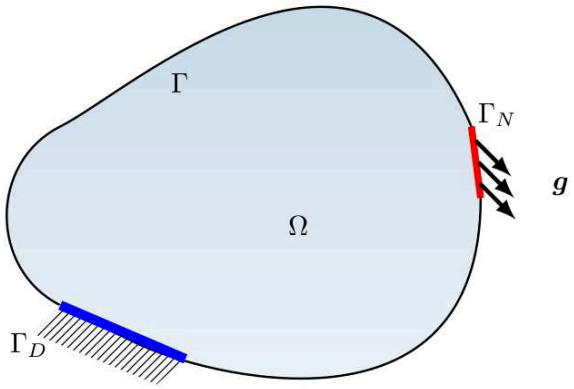


Figure 4.2: Boundary conditions on the structure Ω

3. Incremental approach
   This approach involves constructing a Lagrangian for the time discretized plasticity problem [154]. This formulation shall be treated in the Chapter 5, analytically in Section 5.1. The shape derivative shall be determined in Section 5.2.

As shown in Fig.4.1, the variational formulation for plasticity can be constructed via three possible ways. The numerical resolution can be performed on the three formulations as well as by using a radial return algorithm. The three formulations result in three adjoint problems, each of which finally gives a shape derivative. We hypothesize that the shape derivative obtained via the three approaches should be the same, but leave the comparison study for future works. In this thesis, the adjoint formulation via primal formulation and incremental approach is derived in Section 4.3 and Section 5.2, respectively.

For the primal formulation and the incremental approach, we consider rate-independent associative plasticity. Associative flow rule means that the rate of evolution of the plastic strain variables is proportional to normal of the yield surface. If they are not proportional, rather given experimentally, the flow rule is called non-associative flow rule. For this non-associative flow, the primal framework of Section 4.2 and the shape derivative in Section 4.3 are no longer valid. However, the incremental approach can deal with non-associative flow rule and a general yield criterion, hence they have been studied in Section 5.1 and the shape derivative in Section 5.2.

The standard textbooks for plasticity are [85], [162]. The notations used by the mathematical community and the solid mechanics community are totally different and here we try to establish a fair balance between the two.

Some content of this chapter has been documented in an article submitted for publication:
J.Desai, G.Allaire, F.Jouve, C.Mang. *Topology optimization in quasi-static plasticity with hardening using a levet-set method*
This chapter is essentially an elaborated version (mathematically and numerically) of the above article. More precisely, here we give the mathematical details of derivation of the plasticity model, numerical details of the implementation of the approximate plasticity model and its comparison to the radial return algorithm.

4.1 Governing laws

We first present the laws governing plasticity with linear kinematic and isotropic hardening. We then briefly talk about the duality analysis in Subsection 4.1.1, which lays the foundation for the primal form of plasticity in Subsection 4.1.2 and for the dual form of plasticity in Subsection 4.1.3.

As we did in all other chapters, we follow the convention of representing all scalar mathematical entities by lowercase italic text. All vector spaces are denoted by upper case letters, and all vectors, tensors are denoted by lowercase bold text. We let

- $d = 2, 3$ represent the dimension of the problem
- $\Omega \subset \mathbb{R}^d$ be an open bounded set representing a structure

- $[0,T]$ be a bounded time interval of interest
- $\mathcal{M}_s^d$ represent the set of symmetric $d \times d$ matrices
- $\mathbb{I}$ represent the fourth-order identity tensor of dimension d
- $\partial\Omega = \Gamma_D \cup \Gamma_N \cup \Gamma$, union of disjoint boundaries as in Fig.4.2
- $\boldsymbol{n}$ denote the outward normal to $\partial\Omega$
- Γ_D denote the Dirichlet boundary
- $\boldsymbol{f} : \Omega \times [0,T] \to \mathbb{R}^d$ be the time-dependent body force applied in Ω
- $\boldsymbol{g} : \Gamma_N \times [0,T] \to \mathbb{R}^d$ be the time-dependent surface force applied on the Neumann boundary Γ_N
- $\boldsymbol{u} : \Omega \times [0,T] \to \mathbb{R}^d$ denote the displacement field
- $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\left(\nabla \boldsymbol{u} + \nabla^T \boldsymbol{u}\right)$ denote the total strain tensor
- $\boldsymbol{\sigma} : \Omega \times [0,T] \to \mathcal{M}_s^d$ denote the stress tensor
- $\mathbb{C}$ be the fourth order Hooke's tensor
- $\boldsymbol{\varepsilon}_e : \Omega \times [0,T] \to \mathcal{M}_s^d$ be the elastic strain tensor
- $\boldsymbol{\varepsilon}_p : \Omega \times [0,T] \to \mathcal{M}_s^d$ be the plastic strain tensor
- $\gamma : \Omega \times [0,T] \to \mathbb{R}$ be a scalar function representing the accumulated plastic strain
- $\boldsymbol{q} : \Omega \times [0,T] \to \mathcal{M}_s^d$ be the kinematic hardening tensor
- $\boldsymbol{r} : \Omega \times [0,T] \to \mathcal{M}_s^d$ be internal variable corresponding to kinematic hardening
- $g : \Omega \times [0,T] \to \mathbb{R}$ be the scalar force responsible for isotropic hardening.
- $\sigma_Y \in \mathbb{R}^+$ be the yield strength

Plasticity is a quasi-static phenomenon as we now describe (see [86] for more details). The structure when subjected to a time-dependent body force $\boldsymbol{f}$ and a time-dependent surface force $\boldsymbol{g}$, satisfies the momentum balance equation:

$$\begin{aligned} \operatorname{div}(\boldsymbol{\sigma}) + \boldsymbol{f} &= \boldsymbol{0} && \text{in } \Omega \times (0,T], && (4.1a)\\ \boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N \times (0,T], && (4.1b)\\ \boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma \times (0,T], && (4.1c)\\ \boldsymbol{u} &= \boldsymbol{0} && \text{on } \Gamma_D \times (0,T]. && (4.1d) \end{aligned}$$

Plastic deformations occur when the magnitude of the stress $\boldsymbol{\sigma}$ exceeds the yield strength $\sigma_Y \in \mathbb{R}^+$, a material parameter determined experimentally. The plastic flow is characterized by two variables, the plastic strain tensor $\boldsymbol{\varepsilon}_p$ and the scalar function γ representing the accumulated plastic strain. The total strain tensor of the structure is defined as

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T\right).$$

We need a relation between the total strain tensor $\boldsymbol{\varepsilon}$ and the plastic strain tensor $\boldsymbol{\varepsilon}_p$. Experimentally, it is known that an isotropic material subjected to uni-axial loading deforms according to a stress-strain relation as shown in Fig.4.3. The total strain $\boldsymbol{\varepsilon}$ starts increasing from 0 and remains linear with respect to the stress until the yieldstrength σ_Y (point A), after which it starts accumulating plastic strain. At any point along the curve A-B, the total strain $\boldsymbol{\varepsilon}$ can be decomposed as

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_p.$$

In the case of bi-axial and tri-axial loading, the elastic limit measured by σ_Y becomes a convex closed curve and a convex closed surface respectively.

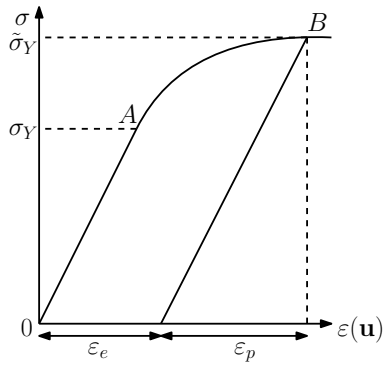
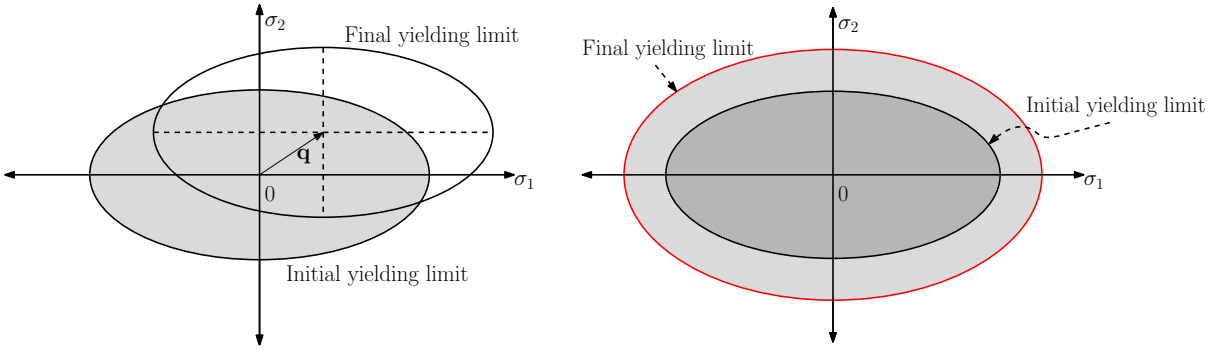


Figure 4.3: Stress-strain relation for uni-axial tensile loading



(a) Kinematic hardening

(b) Isotropic hardening

Figure 4.4: σ_1 and σ_2 denote the principle stresses

Hardening occurs when the plastic flow is followed by a change in yield strength σ_Y and thus a shift in the elastic limit. For instance, in the uni-axial loading case the yieldstrength σ_Y displaces to $\tilde{\sigma}_Y$ (point B) as shown in Fig.4.3. In bi-axial or tri-axial loading, this shift of elastic limit can be such that either the center of the curve/surface shifts, or the diameter of the curve/surface increases. The former being kinematic hardening (shown in Fig.4.4a) and the latter being isotropic hardening (shown in Fig.4.4b) are considered in this thesis. The concept of kinematic hardening was introduced by Prager [149] and developed further by Ziegler [191].

The hardening is modeled by a stress-like hardening tensor $\boldsymbol{q} : \Omega \times [0,T] \to \mathcal{M}^d_s$, a scalar force $g : \Omega \times [0,T] \to \mathbb{R}$, and the corresponding internal variable, $\boldsymbol{r} : \Omega \times [0,T] \to \mathcal{M}^d_s$, $\gamma : \Omega \times [0,T] \to \mathbb{R}$, respectively. As shown in Fig.4.4a, the stress-like tensor $\boldsymbol{q}$ is in fact the shift in the center of the elastic domain.

To define the structure's elastic limit, we consider the von Mises yield criterion [162]

$$f(\boldsymbol{\sigma}, \boldsymbol{q}, g) = |\boldsymbol{\sigma}^D - \boldsymbol{q}^D| + \sqrt{\frac{2}{3}}\,(g - \sigma_Y) \leq 0, \tag{4.2}$$

where the superscript D denotes the deviatoric part of a tensor and $\sigma_Y \in \mathbb{R}^+$ is the yield strength. This criterion defines the elastic domain

$$\mathbb{E} = \{(\boldsymbol{\sigma}, \boldsymbol{q}, g) : f(\boldsymbol{\sigma}, \boldsymbol{q}, g) \leq 0\},$$

which, by definition, is convex. Other criteria such as Tresca or Mohr-Coulomb [166] are not considered in this thesis.

The structure is made of an isotropic material, with Hooke's tensor given by

$$\mathbb{C} = \lambda \mathbf{1} \otimes \mathbf{1} + 2\mu \mathbb{I},$$

where λ, μ are Lamé constants. We place ourselves in the framework of associated plasticity, namely, the plastic flow rate is proportional to the normal of the elastic domain. We first state the second law of thermodynamics

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \dot{\psi} \geq 0, \tag{4.3}$$

where the overdot denotes differentiation with respect to time and ψ is the Helmholtz free energy, given by the sum

$$\psi = \psi(\boldsymbol{\varepsilon}_e, \boldsymbol{r}, \gamma) = \hat{\psi}_e(\boldsymbol{\varepsilon}_e) + \hat{\psi}_p(\boldsymbol{r}, \gamma),$$

where the elastic and plastic energies are respectively defined as

$$\hat{\psi}_e(\boldsymbol{\varepsilon}_e) = \frac{1}{2}\mathbb{C}\boldsymbol{\varepsilon}_e : \boldsymbol{\varepsilon}_e \quad \text{and} \quad \hat{\psi}_p(\boldsymbol{r}, \gamma) = \frac{1}{2}\mathbb{H}\boldsymbol{r} : \boldsymbol{r} + \frac{1}{2}E_{iso}\gamma^2,$$

where $\mathbb{H}$ is the hardening tensor and $E_{iso} \geq 0$ is a material parameter. Due to thermodynamic reasons, the energy $\hat{\psi}_p$ is assumed to be independent of the elastic strain $\boldsymbol{\varepsilon}_e$. On the other hand the stress is assumed to be $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\varepsilon}_e)$. Using these definitions, the second law (4.3) is re-written as

$$\left(\boldsymbol{\sigma} - \frac{\partial \hat{\psi}_e}{\partial \boldsymbol{\varepsilon}_e}\right) : \dot{\boldsymbol{\varepsilon}} + \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p - \frac{\partial \hat{\psi}_p}{\partial \boldsymbol{\varepsilon}_p} : \dot{\boldsymbol{\varepsilon}}_p - \frac{\partial \hat{\psi}_p}{\partial \boldsymbol{r}} : \dot{\boldsymbol{r}} - \frac{\partial \hat{\psi}_p}{\partial \gamma}\dot{\gamma} \geq 0. \tag{4.4}$$

Using Coleman-Noll arguments [50], we deduce

$$\boldsymbol{\sigma} = \frac{\partial \hat{\psi}_e}{\partial \boldsymbol{\varepsilon}_e} = \mathbb{C}\boldsymbol{\varepsilon}_e = \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p). \tag{4.5}$$

Now, the power dissipation function $\mathcal{D}$ is introduced as the difference between the external power and the rate of change of Helmholtz free energy

$$\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p + \boldsymbol{q} : \dot{\boldsymbol{r}} + g\dot{\gamma} \tag{4.6}$$

where

$$\boldsymbol{q} = -\frac{\partial \hat{\psi}_p}{\partial \boldsymbol{r}} = -\mathbb{H}\boldsymbol{r} \quad \text{and} \quad g = -\frac{\partial \hat{\psi}_p}{\partial \gamma} = -E_{iso}\gamma. \tag{4.7}$$

As stated in the introduction, we deal with linear kinematic and linear isotropic hardening. This comes from the above choice of plastic energy $\hat{\psi}_p$. In the above, the energy stored due plastic strain γ is quadratic, but other nonlinear choices may also be considered [139, 159].

Substituting $\mathcal{D}$ in (4.4), we get

$$\mathcal{D} \geq 0.$$

This is exactly Hill's principle (or second law of thermodynamics) and is equivalent to the Drucker-Illyushin's principle of maximum work which states that for any stress state $(\boldsymbol{\sigma}, \boldsymbol{q}, g)$ in $\mathbb{E}$, the plastic flow variables $(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{r}}, \dot{\gamma})$ must satisfy

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}_p + \boldsymbol{q} : \dot{\boldsymbol{r}} + g\dot{\gamma} \geq \boldsymbol{\tau} : \dot{\boldsymbol{\varepsilon}}_p + \boldsymbol{p} : \dot{\boldsymbol{r}} + k\dot{\gamma} \qquad \forall (\boldsymbol{\tau}, \boldsymbol{p}, k) \in \mathbb{E}. \tag{4.8}$$

Since the set $\mathbb{E}$ is invariant by addition of a multiple of the identity tensor to $\boldsymbol{\sigma}$ and $\boldsymbol{q}$, (4.8) implies that necessarily the trace of $\dot{\boldsymbol{\varepsilon}}_p + \dot{\boldsymbol{r}}$ vanishes. Furthermore, (4.8) yields the following characterization of $\mathcal{D}$

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{r}}, \dot{\gamma}) = \sup_{(\boldsymbol{\tau}, \boldsymbol{p}, k) \in \mathbb{E}} \left(\boldsymbol{\tau} : \dot{\boldsymbol{\varepsilon}}_p + \boldsymbol{p} : \dot{\boldsymbol{r}} + k\dot{\gamma}\right), \tag{4.9}$$

where the supremum is attained at $(\boldsymbol{\sigma}, \boldsymbol{q}, g)$. Evidently, there are two cases, the supremum $(\boldsymbol{\sigma}, \boldsymbol{q}, g)$ can lie inside the convex set $\mathbb{E}$ or on the boundary of the set $\mathbb{E}$. If $(\boldsymbol{\sigma}, \boldsymbol{q}, g)$ lies inside the convex set $\mathbb{E}$ and satisfies (4.9), then necessarily $\dot{\boldsymbol{\varepsilon}}_p = \boldsymbol{0}$, $\dot{\boldsymbol{r}} = \boldsymbol{0}$, $\dot{\gamma} = 0$. If $(\boldsymbol{\sigma}, \boldsymbol{q}, g)$ lies on the boundary $\mathbb{E}$ and satisfies (4.9), then (4.10) must hold (this can be shown formally by contradiction). In all, the maximization (4.9) ensures that the normality law is satisfied [86]

$$\begin{aligned} f(\boldsymbol{\sigma}, \boldsymbol{q}, g) < 0 &\implies \dot{\boldsymbol{\varepsilon}}_p = \boldsymbol{0},\ \dot{\boldsymbol{r}} = \boldsymbol{0},\ \dot{\gamma} = 0 \\ f(\boldsymbol{\sigma}, \boldsymbol{q}, g) = 0 &\implies \dot{\boldsymbol{\varepsilon}}_p = \zeta\partial_{\boldsymbol{\sigma}} f,\ \dot{\boldsymbol{r}} = \zeta\partial_{\boldsymbol{q}} f,\ \dot{\gamma} = \zeta\partial_g f, \end{aligned} \tag{4.10}$$

where ζ is a Lagrange multiplier satisfying

$$\zeta \geq 0 \quad \text{and} \quad \zeta f(\boldsymbol{\sigma}, \boldsymbol{q}, g) = 0.$$

The derivatives of f (normal to the elastic domain) are given by

$$\partial_{\boldsymbol{\sigma}} f = \frac{\boldsymbol{\sigma}^D - \boldsymbol{q}^D}{|\boldsymbol{\sigma}^D - \boldsymbol{q}^D|}, \quad \partial_{\boldsymbol{q}} f = -\frac{\boldsymbol{\sigma}^D - \boldsymbol{q}^D}{|\boldsymbol{\sigma}^D - \boldsymbol{q}^D|} \quad \text{and} \quad \partial_g f = \sqrt{\frac{2}{3}}.$$

The multiplier ζ is determined by imposing the consistency condition $\dot{f} = 0$ [162] and in our case (of linear isotropic and kinematic hardening) an analytic formula is available in the plastic zone (where $f = 0$)

$$\zeta = \frac{\partial_{\boldsymbol{\sigma}} f : \dot{\boldsymbol{\sigma}}}{\sqrt{\frac{2}{3}} E_{iso} + \mathbb{H}\partial_{\boldsymbol{\sigma}} f : \partial_{\boldsymbol{\sigma}} f}$$

(see (4.105) for a proof). From (4.10), we get $\dot{\boldsymbol{\varepsilon}}_p = -\dot{\boldsymbol{r}}$. Assuming that the plastic variables $\boldsymbol{\varepsilon}_p$ and $\boldsymbol{r}$ are zero at the initial time instant, we deduce $\boldsymbol{\varepsilon}_p = -\boldsymbol{r}$ for all time t. The internal variable $\boldsymbol{r}$ has thus been characterized and $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\boldsymbol{r}}, \dot{\gamma}) = \mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$. Hence, (4.9) can be written as

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) = \sup_{(\boldsymbol{\tau}, \boldsymbol{p}, k) \in \mathbb{E}} ((\boldsymbol{\tau} - \boldsymbol{p}) : \dot{\boldsymbol{\varepsilon}}_p + k\dot{\gamma}) \tag{4.11}$$

and the equations in (4.7) can be re-written as

$$\boldsymbol{q} = \mathbb{H}\boldsymbol{\varepsilon}_p \quad \text{and} \quad g = -E_{iso}\gamma. \tag{4.12}$$

As seen in the above, the kinematic hardening law (first expression) and the isotropic hardening law (second expression) are linear. The material tensors $\mathbb{C}$ and $\mathbb{H}$ are assumed to be coercive, i.e., $\exists\ c_0 > 0$, $\exists\ h_0 > 0$ such that, $\forall \boldsymbol{\xi} \in \mathcal{M}_s^d$,

$$\mathbb{C}\boldsymbol{\xi} : \boldsymbol{\xi} \geq c_0 |\boldsymbol{\xi}|^2 \quad \text{and} \quad \mathbb{H}\boldsymbol{\xi} : \boldsymbol{\xi} \geq h_0 |\boldsymbol{\xi}|^2. \tag{4.13}$$

Using the relation (4.11), one can establish the equivalence [86]

$$(\boldsymbol{\sigma} - \boldsymbol{q}, g) \in \partial\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \iff (\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \in \partial\mathcal{D}^*(\boldsymbol{\sigma} - \boldsymbol{q}, g), \tag{4.14}$$

where $\mathcal{D}^*(\boldsymbol{\sigma} - \boldsymbol{q}, g)$ is the Legendre-Fenchel conjugate of $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$. The first expression in the above leads to the so-called dual form of plasticity and the second expression leads to the so-called primal form of plasticity. Thus, there are two distinct yet equivalent approaches to formulate the plasticity problem: dual and primal. The equivalence (4.14) using (4.11) can be established using the theory of convex analysis of gauge and indicator function we expound in the next subsection 4.1.1 (see Theorem 13).

4.1.1 Duality analysis

For this subsection, we consider a reflexive Banach space X and a convex subset $K \subset X$. The space X admits a dual space X^*. The following definitions and theorems are classical and we refer the interested reader to [86] for more details.

Definition 2. *We consider a function $f : X \to \mathbb{R}$.*

1. *The Legendre-Fenchel conjugate of $f(x)$ is given by $f^* : X^* \to \mathbb{R}$ as*

$$f^*(x^*) = \sup_{x \in X} \left(\langle x^*, x \rangle - f(x) \right).$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between X and X^*

2. *If $f(x)$ is convex, its subdifferential ∂f is defined as*

$$\partial f(x) = \{x^* \in X^* : f(y) \geq f(x) + \langle x^*, y - x \rangle \qquad \forall y \in X\}. \tag{4.15}$$

If $f(x)$ is smooth, the sub-differential admits a unique element such that

$$\partial f(x) = \{\nabla f(x)\}.$$

Definition 3. *Given a set $K \subset X$, we define*

1. *Indicator function* $I_K(x)$:

$$I_K(x) = \begin{cases} 0 & x \in K \\ \infty & x \notin K. \end{cases} \tag{4.16}$$

2. *Support function* $H_K(x^*)$:

$$H_K(x^*) = \sup_{x \in K} \langle x^*, x \rangle, \tag{4.17}$$

where $x^* \in X^*$

3. *If* K *is convex, its normal cone is given by*

$$\partial N_K(x) = \{x^* \in X^* : \langle x^*, y - x \rangle \leq 0 \qquad \forall y \in K\}. \tag{4.18}$$

One can quite easily verify that the support function and the indicator function are the Legendre-Fenchel conjugate of each other,

$$I_K^*(x^*) = H_K(x^*).$$

Moreover, the indicator function is reflexive with respect to the support function as

$$I_K^{**}(x) = I_K(x).$$

Lemma 6. *The sub-differential of* $I_K(x)$ *is related to the normal cone* $N_K(x)$ *as*

$$\partial I_K(x) = \begin{cases} \emptyset & x \notin K \\ N_K(x) & x \in K \end{cases}.$$

Proof. Using the definition of sub-differential in (4.18), we have

$$\partial I_K(x) = \{x^* \in X^* : I_K(y) \geq I_K(x) + \langle x^*, y - x \rangle \qquad \forall y \in X\}.$$

We first let $x \in K$ in the above. This gives

$$\forall y \in K, \qquad 0 \geq \langle x^*, y - x \rangle \qquad \text{and} \quad \forall y \notin K, \qquad \infty \geq \langle x^*, y - x \rangle.$$

Taking the union of the two cases, we have

$$\langle x^*, y - x \rangle \leq 0 \qquad \forall y \in X, \forall x \in K,$$

which is exactly the definition of the normal cone in (3). We then consider $x \notin K$. We prove by contradiction. We suppose that $\partial I_K(x) \neq \emptyset$. Then there exists $x^* \in X^*$ satisfying the sub-differential definition:

$$\forall y \in X, \qquad I_K(y) \geq I_K(x) + \langle x^*, y - x \rangle = \infty + \langle x^*, y - x \rangle.$$

It suffices to take $y \in K$ to show a contradiction, leading to the desired result. □

The following is an important theorem of this subsection.

Theorem 12. *Let* $K \subset X$ *be a convex set and* $I_K(x)$ *and* $H_K(x)$ *be functions defined in* (4.16)-(4.17). *Then for* $x \in K$, *we have*

$$x^* \in \partial I_K(x) \iff x \in \partial H_K(x^*).$$

Proof. We first consider $x^* \in \partial I_K(x)$. Using Lemma 6, the couple x^* and x satisfy

$$\langle x^*, y \rangle \leq \langle x^*, x \rangle \qquad \forall y \in K.$$

Taking the supremum over $y \in K$,

$$H_K(x) = \sup_{x \in K} \langle x^*, y \rangle \leq \langle x^*, x \rangle.$$

Adding $\langle y^*, x \rangle$ on both sides, we get

$$H_K(x) + \langle y^* - x^*, x \rangle \leq \langle y^*, x \rangle \leq \sup_{x \in K} \langle y^*, x \rangle = H_K(y^*).$$

Using definition (4.18) we prove $x \in \partial H_K(x^*)$.

Conversely, we assume $x \in \partial H_K(x^*)$. Then by definition

$$\langle x^*, x\rangle - H_K(x) \leq \langle y^*, x\rangle - H_K(y^*) \qquad \forall y^* \in X^*.$$

Taking the supremum on the right, we get the Legendre-Fenchel conjugate $I_K(x)$. Since $I_K(x) = 0$ for $x \in K$, the above reduces to

$$H_K(x) \leq \langle x^*, x\rangle,$$

or for all $y^* \in X^*$,

$$\langle y^*, x\rangle \leq \langle x^*, x\rangle$$

implying $x^* \in \partial I_K(x)$. □

We end this subsection with a theorem that establishes the equivalence (4.14) from (4.11).

Theorem 13. *Let* $(\boldsymbol{\sigma} - \boldsymbol{q}, g) \in \mathcal{M}_s^d \times \mathbb{R}$, $(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \in \mathcal{M}_s^d \times \mathbb{R}$, $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$ *be a function defined in* (4.11). *Then we get*

$$(\boldsymbol{\sigma} - \boldsymbol{q}, g) \in \partial\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \iff (\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \in \partial\mathcal{D}^*(\boldsymbol{\sigma} - \boldsymbol{q}, g), \tag{4.19}$$

Proof. Taking

- the space $X = \mathcal{M}_s^d \times \mathbb{R}$,
- the convex set
$$K = \left\{ (\boldsymbol{\sigma} - \boldsymbol{q}, g) \in X,\ |\boldsymbol{\sigma}^D - \boldsymbol{q}^D| + \sqrt{\frac{2}{3}}\,(\sigma_Y - g) \leq 0,\ g \leq 0 \right\},$$
- the expression in (4.11)
$$\mathcal{D} = \sup_{(\boldsymbol{\tau},\boldsymbol{p},k)\in\mathbb{E}} ((\boldsymbol{\tau} - \boldsymbol{p}) : \dot{\boldsymbol{\varepsilon}}_p + k\dot{\gamma}) = \sup_{(\boldsymbol{\tau}-\boldsymbol{p},k)\in K} ((\boldsymbol{\tau} - \boldsymbol{p}) : \dot{\boldsymbol{\varepsilon}}_p + k\dot{\gamma}),$$
- and the support function $H \equiv \mathcal{D}$

in the Theorem 12, one can obtain the relation (4.19). □

As shown in Theorem 13, using (4.11) one can derive

$$(\boldsymbol{\sigma} - \boldsymbol{q}, g) \in \partial\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}).$$

This is same as stating that the $(\boldsymbol{\sigma} - \boldsymbol{q}, g)$ belong to the sub differential of $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$. Using the sub differential definition (4.18), we get

$$\mathcal{D}(\boldsymbol{\varepsilon}_q, \mu) \geq \mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) + (\boldsymbol{\sigma} - \boldsymbol{q}) : (\boldsymbol{\varepsilon}_q - \dot{\boldsymbol{\varepsilon}}_p) + g(\mu - \dot{\gamma}) \quad \forall \boldsymbol{\varepsilon}_q \in \mathcal{M}_s^d, \mu \in \mathbb{R}. \tag{4.20}$$

The dual form and the primal form can be derived easily from the above. For the primal form, it suffices to find an expression of $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$ in terms of $\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}$. For the dual form, one can simply use the definition of $\mathcal{D}$ (4.6) and differentiate the stress-strain relations (4.5), (4.12), with respect to time.

4.1.2 Primal form

The primal variables are $(\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)$. We remind the reader that the dissipation function is given by

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) = \sup_{(\boldsymbol{\sigma},\boldsymbol{q},g)\in\mathbb{E}} ((\boldsymbol{\sigma} - \boldsymbol{q}) : \dot{\boldsymbol{\varepsilon}}_p + g\dot{\gamma}). \tag{4.21}$$

The above is equivalent to stating that out of the possible stress states $(\boldsymbol{\sigma}, \boldsymbol{q}, g)$, the structure finds the stress state that maximizes the energy dissipated. The expression (4.20) is a precursor to the primal form. Since we work with primal variables, we need an expression of $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$ in terms of the primal variables. This expression depends on the yield criterion chosen. As stated earlier, we work with von Mises yield criterion. For this criterion, the expression is given in the following result [153, 87].

Lemma 7. *The dissipation function $\mathcal{D}$ satisfies*

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p,\dot\gamma)=\begin{cases}\sqrt{\frac{2}{3}}\sigma_Y|\dot{\boldsymbol{\varepsilon}}_p| & \text{if}\quad \sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p|\le\dot\gamma,\\ \infty & \text{if}\quad \sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p|>\dot\gamma.\end{cases}\tag{4.22}$$

As a consequence the domain of $\mathcal{D}$ is defined by

$$dom\mathcal{D}=\left\{(\dot{\boldsymbol{\varepsilon}}_p,\dot\gamma),\ \sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p|\le\dot\gamma\ \ a.e.\ \ in\ \Omega\right\}.\tag{4.23}$$

Proof. We compute the maximum in (4.11):

$$\begin{aligned}\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p,\dot\gamma)&=\sup_{(\boldsymbol{\sigma},\boldsymbol{q},g)\in\mathbb{E}}\{(\boldsymbol{\sigma}-\boldsymbol{q})\colon\dot{\boldsymbol{\varepsilon}}_p+g\dot\gamma\}\\&=\sup_{(\boldsymbol{\sigma},\boldsymbol{q},g)\in\mathbb{E}}\{(\boldsymbol{\sigma}^D-\boldsymbol{q}^D)\colon\dot{\boldsymbol{\varepsilon}}_p+g\dot\gamma\}\qquad\text{(as }\dot{\boldsymbol{\varepsilon}}_p\text{ has zero trace)}\\&=\sup_{(\boldsymbol{\sigma},\boldsymbol{q},g)\in\mathbb{E}}\{|\boldsymbol{\sigma}^D-\boldsymbol{q}^D||\dot{\boldsymbol{\varepsilon}}_p|+g\dot\gamma\}\qquad\text{(maximum value is attained if the two tensors are parallel)}\\&=\sup_{g\le0}\left\{\sqrt{\frac{2}{3}}(\sigma_Y-g)|\dot{\boldsymbol{\varepsilon}}_p|+g\dot\gamma\right\}\qquad\text{(using (4.2))}\\&=\sup_{g\le0}\left\{\sqrt{\frac{2}{3}}\sigma_Y|\dot{\boldsymbol{\varepsilon}}_p|+g\left(\dot\gamma-\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p|\right)\right\}\end{aligned}$$

which yields (4.22). □

Evidently, $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p,\dot\gamma)=\sqrt{\frac{2}{3}}\sigma_Y|\dot{\boldsymbol{\varepsilon}}_p|$ on dom$\mathcal{D}$. Eventually, we obtain the primal form of plasticity (that shall be treated in this chapter):

$$\begin{aligned}\boldsymbol{\sigma}&=\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u})-\boldsymbol{\varepsilon}_p)&&\text{in }\Omega\times(0,T],\\&&&\qquad(4.24)\\\operatorname{div}(\boldsymbol{\sigma})+\boldsymbol{f}&=\boldsymbol{0}&&\text{in }\Omega\times(0,T],\\\boldsymbol{\sigma}\cdot\boldsymbol{n}&=\boldsymbol{g}&&\text{on }\Gamma_N\times(0,T],\\\boldsymbol{\sigma}\cdot\boldsymbol{n}&=\boldsymbol{0}&&\text{on }\Gamma\times(0,T],\\\boldsymbol{u}&=\boldsymbol{0}&&\text{on }\Gamma_D\times(0,T],\\\sqrt{\frac{2}{3}}\sigma_Y|\boldsymbol{\varepsilon}_q|\ge\sqrt{\frac{2}{3}}\sigma_Y|\dot{\boldsymbol{\varepsilon}}_p|&+(\boldsymbol{\sigma}-\mathbb{H}\boldsymbol{\varepsilon}_p):(\boldsymbol{\varepsilon}_q-\dot{\boldsymbol{\varepsilon}}_p)-E_{iso}\gamma(\mu-\dot\gamma)\quad\forall(\boldsymbol{\varepsilon}_q,\mu)\in\mathrm{dom}\mathcal{D}&&\text{on }\Omega\times(0,T].\\&&&\qquad(4.25)\end{aligned}$$

The inequality (4.25) is obtained by injecting (4.12) in (4.20).

Very often, the partial differential equations (4.24) are solved in conjunction with the ordinary differential equations (4.10). But here, we solve (4.24) coupled to the inequation (4.25). This coupling, which is purely in terms of the variables $(\boldsymbol{u},\boldsymbol{\varepsilon}_p,\gamma)$ results in the so-called primal formulation. This formulation is treated analytically in Section 4.2 and the corresponding shape derivative is computed in Section 4.3. While the primal formulation illuminates the theoretical properties of the variables $(\boldsymbol{u},\boldsymbol{\varepsilon}_p,\gamma)$, it is hard to implement numerically. Typically, one resorts to the radial return algorithm, explained in Section 4.4.2.

4.1.3 Dual form

The dual variables are $(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{q},g)$. To obtain the dual form, it suffices to substitute $\dot{\boldsymbol{r}}=-\dot{\boldsymbol{\varepsilon}}_p$ in the expression of $\mathcal{D}$ (4.6)

$$\mathcal{D}=(\boldsymbol{\sigma}-\boldsymbol{q})\colon\dot{\boldsymbol{\varepsilon}}_p+g\dot\gamma$$

and differentiate the stress-strain relations (4.5) and (4.12) with respect to time

$$\dot{\boldsymbol{\sigma}}=\mathbb{C}(\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}})-\dot{\boldsymbol{\varepsilon}}_p)$$

$$\dot{\boldsymbol{\varepsilon}}_p = \mathbb{H}^{-1}\dot{\boldsymbol{q}} \quad \text{and} \quad \dot{g} = -E_{iso}\dot{\gamma}.$$

This dual form can be obtained by eliminating $\dot{\boldsymbol{\varepsilon}}_p$ and$\dot{\gamma}$ in the rule (4.8) to obtain

$$(\mathbb{C}\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) - \dot{\boldsymbol{\sigma}}) : \mathbb{C}^{-1}(\boldsymbol{\tau} - \boldsymbol{\sigma}) + \dot{\boldsymbol{q}} : \mathbb{H}^{-1}(\boldsymbol{p} - \boldsymbol{q}) + \frac{(g-k)\dot{g}}{E_{iso}} \leq 0 \qquad \forall (\boldsymbol{\tau}, \boldsymbol{p}, k) \in \mathbb{E}. \tag{4.27}$$

The dual inequation (4.27) solves for $(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{q}, g)$ whereas the primal inequation (4.25) solves for $(\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)$. The inequation (4.27) integrated over Ω results in the so-called dual variational formulation. The integration of (4.20) over Ω results in the so-called primal formulation. The variables $(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{q}, g)$ are called dual variables and the variables $(\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)$ are called primal variables.

The dual variational formulation has been introduced and studied [69]. It remains well-posed under the assumption that the kinematic hardening is non-zero [86]. The same holds true for the primal formulation. We do not resort to this dual formulation because of the difficulties arising in the finite element implementation. One has to resort to a stress formulations, which needs complex polynomial spaces for the stress variables [155]. The analytical and numerical treatment of the primal formulation being much easier than that of the dual formulation, we have chosen the former in this study.

The mathematical analysis performed in Sections (4.2)-(5.1) holds true for anisotropic materials as long as the Hooke's tensor is positive definite. But all of our numerical simulations are performed for isotropic materials only.

4.2 Variational formulation using Primal form

Integrating the equations equations and inequation in (4.25), we derive the primal variational formulation (4.35). This formulation is further subject to penalization and then to regularization in order to make it differentiable. This section closes with some statements about the well-posedness and the convergence of the solution of the penalized-regularized plasticity model towards the exact solution.

We define the displacement space

$$V = \{\boldsymbol{u} \in H^1(\Omega)^d,\ \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_D\}$$

and the space of of plastic strain Q as

$$Q = \{\boldsymbol{\varepsilon}_q \in L^2(\Omega)^{d\times d},\ \boldsymbol{\varepsilon}_q = \boldsymbol{\varepsilon}_q^T,\ \mathrm{tr}(\boldsymbol{\varepsilon}_q) = \boldsymbol{0} \text{ a.e. in } \Omega\}. \tag{4.28}$$

We then define the product space

$$Z = V \times Q \times L^2(\Omega), \tag{4.29}$$

where we seek the solution $\boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)$. The space Z is a Hilbert space equipped with the scalar product, for $\boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)$ and $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{\varepsilon}_q, \mu)$,

$$\langle \boldsymbol{w}, \boldsymbol{z} \rangle = \int_\Omega \boldsymbol{u} \cdot \boldsymbol{v} \ \ dx + \int_\Omega \boldsymbol{\varepsilon}_p : \boldsymbol{\varepsilon}_q \ \ dx + \int_\Omega \gamma\mu \ \ dx. \tag{4.30}$$

Let Z^* be the dual space of Z. The forces are assumed to be smooth as

$$\boldsymbol{f} \in H^1([0,T], L^2(\Omega)^d) \quad \text{and} \quad \boldsymbol{g} \in H^1([0,T], L^2(\Gamma_N)^d).$$

Indeed, since $H^1([0,T], H) \subset \mathcal{C}^0([0,T], H)$ for any Hilbert space H, at any time t the forces $\boldsymbol{f}(t)$ and $\boldsymbol{g}(t)$ are well defined. We introduce a bilinear form $a : Z \times Z \to \mathbb{R}$,

$$a(\boldsymbol{w}, \boldsymbol{z}) = \int_\Omega (\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p) : (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\varepsilon}_q) + \boldsymbol{\varepsilon}_q : \mathbb{H}\boldsymbol{\varepsilon}_p + E_{iso}\gamma\mu) \ \ dx, \tag{4.31}$$

and a linear form $l_t : Z \to \mathbb{R}$ such that

$$l_t(\boldsymbol{z}) = \int_\Omega \boldsymbol{f}(t) \cdot \boldsymbol{v} \ \ dx + \int_{\Gamma_N} \boldsymbol{g}(t) \cdot \boldsymbol{v} \ ds, \tag{4.32}$$

with the forces $\boldsymbol{f}(t) \in L^2(\Omega)^d$, $\boldsymbol{g}(t) \in L^2(\Gamma_N)^d$ and a nonlinear convex functional $j : Z \to \mathbb{R}$ such that

$$j(\boldsymbol{z}) = \int_\Omega \mathcal{D}(\boldsymbol{\varepsilon}_q, \mu) \ \ dx, \tag{4.33}$$

where $\mathcal{D}(\boldsymbol{\varepsilon}_q, \mu)$ is defined by (4.22). This functional $j(\cdot)$ is convex and lower semi-continuous on Z and it is Lipschitz continuous on the convex set $K \subset Z$ defined as

$$K = V \times \mathrm{dom}\mathcal{D},$$

where $\mathrm{dom}\mathcal{D}$ is defined by (4.23). The admissible plastic flow rates $\dot{\boldsymbol{\varepsilon}}_p$, $\dot{\gamma}$ belong to the convex set $\mathrm{dom}\mathcal{D}$.

Lemma 8. *The bilinear form $a(\cdot,\cdot)$ defined in (4.31) is coercive on Z.*

Proof. From (4.31) with $\boldsymbol{z} = \boldsymbol{w} \in Z$, and for any $s \in (0,1)$, we get

$$\begin{aligned}
a(\boldsymbol{w},\boldsymbol{w}) &= \int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p) : (\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p)\ dx + \int_\Omega \mathbb{H}\boldsymbol{\varepsilon}_p : \boldsymbol{\varepsilon}_p\ dx + \int_\Omega E_{iso}\gamma^2\ dx \\
&\geq c_0 \left\|\boldsymbol{\varepsilon}(\boldsymbol{u})\right\|^2_{L^2(\Omega)} + (c_0 + h_0)\left\|\boldsymbol{\varepsilon}_p\right\|^2_{L^2(\Omega)} - \int_\Omega 2\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}_p\ dx + E_{iso}\left\|\gamma\right\|^2_{L^2(\Omega)} \\
&\geq (c_0 - c_0(1-s))\left\|\boldsymbol{\varepsilon}(\boldsymbol{u})\right\|^2_{L^2(\Omega)} + \left(c_0 + h_0 - \frac{c_0}{1-s}\right)\left\|\boldsymbol{\varepsilon}_p\right\|^2_{L^2(\Omega)} + E_{iso}\left\|\gamma\right\|^2_{L^2(\Omega)} \quad s \in (0,1) \\
&= c_0 s \left\|\boldsymbol{\varepsilon}(\boldsymbol{u})\right\|^2_{L^2(\Omega)} + \left(h_0 - \frac{c_0 s}{1-s}\right)\left\|\boldsymbol{\varepsilon}_p\right\|^2_{L^2(\Omega)} + E_{iso}\left\|\gamma\right\|^2_{L^2(\Omega)} \\
&\geq \min\left(c_0 s, \left(h_0 - \frac{c_0 s}{1-s}\right), E_{iso}\right)\left(\left\|\boldsymbol{\varepsilon}(\boldsymbol{u})\right\|^2_{L^2(\Omega)} + \left\|\boldsymbol{\varepsilon}_p\right\|^2_{L^2(\Omega)} + \left\|\gamma\right\|^2_{L^2(\Omega)}\right).
\end{aligned}$$

We choose $s = \frac{h_0}{2c_0 + h_0}$ in order to make the right hand side positive for all $\boldsymbol{w} \in Z$. Finally using Korn's inequality, this proves the coercivity of $a(\cdot,\cdot)$ on Z. □

In order to obtain the primal formulation of (4.24) and (4.25), we multiply (4.1) by $\boldsymbol{v} - \dot{\boldsymbol{u}}$, use (4.5) and integrate the product over Ω by parts to obtain

$$\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p) : (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}))\ dx = \int_\Omega \boldsymbol{f}(t)\cdot(\boldsymbol{v} - \dot{\boldsymbol{u}})\ dx + \int_{\Gamma_N} \boldsymbol{g}(t)\cdot(\boldsymbol{v} - \dot{\boldsymbol{u}})ds \qquad \forall \boldsymbol{v} \in V, \tag{4.34}$$

We then integrate (4.25) over Ω, add (4.34) to it and obtain the variational inequality, for any $\boldsymbol{z} \in K$,

$$\begin{aligned}
\int_\Omega \sqrt{\frac{2}{3}}\sigma_Y |\boldsymbol{\varepsilon}_q|\ dx \geq \int_\Omega \sqrt{\frac{2}{3}}\sigma_Y |\dot{\boldsymbol{\varepsilon}}_p|\ dx + \int_\Omega \boldsymbol{f}(t)\cdot(\boldsymbol{v} - \dot{\boldsymbol{u}})\ dx + \int_{\Gamma_N} \boldsymbol{g}(t)\cdot(\boldsymbol{v} - \dot{\boldsymbol{u}})\ ds \\
- \int_\Omega \left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p) : (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\varepsilon}_q - \boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}) + \dot{\boldsymbol{\varepsilon}}_p) + \mathbb{H}\boldsymbol{\varepsilon}_p : (\boldsymbol{\varepsilon}_q - \dot{\boldsymbol{\varepsilon}}_p) + E_{iso}\gamma(\mu - \dot{\gamma})\right)\ dx.
\end{aligned}$$

We complement this variational inequality with the following initial conditions

$$\boldsymbol{u}(0) = \boldsymbol{0}, \quad \boldsymbol{\varepsilon}_p(0) = \boldsymbol{0}, \quad \gamma(0) = 0 \qquad \text{in } \Omega.$$

To prove existence and uniqueness of a solution, we rely on theorem 4.3 in [87] which requires some additional regularity in time for the solution. Therefore, we assume that the forces satisfy

$$\boldsymbol{f}(0) = \boldsymbol{0} \quad \text{in } \Omega \quad \text{and} \quad \boldsymbol{g}(0) = \boldsymbol{0} \quad \text{on } \Gamma_N.$$

Using the linear forms and the nonlinear functional defined earlier, we obtain the primal form of the plasticity problem (4.24) and (4.25): find $\boldsymbol{w}(t) = (\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)(t)$ with $\boldsymbol{w}(0) = \boldsymbol{0}$ such that $\dot{\boldsymbol{w}}(t) \in K$ (for almost all $t \in (0,T)$) and

$$a(\boldsymbol{w}, \boldsymbol{z} - \dot{\boldsymbol{w}}) + j(\boldsymbol{z}) - j(\dot{\boldsymbol{w}}) \geq l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}) \quad \forall \boldsymbol{z} \in K. \tag{4.35}$$

As a result of Theorem 4.3 in [87] mentioned below, the variational inequality (4.35) is well-posed.

Theorem 14. *[87] Let Z be a Hilbert space; $K \subset Z$ be a nonempty, closed, convex cone; $a : Z \times Z \to \mathbb{R}$ a continuous bilinear form that is symmetric and coercive; $j : K \to \mathbb{R}$ non-negative, convex, positively homogeneous, Lipschitz continuous form; $l_t \in H^1([0,T], Z^*)$ with $l_0(\cdot) = 0$. Then there exists a unique $\boldsymbol{w} \in H^1([0,T], Z)$ satisfying (4.35).*

Remark 5. *In the absence of kinematic hardening or $h_0 = 0$, we cannot show the coercivity of $a(\cdot,\cdot)$ and thus the well-posedness of the problem (4.35). Thus by taking hardening parameters to be zero (E_{iso} and $\mathbb{H}$) in the problem (4.35), one cannot study the perfect plasticity problem.*

Equation (4.35) is not shape-differentiable [133, 171] in the classical sense and we will approximate it by a smooth variational equation in Subsection 4.2.2.

Remark 6. *The sub-differential of $j(\cdot)$ admits the form*

$$\partial j(\dot{\boldsymbol{w}}) = \{\exists \boldsymbol{w}^* \in Z |\ j(\boldsymbol{z}) \geq j(\dot{\boldsymbol{w}}) + \langle \boldsymbol{w}^*, \boldsymbol{z} - \dot{\boldsymbol{w}} \rangle \quad \forall \boldsymbol{z} \in Z\}. \tag{4.36}$$

$j(\cdot)$ being non-smooth in its argument, the derivative $\nabla j(\dot{\boldsymbol{w}})$ (or $\boldsymbol{w}^*$ in the above) is not uniquely-defined only at $\dot{\boldsymbol{w}} = \boldsymbol{0}$. In the inequality in (4.36), the term on the left is convex and is bounded by a linear term on the right for all Z. In addition, for $\boldsymbol{z} = \dot{\boldsymbol{w}}$, we have an equality, which implies that the bound on the right is a tangent. We re-write (4.35) as

$$j(\boldsymbol{z}) - j(\dot{\boldsymbol{w}}) \geq a(\boldsymbol{w}, \dot{\boldsymbol{w}} - \boldsymbol{z}) + \langle l, \boldsymbol{z} - \dot{\boldsymbol{w}} \rangle \quad \forall \boldsymbol{z} \in K.$$

Again, the convex term on the left and is bounded optimally by the tangent on the right. Since the tangent is at the point $\boldsymbol{z} = \dot{\boldsymbol{w}}$, we find that (4.35) is indeed equivalent to finding $\boldsymbol{w}$ such that

$$a(\boldsymbol{w}, \dot{\boldsymbol{w}} - \boldsymbol{z}) + \langle \boldsymbol{w}^*, \dot{\boldsymbol{w}} - \boldsymbol{z} \rangle = \langle l, \dot{\boldsymbol{w}} - \boldsymbol{z} \rangle \qquad \forall \boldsymbol{z} \in K.$$

Since, $\dot{\boldsymbol{w}} + \boldsymbol{z} \in K$, we can replace $\boldsymbol{z}$ in above by $\dot{\boldsymbol{w}} + \boldsymbol{z}$ and re-write the above as

$$a(\boldsymbol{w}, \boldsymbol{z}) + \langle \boldsymbol{w}^*, \boldsymbol{z} \rangle = \langle l, \boldsymbol{z} \rangle \qquad \forall \boldsymbol{z} \in K. \tag{4.37}$$

In the proceeding subsection, we approximate Eqn.(4.37) by Eqn.(4.56).

Equation (4.35) is not differentiable as $j(\cdot)$ is non-smooth or $\boldsymbol{w}^*$ being not uniquely defined. We seek to approximate the above problem or its equivalent (4.35) by a smooth variational equation.

4.2.1 Penalization

We approximate the problem (4.35) posed on the convex set K by a problem posed on the vector space Z by penalizing the constraint $\boldsymbol{z}(t) \in K$. We introduce a penalization constant $0 < \epsilon \ll 1$ and modify the dissipation function $\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$ to $\mathcal{D}_\epsilon(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma})$ as

$$\mathcal{D}_\epsilon(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) = \sqrt{\frac{2}{3}}\sigma_Y \left(|\dot{\boldsymbol{\varepsilon}}_p| + \frac{1}{\epsilon} \max\left(\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p| - \dot{\gamma}, 0 \right) \right). \tag{4.38}$$

Because of the above penalization, we indeed allow the stress state $\boldsymbol{\sigma}$ to exceed the yield limit, exactly like in the case of viscoplasticity. In viscoplasticity, the plastic flow is dependent on the rate of loading. Moreover, one considers the dissipation energy function $\mathcal{D}$ written in the dual form and the von Mises yield criterion, written in terms of dual variables $\boldsymbol{\sigma}, \boldsymbol{q}, g$ is penalized (see [162], equation 7.5b), leading to

$$\mathcal{D}_\epsilon = (\boldsymbol{\sigma} - \boldsymbol{q}) : \dot{\boldsymbol{\varepsilon}}_p + g\dot{\gamma} - \frac{1}{\epsilon} \max(f(\boldsymbol{\sigma}, \boldsymbol{q}), 0)^2$$

where ϵ here corresponds to the viscosity. While the viscoplasticity approach is similar to the penalization (4.38), the exact correspondance between the two is not clear to us.

As seen in (4.38), the exponent of the penalization term is one. We can also consider an exponent two (quadratic penalization) or more. However, we choose the exponent to be one so that we can show existence, uniqueness and convergence of the approximate solution with ease. The penalty-based methods for plasticity have been proven to converge only for a finite dimensional problem [119]. Here we demonstrate the convergence for the infinite dimensional problem. Penalty approaches to solve viscoplasticity have already been studied [192].

We now modify $j(\cdot)$ to $j_\epsilon : Z \to \mathbb{R}$ as

$$j_\epsilon(\dot{\boldsymbol{w}}) = \int_\Omega \mathcal{D}_\epsilon(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \ dx.$$

Problem (4.35) now becomes: find $\boldsymbol{w}_\epsilon(t) \in Z$ such that $\boldsymbol{w}_\epsilon(0) = \boldsymbol{0}$, $\dot{\boldsymbol{w}}_\epsilon(t) \in Z$ and

$$a(\boldsymbol{w}_\epsilon, \boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) + j_\epsilon(\boldsymbol{z}) - j_\epsilon(\dot{\boldsymbol{w}}_\epsilon) \geq l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) \qquad \forall \boldsymbol{z} \in Z. \tag{4.39}$$

The above penalized problem is well-posed as the following theorem shows.

Theorem 15. *The problem (4.39) admits a unique solution* $\boldsymbol{w}_\epsilon \in H^1([0,T],Z)$.

Proof. The bilinear form $a(\cdot,\cdot)$ is coercive in Z (as shown in Lemma 8), and the nonlinearity $\mathcal{D}_\epsilon(\cdot)$ is positively homogeneous, Lipschitz continuous and convex as $\forall \dot{\boldsymbol{w}} = (\dot{\boldsymbol{u}}, \dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}),\ \boldsymbol{z} = (\dot{\boldsymbol{v}}, \dot{\boldsymbol{\varepsilon}}_q, \dot{\mu}) \in Z,\ t \in [0,1]$,

$$
\begin{aligned}
&\frac{1}{\sigma_Y}\sqrt{\frac{3}{2}}\mathcal{D}_\epsilon(t\dot{\boldsymbol{\varepsilon}}_p + (1-t)\dot{\boldsymbol{\varepsilon}}_q, t\dot{\gamma} + (1-t)\dot{\mu}) \\
&= |t\dot{\boldsymbol{\varepsilon}}_p + (1-t)\dot{\boldsymbol{\varepsilon}}_q| + \frac{1}{\epsilon}\max\left(\sqrt{\frac{2}{3}}|t\dot{\boldsymbol{\varepsilon}}_p + (1-t)\dot{\boldsymbol{\varepsilon}}_q| - t\dot{\gamma} - (1-t)\dot{\mu}, 0\right) \\
&\leq t|\dot{\boldsymbol{\varepsilon}}_p| + (1-t)|\dot{\boldsymbol{\varepsilon}}_q| + \frac{1}{\epsilon}\max\left(t\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p| + (1-t)\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_q| - t\dot{\gamma} - (1-t)\dot{\mu}, 0\right) \\
&= t|\dot{\boldsymbol{\varepsilon}}_p| + (1-t)|\dot{\boldsymbol{\varepsilon}}_q| + \frac{1}{\epsilon}\left(\max\left(t\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p| - t\dot{\gamma}, 0\right) + \max\left((1-t)\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_q| - (1-t)\dot{\mu}, 0\right)\right) \\
&= \frac{1}{\sigma_Y}\sqrt{\frac{3}{2}}\left(t\mathcal{D}_\epsilon(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) + (1-t)\mathcal{D}_\epsilon(\dot{\boldsymbol{\varepsilon}}_q, \dot{\mu})\right).
\end{aligned}
\tag{4.40}
$$

Finally $j_\epsilon(\cdot)$ is non-negative, convex, positively homogeneous and Lipschitz continuous on Z. With these properties, (4.39) admits a unique solution $\boldsymbol{w}_\epsilon \in H^1([0,T],Z)$ using Theorem 14. □

Remark 7. *In the absence of kinematic hardening or* $h_0 = 0$, *we cannot show the coercivity of* $a(\cdot,\cdot)$ *and thus the well-posedness of the problem (4.39).*

Now we split $j_\epsilon(\boldsymbol{z})$ in two as

$$
j_\epsilon(\boldsymbol{z}) = j_1(\boldsymbol{z}) + \frac{1}{\epsilon}j_2(\boldsymbol{z}) \qquad \text{such that} \tag{4.41}
$$

$$
j_1(\boldsymbol{z}) = \sqrt{\frac{2}{3}}\sigma_Y\int_\Omega |\boldsymbol{\varepsilon}_q|\, dx \qquad \text{and} \qquad j_2(\boldsymbol{z}) = \sqrt{\frac{2}{3}}\sigma_Y\int_\Omega \max\left(\sqrt{\frac{2}{3}}|\boldsymbol{\varepsilon}_q| - \mu, 0\right) dx.
$$

We state an important Lemma that shall be used several times in this section.

Lemma 9. *The functionals* $\int_0^T j_\epsilon(\cdot)dt$, $\int_0^T j_1(\cdot)dt$, $\int_0^T j_2(\cdot)dt$ *are weakly lower semi-continuous in* $L^1([0,T],Z)$.

Proof. Since $\int_0^T j_\epsilon(\cdot)dt$, $\int_0^T j_1(\cdot)dt$, $\int_0^T j_2(\cdot)dt$ are lower semi-continuous in $L^1([0,T],Z)$ and convex, they are weakly lower semi-continuous in $L^1([0,T],Z)$. □

The solutions to (4.39), $\boldsymbol{w}_\epsilon$ converge weakly and strongly to $\boldsymbol{w}$ (solution to (4.35)) as $\epsilon \longrightarrow 0$ because of the Lemma 9 that exploits the convexity of $j_\epsilon(\cdot)$. We first state the theorem concerning the weak convergence of $\boldsymbol{w}_\epsilon$ and later state the theorem concerning the strong convergence.

Theorem 16. *The sequence of solutions* $\boldsymbol{w}_\epsilon$ *to (4.39) admit a subsequence that satisfies as* $\epsilon \longrightarrow 0$,

$$
\boldsymbol{w}_\epsilon \overset{*}{\rightharpoonup} \boldsymbol{w} \quad \text{in } L^\infty([0,T],Z) \quad \text{and} \quad \dot{\boldsymbol{w}}_\epsilon \rightharpoonup \dot{\boldsymbol{w}} \quad \text{in } L^2([0,T],Z),
$$

where $\boldsymbol{w}$ *is the solution to (4.35). Moreover as* $\epsilon \longrightarrow 0$,

$$
\boldsymbol{w}_\epsilon \longrightarrow \boldsymbol{w} \quad \text{in } L^\infty([0,T],Z).
$$

Proof. We discretize the time interval $[0,T]$ in $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ with $t_n - t_{n-1} = \delta t$. For $l_t \in Z^*$(defined in (4.32)), let $l_n = l_{t_n}$, $\Delta l_n = l_n - l_{n-1}$. Starting from $\boldsymbol{w}_0 = \boldsymbol{0}$ we construct a sequence $\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \Delta\boldsymbol{w}_n$ and a function $\boldsymbol{w}_\epsilon^{\delta t} \in \mathcal{C}^0([0,T],K)$ such that $\boldsymbol{w}_\epsilon^{\delta t}(t) = \boldsymbol{w}_n + (t-t_n)/\delta t\Delta\boldsymbol{w}_n$ for $t \in [t_n, t_{n+1}]$, where $\Delta\boldsymbol{w}_n \in Z$ is the solution of

$$
a(\Delta\boldsymbol{w}_n, \boldsymbol{z} - \Delta\boldsymbol{w}_n) + j_\epsilon(\boldsymbol{z}) - j_\epsilon(\Delta\boldsymbol{w}_n) \geq l_n(\boldsymbol{z} - \Delta\boldsymbol{w}_n) - a(\boldsymbol{w}_n, \boldsymbol{z} - \Delta\boldsymbol{w}_n) \qquad \forall \boldsymbol{z} \in Z. \tag{4.42}
$$

The above is a time-discretized version of (4.39). The problem (4.99) is a variational inequation of the second kind and admits a unique solution [80]. Lemma 4.1 in [87] gives the bound

$$\|\Delta \boldsymbol{w}_n\|_Z \leq C \|\Delta l_n\|_{Z^*},$$

where the constant C is independent of ϵ and δt. Then, using Lemma 4.2 in [87], one can deduce that as $\delta t \longrightarrow 0$, $\boldsymbol{w}_\epsilon^{\delta t} \overset{*}{\rightharpoonup} \boldsymbol{w}_\epsilon$ in $L^\infty([0,T],Z)$, $\|\boldsymbol{w}_\epsilon\|_{L^\infty([0,T],Z)} \leq C_1$ and $\|\dot{\boldsymbol{w}}_\epsilon\|_{L^2([0,T],Z)} \leq C_2$, where the constants C_1 and C_2 are independent of ϵ and δt. These two bounds imply the existence of a subsequence $\boldsymbol{w}_\epsilon$ and a limit $\boldsymbol{w}^* \in H^1([0,T],Z)$ such that $\boldsymbol{w}_\epsilon \overset{*}{\rightharpoonup} \boldsymbol{w}^*$ in $L^\infty([0,T],Z)$ and $\dot{\boldsymbol{w}}_\epsilon \rightharpoonup \dot{\boldsymbol{w}}^*$ in $L^2([0,T],Z)$. Since T is finite, this implies that $\dot{\boldsymbol{w}}_\epsilon \rightharpoonup \dot{\boldsymbol{w}}^*$ in $L^1([0,T],Z)$. It remains to show that this limit $\boldsymbol{w}^*$ is equal to the solution $\boldsymbol{w}$ of (4.35). For that, we integrate (4.39) from $t=0$ to T and take $\boldsymbol{z} = \boldsymbol{z}_0 \in K$ (an arbitrary element) to obtain for all $\epsilon > 0$,

$$\begin{aligned}
\int_0^T j_\epsilon(\dot{\boldsymbol{w}}_\epsilon)dt &\leq \int_0^T \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z}_0 - \dot{\boldsymbol{w}}_\epsilon) + j_\epsilon(\boldsymbol{z}_0) - l_t(\boldsymbol{z}_0 - \dot{\boldsymbol{w}}_\epsilon)\right) dt \\
&= \int_0^T \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z}_0 - \dot{\boldsymbol{w}}_\epsilon) + j(\boldsymbol{z}_0) - l_t(\boldsymbol{z}_0 - \dot{\boldsymbol{w}}_\epsilon)\right) dt \\
&\leq \left(\|\boldsymbol{w}_\epsilon\|_{L^\infty([0,T],Z)} + \|l_t\|_{L^\infty([0,T],Z^*)}\right) \left(\|\boldsymbol{z}_0\|_Z T + \|\dot{\boldsymbol{w}}_\epsilon\|_{L^2([0,T],Z)} \sqrt{T}\right) + C_0 \|\boldsymbol{z}_0\|_Z \\
&\leq C,
\end{aligned}$$

where C is a constant independent of ϵ. Using Lemma 9 and the non-negativity of j_1 and j_2 (defined in (4.41)), we have

$$\begin{aligned}
\int_0^T j_2(\dot{\boldsymbol{w}}^*)dt \leq \liminf_{\epsilon\to 0} \int_0^T j_2(\dot{\boldsymbol{w}}_\epsilon)dt &\leq \lim_{\epsilon\to 0} \epsilon \left(\int_0^T \left(j_\epsilon(\dot{\boldsymbol{w}}_\epsilon) - j_1(\dot{\boldsymbol{w}}_\epsilon)\right) dt\right) \\
&\leq \lim_{\epsilon\to 0} \epsilon \left(C - \int_0^T j_1(\dot{\boldsymbol{w}}_\epsilon)dt\right) \leq \lim_{\epsilon\to 0} C\epsilon = 0.
\end{aligned}$$

Thus, $\displaystyle\int_0^T j_2(\dot{\boldsymbol{w}}^*)dt = 0$, so that $j_2(\dot{\boldsymbol{w}}^*) = 0$ a.e. in $[0,T]$ and finally $\dot{\boldsymbol{w}}^*(t) \in K$. Since $\boldsymbol{w}_\epsilon \in L^\infty([0,T],Z) \subset L^2([0,T],Z)$ and $\dot{\boldsymbol{w}}_\epsilon \in L^2([0,T],Z)$, the solution $\boldsymbol{w}_\epsilon \in H^1([0,T],Z)$. Using the injection $H^1([0,T],Z) \hookrightarrow \mathcal{C}^0([0,T],Z)$ and the bounds on $\boldsymbol{w}_\epsilon$, we get $\boldsymbol{w}_\epsilon \rightharpoonup \boldsymbol{w}^*$ in $\mathcal{C}^0([0,T],Z)$ and $\boldsymbol{w}_\epsilon(T)$ being well defined. We then use $\boldsymbol{w}_\epsilon(0) = \boldsymbol{0}$ to obtain

$$\frac{1}{2} a(\boldsymbol{w}_\epsilon(T), \boldsymbol{w}_\epsilon(T)) = \frac{1}{2}\int_0^T \frac{d}{dt} a(\boldsymbol{w}_\epsilon, \boldsymbol{w}_\epsilon)dt = \int_0^T a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_\epsilon)dt.$$

The bilinear form $w \mapsto a(w,w)(T)$ is convex, proper and lower semi-continuous on $\mathcal{C}^0([0,T],Z) \to \mathbb{R}$, thus weakly lower semi-continuous in $\mathcal{C}^0([0,T],Z)$. This allows us to pass to the limit $\epsilon \longrightarrow 0$ in the above and obtain

$$\int_0^T a(\boldsymbol{w}^*, \dot{\boldsymbol{w}}^*)dt = \frac{1}{2} a(\boldsymbol{w}^*(T), \boldsymbol{w}^*(T)) \leq \liminf_{\epsilon\longrightarrow 0} \frac{1}{2} a(\boldsymbol{w}_\epsilon(T), \boldsymbol{w}_\epsilon(T)) = \liminf_{\epsilon\longrightarrow 0} \int_0^T a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_\epsilon)dt. \tag{4.43}$$

Now we consider (4.39) for $\boldsymbol{z} \in K$,

$$\begin{aligned}
0 &\leq \int_0^T \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) + j(\boldsymbol{z}) - j_\epsilon(\dot{\boldsymbol{w}}_\epsilon) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon)\right) dt \\
\int_0^T j_1(\dot{\boldsymbol{w}}_\epsilon)dt \leq \int_0^T j_\epsilon(\dot{\boldsymbol{w}}_\epsilon)dt &\leq \int_0^T \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) + j(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon)\right) dt \\
\int_0^T j_1(\dot{\boldsymbol{w}}^*)dt \leq \liminf_{\epsilon\to 0} \left(\int_0^T j_1(\dot{\boldsymbol{w}}_\epsilon)dt\right) &\leq \lim_{\epsilon\to 0} \int_0^T \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) + j(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon)\right) dt \quad \text{(using (9))} \\
\int_0^T j_1(\dot{\boldsymbol{w}}^*)dt &\leq \int_0^T \left(a(\boldsymbol{w}^*, \boldsymbol{z} - \dot{\boldsymbol{w}}^*) + j(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}^*)\right) dt \quad \text{(using (4.43) and } \dot{\boldsymbol{w}}_\epsilon \rightharpoonup \dot{\boldsymbol{w}}^*).
\end{aligned}$$

Since $\dot{\boldsymbol{w}}^*(t) \in K$, we have $j_1(\dot{\boldsymbol{w}}^*) = j(\dot{\boldsymbol{w}}^*)$ and

$$\int_0^T \left(a(\boldsymbol{w}^*, \boldsymbol{z} - \dot{\boldsymbol{w}}^*) + j(\boldsymbol{z}) - j(\dot{\boldsymbol{w}}^*) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}^*)\right) dt \geq 0 \qquad \forall \boldsymbol{z} \in K.$$

By a standard procedure of passing to the pointwise inequality [69], we find from the above inequality that $\boldsymbol{w}^*$ satisfies (4.35). Since the solution to (4.35) is unique, we have that the solution $\boldsymbol{w}^*(t) = \boldsymbol{w}(t)$ and the entire sequence converges.

Next we prove that this sequence $\boldsymbol{w}_\epsilon$ actually converges strongly to $\boldsymbol{w}$ in $L^\infty([0,T], Z)$ and not merely weakly-star. For any time $t_0 \in [0,T]$, we have

$$\begin{aligned}
\frac{1}{2}\left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_Z^2(t_0) + \int_0^{t_0} j_1(\dot{\boldsymbol{w}}_\epsilon) dt &\leq \int_0^{t_0} \left(\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_Z^2 + j_\epsilon(\dot{\boldsymbol{w}}_\epsilon)\right) dt \\
&= \int_0^{t_0} \left(a(\boldsymbol{w} - \boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}} - \dot{\boldsymbol{w}}_\epsilon) + j_\epsilon(\dot{\boldsymbol{w}}_\epsilon)\right) dt \\
&= \int_0^{t_0} \left(a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_\epsilon) + j_\epsilon(\dot{\boldsymbol{w}}_\epsilon) - a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}) - a(\boldsymbol{w}, \dot{\boldsymbol{w}}_\epsilon) + a(\boldsymbol{w}, \dot{\boldsymbol{w}})\right) dt \\
&\leq \int_0^{t_0} \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z}) + j_\epsilon(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) - a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}) - a(\boldsymbol{w}, \dot{\boldsymbol{w}}_\epsilon) + a(\boldsymbol{w}, \dot{\boldsymbol{w}})\right) dt \\
&= \int_0^{t_0} \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z} - \dot{\boldsymbol{w}}) + j_\epsilon(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) - a(\boldsymbol{w}, \dot{\boldsymbol{w}}_\epsilon) + a(\boldsymbol{w}, \dot{\boldsymbol{w}})\right) dt \qquad \forall \boldsymbol{z}(t) \in Z.
\end{aligned}$$

Choosing $\boldsymbol{z}(t) \in K \subset Z$ yields the bound

$$\frac{1}{2}\left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_Z^2(t_0) + \int_0^{t_0} j_1(\dot{\boldsymbol{w}}_\epsilon) dt \leq \int_0^{t_0} \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z} - \dot{\boldsymbol{w}}) + j(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) - a(\boldsymbol{w}, \dot{\boldsymbol{w}}_\epsilon) + a(\boldsymbol{w}, \dot{\boldsymbol{w}})\right) dt.$$

Passing to the limit $\epsilon \to 0$, with $\boldsymbol{w}_\epsilon \rightharpoonup \boldsymbol{w}$ and $\boldsymbol{w} \in K$, Lemma 9 leads to

$$\begin{aligned}
\int_0^{t_0} j(\dot{\boldsymbol{w}}) dt = \int_0^{t_0} j_1(\dot{\boldsymbol{w}}) dt &\leq \liminf_{\epsilon \to 0} \int_0^{t_0} j_1(\dot{\boldsymbol{w}}_\epsilon) dt \\
&\leq \lim_{\epsilon \to 0} \left(\int_0^{t_0} \left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z} - \dot{\boldsymbol{w}}) + j(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon) - a(\boldsymbol{w}, \dot{\boldsymbol{w}}_\epsilon) + a(\boldsymbol{w}, \dot{\boldsymbol{w}})\right) dt - \frac{1}{2}\left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_Z^2(t_0)\right) \\
&\leq \int_0^{t_0} \left(a(\boldsymbol{w}, \boldsymbol{z} - \dot{\boldsymbol{w}}) + j(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}})\right) dt - \frac{1}{2}\lim_{\epsilon \to 0}\left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_Z^2(t_0).
\end{aligned}$$

Therefore,

$$\lim_{\epsilon \to 0} \frac{1}{2}\left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_Z^2(t_0) \leq \int_0^{t_0} \left(a(\boldsymbol{w}, \boldsymbol{z} - \dot{\boldsymbol{w}}) + j(\boldsymbol{z}) - j(\dot{\boldsymbol{w}}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}})\right) dt.$$

Taking $\boldsymbol{z} = \dot{\boldsymbol{w}}$ shows that the above limit is zero, establishing the strong convergence $\boldsymbol{w}_\epsilon \longrightarrow \boldsymbol{w}$ in $L^\infty([0,T], Z)$. □

4.2.2 Penalization and regularization

The nonlinearity $j(\boldsymbol{z})$ being unbounded for $\boldsymbol{z} \notin K$, (4.35) is not differentiable with respect to parameters like the shape of the domain [133, 171]. On the contrary, $j_\epsilon(\boldsymbol{z})$ is now bounded on the full space Z, so one should be able to differentiate the penalized formulation (4.39). However $j_\epsilon(\boldsymbol{z})$ is still non-smooth because of the maximum operator and the norm of the plastic tensor. We therefore need to regularize the nonlinearity $j_\epsilon(\cdot)$.

We introduce a regularization constant $0 < \eta \ll 1$. The expression in (4.38) has two kinds of non-smoothness: $\max(\cdot)$ and $|\cdot|$ (the Euclidean norm), we regularize each with operators $M_\eta : L^2(\Omega) \to L^2(\Omega)$ and $N_\eta : Q \to L^2(\Omega)$ respectively, defined as

$$M_\eta(\gamma) = \frac{1}{2}\left(\gamma + \sqrt{\gamma^2 + \left(\frac{\sigma_Y \eta}{TE}\right)^2}\right), \quad N_\eta(\boldsymbol{\varepsilon}_p) = \sqrt{\boldsymbol{\varepsilon}_p : \boldsymbol{\varepsilon}_p + \left(\frac{\sigma_Y \eta}{TE}\right)^2},$$

where T is the final time, σ_Y is the yield strength and E is the Young's modulus. In the above, the factor η is multiplied by $\frac{\sigma_Y}{TE}$ so as to ensure that the regularization is coherent with the order of the solution $\dot{\boldsymbol{\varepsilon}}_p$. For the ease of numerical implementation, we choose a globally smooth regularization rather than a piecewise regularization. We now smoothen the function (4.38) as

$$\mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}) = \sqrt{\frac{2}{3}}\sigma_Y \left(N_\eta\left(\dot{\boldsymbol{\varepsilon}}_p\right) + \frac{1}{\epsilon} M_\eta \left(\sqrt{\frac{2}{3}} N_\eta(\dot{\boldsymbol{\varepsilon}}_p) - \dot{\gamma} \right) \right), \tag{4.44}$$

and define $j_{\epsilon,\eta} : Z \to \mathbb{R}$ in the same manner as before,

$$j_{\epsilon,\eta}(\dot{\boldsymbol{w}}) = \int_\Omega \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}) \;\; dx.$$

Lemma 10. *The function $j_{\epsilon,\eta}(\cdot)$ is convex, lower semi-continuous and satisfies*

$$|j_{\epsilon,\eta}(\boldsymbol{z}) - j_\epsilon(\boldsymbol{z})| \leq C\eta \left\|\boldsymbol{z}\right\|_Z \qquad \text{and} \qquad j_\epsilon(\boldsymbol{z}) < j_{\epsilon,\eta}(\boldsymbol{z}) \qquad \forall \eta > 0,\ \boldsymbol{z} \in Z, \tag{4.45}$$

where C is a constant independent of η.

The proof to the above lemma being straight-forward, is left to the reader. We consider a new problem: find $\boldsymbol{w}_{\epsilon,\eta} \in Z$ such that $\boldsymbol{w}_{\epsilon,\eta}(0) = \boldsymbol{0}$, $\dot{\boldsymbol{w}}_{\epsilon,\eta}(t) \in Z$ and

$$a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta}) + j_{\epsilon,\eta}(\boldsymbol{z}) - j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}) \geq l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta}) \qquad \forall \boldsymbol{z} \in Z. \tag{4.46}$$

The above inequation is shown to be well-posed in the following theorem. The proof is inspired from that of Theorem 4.3 in [87]. One cannot apply directly Theorem 14 because the functional $j_{\epsilon,\eta}$ is not positively homogeneous.

Theorem 17. *The variational inequality (4.46) admits a unique solution $\boldsymbol{w}_{\epsilon,\eta} \in H^1([0,T], Z)$.*

Proof. Since we cannot apply directly Theorem 14 (or Theorem 4.3 in [87]), we modify it to adapt to our case where $j_{\epsilon,\eta}$ positively homogeneous. Again, the time interval $[0,T]$ is discretized in $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ with $t_n - t_{n-1} = \delta t$, sufficiently small, given by (4.51). For $l_t \in Z^*$ (defined in (4.32)), define $l_n = l_{t_n}$, and $\Delta l_n = (l_n - l_{n-1}) \in Z^*$. For any $\delta t \in \mathbb{R}^+$ define $j^{\delta t}_{\epsilon,\eta} : Z \to \mathbb{R}$ as

$$j^{\delta t}_{\epsilon,\eta}(\boldsymbol{z}) = \delta t j_{\epsilon,\eta}\left(\frac{\boldsymbol{z}}{\delta t}\right). \tag{4.47}$$

Since $j_{\epsilon,\eta}(\boldsymbol{z})$ is convex, $j^{\delta t}_{\epsilon,\eta}(\boldsymbol{z})$ is convex. Starting from $\boldsymbol{w}_0 = \boldsymbol{0}$ we construct a sequence $\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \Delta\boldsymbol{w}_n$ and a function $\boldsymbol{w}^{\delta t}_{\epsilon,\eta} \in \mathcal{C}^0([0,T], K)$ such that $\boldsymbol{w}^{\delta t}_{\epsilon,\eta}(t) = \boldsymbol{w}_n + (t - t_n)/\delta t \Delta\boldsymbol{w}_n$ for $t \in [t_n, t_{n+1}]$, where $\Delta\boldsymbol{w}_n \in Z$ is the solution of

$$a(\Delta\boldsymbol{w}_n, \boldsymbol{z} - \Delta\boldsymbol{w}_n) + j^{\delta t}_{\epsilon,\eta}(\boldsymbol{z}) - j^{\delta t}_{\epsilon,\eta}(\Delta\boldsymbol{w}_n) \geq l_n(\boldsymbol{z} - \Delta\boldsymbol{w}_n) - a(\boldsymbol{w}_{n-1}, \boldsymbol{z} - \Delta\boldsymbol{w}_n) \qquad \forall \boldsymbol{z} \in Z \tag{4.48}$$

Since $a(\boldsymbol{w}_{n-1}, \cdot)$ defines a continuous linear form on Z, the right hand side of the above is a continuous linear form on Z. With $j^{\delta t}_{\epsilon,\eta}(\cdot)$ being convex and lower semi continuous, (4.48) is indeed a variational inequation of the 2nd kind and admits a unique solution $\Delta\boldsymbol{w}_n \in Z$ [80]. In order to obtain a bound on $\Delta\boldsymbol{w}_n$, we substitute $\boldsymbol{z} = \boldsymbol{0}$ in (4.48) to obtain

$$a(\Delta\boldsymbol{w}_n, \Delta\boldsymbol{w}_n) \leq l_n(\Delta\boldsymbol{w}_n) - a(\boldsymbol{w}_{n-1}, \Delta\boldsymbol{w}_n) - j^{\delta t}_{\epsilon,\eta}(\Delta\boldsymbol{w}_n) + j^{\delta t}_{\epsilon,\eta}(\boldsymbol{0}). \tag{4.49}$$

Using definitions (4.44) and (4.47), we get

$$j^{\delta t}_{\epsilon,\eta}(\boldsymbol{0}) = \delta t j_{\epsilon,\eta}(\boldsymbol{0}) = \sqrt{\frac{2}{3}} \frac{\sigma_Y^2 \delta t \eta}{TE} \left(1 + \frac{1}{2\epsilon}(1 + \sqrt{2}) \right) \leq C_Y \delta t \eta \left(1 + \frac{k}{\epsilon} \right)$$

where $C_Y = \sqrt{2/3}\sigma_Y^2/(TE)$ and $k = (1 + \sqrt{2})/2$. Then substitute $\boldsymbol{z} = \Delta\boldsymbol{w}_n + \Delta\boldsymbol{w}_{n-1}$ in (4.48), written for index $n-1$, to get

$$0 \leq j^{\delta t}_{\epsilon,\eta}(\Delta\boldsymbol{w}_n + \Delta\boldsymbol{w}_{n-1}) - j^{\delta t}_{\epsilon,\eta}(\Delta\boldsymbol{w}_{n-1}) - l_{n-1}(\Delta\boldsymbol{w}_n) + a(\boldsymbol{w}_{n-1}, \Delta\boldsymbol{w}_{n-1}). \tag{4.50}$$

Adding (4.49) and (4.50), and using the convexity of $j^{\delta t}_{\epsilon,\eta}(\cdot)$, leads to

$$\begin{aligned}
a(\Delta \boldsymbol{w}_n, \Delta \boldsymbol{w}_n) - C_Y \delta t \eta \left(1+\frac{k}{\epsilon}\right) &\leq l_n(\Delta \boldsymbol{w}_n) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_n) + j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_n + \Delta \boldsymbol{w}_{n-1}) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_{n-1}) - l_{n-1}(\Delta \boldsymbol{w}_n) \\
&\leq \Delta l_n(\Delta \boldsymbol{w}_n) + \frac{1}{2}\left(j^{\delta t}_{\epsilon,\eta}(2\Delta \boldsymbol{w}_n) + j^{\delta t}_{\epsilon,\eta}(2\Delta \boldsymbol{w}_{n-1})\right) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_n) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_{n-1}) \\
&= \Delta l_n(\Delta \boldsymbol{w}_n) + \left(j^{\delta t}_{\epsilon,2\eta}(\Delta \boldsymbol{w}_n) + j^{\delta t}_{\epsilon,2\eta}(\Delta \boldsymbol{w}_{n-1})\right) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_n) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_{n-1}) \\
&\leq \Delta l_n(\Delta \boldsymbol{w}_n) + \left(j^{\delta t}_{\epsilon,2\eta}(\Delta \boldsymbol{w}_n) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_n)\right) + \left(j^{\delta t}_{\epsilon,2\eta}(\Delta \boldsymbol{w}_{n-1}) - j^{\delta t}_{\epsilon,\eta}(\Delta \boldsymbol{w}_{n-1})\right).
\end{aligned}$$

Using the fact that

$$|j^{\delta t}_{\epsilon,2\eta}(\boldsymbol{w}) - j^{\delta t}_{\epsilon,\eta}(\boldsymbol{w})| \leq C_Y \delta t \eta \left(1+\frac{k}{\epsilon}\right) \qquad \forall\, \boldsymbol{w} \in Z,$$

and that $\alpha \left\|\boldsymbol{w}_n\right\|^2_Z \leq a(\boldsymbol{w}_n, \boldsymbol{w}_n)$, we obtain the bound

$$\begin{aligned}
\alpha \left\|\Delta \boldsymbol{w}_n\right\|^2_Z &\leq \Delta l_n(\Delta \boldsymbol{w}_n) + 3 C_Y \delta t \eta \left(1+\frac{k}{\epsilon}\right) \\
&\leq \left\|\Delta l_n\right\|_{Z^*} \left\|\Delta \boldsymbol{w}_n\right\|_Z + 3 C_Y \delta t \eta \left(1+\frac{k}{\epsilon}\right) \\
&\leq \frac{1}{2\alpha}\left\|\Delta l_n\right\|^2_{Z^*} + \frac{\alpha}{2}\left\|\Delta \boldsymbol{w}_n\right\|^2 + 3 C_Y \delta t \eta \left(1+\frac{k}{\epsilon}\right)
\end{aligned}$$

from which, choosing a time step δt sufficiently small, compared to η, ϵ, such that

$$\delta t < \frac{\epsilon \left\|\Delta l_n\right\|^2_{Z^*}}{6 C_Y \eta \left(\epsilon + k\right)}, \tag{4.51}$$

we deduce

$$\left\|\Delta \boldsymbol{w}_n\right\|_Z \leq \frac{\sqrt{2}}{\alpha}\left\|\Delta l_n\right\|_{Z^*}.$$

Summing these bounds over the index n leads to

$$\sum_{n=1}^{N} \left\|\Delta \boldsymbol{w}_n\right\|^2_Z \leq \frac{2\delta t^2}{\alpha^2} \sum_{n=1}^{N} \left\|\Delta l_n/\delta t\right\|^2_{Z^*} \quad \text{and} \quad \max_{1\leq n\leq N} \left\|\boldsymbol{w}_n\right\|_Z \leq \frac{\sqrt{2}}{\alpha}\delta t \sum_{n=1}^{N} \left\|\Delta l_n/\delta t\right\|_{Z^*}. \tag{4.52}$$

Now it remains to prove that the interpolation $\boldsymbol{w}^{\delta t}_{\epsilon,\eta}$ of the discrete solutions $\boldsymbol{w}_n$ admits a limit $\boldsymbol{w}_{\epsilon,\eta}$, solution to (4.46), as δt goes to 0. In (4.48) we replace $\boldsymbol{z}$ by $\delta t \boldsymbol{z}$, define $\delta \boldsymbol{w}_n = \Delta \boldsymbol{w}_n/\delta t$ and divide by δt to arrive at

$$a(\boldsymbol{w}_n, \boldsymbol{z} - \delta \boldsymbol{w}_n) + j_{\epsilon,\eta}(\boldsymbol{z}) - j_{\epsilon,\eta}(\delta \boldsymbol{w}_n) \geq l_n(\boldsymbol{z} - \delta \boldsymbol{w}_n) \qquad \forall \boldsymbol{z} \in Z. \tag{4.53}$$

Starting from (4.53), one can follow exactly the same steps as in the proof of Theorem 14 in [87] and arrive at

$$\begin{aligned}
0 \leq &\int_0^T \left(a(\boldsymbol{w}^{\delta t}_{\epsilon,\eta}, \boldsymbol{z} - \dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta}) + j_{\epsilon,\eta}(\boldsymbol{z}) - j_{\epsilon,\eta}(\dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta}) - l_{\delta t}(\boldsymbol{z} - \dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta})\right) dt \\
&- \frac{\delta t}{2} j_{\epsilon,\eta}(\boldsymbol{z}_1) + C\delta t \int_0^T \left\|\dot{l}_{\delta t}\right\|^2_{Z^*} dt + C_Y \delta t \eta \left(1+\frac{k}{\epsilon}\right),
\end{aligned} \tag{4.54}$$

where $l_{\delta t}(t) = l_{n-1} + \Delta l_n \left(t - t_{n-1}\right)/\delta t$ is the interpolation of the source term for $t \in [t_{n-1}, t_n]$. Substituting $\dot{l}_{\delta t}$ to $\Delta l_n/\delta t$ and $\dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta}$ to $\Delta \boldsymbol{w}_n$ in (4.52) leads to

$$\left\|\dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta}\right\|_{L^2([0,T],Z)} \leq C_1 \qquad \text{and} \qquad \left\|\boldsymbol{w}^{\delta t}_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} \leq C_2. \tag{4.55}$$

These uniform bounds imply the existence of a subsequence $\delta t \longrightarrow 0$ such that

$$\boldsymbol{w}^{\delta t}_{\epsilon,\eta} \overset{*}{\rightharpoonup} \boldsymbol{w}_{\epsilon,\eta} \quad \text{in } L^\infty([0,T], Z) \qquad \text{and} \qquad \dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta} \rightharpoonup \dot{\boldsymbol{w}}_{\epsilon,\eta} \quad \text{in } L^2([0,T], Z).$$

Again, we have that $\dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta} \rightharpoonup \dot{\boldsymbol{w}}_{\epsilon,\eta}$ in $L^1([0,T], Z)$. Since $j_{\epsilon,\eta}(\cdot)$ is convex and weakly lower semi continuous, we can pass to the limit in (4.54) to get

$$0 \leq \int_0^T \left(a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta}) + j_{\epsilon,\eta}(\boldsymbol{z}) - j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta})\right) dt.$$

Consequently, $\boldsymbol{w}_{\epsilon,\eta}$ is a solution of problem (4.46). Uniqueness of the solution $\boldsymbol{w}_{\epsilon,\eta}$ of (4.46) is classical. □

We now convert the variational inequation (4.46) into an equation. Since the function $\mathcal{D}_{\epsilon,\eta}$ is smooth, we can define its gradient

$$\nabla_z \mathcal{D}_{\epsilon,\eta}(\boldsymbol{w}) = \left(\frac{\partial \mathcal{D}_{\epsilon,\eta}(\boldsymbol{w})}{\partial \boldsymbol{u}}, \frac{\partial \mathcal{D}_{\epsilon,\eta}(\boldsymbol{w})}{\partial \boldsymbol{\varepsilon}_p}, \frac{\partial \mathcal{D}_{\epsilon,\eta}(\boldsymbol{w})}{\partial \gamma} \right).$$

Lemma 11. *The variational inequality (4.46) is equivalent to the variational formulation: find $\boldsymbol{w}_{\epsilon,\eta}(t) \in Z$ such that $\boldsymbol{w}_{\epsilon,\eta}(0) = \mathbf{0}$, $\dot{\boldsymbol{w}}_{\epsilon,\eta}(t) \in Z$ and*

$$a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z}) + \langle \nabla_z \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}), \boldsymbol{z} \rangle = l_t(\boldsymbol{z}) \qquad \forall \boldsymbol{z} \in Z, \tag{4.56}$$

where $\langle , \rangle$ is the scalar product defined by (4.30).

Proof. By definition of the convexity of $j_{\epsilon,\eta}$ we get

$$j_{\epsilon,\eta}(\boldsymbol{z}) - j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}) \geq \langle \nabla_z \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}), \boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta} \rangle \quad \forall \boldsymbol{z} \in Z.$$

The right hand side in the above is the tangent hyperplane to $j_{\epsilon,\eta}$ at $\boldsymbol{z} = \dot{\boldsymbol{w}}_{\epsilon,\eta}$. On the other hand, (4.46) can be written as

$$j_{\epsilon,\eta}(\boldsymbol{z}) - j_{\epsilon,\eta}(\dot{\boldsymbol{w}}) \geq a(\boldsymbol{w}_{\epsilon,\eta}, \dot{\boldsymbol{w}}_{\epsilon,\eta} - \boldsymbol{z}) + l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta}) \quad \forall \boldsymbol{z} \in Z.$$

Again, the right hand side in the above above is affine in $\boldsymbol{z}$ and it vanishes at $\boldsymbol{z} = \dot{\boldsymbol{w}}_{\epsilon,\eta}$, implying that it is also tangent at $\boldsymbol{z} = \dot{\boldsymbol{w}}_{\epsilon,\eta}$. Since $j_{\epsilon,\eta}$ is smooth, the two tangent hyperplanes must be equal

$$a(\boldsymbol{w}_{\epsilon,\eta}, \dot{\boldsymbol{w}}_{\epsilon,\eta} - \boldsymbol{z}) + l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta}) = \langle \nabla_z \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}), \boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta} \rangle \qquad \forall \boldsymbol{z} \in Z.$$

Replacing $\boldsymbol{z}$ in the above by $\dot{\boldsymbol{w}}_{\epsilon,\eta} + \boldsymbol{z} \in Z$, we deduce (4.56). □

Equation (4.56) is our approximation of the plasticity problem (4.35) we treat for the shape derivative in Section 4.3. We call it the state equation and its solution, the state solution. As expected, for a fixed ϵ, one can prove the convergence of the sequence $\boldsymbol{w}_{\epsilon,\eta}$ of solutions to (4.46) to the solution $\boldsymbol{w}_\epsilon$ to (4.39) as $\eta \longrightarrow 0$. We content ourselves in proving a weak convergence.

Theorem 18. *The sequence of solutions $\boldsymbol{w}_{\epsilon,\eta}$ to (4.46) satisfies*

$$\eta \longrightarrow 0, \qquad \boldsymbol{w}_{\epsilon,\eta} \overset{*}{\rightharpoonup} \boldsymbol{w}_\epsilon \quad \text{in } L^\infty([0,T], Z) \quad \text{and} \quad \dot{\boldsymbol{w}}_{\epsilon,\eta} \rightharpoonup \dot{\boldsymbol{w}}_\epsilon \quad \text{in } L^2([0,T], Z),$$

where $\boldsymbol{w}_\epsilon$ is the solution to (4.39).

Proof. In the proof of Theorem 17 the bounds (4.55) imply

$$\|\boldsymbol{w}_{\epsilon,\eta}\|_{L^\infty([0,T],Z)} \leq \liminf_{\delta t \to 0} \left\|\boldsymbol{w}^{\delta t}_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} \leq C_1 \text{ and } \|\dot{\boldsymbol{w}}_{\epsilon,\eta}\|_{L^2([0,T],Z)} \leq \liminf_{\delta t \to 0} \left\|\dot{\boldsymbol{w}}^{\delta t}_{\epsilon,\eta}\right\|_{L^2([0,T],Z)} \leq C_2,$$

with the constants C_1 and C_2 being independent of $\epsilon, \eta, \delta t$ (provided that δt is sufficiently small). Thus one can extract a subsequence $\boldsymbol{w}_{\epsilon,\eta}$ and there exists a limit $\boldsymbol{w}^*_\epsilon$ such that, as $\eta \longrightarrow 0$,

$$\boldsymbol{w}_{\epsilon,\eta} \overset{*}{\rightharpoonup} \boldsymbol{w}^*_\epsilon \quad \text{in } L^\infty([0,T], Z) \qquad \text{and} \qquad \dot{\boldsymbol{w}}_{\epsilon,\eta} \rightharpoonup \dot{\boldsymbol{w}}^*_\epsilon \quad \text{in } L^2([0,T], Z).$$

Using the above two weak convergences, the same arguments as for (4.43) leads to

$$\int_0^T a(\boldsymbol{w}^*_\epsilon, \dot{\boldsymbol{w}}^*_\epsilon) dt = \frac{1}{2} a(\boldsymbol{w}^*_\epsilon(T), \boldsymbol{w}^*_\epsilon(T)) \leq \liminf_{\epsilon \longrightarrow 0} \frac{1}{2} a(\boldsymbol{w}_{\epsilon,\eta}(T), \boldsymbol{w}_{\epsilon,\eta}(T)) = \liminf_{\epsilon \longrightarrow 0} \int_0^T a(\boldsymbol{w}_{\epsilon,\eta}, \dot{\boldsymbol{w}}_{\epsilon,\eta}) dt. \tag{4.57}$$

In order to establish the equality between $\boldsymbol{w}^*_\epsilon$ and $\boldsymbol{w}_\epsilon$, we re-write (4.46) as

$$\int_0^T \left(a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta}) + j_{\epsilon,\eta}(\boldsymbol{z}) - l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_{\epsilon,\eta}) \right) dt \geq \int_0^T j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}) dt \geq \int_0^T j_\epsilon(\dot{\boldsymbol{w}}_{\epsilon,\eta}) dt,$$

where we use Lemma 10 in the last inequality. The first inequality in (4.45) implies the convergence $j_{\epsilon,\eta}(\boldsymbol{z}) \longrightarrow j_\epsilon(\boldsymbol{z})$ as $\eta \longrightarrow 0$ for all $\boldsymbol{z} \in Z$. Letting η go to 0, the weak convergence $\boldsymbol{w}_{\epsilon,\eta} \rightharpoonup \boldsymbol{w}_\epsilon^*$ and (4.45) leads to

$$\begin{aligned}\int_0^T j_\epsilon(\dot{\boldsymbol{w}}_\epsilon^*)dt &\leq \liminf_{\eta\to 0}\left(\int_0^T j_\epsilon(\dot{\boldsymbol{w}}_{\epsilon,\eta})dt\right) \qquad \text{(by Lemma 9)}\\ &\leq \lim_{\eta\to 0}\left(\int_0^T \left(a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z}-\dot{\boldsymbol{w}}_{\epsilon,\eta}) + j_{\epsilon,\eta}(\boldsymbol{z}) - l_t(\boldsymbol{z}-\dot{\boldsymbol{w}}_{\epsilon,\eta})\right) dt\right)\\ &= \int_0^T \left(a(\boldsymbol{w}_\epsilon^*, \boldsymbol{z}-\dot{\boldsymbol{w}}_\epsilon^*) + j_\epsilon(\boldsymbol{z}) - l_t(\boldsymbol{z}-\dot{\boldsymbol{w}}_\epsilon^*)\right) dt \qquad \text{(using (4.57))}.\end{aligned}$$

Since the solution to (4.39) is unique, we deduce $\boldsymbol{w}_\epsilon^* = \boldsymbol{w}_\epsilon$ and the entire sequence converges. □

It remains to show the strong convergence $\boldsymbol{w}_{\epsilon,\eta} \longrightarrow \boldsymbol{w}$ in $L^\infty([0,T],Z)$ as $\epsilon \longrightarrow 0$ and $\eta \longrightarrow 0$.

Theorem 19. *As $\epsilon \longrightarrow 0$ and $\eta \longrightarrow 0$, the sequence of solutions $\boldsymbol{w}_{\epsilon,\eta} \longrightarrow \boldsymbol{w}$ in $L^\infty([0,T],Z)$.*

Proof. As was performed in the proof of Theorem 16, for a time instant $t_0 \in [0,T]$

$$\begin{aligned}&\frac{1}{2}\left\|\boldsymbol{w}_\epsilon - \boldsymbol{w}_{\epsilon,\eta}\right\|_Z^2(t_0) + \int_0^{t_0} j_\epsilon(\dot{\boldsymbol{w}}_{\epsilon,\eta})dt\\ &\qquad\leq \int_0^{t_0}\left(a(\boldsymbol{w}_\epsilon - \boldsymbol{w}_{\epsilon,\eta}, \dot{\boldsymbol{w}}_\epsilon - \dot{\boldsymbol{w}}_{\epsilon,\eta}) + j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta})\right)dt\\ &\qquad= \int_0^{t_0}\left(a(\boldsymbol{w}_{\epsilon,\eta}, \dot{\boldsymbol{w}}_{\epsilon,\eta}) + j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}) - a(\boldsymbol{w}_{\epsilon,\eta}, \dot{\boldsymbol{w}}_\epsilon) - a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_{\epsilon,\eta}) + a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_\epsilon)\right)dt\\ &\qquad\leq \int_0^{t_0}\left(a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z}) + j_{\epsilon,\eta}(\boldsymbol{z}) - \langle l, \boldsymbol{z}-\dot{\boldsymbol{w}}_{\epsilon,\eta}\rangle - a(\boldsymbol{w}_{\epsilon,\eta}, \dot{\boldsymbol{w}}_\epsilon) - a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_{\epsilon,\eta}) + a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_\epsilon)\right)dt\\ &\qquad= \int_0^{t_0}\left(a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z}-\dot{\boldsymbol{w}}_\epsilon) + j_{\epsilon,\eta}(\boldsymbol{z}) - \langle l, \boldsymbol{z}-\dot{\boldsymbol{w}}_{\epsilon,\eta}\rangle - a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_{\epsilon,\eta}) + a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_\epsilon)\right)dt \qquad \forall \boldsymbol{z} \in Z.\end{aligned}$$

Passing to the limit as $\eta \to 0$, we get

$$\begin{aligned}&\liminf_{\eta\to 0}\left(\frac{1}{2}\left\|\boldsymbol{w}_\epsilon - \boldsymbol{w}_{\epsilon,\eta}\right\|_Z^2(t_0)\right) + \int_0^{t_0} j_\epsilon(\dot{\boldsymbol{w}}_\epsilon)dt\\ &\qquad\leq \liminf_{\eta\to 0}\left(\frac{1}{2}\left\|\boldsymbol{w}_\epsilon - \boldsymbol{w}_{\epsilon,\eta}\right\|_Z^2(t_0) + \int_0^{t_0} j_\epsilon(\dot{\boldsymbol{w}}_{\epsilon,\eta})dt\right)\\ &\qquad\leq \lim_{\eta\to 0}\int_0^{t_0}\left(a(\boldsymbol{w}_{\epsilon,\eta}, \boldsymbol{z}-\dot{\boldsymbol{w}}_\epsilon) + j_{\epsilon,\eta}(\boldsymbol{z}) - \langle l, \boldsymbol{z}-\dot{\boldsymbol{w}}_{\epsilon,\eta}\rangle - a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_{\epsilon,\eta}) + a(\boldsymbol{w}_\epsilon, \dot{\boldsymbol{w}}_\epsilon)\right)dt\\ &\qquad= \int_0^{t_0}\left(a(\boldsymbol{w}_\epsilon, \boldsymbol{z}-\dot{\boldsymbol{w}}_\epsilon) + j_\epsilon(\boldsymbol{z}) - \langle l, \boldsymbol{z}-\dot{\boldsymbol{w}}_\epsilon\rangle\right)dt.\end{aligned}$$

Taking $\boldsymbol{z} = \dot{\boldsymbol{w}}_\epsilon$ we get

$$\liminf_{\eta\to 0}\left(\frac{1}{2}\left\|\boldsymbol{w}_\epsilon - \boldsymbol{w}_{\epsilon,\eta}\right\|_Z^2(t_0)\right) = 0$$

establishing $\boldsymbol{w}_{\epsilon,\eta} \longrightarrow \boldsymbol{w}_\epsilon$ in $L^\infty([0,T],Z)$. Finally, by a simple application of the triangular inequality, we get

$$\left\|\boldsymbol{w} - \boldsymbol{w}_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} \leq \left\|\boldsymbol{w}_\epsilon - \boldsymbol{w}_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} + \left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_{L^\infty([0,T],Z)}$$

which at the limit $\epsilon \longrightarrow 0$ and $\eta \longrightarrow 0$ gives

$$\lim_{\epsilon\to 0,\eta\to 0}\left\|\boldsymbol{w} - \boldsymbol{w}_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} \leq \lim_{\epsilon\to 0,\eta\to 0}\left(\left\|\boldsymbol{w}_\epsilon - \boldsymbol{w}_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} + \left\|\boldsymbol{w} - \boldsymbol{w}_\epsilon\right\|_{L^\infty([0,T],Z)}\right) = 0.$$

□

4.3 Shape derivative for Primal form

In this section, to simplify the notations, we drop the indices ϵ and η, and simply write $\boldsymbol{w}$ instead of $\boldsymbol{w}_{\epsilon,\eta}$. We minimize an objective function $J(\Omega)$ defined as

$$J(\Omega) = \int_0^T \left(\int_\Omega m(\boldsymbol{w}(\Omega))\ dx + \int_{\Gamma_N} p(\boldsymbol{w}(\Omega)) ds \right) dt, \tag{4.58}$$

where $\boldsymbol{w}(\Omega)$ is solution to the state equation (4.56) and the integrands $m(\cdot)$ and $p(\cdot)$ are assumed to be smooth functions at least of class $\mathcal{C}^1$. In addition we assume a quadratic growth condition on $m(\cdot)$ and $p(\cdot)$ such that the objective function is well-defined and the adjoint equation (4.62) is well-posed. This objective can represent a mechanical property such as the total compliance, total power, elastic energy, plastic energy, etc. as well as a geometric property such as the volume. An industrially relevant objective is the total compliance, given by

$$J(\Omega) = \int_0^T \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{u}(\Omega)\ \ ds\ dt. \tag{4.59}$$

In practice, the shape Ω is designed inside a pre-fixed design space $D \subset \mathbb{R}^d$. As shown in Fig.2.4, the blue region represents the shape Ω, and the blue and grey areas represent the design space D. We define the space of admissible shapes $\mathcal{U}_{ad}$ as

$$\mathcal{U}_{ad} = \left\{ \Omega \subset D,\ \int_\Omega dx = V_f \right\}, \tag{4.60}$$

where Ω is an open set and V_f is a target volume. The optimization problem then reads

$$\min_{\Omega \in \mathcal{U}_{ad}} J(\Omega).$$

The question of existence of optimal shapes Ω is a delicate one and we shall not dwell into it (see [93] for a discussion). Rather, we content ourselves with computing numerical minimizers, using a gradient-descent method.

Since the regularized nonlinearity $j_{\epsilon,\eta}(\cdot)$ is $\mathcal{C}^\infty$, it is possible to compute the shape derivative of the objective function $J(\Omega)$ defined by (4.58).

Theorem 20. *Let $\Omega \subset \mathbb{R}^d$ be a smooth bounded open set. Let $\boldsymbol{f} \in \mathcal{C}^0([0,T], H^1(\mathbb{R}^d)^d)$, $\boldsymbol{g} \in \mathcal{C}^0([0,T], H^2(\mathbb{R}^d)^d)$ and $\boldsymbol{w}(\Omega) \in H^1([0,T], Z)$ the solution to (4.56). Then the shape derivative of $J(\Omega)$ along $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, $J'(\Omega)(\boldsymbol{\theta})$ is given by*

$$J'(\Omega)(\boldsymbol{\theta}) = \int_0^T \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} \Big(m(\boldsymbol{w}) + \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p) : (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\varepsilon}_q) + \boldsymbol{\varepsilon}_q : \mathbb{H}\boldsymbol{\varepsilon}_p + E_{iso}\gamma\mu + \nabla_z \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}) \cdot \boldsymbol{z} - l_t(\boldsymbol{z}) \Big) ds\, dt, \tag{4.61}$$

where $\boldsymbol{z}(\Omega) \in H^1([0,T], Z)$ is the solution to the adjoint problem, with the final condition $\boldsymbol{z}(T) = \boldsymbol{0}$,

$$a(\boldsymbol{z}, \boldsymbol{\varphi}) - \left\langle \frac{d}{dt} \left(\nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}) \boldsymbol{z} \right), \boldsymbol{\varphi} \right\rangle = -\langle \nabla_z m(\boldsymbol{w}), \boldsymbol{\varphi} \rangle - \int_{\Gamma_N} \nabla_z p(\boldsymbol{w}) \boldsymbol{\varphi}\ ds \qquad \forall t \in [0,T),\ \forall \boldsymbol{\varphi} \in Z, \tag{4.62}$$

which is assumed to be well-posed (recall that $\langle , \rangle$ is the scalar product defined by (4.30) in Z).

Proof. The idea of the proof is classical and assuming that the adjoint equation is well-posed, it relies on Céa's technique [53]. Define three spaces $\tilde{V}$, $\tilde{Q}$ and $\tilde{Z} = \tilde{V} \times \tilde{Q} \times L^2(\mathbb{R}^d)$ (which are similar to those in (4.29) except that Ω is replaced by $\mathbb{R}^d$) by

$$\tilde{V} = \{ \boldsymbol{u} \in H^1(\mathbb{R}^d)^d,\ \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_D \} \quad \text{and} \quad \tilde{Q} = \{ \boldsymbol{\varepsilon}_q \in L^2(\mathbb{R}^d)^{d\times d},\ \operatorname{tr}(\boldsymbol{\varepsilon}_q) = 0 \text{ a.e. in } \mathbb{R}^d \}. \tag{4.63}$$

For $\tilde{\boldsymbol{w}} = (\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{\varepsilon}}_p, \tilde{\gamma}) \in H^1([0,T], \tilde{Z})$, $\tilde{\boldsymbol{z}} = (\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{\varepsilon}}_q, \tilde{\mu}) \in H^1([0,T], \tilde{Z})$ (the Lagrange multiplier for the state equation (4.56)) and $\tilde{\boldsymbol{\lambda}} \in L^2(\mathbb{R}^d)^d$ (the Lagrange multiplier for the initial condition $\tilde{\boldsymbol{w}}(0) = \boldsymbol{0}$), define a Lagrangian by

$$\begin{aligned} \mathcal{L}(\Omega, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}, \tilde{\boldsymbol{\lambda}}) = & \int_0^T \left(\int_\Omega m(\tilde{\boldsymbol{w}})\ dx + \int_{\Gamma_N} p(\tilde{\boldsymbol{w}}) ds \right) dt \\ & + \int_0^T \left(a(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}) - l_t(\tilde{\boldsymbol{z}}) + \langle \nabla_z \mathcal{D}_{\epsilon,\eta}(\dot{\tilde{\boldsymbol{w}}}), \tilde{\boldsymbol{z}} \rangle \right)\ dt + \int_\Omega \tilde{\boldsymbol{\lambda}} \cdot \tilde{\boldsymbol{w}}(0)\, dx. \end{aligned} \tag{4.64}$$

We remark that here the variables $\tilde{\boldsymbol{w}}(t), \tilde{\boldsymbol{z}}(t)$ and $\tilde{\boldsymbol{\lambda}}$ are defined on the full space $\mathbb{R}^d$ and are thus independent of Ω. Although $\tilde{\boldsymbol{u}}(t)$ and $\tilde{\boldsymbol{v}}(t)$ are required to vanish on Γ_D, they do not depend on Ω since Γ_D is a fixed boundary. Therefore, writing the optimality conditions applied to the Lagrangian (4.64), namely that its partial derivatives with respect to the independent variables $(\Omega, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\lambda})$ vanishes, yields the state equation, the adjoint equation and the shape derivative.

When the Lagrangian (4.64) is differentiated with respect to the adjoint variable $\tilde{\boldsymbol{z}}$, along $\boldsymbol{\varphi} \in H^1([0,T], \tilde{Z})$, and equated to zero, followed by the substitution $\tilde{\boldsymbol{w}} = \boldsymbol{w}$, we get

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{z}}(\boldsymbol{\varphi}) = \int_0^T \left(a(\boldsymbol{w}, \boldsymbol{\varphi}) + \langle \nabla_z \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}), \boldsymbol{\varphi}\rangle - l_t(\boldsymbol{\varphi})\right)\, dt = 0 \qquad \forall \boldsymbol{\varphi} \in H^1([0,T], \tilde{Z}).$$

Since the bilinear form $a(\cdot,\cdot)$ and the linear forms in the above are defined only on Ω, we can replace $\tilde{Z}$ by Z. Differentiating (4.64) with respect to $\tilde{\boldsymbol{\lambda}}$ at $\tilde{\boldsymbol{w}} = \boldsymbol{w}$, equating it to zero, we deduce the initial condition $\boldsymbol{w}(0) = \boldsymbol{0}$ a.e. on Ω. We thus recover the state equation (4.56). Next, we differentiate the Lagrangian (4.64) with respect to $\tilde{\boldsymbol{w}}$ along $\boldsymbol{\varphi} \in H^1([0,T], \tilde{Z})$ and equate it to zero at $\tilde{\boldsymbol{w}} = \boldsymbol{w}$, $\tilde{\boldsymbol{z}} = \boldsymbol{z}$, $\tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda}$, to get

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \boldsymbol{w}}(\boldsymbol{\varphi}) = &\int_0^T \int_\Omega \nabla_z m(\boldsymbol{w})\boldsymbol{\varphi}\, dx\, dt + \int_0^T \int_{\Gamma_N} \nabla_z p(\boldsymbol{w})\boldsymbol{\varphi}\, ds\, dt \\ &+ \int_0^T a(\boldsymbol{\varphi}, \boldsymbol{z}) dt + \int_0^T \langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})\dot{\boldsymbol{\varphi}}, \boldsymbol{z}\rangle dt + \int_\Omega \boldsymbol{\lambda}\cdot\boldsymbol{\varphi}(0)\, dx = 0 \qquad \forall \boldsymbol{\varphi} \in H^1([0,T], \tilde{Z}).\end{aligned}$$

Using the symmetry of the second derivative $\nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})$, and integrating by parts in time, we deduce

$$\begin{aligned}&\int_0^T \int_\Omega \nabla_z m(\boldsymbol{w})\boldsymbol{\varphi}\, dx\, dt + \int_0^T \int_{\Gamma_N} \nabla_z p(\boldsymbol{w})\boldsymbol{\varphi}\, ds\, dt + \int_0^T a(\boldsymbol{\varphi}, \boldsymbol{z}) dt + \langle \boldsymbol{\varphi}, \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})\boldsymbol{z}\rangle\big|_{t=T} \\ &\quad - \langle \boldsymbol{\varphi}, \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})\boldsymbol{z}\rangle\big|_{t=0} - \int_0^T \left\langle \boldsymbol{\varphi}, \frac{d}{dt}(\nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})\boldsymbol{z})\right\rangle dt + \int_\Omega \boldsymbol{\lambda}\cdot\boldsymbol{\varphi}(0)\, dx = 0 \quad \forall \boldsymbol{\varphi} \in H^1([0,T], \tilde{Z}).\end{aligned}$$

Since all integrals in the above are defined only on Ω, we can replace $\tilde{Z}$ by Z. Varying the test function $\boldsymbol{\varphi}$, we derive the following adjoint equation:

$$\boldsymbol{\lambda} = \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})\boldsymbol{z}\big|_{t=0}, \quad \boldsymbol{z}(T) = \boldsymbol{0} \quad \text{and}$$

$$a(\boldsymbol{z}, \boldsymbol{\varphi}) - \left\langle \frac{d}{dt}\left(\nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})\boldsymbol{z}\right), \boldsymbol{\varphi}\right\rangle = -\langle \nabla_z m(\boldsymbol{w}), \boldsymbol{\varphi}\rangle - \int_{\Gamma_N} \nabla_z p(\boldsymbol{w})\boldsymbol{\varphi}\, ds \qquad t \in [0,T),\ \forall\ \boldsymbol{\varphi} \in Z.$$

Finally, using the relation $J(\Omega) = \mathcal{L}(\Omega, \boldsymbol{w}, \tilde{\boldsymbol{z}}, \tilde{\boldsymbol{\lambda}})$, we determine the shape derivative $J'(\Omega)(\boldsymbol{\theta})$ for any $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ by

$$J'(\Omega)(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{\theta}) + \frac{\partial \mathcal{L}}{\partial \boldsymbol{w}}\left(\frac{\partial \boldsymbol{w}}{\partial \Omega}(\boldsymbol{\theta})\right),$$

because $\tilde{\boldsymbol{z}}$ and $\tilde{\boldsymbol{\lambda}}$ do not depend on Ω. Now, replacing them by their precise values $\boldsymbol{z}$ and $\boldsymbol{\lambda}$, given by the adjoint problem, the last term cancels to get

$$J'(\Omega)(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{\theta})$$

and formula (4.61) is deduced by application of Lemma 1. □

4.3.1 Well-posedness of the time discretized version of the adjoint equation (4.62)

In the previous proof, we assumed that the adjoint equation (4.62) was well-posed. The adjoint problem (4.62) is a linear backward parabolic equation with a final condition at $t = T$. The right hand side of (4.62) involves the derivative of the objective function which is assumed to satisfy a growth condition that renders it well-defined. The only difficult point is that the time derivative of $\boldsymbol{z}$ is multiplied by the Hessian operator of the convex dissipation function. If we knew that this operator is coercive, then existence and uniqueness would be easy (assuming further that $\dot{\boldsymbol{w}}$ is a smooth function). In full generality, the analysis for the time-continuous adjoint problem (4.62) is very complicated. However, if we consider a time-discretized version

of (4.62), then the analysis is much simpler as we shall now show. We split the time interval $[0,T]$ in N intervals of length δt. We denote the solution of the state problem (4.56), $\boldsymbol{w}(t)$ evaluated at time instant $t_n = n\delta t$ by $\boldsymbol{w}_n = (\boldsymbol{u}_n, \boldsymbol{\varepsilon}_{p,n}, \gamma_n)$. Similarly, $\dot{\boldsymbol{w}}_n = (\dot{\boldsymbol{u}}_n, \dot{\boldsymbol{\varepsilon}}_{p,n}, \dot{\gamma}_n)$ denotes the time derivative $\dot{\boldsymbol{w}}(t)$ at time instant t_n. On the other hand, $\boldsymbol{z}_n$ denotes an approximation of the adjoint state (4.62) at time t_n defined as the solution to the system below: for $\boldsymbol{z}_N = 0$, find a family $\boldsymbol{z}_n \in Z$, $N-1 \geq n \geq 0$, such that

$$a(\boldsymbol{\varphi}, \boldsymbol{z}_n) + \frac{1}{\delta t}\left\langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\boldsymbol{z}_n - \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{n+1})\boldsymbol{z}_{n+1}, \boldsymbol{\varphi}\right\rangle = -\langle \nabla_z m(\boldsymbol{w}_{n+1}), \boldsymbol{\varphi}\rangle - \int_{\Gamma_N} \nabla_z p(\boldsymbol{w}_{n+1}) \cdot \boldsymbol{\varphi}\, ds \qquad \forall \boldsymbol{\varphi} \in Z. \quad (4.65)$$

Theorem 21. *We assume that $\dot{\boldsymbol{w}}_n \in V \times L^\infty(\Omega)^{d\times d} \times L^2(\Omega)$ and $\epsilon > 0$, $\eta > 0$. Then the time-discretized adjoint problem (4.83) admits a unique solution $\boldsymbol{z}_n \in Z$, $n = N-1, \cdots, 1, 0$.*

Proof. Every equation in the system (4.83) is linear in $\boldsymbol{z}_n$. The form $a : Z \times Z \to \mathbb{R}$ is bilinear, symmetric, bounded and coercive as shown in Lemma 8 for $h_0 > 0$. In what follows, we show that the adjoint equation is well posed even for $h_0 = 0$. The bilinear form $\delta t a(\cdot,\cdot) + \langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\cdot,\cdot\rangle$ is symmetric and bounded, and in order to demonstrate its coercivity, we consider it for all $\boldsymbol{z}, \boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{\varepsilon}_q, \mu) \in Z$,

$$\delta t a(\boldsymbol{z},\boldsymbol{z}) + \langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\boldsymbol{z},\boldsymbol{z}\rangle = \delta t \left(\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\varepsilon}_q) : (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\varepsilon}_q)\ dx + \int_\Omega \mathbb{H}\boldsymbol{\varepsilon}_q : \boldsymbol{\varepsilon}_q\ dx + \int_\Omega E_{iso}\gamma^2\ dx \right) + \langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\boldsymbol{z},\boldsymbol{z}\rangle.$$

We write the expression of $\langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\boldsymbol{z},\boldsymbol{z}\rangle$ (which is the second derivative of $\mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})$ along two directions $\boldsymbol{z},\boldsymbol{z}$):

$$\langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\boldsymbol{z},\boldsymbol{z}\rangle = \int_\Omega \sqrt{\frac{2}{3}}\sigma_Y \left(N_\eta''(\dot{\boldsymbol{\varepsilon}}_p)\boldsymbol{\varepsilon}_q^2 + \frac{1}{\epsilon} M_\eta'' \left(\sqrt{\frac{2}{3}} N_\eta(\dot{\boldsymbol{\varepsilon}}_p) - \dot{\gamma} \right) \left(\sqrt{\frac{2}{3}} N_\eta'(\dot{\boldsymbol{\varepsilon}}_p)\boldsymbol{\varepsilon}_q - \mu \right)^2 + \frac{1}{\epsilon}\sqrt{\frac{2}{3}} M_\eta' \left(\sqrt{\frac{2}{3}} N_\eta(\dot{\boldsymbol{\varepsilon}}_p) - \dot{\gamma} \right) N_\eta''(\dot{\boldsymbol{\varepsilon}}_p)\boldsymbol{\varepsilon}_q^2 \right) dx \quad (4.66)$$

$$\text{where} \qquad N_\eta''(\dot{\boldsymbol{\varepsilon}}_p)\boldsymbol{\varepsilon}_q^2 = \left(\frac{\boldsymbol{\varepsilon}_q^2}{(\dot{\boldsymbol{\varepsilon}}_p^2 + \tilde{\eta}^2)^{1/2}} - \frac{(\dot{\boldsymbol{\varepsilon}}_p : \boldsymbol{\varepsilon}_q)^2}{(\dot{\boldsymbol{\varepsilon}}_p^2 + \tilde{\eta}^2)^{3/2}} \right) \quad \text{and} \quad \tilde{\eta} = \frac{\sigma_Y \eta}{TE}.$$

By construction $M_\eta'(\cdot),\ M_\eta''(\cdot) \geq 0$. Moreover

$$N_\eta''(\dot{\boldsymbol{\varepsilon}}_p)\boldsymbol{\varepsilon}_q^2 \geq \left(\frac{\boldsymbol{\varepsilon}_q^2}{(\dot{\boldsymbol{\varepsilon}}_p^2 + \tilde{\eta}^2)^{1/2}} - \frac{\dot{\boldsymbol{\varepsilon}}_p^2\ \boldsymbol{\varepsilon}_q^2}{(\dot{\boldsymbol{\varepsilon}}_p^2 + \tilde{\eta}^2)^{3/2}} \right) = \frac{\tilde{\eta}^2 \boldsymbol{\varepsilon}_q^2}{(\dot{\boldsymbol{\varepsilon}}_p^2 + \tilde{\eta}^2)^{3/2}}.$$

The second derivative (4.66) can then be bounded from below by

$$\begin{aligned}\langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\boldsymbol{z},\boldsymbol{z}\rangle &\geq \int_\Omega \sqrt{\frac{2}{3}}\sigma_Y \boldsymbol{\varepsilon}_q^2 \left(\frac{\tilde{\eta}^2}{(\dot{\boldsymbol{\varepsilon}}_{p,n}^2 + \tilde{\eta}^2)^{3/2}} + \frac{1}{\epsilon}\sqrt{\frac{2}{3}} M_\eta' \left(\sqrt{\frac{2}{3}} N_\eta(\dot{\boldsymbol{\varepsilon}}_{p,n}) - \dot{\gamma}_n \right) \frac{\tilde{\eta}^2}{(\dot{\boldsymbol{\varepsilon}}_{p,n}^2 + \tilde{\eta}^2)^{3/2}} \right) dx \\ &\geq \int_\Omega \sqrt{\frac{2}{3}} \frac{\sigma_Y \tilde{\eta}^2 \boldsymbol{\varepsilon}_q^2}{(\dot{\boldsymbol{\varepsilon}}_{p,n}^2 + \tilde{\eta}^2)^{3/2}} dx.\end{aligned} \quad (4.67)$$

Then, performing a similar calculation as in Lemma 8, we get for $s \in (0,1)$

$$\delta t a(\boldsymbol{z},\boldsymbol{z}) + \langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}_n)\boldsymbol{z},\boldsymbol{z}\rangle \geq \delta t \left(c_0 s \left\|\boldsymbol{\varepsilon}(\boldsymbol{v})\right\|_{L^2(\Omega)}^2 + \left(h_0 - \frac{c_0 s}{1-s} \right) \left\|\boldsymbol{\varepsilon}_q\right\|_{L^2(\Omega)}^2 + E_{iso} \left\|\mu\right\|_{L^2(\Omega)}^2 \right) + \sqrt{\frac{2}{3}}\sigma_Y \min_{x\in\Omega} \left(\frac{\tilde{\eta}^2}{(\dot{\boldsymbol{\varepsilon}}_{p,n}^2 + \tilde{\eta}^2)^{3/2}} \right) \left\|\boldsymbol{\varepsilon}_q\right\|_{L^2(\Omega)}^2.$$

Denote $C = \sqrt{\frac{2}{3}}\sigma_Y \min_{x\in\Omega}\left(\frac{\tilde{\eta}^2}{(\dot{\boldsymbol{\epsilon}}^2_{p,n}+\tilde{\eta}^2)^{3/2}}\right)$, which is finite since $\eta > 0$. By the assumption $\dot{\boldsymbol{\varepsilon}}_{p,n} \in L^\infty(\Omega)^{d\times d}$ we have that $C > 0$. If $h_0 > 0$, we take $s = \frac{h_0}{2c_0+h_0}$, while if $h_0 = 0$, we take $s = \frac{C}{2c_0\delta t+C}$ and find the left hand side in the above to be coercive. The adjoint equation (4.83) thus admits a unique solution $\boldsymbol{z}_n \in Z,\ n = N-1,\cdots,1,0$. □

Remark 8. *As shown in the previous theorem, the approximate dissipation function (4.44) is so constructed such that $\langle\nabla^2_z\mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}})\boldsymbol{z},\boldsymbol{z}\rangle$ is positive for all non zero $\boldsymbol{z}\in Z$. This ensures the well-posedness of the system (4.83) even for $h_0 = 0$. This is remarkable because the adjoint system (4.83) is well posed even when the state equation (4.56) cannot be shown to be well-posed.*

4.3.2 Formal limit analysis of the adjoint equation

At this stage, one may ask what happens in the adjoint (4.62) when the parameters $\epsilon,\eta \longrightarrow 0$. In this subsection, we answer this question in a formal manner. Note that the adjoint equation is the adjoint operator of the linearized equation corresponding to the nonlinear problem (4.46). We have been able to pass to the limit $\epsilon,\eta\longrightarrow 0$ in the nonlinear problem (4.46) (end of Section 4.2), and obtain the formulation (4.35). Instead of passing to the limit $\epsilon,\eta\longrightarrow 0$ in (4.62), one may equivalently linearize (4.35). This linearization is not possible in a classical sense as

1. it contains a non-smooth functional $j(\cdot)$
2. it is posed on a convex set K.

The first difficulty is addressed using the sub-differential of $j(\cdot)$. The second difficulty may be addressed using the notion of conical derivative, which is well studied for inequality of the first kind [171]. It finds application in the obstacle problem and contact mechanics [128]. The conical derivative for a particular inequality of the second kind is studied in [171] and an application to the continuous viscoplasticity problem is considered in [170].

The plasticity problem (4.35) is an inequality neither of the first kind nor of the second. Its time-discretized version however classifies as an inequality of the second kind. To the best of our knowledge, calculation of the conical derivative for neither the continuous plasticity problem (4.35) nor its time-discretized version has been performed yet and remains outside the scope of this thesis. Nevertheless, we shall determine a shape derivative for a simplified version of problem (4.35) and in a formal manner because of the very strong assumption (see **P.1** below).

We remind the reader that in problem (4.35) the solution and the forces are assumed to vanish at time $t = 0$. An incremental body force $\boldsymbol{f}$ and surface load $\boldsymbol{g}$ are applied. For $\boldsymbol{w}_0 = \boldsymbol{0}$, the solution at the end of every time-instant is computed using so that the problem reads: for

$$\boldsymbol{w}_{n+1} = \boldsymbol{w}_n + \Delta\boldsymbol{w}_n, \qquad n = 0,1,\cdots,N-1 \tag{4.68}$$

such that every increment $\Delta\boldsymbol{w}_n = (\Delta\boldsymbol{u}_n, \Delta\boldsymbol{\varepsilon}_{p,n}, \Delta\gamma_n) \in K$ satisfies

$$a(\Delta\boldsymbol{w}_n, \boldsymbol{z}-\Delta\boldsymbol{w}_n) + j(\boldsymbol{z}) - j(\Delta\boldsymbol{w}_n) \geq l_n(\boldsymbol{z}-\Delta\boldsymbol{w}_n) - a(\boldsymbol{w}_{n-1}, \boldsymbol{z}-\Delta\boldsymbol{w}_n) \qquad \forall\boldsymbol{z}\in K,\ n=1,\cdots,N, \tag{4.69}$$

where $\boldsymbol{z} = (\boldsymbol{v},\boldsymbol{\varepsilon}_q,\mu)$ and $l_n = l_{t_n}$ (see Def.(4.32)). The above is an inequation of the second category posed on a convex set K and admits a unique solution $\Delta\boldsymbol{w}_n \in Z$ [80]. The problem (4.69) can be equivalently be formulated as the minimization of $\mathcal{J}_n : K \to \mathbb{R}$

$$\min_{\boldsymbol{z}\in K} \mathcal{J}_n(\boldsymbol{z}), \tag{4.70}$$

$$\text{where} \quad \mathcal{J}_n(\boldsymbol{z}) = \left(\frac{1}{2}a(\boldsymbol{z},\boldsymbol{z}) + j(\boldsymbol{z}) - l_n(\boldsymbol{z}) + a(\boldsymbol{w}_{n-1},\boldsymbol{z})\right). \tag{4.71}$$

Since the functional $\mathcal{J}_n(\cdot)$ is convex on its convex domain K, it again admits a unique solution $\Delta\boldsymbol{w}_n \in Z$.

We seek an adjoint problem for a shape optimization problem with (4.69) as a constraint. For that, we need to convert the inequation (4.69) into an equation. Like in the case for contact mechanics in Section 3.36, there are two possibilities:

- adding more constraints to the space Q(see (4.73))

- reformulate the problem (4.69) using Lagrange multipliers.

Here, consider the first possibility only. We first introduce a set in Ω where the plastic flow rate is positive:

$$\Omega_{p,n} = \{\boldsymbol{x} \in \Omega |\ \Delta\boldsymbol{\varepsilon}_{p,n}(\boldsymbol{x}) \neq \mathbf{0} \quad \text{on } \Omega\}\,, \tag{4.72}$$

and then a space $Q_{p,n}$

$$Q_{p,n} = \{\boldsymbol{\varepsilon}_q \in Q |\ \boldsymbol{\varepsilon}_q(\boldsymbol{x}) = \mathbf{0} \quad \forall \boldsymbol{x} \in \Omega \backslash \Omega_{p,n}\}. \tag{4.73}$$

Both the set $\Omega_{p,n}$ and the space $Q_{p,n}$ are dependent on the time step index n (and independent of $\boldsymbol{w}_n$). Moreover, we suppose that set $\Omega_{p,n}$ is open and Lipschitz. To simplify the notations, let

$$\boldsymbol{\sigma}_{n+1} = \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1}) - \boldsymbol{\varepsilon}_{p,n+1}) \quad \text{and} \quad \boldsymbol{q}_{n+1} = \mathbb{H}\boldsymbol{\varepsilon}_{p,n+1}. \tag{4.74}$$

We now consider the minimization of (4.71) over a smaller space $V \times Q_{p,n}$:

$$\min_{\boldsymbol{z} \in V \times Q_{p,n} \times L^2(\Omega)} \mathcal{J}_n(\boldsymbol{z}) \tag{4.75}$$

$$\text{under the constraint } \mu = \sqrt{\frac{2}{3}}|\boldsymbol{\varepsilon}_q|. \tag{4.76}$$

This problem is well posed as the following theorem shows.

Theorem 22. *Let $\Omega_{p,n}$ and $Q_{p,n}$ be as defined in (4.72) and (4.73) respectively and let the set $\Omega_{p,n}$ be open and Lipschitz. There exists a unique solution $\Delta\boldsymbol{w}_n^* = (\Delta\boldsymbol{u}_n^*, \Delta\boldsymbol{\varepsilon}_{p,n}^*, \mu_n^*) \in V \times Q_{p,n} \times L^2(\Omega)$ to the problem (4.75). In addition, if $a(\cdot,\cdot)$ is coercive, then*

$$\Delta\boldsymbol{w}_n^* = \boldsymbol{w}_n,$$

where $\boldsymbol{w}_n$ is the solution to problem (4.70) and $\Delta\boldsymbol{w}_n^*$ additionally satisfies

$$\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}) \colon (\boldsymbol{\varepsilon}(\boldsymbol{u}_n^*) - \boldsymbol{\varepsilon}_{p,n}^*) dx = \int_\Omega \boldsymbol{f}_n \cdot \boldsymbol{v}\ dx + \int_{\Gamma_N} \boldsymbol{g}_n \cdot \boldsymbol{v}\ ds \qquad \forall \boldsymbol{v} \in V, \tag{4.77a}$$

$$\int_{\Omega_{p,n}} \boldsymbol{\varepsilon}_q \colon \left(\boldsymbol{\sigma}_n^* + \boldsymbol{q}_n^* - \sqrt{\frac{2}{3}} \frac{\Delta\boldsymbol{\varepsilon}_{p,n}^*}{|\Delta\boldsymbol{\varepsilon}_{p,n}^*|} (\sigma_Y + E_{iso}\gamma_n^*) \right) dx = 0 \qquad \forall \boldsymbol{\varepsilon}_q \in Q_{p,n}, \tag{4.77b}$$

where $(\boldsymbol{u}_n^*, \boldsymbol{\varepsilon}_{p,n}^*, \gamma_n^*) = (\boldsymbol{u}_{n-1}, \boldsymbol{\varepsilon}_{p,n-1}, \gamma_{n-1}) + (\Delta\boldsymbol{u}_n^*, \Delta\boldsymbol{\varepsilon}_{p,n}^*, \Delta\gamma_n^*)$ and $\boldsymbol{\sigma}_n^* = \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n^*) - \boldsymbol{\varepsilon}_{p,n}^*)$, $\boldsymbol{q}_n^* = -\mathbb{H}\boldsymbol{\varepsilon}_{p,n}^*$.

Proof. To simplify notations, let $S_n = V \times Q_{p,n} \times L^2(\Omega)$ and let $\boldsymbol{z} = (\boldsymbol{v}, \boldsymbol{\varepsilon}_q, \mu)$. Substituing the constraint $\mu = \sqrt{\frac{2}{3}}|\boldsymbol{\varepsilon}_q|$ in the functional (4.71), we get the minimization problem

$$\begin{aligned} &\min_{\boldsymbol{z} \in S_n} \mathcal{J}(\boldsymbol{z}) \\ \text{s.t. } \mathcal{J}(\boldsymbol{z}) = &\frac{1}{2}\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{v} + 2\boldsymbol{u}_{n-1}) - \boldsymbol{\varepsilon}_q - 2\boldsymbol{\varepsilon}_{p,n-1}) \colon (\boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{\varepsilon}_q) dx - l_n(\boldsymbol{z}) \\ &+ \frac{1}{2}\int_\Omega \mathbb{H}(\boldsymbol{\varepsilon}_q + 2\boldsymbol{\varepsilon}_{p,n-1}) \colon \boldsymbol{\varepsilon}_q dx + \frac{1}{2}\int_{\Omega_{p,n}} E_{iso}\left(\frac{2}{3} + \sqrt{\frac{8}{3}}\frac{\gamma_{n-1}}{|\boldsymbol{\varepsilon}_q|}\right)|\boldsymbol{\varepsilon}_q|^2\, dx + \int_{\Omega_{p,n}} \sqrt{\frac{2}{3}}\sigma_Y |\boldsymbol{\varepsilon}_q|\, dx \end{aligned} \tag{4.78}$$

Thus, the functional $\mathcal{J}(\boldsymbol{z})$ can be expressed purely in terms of $\boldsymbol{v}, \boldsymbol{\varepsilon}_q$. Thus $\mathcal{J}(\boldsymbol{z})$ (4.78) only depends $\boldsymbol{v}, \boldsymbol{\varepsilon}_q$ and the the the minimization above takes only over $V \times Q_{p,n}$. One can quite easily verify that the function $\mathcal{J}(\boldsymbol{z})$ (4.78) is convex with respect to $\boldsymbol{v}, \boldsymbol{\varepsilon}_q$ and the set $V \times Q_{p,n}$ is convex. Thus the above minimization admits a unique solution $\Delta\boldsymbol{w}_n^* = (\Delta\boldsymbol{u}_n^*, \Delta\boldsymbol{\varepsilon}_{p,n}^*, \Delta\gamma_n^*) \in S_n$ (where $\Delta\gamma_n^*$ satisfies the constraint (4.76)). Let $C_n \subset S_n$ be a set where the constraint (4.76) is satisfied. Then the solution $\Delta\boldsymbol{w}_n^*$ also satisfies

$$a(\Delta\boldsymbol{w}_n^*, \boldsymbol{z} - \Delta\boldsymbol{w}_n^*) + j(\boldsymbol{z}) - j(\Delta\boldsymbol{w}_n^*) \geq l_n(\boldsymbol{z} - \Delta\boldsymbol{w}_n^*) - a(\boldsymbol{w}_{n-1}, \boldsymbol{z} - \Delta\boldsymbol{w}_n^*) \qquad \forall \boldsymbol{z} \in C_n. \tag{4.79}$$

We now establish the equivalence between $\Delta\boldsymbol{w}_n^*$ and $\Delta\boldsymbol{w}_n$. One can easily verify that $\Delta\boldsymbol{w}_n \in C_n$ and that $\Delta\boldsymbol{w}_n^* \in K$. Substituting $\boldsymbol{z} = \Delta\boldsymbol{w}_n$ in (4.79) and $\boldsymbol{z} = \Delta\boldsymbol{w}_n^*$ in (4.69), and adding the two resulting inequations, we get

$$a(\Delta\boldsymbol{w}_n - \Delta\boldsymbol{w}_n^*, \Delta\boldsymbol{w}_n - \Delta\boldsymbol{w}_n^*) \leq 0.$$

Given the coercivity of $a(\cdot,\cdot)$, we have

$$\Delta \boldsymbol{w}_n = \Delta \boldsymbol{w}_n^*.$$

In (4.78), the integrals over $\Omega_{p,n}$ are differentiable in the classical sense if $\boldsymbol{\varepsilon}_q(\boldsymbol{x}) \neq \mathbf{0}$ over $\Omega_{p,n}$. We know that our minimizer $\boldsymbol{\varepsilon}_p$ is non-zero over $\Omega_{p,n}$ by definition (4.73). We can thus differentiate the functional $\mathcal{J}(\cdot)$ at $\boldsymbol{z} = \Delta \boldsymbol{w}^*$, equate it to zero and obtain (4.77a)-(4.77b). □

Remark 9. *No derivatives are involved in* (4.77b) *so it implies*

$$\boldsymbol{\sigma}_n^{*D} + \boldsymbol{q}_n^{*D} - \sqrt{\frac{2}{3}} \frac{\Delta \boldsymbol{\varepsilon}_{p,n}^*}{|\Delta \boldsymbol{\varepsilon}_{p,n}^*|} (\sigma_Y + E_{iso}\gamma_n^*) = 0 \quad \textit{and} \quad |\boldsymbol{\sigma}_n^{*D} + \boldsymbol{q}_n^{*D}| = \sqrt{\frac{2}{3}} (\sigma_Y + E_{iso}\gamma_n^*) \tag{4.80}$$

since $\boldsymbol{\varepsilon}_q$ is trace-free. In other words, the normality law is respected and the yield limit is attained wherever the plastic flow rate is positive. The inverse may not be true as one may attain the yield limit and still have zero plastic flow rate. Problem (4.69) *takes Ω and the forces $\boldsymbol{f}, \boldsymbol{g}$ as input whereas problem* (4.77a)-(4.77b) *takes $\Omega, \boldsymbol{f}, \boldsymbol{g}$ well as $\Omega_{p,n}$ (and hence $Q_{p,n}$) as input.*

Thus we have two equations (4.77a)-(4.77b) to completely describe the problem (4.69). We note that even in this form, the problem remains tough to resolve as the domain of integration in (4.77b) depends on $\Delta \boldsymbol{w}_n$ that is an unknown of the problem.

For the shape optimization problem, we consider the time-discretized version of the objective function (4.58),

$$J(\Omega) = \sum_{n=1}^{N} \delta t \left(\int_\Omega m(\boldsymbol{w}_n(\Omega)) dx + \int_{\Gamma_N} p(\boldsymbol{w}_n(\Omega)) ds \right), \tag{4.81}$$

where δt is the fixed time interval and $\boldsymbol{w}_n$ solves (4.77a)-(4.77b). To determine the shape derivative for the functional (4.81), we state the following theorem. We compute the shape derivative of (4.81) under the following strong assumption.

P.1 When Ω is perturbed to $(I_d + \boldsymbol{\theta})\Omega$, with a small vector field $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, the corresponding solution $\boldsymbol{w}_{n,\boldsymbol{\theta}} \equiv \boldsymbol{w}_n((I_d + \boldsymbol{\theta})\Omega)$ of problem (4.70) (for all $1 \leq n \leq N$,) is differentiable with respect to $\boldsymbol{\theta}$ and the plastic zone of $\boldsymbol{w}_{n,\boldsymbol{\theta}}$, $\Omega_{n,p,\boldsymbol{\theta}}$ (see Def.4.73) is perturbed to $(I_d + \boldsymbol{\theta}_p)\Omega_{p,n}$, where $\Omega_{p,n}$ is the plastic zone of $\boldsymbol{w}_n(\Omega)$ and $\boldsymbol{\theta}_p$ is a vector field which smoothly depends on $\boldsymbol{\theta}$ and $\boldsymbol{w}_n$, while $\Omega_{p,n}$ is an open Lipschitz set.

The (A) assumption implies that the plastic zone does not change its topology, meaning that there is no creation of new plastic zones or creation of elastic zones inside the plastic zone.

Theorem 23. *Let $\Omega \subset \mathbb{R}^d$ be a smooth and bounded open set, $\boldsymbol{f}_n = \boldsymbol{f}(t_n) \in H^1(\mathbb{R}^d)^d$ and $\boldsymbol{g}_n = \boldsymbol{g}(t_n) \in H^2(\mathbb{R}^d)^d$ be smooth loads. Assume that the solution to* (4.77a)-(4.77b), *$\Delta \boldsymbol{w}_n = (\Delta \boldsymbol{u}_n, \Delta \boldsymbol{\varepsilon}_{p,n}, \Delta \gamma_n) \in V \times Q \times L^2(\Omega)$, is smooth, namely belongs to $H^2(\Omega)^d \times H^1(\Omega)^{d\times d} \times L^2(\Omega)$. Assume that the integrand $p(\cdot)$, in the cost function* (4.81), *does not depend on $\Delta \boldsymbol{\varepsilon}_{p,n}$ and $\Delta \gamma_n$ and that the zone in Ω where plastic rate is positive, $\Omega_{p,n}$ grows with respect to n i.e., $\Omega_{p,1} \subset \Omega_{p,2} \subset \cdots \subset \Omega_{p,N}$.*

Then under the assumption ***P.1****, the shape derivative of* (4.81), *in the direction $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, is given by*

$$J'(\Omega)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n} \sum_{n=1}^{N} (m(\boldsymbol{w}_n) + \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}_n) : (\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n}) - \boldsymbol{f}_n \cdot \boldsymbol{v}_n)\, ds, \tag{4.82}$$

where $\boldsymbol{\sigma}, \boldsymbol{q}$ are defined in (4.74) *and $\boldsymbol{z}_n = (\boldsymbol{v}_n, \boldsymbol{\varepsilon}_{q,n}, \mu_n) \in Z$ is the adjoint variables satisfying $\forall \boldsymbol{\varphi} \in V$,*

$$\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{\varphi}) : (\boldsymbol{\varepsilon}(\boldsymbol{v}_n) - \boldsymbol{\varepsilon}_{q,n}) dx = -\left(\int_\Omega \partial_{\boldsymbol{u}_n} m(\boldsymbol{w}_n(\Omega)) \boldsymbol{\varphi}\, \check{a} dx + \int_{\Gamma_N} \partial_{\boldsymbol{u}_n} p(\boldsymbol{w}_n(\Omega)) \boldsymbol{\varphi}\, ds \right) \qquad \forall \boldsymbol{\varphi} \in V \tag{4.83a}$$

$$\begin{aligned}
\boldsymbol{\varepsilon}_{q,n} = \Bigg(& \mathbb{C} + \mathbb{H} + \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_n) \left(\frac{\mathbb{I}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} \right. \\
& \left. - \frac{\boldsymbol{\varepsilon}_{p,n} \otimes \boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|^3} \right) \Bigg)^{-1} \Bigg(-\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}_n) + \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_{n+1}) \left(\frac{\boldsymbol{\varepsilon}_{q,n+1}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \frac{(\boldsymbol{\varepsilon}_{p,n} : \boldsymbol{\varepsilon}_{q,n+1})\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|^3} \right) \\
& + \sqrt{\frac{2}{3}} \mu_n \frac{\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} - \sqrt{\frac{2}{3}} \mu_{n+1} \frac{\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \partial_{\boldsymbol{\varepsilon}_{p,n}} m(\boldsymbol{w}_n) \Bigg)^D
\end{aligned} \tag{4.83b}$$

$$\mu_n = \mu_{n+1} - \partial_{\gamma_n} m(\boldsymbol{w}_n) \qquad \text{in } \Omega\backslash\Omega_p \tag{4.83c}$$

$$\mu_n = \mu_{n+1} - \partial_{\gamma_n} m(\boldsymbol{w}_n) + \sqrt{\frac{2}{3}} E_{iso} \frac{\boldsymbol{\varepsilon}_{q,n} : (\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1})}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} \qquad \text{in } \Omega_p$$

Proof. The main idea is to define a Lagrangian for the simpler equations (4.77a)-(4.77b) rather than for the variational inequality (4.69). As usual, we apply Céa's method in order to determine the shape derivative by constructing the Lagrangian

$$\begin{aligned}
&\mathcal{L}\left(\{\tilde{\boldsymbol{w}}_n, \tilde{\boldsymbol{z}}_n, \tilde{\boldsymbol{\lambda}}_n \tilde{\Omega}_{p,n}\}_{n=1,\cdots,N}, \Omega\right) \\
&= \sum_{n=1}^{N} \delta t \left(\int_\Omega m(\tilde{\boldsymbol{w}}_n) dx + \int_{\Gamma_N} p(\tilde{\boldsymbol{w}}_n) ds \right) \\
&\quad + \sum_{n=1}^{N} \delta t \left(\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}}_n) : (\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_n) - \tilde{\boldsymbol{\varepsilon}}_{p,n}) dx - \int_\Omega \boldsymbol{f}_n \cdot \tilde{\boldsymbol{v}}_n \, dx - \int_{\Gamma_N} \boldsymbol{g}_n \cdot \tilde{\boldsymbol{v}}_n \, ds \right) \\
&\quad + \sum_{n=1}^{N} \delta t \left(\int_{\tilde{\Omega}_{p,n}} \tilde{\boldsymbol{\varepsilon}}_{q,n} : \left(\tilde{\boldsymbol{\sigma}}_n^D + \tilde{\boldsymbol{q}}_n^D - \sqrt{\frac{2}{3}} (\sigma_Y + E_{iso} \tilde{\gamma}_n) \frac{\tilde{\boldsymbol{\varepsilon}}_{p,n} - \tilde{\boldsymbol{\varepsilon}}_{p,n-1}}{|\tilde{\boldsymbol{\varepsilon}}_{p,n} - \tilde{\boldsymbol{\varepsilon}}_{p,n-1}|} \right) dx \right) \\
&\quad + \sum_{n=1}^{N} \int_{\Omega\backslash\tilde{\Omega}_{p,n}} \tilde{\boldsymbol{\lambda}}_n : (\tilde{\boldsymbol{\varepsilon}}_{p,n} - \tilde{\boldsymbol{\varepsilon}}_{p,n-1}) dx + \sum_{n=1}^{N} \int_\Omega \tilde{\mu}_n \left(\tilde{\gamma}_n - \tilde{\gamma}_{n-1} - \sqrt{\frac{2}{3}} |\tilde{\boldsymbol{\varepsilon}}_{p,n} - \tilde{\boldsymbol{\varepsilon}}_{p,n-1}| \right) dx
\end{aligned} \tag{4.84}$$

where

- $\tilde{\boldsymbol{w}}_n = (\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{\varepsilon}}_{p,n}, \tilde{\gamma}) \in \tilde{V} \times \tilde{Q} \times L^2(\mathbb{R})$ (defined in (4.63))
- $\tilde{\boldsymbol{z}}_n = (\tilde{\boldsymbol{v}}_n, \tilde{\boldsymbol{\varepsilon}}_{q,n}, \tilde{\mu}) \in \tilde{V} \times \tilde{Q} \times L^2(\mathbb{R})$ is the adjoint variable
- $\tilde{\boldsymbol{\sigma}}_n = \mathbb{C}(\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_n) - \tilde{\boldsymbol{\varepsilon}}_{p,n}), \tilde{\boldsymbol{q}}_n = -\mathbb{H}\tilde{\boldsymbol{\varepsilon}}_{p,n}$ (similar to the definition (4.74))
- $\tilde{\Omega}_{p,n}$ is the zone where plastic rate is positive, assumed to be independent of $\tilde{w}_n$ here
- $\tilde{\boldsymbol{\lambda}}_n$ is the Lagrange multiplier that penalizes $\tilde{\boldsymbol{\varepsilon}}_{p,n} = \tilde{\boldsymbol{\varepsilon}}_{p,n-1}$ in the zone where the plastic flow is not positive

The variables $\tilde{\boldsymbol{w}}_n$ and $\tilde{\boldsymbol{z}}_n$ vanish on Γ_D, which is a fixed set, so it does not cause any problem for differentiating the Lagrangian (4.84).

We now compute the optimality condition for the Lagrangian (4.84). The optimal variables are denoted by $(\boldsymbol{w}_n, \boldsymbol{z}_n, \boldsymbol{\lambda}_n)$, $n = 1, \cdots, N$. Since the Lagrangian is linear with respect to $\tilde{\boldsymbol{z}}$ and $\tilde{\boldsymbol{\lambda}}$, it is easy to compute its partial derivatives with respect to these two variables. Equating to zero the partial derivative for $\tilde{\boldsymbol{z}}$ in the direction $(\boldsymbol{\varphi}, \boldsymbol{\psi}, \xi) \in V \times Q \times L^2(\Omega)$ yields

$$\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}) : (\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n}) dx = \int_\Omega \boldsymbol{f}_n \cdot \boldsymbol{v} \, dx + \int_{\Gamma_N} \boldsymbol{g}_n \cdot \boldsymbol{v} \, ds \qquad \forall \boldsymbol{v} \in V, \tag{4.85a}$$

$$\int_{\tilde{\Omega}_{p,n}} \boldsymbol{\varepsilon}_q : \left(\boldsymbol{\sigma}_n + \boldsymbol{q}_n - \sqrt{\frac{2}{3}} \frac{\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} (\sigma_Y + E_{iso}\gamma_n) \right) dx = 0 \qquad \forall \boldsymbol{\varepsilon}_q \in Q_{p,n}, \tag{4.85b}$$

$$\int_{\Omega\backslash\tilde{\Omega}_{p,n}} \xi(\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}) dx = 0 \qquad \forall \xi \in L^2(\Omega). \tag{4.85c}$$

Equating to zero the partial derivative for $\tilde{\boldsymbol{\lambda}}_n$ in the direction $\boldsymbol{\psi} \in Q$ leads to

$$\int_{\Omega\backslash\tilde{\Omega}_{p,n}} \boldsymbol{\psi} : (\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}) dx = 0 \qquad \forall \boldsymbol{\psi} \in Q, \tag{4.86}$$

which implies that $\boldsymbol{\varepsilon}_{p,n} = \boldsymbol{\varepsilon}_{p,n-1}$ in $\Omega\backslash\tilde{\Omega}_{p,n}$. Choosing $\tilde{\Omega}_{p,n} = \Omega_{p,n}$ one can check that the optimality conditions (4.85) and (4.86) are precisely the state equations (4.77a)-(4.77b).

The adjoint equations (4.83) are obtained by writing the optimality condition of the Lagrangian (4.84) with respect to $\tilde{\boldsymbol{w}}_n$ in the direction $(\boldsymbol{\varphi}, \boldsymbol{\psi}, \xi) \in V \times Q \times L^2(\Omega)$. The adjoint $\boldsymbol{z}_n$ is evaluated at the state $\boldsymbol{w}_n$,

$\boldsymbol{\lambda}_n$ and $\tilde{\Omega}_{p,n} = \Omega_{p,n}$. The adjoint problem amounts to find $\boldsymbol{z}_n = (\boldsymbol{v}_n, \boldsymbol{\varepsilon}_{q,n}, \mu_n) \in V \times Q \times L^2(\Omega)$, $1 \le n \le N$ such that

$$\int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{\varphi}) : (\boldsymbol{\varepsilon}(\boldsymbol{v}_n) - \boldsymbol{\varepsilon}_{q,n}) dx = -\left(\int_\Omega \partial_{\boldsymbol{u}_n} m(\boldsymbol{w}_n)\boldsymbol{\varphi}\, \check{a} dx + \int_{\Gamma_N} \partial_{\boldsymbol{u}_n} p(\boldsymbol{w}_n)\boldsymbol{\varphi}\, ds \right) \qquad \forall \boldsymbol{\varphi} \in V,$$

$$\begin{aligned}
&\int_{\Omega_{p,n}} \boldsymbol{\psi} : \left(\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}_n) + (\mathbb{C}+\mathbb{H})\boldsymbol{\varepsilon}_{q,n} + \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_n) \left(\frac{\boldsymbol{\varepsilon}_{q,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} \right.\right. \\
&\left.\left. - \frac{(\boldsymbol{\varepsilon}_{p,n} : \boldsymbol{\varepsilon}_{q,n})\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|^3} \right) \right) dx - \int_{\Omega_{p,n+1}} \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso})\gamma_{n+1}\boldsymbol{\psi} : \left(\frac{\boldsymbol{\varepsilon}_{q,n+1}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \frac{(\boldsymbol{\varepsilon}_{p,n} : \boldsymbol{\varepsilon}_{q,n+1})\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|^3} \right) dx \\
&+ \int_{\Omega\setminus\Omega_{p,n}} \boldsymbol{\lambda}_n : \boldsymbol{\psi} dx - \int_{\Omega\setminus\Omega_{p,n+1}} \boldsymbol{\lambda}_{n+1} : \boldsymbol{\psi} dx - \int_{\Omega_{p,n}} \sqrt{\frac{2}{3}}\mu_n \frac{\boldsymbol{\psi} : \boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} dx \\
&+ \int_{\Omega_{p,n+1}} \sqrt{\frac{2}{3}}\mu_{n+1} \frac{\boldsymbol{\psi} : \boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} dx = -\int_\Omega \partial_{\boldsymbol{\varepsilon}_{p,n}} m(\boldsymbol{w}_n)\boldsymbol{\psi}\, \check{a} dx \qquad \forall \boldsymbol{\psi} \in Q.
\end{aligned} \tag{4.87a}$$

$$\int_\Omega (\mu_n - \mu_{n+1})\xi\, dx - \int_{\Omega_{p,n}} \sqrt{\frac{2}{3}} E_{iso} \frac{\boldsymbol{\varepsilon}_{q,n} : (\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1})}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} \xi\, dx = -\int_\Omega \partial_{\gamma_n} m(\boldsymbol{w}_n)\xi\, \check{a} dx \qquad \forall \xi \in L^2(\Omega) \tag{4.87b}$$

Since the second equation (4.87a) above does not involve any derivative on the test function $\boldsymbol{\psi}$, we can equate the coefficient of $\boldsymbol{\psi}$ to zero. Equation (4.87a) contains integrals on several domains. Here, we use the assumption (B), with which we have the inclusion $\Omega_{p,n} \subset \Omega_{p,n+1}$. Choosing a test function $\boldsymbol{\psi} = \mathbf{0}$ on $\Omega\setminus\Omega_{p,n}$, we can equate the integral on $\Omega_{p,n}$ to zero in (4.87a), and obtain

$$\begin{aligned}
&\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}_n) + (\mathbb{C}+\mathbb{H})\boldsymbol{\varepsilon}_{q,n} + \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_n) \left(\frac{\boldsymbol{\varepsilon}_{q,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} - \frac{(\boldsymbol{\varepsilon}_{p,n} : \boldsymbol{\varepsilon}_{q,n})\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|^3} \right) - \sqrt{\frac{2}{3}}\mu_n \frac{\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} \\
&= \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_{n+1}) \left(\frac{\boldsymbol{\varepsilon}_{q,n+1}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \frac{(\boldsymbol{\varepsilon}_{p,n} : \boldsymbol{\varepsilon}_{q,n+1})\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|^3} \right) - \sqrt{\frac{2}{3}}\mu_{n+1} \frac{\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \partial_{\boldsymbol{\varepsilon}_{p,n}} m(\boldsymbol{w}_n)
\end{aligned}$$

Using tensor contraction property, we can separate $\boldsymbol{\varepsilon}_{q,n}$ in the above as

$$\begin{aligned}
&\left(\mathbb{C} + \mathbb{H} + \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_n) \left(\frac{\mathbb{I}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} - \frac{\boldsymbol{\varepsilon}_{p,n} \otimes \boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|^3} \right) \right) \boldsymbol{\varepsilon}_{q,n} \\
&= -\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{v}_n) + \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_{n+1}) \left(\frac{\boldsymbol{\varepsilon}_{q,n+1}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \frac{(\boldsymbol{\varepsilon}_{p,n} : \boldsymbol{\varepsilon}_{q,n+1})\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|^3} \right) \\
&+ \sqrt{\frac{2}{3}}\mu_n \frac{\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|} - \sqrt{\frac{2}{3}}\mu_{n+1} \frac{\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \partial_{\boldsymbol{\varepsilon}_{p,n}} m(\boldsymbol{w}_n),
\end{aligned}$$

resulting in the expression for $\boldsymbol{\varepsilon}_{q,n}$ in (4.107). Since $\boldsymbol{\varepsilon}_{q,n} \in Q$, it must be trace-free. To ensure this, we find an expression for $\boldsymbol{\varepsilon}_{q,n}$ from the above, and then take the deviator on both sides, resulting in (4.107).

Using the assumption (B), we also have $\Omega\setminus\Omega_{p,n+1} \subset \Omega\setminus\Omega_{p,n}$. Choosing a test function $\boldsymbol{\psi} = 0$ on $\Omega\setminus\Omega_{p,n}$ in (4.87a), we get

$$\boldsymbol{\lambda}_n = -\partial_{\boldsymbol{\varepsilon}_{p,n}} m(\boldsymbol{w}_n) \quad \text{in } \Omega_{p,n+1}\setminus\Omega_{p,n}$$

$$\begin{aligned}
\boldsymbol{\lambda}_n = \boldsymbol{\lambda}_{n+1} + \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_{n+1}) &\left(\frac{\boldsymbol{\varepsilon}_{q,n+1}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} \right. \\
&\left. - \frac{(\boldsymbol{\varepsilon}_{p,n} : \boldsymbol{\varepsilon}_{q,n+1})\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|^3} \right) - \sqrt{\frac{2}{3}}\mu_{n+1} \frac{\boldsymbol{\varepsilon}_{p,n}}{|\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}|} - \partial_{\boldsymbol{\varepsilon}_{p,n}} m(\boldsymbol{w}_n) \quad \text{in } \Omega\setminus\Omega_{p,n+1}
\end{aligned}$$

Using similar manipulation on the last adjoint equation (4.87b), we get (4.83c).

To simplify notations, we let $\mathbb{N} = \{1, 2, \cdots, N\}$. Since $\boldsymbol{w}_n$ and $\tilde{\Omega}_{p,n} = \Omega_{p,n}$, the solution of problem (4.70), satisfies the state equations (4.77a)-(4.77b), we have, for any $\tilde{\boldsymbol{z}}_n, \tilde{\boldsymbol{\lambda}}_n$,

$$\mathcal{L}\left(\{\boldsymbol{w}_n, \tilde{\boldsymbol{z}}_n, \tilde{\boldsymbol{\lambda}}_n, \Omega_{p,n}\}_{n\in\mathbb{N}}, \Omega\right) = J(\Omega).$$

Now we differentiate both sides in the above with respect to the shape Ω in the direction of a vector field $\boldsymbol{\theta}$. Because of assumption **P.1**, when $\boldsymbol{\theta}$ moves the domain Ω to $(I_d + \boldsymbol{\theta})\Omega$, the plastic zone $\Omega_{\boldsymbol{w}}$ is displaced to $(I_d + \boldsymbol{\theta}_p)\Omega_{\boldsymbol{w}}$, with another vector field $\boldsymbol{\theta}_p$. Thus, by the chain rule lemma, we obtain

$$J'(\Omega)(\boldsymbol{\theta}) = \left\langle \frac{\partial \mathcal{L}}{\partial \Omega}\left(\{\boldsymbol{w}_n, \tilde{\boldsymbol{z}}_n, \tilde{\boldsymbol{\lambda}}_n, \Omega_{p,n}\}_{n\in\mathbb{N}}, \Omega\right), \boldsymbol{\theta}\right\rangle + \sum_{n=1}^{N}\left(\left\langle \frac{\partial \mathcal{L}}{\partial \Omega_{p,n}}\left(\{\boldsymbol{w}_n, \tilde{\boldsymbol{z}}_n, \tilde{\boldsymbol{\lambda}}_n, \Omega_{p,n}\}_{n\in\mathbb{N}}, \Omega\right), \boldsymbol{\theta}_p\right\rangle + \left\langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{w}_n}\left(\{\boldsymbol{w}_n, \tilde{\boldsymbol{z}}_n, \tilde{\boldsymbol{\lambda}}_n, \Omega_{p,n}\}_{n\in\mathbb{N}}, \Omega\right), \frac{\partial \boldsymbol{w}_n}{\partial \Omega}(\boldsymbol{\theta})\right\rangle\right). \tag{4.89}$$

Now, substituting $\tilde{\boldsymbol{z}}_n = \boldsymbol{z}_n$, $\tilde{\boldsymbol{\lambda}}_n = \boldsymbol{\lambda}_n, 1 \le n \le N$, and using the adjoint equation (4.83), the last line of (4.89) vanishes because it is the adjoint variational formulation. We obtain

$$J'(\Omega)(\boldsymbol{\theta}) = \left\langle \frac{\partial \mathcal{L}}{\partial \Omega}\left(\{\boldsymbol{w}_n, \boldsymbol{z}_n, \boldsymbol{\lambda}_n, \Omega_{p,n}\}_{n\in\mathbb{N}}, \Omega\right), \boldsymbol{\theta}\right\rangle + \sum_{n=1}^{N}\left(\int_{\partial\Omega_{p,n}} (\boldsymbol{\theta}_p \cdot \boldsymbol{n})\, \boldsymbol{\varepsilon}_{q,n} : \left(\boldsymbol{\sigma}_n^D + \boldsymbol{q}_n^D - \sqrt{\frac{2}{3}}(\sigma_Y + E_{iso}\gamma_n)\frac{\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}}{|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|}\right) ds + \int_{\partial(\Omega\setminus\tilde{\Omega}_{p,n})} (\boldsymbol{\theta}_p \cdot \boldsymbol{n})\boldsymbol{\lambda}_n : (\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}) ds\right).$$

Given that $\boldsymbol{\varepsilon}_{p,n} = \boldsymbol{\varepsilon}_{p,n-1}$ in $\Omega\setminus\Omega_{\boldsymbol{w}}$ and the assumption that $\boldsymbol{\varepsilon}_{p,n} \in H^1(\Omega)^{d\times d}$, we deduce that the last integral vanishes. Furthermore, the smoothness assumption $(\boldsymbol{u}_n, \boldsymbol{\varepsilon}_{q,n}, \gamma_n) \in H^2(\Omega)^d \times H^1(\Omega)^{d\times d} \times L^2(\Omega)$ implies that the yield strength is attained even on $\partial\Omega_{p,n}$, so the penultimate integral vanishes too. Thus we find

$$J'(\Omega)(\boldsymbol{\theta}) = \left\langle \frac{\partial \mathcal{L}}{\partial \Omega}\left(\{\boldsymbol{w}_n, \boldsymbol{z}_n, \boldsymbol{\lambda}_n, \Omega_{p,n}\}_{n\in\mathbb{N}}, \Omega\right), \boldsymbol{\theta}\right\rangle.$$

Finally, using Lemma 1 leads to formula (4.82).

□

Remark 10. *Note that the variation $\boldsymbol{\theta}_p$ of the plastic zone, which is assumed to exist in assumption **P.1**, does not play any role in the shape derivative (4.82). Theorem 29 is a mathematically clean version of many results in the engineering literature (often in a discretized setting) where the equations are derived without taking into account the non-differentiability issues. Of course, our result relies on a very strong assumption which is the price to pay to deduce a formula as simple as (4.82).*

Remark 11. *In the time-discretized state equation (4.69), the variables $(\boldsymbol{u}_n, \boldsymbol{\varepsilon}_{p,n}, \gamma_n)$ are coupled. In the same spirit, in the adjoint equation (4.83), the variables $(\boldsymbol{v}_n, \boldsymbol{\varepsilon}_{q,n}, \mu_n)$ are coupled. This is unfortunate in order to resolve the system numerically, one ends up with a large system of equations.*

4.4 Numerical Implementation

In this section, we talk about the numerical aspects of the resolution of the state and the adjoint equations, along with shape optimization algorithm.

We first discuss the numerical resolution of the plasticity problem using state equation (4.56) in Subsection 4.4.1. In particular, we talk about the high numerical resolution costs involved. We then present and justify using the radial return algorithm for the plasticity problem in Subsection 4.4.2. We then briefly talk about resolution of the adjoint equation in Subsection 4.4.3. And finally present the shape optimization algorithm in Subsections 4.4.4 and 4.4.5.

The domain Ω is discretized using a simplicial unstructured mesh and the space Z, defined by (4.29), is discretized as Z^h, using the finite element framework

$$Z^h = \mathbb{P}^1(\Omega)^d \times \mathbb{P}^0(\Omega)^{d\times d} \times \mathbb{P}^0(\Omega). \tag{4.90}$$

Here, we have discretized the continuous space Q (defined in (4.28)) using $\mathbb{P}^0(\Omega)^{d\times d}$. The functions (tensors) $\boldsymbol{\tau} \in \mathbb{P}^0(\Omega)^{d\times d}$ are not trace-free (unlike the functions in Q). In order to incorporate the trace-free condition numerically, we work only with the deviator of the functions in $\mathbb{P}^0(\Omega)^{d\times d}$. The space K is discretized as K^h, defined by

$$K^h = \left\{(\boldsymbol{u}, \boldsymbol{\varepsilon}_{P_p las}, \gamma) \in Z^h,\ \sqrt{\frac{2}{3}}|\boldsymbol{\varepsilon}_{P_p las}| \le \gamma \text{ a.e. in } \Omega\right\}. \tag{4.91}$$

The maximal mesh size is denoted by $h_{\max}$, the minimal mesh size by $h_{\min}$ and the number of mesh vertices is N_v. We assume the mesh to be regular, or $h_{\max}$ and $h_{\min}$ to be of the same order. The space-time discretized state solution is $\tilde{\boldsymbol{w}}(t) \in Z^h$ and the space-time discretized adjoint solution is $\tilde{\boldsymbol{z}}(t) \in Z^h$. The time interval $[0, T]$ is discretized in N intervals of length δt. We label the time at the end of n-th time interval as $t_n,\ n = 1, 2, \cdots, N$. All of our numerical experiments are performed on an open-source software `FreeFEM` [92]. For all the numerical simulations in this section and the next Section 4.5, we consider mild-steel, with the properties given in Table 4.1. The value of the kinematic hardening tensor $\mathbb{H}$ shall be specified for every test case. For all test cases, the body force $\boldsymbol{f}(t)$ is assumed to be zero.

E	ν	σ_Y	E_{iso}
210GPa	0.3	279MPa	712MPa

Table 4.1: Mild steel properties

4.4.1 Resolution of the plasticity formulation

The space-discretized version of the problem (4.35) reads: find $\boldsymbol{w}^h \in Z^h$ such that

$$a(\boldsymbol{w}^h, \boldsymbol{z}^h - \dot{\boldsymbol{w}}^h) + j(\boldsymbol{z}^h) - j(\dot{\boldsymbol{w}}^h) \geq \boldsymbol{l}_t(\boldsymbol{z}^h - \dot{\boldsymbol{w}}^h) \qquad \forall \boldsymbol{z}^h \in K^h. \tag{4.92}$$

and the space-discretized version of the problem (4.56) reads: find $\boldsymbol{w}^h_{\epsilon,\eta} \in Z^h$ such that

$$a(\boldsymbol{w}^h_{\epsilon,\eta}, \boldsymbol{z}^h) + \langle \nabla_z \mathcal{D}_{\epsilon,\eta}(\dot{\boldsymbol{w}}^h_{\epsilon,\eta}), \boldsymbol{z}^h \rangle = \langle \boldsymbol{l}^h, \boldsymbol{z}^h \rangle \qquad \forall \boldsymbol{z}^h \in Z^h. \tag{4.93}$$

Let $\tilde{\boldsymbol{w}}(t)$ be the approximate solution to the time-discretized version of (4.93). We let the value of $\tilde{\boldsymbol{w}}(t)$ at every time instant t_n be $\tilde{\boldsymbol{w}}_n$ and its time derivative be $\delta\tilde{\boldsymbol{w}}_n$ and use the affine interpolation

$$\tilde{\boldsymbol{w}}(t) = \tilde{\boldsymbol{w}}_n + \delta\tilde{\boldsymbol{w}}_n(t - t_n)$$

for $t \in [t_n, t_{n+1}]$. We solve problem (4.93) for $\delta\tilde{\boldsymbol{w}}_n$ using an explicit time integration scheme

$$\kappa a(\delta\tilde{\boldsymbol{w}}_n, \boldsymbol{z}) + \langle \nabla_z \mathcal{D}_{\epsilon,\eta}(\delta\tilde{\boldsymbol{w}}_n), \boldsymbol{z} \rangle = \langle \boldsymbol{l}^h, \boldsymbol{z} \rangle - a(\tilde{\boldsymbol{w}}_n, \boldsymbol{z}) \qquad \forall \boldsymbol{z} \in Z^h. \tag{4.94}$$

The above problem is no longer quasi-static in nature as there is dependence on the time interval length κ. Instead of determining the value of $\tilde{\boldsymbol{w}}(t)$ at every time instant, we determine the increments $\kappa\delta\tilde{\boldsymbol{w}}_n$. Since the nonlinearity $j_{\epsilon,\eta}(\cdot)$ is $\mathcal{C}^\infty$ smooth in its argument, the Newton-Raphson algorithm can be applied to (4.94). The choice of the parameters $\epsilon,\ \eta$ determine the ease of convergence of Newton-Raphson. If the parameters are too small, the algorithm might not converge at all. We want to choose the parameters as big as possible and yet not compromise on the error in the discretized solution $\tilde{\boldsymbol{w}}(t)$. We thus need an estimate of the error arising from the penalization and regularization. To this effect, we consider the solution $\boldsymbol{w}^h_{\epsilon,\eta}$ to the equation (4.93) and ignore the errors arising from the time-discretizing the equation (4.93). The Theorem 24 give the desired estimate.

Theorem 24. *Let $\boldsymbol{w} \in Z$ be the solution to the plasticity problem (4.35) and let $\boldsymbol{w}^h_{\epsilon,\eta} = (\boldsymbol{u}^h, \boldsymbol{\varepsilon}^h_p, \gamma^h) \in \mathbb{P}^1(\Omega^h)^d \times \mathbb{P}^0(\Omega^h)^{d\times d} \times \mathbb{P}^0(\Omega^h) = Z^h$ be the solution to the problem (4.93). Then the following error estimate holds*

$$\left\|\boldsymbol{w} - \boldsymbol{w}^h_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} \leq C_1\sqrt{\frac{h_{\max}}{L}} + C_2\sqrt{\frac{\epsilon}{h_{\max}^{d/2}}} + C_3\sqrt{\frac{\eta}{\epsilon}}, \tag{4.95}$$

where C_1, C_2, C_3 are constants that dependent on the geometry and the material, L is the characteristic length of the structure and d is the dimension of the problem.

We first mention a useful Lemma 12 and then present proof to the above theorem. The proof is technical, but the idea of the proof is simple. We split the difference $\boldsymbol{w} - \boldsymbol{w}^h_{\epsilon,\eta}$ as

$$\left\|\boldsymbol{w} - \boldsymbol{w}^h_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)} \leq \left\|\boldsymbol{w} - \boldsymbol{w}^h\right\|_{L^\infty([0,T],Z)} + \left\|\boldsymbol{w}^h - \boldsymbol{w}^h_{\epsilon,\eta}\right\|_{L^\infty([0,T],Z)}$$

where $\boldsymbol{w}^h \in Z^h$ is the solution to (4.92). A bound on the term $\boldsymbol{w} - \boldsymbol{w}^h$ has in been derived in [87]. We derive a bound for the remaining part in the proof below. The solution $\boldsymbol{w}^h$ satisfies a few properties that shall be used for the Theorem 24 later.

Lemma 12. *Let $\boldsymbol{w}^h(t) \in Z^h$ be the solution to the problem (4.92). Then the following holds:*

$$\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}^h) - \boldsymbol{\varepsilon}_p^h) : \dot{\boldsymbol{\varepsilon}}_p^h \, dx - \int_\Omega \mathbb{H}\boldsymbol{\varepsilon}_p : \dot{\boldsymbol{\varepsilon}}_p^h dx - \int_\Omega E_{iso}\gamma^h \dot{\gamma}^h dx - j(\dot{\boldsymbol{w}}^h) = 0.$$

Proof. We substitute $\boldsymbol{z} = (\dot{\boldsymbol{u}}^h, \boldsymbol{0}, 0) \in K^h$ in (4.92) and obtain

$$\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}^h) - \boldsymbol{\varepsilon}_p^h) : \dot{\boldsymbol{\varepsilon}}_p^h \, dx - \int_\Omega \mathbb{H}\boldsymbol{\varepsilon}_p : \dot{\boldsymbol{\varepsilon}}_p^h dx - \int_\Omega E_{iso}\gamma^h \dot{\gamma}^h dx - j(\dot{\boldsymbol{w}}^h) \geq 0. \tag{4.96}$$

We then substitute $\boldsymbol{z} = (\dot{\boldsymbol{u}}^h, 2\dot{\boldsymbol{\varepsilon}}_p^h, 2\dot{\gamma}^h) \in K^h$ in (4.92) and obtain

$$-\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}^h) - \boldsymbol{\varepsilon}_p^h) : \dot{\boldsymbol{\varepsilon}}_p^h \, dx + \int_\Omega \mathbb{H}\boldsymbol{\varepsilon}_p : \dot{\boldsymbol{\varepsilon}}_p^h dx + \int_\Omega E_{iso}\gamma^h \dot{\gamma}^h dx + j(2\dot{\boldsymbol{w}}^h) - j(\dot{\boldsymbol{w}}^h) \geq 0.$$

Using the positive homogenity of $j(\cdot)$ and multipliying by -1 on both sides in the above, we get

$$\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}^h) - \boldsymbol{\varepsilon}_p^h) : \dot{\boldsymbol{\varepsilon}}_p^h \, dx - \int_\Omega \mathbb{H}\boldsymbol{\varepsilon}_p : \dot{\boldsymbol{\varepsilon}}_p^h dx - \int_\Omega E_{iso}\gamma^h \dot{\gamma} dx - j(\dot{\boldsymbol{w}}^h) \leq 0. \tag{4.97}$$

From (4.96) and (4.97), we deduce

$$\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}^h) - \boldsymbol{\varepsilon}_p^h) : \dot{\boldsymbol{\varepsilon}}_p^h \, dx - \int_\Omega \mathbb{H}\boldsymbol{\varepsilon}_p : \dot{\boldsymbol{\varepsilon}}_p^h dx - \int_\Omega E_{iso}\gamma^h \dot{\gamma}^h dx - j(\dot{\boldsymbol{w}}^h) = 0.$$

We remark that the above can also be seen as the space-discretized form of (4.6) integrated over Ω. □

We now demonstrate the proof to the Theorem 24.

Proof. We consider the space discretized version of the problem (4.46) (equivalent to (4.93)): find $\boldsymbol{w}_{\epsilon,\eta}^h(t) \in Z^h$ s.t.

$$a(\boldsymbol{w}^h, \boldsymbol{z} - \dot{\boldsymbol{w}}^h) + j_{\epsilon,\eta}(\boldsymbol{z}) - j_{\epsilon,\eta}(\dot{\boldsymbol{w}}^h) \geq l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}^h) \qquad \forall \boldsymbol{z} \in Z^h, \tag{4.98}$$

and the space discretized version of the problem (4.39): find $\boldsymbol{w}_\epsilon^h(t) = (\boldsymbol{u}_\epsilon^h, \boldsymbol{\varepsilon}_{p,\epsilon}^h, \gamma_\epsilon^h)(t) \in K^h$ s.t.

$$a(\boldsymbol{w}_\epsilon^h, \boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon^h) + j_\epsilon(\boldsymbol{z}) - j_\epsilon(\dot{\boldsymbol{w}}_\epsilon^h) \geq l_t(\boldsymbol{z} - \dot{\boldsymbol{w}}_\epsilon^h) \qquad \forall \boldsymbol{z} \in Z^h. \tag{4.99}$$

The above problem is well posed and admits a unique solution $\boldsymbol{w}_\epsilon^h \in H^1([0,1], Z^h)$ using Theorem 15 with Z replaced by Z^h. We use the two solutions $\boldsymbol{w}^h, \boldsymbol{w}_\epsilon^h$ along with the triangular inequality to split $\boldsymbol{w} - \boldsymbol{w}_{\epsilon,\eta}^h$ as

$$\left\|\boldsymbol{w} - \boldsymbol{w}_{\epsilon,\eta}^h\right\|_{L^\infty([0,T],Z)} \leq \left\|\boldsymbol{w} - \boldsymbol{w}^h\right\|_{L^\infty([0,T],Z)} + \left\|\boldsymbol{w}^h - \boldsymbol{w}_\epsilon^h\right\|_{L^\infty([0,T],Z)} + \left\|\boldsymbol{w}_\epsilon^h - \boldsymbol{w}_{\epsilon,\eta}^h\right\|_{L^\infty([0,T],Z)}. \tag{4.100}$$

The right hand side in the above has three terms. We will determine a bound for each. As shown in [87], the Z^h finite element approximation of the function $\boldsymbol{w}$ satisfies

$$\left\|\boldsymbol{w} - \boldsymbol{w}^h\right\|_{L^\infty([0,T],Z)} \leq C_1 \sqrt{\frac{h_{\max}}{L}},$$

where C_1 is a constant. Now we consider the second term in the right hand side of (4.100) multiplied by $\alpha/2$, (α being the coercivity constant $a(\cdot,\cdot)$)

$$\begin{aligned}
\frac{\alpha}{2}\left\|\boldsymbol{w}^h-\boldsymbol{w}_\epsilon^h\right\|_Z^2(t_0) &= \frac{\alpha}{2}\int_0^{t_0}\frac{d}{dt}\left\|\boldsymbol{w}^h-\boldsymbol{w}_\epsilon^h\right\|_Z^2 dt \qquad \text{(using } \boldsymbol{w}(t)=\boldsymbol{w}^h(t)=0 \text{ at } t=0)\\
&\le \frac{1}{2}\int_0^{t_0}\frac{d}{dt}a(\boldsymbol{w}^h-\boldsymbol{w}_\epsilon^h,\boldsymbol{w}^h-\boldsymbol{w}_\epsilon^h)dt\\
&= \int_0^{t_0}a(\boldsymbol{w}^h-\boldsymbol{w}_\epsilon^h,\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)dt\\
&= \int_0^{t_0}\left(a(\boldsymbol{w}^h,\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)-a(\boldsymbol{w}_\epsilon^h,\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)\right)dt\\
&\le \int_0^{t_0}\Big(a(\boldsymbol{w}^h,\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)+j_\epsilon(\dot{\boldsymbol{w}}^h)-j_\epsilon(\dot{\boldsymbol{w}}_\epsilon^h)\\
&\qquad -l_t(\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)\Big)dt \qquad \text{(using (4.99) with } \boldsymbol{z}=\dot{\boldsymbol{w}}^h\in Z^h\ )\\
&\le \int_0^{t_0}\left(a(\boldsymbol{w}^h,\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)+j(\dot{\boldsymbol{w}}^h)-j_\epsilon(\dot{\boldsymbol{w}}_\epsilon^h)-l_t(\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)\right)dt \quad \text{(since } \dot{\boldsymbol{w}}^h\in K^h)\\
&= \int_0^{t_0}\int_\Omega\left(\boldsymbol{\sigma}^h\colon(\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^h-\dot{\boldsymbol{u}}_\epsilon^h)-\dot{\boldsymbol{\varepsilon}}_p^h+\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h)+E_{iso}\gamma^h(\dot\gamma^h-\dot\gamma_\epsilon^h)+\mathbb{H}\boldsymbol{\varepsilon}_p^h\colon(\dot{\boldsymbol{\varepsilon}}_p^h-\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h)\right)dx\ dt\\
&\quad+\int_0^{t_0}\left(j(\dot{\boldsymbol{w}}^h)-j_\epsilon(\dot{\boldsymbol{w}}_\epsilon^h)-l_t(\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h)\right)dt \quad \text{(where } \boldsymbol{\sigma}^h=\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}^h)-\boldsymbol{\varepsilon}_p^h)).
\end{aligned}$$

Since

$$\int_\Omega\boldsymbol{\sigma}^h\colon\boldsymbol{\varepsilon}(\dot{\boldsymbol{u}}^h-\dot{\boldsymbol{u}}_\epsilon^h)dx=\int_\Omega f\cdot(\dot{\boldsymbol{u}}^h-\dot{\boldsymbol{u}}_\epsilon^h)dx+\int_{\Gamma_N}\boldsymbol{g}\cdot(\dot{\boldsymbol{u}}^h-\dot{\boldsymbol{u}}_\epsilon^h)ds=l_t(\dot{\boldsymbol{w}}^h-\dot{\boldsymbol{w}}_\epsilon^h),$$

we get

$$\begin{aligned}
\frac{\alpha}{2}\left\|\boldsymbol{w}^h-\boldsymbol{w}_\epsilon^h\right\|_Z^2(t_0)+\int_0^{t_0}j_\epsilon(\dot{\boldsymbol{w}}_\epsilon^h)dt &\le \int_0^{t_0}\int_\Omega\left(\boldsymbol{\sigma}^h\colon(\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h-\dot{\boldsymbol{\varepsilon}}_p^h)+E_{iso}\gamma^h(\dot\gamma^h-\dot\gamma_\epsilon^h)+\mathbb{H}\boldsymbol{\varepsilon}_p^h\colon(\dot{\boldsymbol{\varepsilon}}_p^h-\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h)\right)dx\ dt\\
&\quad+\int_0^{t_0}j(\dot{\boldsymbol{w}}^h)dt\\
&= \int_0^{t_0}\int_\Omega\left((\boldsymbol{\sigma}^h-\mathbb{H}\boldsymbol{\varepsilon}_p^h)\colon\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h-E_{iso}\gamma^h\dot\gamma_\epsilon^h\right)dx\ dt \qquad \text{(using Lemma (12))}\\
&= \int_0^{t_0}\int_\Omega\Big((\boldsymbol{\sigma}^h-\mathbb{H}\boldsymbol{\varepsilon}_p^h)^D\colon\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h\\
&\qquad -E_{iso}\gamma^h\dot\gamma_\epsilon^h\Big)dx\ dt \qquad \text{(since } \dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h \text{ is trace-free)}\\
&\le \int_0^{t_0}\int_\Omega\left(\left|(\boldsymbol{\sigma}^h-\mathbb{H}\boldsymbol{\varepsilon}_p^h)^D\right|\left|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h\right|-E_{iso}\gamma^h\dot\gamma_\epsilon^h\right)dx\ dt.
\end{aligned}$$

We use the split $j_\epsilon(\cdot)=j_1(\cdot)+\frac{1}{\epsilon}j_2(\cdot)$ defined in (4.41) and obtain a bound as

$$\begin{aligned}
&\frac{\alpha}{2}\left\|\boldsymbol{w}^h-\boldsymbol{w}_\epsilon^h\right\|_Z^2(t_0)+\frac{1}{\epsilon}\int_0^{t_0}j_2(\dot{\boldsymbol{w}}_\epsilon^h)dt\\
&\le \int_0^{t_0}\int_\Omega\left(\left|(\boldsymbol{\sigma}^h-\mathbb{H}\boldsymbol{\varepsilon}_p^h)^D\right|\left|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h\right|+E_{iso}\gamma^h\left(\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h|-\dot\gamma_\epsilon^h-\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h|\right)-j_1(\dot{\boldsymbol{w}}_\epsilon^h)\right)dx\ dt\\
&= \int_0^{t_0}\int_\Omega\left(\left|(\boldsymbol{\sigma}^h-\mathbb{H}\boldsymbol{\varepsilon}_p^h)^D\right|\left|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h\right|+E_{iso}\gamma^h\left(\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h|-\dot\gamma_\epsilon^h-\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h|\right)-\sqrt{\frac{2}{3}}\sigma_Y|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h|\right)dx\ dt\\
&= \int_0^{t_0}\int_\Omega\left(\left|(\boldsymbol{\sigma}^h-\mathbb{H}\boldsymbol{\varepsilon}_p^h)^D\right|-\sqrt{\frac{2}{3}}(\sigma_Y+E_{iso}\gamma^h)\right)\left|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h\right|+E_{iso}\gamma^h\left(\sqrt{\frac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_{p,\epsilon}^h|-\dot\gamma_\epsilon^h\right)dx\ dt.
\end{aligned}$$

Since the discrete solution $\boldsymbol{w}^h$ satisfies the von Mises criterion (consider the space-discretized version of the derivation of Subsection 4.1.2)

$$\left|(\boldsymbol{\sigma}^h-\mathbb{H}\boldsymbol{\varepsilon}_p^h)^D\right|-\sqrt{\frac{2}{3}}(\sigma_Y+E_{iso}\gamma^h)\le 0.$$

In addition, since $\dot{\gamma}^h(t) \in K^h,\ \dot{\gamma}^h \geq 0$ and $\gamma^h(0) = 0$ implies $\gamma^h \geq 0$, we have

$$\gamma^h \left(\sqrt{\frac{2}{3}} |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon \right) \leq \gamma^h \max \left(0, \sqrt{\frac{2}{3}} |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon \right).$$

We get the bound

$$\begin{aligned}
&\frac{\alpha}{2} \left\| \boldsymbol{w}^h - \boldsymbol{w}^h_\epsilon \right\|^2_Z (t_0) + \frac{1}{\epsilon} \int_0^{t_0} j_2(\dot{\boldsymbol{w}}^h_\epsilon) dt \\
&\quad \leq \int_0^{t_0} \int_\Omega E_{iso} \gamma^h \max \left(0, \sqrt{\frac{2}{3}} |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon \right) dx \ dt \\
&\quad \leq \int_0^{t_0} \left(\frac{\epsilon}{2s} \left\| E_{iso} \gamma^h \right\|^2_{L^2(\Omega)} + \frac{s}{2\epsilon} \left\| \max \left(0, \sqrt{\frac{2}{3}} |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon \right) \right\|^2_{L^2(\Omega)} \right) dt \qquad \text{for any } s > 0.
\end{aligned}$$

Since all norms are equivalent in finite dimension,

$$\left\| \max \left(0, \sqrt{\frac{2}{3}} |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon \right) \right\|_{L^2(\Omega)} \leq \frac{1}{h^{d/2}_{\min}} \left\| \max \left(0, \sqrt{\frac{2}{3}} |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon \right) \right\|_{L^1(\Omega)}.$$

Moreover, if ϵ is sufficiently small, we can safely assume

$$\left\| \max \left(0, \sqrt{\frac{2}{3}} |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon \right) \right\|_{L^1(\Omega)} < 1.$$

Using the two bounds and choosing $s = h^{d/2}_{\min}$, we get

$$\begin{aligned}
\frac{\alpha}{2} \left\| \boldsymbol{w}^h - \boldsymbol{w}^h_\epsilon \right\|^2_Z (t_0) + \frac{1}{\epsilon} \int_0^{t_0} j_2(\dot{\boldsymbol{w}}^h_\epsilon) dt &\leq \int_0^{t_0} \left(\frac{\epsilon}{2h^{d/2}_{\min}} \left\| E_{iso} \gamma^h \right\|^2_{L^2(\Omega)} + \frac{1}{2\epsilon} \left\| \max(0, |\dot{\boldsymbol{\varepsilon}}^h_{p,\epsilon}| - \dot{\gamma}^h_\epsilon) \right\|_{L^1(\Omega)} \right) dt \\
&= \int_0^{t_0} \left(\frac{\epsilon}{2h^{d/2}_{\min}} \left\| E_{iso} \gamma^h \right\|^2_{L^2(\Omega)} + \frac{1}{2\epsilon} j_2(\dot{\boldsymbol{w}}^h_\epsilon) \right) dt
\end{aligned}$$

We then get the bound

$$\begin{aligned}
\frac{\alpha}{2} \left\| \boldsymbol{w}^h - \boldsymbol{w}^h_\epsilon \right\|^2_Z (t_0) &\leq \frac{\alpha}{2} \left\| \boldsymbol{w}^h - \boldsymbol{w}^h_\epsilon \right\|^2_Z (t_0) + \frac{1}{2\epsilon} \int_0^{t_0} j_2(\dot{\boldsymbol{w}}^h_\epsilon) dt \\
&\leq \int_0^{t_0} \frac{\epsilon}{2h^{d/2}_{\min}} \left\| E_{iso} \gamma^h \right\|^2_{L^2(\Omega)} dt.
\end{aligned}$$

Since the finite element mesh is assumed to be regular, we let $h_{\min} = C_{\Omega_h} h_{\max}$ where C_{Ω_h} is a constant of order unity. Simplifying,

$$\left\| \boldsymbol{w}^h - \boldsymbol{w}^h_\epsilon \right\|_Z (t_0) \leq \left\| \boldsymbol{w}^h - \boldsymbol{w}^h_\epsilon \right\|_{L^\infty([0,T],Z)} \leq C_2 \sqrt{\frac{\epsilon}{h^{d/2}_{\max}}} \qquad \text{where} \qquad C_2 = \sqrt{\frac{1}{C^{d/2}_{\Omega_h} \alpha} \int_0^T \left\| E_{iso} \gamma^h \right\|^2_{L^2(\Omega)} dt}.$$

For the third term in the left hand side of (4.100), we perform a similar manipulation

$$\begin{aligned}
\frac{\alpha}{2}\left\|\boldsymbol{w}_{\epsilon}^{h}-\boldsymbol{w}_{\epsilon,\eta}^{h}\right\|_{Z}^{2}(t_0) &= \frac{\alpha}{2}\int_0^{t_0}\frac{d}{dt}\left\|\boldsymbol{w}_{\epsilon}^{h}-\boldsymbol{w}_{\epsilon,\eta}^{h}\right\|_{Z}^{2}dt \qquad (\text{using } \boldsymbol{w}_{\epsilon}^{h}=\boldsymbol{w}_{\epsilon,\eta}^{h}=0 \text{ at } t=0)\\
&\leq \int_0^{t_0}\frac{d}{dt}a(\boldsymbol{w}_{\epsilon}^{h}-\boldsymbol{w}_{\epsilon,\eta}^{h},\boldsymbol{w}_{\epsilon}^{h}-\boldsymbol{w}_{\epsilon,\eta}^{h})dt\\
&= \frac{1}{2}\int_0^{t_0}a(\boldsymbol{w}_{\epsilon}^{h}-\boldsymbol{w}_{\epsilon,\eta}^{h},\dot{\boldsymbol{w}}_{\epsilon}^{h}-\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})dt\\
&\leq \int_0^{t_0}\left(a(\boldsymbol{w}_{\epsilon}^{h},\dot{\boldsymbol{w}}_{\epsilon}^{h}-\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})-a(\boldsymbol{w}_{\epsilon,\eta}^{h},\dot{\boldsymbol{w}}_{\epsilon}^{h}-\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})\right)dt\\
&\leq \int_0^{t_0}\Big(j_{\epsilon}(\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})-j_{\epsilon}(\dot{\boldsymbol{w}}_{\epsilon}^{h})-l_t(\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h}-\dot{\boldsymbol{w}}_{\epsilon}^{h})\\
&\qquad -a(\boldsymbol{w}_{\epsilon,\eta}^{h},\dot{\boldsymbol{w}}_{\epsilon}^{h}-\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})\Big)dt \quad (\text{taking } \boldsymbol{z}=\boldsymbol{w}_{\epsilon,\eta}^{h} \text{ in (4.99)})\\
&\leq \int_0^{t_0}\left(j_{\epsilon}(\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})-j_{\epsilon}(\dot{\boldsymbol{w}}_{\epsilon}^{h})+j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})-j_{\epsilon,\eta}(\dot{\boldsymbol{w}}_{\epsilon,\eta}^{h})\right)dt \quad (\text{taking } \boldsymbol{z}=\boldsymbol{w}_{\epsilon}^{h} \text{ in (4.98)}).
\end{aligned}$$

Here, we recall the property: for all $\boldsymbol{w}\in Z^h$

$$|j_{\epsilon}(\boldsymbol{w})-j_{\epsilon,\eta}(\boldsymbol{w})|\leq \frac{\sigma_Y^2}{ET}\sqrt{\frac{2}{3}}\left(\eta+\frac{\eta}{\epsilon}\right)\int_{\Omega}dx.$$

Resulting in

$$\left\|\boldsymbol{w}_{\epsilon}^{h}-\boldsymbol{w}_{\epsilon,\eta}^{h}\right\|_{Z}(t_0)\leq\left\|\boldsymbol{w}_{\epsilon}^{h}-\boldsymbol{w}_{\epsilon,\eta}^{h}\right\|_{L^{\infty}([0,T],Z)}\leq C_3\sqrt{\left(\eta+\frac{\eta}{\epsilon}\right)} \quad \text{where} \quad C_3=\sqrt{\frac{2\sigma_Y^2}{E}\sqrt{\frac{2}{3}}\int_{\Omega}dx}.$$

Combining the three estimates, we have the desired result (4.95). □

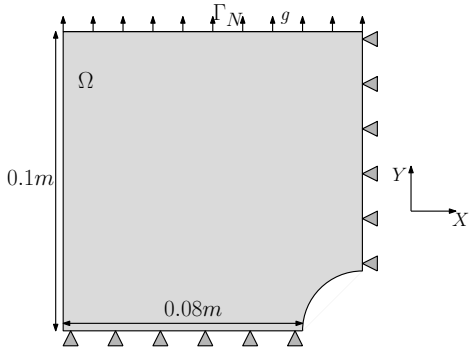
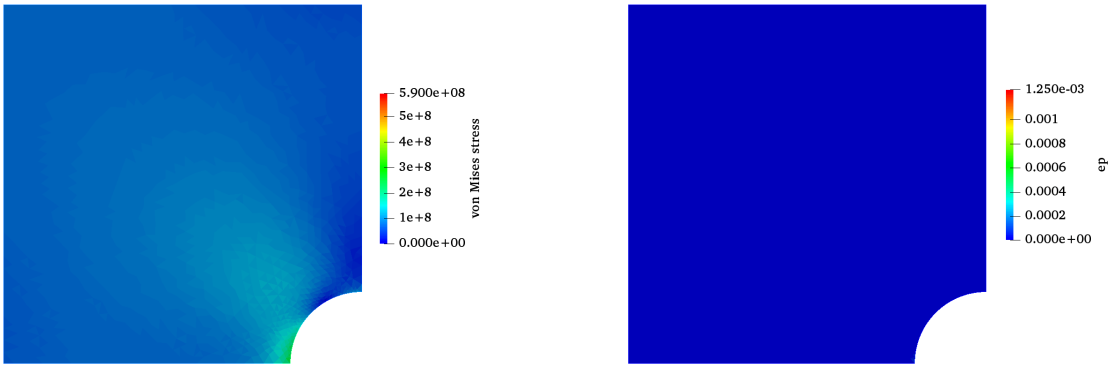


Figure 4.5: Square with a hole, treated in symmetry

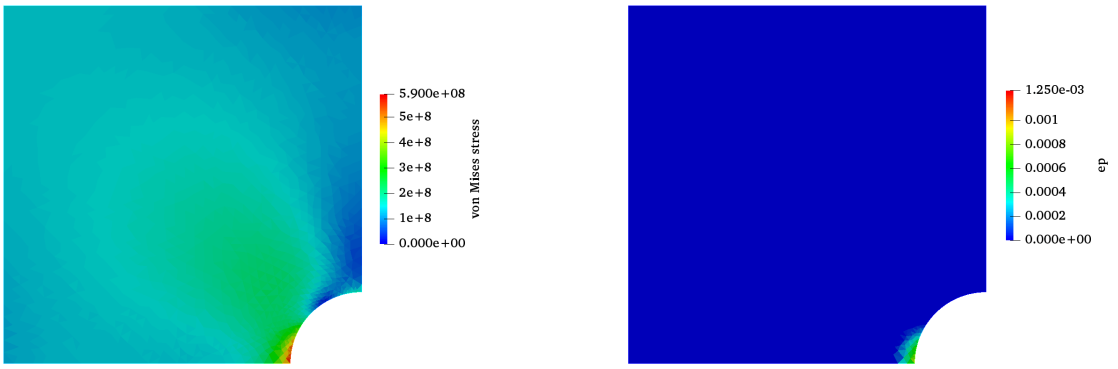
As the Theorem (24) shows, if one seeks an error of the order $\sqrt{\frac{h_{\max}}{L}}$, one ought to choose

$$\epsilon=\left(\frac{h_{\min}}{L}\right)^{1+d/2} \quad \text{and} \quad \eta=\epsilon^2. \tag{4.101}$$

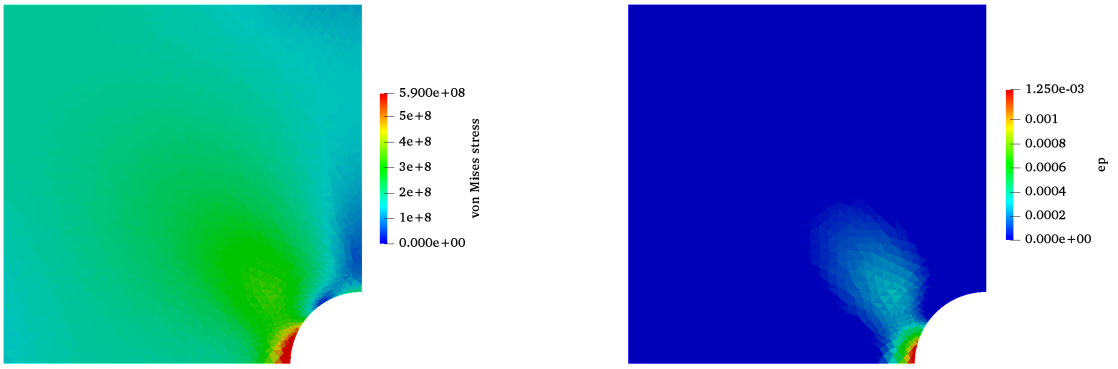
Having the right choice of ϵ and η, we now present a standard test-case of square with a hole in the middle subjected to tension. For the numerical simulation, we consider only one fourth of the square as shown in Fig.4.5. As can been seen in the Fig.4.5, the square is constrained to move along the horizontal on its vertical edge on the right, and constrained to move along the vertical along on its horizontal edge at the bottom. A monotonically increasing surface force $\boldsymbol{g}=(0,250t)\text{MN/m}$, $t\in[0,1]$ is applied. Using (4.101), we choose $\epsilon=2.5\times10^{-3},\eta=6.2\times10^{-6}$. The time step $\delta t=0.05$. The material properties shown in Table 4.1 are considered. We let $\mathbb{H}=0.5E\mathbb{I}\text{MPa}$. Figure 4.6 shows the von Mises stress and the y component of the plastic deformation $\boldsymbol{\varepsilon}_p$ computed at several time instants. As seen in the Fig.4.6, the regularization-penalization



(a) $t = 0.45$



(b) $t = 0.75$



(c) $t = 1$

Figure 4.6: von Mises stress $|\boldsymbol{\sigma}^D|$ on the left and y component of $\boldsymbol{\varepsilon}_p$ on the right, last time step $t = 1$ for $\epsilon = 2.5 \times 10^{-3}, \eta = 6.2 \times 10^{-6}$

error seem to be negligible, thanks to the appropriate choice (4.101). If these parameters were chosen larger, one might commit significant errors (see Fig.4.7). Here we make a very important observation. As seen in (4.101), the parameters ϵ and η might end up being too small if the mesh is too refined. Convergence of Newton-Raphson may not be then attained in a straight-forward way. The overall computational time is high because

- one ought to begin the Newton algorithm with a couple (ϵ, η) that is relatively big in magnitude, converge the Newton algorithm, decrease their magnitudes by a factor, and iterate
- the stiffness matrix size for resolution of (4.93) (with N_v vertices) is $10N_v \times 10N_v$ in $2D$ and $17N_v \times 17N_v$ in $3D$
- the stiffness matrix has penalization terms and thus ill-conditioned

Thus one is able to converge the Newton algorithm, except, at a high computational cost. In commercial software, the typical finite-element solver is based on the well-known radial-return algorithm. The radial-return algorithm is equivalent to resolving (4.92) via an implicit time integration scheme, resulting in a solution $\boldsymbol{w}_r \in Z^h$ [162]. This solution is the limit of $\boldsymbol{w}_{\epsilon,\eta}$ as $\epsilon \longrightarrow 0, \eta \longrightarrow 0$. Numerical resolution using radial-return algorithm takes far lesser time. Moreover, moreover $\boldsymbol{w}_r$ is not very different from $\boldsymbol{w}_{\epsilon,\eta}$ for well-chosen ϵ, η (see Fig.4.7). Hence, we prefer to stick to the radial return algorithm for all of our numerical experiments further.

4.4.2 Radial return algorithm

Radial return algorithm is most widely used numerical approach for solving the plasticity problem. This approach is very old, the first paper dating back to 1963, by Wilkins [185]. Initially proposed for perfect plasticity, this approach was later developed to incorporate hardening. The idea of radial return algorithm is to solve the plasticity equations by constructing a tangent matrix for the time-discretized implicit Euler scheme of (4.103) [165]. Here we briefly outline the simplest radial return algorithm, we implement on `FreeFEM`. An interested reader can refer to [162] for more details.

We re-write the dissipation energy function (4.21),

$$\mathcal{D}(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) = \min_{(\boldsymbol{\sigma},\boldsymbol{q},g)\in\mathbb{E}} ((\boldsymbol{\sigma}-\boldsymbol{q}) : \dot{\boldsymbol{\varepsilon}}_p + g\dot{\gamma})$$

The above minimization leads to

$$(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \in N_{\mathbb{E}}(\boldsymbol{\sigma}-\boldsymbol{q}, g),$$

(for a proof, see Theorem 13). Since the convex set $\mathbb{E}$ is defined by the von Mises function f, which is differentiable, the above leads to

$$\dot{\boldsymbol{\varepsilon}}_p = \zeta \frac{\partial f}{\partial \boldsymbol{\sigma}},\ \dot{\gamma} = \zeta \frac{\partial f}{\partial g},\ \zeta \geq 0 \text{ and } \zeta f = 0 \tag{4.102}$$

where ζ is the Lagrange multiplier and f is the von Mises yield function. For more details on the equivalence between (4.21) and (4.102), the reader may refer to [86], Chapter 4. The law (4.102) above is an associative flow law as the flow rates $\dot{\boldsymbol{\varepsilon}}_p$ and $\dot{\gamma}$ are derivatives of the yield function. On the contrary, if

$$\dot{\boldsymbol{\varepsilon}}_p = \zeta \boldsymbol{r}(\boldsymbol{\sigma}, \boldsymbol{q}),\ \dot{\gamma} = \zeta h(\boldsymbol{\sigma}, \boldsymbol{q}),\ \zeta \geq 0 \text{ and } \zeta f = 0$$

where the functions $\boldsymbol{r}(\boldsymbol{\sigma}, \boldsymbol{q})$ and $h(\boldsymbol{\sigma}, \boldsymbol{q})$ are determined experimentally and not necessarily related to the yield limit, then the law is called a non-associative flow law. They have been studied by J.L. Chaboche [43]. The complete set of equations for the associative flow is then given as

$$\begin{aligned}
-\text{div}(\boldsymbol{\sigma}) &= \boldsymbol{b} && \text{in } \Omega \times (0,T] \\
\boldsymbol{\sigma} &= \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p) && \text{in } \Omega \times (0,T] \\
f(\boldsymbol{\sigma}, \boldsymbol{q}, g) &= |\boldsymbol{\sigma}^D - \boldsymbol{q}^D| + \sqrt{\frac{2}{3}}(g - \sigma_Y) \leq 0 && \text{in } \Omega \times [0,T) \\
\boldsymbol{q} &= -\mathbb{H}\boldsymbol{\varepsilon}_p && \text{in } \Omega \times [0,T) \\
g &= -E_{iso}\gamma && \text{in } \Omega \times [0,T) \\
\dot{\gamma} &= \sqrt{\tfrac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p| && \text{in } \Omega \times [0,T) \\
\dot{\boldsymbol{\varepsilon}}_p &= \zeta \frac{\partial f}{\partial \boldsymbol{\sigma}},\ \dot{\gamma} = \zeta \frac{\partial f}{\partial g},\ \zeta \geq 0,\ f \leq 0 \text{ and } \zeta f = 0 && \text{in } \Omega \times [0,T) \\
\boldsymbol{u} &= 0 && \text{on } \Gamma_D \times [0,T) \\
\boldsymbol{\sigma}\boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N \times [0,T).
\end{aligned} \tag{4.103}$$

The system of equations (4.103) is equivalent to the system of equations (5.1) studied for the incremental approach in Section 5.1. The zone in Ω where $\zeta > 0$ is where the plastic flow rate $|\dot{\boldsymbol{\varepsilon}}_p| > 0$ and $\dot{\gamma} > 0$, and $f = 0$. In this zone, f can only remain zero. This is because if it decreases in value, we would have a positive plastic flow rate without attaining the yield limit. Hence, f must remain zero,

$$\dot{f}(\boldsymbol{\sigma}(t), \boldsymbol{q}(t), g(t)) = 0 \qquad \text{wherever} \qquad \zeta > 0.$$

This condition is known as the consistency condition.This condition helps to determine the Lagrange multiplier ζ explicitly. And using the expression for the multiplier ζ, one can determine an explicit relationship between $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\varepsilon}}(\boldsymbol{u})$. We first expand the consistency condition:

$$\dot{f}(\boldsymbol{\sigma},\boldsymbol{q},g)=0 \implies (\dot{\boldsymbol{\sigma}}^D-\dot{\boldsymbol{q}}^D):\partial_{\boldsymbol{\sigma}}f+\sqrt{\frac{2}{3}}\dot{g}=0 \qquad \text{where} \qquad \partial_{\boldsymbol{\sigma}}f=\frac{\partial f}{\partial\boldsymbol{\sigma}}=\frac{(\boldsymbol{\sigma}^D-\boldsymbol{q}^D)}{|\boldsymbol{\sigma}^D-\boldsymbol{q}^D|}. \tag{4.104}$$

Differentiating the stress-strain relations in (4.103) with respect to time and using the relations derived from (4.103):

$$\dot{\boldsymbol{q}}=\mathbb{H}\dot{\boldsymbol{\varepsilon}}_p, \qquad \dot{g}=-E_{iso}\zeta, \qquad \text{along with } \dot{\boldsymbol{\varepsilon}}_p=\zeta\partial_{\boldsymbol{\sigma}}f,$$

we get the expression for ζ

$$\zeta=\mathbb{Z}\;\partial_{\boldsymbol{\sigma}}f:\dot{\boldsymbol{\sigma}} \qquad \text{where} \qquad \mathbb{Z}=\left(\sqrt{\frac{2}{3}}E_{iso}+\mathbb{H}\partial_{\boldsymbol{\sigma}}f:\partial_{\boldsymbol{\sigma}}f\right)^{-1}. \tag{4.105}$$

Finally, we inject the above relation into the time derivative of $\boldsymbol{\sigma}$:

$$\begin{aligned}\dot{\boldsymbol{\sigma}}=\mathbb{C}\,(\dot{\boldsymbol{\varepsilon}}(\boldsymbol{u})-\dot{\boldsymbol{\varepsilon}}_p)&=\mathbb{C}\,(\dot{\boldsymbol{\varepsilon}}(\boldsymbol{u})-\zeta\partial_{\boldsymbol{\sigma}}f)\\&=\mathbb{C}\,(\dot{\boldsymbol{\varepsilon}}(\boldsymbol{u})-\mathbb{Z}\;\partial_{\boldsymbol{\sigma}}f:\dot{\boldsymbol{\sigma}}\;\partial_{\boldsymbol{\sigma}}f)=\mathbb{C}\,(\dot{\boldsymbol{\varepsilon}}(\boldsymbol{u})-\mathbb{Z}\;\partial_{\boldsymbol{\sigma}}f\otimes\partial_{\boldsymbol{\sigma}}f\;\dot{\boldsymbol{\sigma}})\,.\end{aligned}$$

Which on simplification results in

$$\dot{\boldsymbol{\sigma}}=\mathbb{D}\dot{\boldsymbol{\varepsilon}}(\boldsymbol{u}) \qquad \text{where} \qquad \mathbb{D}=(\mathbb{I}+\mathbb{C}\mathbb{Z}\;\partial_{\boldsymbol{\sigma}}f\otimes\partial_{\boldsymbol{\sigma}}f)^{-1}\,\mathbb{C} \qquad \text{if } \zeta>0 \tag{4.106a}$$

$$\dot{\boldsymbol{\sigma}}=\mathbb{C}\dot{\boldsymbol{\varepsilon}}(\boldsymbol{u}) \qquad \text{if } \zeta=0. \tag{4.106b}$$

The matrix $\mathbb{D}$ is the very well-known tangent matrix. Thus the consistency condition is satisfied if the relation (4.106) is holds. We now have all the ingredients to time-discretize the system (4.103) using an implicit Euler scheme, and solve it using the following algorithm 4. This is known as the radial return

Algorithm 4 Radial return algorithm

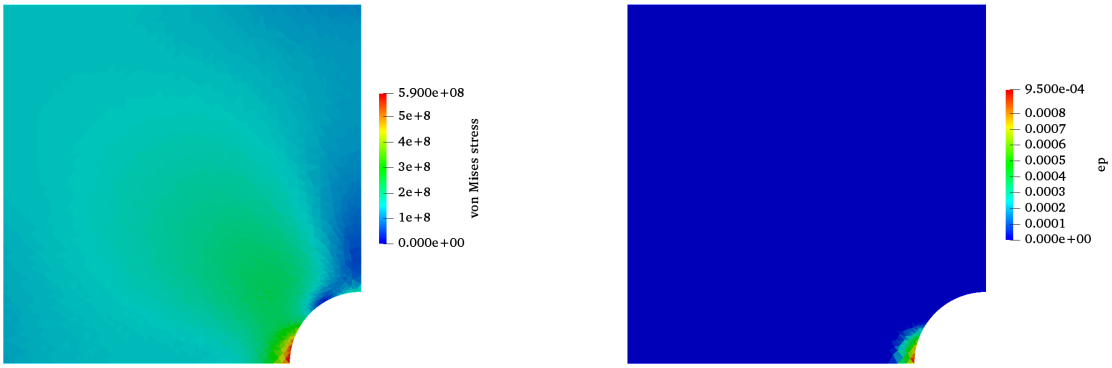
At a time instant t_i, let $\Delta\boldsymbol{\varepsilon}$ be a test strain,
For $n=1,\cdots,$ (repeat the following until convergence)

- Algebraic update of functions
  1. $\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1})=\boldsymbol{\varepsilon}(\boldsymbol{u}_n)+\Delta\boldsymbol{\varepsilon}$
  2. $\boldsymbol{\sigma}^D_{n+1}=\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1})^D-\boldsymbol{\varepsilon}_{p,n}$
  3. $\hat{\boldsymbol{n}}=\boldsymbol{\sigma}^D_{n+1}-\boldsymbol{q}^D_n$
  4. Compute $f=f(\boldsymbol{\sigma}_{n+1},\boldsymbol{q}_n,-E_{iso}\gamma_n)$
  5. For the zone where $f\geq 0$, compute ζ using 4.105 and set $\zeta=0$ in the zone where $f<0$
  6. $\boldsymbol{\varepsilon}_{p,n+1}=\boldsymbol{\varepsilon}_{p,n}+\zeta\hat{\boldsymbol{n}}$
  7. $\boldsymbol{q}_{n+1}=\boldsymbol{q}_n+\zeta\mathbb{H}\hat{\boldsymbol{n}}$
  8. $\gamma_{n+1}=\gamma_n+\sqrt{\frac{2}{3}}|\hat{\boldsymbol{n}}|$
  9. $\boldsymbol{\sigma}_{n+1}=\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1})-\boldsymbol{\varepsilon}_{p,n+1})$
- Compute $\mathbb{D}$ using (4.106) and solve for $\delta\boldsymbol{u}$

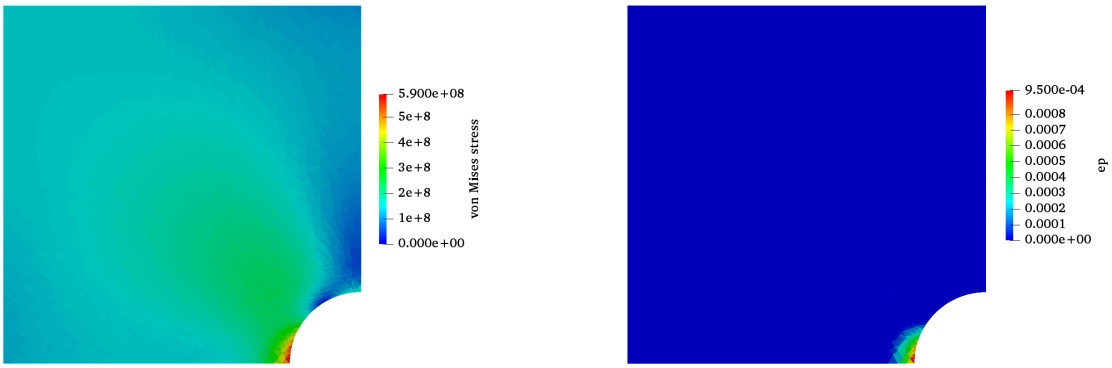
$$\int_\Omega \mathbb{D}\boldsymbol{\varepsilon}(\delta\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{\varphi})dx=\int_\Omega\boldsymbol{\sigma}_{n+1}:\boldsymbol{\varepsilon}(\boldsymbol{\varphi})dx-\int_{\Gamma_N}\boldsymbol{g}(t_i)\cdot\boldsymbol{\varphi}\,ds \qquad \forall\boldsymbol{\varphi}\in\mathbb{P}^1(\Omega)^d$$

- If $\|\delta\boldsymbol{u}\|_{L^2(\Omega)}\leq tol$, exit,
  else set $\Delta\boldsymbol{\varepsilon}=\boldsymbol{\varepsilon}(\delta\boldsymbol{u})$, $n=n+1$

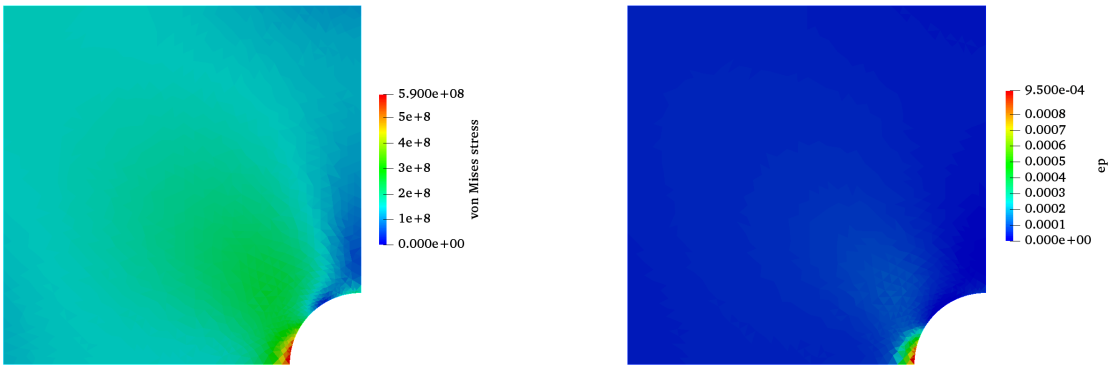
algorithm. The iterations in algorithm 4 can be shown to converge with a superlinear rate.



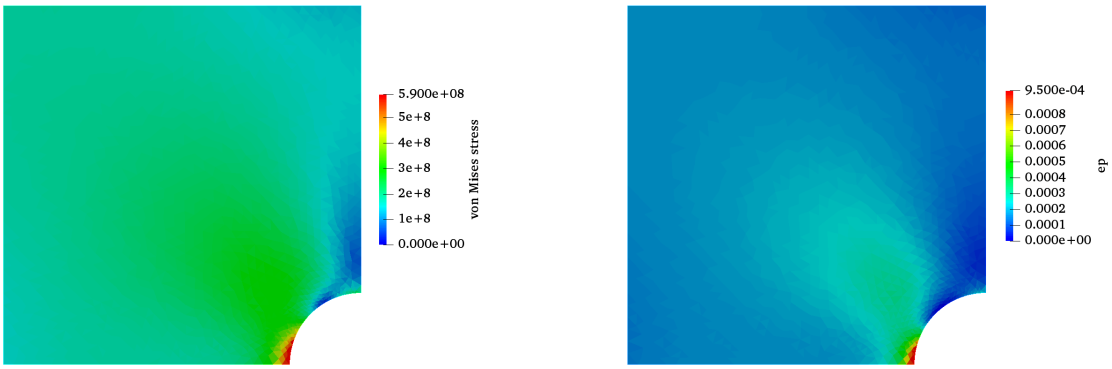
(a) Radial return algorithm



(b) $\epsilon = 2.5 \times 10^{-3}, \eta = 6.2 \times 10^{-6}$



(c) $\epsilon = 3.9 \times 10^{-2}, \eta = 1.59 \times 10^{-3}$



(d) $\epsilon = 7.9 \times 10^{-2}, \eta = 6.3 \times 10^{-3}$

Figure 4.7: von Mises stress $|\boldsymbol{\sigma}^D|$ on the left and second component of $\boldsymbol{\varepsilon}_p$ on the right, last time step $t = 1$

The radial return algorithm ensures that the consistency condition written with respect to the time-

continuous problem is satisfied [138]. One can also perform the analysis as in (4.104)-(4.106) for the time-discretized plasticity problem and end up with slightly different tangent matrix D_n. This tangent matrix was derived in [164] and with it, the radial return algorithm was shown to converge with a quadratic speed.

In the succeeding subsections, we denote the solution (primal variables) obtain via radial return algorithm by $\boldsymbol{w}_r(t)$ to distinguish it from the solution obtained by solving (4.93). From the algorithm 4, one can easily verify that the condition $\Delta\gamma_n = \sqrt{\frac{2}{3}}|\Delta\boldsymbol{\varepsilon}_{p,n}|$ is satisfied. Hence the solution $\boldsymbol{w}_r(t_n) \in K^h$ (see Def.(4.91)).

We now compare the solution obtained via radial return to the solution obtained for the penalized regularized formulation (4.93). For the comparison, we shall consider the standard test-case of square with a hole in the middle subjected to tension (same as in Fig.4.5, considered in the previous subsection). A monotonically increasing force $\boldsymbol{g} = [0, 200t]\text{MN/m}$, $t \in [0, 1]$ is applied on the topmost edge. The time step $\delta t = 0.05$. The material properties in Table 4.1 are considered along with $\mathbb{H} = 0.5E\mathbb{I}\text{MPa}$. In Fig.4.7, we plot the von Mises stress along with the plastic strain (y component) at the last time instant $t = 1$, obtained via radial return algorithm. We also plot the same obtained by resolving the Eqn.4.93 for different choices of η, ϵ. We do this in order to see the convergence of the approximate solution as the regularization and penalization parameters tend to zero.

In Fig.4.7, we indeed see that when ϵ is relatively large (3.9×10^{-2} or 7.9×10^{-2}), the error in the plastic strain is non-negligible and one can clearly make out the big differences with respect to the solution obtained via radial return algorithm. However, for $\epsilon = 2.5 \times 10^{-3}, \eta = 6.2 \times 10^{-6}$ (the choice (4.101)), the solution seems to be identical to the one obtained by radial return algorithm. Thus justifying using radial return algorithm instead of the approximation (4.93).

4.4.3 Resolution of adjoint system

We denote by $\tilde{\boldsymbol{z}}_n = \tilde{\boldsymbol{z}}(t_n)$ the discrete values of the adjoint, which is linearly interpolated in time on each sub-interval. We further discretize in space the time-discrete adjoint system (4.83) which was studied in Subsection 4.3.1 (and proved to be well-posed). The space-time discretized adjoint problem is defined by: $\tilde{\boldsymbol{z}}_N = 0$ and, for $n = N-1, \cdots, 1, 0$, find the solution $\tilde{\boldsymbol{z}}_n \in Z^h$ of

$$\begin{aligned}\delta t\langle \nabla_z m(\tilde{\boldsymbol{w}}_{n+1}), \boldsymbol{\varphi}\rangle + \delta t \int_{\Gamma_N} \nabla_z p(\tilde{\boldsymbol{w}}_{n+1})\boldsymbol{\varphi}\, ds + \delta t a(\boldsymbol{\varphi}, \tilde{\boldsymbol{z}}_n) \\ + \left\langle \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\delta\tilde{\boldsymbol{w}}_n)\tilde{\boldsymbol{z}}_n - \nabla_z^2 \mathcal{D}_{\epsilon,\eta}(\delta\tilde{\boldsymbol{w}}_{n+1})\tilde{\boldsymbol{z}}_{n+1}, \boldsymbol{\varphi}\right\rangle = 0 \qquad \forall \boldsymbol{\varphi} \in Z^h. \quad (4.107)\end{aligned}$$

This system is going backward in time. One ought to solve the state equation (4.92) until the last time step, store the solutions $\tilde{\boldsymbol{w}}_n$ for every time-step and retrieve the solutions one by one starting from the last time step. This is thus quite heavy in terms of memory requirement for numerical simulations. Finally, the time discretized shape derivative reads

$$\begin{aligned}J'(\Omega)(\boldsymbol{\theta}) = \sum_{n=0}^{N-1} \delta t \int_\Gamma \boldsymbol{\theta}\cdot\boldsymbol{n} \Big(m(\tilde{\boldsymbol{w}}_n) + \mathbb{C}(\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_{n+1}) - \tilde{\boldsymbol{\varepsilon}}_{p,n+1}) : (\boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}}_n) - \tilde{\boldsymbol{\varepsilon}}_{q,n}) + \mathbb{H}\tilde{\boldsymbol{\varepsilon}}_{p,n+1} : \tilde{\boldsymbol{\varepsilon}}_{q,n} \\ + E_{iso}\tilde{\gamma}_{n+1}\tilde{\mu}_n + \nabla_z \mathcal{D}_{\epsilon,\eta}\left(\frac{\tilde{\boldsymbol{w}}_{n+1} - \tilde{\boldsymbol{w}}_n}{\delta t}\right)\tilde{\boldsymbol{z}}_n - \boldsymbol{f}(t_n)\cdot\tilde{\boldsymbol{v}}_n \Big)\, ds, \end{aligned} \quad (4.108)$$

where $(\tilde{\boldsymbol{u}}_n, \tilde{\boldsymbol{\varepsilon}}_{p,n}, \tilde{\gamma}_n) = \tilde{\boldsymbol{w}}_n$ and $(\tilde{\boldsymbol{v}}_n, \tilde{\boldsymbol{\varepsilon}}_{q,n}, \tilde{\mu}_n) = \tilde{\boldsymbol{z}}_n$.

In numerical practice, we choose L as the characteristic length of the domain D (instead of the structure Ω as in Theorem 24) and choose the values of ϵ, η for penalization and regularization according to (4.101).

Remark 12. *As mentioned in Subsection 4.4.2, we replace $\tilde{\boldsymbol{w}}_n$ in the adjoint equation (4.107) and in the shape derivative (4.108) by the solution obtained via radial return, $\boldsymbol{w}_r(t_n) \in K^h$, which does not take into account the penalization and regularization. In formula (4.108) of the shape derivative, and more precisely in the term $\nabla_z \mathcal{D}_{\epsilon,\eta}$, we neglect the contribution $\frac{1}{\epsilon}M'_\eta(\cdot)$. The reason for this is because we replace the penalized solution $\tilde{\boldsymbol{w}}(t)$ by the non-penalized one $\boldsymbol{w}_r(t)$. For the penalized solution, the contribution $\frac{1}{\epsilon}M'_\eta\left(\frac{\tilde{\boldsymbol{w}}_{n+1} - \tilde{\boldsymbol{w}}_n}{\delta t}\right)$ is of order $\mathcal{O}(1)$ since it satisfies the problem (4.56). However, the same term is of order $\mathcal{O}(1/\epsilon)$ for the non-penalized solution because the regularization $M_\eta(s)$ of $\max(0, s)$ is not exactly zero for negative values of s as shown in Fig.4.8. To avoid this numerical artifact we found it more efficient to just cancel this term in (4.108).*

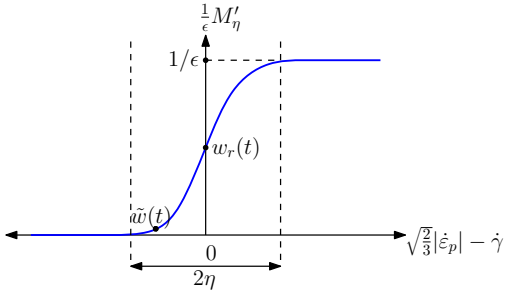


Figure 4.8: Regularized Heaviside function

4.4.4 Regularization and extension of the shape derivative

As usual, we resort to the level-set method, described in Section 2.2, Chapter 2. We need to regularize the shape derivative. The optimal shape we look for may not have a smooth boundary. Hence neither the shape derivative nor the normal to the boundary Γ may be smooth enough. In such a case, it is imperative to regularize the shape derivative [40], [8] in such a way that the decent direction is still respected. One possibility is to consider the H^1 scalar product instead of the L^2 scalar product by finding a function $dj(\Omega) \in H^1(D)$ such that

$$\int_D \left(h_{\min}^2 \nabla dj(\Omega) \cdot \nabla \varphi + dj(\Omega)\varphi\right) dx = \int_\Gamma j(\Omega)\, \varphi \, dx \qquad \forall \varphi \in H^1(D), \tag{4.109}$$

where $h_{\min}$ is the fixed minimal mesh size, and the function $j(\Omega)$ is defined by formula (4.108) with

$$J'(\Omega)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta} \cdot \boldsymbol{n}\, j(\Omega)\, ds. \tag{4.110}$$

Since we have chosen $\mathbb{P}^1$ basis elements for the displacement vector and the plastic strain, the shape derivative in (4.108) is $\mathbb{P}^0$ smooth and so $j(\Omega) \in \mathbb{P}^0(\Omega)$. Thus, it is enough to discretize (4.109) with $\mathbb{P}^1$ finite elements, so that $dj(\Omega) \in \mathbb{P}^1(D)$.

4.4.5 Shape optimization algorithm

We consider the shape optimization problem

$$\min_{\Omega \in \mathcal{U}_{ad}} J(\Omega),$$

where we remind the reader that $\mathcal{U}_{ad}$ is the space of admissible spaces inside the design space D (see Fig.2.4). In order to devise an optimization strategy taking the volume constraint into account, we introduce a Lagrangian $\mathcal{L}(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}, \Omega, \lambda)$ defined as

$$\mathcal{L}(\tilde{\boldsymbol{w}}, \tilde{\boldsymbol{z}}, \Omega, \lambda) = \frac{J(\Omega)}{C_1} + \frac{\lambda}{C_2}\left(V_\Omega - V_f\right) \tag{4.111}$$

where λ is the Lagrange multiplier for the volume constraint, V_{Ω^h} is the volume of the optimizable domain, and C_1, C_2 are two normalization constants. The optimizable domain is D unless otherwise mentioned. Starting from some initial shape Ω_0, these constants are chosen as

$$C_1 = \int_{\partial\Omega_0} |j(\Omega_0)|\, dx, \quad C_2 = |V_{\Omega_0} - V_f|\,, \tag{4.112}$$

where j is the integrand of the shape derivative, defined in (4.110), and the initial volume is usually larger than the target volume V_f. In the context of a gradient algorithm, the descent step is also a pseudo-time step for the level-set transport equation (2.17), denoted by τ, which we choose as

$$\tau = \frac{h_{\min}}{2}, \tag{4.113}$$

where $h_{\min}$ is the minimal mesh size at the first iteration. In this chapter, the objective function $J(\Omega)$ has been normalized using C_1 and the pseudo-time step is constant, unlike in Chapter 2 and in Chapter 6 (where the normalization constants are updated every iteration). This ensures that the shape optimization evolves very slowly. This slow evolution is crucial to minimize oscillations in the convergence curve, arising from the non-smooth nature of the plasticity problem.

The Lagrange multiplier is updated using the gradient ascent as

$$\lambda_{i+1} = \lambda_i + \frac{\tau}{C_2}\left(V_{\Omega_i} - V_f\right). \tag{4.114}$$

The velocity of advection to solve transport equation (2.17) with the initial condition ϕ_i is then defined as

$$\boldsymbol{\theta}_i = \left(\frac{dj(\Omega_i)}{C_1} + \frac{\lambda_{i+1}}{C_2}\right)\boldsymbol{n}. \tag{4.115}$$

The number of gradient descent iterations is $I_{\max} = 200$. The volume constraint is not enforced at each iteration but the volume will converge to its target value by applying a gradient algorithm to the Lagrange multiplier with the same step τ. When the volume is close to the volume target, $|V_{\Omega_{i+1}} - V_f| \leq 10^{-1}V_f$, we apply a projection algorithm by updating the level-set ϕ_{i+1} using the (2.31) in an iterative manner until

$$|V_{\Omega_{i+1}} - V_f| \leq 10^{-4}V_f. \tag{4.116}$$

Using the updated level-set, the design space D is remeshed (the second last step of algorithm (5)). After remeshing, the volume tolerance (4.116) is no longer satisfied. One has to move the mesh points of the remeshed shape Ω_{i+1} using the `lag 0` option of `MMG` to ensure that the volume tolerance (4.116) is satisfied.

Remark 13. *We did not use the option `lag 0` of `MMG` in the other chapters because there, the governing model (linear elasticity, contact and fracture) is not as sensitive to remeshing and volume of the structure as the plasticity model is.*

This subsection can thus be condensed to the algorithm 5.

Algorithm 5 Repeat over $i = 0, \cdots, I_{max}$

1. Solve for $\tilde{\boldsymbol{w}}$ using the radial return algorithm on the mesh of Ω_i starting from t_1 until the last time t_N.
2. Solve for the adjoint $\tilde{\boldsymbol{z}}$ using (4.107) on the mesh of Ω_i starting from the last time t_N until t_1.
3. Compute the regularized shape derivative $dj(\Omega_i)$ by solving (4.109) with the right hand side (4.108).
4. Update the Lagrange multiplier λ_{i+1} using (4.114)
5. Solve the transport equation (2.17) with the velocity given by (4.115) for the pseudo-time step τ given by (4.113) to obtain the new level-set function $\tilde{\phi}_{i+1}$
6. Re-initialize $\tilde{\phi}_{i+1}$ to the signed distance function ϕ_{i+1} (defining a new shape Ω_{i+1})
7. Compute the volume $V_{\Omega_{i+1}}$. If $|V_{\Omega_{i+1}} - V_f| \leq 10^{-1}V_f$, apply the projection algorithm (2.31) until (4.116)
8. Remesh the box D using `MMG` [61] to obtain the body fitted mesh of the new shape Ω_{i+1}
9. Move the mesh nodes of Ω_{i+1} using the `lag 0` option of `MMG` to ensure that (4.116) is satisfied

4.5 Numerical Results

This section displays 2D and 3D optimization results with three minimization criteria: total compliance (4.59), total energy (4.119) and plastic energy (4.120). In each case a volume constraint $|\Omega| = V_f$ is imposed and the optimization algorithm 5 is applied. The solid is assumed to be mild steel with the properties: $E = 210GPa$, $\nu = 0.3$, $\sigma_Y = 279MPa$, $E_{iso} = 712MPa$. For all test-cases in this section except the one corresponding to (4.121), we consider a force $\boldsymbol{g}$ that increases from zero to a final value in one second in a constant direction with a time step $\delta t = 0.05$. The time-discretized adjoint equation (4.107) is solved using ϵ, η given in (4.101).

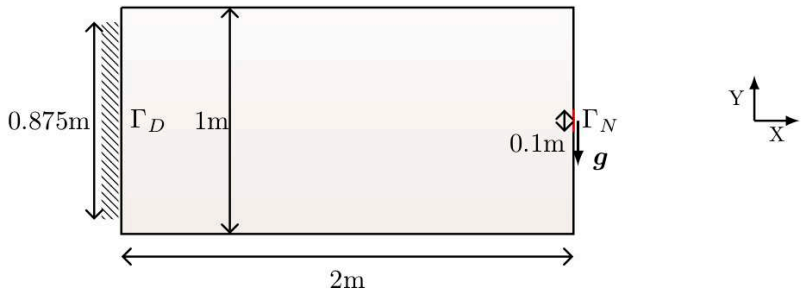
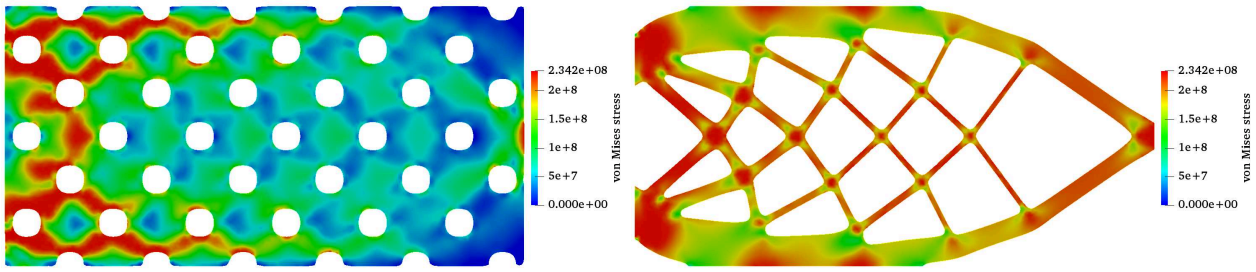


Figure 4.9: 2D Cantilever boundary conditions

4.5.1 2D Cantilever

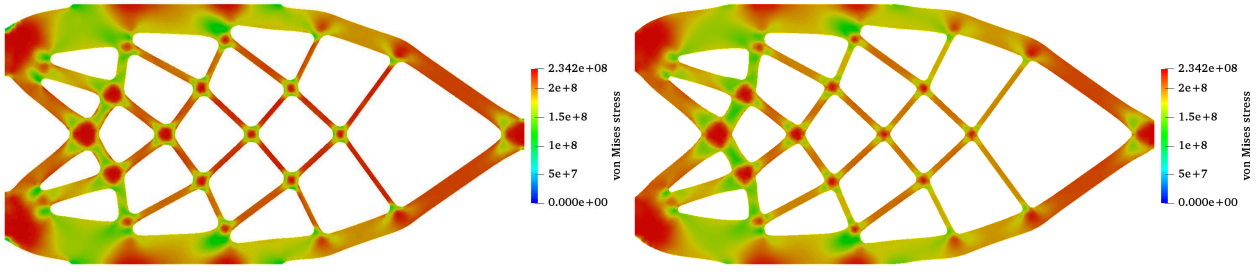
We study a $2m \times 1m$ 2D cantilever beam which is partially clamped on the left side (there is a small difference between the size of the Dirichlet boundary condition and the left edge of the beam), while a vertical concentrated force is applied at the middle of the right side of the beam (see Fig.4.9). The reason to not completely clamp the left side of the cantilever beam is to allow the shape to move around Γ_D and to avoid potential plastic zone which often appears around the Dirichlet boundary condition. A target volume $V_f = 0.7m^2$ is imposed. Based on the quasi-static assumption, the rate of force increment has no impact on the solution at the final time instant $t = 1$. However, the rate does impact the objective function (4.59). If the force grows faster in the beginning and then slowly after the onset of plasticity, the objective function is influenced more by the plastic flow. To see a greater impact of the plastic flow on the shape derivative (and hence the shape), we choose

$$\boldsymbol{g} = (0, 220 \min(1.5t, 1))MN/m,\ t \in [0, 1]s. \tag{4.117}$$



(a) Initialization with small holes placed periodically

(b) Optimized for (4.59), linear elasticity

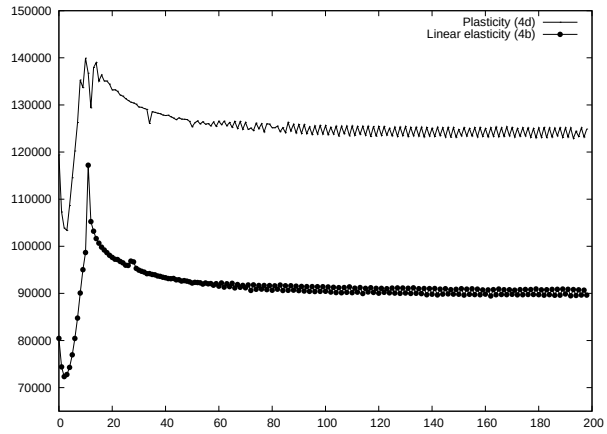


(c) Optimized for (4.59), $E_{iso} = 712MPa$, $\mathbb{H} = 0$

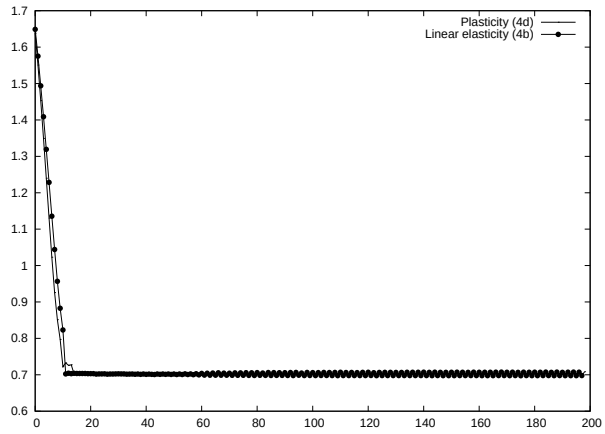
(d) Optimized for (4.59), $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$

Figure 4.10: Von Mises stress at $t = 1$s corresponding to various shapes for a target volume $V_f = 0.7m^2$ and force (4.117)

The parameters of the remeshing tool `MMG` are fixed to $h_{\min} = 0.01m$ (minimal mesh size), $h_{\max} = 0.02m$ (maximal mesh size). First, we minimize the total compliance (4.59). The initial shape and the final shapes



(a) Total compliance (4.59) vs iterations



(b) Volume vs iterations

Figure 4.11: Convergence history corresponding to shapes (4.10b) and (4.10d)

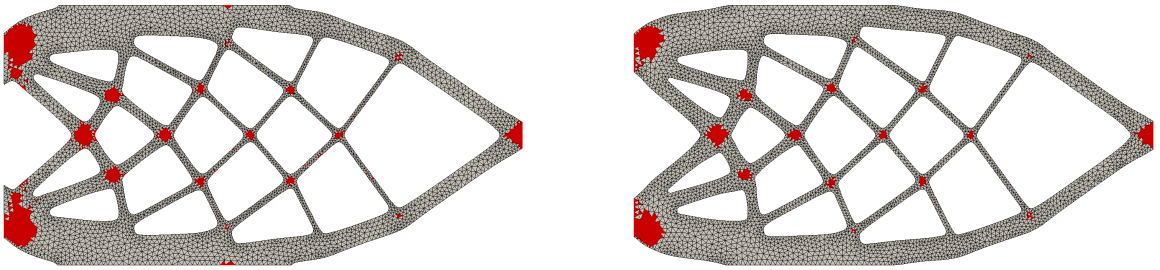


Figure 4.12: Plastic zones ($\gamma > 0$) at $t = 1$s computed for shapes (4.10b) and (4.10d)

	Shape (4.10b)	Shape (4.10d)
Comp. (4.59) for linear elasticity	89, 131	90, 428
Comp. (4.59) for plasticity ($E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$)	126, 555	123, 172

Table 4.2: 2D Cantilever shape comparison for force (4.117)

for the linear elasticity and plasticity models are shown in Fig. 4.10. Let us first note that the presence, or not, of the hardening tensor $\mathbb{H}$ does not change much the resulting optimized shape in Figures 4.10c and 4.10d. As can be seen on Figures 4.10b and 4.10d, the optimized shapes for linear elasticity or plasticity are very similar. The only slight difference is near the Dirichlet boundary condition, where the bars are thicker for the plasticity case. It turns out that the displacement for linear elasticity is numerically very close to the one for plasticity. Although the plastic deformation $\boldsymbol{\varepsilon}_p$ does contribute to the shape derivative for the plasticity case, it does not induce a different topology, compared to the elasticity case. The convergence history for the total compliance is depicted in Fig. 4.11.

To quantitatively compare the two optimized shapes in Figures 4.10b (elasticity) and 4.10d (plasticity), we perform a plasticity computation for both of them with $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$ and the force (4.117). The plastic zones (where $\gamma > 0$) at time $t = 1$s along with the mesh are plotted in Fig.4.12 and the total compliance (4.59) is noted in Table 4.2. In Fig.4.12, we observe that the plastic zones are slightly smaller for (4.10d) compared to (4.10b). As seen in Table 4.2, the total compliance for the cantilever beam obtained for plasticity is 2.75% lesser than the one obtained for the linear elasticity case. While this improvement is pertinent, it is not very impressive. On the other hand, Table 4.2 confirms that Figure 4.10b is (slightly) better than Figure 4.10d for the linear elasticity.

Next, we investigate if a few parameters of the previous test-case (external force, optimization criteria or initialization) results in a drastic change of the plastic zone. Specifically we investigate three variations.

1. Increase the external force to

$$\boldsymbol{g} = (0, 400\min(1.5t, 1))MN/m,\ t \in [0, 1]s \tag{4.118}$$

such that the entire shape undergoes a plastic deformation.

2. Consider two new criteria for minimization: total energy

$$J(\Omega) = \int_0^T \int_\Omega \frac{1}{2}\left(\mathbb{C}\boldsymbol{\varepsilon}_e : \boldsymbol{\varepsilon}_e + \mathbb{H}\boldsymbol{\varepsilon}_p : \boldsymbol{\varepsilon}_p + E_{iso}\gamma^2\right) dx \tag{4.119}$$

and energy due to kinematic hardening

$$J(\Omega) = \int_0^T \int_\Omega \frac{1}{2}\mathbb{H}\boldsymbol{\varepsilon}_p : \boldsymbol{\varepsilon}_p dx, \tag{4.120}$$

in addition to the total compliance criterion (4.59).

3. Consider three different initializations (as shown in Fig.4.13) for total compliance minimization.

The shapes obtained for the three different initializations are plotted in Fig.4.13, their corresponding compliances (4.59) are presented in Table 4.3, and their convergence histories are depicted in Fig.4.14. As expected, we obtained three different topologies. In Fig.4.13, we observe that plastic deformation occurs everywhere in the optimal shapes. This was not expected as yielding should have resulted in a high accumulated plastic deformation and hence a high total compliance. However what actually happens is that, when the shapes reach the yield point, hardening occurs. Once the shape hardens, its load bearing capacity increases. Hence the optimal shapes are the ones that struggle a balance between hardening and plastic deformation. Consequently, it is unrealistic to expect dramatic reduction in the size of plastic zones. As seen in Table 4.3, the cantilever beam is best optimized if initialized by the solution obtained for the linear elasticity case (Figure 4.13c). In Fig.4.14, we see almost no decrease in the objective function for the shape of Figure 4.13f. This means that the shape obtained for the linear elasticity case is almost optimal for plasticity.

The shapes obtained for different objective functions, namely total energy (4.119) and plastic energy (4.120), are plotted in Fig. 4.15. The shapes in Fig.4.15a and Fig.4.15b are similar to the previous shapes of Figures 4.13d and 4.13f respectively. In both cases they were initialized with Figure 4.13c. Again, the size of the plastic zone (where $\gamma > 0$) has not decreased. We believe it is because plastic zones are hardened zones and, as a result, are necessary for minimizing the total energy or the plastic energy.

	Shape (4.13b)	Shape (4.13d)	Shape (4.13f)
Comp. (4.59) for linear elasticity	423, 424	410, 188	404, 180
Comp. (4.59) for plasticity ($E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$)	608, 714	578, 630	558, 156

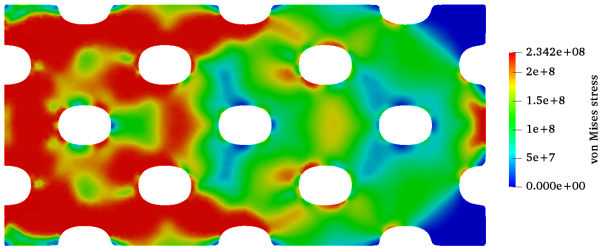
Table 4.3: 2D Cantilever shape comparison for force (4.118)

4.5.2 2D Wedge

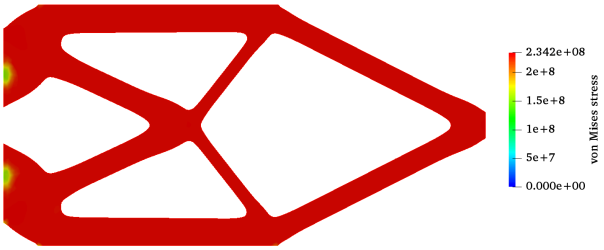
We study a 2D wedge (see Fig.4.16) which is fixed on its leftmost leg, has a vanishing vertical displacement on its rightmost leg, and is loaded on the middle of its upper boundary. A target volume $V_f = 0.2m^2$ is imposed. As before, the force grows in the beginning and then remains constant

$$\boldsymbol{g} = (0, 500\min(1.5t, 1))MN/m,\ t \in [0, 1]s.$$

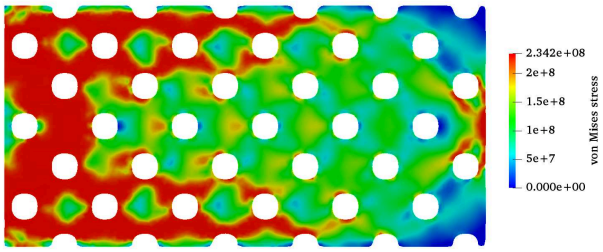
The parameters of the remeshing tool `MMG` are fixed to $h_{\min} = 0.005m$ (minimal mesh size), $h_{\max} = 0.01m$ (maximal mesh size). Isotropic hardening combined with kinematic hardening is considered using the parameters $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$. Two initializations are considered for the minimization of total compliance (4.59): one consisting of periodically distributed holes (see Figure 4.17a), the other being the optimal shape for compliance minimization in linear elasticity (see Figure 4.17c). The two initializations result in two different shapes as shown in Figures 4.17b and 4.17d. The corresponding convergence histories are plotted in Fig.4.18. As can be checked on Fig.4.18, the shape optimized for linear elasticity performs



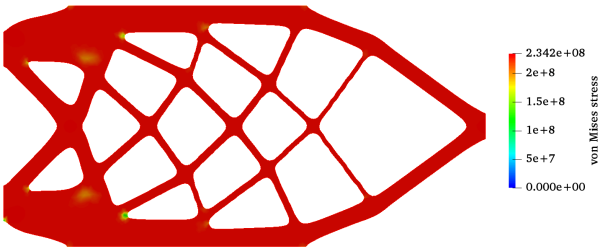
(a) Periodically distributed big holes



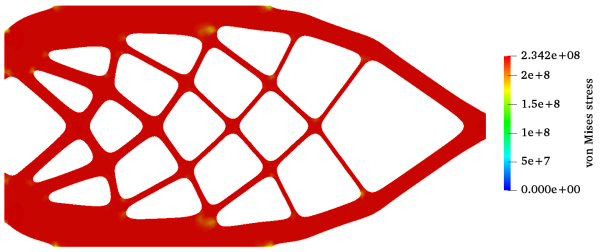
(b) Optimized shape, initialized from (4.13a)



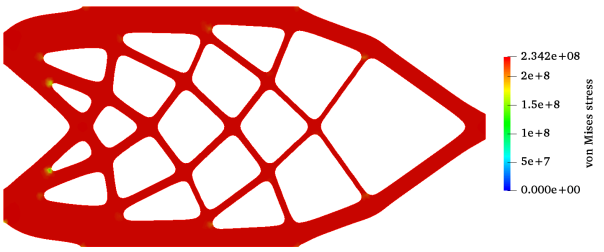
(c) Periodically distributed small holes



(d) Optimized shape, initialized from (4.13c)



(e) Shape obtained by minimizing compliance for linear elasticity



(f) Optimized shape, initialized from (4.13e)

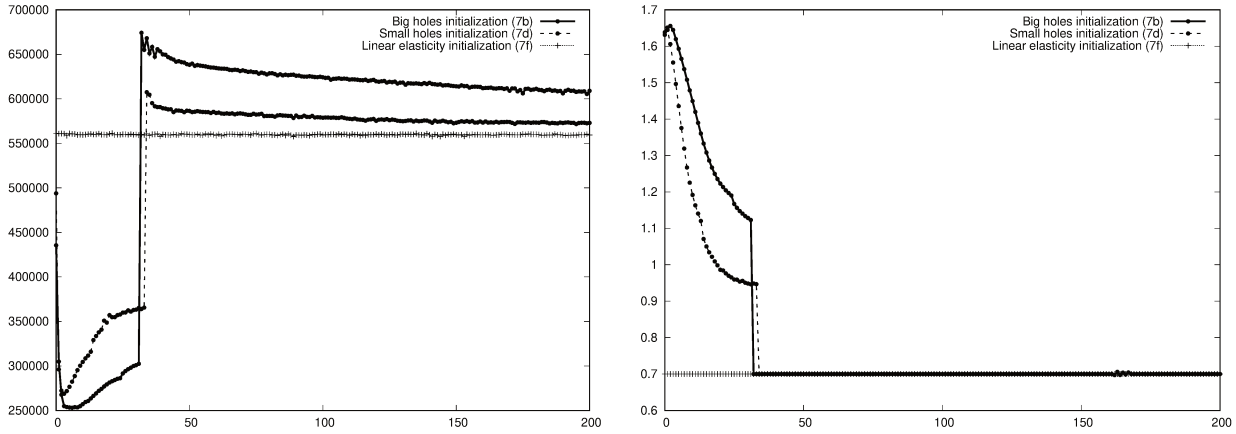
Figure 4.13: Von Mises stress at $t = 1$s for the initial shapes (on the left) and optimized shapes for total compliance (4.59) (on the right), with $V_t = 0.7m^2$, $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$ and force (4.118).

better in terms of total compliance than the shape optimized for plasticity (Figure 4.17d), starting from a periodically perforated initialization. Once again, it stresses the importance of the initialization. In all cases, the optimized shapes undergo plastic deformation everywhere in the solid.

Finally we consider the case of a force whose direction changes in time, defined as

$$\boldsymbol{g} = \left(80\cos\left(\frac{\pi t}{2}\right)|\sin(3\pi t)|, -80\sin\left(\frac{\pi t}{2}\right)|\sin(3\pi t)|\right) MPa \qquad \text{for } t \in [0,1]s. \tag{4.121}$$

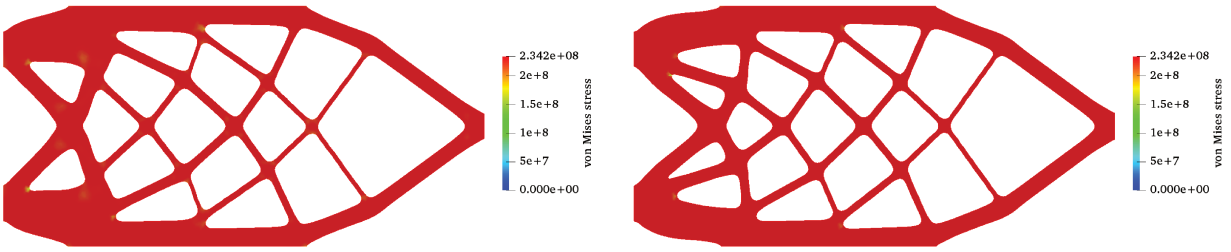
To get an intuitive idea of the direction of forcing, we also plot the force vectors for a few time steps in Fig.4.19. Only for this test-case, the time step for plasticity is taken smaller, $\delta t = 0.01s$. It implies that there are at least 100 time steps for solving the plasticity equations and the adjoint system. The initial shape is the same as in Fig.4.17a. We minimize the total compliance (4.59) for plasticity as well as for linear elasticity. For the linear elasticity case, the displacement vector $\boldsymbol{u}$ is computed for the force (4.121) at every time-step assuming quasi-static evolution. Because of this assumption, there is no dependence of $\boldsymbol{u}$ on the forcing trajectory (this could also be seen as a multiple loading test-case). However in the case of plasticity, the forcing trajectory plays an important role in influencing $\boldsymbol{u}$ as the shape undergoes plastic deformation at every time-step. This test-case is thus indicative of the role, the forcing trajectory plays in



(a) Total compliance (4.59) vs iterations

(b) Volume vs iterations

Figure 4.14: Convergence history for the shapes (4.13b), (4.13d) and (4.13f)



(a) Total energy (4.119) minimization

(b) Plastic energy (4.120) minimization

Figure 4.15: Von Mises stress at $t = 1$s for optimized shapes, initialized from (4.13c), with $V_t = 0.7m^2$, $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$ and force (4.118).

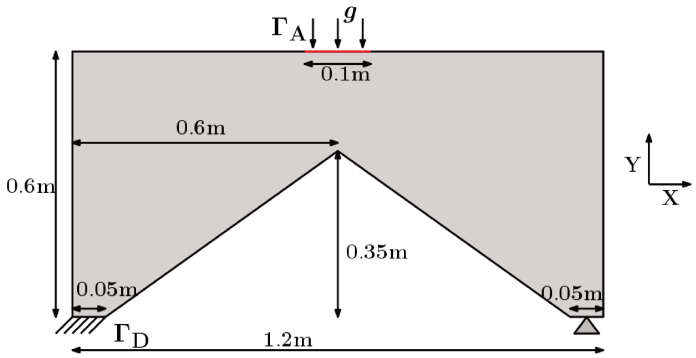
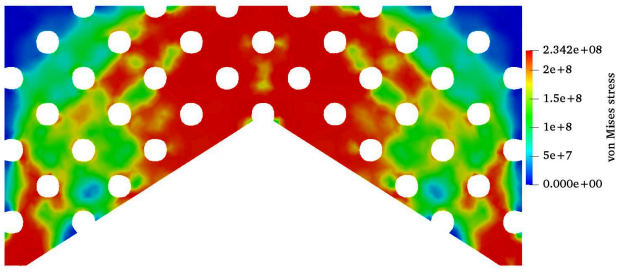


Figure 4.16: 2D wedge boundary conditions

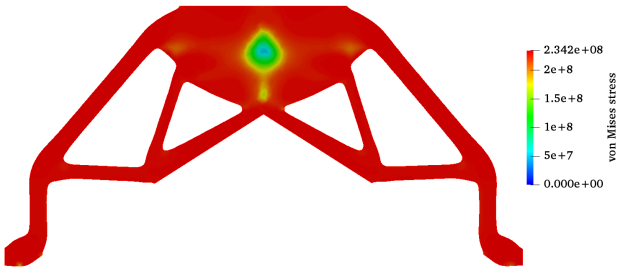
shape optimization for plasticity. The shape optimized for the force (4.121) in linear elasticity is plotted in Fig.4.20a and for combined hardening in Fig.4.20b. Table 4.4 compares the two shapes (4.20a) and (4.20b). As anticipated, the shape of Figure 4.20b performs better for plasticity.

4.5.3 3D Cantilever

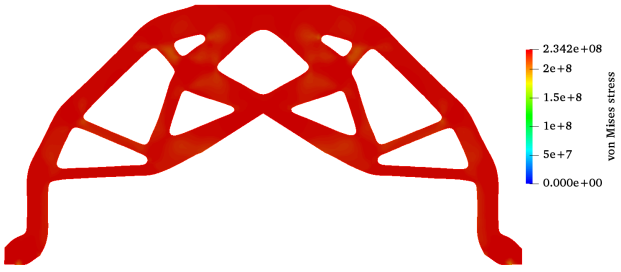
We now consider the minimization of the total compliance (4.59) for a 3D cantilever beam of dimensions $5m \times 2.4m \times 3m$, as shown in Fig.4.21. The cantilever beam is fixed on its leftmost side, loaded downwards on a circular region of radius $0.1m$ on its rightmost side with $\boldsymbol{g} = (0, 5000t, 0)MN/m$ where $t \in [0, 1]s$.



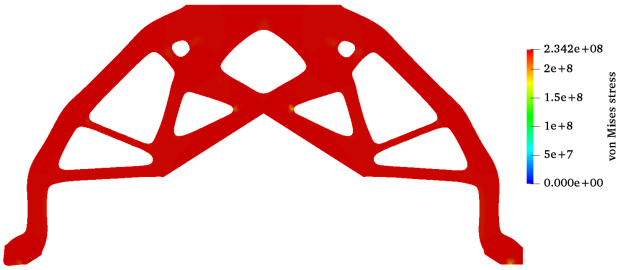
(a) Periodically distributed small holes



(b) Optimized shape, initialized from (4.17a)

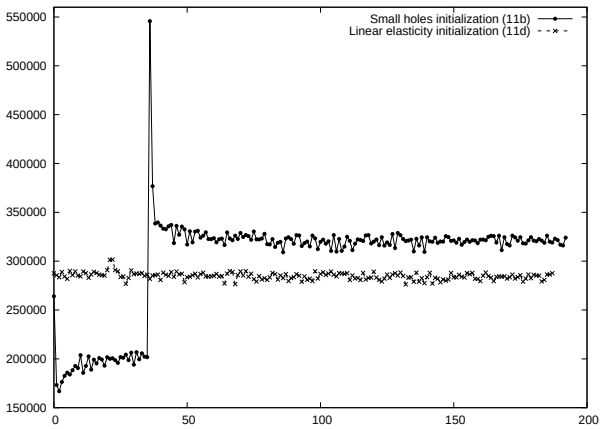


(c) Shape obtained by minimizing compliance for linear elasticity

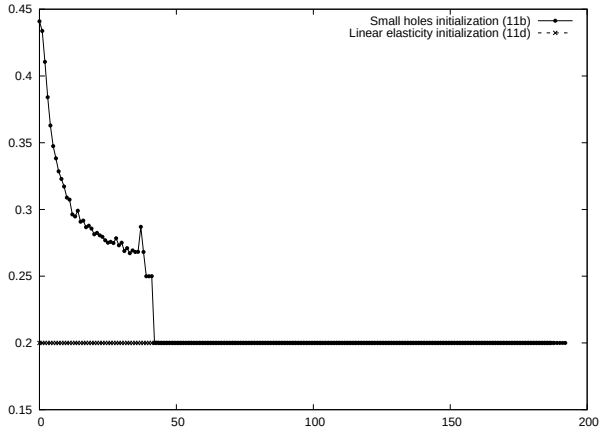


(d) Optimized shape, initialized from (4.17c)

Figure 4.17: Von Mises stress at $t = 1$s for initialized shapes (on the left), optimized shapes for total compliance (4.59) (on the right), with $V_t = 0.2m^2$, $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$



(a) Total compliance (4.59) vs iterations



(b) Volume vs iterations

Figure 4.18: Convergence history for the shapes (4.17b) and (4.17d).

	Shape (4.20a)	Shape (4.20b)
Comp. (4.59) for linear elasticity	19,566	19,915
Comp. (4.59) for plasticity ($E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$)	24,094	23,313

Table 4.4: 2D wedge shape comparison for force (4.121)

For this test-case we consider combined hardening with $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$ and a target volume $V_f = 12m^3$. The parameters of `MMG` are set to $h_{\min} = 0.04m$, and $h_{\max} = 0.12m$. We initialize the shape optimization with a perforated shape as in Fig.4.22a. Learning from the previous test-cases, we

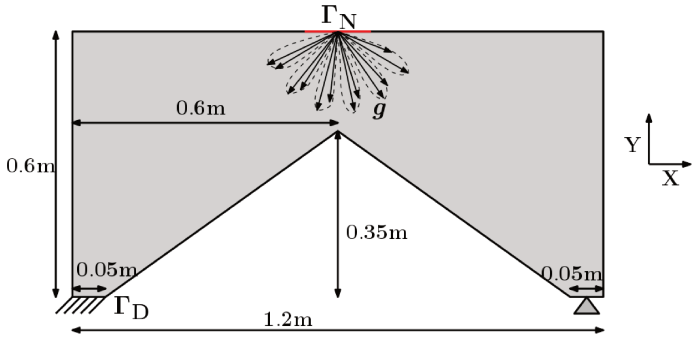
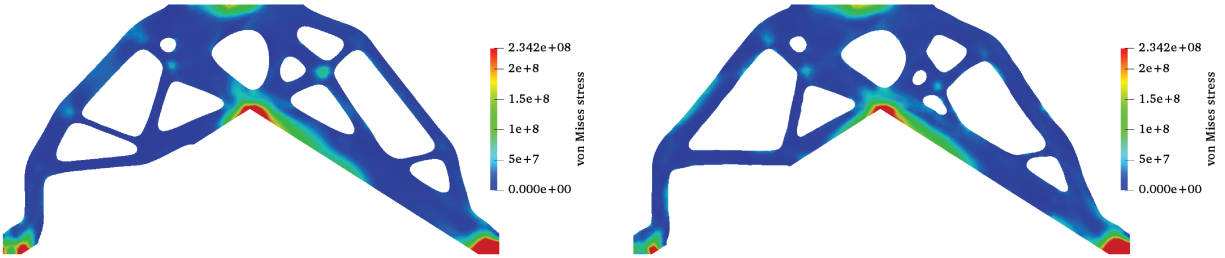


Figure 4.19: Rotating force (4.121) applied to the wedge



(a) Linear elasticity

(b) Plasticity, $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$

Figure 4.20: Von Mises stress at $t = 1$s for the shape optimized for total compliance (4.59), force (4.121), $V_f = 0.2m^2$.

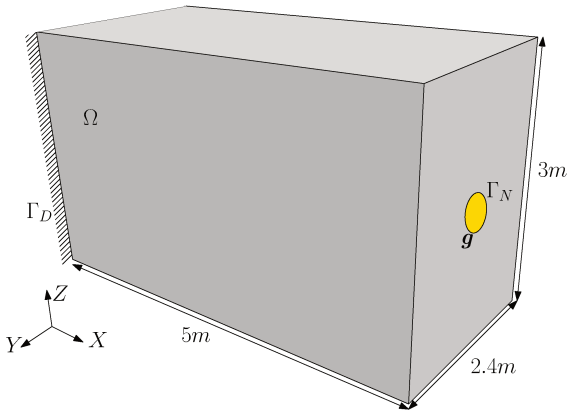
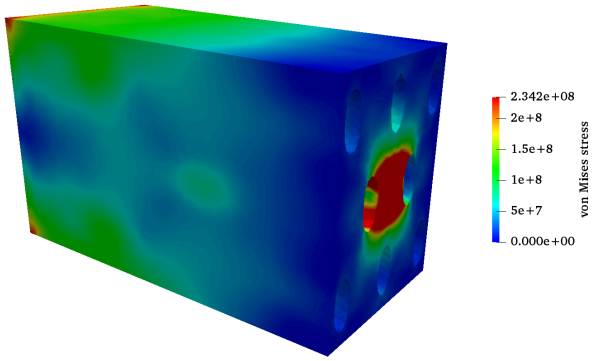
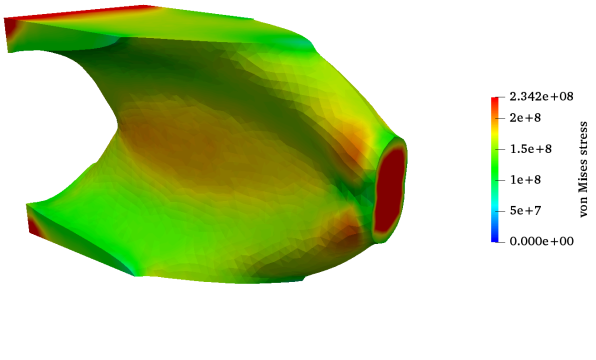


Figure 4.21: 3D cantilever boundary conditions

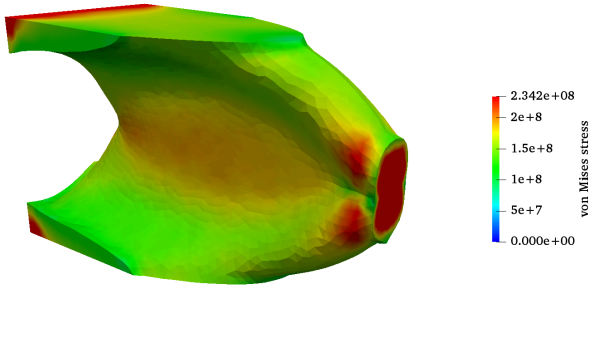
also also initialize with the shape obtained after minimizing compliance for linear elasticity (see Fig.4.22c). The optimization from initialization in Fig.4.22a is run for longer, 250 iterations instead of 200 iterations as in the other test-cases. This is because an initialization with holes is far from the optimum and it take longer to converge to a form with plate-like features (which is known to be optimal for maximizing rigidity). As seen in Fig.4.22, the two initializations result in the same shape (see Fig.4.22b-4.22d). Their corresponding convergence histories are plotted in Fig.4.23. The shapes (4.22b), (4.22c) and (4.22d) are compared quantitatively in Table 4.5. As seen in Fig.4.23, it takes a long time for the shape in Fig.4.22a to converge, whereas the shape in Fig.4.22c converges in the first few iterations. Consequently, we conclude that it is often advantageous to first optimize the shape for linear elasticity, and then use the optimized shape as initialization to minimize for plasticity.



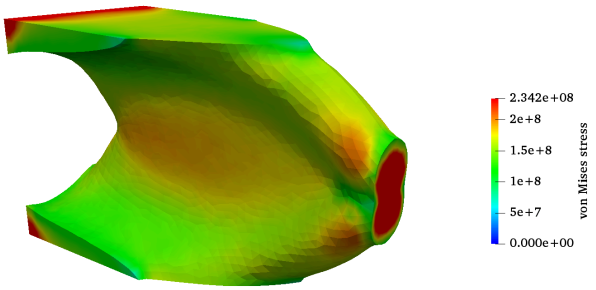
(a) Periodically distributed big holes



(b) Optimized shape, initialized from (4.22a)

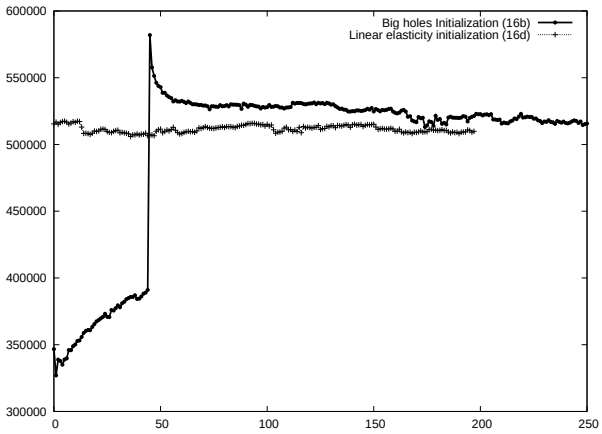


(c) Shape obtained by minimizing compliance for linear elasticity

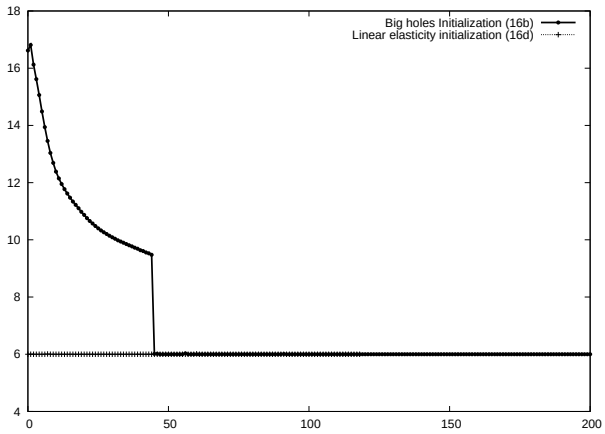


(d) Optimized shape, initialized from (4.22c)

Figure 4.22: Von Mises stress at $t = 1$s for the initial shapes (on the left) and the optimized shapes for total compliance (4.59) (on the right), with $V_t = 12m^3$, $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$.



(a) Total compliance (4.59) vs iterations



(b) Volume vs iterations

Figure 4.23: Convergence history for shapes (4.22b) and (4.22d)

	Shape (4.22c)	Shape (4.22b)	Shape (4.22d)
Comp. (4.59) in linear elasticity	453, 774	456, 363	451, 848
Comp. (4.59) in plasticity ($E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$)	515, 452	515, 246	507, 319

Table 4.5: 3D Cantilever shape comparison

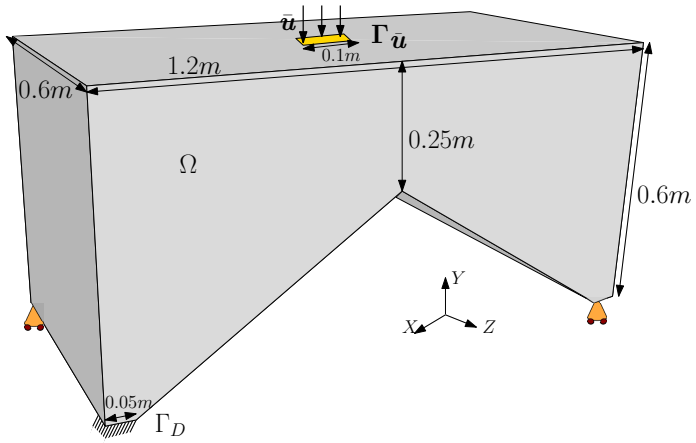


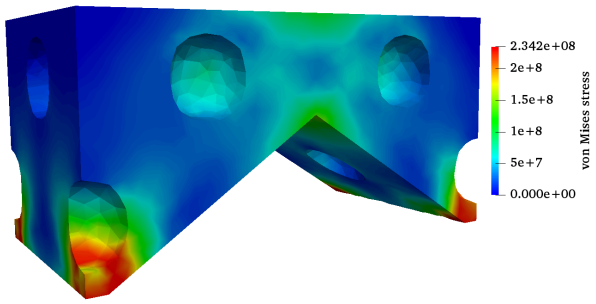
Figure 4.24: 3D wedge boundary conditions

4.5.4 3D Wedge

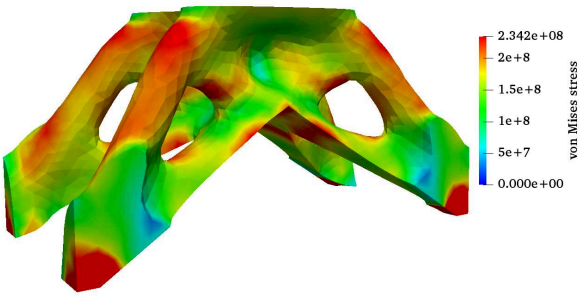
We now consider a 3D wedge of dimensions $1.2m \times 0.6m \times 0.6m$ as shown in Fig.4.24. The geometry is supported on four square surfaces each being $0.05m \times 0.05m$, three of which can be seen in the Fig.4.24. The wedge is clamped along all the three axes on one surface and only along y-direction on the remaining three surfaces. The wedge is forced on a square surface on the topmost plane with $\boldsymbol{g} = (0, -500t, 0)MN/m$ where $t \in [0, 1]$s. The parameters of `MMG` are set to $h_{\min} = 0.012m$, and $h_{\max} = 0.032m$. We consider combined hardening with $E_{iso} = 712MPa, \mathbb{H} = 105\mathbb{I}MPa$ and impose a target volume of $V_f = 0.07m^3$. Optimized shapes for linear elasticity and plasticity are displayed in Fig.4.25. Again, we consider two initializations: one with periodically distributed holes and one obtained by minimizing compliance for linear elasticity. It yields two topologically different optimized shapes as shown in Fig.4.25d. As can be seen in Table 4.6, the shape (4.25d) outperforms the shape (4.25b) in terms of the objective (4.59) in plasticity as well as in linear elasticity.

	Shape (4.25b)	Shape (4.25d)
Comp. (4.59) in linear elasticity	4,843	4,387
Comp. (4.59) in plasticity ($E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$)	5,092	4,547

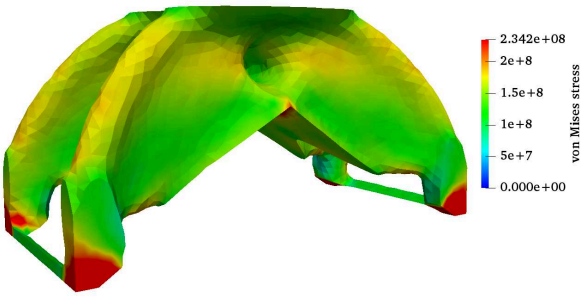
Table 4.6: 3D wedge shape comparison



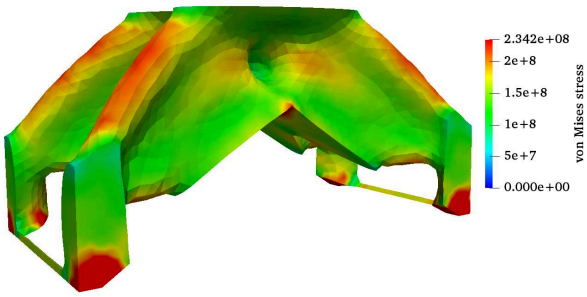
(a) Periodically distributed big holes



(b) Optimized shape, initialized from (4.25a)



(c) Shape obtained by minimizing compliance for linear elasticity



(d) Optimized shape, initialized from (4.25c)

Figure 4.25: Von Mises stress at $t = 1$s for the initial shapes (on the left) and the optimized shapes for total compliance (4.59) (on the right), with $V_t = 0.07m^3$, $E_{iso} = 712MPa$, $\mathbb{H} = 105\mathbb{I}MPa$.

Chapter 5

Plasticity: Incremental approach

Contents

For an evolution partial differential equation discretized in space-time and written in the matricial form, an adjoint method has been developed to determine the shape derivative [172]. One can think of applying this method to the space-time discretized version of the plasticity problem (4.103). The adjoint method assumes that the discretized problem is differentiable. However, neither the plasticity problem (4.103) nor its space-time discretized version (treated in algorithm 4) is differentiable with respect to the dual variables $(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{q}, g)$ or the primal variables $(\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)$. This is because the yield-limit is non-smooth with respect to $\boldsymbol{\sigma}$ and $\boldsymbol{q}$. One can nevertheless ignore the non-differentiabilty and still proceed to determine the shape derivative for the problem (4.103) [116]. We, however, prefer to take into account the non-differentiability issues and treat the time-discretized form of the plasticity problem.

In the previous chapter, we studied a variational inequation for the time-continuous formulation of the plasticity problem in the primal form. Later, via penalization and regularization, the inequation was approximated by an equation. While this simplifies the determination of the shape derivative (in Section 4.3), there are three drawbacks with this approach

1. For the numerical implementation of the approximate primal problem (4.46), one has to solve for the primal variables $(\boldsymbol{u}, \boldsymbol{\varepsilon}_p, \gamma)$ simultaneously, which is computationally expensive
2. The solution to (4.46) computed numerically has errors due to penalization and regularization.
3. The primal equation (4.35) is limited to the von Mises yield criterion with an associated flow-rule, and its extension to a general yield criterion with a non-associated plasticity is not evident.

To overcome the above three drawbacks, while considering the time-discretized version of the plasticity problem, we apply the incremental approach. In this incremental approach, the plasticity problem written in form (5.1), is viewed as an optimization problem [154]. The presentation here is given for the von Mises criteria with a associated flow-rule, but it can easily be extended to a general criteria yield criterion with a non-associated flow rule.

The optimization is carried out using a Lagrangian approach. Using the Lagrangian (5.4) and the projector operator (5.5), a variational equation (5.6a) is derived. This variational equation is solely in terms of $\boldsymbol{u}$. The other primal variables can be derived from $\boldsymbol{u}$ using (5.6b), (5.7). In the previous chapter, we could not use the projector operator because we dealt with the time-continuous version of the plasticity problem. The shape derivative for the variational equation (5.6) is derived in Section 5.2.

No numerical results are presented. The incremental formulation used in this chapter is inspired from an earlier PhD thesis [154], where the author takes into account kinematic hardening, while isotropic hardening is not considered. The extension of the work in [154] to the case where both the types of hardening could be taken into account is not that evident. One possibility of extension that we found recently is documented in this chapter (see Section 5.1), while the numerical verification of the shape derivative (computed in Section 5.2) is a subject of future works.

In the primal approach, no time-discretization is involved while computing the shape derivative or the adjoint equation (see Section 4.3), unlike for the incremental approach. Since the time-continuous plasticity problem can be treated, we gave preference to the primal approach (for the theoretical analysis and the numerical analysis), over the incremental approach.

The notations used in this Chapter are exactly the same as the ones in Chapter 4 and we advice the reader to refer to Section 4.1 for notations.

5.1 Variational formulation

As discussed in the Section 4.1, the equations govering plasticity are

1. the momentum balance equations (4.1)
2. the stress-strain relations (4.5), (4.12)
3. the von Mises yield criterion (4.2)
4. the Legendre-Fenchel equivalence (4.14) (dual or primal)

In the incremental approach, we consider equations (4.1), (4.5), (4.12) and (4.2). The equivalence (4.14) is reformulated as we now show. We recall the definition of the convex set of admissible stress states $\mathbb{E}$:

$$\mathbb{E} = \{(\boldsymbol{\sigma}, \boldsymbol{q}, g) : f(\boldsymbol{\sigma}, \boldsymbol{q}, g) \leq 0\},$$

where we consider the von Mises yield criteria

$$f(\boldsymbol{\sigma}, \boldsymbol{q}, g) = |\boldsymbol{\sigma}^D - \boldsymbol{q}^D| + \sqrt{\tfrac{2}{3}}(g - \sigma_Y).$$

We let $\chi_{\mathbb{E}}(\boldsymbol{\sigma} - \boldsymbol{q}, g)$ be the indicator function of $\mathbb{E}$ defined as

$$\chi_{\mathbb{E}}(\boldsymbol{\sigma} - \boldsymbol{q}, g) = \begin{cases} 0, & (\boldsymbol{\sigma}, \boldsymbol{q}, g) \in \mathbb{E} \\ \infty, & (\boldsymbol{\sigma}, \boldsymbol{q}, g) \notin \mathbb{E} \end{cases}.$$

The dissipation function $\mathcal{D}$ is a support function (see Def.(4.17)) and its Legendre-Fenchel conjugate is the indicator function $\chi_{\mathbb{E}}$ (see Subsection 4.1.1). Using this fact, we get (4.14), re-written here

$$(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \in \partial \mathcal{D}^*(\boldsymbol{\sigma} - \boldsymbol{q}, g) \iff (\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \in \partial \chi_{\mathbb{E}}(\boldsymbol{\sigma} - \boldsymbol{q}, g)$$

The sub-differential of the indicator function $\chi_{\mathbb{E}}$ is given by the normal cone of $\mathbb{E}$ (see Lemma 6). This implies

$$(\dot{\boldsymbol{\varepsilon}}_p, \dot{\gamma}) \in N_{\mathbb{E}}(\boldsymbol{\sigma} - \boldsymbol{q}, g).$$

Computing the normal to $\mathbb{E}$ (by differentiating the von Mises yield function), one can quite easily show that

$$\dot{\gamma} = \sqrt{\tfrac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p|.$$

This results in the following plasticity problem:

$$\begin{aligned}
-\mathrm{div}(\boldsymbol{\sigma}) &= \boldsymbol{f} && \text{in } \Omega \times [0,T) \\
\boldsymbol{u} &= \boldsymbol{0} && \text{on } \Gamma_D \times [0,T) \\
\boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{g} && \text{on } \Gamma_N \times [0,T) \\
\boldsymbol{\sigma} \cdot \boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma \times [0,T) \\
\boldsymbol{\sigma} &= \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}_p) && \text{in } \Omega \times [0,T) \\
\boldsymbol{q} &= -\mathbb{H}\boldsymbol{\varepsilon}_p && \text{in } \Omega \times [0,T) \\
g &= -E_{iso}\gamma && \text{in } \Omega \times [0,T) \\
(\boldsymbol{\sigma}, \boldsymbol{q}, g) &\in \mathbb{E} && \text{in } \Omega \times [0,T) \\
\dot{\boldsymbol{\varepsilon}}_p &\in \partial_1 \chi_{\mathbb{E}}(\boldsymbol{\sigma} - \boldsymbol{q}, g) && \text{in } \Omega \times [0,T) \\
\dot{\gamma} &= \sqrt{\tfrac{2}{3}}|\dot{\boldsymbol{\varepsilon}}_p| && \text{in } \Omega \times [0,T),
\end{aligned} \tag{5.1}$$

where ∂_1 denotes the sub differential with respect to the first argument. Now we apply the incremental approach, by discretizing the about equations in time. This allows us to formulate every time step as an optimization problem (using a Lagrangian approach) and solve it using a variational equation. The time interval $[0,T]$ is discretized to N intervals, each being of length δt. Let t_n denote the end of the n-th time interval. This incremental approach has been studied at length for plasticity with kinematic hardening, Drucker-Prager model and viscoplasticity in [154].

We extend the approach in [154] to take into account isotropic hardening by time-discretizing (5.1) as

$$\begin{aligned}
-\text{div}(\boldsymbol{\sigma}_n) &= \boldsymbol{f}(t_n) && \text{in } \Omega \times [0,T) \\
\boldsymbol{u}_n &= \boldsymbol{0} && \text{on } \Gamma_D \times [0,T) \\
\boldsymbol{\sigma}_n \cdot \boldsymbol{n} &= \boldsymbol{g}(t_n) && \text{on } \Gamma_N \times [0,T) \\
\boldsymbol{\sigma}_n \cdot \boldsymbol{n} &= \boldsymbol{0} && \text{on } \Gamma \times [0,T) \\
\boldsymbol{\sigma}_n &= \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n}) && \text{in } \Omega \times [0,T) \\
\boldsymbol{q}_n &= -\mathbb{H}_n \boldsymbol{\varepsilon}_{p,n} && \text{in } \Omega \times [0,T) \\
g_n &= -E_{iso}\gamma_{n-1} && \text{in } \Omega \times [0,T) \\
(\boldsymbol{\sigma}_n, \boldsymbol{q}_n) &\in \mathbb{E}_{n-1} && \text{in } \Omega \times [0,T) \\
\boldsymbol{\varepsilon}_{p,n} &\in \boldsymbol{\varepsilon}_{p,n-1} + \delta t\ \partial_1 \chi_{\mathbb{E}_{n-1}}(\boldsymbol{\sigma}_n - \boldsymbol{q}_n, g_n) && \text{in } \Omega \times [0,T) \\
\gamma_n &= \gamma_{n-1} + \sqrt{\tfrac{2}{3}}|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}| && \text{in } \Omega \times [0,T)
\end{aligned} \tag{5.2}$$

where the subscript n denotes the time step and the elastic domain $\mathbb{E}_{n-1}$ is now a function of the time step,

$$\mathbb{E}_{n-1} = \left\{ (\boldsymbol{\sigma}, \boldsymbol{q}) : \quad |\boldsymbol{\sigma}^D - \boldsymbol{q}^D| - \sqrt{\tfrac{2}{3}}(E_{iso}\gamma_{n-1} + \sigma_Y) \leq 0 \right\} \tag{5.3}$$

For the system of equations (5.2), it suffices to solve for $(\boldsymbol{u}_n, \boldsymbol{\sigma}_n, \boldsymbol{q}_n)$ to compute all other variables. In the time-discretization (5.2), the isotropic hardening force g_n is updated in an explicit manner; whereas the kinematic hardening $\boldsymbol{q}_n$ is updated in an implicit manner. Such a discretization is thus explicit with respect to the isotropic hardening parameters and implicit with respect to the kinematic hardening parameters.

The first time step $n = 1$ of the problem (5.2) in the absence of hardening ($\mathbb{H} = \boldsymbol{0}, E_{iso} = 0$) is known as the Hencky's problem [110]. The shape derivative for the Hencky's problem was derived in [130, 128] via Perzyna's regularization.

For the mathematical analysis of the above problem (5.2), we define the space of stress:

$$P = \{\boldsymbol{\sigma} \in L^2(\Omega)^{d\times d},\ \boldsymbol{\sigma}(\boldsymbol{x}) \in \mathcal{M}_s^d \text{ a.e. in } \Omega\}.$$

We remind the reader that the material tensors $\mathbb{C}$ and $\mathbb{H}$ are coercive with a constant h_0 (4.13). We introduce a bilinear form $A : P \times P \mapsto \mathbb{R}$

$$A(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \int_\Omega \mathbb{C}^{-1}\boldsymbol{\sigma} : \boldsymbol{\tau}\ dx,$$

and $C : P \times P \mapsto \mathbb{R}$

$$C(\boldsymbol{q}, \boldsymbol{\tau}) = \int_\Omega \mathbb{H}^{-1}\boldsymbol{q} : \boldsymbol{\tau}\ dx,$$

and a linear form $l_n^I : V \mapsto \mathbb{R}$

$$l_n^I(\boldsymbol{u}) = \int_\Omega \boldsymbol{f}(t_n) \cdot \boldsymbol{u}\ dx + \int_{\Gamma_N} \boldsymbol{g}(t_n) \cdot \boldsymbol{u}\ ds.$$

The linear form l_n^I is similar to the one defined in the previous chapter (4.32), except it is defined only over V instead of Z. Using the above bilinear and linear forms, the equations (5.2) can be now be viewed as an optimization problem, written using the Lagrangian in the following Lemma.

Lemma 13. *For $(\tilde{\boldsymbol{\sigma}}, \tilde{\boldsymbol{q}}, \tilde{\boldsymbol{u}}) \in P \times P \times V$, let a Lagrangian be defined as*

$$L(\tilde{\boldsymbol{\sigma}}, \tilde{\boldsymbol{q}}, \tilde{\boldsymbol{u}}) = \frac{1}{2}A(\tilde{\boldsymbol{\sigma}}, \tilde{\boldsymbol{\sigma}}) + \int_\Omega \tilde{\boldsymbol{\sigma}} : \boldsymbol{\varepsilon}_{p,n-1} dx + \delta t \chi_{\mathbb{E}_{n-1}}(\tilde{\boldsymbol{\sigma}} + \tilde{\boldsymbol{q}}) + \frac{1}{2}C(\tilde{\boldsymbol{q}}, \tilde{\boldsymbol{q}}) - \int_\Omega \tilde{\boldsymbol{\sigma}} : \boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}})\ dx - l_n^I(\tilde{\boldsymbol{u}}). \tag{5.4}$$

The optimum to the problem

$$\sup_{\boldsymbol{u}\in V} \inf_{(\boldsymbol{\sigma},\boldsymbol{q})\in P\times P} L(\boldsymbol{\sigma}, \boldsymbol{q}, \boldsymbol{u}),$$

given by $(\boldsymbol{\sigma}_n, \boldsymbol{q}_n, \boldsymbol{u}_n)$ satisfies the system (5.2).

Proof. Note that in the Lagrangian L (5.4), the indicator function $\chi_{\mathbb{E}_{n-1}}$ is independent of g_n. This is because the elastic domain $\mathbb{E}_{n-1}$ is independent of g_n (see (5.3)). The indicator function $\chi_{\mathbb{E}_{n-1}}$ must remain finite so that the L remains finite. Thus the optimum $(\boldsymbol{\sigma}_n, \boldsymbol{q}_n) \in \mathbb{E}_{n-1}$. The remaining proof is straight-forward and given in Proposition 5.4 in [154]. □

At this stage, we introduce a yield limit $\tilde{f} : \mathcal{M}_s^d \times \mathbb{R} \to \mathbb{R}$ given by

$$\tilde{f}(\boldsymbol{\eta}, \gamma) = |\boldsymbol{\eta}^D| - \sqrt{\tfrac{2}{3}}(E_{iso}\gamma + \sigma_Y).$$

This yield limit has the same expression as in (4.2), except the first two arguments in f are replaced by one single argument. For the above yield limit, let $\mathcal{P} : \mathcal{M}_s^d \times \mathbb{R} \mapsto \mathbb{E}$ represent the orthogonal projection operator on the convex set $\mathbb{E}$. For every $(\boldsymbol{\eta}, \gamma) \in \mathcal{M}_s^d \times \mathbb{R}$, $\mathcal{P}(\boldsymbol{\eta}, \gamma)$ can be expressed explicitly as

$$\mathcal{P}(\boldsymbol{\eta}, \gamma) = \boldsymbol{\eta} - \max(\tilde{f}(\boldsymbol{\eta}, \gamma), 0)\frac{\boldsymbol{\eta}^D}{|\boldsymbol{\eta}^D|}. \tag{5.5}$$

Introduce two new tensors $\mathbb{N}$ and $\mathbb{D}$:

$$\mathbb{N} = \mathbb{C} + \mathbb{H}, \qquad \mathbb{D} = (\mathbb{C}^{-1} + \mathbb{H}^{-1})^{-1}.$$

We finally present the variational equation for plasticity in the following theorem.

Theorem 25. *The Lagrangian* (5.4) *admits a unique optimum* $(\boldsymbol{\sigma}_n, \boldsymbol{q}_n, \boldsymbol{u}_n)$, *which satisfies*

$$\int_\Omega R\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n-1}), \gamma_{n-1}\right) : \boldsymbol{\varepsilon}(\boldsymbol{v})\ dx = l_n^I(\boldsymbol{v}) \qquad \forall \boldsymbol{v} \in V, \tag{5.6a}$$

$$\boldsymbol{\sigma}_n = R\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n-1}), \gamma_{n-1}\right), \qquad \boldsymbol{q}_n = \mathbb{H}\boldsymbol{\varepsilon}_{p,n}, \qquad \boldsymbol{\varepsilon}_{p,n} = \boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \mathbb{C}^{-1}\boldsymbol{\sigma}_n, \tag{5.6b}$$

$$\textit{where} \qquad R(\boldsymbol{\eta}, \gamma) = \begin{cases} \mathcal{P}(\boldsymbol{\eta}, \gamma) & \textit{if } |\mathbb{H}| = 0 \\ \mathbb{D}\mathbb{C}\boldsymbol{\eta} + \mathbb{N}\mathbb{C}^{-1}\mathcal{P}(\boldsymbol{\eta}, \gamma) & \textit{if } |\mathbb{H}| \neq 0. \end{cases}$$

and the isotropic hardening variable is updated using

$$\gamma_n = \gamma_{n-1} + \sqrt{\tfrac{2}{3}}|\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}|. \tag{5.7}$$

Proof. See Proposition 5.4 and Theorem 3.6 of [154]. □

For numerical implementation, one has to solve (5.6a) for $\boldsymbol{u}_n$ and then update $\boldsymbol{\varepsilon}_{p,n}$ (using (5.6b)) and γ_n (using (5.7)).

The Lagrangian for the perfect plasticity ((5.6) with $E_{iso} = 0, |\mathbb{H}| = 0$) admits a unique minimum given that the force $\boldsymbol{f}_n$ satisfies the safe-load condition. The solution to the stress problem is not unique and there can be an infinity of solutions. The safe load condition states that out of all the solutions possible, there exists at least one solution such that the stress is completely elastic. Since the set is convex, one can additionally show that there would then exist at least one solution such that the plastification is not everywhere.

In the next subsection, we will have to differentiate Eqn.(5.6) with respect to its solution $(\boldsymbol{u}_n, \boldsymbol{\varepsilon}_{p,n}, \gamma_n)$. However, the nonlinear projection operator $\mathcal{P}(\cdot, \cdot)$ is not smooth in its argument and must be regularized. In order to do so, we introduce a parameter $0 < \epsilon \ll 1$ and two operators $M_\epsilon : \mathbb{R} \to \mathbb{R}$, $N_\epsilon(\cdot) : \mathcal{M}_s^d \to \mathbb{R}$ defined as

$$M_\epsilon(\eta) = \frac{1}{2}\left(\eta + \sqrt{\eta^2 + \epsilon^2}\right), \qquad N_\epsilon(\boldsymbol{\eta}) = \sqrt{\boldsymbol{\eta} : \boldsymbol{\eta} + \epsilon^2}.$$

Using $N_\epsilon(\cdot)$, we first regularize the yield limit $\tilde{f}(\boldsymbol{\eta}, \gamma)$ to $\tilde{f}_\epsilon : \mathcal{M}_s^d \times \mathbb{R} \mapsto \mathbb{R}$,

$$\tilde{f}_\epsilon(\boldsymbol{\eta}, \gamma) = N_\epsilon(\boldsymbol{\eta}^D) - \sqrt{\tfrac{2}{3}}(E_{iso}\gamma + \sigma_Y).$$

We then regularize $\mathcal{P}(\cdot, \cdot)$ to $\mathcal{P}_\epsilon : \mathcal{M}_s^d \times \mathbb{R} \mapsto \mathbb{R}$,

$$\mathcal{P}_\epsilon(\boldsymbol{\eta}, \gamma) = \boldsymbol{\eta} - M_\epsilon(\tilde{f}_\epsilon(\boldsymbol{\eta}, \gamma))N_\epsilon'(\boldsymbol{\eta}^D)\boldsymbol{\eta}^D. \tag{5.8}$$

Finally, we regularize $R(\cdot,\cdot)$ to $R_\epsilon : \mathcal{M}_s^d \times \mathbb{R} \mapsto \mathbb{R}$ given by

$$R_\epsilon(\boldsymbol{\eta},\gamma) = \begin{cases} \mathcal{P}_\epsilon(\boldsymbol{\eta},\gamma) & \text{if } |\mathbb{H}| = 0 \\ \mathbb{D}\mathbb{C}\boldsymbol{\eta} + \mathbb{N}\mathbb{C}^{-1}\mathcal{P}_\epsilon(\boldsymbol{\eta},\gamma) & \text{if } |\mathbb{H}| \neq 0. \end{cases} \tag{5.9}$$

The regularized operator $R_\epsilon(\cdot,\cdot)$ is Gâteaux differentiable since it is $\mathcal{C}^\infty$ smooth in its arguments. The regularized plasticity problem that shall be treated in the next section reads

$$\int_\Omega R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n-1}), \gamma_{n-1}\right) : \boldsymbol{\varepsilon}(\boldsymbol{v})\ dx = l_n^I(\boldsymbol{v}), \tag{5.10}$$

$$\boldsymbol{\varepsilon}_{p,n} = \boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \mathbb{C}^{-1}R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n-1}), \gamma_{n-1}\right) \qquad \text{and} \qquad \gamma_n = \gamma_{n-1} + \sqrt{\tfrac{2}{3}}N_\epsilon(\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}).$$

The solution $(\boldsymbol{u}_n, \boldsymbol{\varepsilon}_{p,n}, \gamma_n)$ to the above depends on ϵ and we drop the subscript ϵ for simplifying the notations.

5.2 Shape derivative

Like in section (4.1), we let the state solution $\boldsymbol{w}_n = (\boldsymbol{u}_n, \boldsymbol{\varepsilon}_{p,n}, \gamma_n)$, and the adjoint solution $\boldsymbol{z}_n = (\boldsymbol{v}_n, \boldsymbol{\varepsilon}_{q,n}, \mu_n)$, $(n = 1, 2, \cdots, N)$. We denote the solution over all time steps as $\boldsymbol{w} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_N\}$, $\boldsymbol{z} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_N\}$. We minimize an objective functional $J(\Omega)$ defined as

$$J(\Omega) = \sum_{n=1}^{N} \kappa \left(\int_\Omega m(\boldsymbol{w}_n(\Omega))\ dx + \int_{\Gamma_N} p(\boldsymbol{w}_n(\Omega)) ds \right) dt \tag{5.11}$$

where $\boldsymbol{w}_n(\Omega)$ is solution to the problem (5.10) and $m(\boldsymbol{w}(\Omega))$, $p(\boldsymbol{w}(\Omega))$ are assumed to be at least $\mathcal{C}^0$ smooth like in section 3.

Theorem 26. *Let Ω be a smooth bounded open set. Let $l_n^I \in H^{-1}([0,T], Z^*)$ and $\boldsymbol{w}_n \in H^1([0,T], Z)$ be the solution to (5.2). Then the shape derivative of $J(\Omega)$ defined along $\boldsymbol{\theta} \in W_0^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$, $J'(\Omega)(\boldsymbol{\theta})$ is given by*

$$\begin{aligned} J'(\Omega)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta}\cdot\boldsymbol{n} \sum_{n=1}^{N} \Big(& m(\boldsymbol{w}_n) + R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n-1}), \gamma_{n-1}\right) : (\boldsymbol{\varepsilon}(\boldsymbol{v}_n) - \boldsymbol{\varepsilon}_{q,n}) \\ & + \mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n}) : \boldsymbol{\varepsilon}_{q,n} + \mu_n \left(\gamma_n - \gamma_{n-1} - \sqrt{\tfrac{2}{3}}N_\epsilon(\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}) \right) - l_n^I(\boldsymbol{z}_n) \Big) ds \end{aligned} \tag{5.12}$$

where $\boldsymbol{w}_n$ is the solution (5.10) and $\boldsymbol{z}_n = (\boldsymbol{v}_n, \boldsymbol{\varepsilon}_{q,n}, \mu_n)$ is the solution to the adjoint problem: $\boldsymbol{z}_N = 0$,

$$\mu_n = \mu_{n+1} - \partial_{\gamma_n} R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1}) - \boldsymbol{\varepsilon}_{p,n}), \gamma_n\right) : (\boldsymbol{\varepsilon}(\boldsymbol{v}_{n+1}) - \boldsymbol{\varepsilon}_{q,n+1}), \tag{5.13a}$$

$$\begin{aligned} & \int_\Omega \sqrt{\tfrac{2}{3}} \left(\mu_n N_\epsilon'(\boldsymbol{\varepsilon}_{p,n} - \boldsymbol{\varepsilon}_{p,n-1}) - \mu_{n+1} N_\epsilon'(\boldsymbol{\varepsilon}_{p,n+1} - \boldsymbol{\varepsilon}_{p,n}) \right) : \boldsymbol{\varphi}\ dx \\ & - \int_\Omega \partial_{\boldsymbol{\varepsilon}_{p,n}} R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1}) - \boldsymbol{\varepsilon}_{p,n}), \gamma_n\right) \mathbb{C}\boldsymbol{\varphi} : (\boldsymbol{\varepsilon}(\boldsymbol{v}_{n+1}) - \boldsymbol{\varepsilon}_{q,n+1})\ dx \\ & - \int_\Omega \mathbb{C}\boldsymbol{\varepsilon}_{q,n} : \boldsymbol{\varphi} \check{a} dx + \int_\Omega \partial_{\boldsymbol{\varepsilon}_{p,n}} m(\boldsymbol{w}_n) : \boldsymbol{\varphi}\ dx + \int_{\Gamma_N} \partial_{\boldsymbol{\varepsilon}_{p,n}} p(\boldsymbol{w}_n) : \boldsymbol{\varphi}\ ds = 0, \end{aligned}$$

$$\begin{aligned} & \textit{and} \quad \int_\Omega \partial_{\boldsymbol{u}_n} R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n) - \boldsymbol{\varepsilon}_{p,n-1}), \gamma_{n-1}\right) \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{\varphi}) : (\boldsymbol{\varepsilon}(\boldsymbol{v}_n) - \boldsymbol{\varepsilon}_{q,n})\ dx + \int_\Omega \mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{\varphi}) : \boldsymbol{\varepsilon}_{q,n}\ dx \\ & + \int_\Omega \partial_{\boldsymbol{u}_n} m(\boldsymbol{w}_n) : \boldsymbol{\varphi}\ dx + \int_{\Gamma_N} \partial_{\boldsymbol{u}_n} p(\boldsymbol{w}_n) : \boldsymbol{\varphi}\ ds = 0, \qquad n = N-1, \cdots, 1. \end{aligned}$$

Proof. Since we are dealing with an equation (5.10), we can easily apply Céa's method. According to Céa's method, we define a Lagrangian (for the shape derivative) in a formal manner as

$$\begin{aligned}\mathcal{L}(\hat{\boldsymbol{w}},\hat{\boldsymbol{z}},\Omega) = &\sum_{n=1}^{N}\kappa\left(\int_\Omega m(\tilde{\boldsymbol{w}}_n)dx + \int_{\Gamma_N} p(\tilde{\boldsymbol{w}}_n)ds\right)\\ &+\sum_{n=1}^{N}\kappa\left(\int_\Omega R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_n)-\tilde{\boldsymbol{\varepsilon}}_{p,n-1}),\tilde{\gamma}_{n-1}\right):(\boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}}_n)-\tilde{\boldsymbol{\varepsilon}}_{q,n})\ dx - l_n^I(\tilde{\boldsymbol{z}}_n)\right)\\ &+\sum_{n=1}^{N}\kappa\left(\int_\Omega \mathbb{C}(\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_n)-\tilde{\boldsymbol{\varepsilon}}_{p,n}):\tilde{\boldsymbol{\varepsilon}}_{q,n}\ dx\right) + \sum_{n=1}^{N}\kappa\int_\Omega \tilde{\mu}_n\left(\tilde{\gamma}_n-\tilde{\gamma}_{n-1}-\sqrt{\tfrac{2}{3}}N_\epsilon(\tilde{\boldsymbol{\varepsilon}}_{p,n}-\tilde{\boldsymbol{\varepsilon}}_{p,n-1})\right)dx,\end{aligned} \tag{5.14}$$

where $\tilde{\boldsymbol{w}}_n = (\tilde{\boldsymbol{u}},\tilde{\boldsymbol{\varepsilon}}_{p,n},\tilde{\gamma}_n)\in\tilde{Z}$ (see Def.(4.63)), $(\tilde{\boldsymbol{w}}_1,\tilde{\boldsymbol{w}}_2,\cdots,\tilde{\boldsymbol{w}}_N)=\hat{\boldsymbol{w}}$ denotes the state variable, and likewise $\tilde{\boldsymbol{z}}_n=(\tilde{\boldsymbol{v}},\tilde{\boldsymbol{\varepsilon}}_{q,n},\tilde{\mu}_n)\in\tilde{Z}$, $(\tilde{\boldsymbol{z}}_1,\tilde{\boldsymbol{z}}_2,\cdots,\tilde{\boldsymbol{z}}_N)=\hat{\boldsymbol{z}}$ denotes the adjoint variable. As usual, we apply the optimality conditions on the Lagrangian in (5.14) to obtain the state equation (5.10), the adjoint equation and the shape derivative.

In order to obtain the adjoint equation, we first differentiate (5.14) with respect to $\tilde{\gamma}_n$ along $\varphi\in L^2(\Omega)$, equate it to zero at $(\hat{\boldsymbol{w}},\hat{\boldsymbol{z}})=(\boldsymbol{w},\boldsymbol{z})$ and obtain

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\gamma_n}\varphi &= \kappa\int_\Omega(\mu_n-\mu_{n+1})\varphi\ dx + \kappa\int_\Omega \partial_{\gamma_n}R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1})-\boldsymbol{\varepsilon}_{p,n}),\gamma_n\right)\varphi:(\boldsymbol{\varepsilon}(\boldsymbol{v}_{n+1})-\boldsymbol{\varepsilon}_{q,n+1})dx\\ &= 0.\end{aligned} \tag{5.15}$$

Since the above holds for all $\varphi\in L^2(\Omega)$ and there are no derivatives involved in φ, we obtain

$$\mu_n = \mu_{n+1} - \partial_{\gamma_n}R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1})-\boldsymbol{\varepsilon}_{p,n}),\gamma_n\right):(\boldsymbol{\varepsilon}(\boldsymbol{v}_{n+1})-\boldsymbol{\varepsilon}_{q,n+1}).$$

Next, we differentiate (5.14) with respect to $\tilde{\boldsymbol{\varepsilon}}_{p,n}$ along $\boldsymbol{\varphi}\in Q$, equate it to zero at $(\hat{\boldsymbol{w}},\hat{\boldsymbol{z}})=(\boldsymbol{w},\boldsymbol{z})$ and obtain

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\boldsymbol{\varepsilon}_{p,n}}\boldsymbol{\varphi} &= -\kappa\int_\Omega\sqrt{\tfrac{2}{3}}\left(\mu_nN_\epsilon'(\boldsymbol{\varepsilon}_{p,n}-\boldsymbol{\varepsilon}_{p,n-1})-\mu_{n+1}N_\epsilon'(\boldsymbol{\varepsilon}_{p,n+1}-\boldsymbol{\varepsilon}_{p,n})\right):\boldsymbol{\varphi}\ dx\\ &\quad-\kappa\int_\Omega\partial_{\boldsymbol{\varepsilon}_{p,n}}R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_{n+1})-\boldsymbol{\varepsilon}_{p,n}),\gamma_n\right)\mathbb{C}\boldsymbol{\varphi}:(\boldsymbol{\varepsilon}(\boldsymbol{v}_{n+1})-\boldsymbol{\varepsilon}_{q,n+1})\ dx\\ &\quad-\kappa\int_\Omega\mathbb{C}\boldsymbol{\varepsilon}_{q,n}:\boldsymbol{\varphi}\breve{a}dx+\kappa\int_\Omega\partial_{\boldsymbol{\varepsilon}_{p,n}}m(\boldsymbol{w}_n):\boldsymbol{\varphi}\ dx+\kappa\int_{\Gamma_N}\partial_{\boldsymbol{\varepsilon}_{p,n}}p(\boldsymbol{w}_n):\boldsymbol{\varphi}\ ds\\ &= 0\end{aligned}$$

Finally, we differentiate (5.14) with respect to $\boldsymbol{u}_n$ along $\boldsymbol{\varphi}\in V$, equate it to zero at $(\hat{\boldsymbol{w}},\hat{\boldsymbol{z}})=(\boldsymbol{w},\boldsymbol{z})$ and obtain

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\boldsymbol{u}_n}\boldsymbol{\varphi} &= \kappa\int_\Omega\partial_{\boldsymbol{u}_n}R_\epsilon\left(\mathbb{C}(\boldsymbol{\varepsilon}(\boldsymbol{u}_n)-\boldsymbol{\varepsilon}_{p,n-1}),\gamma_{n-1}\right)\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{\varphi}):(\boldsymbol{\varepsilon}(\boldsymbol{v}_n)-\boldsymbol{\varepsilon}_{q,n})\ dx\\ &\quad+\kappa\int_\Omega\mathbb{C}\boldsymbol{\varepsilon}(\boldsymbol{\varphi}):\boldsymbol{\varepsilon}_{q,n}\ dx+\kappa\int_\Omega\partial_{\boldsymbol{u}_n}m(\boldsymbol{w}_n):\boldsymbol{\varphi}\ dx+\kappa\int_{\Gamma_N}\partial_{\boldsymbol{u}_n}p(\boldsymbol{w}_n):\boldsymbol{\varphi}\ ds=0.\end{aligned}$$

□

To determine the shape derivative (5.12), it suffices to differentiate the relation $J(\Omega)=\mathcal{L}(\boldsymbol{w}(\Omega),\hat{\boldsymbol{z}},\Omega)$ with respect to Ω, use the adjoint equations (5.13) and apply the Lemma 1.

Remark 14. *In the incremental formulation (5.10), the variables $(\boldsymbol{u}_n,\boldsymbol{\varepsilon}_{p,n},\gamma_n)$ are decoupled. One first determines $\boldsymbol{u}_n$ by solving the variational equation and then updates explicitly $\boldsymbol{\varepsilon}_{p,n}$, followed by γ_n.*

In the same spirit, in the adjoint equation (5.13), the variables $(\boldsymbol{v}_n,\boldsymbol{\varepsilon}_{q,n},\mu_n)$ are decoupled. However, the sequence of determining the variables is inversed, one first determines μ_n, then $\boldsymbol{\varepsilon}_{q,n}$ followed by $\boldsymbol{v}_n$.

Chapter 6

Fracture

Contents

6.1 Introduction

The integrity of a mechanical structure is of paramount importance to ensure safety. Structures subjected to a critical external load may undergo high stress, material damage, crack and an ultimate collapse. Shape and topology optimization to minimize damage can facilitate engineers to conceive robust structures, less susceptible to crack and hence with a longer life. The goal of the present work is to design a level-set based algorithm for optimizing structures which are not prone to fracture.

The first step is to choose a convenient fracture model. In the framework of brittle fracture mechanics, we consider the so-called Francfort-Marigo model [77], in the spirit of the pioneering energetic approach of Griffith [84]. This model relies on the quasi-static minimization of a total energy which is the sum of a linear elasticity energy and a fracture energy, under an irreversibility constraint. Its mathematical formulation was inspired by the Mumford-Shah energy functional for image segmentation [136], which is a simpler model with a scalar unknown and without time variable. The original Francfort-Marigo model belongs to the class of

free discontinuity problems (the crack is an unknown of the problem), which are notoriously difficult to solve numerically. Therefore, the same authors propose to approximate their original model with a regularized damage model [35], following a Gamma-convergence process, first suggested by Ambrosio and Tortorelli for the Mumford-Shah functional [17]. This Ambrosio-Tortorelli regularization features a small regularizing parameter $\ell > 0$. In the case of the Mumford-Shah functional, by Gamma-convergence techniques it was proved in [17], [36] that, as ℓ converges to zero, the minimizers of the regularized model converge to those of the original free discontinuity problem. A similar convergence holds true for a time-discretized version of the Francfort-Marigo model [44, 79].

In the regularization of this free discontinuity problem, the parameter ℓ can be viewed as a purely mathematical artifact, used to approach the exact minimization problem. However, this regularization has also a very clear physical interpretation as damage model with an internal length scale which is precisely the regularization parameter ℓ. Recall that a damage model features a damage variable α, measuring the extent of degradation of the material and varying continuously between 0 (no damage) and 1 (fully damage), and an elastic Hooke's tensor which is a decreasing function of α. Additional ingredients are, like in the fracture case, an irreversibility constraint and an energy criterion which determines the onset of damage. How the original fracture or free discontinuity problem is recast as a damage problem is precisely explained in [35] and [148]. This approach is also called sometimes a smeared interface approach or a phase-field approach. Its main advantage is its simplicity of numerical implementation using the finite element method. Furthermore, it can easily detect initiation, branching and coalescence of cracks without the need of meshing the crack path. In the sequel we shall indifferently call this model a damage or a fracture model.

The second step is to choose a shape and topology optimization method. There are mainly two classes of algorithms. The first class, that of so-called density methods, represents or approximates shapes by a continuous density function. The second class captures shapes by means of an auxiliary function which could be a characteristic function, a phase-field function or a level-set function. Here, we follow the level-set framework, introduced by Osher and Sethian [144]. The level-set approach was combined with the notion of shape derivative in [13, 184] to make it a successful method for shape and topology optimization of structures. We refer to [8] for a review of the level-set method in structural optimization and for further references. Combining the Francfort-Marigo fracture model with the level-set method, we propose a shape and topology optimization algorithm for preventing crack initiation and propagation in solid structures.

There has already been some works on this topic. In the level-set framework we are only aware of [186] which optimizes the configuration of composite materials in a phase-field based fracture model. We differ from [186] in many aspects: they do not use a level-set equation but rather a reaction-diffusion equation, they do not use a shape derivative but instead a topological gradient for a simplified model with a fixed damage field and finally they do not optimize the overall shape but just the inclusion's shape inside a composite structure. There are more works on topology optimization using SIMP (solid isotropic material with penalization) applied to various fracture models. Topology optimization using SIMP was performed for a fracture model in [101] or for a damage model to reinforce concrete in [20]. Path dependency of the damage model was taken into account in SIMP [96]. Phase-field model of fracture was considered for topology optimization using extended BESO (Bi-directional Evolutionary Structural Optimization) [54]. Fracture governed by the phase-field model is considered for maximizing the fracture resistance of periodic composite [54].

A common feature in all the previous works using SIMP is that several material properties must be approximated for the mixture of material and void, corresponding to the density variable. This is classical for the Young's modulus [32] but more delicate for other properties like, for example, the fracture toughness. Usually, these material properties are approximated by multiplying their values by the density raised to a certain exponent. This exponent is different for every property and ought to be chosen in an ad-hoc manner, ensuring numerical stability. If the optimized shape has intermediate densities, the interpretation of damage is quite artificial and may cause numerical difficulties. On the contrary, in the level-set framework, since the material properties are never approximated, such artificial damaged zones are avoided.

The content of this chapter is as follows. Section 6.2 is devoted to the presentation of the Francfort-Marigo damage model. Although this model has nice properties, it features an irreversibility constraint (a damage region cannot heal and be again undamaged) which makes it a variational inequality, instead of a more standard variational equality. Unfortunately, the adjoint method for computing sensitivities or derivatives of an objective function is extremely involved and not practical for variational inequalities since it involves the notion of conical derivative [134, 171]. Therefore, we penalize the irreversibility constraint to transform the variational inequality into a more convenient variational equality which is amenable to the adjoint method.

Section 6.3 states the optimization problem and delivers its shape derivative (see Proposition 1), relying on the adjoint method and the well-known Hadamard and Céa's methods. There is again a subtle point here. The state equation corresponds to the minimization of a total energy depending on two variables, the elastic displacement $\boldsymbol{u}$ and the damage variable α. It turns out that this energy is not convex with respect to $(\boldsymbol{u}, \alpha)$. As usual, the adjoint equation is the adjoint linearization of the optimality condition for this energy minimization. Therefore, the adjoint equation features a linear operator which is the transpose of the Hessian (or second-order derivative matrix) of this non-convex energy. As such, it is not a coercive operator and it is not clear at all that the adjoint equation admits a solution. Nevertheless, under a technical assumption we are able to prove the existence and uniqueness of a solution of the time-discretized version of the adjoint problem (see Theorem 28).

Section 6.4 presents the discretization of the damage model and of the adjoint problem in space and time, and the level-set algorithm. A key ingredient for the sequel is that, the damage model being non-convex, it is solved with a so-called backtracking algorithm, due to [35], which is able to escape from local minima in the total energy minimization. Our finite element analyses are performed within the `FreeFEM` software [92]. For the finer 3D meshes we use the parallel computational capacities of `FreeFEM` which rely on domain decomposition and the `PETSc` package [22]. The level-set transport equation is solved with the `advect` library [39]. One original feature of our work is that we rely on body-fitted meshes at each iteration, using the remeshing software `MMG` [61]. Having a body-fitted mesh of the structure is crucial for an accurate evaluation of the damage. Indeed, there is no ersatz or weak material, so damage does not interact with this fictitious phase. Furthermore, since damage typically occurs in region of high stresses, those are more precisely computed with a body-fitted mesh. We also outline the details of our gradient-based algorithm for shape and topology optimization.

Finally, Section 6.5 is concerned with 2D and 3D numerical test cases. The objective function is the so-called total compliance (integrated in time), which is minimized under a volume constraint. These examples illustrate the efficiency of our proposed shape and topology optimization algorithm to obtain crack-free optimal structures. One specific difficulty in optimization of damage of fracture models is that these phenomena may be discontinuous: a small increase in the loadings, a small change in the structure's geometry may cause the sudden occurrence of a not-so-small crack and thus a large increase of the objective function (see the discussion in Remark 26). Nevertheless, our algorithm is able to sustain these large oscillations and, in the end, converge smoothly to optimal undamaged structures.

The content of this chapter can be found in an article, submitted for publication:
J.Desai, G.Allaire, F.Jouve. *Topology optimization of structures undergoing fracture*
This chapter is an elaborated version of the above article as it contains more details on the analytical and the numerical treatment of the fracture problem, along with additional numerical test cases in results Section 6.5.

6.2 Fracture model

This section is devoted to a presentation of the Francfort-Marigo model [77] of brittle fracture which relies on a mechanical energy minimization. Their original model was based on a representation of the fracture as a line (in 2D) or surface (in 3D) of displacement discontinuity but they proposed [35] a very efficient approximation using the Ambrosio-Tortorelli regularization, which can be viewed as a gradient damage model [148]. Over the last two decades, several researchers have worked on different aspects of this regularization, cited in the following presentation.

6.2.1 Governing laws and variational formulation

We let

- $d = 2, 3$ be the dimension of the problem
- $\Omega \subset \mathbb{R}^d$ be an open bounded set representing the structure in Fig.6.1
- $[0, T]$ be the bounded time interval of interest
- $\partial\Omega = \Gamma_D \cup \Gamma_N \cup \Gamma \cup \Gamma_{\tilde{\boldsymbol{u}}}$, union of disjoint boundaries in Fig.6.1
- $\boldsymbol{n}$ denote the outward normal to $\partial\Omega$

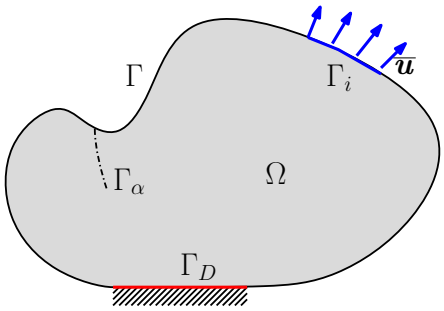


Figure 6.1: Boundary conditions and the unknown crack Γ_α

- Γ denote the free boundary (no traction)
- Γ_D denote the Dirichlet boundary
- $\Gamma_{\bar{\boldsymbol{u}}}$ be the boundary subjected to an imposed displacement $\bar{\boldsymbol{u}} : \Gamma_{\bar{\boldsymbol{u}}} \times [0,T] \mapsto \mathbb{R}^d$, typically $\bar{\boldsymbol{u}} \in \mathcal{C}^0([0,T], H^2(\mathbb{R}^d)^d)$
- $\boldsymbol{f} : \Omega \times [0,T] \mapsto \mathbb{R}^d$ be the body force applied in Ω, and $\boldsymbol{f}(t) \in L^2(\Omega)^d$
- $\boldsymbol{g} : \Gamma_N \times [0,T] \mapsto \mathbb{R}^d$ be the surface force applied on the Neumann boundary Γ_N, and $\boldsymbol{g}(t) \in H^1(\Omega)^d$
- $\boldsymbol{u} : \Omega \times [0,T] \mapsto \mathbb{R}^d$ be the displacement vector field
- $\boldsymbol{\varepsilon}(\boldsymbol{u}) = \frac{1}{2}\left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T\right)$ denote the second order strain tensor
- $\mathbb{C}_0$ be the fourth order Hooke's tensor
- σ_M denote the ultimate tensile strength of the material
- $\alpha : \Omega \times [0,T] \mapsto [0,1]$ be the damage variable
- G_c be the fracture toughness
- ℓ be the characteristic length (measuring the thickness of the damaged zone)
- $w(\alpha)$ be the energy dissipation function
- $\mathcal{M}_s^d$ denote the set of symmetric $d \times d$ matrices.

The fourth order Hooke's tensor $\mathbb{C}_0$ is assumed to be coercive on the set $\mathcal{M}_s^d$, i.e., $\exists\ c_{\min} > 0,\ c_{\max} > 0$ such that, $\forall \boldsymbol{\xi} \in \mathcal{M}_s^d$,

$$c_{\min}|\boldsymbol{\xi}|^2 \leq \mathbb{C}_0 \boldsymbol{\xi} : \boldsymbol{\xi} \leq c_{\max}|\boldsymbol{\xi}|^2.$$

The rate of evolution of the body force $\boldsymbol{f}$, the surface force $\boldsymbol{g}$ and the imposed displacement $\bar{\boldsymbol{u}}$ is assumed to be negligible, resulting in a quasi static evolution of the structure, and hence an acceleration $\ddot{\boldsymbol{u}} = \boldsymbol{0}$. As the loading increases with time t and the elastic energy in the structure exceeds a critical elastic energy density, the structure undergoes damage, which is measured with the damage variable $\alpha : \Omega \times [0,T] \mapsto [0,1]$. The value $\alpha = 0$ corresponds to no damage and $\alpha = 1$ to a complete damage. Damage is characterized by deterioration of the stiffness and is modeled by assuming the Hooke's tensor $\mathbb{C}(\alpha)$ to be a convex function of the damage variable α, such that

$$\mathbb{C}(0) = \mathbb{C}_0, \quad \mathbb{C}(1) = \boldsymbol{0} \quad \text{and} \quad \mathbb{C}'(\alpha)\boldsymbol{\xi} : \boldsymbol{\xi} < 0\ \forall \boldsymbol{\xi} \in \mathcal{M}_s^d. \tag{6.1}$$

The second expression in the above states that when the damage variable α attains unity, the stiffness becomes zero and structure undergoes a crack. The third expression in the above states that, when α increases, the stiffness must decrease.

The study of the damage model requires some functional spaces. The space of admissible displacements is given by

$$V = \{\boldsymbol{u} \in H^1(\Omega)^d : \boldsymbol{u} = \boldsymbol{0} \ \text{ on } \Gamma_D\}. \tag{6.2}$$

For $v = \dot{\bar{\boldsymbol{u}}}$, the affine space of admissible velocities is defined as

$$C_v = \{\boldsymbol{v} \in V : \boldsymbol{v} = v, \text{ on } \Gamma_{\bar{\boldsymbol{u}}}\}. \tag{6.3}$$

The subspace of functions in $H^1(\Omega)$ which vanish on $\Gamma_D, \Gamma_{\bar{\boldsymbol{u}}}$ and Γ_N is denoted by

$$H^1_D(\Omega) = \{\beta \in H^1(\Omega) : \ \beta = 0 \ \text{ on } \Gamma_D \cup \Gamma_{\bar{\boldsymbol{u}}} \cup \Gamma_N\}. \tag{6.4}$$

The convex set of admissible damage is

$$\mathcal{D}_1 = \{\alpha \in H^1_D(\Omega) : \ 0 \le \alpha(x) \le 1 \ \ a.e.\ x \in \Omega\} \tag{6.5}$$

and the convex set of admissible damage evolution rate ($\gamma = \dot{\alpha}$) is

$$\mathcal{D} = \{\gamma \in H^1_D(\Omega) : \ \gamma(x) \ge 0 \ \ a.e.\ x \in \Omega\}, \tag{6.6}$$

To simplify notations, we define the product space

$$Z = V \times H^1_D(\Omega).$$

The initial condition of the model is

$$(\boldsymbol{u}(0), \alpha(0)) = (\boldsymbol{0}, \alpha_0) \in V \times \mathcal{D}_1. \tag{6.7}$$

Remark 15. *Typically, the set $\mathcal{D}_1$ of admissible damage variable and the set $\mathcal{D}$ of admissible damage evolution rate are defined for functions in $H^1(\Omega)$. In (6.5) and (6.6) we rather choose $H^1_D(\Omega)$ because we wish to forbid crack formation on $\Gamma_D, \Gamma_{\bar{\boldsymbol{u}}}$ and Γ_N. This is because the creation of cracks on these boundaries require only half the energy needed for cracks occurring inside the structure Ω and are thus artificial.*

Following [16, 148] we introduce the elastic energy and the damage energy respectively as

$$\mathcal{E}(\boldsymbol{u}, \alpha) = \frac{1}{2}\int_\Omega \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u})\, dx, \tag{6.8}$$

$$\mathcal{H}(\alpha) = \frac{G_c}{4c_w}\left(\int_\Omega \ell |\nabla \alpha|^2 dx + \int_\Omega \frac{w(\alpha)}{\ell} dx\right), \quad \text{where} \quad c_w = \int_0^1 \sqrt{w(\xi)} d\xi, \tag{6.9}$$

where G_c is the fracture toughness of the material, $\ell > 0$ is the characteristic length, measuring the thickness of the damaged zone around the fracture, and $w(\alpha)$ is the fracture energy density [51], assumed to satisfy

$$w(0) = 0,\ w(1) = 1, \text{ and } w'(\alpha) \ge 0 \text{ for } 0 \le \alpha \le 1.$$

The sum of the integrands in the above two energies defines the energy density W_ℓ

$$W_\ell(\boldsymbol{u}, \alpha) = \frac{1}{2}\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) + \frac{G_c}{4c_w}\left(\ell |\nabla \alpha|^2 + \frac{w(\alpha)}{\ell}\right). \tag{6.10}$$

The total energy of the structure is then given by

$$\mathcal{P}(\boldsymbol{u}, \alpha) = \int_\Omega W_\ell(\boldsymbol{u}, \alpha)\ dx - \int_\Omega \boldsymbol{f} \cdot \boldsymbol{u}\ dx - \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{u}\ ds. \tag{6.11}$$

The Francfort-Marigo regularized fracture model amounts to minimize the total energy $\mathcal{P}(\boldsymbol{u}, \alpha)$ among all fields $(\boldsymbol{u}(t), \alpha(t)) \in V \times \mathcal{D}_1$, for $t \in (0, T]$, such that $(\dot{\boldsymbol{u}}(t), \dot{\alpha}(t)) \in C_v \times \mathcal{D}$ and with the initial condition (6.7). In particular, the definition of the space $\mathcal{D}$ contains the irreversibility condition $\dot{\alpha} \ge 0$ and that of the space C_v implies that the boundary condition $\boldsymbol{u}(t) = \bar{\boldsymbol{u}}(t)$ on $\Gamma_{\bar{\boldsymbol{u}}}$ is satisfied at all times. The solution $(\boldsymbol{u}(t), \alpha(t))$ satisfies the energy balance

$$\int_\Omega \dot{W}_\ell(\boldsymbol{u}, \alpha, \nabla\alpha)\ dx = \int_\Omega \boldsymbol{f} \cdot \dot{\boldsymbol{u}}\ dx + \int_{\Gamma_N} \boldsymbol{g} \cdot \dot{\boldsymbol{u}}\ ds + \int_{\Gamma_{\bar{\boldsymbol{u}}}} (\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u})\ \cdot \boldsymbol{n}) \cdot \dot{\bar{\boldsymbol{u}}}\, ds, \tag{6.12}$$

where overdot represents derivative with respect to time. The optimality conditions for the minimization of (6.11), written in a compact form, are [121, 142]: for all $t \in (0, T]$, find $(\boldsymbol{u}(t), \alpha(t)) \in V \times \mathcal{D}_1$, satisfying the initial condition (6.7), such that $(\dot{\boldsymbol{u}}(t), \dot{\alpha}(t)) \in C_v \times \mathcal{D}$, $\boldsymbol{u}(t) = \bar{\boldsymbol{u}}(t)$ on $\Gamma_{\bar{\boldsymbol{u}}}$ and

$$d\mathcal{P}(\boldsymbol{u}, \alpha)(\boldsymbol{v} - \dot{\boldsymbol{u}}, \beta - \dot{\alpha}) \ge 0 \qquad \forall (\boldsymbol{v}, \beta) \in C_v \times \mathcal{D}. \tag{6.13}$$

From (6.13), one can derive the strong form of the quasi-static damage model: find $(\boldsymbol{u}(t), \alpha(t))$, satisfying the initial condition (6.7), such that

$$\dot{\alpha} \geq 0 \qquad \text{in } \Omega \times (0,T], \tag{6.14a}$$

$$\frac{1}{2}\mathbb{C}'(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{u}) - \frac{G_c}{2c_w}\ell\Delta\alpha + \frac{G_c}{4c_w}\frac{w'(\alpha)}{\ell} \geq 0 \qquad \text{in } \Omega \times (0,T], \tag{6.14b}$$

$$\dot{\alpha}\left(\frac{1}{2}\mathbb{C}'(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{u}) - \frac{G_c}{2c_w}\ell\Delta\alpha + \frac{G_c}{4c_w}\frac{w'(\alpha)}{\ell}\right) = 0 \qquad \text{in } \Omega \times (0,T], \tag{6.14c}$$

$$\nabla\alpha\cdot\boldsymbol{n} \geq 0, \quad \dot{\alpha}\nabla\alpha\cdot\boldsymbol{n} = 0 \qquad \text{on } \partial\Omega \times (0,T], \tag{6.14d}$$

$$\alpha = 0 \qquad \text{on } (\Gamma_D \cup \Gamma_{\bar{\boldsymbol{u}}}) \times (0,T], \tag{6.14e}$$

$$\operatorname{div}(\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u})) + \boldsymbol{f} = 0 \qquad \text{in } \Omega \times (0,T], \tag{6.14f}$$

$$\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u})\cdot\boldsymbol{n} = \boldsymbol{g} \qquad \text{on } \Gamma_N \times (0,T] \tag{6.14g}$$

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \qquad \text{on } \Gamma_{\bar{\boldsymbol{u}}} \times (0,T], \tag{6.14h}$$

$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on } \Gamma_D \times (0,T]. \tag{6.14i}$$

Inequality (6.14b) is known as the damage criterion. Equation (6.14c) is the complementary relation which essentially states that the damage criterion is an equality only if the damage evolution rate is positive. The variational formulation (or weak form) of the system of equations (6.14) reads: for all $t \in (0,T]$ find $(\boldsymbol{u}(t), \alpha(t)) \in V \times \mathcal{D}_1$ such that $\boldsymbol{u}(t) = \bar{\boldsymbol{u}}(t)$ on $\Gamma_{\bar{\boldsymbol{u}}}$, $\dot{\alpha} \in \mathcal{D}$

$$\int_\Omega \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{v})\,dx = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\,dx + \int_{\Gamma_N} \boldsymbol{g}\cdot\boldsymbol{v}\,ds \quad \forall \boldsymbol{v} \in V, \tag{6.15a}$$

$$\int_\Omega \frac{1}{2}\mathbb{C}'(\alpha)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{u})\,dx + \int_\Omega \frac{G_c}{2c_w}\ell\nabla\alpha\cdot\nabla\beta\,dx + \int_\Omega \frac{G_c}{4c_w}\frac{w'(\alpha)}{\ell}\beta\,dx \geq 0 \quad \forall \beta \in \mathcal{D}. \tag{6.15b}$$

When the characteristic length ℓ is small enough, and since $\mathbb{C}(1) = \boldsymbol{0}$, the above gradient damage model (6.15) is known to approximate brittle fracture. In Fig.6.1, Γ_α represents a crack which is a priori unknown. The model (6.15) is able to capture a crack Γ_α which is typically a line in 2D and a surface in 3D, where α takes a unit value, which is surrounded by a band of thickness 2ℓ where α is strictly positive. Thus, the length ℓ characterizes the thickness of the region that shall undergo damage.

Two choices of dissipation function $w(\alpha)$ are usually considered and the second one is chosen for the numerical test cases in this chapter:

- `DQ` model (proposed in a different context by Ambrosio and Tortorelli [18])

$$w(\alpha) = \alpha^2, \qquad c_w = \frac{1}{2}. \tag{6.16}$$

  In this case, the damage onsets as soon as the external loading increments from zero.

- `DL` model (proposed in [148])

$$w(\alpha) = \alpha, \qquad c_w = \frac{2}{3}. \tag{6.17}$$

  In this case, the damage onsets only when a critical elastic energy density is exceeded.

The two choices of $w(\alpha)$, (6.17) and (6.16), make $w(\alpha)$ convex. Therefore, the functional $\alpha \mapsto \mathcal{W}_\ell(\boldsymbol{u}, \alpha)$ is convex. Similarly, for a fixed α, $\boldsymbol{u} \mapsto \mathcal{W}_\ell(\boldsymbol{u}, \alpha)$ is convex. This implies that fixing one variable at a time, one can alternately solve the minimization problem (6.15) [34]. However, $\mathcal{W}_\ell(\boldsymbol{u}, \alpha)$ is not convex with respect to the two variables $(\boldsymbol{u}, \alpha)$ together.

Remark 16. *The existence of at least one minimizer of* (6.11) *or, equivalently, of a solution to* (6.15) *in some weak sense was obtained in [79] for the so-called antiplanar shear case. Using Gamma-convergence, it can be proved that, as the length ℓ tends to zero, the global minimizers of* (6.11) *tend to a global minimizer of the original Francfort-Marigo energy*

$$\mathcal{W}_0 = \mathcal{E}(\boldsymbol{u}, \alpha) + G_c \int_{\Gamma_\alpha} ds \tag{6.18}$$

(with a free discontinuity Γ_α modeling fracture) [44, 79], provided that c_w is given by (6.9).

Remark 17. *For the unregularized fracture model (which involves minimization of (6.18)), it is well-known [77] that no surface or body forces can be applied. This is because in the case of complete damage ($\mathbb{C}(1) = \mathbf{0}$), applying a force leads to the breakdown of the structure, which is not physical.*

On the contrary, for the regularized fracture model (6.15), one can indeed apply a surface or a body force [148]. If the magnitude of force is relatively small, one can obtain a solution numerically (with small damaged zones) with ease. However, if the magnitude of applied forces is large, it is quite hard to attain convergence numerically. This is explained by the fact that one possible solution to the regularized problem (6.15) is the one in which the structure breaks, and the Dirichlet boundary Γ_D and the Neumann boundary Γ_N are disconnected, while the total energy is minimized to $-\infty$.

6.2.2 Penalization

The damage criterion (6.15b) is a variational inequality and is not shape-differentiable in the classical sense. For a class of inequalities, called of the first kind and of the second kind, one can determine the so-called conical derivative [134, 171]. Well-known examples of inequality of the first kind include the obstacle-problem and the frictionless contact mechanics problem. The damage inequality (6.15b) without the irreversibility constraint (only the box constraint $\alpha \in [0, 1]$) and for a convex $\mathbb{C}(\alpha)$ classifies as an inequality of the second kind, hence easy to analyze. But with the irreversibility constraint, inequality (6.15b) classifies neither as the first kind nor as the second kind. The analytical treatment of (6.15b) is thus complex and out of the scope of this thesis. Instead, we prefer to convert the inequality (6.15b) into an equation, using penalization. Let ϵ be a penalization factor such that $0 < \epsilon \ll 1$ and let $\mathcal{M}$ be the max function, defined as

$$\mathcal{M}(\beta) = \max(\beta, 0).$$

From now on, we replace the original model (6.15) by the following penalized problem: for all $t \in (0, T]$, find $(\boldsymbol{u}_\epsilon(t), \alpha_\epsilon(t)) \in Z$ such that $\alpha_\epsilon(0) = \alpha_0$, $\boldsymbol{u}_\epsilon(t) = \bar{\boldsymbol{u}}(t)$ on $\Gamma_{\bar{\boldsymbol{u}}}$, and

$$\int_\Omega \frac{G_c}{2c_w}\left(\ell \nabla\alpha_\epsilon \cdot \nabla\beta + \frac{w'(\alpha_\epsilon)\beta}{2\ell}\right) dx + \int_\Omega \frac{1}{2}\mathbb{C}'(\alpha_\epsilon)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_\epsilon) : \boldsymbol{\varepsilon}(\boldsymbol{u}_\epsilon)\, dx + \int_\Omega \frac{G_c}{\epsilon}\left(\mathcal{M}(\alpha_\epsilon - 1) - \mathcal{M}(-\dot{\alpha}_\epsilon)\right)\beta\, dx = 0 \quad \forall \beta \in H^1_D(\Omega), \tag{6.19a}$$

$$\int_\Omega \mathbb{C}(\alpha_\epsilon)\boldsymbol{\varepsilon}(\boldsymbol{u}_\epsilon) : \boldsymbol{\varepsilon}(\boldsymbol{v})\, dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, ds + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\, ds \qquad \forall \boldsymbol{v} \in V. \tag{6.19b}$$

We call (6.19), the state problem, and the solution to it $(\boldsymbol{u}_\epsilon, \alpha_\epsilon)$, the state solution. Note that we implicitly assume in writing (6.19a) that α_ϵ admits a time derivative. In the sequel it is assumed that there exists a unique state solution. As is clear in (6.19a), only the bound constraints $\dot{\alpha}_\epsilon \geq 0$ and $\alpha_\epsilon \leq 1$ are penalized. The constraint $\alpha_\epsilon \geq 0$ is not penalized explicitly, rather implicitly by penalizing $\dot{\alpha}_\epsilon \geq 0$ and defining an initial condition $\alpha_0 \geq 0$. A similar penalization approach was studied numerically [132], where authors penalize only the irreversibility criterion. Our work defers as we penalize the upper bound $\alpha_\epsilon \leq 1$ as well.

Remark 18. *The conversion of (6.15) to (6.19) has been made using the* max *function, but one can also consider a regularized form of the* max *function. This conversion not only simplifies the computation of the shape-derivative, but also helps in the numerical resolution using alternate minimizations (precisely, step 2 of the algorithm 6 using a Newton scheme). Without penalization, the damage model (6.15) can be solved with a sequential quadratic programming solver, capable of taking simple bound constraints into account [34].*

6.2.3 Existence of solution to the time-discretized version of the fracture problem (6.19)

Gamma-convergence results are used to prove that the global minimum of the regularized fracture problem (6.15) converges to the global minimum of the Franfort-Marigo energy [79, 44] as ℓ tends to zero, provided the following normalization holds

$$c_w = \int_0^1 w(\beta) d\beta.$$

However, demonstrating the existence of at least one minimum of (6.11) or equivalently, a solution to (6.15) is not at all straight-forward. For the antiplanar shear case where the displacement $\boldsymbol{u}$ can be simplified using

a scalar variable, the existence of a minimum was shown by A. Giacomini [79]. The existence of a solution was demonstrated by C. Larsen and co-workers [112] for the following dynamic crack-problem:

$$\begin{aligned} \rho\ddot{\boldsymbol{u}} - \operatorname{div}\left(((1-\alpha)^2+\kappa)\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}+k\dot{\boldsymbol{u}})\right) &= \boldsymbol{f} \quad \text{in } \Omega, \\ \left((1-\alpha)^2+\kappa\right)\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}+k\dot{\boldsymbol{u}}) &= \boldsymbol{g} \quad \text{in } \Gamma_N, \\ \alpha = \underset{\beta\in\mathcal{D}_1}{\arg\min}\left(\mathcal{E}(\boldsymbol{u},\beta)+\mathcal{H}(\beta)\right), &\quad \dot{\alpha}>0, \end{aligned} \tag{6.20}$$

along with the irreversibility condition (6.14a) and the boundary conditions (6.14g), (6.14h) and (6.14i). In the above, $\rho, k \in L^\infty(\Omega)$ are uniformly positive functions and a residual stiffness is such that $0 < \kappa \ll \ell$. The quasi-static problem (6.15) can be recovered from the above (6.20) by substituting $\rho = 0, k = 0$ and

$$\mathbb{C}(\alpha) = ((1-\alpha)^2+\kappa)\mathbb{C}_0. \tag{6.21}$$

Since the proof in [112] is facilitated only because of the term $k\dot{\boldsymbol{u}}$ (as this helps to bound the term $\dot{\alpha}$), this proof can thus be not applied to (6.15) (or to its penalized version (6.19)). If $k = 0$, then one cannot show a bound on $\dot{\alpha}$. For the problem (6.15), it is known that α is discontinuous in time, even if the total energy (6.11) is continuous in time [77] and $\dot{\alpha}$ can thus be unbounded.

Moreover, in the dynamic version (6.20), the damage variable α is constrained to take values in $[0, 1]$. This allows the authors to use the Lebesque's dominated convergence theorem [112]. Whereas in our penalized quasi-static formulation (6.19), this constraint is absent and the damage variable can indeed take values in $\mathbb{R}^+$. This complicates further the demonstration of existence of a solution to (6.19). In order to overcome this difficulty, we shall assume that the damage variable is bounded (see Remark 19).

Given the above difficulties in proving the existence of solution to (6.19), it is imperative to slightly simplify the problem (6.19) in order to show existence. We consider a time-discretized version of the equation (6.19). The time interval $[0, T]$ is split in N intervals, each of size κ and the end of interval is denoted by t_n. Let $\boldsymbol{f}_n = \boldsymbol{f}(t_n) \in L^2(\Omega)^d, \boldsymbol{g}_n = \boldsymbol{g}(t_n) \in H^1(\Omega)^d$. Then for $\mathbb{C}(\alpha)$ given by (6.21): for $(\boldsymbol{u}_0, \alpha_0) \in Z$, find $(\boldsymbol{u}_n, \alpha_n) \in Z, 1 \le n \le N$ such that $\boldsymbol{u}_n = \bar{\boldsymbol{u}}(t_n)$ on $\Gamma_{\bar{\boldsymbol{u}}}$, and

$$\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v})\, dx = \int_\Omega \boldsymbol{f}_n \cdot \boldsymbol{v}\, dx + \int_{\Gamma_N} \boldsymbol{g}_n \cdot \boldsymbol{v}\, ds \qquad \forall \boldsymbol{v} \in V \tag{6.22a}$$

and

$$\begin{aligned} &\int_\Omega \frac{G_c}{2c_w}\left(\ell\nabla\alpha_n\cdot\nabla\beta + \frac{w'(\alpha_n)}{2\ell}\right) dx + \int_\Omega \frac{1}{2}\mathbb{C}'(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n):\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\, dx \\ &\quad + \int_\Omega \frac{G_c}{\epsilon}\left(\mathcal{M}(\alpha_n - 1)\beta\, dx - \int_\Omega \mathcal{M}\left(\frac{\alpha_{n-1}-\alpha_n}{\delta t}\right)\beta\right) dx = 0 \qquad \forall \beta \in H^1_D(\Omega). \end{aligned} \tag{6.22b}$$

The above time-discretization shall be treated in all the numerical results of this chapter. Like its continuous counterpart (6.19), the above cannot be shown to admit a unique solution. In the order to show existence of one solution, we shall construct a sequence, converging to a limit, satisfying the above. The construction of the proof is inspired from [112], modulo a few differences.

Theorem 27. *Let $\mathbb{C}(\alpha) = ((1-\alpha)^2+\kappa)\mathbb{C}_0$, the dissipation function $w(\cdot)$ be convex and $\boldsymbol{f}_n \in L^2(\Omega)^d$, $\boldsymbol{g}_n \in H^1(\Omega)^d$.*

Assume that if $\boldsymbol{u}_n \in V$, the solution to (6.22), $\alpha_n \in L^\infty(\Omega)$. Then the problem (6.22) admits at least one family of solutions $(\boldsymbol{u}_n, \alpha_n) \in Z$, for $1 \le n \le N$. In addition, this family of solutions satisfies

$$\|\boldsymbol{u}_n\|_V \le C^u_n(\kappa), \quad \|\nabla\alpha_n\|_{L^2(\Omega)} \le C^a_n(\kappa, \ell) \tag{6.23}$$

where the constants C^u_n and C^a_n do not depend on $\epsilon, \delta t$.

Remark 19. *In the above theorem, we have assumed that $\alpha_n \in L^\infty(\Omega)$. This is not a strong hypothesis as α_n satisfies the elliptic equation (6.22b). Here, one may use the maximum principle or using the De Giorgi's principle [63] to demonstrate that α_n is bounded. How exactly can this be done, can be a subject of further research.*

Proof. In order to simplify the presentation of the proof, we drop index n from the problem (6.22) resulting in: find $(\boldsymbol{u}, \alpha) \in Z$ such that $\boldsymbol{u} = \bar{\boldsymbol{u}}(t_n)$ on $\Gamma_{\bar{\boldsymbol{u}}}$, and

$$\int_\Omega \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v})\, dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, dx + \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v}\, ds \qquad \forall \boldsymbol{v} \in V \tag{6.24a}$$

$$\text{and} \quad \int_\Omega \frac{G_c}{2c_w}\left(\ell\nabla\alpha\cdot\nabla\beta + \frac{w'(\alpha)}{2\ell}\right)\,dx + \int_\Omega \frac{1}{2}\mathbb{C}'(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u})\colon\boldsymbol{\varepsilon}(\boldsymbol{u})\,dx + \int_\Omega \frac{G_c}{\epsilon}\left(\mathcal{M}(\alpha-1)\beta\,dx - \int_\Omega \mathcal{M}\left(\frac{\alpha_{n-1}-\alpha}{\delta t}\right)\beta\right)dx = 0 \qquad \forall\beta\in H^1_D(\Omega), \quad (6.24b)$$

where $(\boldsymbol{u},\alpha)\equiv(\boldsymbol{u}_n,\alpha_n),(\boldsymbol{f},\boldsymbol{g})\equiv(\boldsymbol{f}_n,\boldsymbol{g}_n)$ is understood. The idea of the proof is to construct a sequence $(\boldsymbol{u}^i,\alpha^i)\in Z$ that weakly converges to $(\boldsymbol{u},\alpha)\in Z$ as $i\to\infty$. In order to construct such a sequence we solve: for $(\boldsymbol{u}^0,\alpha^0)=(\boldsymbol{0},0)$ find $(\boldsymbol{u}^i,\alpha^i)\in V\times H^1_D(\Omega), i>0$ such that $\boldsymbol{u}^i=\bar{\boldsymbol{u}}(t_n)$ on $\Gamma_{\bar{\boldsymbol{u}}}$, and

$$\text{and} \quad \int_\Omega \mathbb{C}(\alpha^{i-1})\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\colon\boldsymbol{\varepsilon}(\boldsymbol{v})\,dx = \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\,dx + \int_{\Gamma_N}\boldsymbol{g}\cdot\boldsymbol{v}\,ds \qquad \forall\boldsymbol{v}\in V. \quad (6.25a)$$

$$\frac{G_c\ell}{2c_w}\int_\Omega \nabla\alpha^i\cdot\nabla\beta\,dx + \frac{G_c}{4c_w\ell}\int_\Omega w'(\alpha^i)\beta\,dx + \int_\Omega \mathbb{C}'(\alpha^i)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\,dx + \frac{G_c}{\epsilon}\left(\int_\Omega \mathcal{M}(\alpha^i-1)\beta\,dx - \frac{1}{\delta t}\int_\Omega \mathcal{M}(\alpha_{n-1}-\alpha^i)\beta\,dx\right) = 0 \qquad \forall\beta\in H^1_D(\Omega). \quad (6.25b)$$

The equation (6.25b) can be equivalently formulated as the minimization

$$\alpha^i = \underset{\beta\in H^1_D(\Omega)}{\arg\min}\left(\mathcal{E}(\boldsymbol{u}^i,\beta) + \mathcal{H}(\beta) + \frac{G_c}{2\epsilon}\left(\int_\Omega \mathcal{F}(\beta-1)dx + \int_\Omega \mathcal{F}\left(\frac{\alpha_{n-1}-\beta}{\delta t}\right)dx\right)\right), \quad (6.26)$$

where $\mathcal{E}$ and $\mathcal{H}$ are defined in (6.8) and (6.9) respectively, and $\mathcal{F}'(\cdot)=\mathcal{M}(\cdot)$. Given that $\mathbb{C}(\alpha)\boldsymbol{\xi}:\boldsymbol{\xi}\geq\kappa c_{\min}|\boldsymbol{\xi}|^2>0$, we can easily show that there exists a unique solution $\boldsymbol{u}^i\in V$ to (6.25a). Given the convexity of $\mathbb{C}(\beta)$ and of $w(\beta)$, the two functionals $\mathcal{H}(\beta)$ and $\mathcal{E}(\boldsymbol{u}^i,\beta)$ are convex in β. Alongside, the functional $M_\epsilon(\cdot)$ is convex in its argument. The sum of the functionals on the right in (6.26) is thus convex and is minimized on the entire space $H^1_D(\Omega)$. Hence, there exists a unique solution $\alpha^i\in H^1_D(\Omega)$ that solves (6.26) and (6.25b). We now define $(\delta\boldsymbol{u}^i,\delta\alpha^i)\in V\times H^1_D(\Omega)$ given by

$$\delta\alpha^i = \alpha^{i+1}-\alpha^i \quad \text{and} \quad \delta\boldsymbol{u}^i = \boldsymbol{u}^{i+1}-\boldsymbol{u}^i.$$

We first establish a bound on the solution $\boldsymbol{u}^i_n$. Substituting $\boldsymbol{v}=\delta\boldsymbol{u}_{n-1}$ in (6.25a), we get

$$\begin{aligned}\int_\Omega \boldsymbol{f}\cdot\delta\boldsymbol{u}^{i-1}dx + \int_{\Gamma_N}\boldsymbol{g}\cdot\delta\boldsymbol{u}^{i-1}dx &= \int_\Omega \mathbb{C}(\alpha^i)\boldsymbol{\varepsilon}(\boldsymbol{u}^i):\boldsymbol{\varepsilon}(\delta\boldsymbol{u}^{i-1})dx\\ &= \mathcal{E}(\boldsymbol{u}^i,\alpha^i) - \mathcal{E}(\boldsymbol{u}^{i-1},\alpha^{i-1}) + \mathcal{E}(\delta\boldsymbol{u}^{i-1},\alpha^i)\\ &\quad - \frac{1}{2}\int_\Omega(\mathbb{C}(\alpha^i)-\mathbb{C}(\alpha^{i-1}))\boldsymbol{\varepsilon}(\boldsymbol{u}^{i-1})\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^{i-1})dx\end{aligned}$$

For the last term on the right hand side in the above, we use

$$\mathbb{C}(\alpha^i)-\mathbb{C}(\alpha^{i-1}) = (1-\alpha^i)^2\mathbb{C}_0-(1-\alpha^{i-1})^2\mathbb{C}_0 = (\delta\alpha^{i-1}(\alpha^i+\alpha^{i-1})+2\delta\alpha^{i-1})\mathbb{C}_0 = \mathbb{C}'(\alpha^i)\delta\alpha^{i-1}-(\delta\alpha^{i-1})^2\mathbb{C}_0,$$

and obtain

$$\begin{aligned}\int_\Omega \boldsymbol{f}\cdot\delta\boldsymbol{u}^{i-1}dx + \int_{\Gamma_N}\boldsymbol{g}_n\cdot\delta\boldsymbol{u}^{i-1}dx &= \mathcal{E}(\boldsymbol{u}^i,\alpha^i) - \mathcal{E}(\boldsymbol{u}^{i-1},\alpha^{i-1}) + \mathcal{E}(\delta\boldsymbol{u}^{i-1},\alpha^i)\\ &\quad + \frac{1}{2}\left(\int_\Omega(\delta\alpha^{i-1})^2\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^i)dx\right.\\ &\quad \left. - \int_\Omega \mathbb{C}'(\alpha^i)\delta\alpha^{i-1}\boldsymbol{\varepsilon}(\boldsymbol{u}^{i-1})\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^{i-1})dx\right).\end{aligned} \quad (6.27)$$

Taking $\beta=\delta\alpha^{i-1}$ in (6.25b),

$$\begin{aligned}\int_\Omega \mathbb{C}'(\alpha^i)\delta\alpha^{i-1}\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^i)dx &= \frac{G_c}{\epsilon}\left(-\int_\Omega \mathcal{M}(\alpha^i-1)\delta\alpha^{i-1}dx + \frac{1}{\delta t}\int_\Omega \mathcal{M}(-\delta\alpha^{i-1})\delta\alpha^{i-1}dx\right)\\ &\quad - \frac{G_c}{4c_w\ell}\int_\Omega w'(\alpha^i)\delta\alpha^{i-1}dx - \frac{G_c\ell}{2c_w}\int_\Omega \nabla\alpha^i\cdot\nabla\delta\alpha^{i-1}dx\end{aligned}$$

Substituting the above in (6.27),

$$\begin{aligned}
&\mathcal{E}(\boldsymbol{u}^i,\alpha^i)-\mathcal{E}(\boldsymbol{u}^{i-1},\alpha^{i-1})+\mathcal{E}(\delta\boldsymbol{u}^{i-1},\alpha^i)+\frac{1}{2}\left(\int_\Omega(\delta\alpha^{i-1})^2\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^i)dx\right)\\
&=\int_\Omega \boldsymbol{f}\cdot\delta\boldsymbol{u}^{i-1}dx+\int_{\Gamma_N}\boldsymbol{g}\cdot\delta\boldsymbol{u}^{i-1}dx+\frac{G_c}{\epsilon}\left(-\int_\Omega\mathcal{M}(\alpha^i-1)\delta\alpha^{i-1}dx+\frac{1}{\delta t}\int_\Omega\mathcal{M}(-\delta\alpha^{i-1})\delta\alpha^{i-1}dx\right)\\
&\quad-\frac{G_c}{4c_w\ell}\int_\Omega w'(\alpha^i)\delta\alpha^{i-1}dx-\frac{G_c\ell}{2c_w}\int_\Omega\nabla\alpha^i\cdot\nabla\delta\alpha^{i-1}dx
\end{aligned}$$

Exploiting the convexity of the function $M_\epsilon(\cdot)$ and of the function $w(\cdot)$, we get the bound

$$\begin{aligned}
&\mathcal{E}(\boldsymbol{u}^i,\alpha^i)-\mathcal{E}(\boldsymbol{u}^{i-1},\alpha^{i-1})+\mathcal{E}(\delta\boldsymbol{u}^{i-1},\alpha^i)+\frac{1}{2}\left(\int_\Omega(\delta\alpha^{i-1})^2\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^i)dx+\frac{G_c\ell}{2c_w}\int_\Omega|\nabla\delta\alpha^{i-1}|^2dx\right)\\
&\le\int_\Omega \boldsymbol{f}\cdot\delta\boldsymbol{u}^{i-1}dx+\int_{\Gamma_N}\boldsymbol{g}\cdot\delta\boldsymbol{u}^{i-1}dx+\frac{G_c}{\epsilon}\int_\Omega(\mathcal{F}(\alpha^{i-1}-1)-\mathcal{F}(\alpha^i-1))dx+\mathcal{H}(\alpha^{i-1})-\mathcal{H}(\alpha^i)
\end{aligned}$$

Summing the above from $i=1,\cdots,\mathcal{I}$, using the fact $(\boldsymbol{u}^0,\alpha^0)=(\boldsymbol{0},0)$ and the positivity of $\mathcal{H}(\cdot),M_\epsilon(\cdot)$, we get

$$\begin{aligned}
&\mathcal{E}(\boldsymbol{u}^{\mathcal{I}},\alpha^{\mathcal{I}})+\mathcal{H}(\alpha^{\mathcal{I}})+\frac{G_c}{\epsilon}\int_\Omega\mathcal{F}(\alpha^{\mathcal{I}}-1)dx+\sum_{i=1}^{\mathcal{I}}\left(\mathcal{E}(\delta\boldsymbol{u}^{i-1},\alpha^{i-1})+\int_\Omega(\delta\alpha^{i-1})^2\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^i)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}^i)dx\right.\\
&\qquad\left.+\frac{G_c\ell}{2c_w}\int_\Omega|\nabla\delta\alpha^{i-1}|^2dx\right)\le\sum_{i=1}^{\mathcal{I}}\left(\int_\Omega\boldsymbol{f}\cdot\delta\boldsymbol{u}^{i-1}dx+\int_{\Gamma_N}\boldsymbol{g}\cdot\delta\boldsymbol{u}^{i-1}dx\right)
\end{aligned}\tag{6.28}$$

The term on the right in the above can bounded as

$$\begin{aligned}
\sum_{i=1}^{\mathcal{I}}\left(\int_\Omega\boldsymbol{f}\cdot\delta\boldsymbol{u}^{i-1}dx+\int_{\Gamma_N}\boldsymbol{g}\cdot\delta\boldsymbol{u}^{i-1}dx\right)&=\int_\Omega\boldsymbol{f}\cdot\boldsymbol{u}^{\mathcal{I}}dx+\int_{\Gamma_N}\boldsymbol{g}\cdot\boldsymbol{u}^{\mathcal{I}}dx\\
&\le\left(\|\boldsymbol{f}\|_{L^2(\Omega)^d}+C_{tr}\|\boldsymbol{g}\|_{L^2(\Gamma_N)}\right)\|\boldsymbol{u}^{\mathcal{I}}\|_{L^2(\Omega)^d}\\
&\le C_kC_p\|\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})\|_{L^2(\Omega)^d}\left(\|\boldsymbol{f}\|_{L^2(\Omega)^d}+C_{tr}\|\boldsymbol{g}\|_{L^2(\Gamma_N)}\right),
\end{aligned}$$

where C_k Korn's inequality constant and C_p is Poincaré's inequality constant. The above yields

$$\begin{aligned}
\kappa\|\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})\|^2_{L^2(\Omega)^d}&\le\mathcal{E}(\boldsymbol{u}^{\mathcal{I}},\alpha^{\mathcal{I}})\\
&\le\|\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})\|_{L^2(\Omega)^d}C_kC_p\left(\|\boldsymbol{f}\|_{L^2(\Omega)^d}+C_{tr}\|\boldsymbol{g}\|_{L^2(\Gamma_N)}\right).
\end{aligned}$$

Thus, we deduce for all $\mathcal{I}>0$,

$$\begin{aligned}
\|\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})\|_{L^2(\Omega)^d}&\le\frac{C_1}{\kappa}\\
\text{where } C_1=C_kC_p&\left(\|\boldsymbol{f}\|_{L^2(\Omega)^d}+C_{tr}\|\boldsymbol{g}\|_{L^2(\Gamma_N)}\right).
\end{aligned}\tag{6.29}$$

It is important to note that the bound C_1 does not depend on penalization parameter ϵ, the time step δt or the characteristic length ℓ. Using the Korn's inequality, (6.30) results in the first bound in (6.23),

$$\|\boldsymbol{u}\|_V=C_n^u(\kappa)=\frac{C_kC_1}{\kappa},\tag{6.30}$$

where C_n^u depends on the index n due to the forces $\boldsymbol{f}_n,\boldsymbol{g}_n$. We then establish a bound $\alpha^{\mathcal{I}}$. Since all the terms on the left in (6.28) are positive, we observe that

$$\frac{G_c\ell}{4c_w}\|\nabla\alpha^{\mathcal{I}}\|^2_{L^2(\Omega)}=\mathcal{H}(\alpha^{\mathcal{I}})\le\sum_{i=1}^{\mathcal{I}}\left(\int_\Omega\boldsymbol{f}\cdot\delta\boldsymbol{u}^{i-1}dx+\int_{\Gamma_N}\boldsymbol{g}\cdot\delta\boldsymbol{u}^{i-1}dx\right)\le\frac{C_1^2}{\kappa},$$

simplifying to the second bound in (6.23)

$$\left\|\nabla\alpha^{\mathcal{I}}\right\|^2_{L^2(\Omega)} \leq C_n^a(\kappa,\ell)^2 = 2C_1\sqrt{\frac{c_w}{\kappa G_c\ell}}, \tag{6.31}$$

for all $\mathcal{I}>0$. Here, the bound C_n^a does not depend on ϵ, δt or ℓ, but it depends on κ,ℓ and the index n (due to C_1). Since the estimates (6.30)-(6.31) hold for all $\mathcal{I}>0$, we can state that there exists a limit $(\boldsymbol{u},\alpha)\in Z$ and a subsequence $(\boldsymbol{u}^{\mathcal{I}},\alpha^{\mathcal{I}})$ (we keep the index $\mathcal{I}$ to simplify notation) such that as $\mathcal{I}\longrightarrow\infty$:

$$\boldsymbol{u}^{\mathcal{I}}\rightharpoonup\boldsymbol{u} \quad \text{in} \quad V, \quad \text{and} \quad \alpha^{\mathcal{I}}\rightharpoonup\alpha \quad \text{in} \quad H^1_D(\Omega), \tag{6.32a}$$

$$\alpha^{\mathcal{I}}\longrightarrow\alpha \quad \text{in} \quad L^2(\Omega), \tag{6.32b}$$

where strong convergence follows from Rellich's theorem. Using the hypothesis that the solution $\alpha^{\mathcal{I}}\in L^\infty(\Omega)$, we conveniently state

$$\left\|\alpha^{\mathcal{I}}\right\|_{L^\infty(\Omega)}\leq K,$$

where $K>0$ is some constant. Let $\gamma^{\mathcal{I}}=(1-\alpha^{\mathcal{I}})^2$, $\gamma=(1-\alpha)^2$. Then the above estimate implies

$$0\leq\gamma^{\mathcal{I}}\leq(K-1)^2 \quad \text{a.e. in} \quad \Omega,$$

which, together with the strong convergence (6.32b) results in

$$\gamma^{\mathcal{I}}\longrightarrow\gamma \quad \text{in} \quad L^2(\Omega).$$

We now pass to the limit in the equation (6.25a),

$$\begin{aligned}
\forall\boldsymbol{v}\in V, \quad \int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}\,dx+\int_{\Gamma_N}\boldsymbol{g}\cdot\boldsymbol{v}\,ds &= \lim_{\mathcal{I}\longrightarrow\infty}\int_\Omega \mathbb{C}(\alpha^{\mathcal{I}-1})\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{v})\;dx\\
&= \lim_{\mathcal{I}\longrightarrow\infty}\int_\Omega \gamma^{\mathcal{I}}\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{v})\;dx\\
&= \lim_{\mathcal{I}\longrightarrow\infty}\left(\int_\Omega \gamma\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{v})\;dx\right)\\
&\quad+\lim_{\mathcal{I}\longrightarrow\infty}\left(\int_\Omega (\gamma^{\mathcal{I}}-\gamma)\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{v})\;dx\right)\\
&= \int_\Omega \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{v})\;dx+\lim_{\mathcal{I}\longrightarrow\infty}\left(\int_\Omega (\gamma^{\mathcal{I}}-\gamma)\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{v})\;dx\right),
\end{aligned}$$

where we have used the weak convergence in $\boldsymbol{u}^{\mathcal{I}}$. Now, we look at the last term in the above,

$$\int_\Omega(\gamma^{\mathcal{I}}-\gamma)\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{v})\;dx\leq c_{\max}\left\|(\gamma^{\mathcal{I}}-\gamma)|\boldsymbol{\varepsilon}(\boldsymbol{v})|\right\|_{L^2(\Omega)}\left\|\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})\right\|_{L^2(\Omega)}$$

Since $\gamma^{\mathcal{I}}-\gamma\to0$ a.e. in Ω, $|\gamma^{\mathcal{I}}-\gamma|\leq 2C_s^2C_2^2$, $|\nabla\boldsymbol{v}|^2\in L^1(\Omega)$, then using Lebesque Dominated Theorem, we get

$$\lim_{\mathcal{I}\to\infty}\left(c_{\max}\left\|(\gamma^{\mathcal{I}}-\gamma)|\boldsymbol{\varepsilon}(\boldsymbol{v})|\right\|_{L^2(\Omega)}\left\|\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})\right\|_{L^2(\Omega)}\right)=0. \tag{6.33}$$

We now need to pass to the limit in the equivalent formulation of (6.25b), i.e., (6.26)

$$\left(\mathcal{E}(\boldsymbol{u}^{\mathcal{I}},\alpha^{\mathcal{I}})+\mathcal{H}(\alpha^{\mathcal{I}})+\frac{G_c}{\epsilon}\left(\int_\Omega\mathcal{F}(\alpha^{\mathcal{I}}-1)\,dx+\int_\Omega\mathcal{F}\left(\frac{\alpha_{n-1}-\alpha^{\mathcal{I}}}{\delta t}\right)\,dx\right)\right).$$

For the second term $\mathcal{H}(\beta)$ and the third term containing $\mathcal{F}$, one can easily use the property of lower semi continuity to pass to the limit, since they are convex in $\alpha^{\mathcal{I}}$. For the first term $\mathcal{E}(\boldsymbol{u}^{\mathcal{I}},\beta)$, we use

$$\begin{aligned}
\int_\Omega\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}):\boldsymbol{\varepsilon}(\boldsymbol{u})dx &\leq \liminf_{\mathcal{I}\to\infty}\int_\Omega\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})dx \qquad \text{(given the weak convergence } \boldsymbol{u}^{\mathcal{I}}\rightharpoonup\boldsymbol{u} \text{ in } V)\\
&\leq \lim_{\mathcal{I}\to\infty}\left(\int_\Omega(\mathbb{C}(\alpha)-\mathbb{C}(\alpha^{\mathcal{I}}))\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})dx+\int_\Omega\mathbb{C}(\alpha^{\mathcal{I}})\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})dx\right)\\
&= \lim_{\mathcal{I}\to\infty}\left(\int_\Omega(\gamma-\gamma^{\mathcal{I}})\mathbb{C}_0\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})dx+\int_\Omega\mathbb{C}(\alpha^{\mathcal{I}})\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}}):\boldsymbol{\varepsilon}(\boldsymbol{u}^{\mathcal{I}})dx\right).
\end{aligned}$$

In the above, one can use the convergence (6.33) and derive the desired result. □

In the sequel, to simplify notations, we shall drop all ϵ indices and simply denote by $(\boldsymbol{u},\alpha)$ the solution $(\boldsymbol{u}_\epsilon,\alpha_\epsilon)$ of problem (6.19).

6.3 Optimization problem

6.3.1 Setting of the problem

We minimize an objective functional $J(\Omega)$ given by

$$J(\Omega) = \int_0^T \int_\Omega m(\boldsymbol{u}(\Omega), \alpha(\Omega))\, dx\, dt, \tag{6.34}$$

where $(\boldsymbol{u}(\Omega), \alpha(\Omega))$ is the solution of (6.19) and the function $m(\cdot,\cdot)$ is assumed to be $\mathcal{C}^1$ smooth with quadratic growth and linear growth for its derivative, so as to ensure that the objective function (6.34) is well-defined and the adjoint equation is well-posed. This objective functional represents a mechanical property such as total power, total elastic energy, or total fracture energy. In Section 6.5, we shall maximize the total elastic energy and thus choose

$$m(\boldsymbol{u}, \alpha) = -\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}).$$

The justification for the above choice shall be given at the beginning of Section 6.5.

In practice, the shape Ω must be found inside a pre-fixed design space $D \subset \mathbb{R}^d$. Figure 6.2 shows the

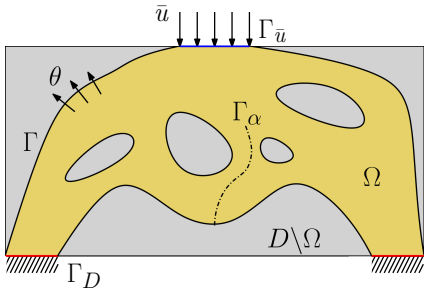


Figure 6.2: Design domain D and shape Ω

shape Ω (in gold) and the design space D (in gold and grey). The crack Γ_α is an unknown of the problem (6.19) that shall be determined for every shape Ω. This crack Γ_α might appear anywhere in the shape Ω and there is no postulated initial crack. The space of admissible shapes $\mathcal{U}_{ad}$ is defined as

$$\mathcal{U}_{ad} = \left\{ \Omega \subset D : \int_\Omega dx = V_f \right\},$$

where Ω is an open set and V_f is the target volume. The minimization problem then reads

$$\min_{\Omega \in \mathcal{U}_{ad}} J(\Omega). \tag{6.35}$$

As is well known, very often there exists an optimal shape only if additional uniform smoothness conditions are imposed to the admissible shapes, that we shall not consider in the sequel. As usual, we content ourselves with computing numerical minimizers, using a gradient-descent method.

6.3.2 Shape derivative computation

To define the adjoint problem for the shape derivative, we introduce the subspace V_0 of V

$$V_0 = \{\boldsymbol{u} \in H^1(\Omega)^d : \boldsymbol{u} = \boldsymbol{0} \text{ on } \Gamma_D \cup \Gamma_{\bar{\boldsymbol{u}}}\}. \tag{6.36}$$

The adjoint variational formulation is defined as: find $(\boldsymbol{v}, \beta) \in H^1([0,T], Z)$, satisfying the final condition $\beta(T) = 0$, such that, for all $t \in [0,T)$,

$$\int_\Omega (\partial_{\boldsymbol{u}} m(\boldsymbol{u}, \alpha)\boldsymbol{\psi} + \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}) + \mathbb{C}'(\alpha)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}))\, dx = 0 \qquad \forall \boldsymbol{\psi} \in V_0, \tag{6.37a}$$

$$\int_\Omega \Big(\partial_\alpha m(\boldsymbol{u},\alpha)\varphi + \mathbb{C}'(\alpha)\varphi\, \boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) + \frac{G_c \ell}{2c_w} \nabla\beta \cdot \nabla\varphi + \left(\frac{1}{2}\mathbb{C}''(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) + \frac{G_c}{4c_w}\frac{w''(\alpha)}{\ell} + \frac{G_c}{\epsilon}\mathcal{M}'(\alpha-1) \right) \beta\varphi \Big)\, dx - \int_\Omega \frac{G_c}{\epsilon} \left(\frac{d}{dt}(\mathcal{M}'(-\dot\alpha)\beta) \right) \varphi\, dx = 0 \qquad \forall \varphi \in H^1_D(\Omega). \quad (6.37b)$$

In the sequel, we assume that there exists a unique solution of the adjoint equation. Note that (6.37b) features the time derivative of the derivative $\mathcal{M}'$ of the maximum function $\mathcal{M}(\beta) = \max(\beta, 0)$. Since $\mathcal{M}'$ is the Heaviside function, its time derivative is a Dirac mass and the precise meaning of the last integral in (6.37b) is unclear. However, if $\mathcal{M}(\beta)$ was a regularization of $\max(\beta, 0)$, then everything makes sense, including the following proposition. In other words, our computation of the shape derivative below is fine for a smooth function $\mathcal{M}(\beta)$ and merely formal for the maximum function. Further time discretization will make clear in which sense the time derivative of the Heaviside function is computed (see Subsection 6.3.3).

Proposition 1. *Let Ω be a smooth bounded open set. Assume that there exists a unique solution $(\boldsymbol{u}, \alpha)$ to (6.19), which belongs to $H^1([0,T], Z)$, and that there exists a unique solution $(\boldsymbol{v}, \beta) \in H^1([0,T], Z)$ of the adjoint equation (6.57). If the state solution $(\boldsymbol{u}, \alpha) \equiv (\boldsymbol{u}(\Omega), \alpha(\Omega))$ is shape-differentiable, then the objective function (6.34) admits a shape derivative, given, for any $\boldsymbol{\theta} \in W^{1,\infty}_0(D, \mathbb{R}^d)$, by*

$$J'(\Omega)(\boldsymbol{\theta}) = \int_0^T \int_\Gamma \boldsymbol{\theta}\cdot\boldsymbol{n} \left(m(\boldsymbol{u},\alpha) + \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{f}\cdot\boldsymbol{v} + \frac{1}{2}\mathbb{C}'(\alpha)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) + \frac{G_c}{2c_w}\left(\ell \nabla\alpha\cdot\nabla\beta + \frac{w'(\alpha)\beta}{2\ell} \right) + \frac{G_c}{\epsilon}\Big(\mathcal{M}(\alpha-1) - \mathcal{M}(-\dot\alpha) \Big)\beta \right) ds\, dt. \quad (6.38)$$

Remark 20. *The uniqueness of the solution of the damage model (6.19) is far from being obvious since this model is the optimality condition for the minimization of a non-convex energy, which thus may have multiple minima. For the same reason, the existence of a solution for the adjoint equation (6.57) is not obvious either because the corresponding operator is not coercive (nevertheless, see Subsection 6.3.3 for a positive result in this direction). Note that we also assume that the solutions are smooth with respect to time since they belong to $H^1([0,T], Z)$ and, in particular, are continuous with respect to time. We use this assumption in the variational formulations (6.19a) and (6.37b), which involve the time derivative of the damage variable α. Unfortunately, as discussed further in Remark 26, it is likely that, in some cases, the solution $(\boldsymbol{u}(\Omega), \alpha(\Omega))$ is discontinuous in time and thus the shape derivative (6.38) is not rigorously justified.*

Proof. The idea of the proof is well-known, based on Céa's method [53]. Introduce a few spaces on the full space $\mathbb{R}^d$ and thus independent of Ω:

$$\tilde V = \{\boldsymbol{v} \in H^1(\mathbb{R}^d)^d : \boldsymbol{v} = \boldsymbol{0} \ \text{ on } \Gamma_D\}, \quad \tilde{\mathcal{C}}_t = \{\boldsymbol{v} \in \tilde V : \boldsymbol{v} = \bar{\boldsymbol{u}}(t) \ \text{ on } \Gamma_{\bar{\boldsymbol{u}}}\},$$
$$\tilde H^1_D(\mathbb{R}^d) = \{\beta \in H^1(\mathbb{R}^d) : \beta = 0 \ \text{ on } \Gamma_D \cup \Gamma_{\bar{\boldsymbol{u}}}\}, \quad \tilde Z_t = \tilde{\mathcal{C}}_t \times \tilde H^1_D(\mathbb{R}^d),$$
$$\tilde V_0 = \{\boldsymbol{v} \in \tilde V : \boldsymbol{v} = \boldsymbol{0} \ \text{ on } \Gamma_{\bar{\boldsymbol{u}}}\}, \quad \tilde Z_0 = \tilde V_0 \times \tilde H^1_D(\mathbb{R}^d). \quad (6.39)$$

For independent variables $\tilde{\boldsymbol{u}}(t), \tilde{\boldsymbol{v}}(t), \tilde\alpha(t), \tilde\beta(t)$ and $\tilde\lambda$, belonging to the spaces

- $(\tilde{\boldsymbol{u}}, \tilde\alpha) \in H^1([0,T], \tilde Z_t)$,
- $(\tilde{\boldsymbol{v}}, \tilde\beta) \in H^1([0,T], \tilde Z_0)$ (the Lagrange multiplier for the state equation (6.19)),
- $\tilde\lambda \in L^2(\mathbb{R}^d)$ (the Lagrange multiplier for the initial condition $\tilde\alpha(0) = \alpha_0$),

define a Lagrangian as

$$\mathcal{L}(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}, \tilde\alpha, \tilde\beta, \tilde\lambda, \Omega) = \int_0^T \int_\Omega m(\tilde{\boldsymbol{u}}, \tilde\alpha) dx dt + \int_0^T \left(\int_\Omega (\mathbb{C}(\tilde\alpha)\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}) : \boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}}) - \boldsymbol{f}\cdot\tilde{\boldsymbol{v}})\, dx - \int_{\Gamma_N} \boldsymbol{g}\cdot\tilde{\boldsymbol{v}}\, ds \right) dt + \int_0^T \int_\Omega \left(\frac{1}{2}\mathbb{C}'(\tilde\alpha)\tilde\beta\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}) : \boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}) + \frac{G_c}{2c_w}\left(\ell\nabla\tilde\alpha\cdot\nabla\tilde\beta + \frac{w'(\tilde\alpha)\tilde\beta}{2\ell} \right) + \frac{G_c}{\epsilon}\Big(\mathcal{M}(\tilde\alpha - 1) - \mathcal{M}(-\dot{\tilde\alpha}) \Big)\tilde\beta \right) dx\, dt + \int_\Omega \tilde\lambda(\tilde\alpha(0) - \alpha_0)\, dx. \quad (6.40)$$

Since the boundaries $\Gamma_D \cup \Gamma_{\tilde{\boldsymbol{u}}}$ are non-optimizable, the variables $\tilde{\boldsymbol{u}}(t), \tilde{\boldsymbol{v}}(t), \tilde{\alpha}(t), \tilde{\beta}(t)$, and $\tilde{\lambda}$ are independent of Ω. When the optimality condition are applied to the Lagrangian (6.40) (that is, its partial derivatives with respect to its independent variables are set to zero), we obtain the state equation (6.19), the adjoint equation (6.57) and the shape derivative (6.38).

At first, differentiating the Lagrangian (6.40) with respect to the adjoint variable $(\tilde{\boldsymbol{v}}, \tilde{\beta})$ in the direction $(\boldsymbol{\psi}, \varphi) \in H^1([0,T], \tilde{Z}_0)$ and equating it to zero at $(\tilde{\boldsymbol{u}}, \tilde{\alpha}) = (\boldsymbol{u}, \alpha)$, we obtain

$$\frac{\partial \mathcal{L}}{\partial \tilde{\beta}}(\varphi) = \int_0^T \left(\int_\Omega \frac{G_c}{2c_w} \left(\ell \nabla\alpha \cdot \nabla\varphi + \frac{w'(\alpha)\varphi}{2\ell} \right) dx + \int_\Omega \frac{1}{2} \mathbb{C}'(\alpha)\varphi \boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) \; dx \right.$$
$$\left. + \int_\Omega \frac{G_c}{\epsilon} \left(\mathcal{M}(\alpha - 1) - \mathcal{M}(-\dot{\alpha}) \right) \varphi \, dx \right) dt = 0 \qquad \forall \varphi \in H^1([0,T], \tilde{H}^1_D(\mathbb{R}^d))) \; ,$$

$$\frac{\partial \mathcal{L}}{\partial \tilde{\boldsymbol{v}}}(\boldsymbol{\psi}) = \int_0^T \left(\int_\Omega \left(\mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) - \boldsymbol{f} \cdot \boldsymbol{v} \right) dx - \int_{\Gamma_N} \boldsymbol{g} \cdot \boldsymbol{v} ds \right) dt = 0 \qquad \forall \boldsymbol{\psi} \in H^1([0,T], \tilde{V}_0).$$

To obtain the initial condition $\alpha(0) = \alpha_0$, it suffices to differentiate (6.40) with respect to $\tilde{\lambda}$ at $\tilde{\alpha} = \alpha$. We thus recover the state equation (6.19).

Second, we differentiate the Lagrangian (6.40) with respect to $(\tilde{\boldsymbol{u}}, \tilde{\alpha})$ to recover the adjoint equation. By definition, $\tilde{\boldsymbol{u}} \in \tilde{\mathcal{C}}_t$, which is an affine space. The admissible perturbations $\boldsymbol{\psi}$ with respect to $\tilde{\boldsymbol{u}}$, must be such that $\tilde{\boldsymbol{u}} + \boldsymbol{\psi} \in \tilde{\mathcal{C}}_t$, hence $\boldsymbol{\psi} \in \tilde{V}_0$. Equating to zero the partial derivative in the direction $(\boldsymbol{\psi}, \varphi) \in H^1([0,T], \tilde{Z}_0)$, and denoting by $(\boldsymbol{v}, \beta)$ its solution for $(\tilde{\boldsymbol{u}}, \tilde{\alpha}) = (\boldsymbol{u}, \alpha)$ and $\tilde{\lambda} = \lambda$, we arrive at

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{u}}(\boldsymbol{\psi}) = \int_0^T \left(\int_\Omega \left(\partial_{\boldsymbol{u}} m(\boldsymbol{u}, \alpha)\boldsymbol{\psi} + \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}) + \mathbb{C}'(\alpha)\beta\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}) \right) dx \right) dt = 0 \qquad \forall \boldsymbol{\psi} \in H^1([0,T], \tilde{V}_0), \tag{6.42a}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha}(\varphi) = \int_0^T \left(\int_\Omega \left(\partial_\alpha m(\boldsymbol{u}, \alpha)\varphi + \mathbb{C}'(\alpha)\varphi\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) + \frac{G_c \ell}{2c_w} \nabla\beta \cdot \nabla\varphi \right. \right. \\ \left. + \left(\frac{1}{2}\mathbb{C}''(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) + \frac{G_c}{4c_w} \frac{w''(\alpha)}{\ell} + \frac{G_c}{\epsilon} \mathcal{M}'(\alpha - 1) \right) \beta\varphi + \frac{G_c}{\epsilon} \mathcal{M}'(-\dot{\alpha})\dot{\varphi}\beta \right) dx \Big) dt \\ + \int_\Omega \lambda\varphi(0) dx = 0 \qquad \forall \varphi \in H^1([0,T], \tilde{H}^1_D(\mathbb{R}^d)). \end{aligned} \tag{6.42b}$$

Varying the test function $\boldsymbol{\psi}$ in (6.42a), we get the boundary condition $\boldsymbol{v}(t) = \boldsymbol{0}$ on $\Gamma_{\tilde{\boldsymbol{u}}}$ and for all $t \in [0,T)$

$$\int_\Omega \left(\partial_{\boldsymbol{u}} m(\boldsymbol{u}, \alpha)\boldsymbol{\psi} + \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}) + \mathbb{C}'(\alpha)\beta\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}) \right) dx = 0 \qquad \forall \boldsymbol{\psi} \in V_0,$$

where we used definition (6.36) of V_0. We have thus derived the adjoint equation (6.37a). Now, to get rid of the time derivative $\dot{\varphi}$ in equation (6.42b), we integrate the term $\mathcal{M}'(-\dot{\alpha})\dot{\varphi}\beta$ by parts with respect to t and obtain

$$\begin{aligned} \int_0^T \left(\int_\Omega \left(\partial_\alpha m(\boldsymbol{u}, \alpha)\varphi + \mathbb{C}'(\alpha)\varphi\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{v}) + \frac{G_c \ell}{2c_w} \nabla\beta \cdot \nabla\varphi \right. \right. \\ \left. + \left(\frac{1}{2}\mathbb{C}''(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) + \frac{G_c}{4c_w} \frac{w''(\alpha)}{\ell} + \frac{G_c}{\epsilon} \mathcal{M}'(\alpha - 1) \right) \beta\varphi \right) dx - \int_\Omega \frac{G_c}{\epsilon} \left(\frac{d}{dt} (\mathcal{M}'(-\dot{\alpha})\beta) \right) \varphi \, dx \Big) dt \\ + \int_\Omega \lambda \, \varphi|_{t=0} \, dx + \int_\Omega \frac{G_c}{\epsilon} \left(\mathcal{M}'(-\dot{\alpha})\beta\varphi|_{t=T} - \mathcal{M}'(-\dot{\alpha})\beta\varphi|_{t=0} \right) dx = 0 \qquad \forall \varphi \in H^1([0,T], \tilde{H}^1_D(\mathbb{R}^d)). \end{aligned} \tag{6.43}$$

This integration by part is legitimate if $\mathcal{M}$ is a smooth function but is purely formal if $\mathcal{M}$ is the maximum function since in such a case $\mathcal{M}'$ is a Heaviside function and its time derivative involves a Dirac function. Varying φ in (6.43), we find that the Lagrange multiplier λ is given by

$$\lambda = \frac{G_c}{\epsilon} \, \mathcal{M}'(-\dot{\alpha})\beta|_{t=0} \, ,$$

and that the adjoint problem (6.37b) for β holds true. Finally, since $J(\Omega) = \mathcal{L}(\boldsymbol{u}, \alpha, \tilde{\boldsymbol{v}}, \tilde{\beta}, \tilde{\lambda}, \tilde{\boldsymbol{\mu}}, \Omega)$, the shape derivative $J'(\Omega)(\boldsymbol{\theta})$ satisfies, for any $\boldsymbol{\theta} \in W_0^{1,\infty}(D, \mathbb{R}^d)$,

$$J'(\Omega)(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{\theta}) + \frac{\partial \mathcal{L}}{\partial \boldsymbol{u}}\left(\frac{\partial \boldsymbol{u}}{\partial \Omega}(\boldsymbol{\theta})\right) + \frac{\partial \mathcal{L}}{\partial \alpha}\left(\frac{\partial \alpha}{\partial \Omega}(\boldsymbol{\theta})\right)$$

as the variables $\tilde{\boldsymbol{v}}, \tilde{\beta}, \tilde{\lambda}$ and $\tilde{\boldsymbol{\mu}}$ are independent of Ω. Substituting these variables by the optimal ones $\boldsymbol{v}, \beta, \lambda$, $\boldsymbol{\mu}$ and using the adjoint equation (6.57), the two last terms in the above formula vanish, resulting in

$$J'(\Omega)(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{\theta}).$$

Consequently, formula (6.38) is deduced by a straightforward application of Lemma 1. □

6.3.3 Time-discretized state and adjoint equations

The adjoint equation (6.57) is a linear backward parabolic equation with a final condition at $t = T$. This equation was assumed to be well-posed in the statement of Theorem 1. There are two difficulties in proving that (6.57) admits a unique solution. First, the bilinear form, involved in (6.57), is the Hessian of the non-convex energy functional (6.10) and thus is not coercive. Second, if $\mathcal{M}$ is the maximum function, then (6.57) features a time derivative which is a Dirac function (the precise meaning of which is unclear). This second issue can be settled upon time discretization, as can be expected. It turns out that, upon a technical assumption (see (6.50) in Theorem 28), the first issue can also be circumvented by a trick similar to what is used for solving Helmholtz equation (this equation is not coercive but its kernel is at most finite dimensional and often reduced to zero).

To construct a time-discretized version of the adjoint problem (6.57), we first time-discretize the state equation (6.19) along with the objective function (6.34). The time interval $[0, T]$ is split in N sub-intervals of length $\delta t = T/N$. Let $(\boldsymbol{u}_n, \alpha_n)$ and $(\boldsymbol{v}_n, \beta_n)$ denote the discrete state and adjoint solutions, respectively, at the end of every n-th time interval. The discrete state is determined using an implicit scheme: initialize $(\boldsymbol{u}_0, \alpha_0) = (\boldsymbol{0}, \alpha_0)$ and, for $0 < n \leq N$, find $(\boldsymbol{u}_n, \alpha_n) \in Z$ such that $\boldsymbol{u}_n = \bar{\boldsymbol{u}}(t_n)$ on $\Gamma_{\bar{\boldsymbol{u}}}$ and

$$\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v})\, dx = \int_\Omega \boldsymbol{f}(t_n) \cdot \boldsymbol{v}\, dx + \int_{\Gamma_N} \boldsymbol{g}(t_n) \cdot \boldsymbol{v}\, ds \qquad \forall \boldsymbol{v} \in V, \tag{6.44a}$$

$$\begin{aligned}\int_\Omega \frac{G_c}{2c_w}\left(\ell \nabla \alpha_n \cdot \nabla \beta + \frac{w'(\alpha_n)}{2\ell}\right) dx + \int_\Omega \frac{1}{2}\mathbb{C}'(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n)\, dx \\ + \int_\Omega \frac{G_c}{\epsilon}\left(\mathcal{M}(\alpha_n - 1)\beta\, dx - \int_\Omega \mathcal{M}\left(\frac{\alpha_{n-1} - \alpha_n}{\delta t}\right)\beta\right) dx = 0 \qquad \forall \beta \in H^1_D(\Omega).\end{aligned} \tag{6.44b}$$

The above problem can be shown to admit at least one solution (see Subsection 6.2.3), while uniqueness of the solution is not guaranteed (like its continuous counterpart (6.19)). Nevertheless, we assume that the solution $(\boldsymbol{u}_n, \alpha_n) \in Z$ is unique. The discretization of the objective function (6.34) reads:

$$J_N(\Omega) = \sum_{n=0}^{N} \delta t \int_\Omega m(\boldsymbol{u}_n, \alpha_n)\, dx. \tag{6.45}$$

Introducing a Lagrangian, as in the proof of Proposition 1, adapted to the above discretization, we obtain the following discrete adjoint problem: initialize $(\boldsymbol{v}_N, \beta_N) = (\boldsymbol{0}, 0)$ and, for $N - 1 \geq n \geq 0$, find $(\boldsymbol{v}_n, \beta_n) \in Z$ such that

$$\int_\Omega \left(\partial_{\boldsymbol{u}_n} m(\boldsymbol{u}_n, \alpha_n)\boldsymbol{\psi} + \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{v}_n) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}) + \mathbb{C}'(\alpha_n)\beta_n \boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{\psi})\right) dx = 0 \qquad \forall \boldsymbol{\psi} \in V_0, \tag{6.46a}$$

$$\begin{aligned}&\int_\Omega \frac{G_c \ell}{2c_w}\nabla\varphi \cdot \nabla\beta_n\, dx + \int_\Omega \left(\mathbb{C}'(\alpha_n)\varphi\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v}_n) + \frac{1}{2}\mathbb{C}''(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n) + \frac{G_c}{4c_w}\frac{w''(\alpha_n)}{\ell}\right)\varphi\beta_n\, dx \\ &+ \int_\Omega \frac{G_c}{\epsilon}\left(\mathcal{M}'(\alpha_n - 1) + \frac{1}{\delta t}\mathcal{M}'(\alpha_{n-1} - \alpha_n)\right)\varphi\beta_n\, dx \\ &+ \int_\Omega \partial_{\alpha_n} m(\boldsymbol{u}_n, \alpha_n)\varphi\, dx = \frac{G_c}{\epsilon\delta t}\int_\Omega \mathcal{M}'(\alpha_n - \alpha_{n+1})\beta_{n+1}\varphi\, dx \quad \forall \varphi \in H^1_D(\Omega).\end{aligned} \tag{6.46b}$$

Remark 21. *When $\mathcal{M}$ is the maximum function, its derivative is the Heaviside function, $\mathcal{M}' = \mathcal{H}$. The value $\mathcal{H}(0)$ is not precisely defined since the Heaviside function is discontinuous at zero. Numerically, we tested the adjoint equation (6.47) (along with the corresponding shape derivative) for values $\mathcal{H}(0) = 0$ and $\mathcal{H}(0) = 1$, and both choices yield the same optimized shape (at least, for the 2D cantilever in Section 6.5).*

Since the variational formulation (6.46) is linear, it can be written in a compact form: $(\boldsymbol{v}_N, \beta_N) = (\mathbf{0}, 0)$ and, for $N - 1 \geq n \geq 0$, find $(\boldsymbol{v}_n, \beta_n) \in Z$ such that

$$a_n(\boldsymbol{v}_n, \beta_n, \boldsymbol{\psi}, \varphi) = f_n(\boldsymbol{\psi}, \varphi) \qquad \forall (\boldsymbol{\psi}, \varphi) \in Z, \tag{6.47}$$

where the symmetric bilinear form $a_n : Z \times Z \mapsto \mathbb{R}$ is defined as

$$\begin{aligned} a_n(\boldsymbol{v}, \beta, \boldsymbol{\psi}, \varphi) = &\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{\psi}) : \boldsymbol{\varepsilon}(\boldsymbol{v})\ dx + \int_\Omega \mathbb{C}'(\alpha_n)\beta\boldsymbol{\varepsilon}(\boldsymbol{\psi}) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx + \int_\Omega \frac{G_c\ell}{2c_w}\nabla\varphi \cdot \nabla\beta\ dx \\ &+ \int_\Omega \mathbb{C}'(\alpha_n)\varphi\boldsymbol{\varepsilon}(\boldsymbol{v}) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n)\, dx + \int_\Omega \left(\frac{1}{2}\mathbb{C}''(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n) + \frac{G_c}{4c_w}\frac{w''(\alpha_n)}{\ell}\right)\varphi\beta\, dx \\ &+ \frac{G_c}{\epsilon}\int_\Omega \left(\mathcal{M}'(\alpha_n - 1) + \frac{1}{\delta t}\mathcal{M}'(\alpha_{n-1} - \alpha_n)\right)\varphi\beta\, dx, \end{aligned} \tag{6.48}$$

and the linear form $f_n : Z \mapsto \mathbb{R}$ is

$$f_n(\boldsymbol{\psi}, \varphi) = -\int_\Omega \partial_{\boldsymbol{u}} m(\boldsymbol{u}_n, \alpha_n)\boldsymbol{\psi}\, dx - \int_\Omega \partial_\alpha m(\boldsymbol{u}_n, \alpha_n)\varphi\, dx + \frac{G_c}{\epsilon\delta t}\int_\Omega \mathcal{M}'(\alpha_n - \alpha_{n+1})\beta_{n+1}\varphi\, dx. \tag{6.49}$$

To the bilinear form a_n is associated an operator $A_n : Z \mapsto Z$, defined by

$$a_n(\boldsymbol{v}, \beta, \boldsymbol{\psi}, \varphi) = \langle A_n(\boldsymbol{v}, \beta), (\boldsymbol{\psi}, \varphi)\rangle. \tag{6.50}$$

To prove the existence of a solution to the time-discretized adjoint equation (6.47), we have to change the assumption (6.1) on $\mathbb{C}(\alpha)$, which cannot be anymore degenerate when damage is complete.

Lemma 14. *Assume $\mathbb{C}(\alpha)$ is convex, decreasing, $\mathbb{C}(0) = \mathbb{C}_0$ and there exists $\kappa > 0$ (a residual stiffness) such that*

$$\mathbb{C}(\alpha)\boldsymbol{\xi} : \boldsymbol{\xi} \geq \kappa|\boldsymbol{\xi}|^2 \qquad \forall \boldsymbol{\xi} \in \mathcal{M}_s^d,\ 0 \leq \alpha \leq 1. \tag{6.51}$$

Assume that $\boldsymbol{u}_n \in W^{1,\infty}(\Omega)^d$. There exists a real number $c > 0$ such that $(A_n + cI)$ is invertible from Z to Z (where I is the identity operator) and its inverse is a compact linear continuous operator.

Proof. It is enough to check the coercivity of the bilinear form

$$a_n(\boldsymbol{v}, \beta, \boldsymbol{\psi}, \varphi) + c\langle(\boldsymbol{v}, \beta), (\boldsymbol{\psi}, \varphi)\rangle,$$

for some constant c, large enough. Since $\mathbb{C}(\alpha)$ is assumed to be convex and, by its definition (6.17) or (6.16), $w(\alpha)$ is convex too, we have

$$\mathbb{C}''(\alpha) \geq 0, \quad w''(\alpha) \geq 0. \tag{6.52}$$

Compute

$$
\begin{aligned}
&a_n(\boldsymbol{\psi},\varphi,\boldsymbol{\psi},\varphi)\\
&= \int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{\psi}):\boldsymbol{\varepsilon}(\boldsymbol{\psi})\ dx + \int_\Omega 2\mathbb{C}'(\alpha_n)\varphi\boldsymbol{\varepsilon}(\boldsymbol{\psi}):\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx\\
&\quad + \int_\Omega \frac{G_c\ell}{2c_w}\nabla\varphi\cdot\nabla\varphi\ dx + \int_\Omega \left(\frac{1}{2}\mathbb{C}''(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n):\boldsymbol{\varepsilon}(\boldsymbol{u}_n) + \frac{G_c}{4c_w}\frac{w''(\alpha_n)}{\ell}\right)\varphi^2 dx\\
&\quad + \frac{G_c}{\epsilon}\int_\Omega \left(\mathcal{M}'(\alpha_n-1) + \frac{1}{\delta t}\mathcal{M}'(\alpha_{n-1}-\alpha_n)\right)\varphi^2\ dx\\
&\geq \kappa\int_\Omega |\boldsymbol{\varepsilon}(\boldsymbol{\psi})|^2 dx + \int_\Omega 2\mathbb{C}'(\alpha_n)\varphi\boldsymbol{\varepsilon}(\boldsymbol{\psi}):\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx\\
&\qquad + \int_\Omega \frac{G_c\ell}{2c_w}|\nabla\varphi|^2 dx + \int_\Omega \left(\frac{1}{2}\mathbb{C}''(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n):\boldsymbol{\varepsilon}(\boldsymbol{u}_n) + \frac{G_c}{4c_w}\frac{w''(\alpha_n)}{\ell}\right)\varphi^2 dx \qquad (\text{since } \mathcal{M}' = \mathcal{H} \geq 0\ )\\
&\geq \kappa\int_\Omega |\boldsymbol{\varepsilon}(\boldsymbol{\psi})|^2 dx + \int_\Omega 2\mathbb{C}'(\alpha_n)\varphi\boldsymbol{\varepsilon}(\boldsymbol{\psi}):\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx + \int_\Omega \frac{G_c\ell}{2c_w}|\nabla\varphi|^2 dx \qquad (\text{using (6.52)}\ )\\
&\geq \kappa\int_\Omega |\boldsymbol{\varepsilon}(\boldsymbol{\psi})|^2\ dx - 2\left\|\mathbb{C}'(\alpha_n)\right\|_{L^\infty(\Omega)}\int_\Omega |\boldsymbol{\varepsilon}(\boldsymbol{\psi})|\ |\varphi\boldsymbol{\varepsilon}(\boldsymbol{u}_n)| dx + \int_\Omega \frac{G_c\ell}{2c_w}|\nabla\varphi|^2 dx\\
&\geq \kappa\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)} - \left\|\mathbb{C}'(\alpha_n)\right\|_{L^\infty(\Omega)}\left(s\left\|\varphi|\boldsymbol{\varepsilon}(\boldsymbol{u}_n)|\right\|^2_{L^2(\Omega)} + \frac{1}{s}\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)}\right) + \frac{G_c\ell}{2c_w}\left\|\nabla\varphi\right\|^2_{L^2(\Omega)}\\
&\qquad\qquad (\text{using Young's inequality with } s>0)\\
&\geq \kappa\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)} - \left\|\mathbb{C}'(\alpha_n)\right\|_{L^\infty(\Omega)}\left(s\left\|\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\right\|_{L^\infty(\Omega)}\left\|\varphi\right\|^2_{L^2(\Omega)} + \frac{1}{s}\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)}\right) + \frac{G_c\ell}{2c_w}\left\|\nabla\varphi\right\|^2_{L^2(\Omega)}\\
&\qquad\qquad (\text{because of our assumption } \boldsymbol{u}_n \in W^{1,\infty}(\Omega)^d\ )\\
&= \left(\kappa - \frac{1}{s}\left\|\mathbb{C}'(\alpha_n)\right\|_{L^\infty(\Omega)}\right)\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)} + \frac{G_c\ell}{2c_w}\left\|\nabla\varphi\right\|^2_{L^2(\Omega)} - s\left\|\mathbb{C}'(\alpha_n)\right\|_{L^\infty(\Omega)}\left\|\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\right\|_{L^\infty(\Omega)}\left\|\varphi\right\|^2_{L^2(\Omega)}.
\end{aligned}
$$

We choose $s = 2\left\|\mathbb{C}'(\alpha_n)\right\|_{L^\infty(\Omega)}/\kappa$ resulting in the bound

$$
\begin{aligned}
a_n(\boldsymbol{\psi},\varphi,\boldsymbol{\psi},\varphi) &\geq \frac{\kappa}{2}\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)} + \frac{G_c\ell}{2c_w}\left\|\nabla\varphi\right\|^2_{L^2(\Omega)} - \frac{2}{\kappa}\left\|\mathbb{C}'(\alpha_n)\right\|^2_{L^\infty(\Omega)}\left\|\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\right\|_{L^\infty(\Omega)}\left\|\varphi\right\|^2_{L^2(\Omega)}\\
&\geq C_1\left(\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)} + \left\|\nabla\varphi\right\|^2_{L^2(\Omega)}\right) - C_2\left\|\varphi\right\|^2_{L^2(\Omega)}\\
&\geq C_1\left(\left\|\boldsymbol{\varepsilon}(\boldsymbol{\psi})\right\|^2_{L^2(\Omega)} + \left\|\nabla\varphi\right\|^2_{L^2(\Omega)}\right) - C_2\left(\left\|\boldsymbol{\psi}\right\|^2_{L^2(\Omega)^d} + \left\|\varphi\right\|^2_{L^2(\Omega)}\right),
\end{aligned}
$$

where $C_1 = \min\left(\frac{\kappa}{2}, \frac{G_c\ell}{2c_w}\right)$ and $C_2 = \frac{2}{\kappa}\left\|\mathbb{C}'(\alpha_n)\right\|^2_{L^\infty(\Omega)}\left\|\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\right\|_{L^\infty(\Omega)}$. Therefore, choosing $c > C_2$ yields the result. □

Lemma 14 implies that the operator A_n has a discrete countably infinite spectrum, like any elliptic operator, although it is not coercive (like the Helmholtz equation). It allows us to state a result about the well-posedness of the time-discretized adjoint equation (6.47).

Theorem 28. *Under the hypotheses of Lemma 14 and assuming that* 0 *does not belong to the spectrum of* A_n*, the time-discretized adjoint equation* (6.47) *admits a unique solution* $(\boldsymbol{v}_n, \beta_n) \in Z$*,* $N-1 \geq n \geq 1$*.*

Remark 22. *In Theorem 28, we made a strong assumption that zero is not in the spectrum of* A_n *or equivalently that the kernel of* A_n *is reduced to zero. Since the spectrum of* A_n *is discrete, it is unlikely that it generically contains the value zero. Even if zero is an eigenvalue of* A_n*, we can always perturb the coefficients in* A_n *(*$\boldsymbol{u}_n$ *and* α_n*) in such a way that the spectrum is perturbed so that zero is not anymore an eigenvalue.*

Proof. The linear form (6.49) is clearly continuous on Z. Therefore, solving the variational formulation (6.47) amounts to solve the linear equation in Z, $A_n(\boldsymbol{v},\beta) = f_n$. By virtue of Lemma 14, A_n admits a family of eigenvectors, which form a Hilbert basis of Z. By the spectral decomposition of A_n this equation has a unique solution if 0 does not belong to the spectrum of A_n. □

6.3.4 Formal analysis of the limit adjoint equation

In the previous subsection, we determined the adjoint equation (6.57) for the approximate damage problem (6.19). The approximation is brought about using the penalization-regularization parameter ϵ. As ϵ tends to zero, we recover the original damage problem (6.15) which is not differentiable. This was exactly the case with contact mechanics (see chapter 3) and plasticity (see chapter 4). Hence, like we did in the previous chapter, here as well, we wish to determine the adjoint equation corresponding to the limit problem (6.15).

In order to do so, there are two possibilities: one can attempt to

1. Pass to the limit in the adjoint equation (6.57)
2. Find the conical derivative of the problem (6.15)
3. Reformulate the problem (6.15), converting it to an equation.

These are exactly the same three options we had in the previous chapters. Here as well, we shall proceed with the third option in the above as it makes it easier to find the adjoint problem. To simplify the presentation, we shall consider the time-discretized version of the damage problem (6.15): for the initial condition $(\mathbf{0}, \alpha_0)$, find $(\boldsymbol{u}_n, \alpha_n) \in Z$, $1 \leq n \leq N$, such that $\boldsymbol{u}_n = \bar{\boldsymbol{u}}(t_n)$ on $\Gamma_{\bar{\boldsymbol{u}}}$ and

$$\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v})\ dx = \int_\Omega \boldsymbol{f}_n \cdot \boldsymbol{v}\ dx + \int_{\Gamma_N} \boldsymbol{g}_n \cdot \boldsymbol{v}\ ds \qquad \forall \boldsymbol{v} \in V \tag{6.53a}$$

$$\text{and } \int_\Omega \mathbb{C}'(\alpha_n)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx + \int_\Omega \frac{G_c}{2c_w}\left(\ell \nabla\alpha_n \cdot \nabla\beta + \frac{w'(\alpha_n)\beta}{2\ell}\right) dx \geq 0 \qquad \forall \beta \in \mathcal{D}. \tag{6.53b}$$

For the shape optimization problem, we then minimize the functional

$$J(\Omega) = \sum_{n=0}^{N} \delta t \left(\int_\Omega m(\boldsymbol{u}_n(\Omega), \alpha_n(\Omega)) dx + \int_{\Gamma_N} p(\boldsymbol{u}_n(\Omega), \alpha_n(\Omega)) ds \right). \tag{6.54}$$

To convert (6.53) into an equation, we introduce the set where the damage variable α_n is positive or equal to unity:

$$\Omega^1_{\alpha_n} = \{\boldsymbol{x} \in \Omega |\, \alpha_n(\boldsymbol{x}) = \alpha_{n-1}(\boldsymbol{x})\}$$
$$\Omega^2_{\alpha_n} = \{\boldsymbol{x} \in \Omega |\, \alpha_n(\boldsymbol{x}) = 1\}.$$

Using these sets, we introduce a space K_{α_n} defined as

$$K_{\alpha_n} = \{\beta \in H^1(\Omega) |\, \beta(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Omega^1_{\alpha_n} \cup \Omega^2_{\alpha_n}\}.$$

Given the complementary condition (6.14c), we know that the damage criterion (6.14b) is exactly satisfied only if the damage evolution rate is positive (hence $\alpha_n(\boldsymbol{x}) > \alpha_{n-1}(\boldsymbol{x})$). In the domain Ω, there can be exactly two regions. The first where $\alpha_n(\boldsymbol{x}) = \alpha_{n-1}(\boldsymbol{x})$ and the second where $\alpha_n(\boldsymbol{x}) > \alpha_{n-1}(\boldsymbol{x})$. Consider a test function that is zero on the first domain and in addition, zero in the domain where $\alpha_n(\boldsymbol{x}) = 1$. Such a test function belongs to K_{α_n} and is non-zero where the damage criterion (6.14b) is an equality. We choose one such test function, multiply it with the damage-criterion (6.14b), integrate over Ω and integrate by parts to obtain the equation (6.55b). This results in the following set of equations: for the initial condition $(\mathbf{0}, \alpha_0)$, find $(\boldsymbol{u}_n, \alpha_n) \in Z$, $1 \leq n \leq N$, such that $\boldsymbol{u}_n = \bar{\boldsymbol{u}}(t_n)$ on $\Gamma_{\bar{\boldsymbol{u}}}$ and

$$\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v})\ dx = \int_\Omega \boldsymbol{f}_n \cdot \boldsymbol{v}\ dx + \int_{\Gamma_N} \boldsymbol{g}_n \cdot \boldsymbol{v}\ ds \qquad \forall \boldsymbol{v} \in V \tag{6.55a}$$

$$\int_\Omega \mathbb{C}'(\alpha_n)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx + \int_\Omega \frac{G_c}{2c_w}\left(\ell \nabla\alpha_n \cdot \nabla\beta\ dx + \frac{w'(\alpha_n)\beta}{2\ell}\right) dx = 0 \qquad \forall \beta \in K_{\alpha_n}, \tag{6.55b}$$

The solution to the problem (6.53) can be shown to satisfy the equations in (6.55). However showing that the solution to (6.55) satisfies the problem (6.53) is not that straight-forward, given the non convexity of the problem. To draw an analogy with the inequality of plasticity, we could show the equivalence between the inequality (4.69) and the equations (4.77a)-(4.77b) (obtained using the same methodology) in Theorem 22, because of the convex nature of the plasticity problem. We nevertheless treat the problem (6.55) for determining the shape derivative. Like we did in the Subsection 4.3.2, we make a strong assumption

B.1 When Ω is perturbed to $(I_d+\boldsymbol{\theta})\Omega$, with a small vector field $\boldsymbol{\theta}\in W_0^{1,\infty}(\mathbb{R}^d,\mathbb{R}^d)$, for all $1\le n\le N$, the corresponding solution $(\boldsymbol{u}_{n,\boldsymbol{\theta}},\alpha_{n,\boldsymbol{\theta}})\equiv(\boldsymbol{u}_n((I_d+\boldsymbol{\theta})\Omega),\alpha_n((I_d+\boldsymbol{\theta})\Omega)$ of problem (6.55) is differentiable with respect to $\boldsymbol{\theta}$ and the damaged zones $\Omega^1_{\alpha_n},\Omega^2_{\alpha_n}$ are perturbed to $(I_d+\boldsymbol{\theta}^1_p)\Omega^1_{\alpha_n},(I_d+\boldsymbol{\theta}^2_p)\Omega^2_{\alpha_n}$, respectively where $\boldsymbol{\theta}^1_p,\boldsymbol{\theta}^2_p$ are vector fields which smoothly depend on $\boldsymbol{\theta}$ and $(\boldsymbol{u}_n,\alpha_n)$, while $\Omega^1_{\alpha_n},\Omega^2_{\alpha_n}$ are open Lipschitz sets.

The above assumption implies that the damaged zone does not change its topology when the shape Ω evolves. This is clearly not true in the context of damage mechanics as we know that given the non smooth nature of problem, slight changes in the topology can lead to appearance or disappearance of cracks. We give limit adjoint problem and the corresponding shape derivative for minimizing the functional (6.54) in the following theorem.

Theorem 29. *Let $\boldsymbol{f}_n=\boldsymbol{f}(t_n)\in H^1(\mathbb{R}^d)^d$, $\boldsymbol{g}_n=\boldsymbol{g}(t_n)\in H^2(\mathbb{R}^d)^d$ and $(\boldsymbol{u}_n,\alpha_n)\in V\times\mathcal{D}$ be the solution to (6.55). Assume that the zone in Ω where damage flow rate is non zero, $\Omega\backslash\Omega^1_n$ grows with respect to n i.e., $\Omega^1_{\alpha_0}\supset\Omega^1_{\alpha_1}\supset\cdots\supset\Omega^1_{\alpha_N}$.*

*Then under the assumption **B.1**, the shape derivative of $J(\Omega)$ in (6.54) along $\boldsymbol{\theta}\in W_0^{1,\infty}(\mathbb{R}^d,\mathbb{R}^d)$, is given by*

$$J'(\Omega)(\boldsymbol{\theta})=\int_\Gamma\boldsymbol{\theta}\cdot\boldsymbol{n}\sum_{i=1}^N\Big(\mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{v}_n)-\boldsymbol{f}_n\cdot\boldsymbol{v}_n+\mathbb{C}'(\alpha_n)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\\+\frac{G_c}{2c_w}\Big(\ell\nabla\alpha_n\cdot\nabla\beta_n\,dx+\frac{w'(\alpha_n)\beta_n}{2\ell}\Big)\Big)\,ds,\tag{6.56}$$

where $(\boldsymbol{v}_n,\beta_n)\in V\times K_{\alpha_n}$ is the adjoint variable satisfying the equation: $(\boldsymbol{v}_N,\beta_N)=(\boldsymbol{0},0),N>n\ge1$ and

$$\int_\Omega\mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{\varphi})\colon\boldsymbol{\varepsilon}(\boldsymbol{v})\,dx+\int_\Omega\mathbb{C}'(\alpha_n)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{\varphi})\,dx\\+\int_\Omega\partial_{\boldsymbol{u}_n}m(\boldsymbol{u}_n,\alpha_n)\cdot\boldsymbol{\varphi}\,dx+\int_{\Gamma_N}\partial_{\boldsymbol{u}_n}p(\boldsymbol{u}_n,\alpha_n)\cdot\boldsymbol{\varphi}\,ăds=0\qquad\forall\boldsymbol{\varphi}\in V,\tag{6.57a}$$

$$\int_\Omega\frac{1}{2}\mathbb{C}''(\alpha_n)\psi\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\,dx+\int_\Omega\frac{G_c}{2c_w}\Big(\ell\nabla\psi\cdot\nabla\beta+\frac{w''(\alpha_n)\psi\beta}{2\ell}\Big)\,dx\\+\int_\Omega\partial_{\alpha_n}m(\boldsymbol{u}_n,\alpha_n)\psi\,dx+\int_{\Gamma_N}\partial_{\alpha_n}p(\boldsymbol{u}_n,\alpha_n)\psi\,ăds=0\qquad\forall\psi\in K_{\alpha_n}.\tag{6.57b}$$

Proof. Let

- $\tilde{\boldsymbol{U}}=\{(\tilde{\boldsymbol{u}}_0,\tilde\alpha_0),(\tilde{\boldsymbol{u}}_1,\tilde\alpha_1),\dots,(\tilde{\boldsymbol{u}}_N,\tilde\alpha_N)\}$, and each element $(\tilde{\boldsymbol{u}}_i,\tilde\alpha_i)\in\tilde Z$ (defined in (6.39))
- $\tilde{\boldsymbol{V}}=\{(\tilde{\boldsymbol{v}}_0,\tilde\beta_0),(\tilde{\boldsymbol{v}}_1,\tilde\beta_1),\dots,(\tilde{\boldsymbol{v}}_N,\tilde\beta_N)\}$, (the adjoint multiplier corresponding to the state equation and each element $(\tilde{\boldsymbol{v}}_i,\tilde\beta_i)\in\tilde Z$)
- $\tilde{\boldsymbol{\lambda}}^1=\{\tilde\lambda^1_0,\tilde\lambda^1_1,\dots,\tilde\lambda^1_N\}$ (the adjoint multiplier to impose the irreversibility constraint, $\lambda^1_n\in L^2(\mathbb{R}^d)$)
- $\tilde{\boldsymbol{\lambda}}^2=\{\tilde\lambda^2_0,\tilde\lambda^2_1,\dots,\tilde\lambda^2_N\}$ (the adjoint multiplier to ensure that the damage variable is less than equal to unity, $\lambda^2_n\in L^2(\mathbb{R}^d)$)
- $\tilde{\boldsymbol{\Omega}}^1=\{\tilde\Omega^1_1,\tilde\Omega^1_2,\dots,\tilde\Omega^1_N\}$ (the domains $\tilde\Omega^1_n\subset\Omega$ for every time step n where the damage evolution rate shall be zero)
- $\tilde{\boldsymbol{\Omega}}^2=\{\tilde\Omega^2_1,\tilde\Omega^2_2,\dots,\tilde\Omega^2_N\}$ (the domains $\tilde\Omega^2_n\subset\Omega$ for every time step n where the damage variable is equal to unity)

A priori, we do not know if the adjoint problem corresponding to the damage problem (6.15) is well-posed or not. We nevertheless assume it to be well-posed and apply Céa's technique [53] by constructing

$$\begin{aligned}\mathcal{L}(\tilde{\boldsymbol{U}},\tilde{\boldsymbol{V}},\tilde{\boldsymbol{\lambda}}^1,\tilde{\boldsymbol{\lambda}}^2,\tilde{\boldsymbol{\Omega}}^1,\tilde{\boldsymbol{\Omega}}^2,\Omega) = &\sum_{i=1}^{N}\delta t\left(\int_\Omega m(\tilde{\boldsymbol{u}}_n,\tilde{\alpha}_n)dx+\int_{\Gamma_N}p(\tilde{\boldsymbol{u}}_n,\tilde{\alpha}_n)ds\right)\\ &+\sum_{i=1}^{N}\delta t\left(\int_\Omega \mathbb{C}(\tilde{\alpha}_n)\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_n)\colon\boldsymbol{\varepsilon}(\tilde{\boldsymbol{v}}_n)\ dx-\int_\Omega \boldsymbol{f}_n\cdot\tilde{\boldsymbol{v}}_n\ dx-\int_{\Gamma_N}\boldsymbol{g}_n\cdot\tilde{\boldsymbol{v}}_n\ ds\right)\\ &+\sum_{i=1}^{N}\delta t\left(\int_\Omega \mathbb{C}'(\tilde{\alpha}_n)\tilde{\beta}_n\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_n)\colon\boldsymbol{\varepsilon}(\tilde{\boldsymbol{u}}_n)\ dx+\int_\Omega\frac{G_c}{2c_w}\left(\ell\nabla\tilde{\alpha}_n\cdot\tilde{\nabla}\tilde{\beta}_n+\frac{w'(\tilde{\alpha}_n)}{2\ell}\tilde{\beta}_n\right)dx\right.\\ &\left.+\int_{\tilde{\Omega}^1_n}\tilde{\lambda}^1_n(\tilde{\alpha}_n-\tilde{\alpha}_{n-1})dx+\int_{\tilde{\Omega}^2_n}\tilde{\lambda}^2_n(\tilde{\alpha}_n-1)dx\right).\end{aligned}\tag{6.58}$$

As usual, applying the optimality conditions to the above leads to the determination of state equation, adjoint equation and the shape derivative. Differentiating the Lagrangian (6.58) with respect to the adjoint multiplier $\tilde{\boldsymbol{V}},\tilde{\boldsymbol{\beta}}$ at $\tilde{\boldsymbol{U}}=\boldsymbol{U}$, we get $1\le n\le N$

$$\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{v})\ dx=\int_\Omega\boldsymbol{f}_n\cdot\boldsymbol{v}\ dx+\int_{\Gamma_N}\boldsymbol{g}_n\cdot\boldsymbol{v}\ ds\qquad\forall\boldsymbol{v}\in\tilde{V}\tag{6.59a}$$

$$\text{and}\ \int_\Omega \mathbb{C}'(\alpha_n)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx+\int_\Omega\frac{G_c}{2c_w}\left(\ell\nabla\alpha_n\cdot\nabla\beta+\frac{w'(\alpha_n)\beta}{2\ell}\right)dx\ge 0\qquad\forall\beta\in\tilde{H}^1_0.\tag{6.59b}$$

Then, differentiating $\mathcal{L}$ (6.58) with respect to $\tilde{\boldsymbol{\lambda}}^1,\tilde{\boldsymbol{\lambda}}^2$ and equating it to zero at $\tilde{\boldsymbol{U}}=\boldsymbol{U}$, we find that

$$1\le n\le N,\quad \alpha_n=\alpha_{n-1}\quad\text{on}\quad\tilde{\Omega}^1_n\quad\text{and}\quad\alpha_n=1\quad\text{on}\quad\tilde{\Omega}^2_n.$$

In order to satisfy the above and the system (6.59), we can simply substitute $\tilde{\Omega}^1_n=\Omega^1_{\alpha_n}$ and $\tilde{\Omega}^2_n=\Omega^2_{\alpha_n}$. We can then replace $\tilde{H}^1_0$ by K_{α_n} in (6.59b) and recover the state equation (6.55). In order to determine the adjoint equation, we differentiate (6.58) with respect to $(\tilde{\boldsymbol{u}}_n,\tilde{\alpha}_n)$, substitute $(\tilde{\boldsymbol{u}}_n,\tilde{\alpha}_n)=(\boldsymbol{u}_n,\alpha_n),(\tilde{\boldsymbol{v}}_n,\tilde{\beta}_n)=(\boldsymbol{v}_n,\beta_n),(\tilde{\Omega}^1_n,\tilde{\Omega}^2_n)=(\Omega^1_{\alpha_n},\Omega^2_{\alpha_n}),\tilde{\Omega}^1_{n+1}=\Omega_{\alpha_{n+1}}$ and obtain

$$\begin{aligned}&\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{\varphi})\colon\boldsymbol{\varepsilon}(\boldsymbol{v})\ dx+\int_\Omega 2\mathbb{C}'(\alpha_n)\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{\varphi})\ dx\\ &+\int_\Omega\partial_{\boldsymbol{u}_n}m(\boldsymbol{u}_n,\alpha_n)\boldsymbol{\varphi}\,dx+\int_{\Gamma_N}\partial_{\boldsymbol{u}_n}p(\boldsymbol{u}_n,\alpha_n)\boldsymbol{\varphi}\,ă ds\qquad\forall\boldsymbol{\varphi}\in\tilde{V},\end{aligned}$$

$$\begin{aligned}&\int_\Omega \mathbb{C}''(\alpha_n)\psi\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx+\int_\Omega\frac{G_c}{2c_w}\ell\nabla\psi\cdot\nabla\beta\ dx+\int_\Omega\frac{G_c}{4c_w}\frac{w''(\alpha_n)\psi}{\ell}\beta\ dx+\int_{\Omega^1_{\alpha_n}}\psi\lambda^1_n-\int_{\Omega^1_{\alpha_{n+1}}}\psi\lambda^1_{n+1}dx\\ &+\int_{\Omega^2_{\alpha_n}}\lambda^2_n\psi\,dx+\int_\Omega\partial_{\alpha_n}m(\boldsymbol{u}_n,\alpha_n)\psi\,dx+\int_{\Gamma_N}\partial_{\alpha_n}p(\boldsymbol{u}_n,\alpha_n)\psi\,ă ds=0\qquad\forall\psi\in\tilde{H}^1_0.\end{aligned}$$

Using the assumption that $\Omega^1_{\alpha_{n+1}}\subset\Omega^1_{\alpha_n}$, and choosing a ψ such that $\psi=0$ on $\Gamma^1_{\alpha_n}\cup\Gamma^2_{\alpha_1}$ we obtain the weak form of the adjoint equation (6.57b): find $\beta\in K_{\alpha_n}$ s.t.

$$\begin{aligned}&\int_\Omega \mathbb{C}''(\alpha_n)\psi\beta\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\colon\boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx+\int_\Omega\frac{G_c\ell}{2c_w}\nabla\psi\cdot\nabla\beta\ dx+\int_\Omega\frac{G_cw''(\alpha_n)\psi\beta}{4c_w\ell}\ dx\\ &+\int_\Omega\partial_{\alpha_n}m(\boldsymbol{u}_n,\alpha_n)\psi\,dx+\int_{\Gamma_N}\partial_{\alpha_n}p(\boldsymbol{u}_n,\alpha_n)\psi\,ă ds=0\qquad\forall\psi\in K_{\alpha_n}.\end{aligned}$$

Now we derive the shape derivative. Let $\boldsymbol{\Omega}^1_\alpha=\{\tilde{\Omega}^1_{\alpha_1},\tilde{\Omega}^1_{\alpha_2},\dots,\tilde{\Omega}^1_{\alpha_N}\}$, $\boldsymbol{\Omega}^2_\alpha=\{\tilde{\Omega}^2_{\alpha_1},\tilde{\Omega}^2_{\alpha_2},\dots,\tilde{\Omega}^2_{\alpha_N}\}$. Then each element of $\boldsymbol{U},\boldsymbol{\Omega}^1_\alpha,\boldsymbol{\Omega}^2_\alpha$ satisfies the equation (6.55), and the Lagrangian (6.58) simplifies to

$$\mathcal{L}(\boldsymbol{U},\tilde{\boldsymbol{V}},\tilde{\boldsymbol{\lambda}}^1,\tilde{\boldsymbol{\lambda}}^2,\boldsymbol{\Omega}^1_\alpha,\boldsymbol{\Omega}^2_\alpha,\Omega)=J(\Omega).$$

We differentiate both sides in the above with respect to the shape Ω in the direction of a vector field $\boldsymbol{\theta}$. Using assumption **B.1**, when $\boldsymbol{\theta}$ moves the domain Ω to $(I_d+\boldsymbol{\theta})\Omega$, the damage zones $\Omega^1_{\alpha_n}, \Omega^2_{\alpha_n}$ are perturbed to $(I_d+\boldsymbol{\theta}^1_p)\Omega^1_{\alpha_n}, (I_d+\boldsymbol{\theta}^2_p)\Omega^2_{\alpha_n}$, ($\boldsymbol{\theta}^1_p, \boldsymbol{\theta}^2_p$ are another vector fields). Applying the chain rule lemma, we thus obtain

$$\begin{aligned} J'(\Omega)(\boldsymbol{\theta}) = \left\langle \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{U}, \tilde{\boldsymbol{V}}, \tilde{\boldsymbol{\lambda}}^1, \tilde{\boldsymbol{\lambda}}^2, \boldsymbol{\Omega}^1_\alpha, \boldsymbol{\Omega}^2_\alpha, \Omega), \boldsymbol{\theta} \right\rangle + \sum_{n=1}^{N} \Bigg(\left\langle \frac{\partial \mathcal{L}}{\partial \Omega^1_{\alpha_n}}(\boldsymbol{U}, \tilde{\boldsymbol{V}}, \tilde{\boldsymbol{\lambda}}^1, \tilde{\boldsymbol{\lambda}}^2, \boldsymbol{\Omega}^1_\alpha, \boldsymbol{\Omega}^2_\alpha, \Omega), \boldsymbol{\theta}^1_p \right\rangle \\ + \left\langle \frac{\partial \mathcal{L}}{\partial \Omega^1_{\alpha_n}}(\boldsymbol{U}, \tilde{\boldsymbol{V}}, \tilde{\boldsymbol{\lambda}}^2, \tilde{\boldsymbol{\lambda}}^2, \boldsymbol{\Omega}^1_\alpha, \boldsymbol{\Omega}^2_\alpha, \Omega), \boldsymbol{\theta}^2_p \right\rangle + \left\langle \frac{\partial \mathcal{L}}{\partial \boldsymbol{u}_n}(\boldsymbol{U}, \tilde{\boldsymbol{V}}, \tilde{\boldsymbol{\lambda}}^2, \tilde{\boldsymbol{\lambda}}^2, \boldsymbol{\Omega}^1_\alpha, \boldsymbol{\Omega}^2_\alpha, \Omega), \frac{\partial \boldsymbol{u}_n}{\partial \Omega}(\boldsymbol{\theta}) \right\rangle \\ + \left\langle \frac{\partial \mathcal{L}}{\partial \alpha_n}(\boldsymbol{U}, \tilde{\boldsymbol{V}}, \tilde{\boldsymbol{\lambda}}^2, \tilde{\boldsymbol{\lambda}}^2, \boldsymbol{\Omega}^1_\alpha, \boldsymbol{\Omega}^2_\alpha, \Omega), \frac{\partial \alpha_n}{\partial \Omega}(\boldsymbol{\theta}) \right\rangle \Bigg). \end{aligned} \tag{6.61}$$

Now, substituting $\tilde{\boldsymbol{v}}_n = \boldsymbol{v}_n, \tilde{\beta}_n = \beta, \tilde{\lambda}^1_n = \lambda^1_n, \tilde{\lambda}^2_n = \lambda^2_n$ (for each n) and using the adjoint equation (6.57), the last two terms of (6.61) vanish. We obtain

$$J'(\Omega)(\boldsymbol{\theta}) = \left\langle \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\Omega}^1_\alpha, \boldsymbol{\Omega}^2_\alpha, \Omega), \boldsymbol{\theta} \right\rangle + \sum_{n=1}^{N} \left(\int_{\partial \Omega^1_n} \lambda^1_n (\alpha_n - \alpha_{n-1}) ds + \int_{\partial \Omega^2_n} \lambda^2_n (\alpha_n - 1) ds \right).$$

Given that $\boldsymbol{\varepsilon}_{p,n} = \boldsymbol{\varepsilon}_{p,n-1}$ in $\Omega \backslash \Omega_{\boldsymbol{w}}$ and the assumption that $\boldsymbol{\varepsilon}_{p,n} \in H^1(\Omega)^{d\times d}$, we deduce that the last integral vanishes. Since $\alpha_n \in H^1(\Omega)$ and $\alpha_n = \alpha_{n-1}$ in Ω^1_n, the trace of the penultimate integral in the above vanishes. With the same reasoning, the last integral vanishes too, resulting in

$$J'(\Omega)(\boldsymbol{\theta}) = \left\langle \frac{\partial \mathcal{L}}{\partial \Omega}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\Omega}^1_\alpha, \boldsymbol{\Omega}^2_\alpha, \Omega), \boldsymbol{\theta} \right\rangle.$$

Finally, using Lemma 1 leads to formula (6.56). □

6.4 Numerical Implementation

In this section, we expound the numerical aspects of the resolution of the state equation (6.19) and adjoint equation (6.57) and the proposed shape optimization algorithm. For all the numerical test cases presented in Section 6.5, the material is chosen to be concrete having the following properties [26]: Young modulus $E = 29$GPa, Poisson ratio $\nu = 0.3$, ultimate tensile strength $\sigma_M = 4.5$MPa and fracture toughness $G_c = 70$MPa. For the dissipation function `DL` in (6.17), the characteristic length ℓ is calculated using the formula [176]

$$\ell = \frac{3 G_c E}{8 \sigma_M^2}, \tag{6.62}$$

and for the dissipation function `DQ` in (6.16), the same is calculated using the formula

$$\ell = \frac{G_c E}{2 \sigma_M^2}, \tag{6.63}$$

The domain Ω is discretized by a simplicial unstructured mesh Ω^h. The mesh is produced by the `MMG` software [61] which features two important input parameters: the minimal and maximal mesh size, denoted by $h_{\min}$ and $h_{\max}$, respectively. The mesh Ω^h is assumed to be uniform in the sense that $h_{\max}$ and $h_{\min}$ are of the same order of magnitude. Following the numerical experiments in [132], for all our numerical test cases the mesh is chosen such that

$$2h_{\min} < \ell. \tag{6.64}$$

Although this choice was proposed in [132] for quadrilateral mesh elements, we follow it for our simplicial meshes. Nevertheless, we have to make one exception with the rule (6.64) in the test case of Subsection 6.5.7 (a realistic column of height $4m$) where we just enforce $h_{\min} < \ell$ in order to have a not too fine mesh which can be treated without resorting to high performance computing. The penalization parameter is chosen to be small, $\epsilon = \mathcal{O}(h^2_{\max})$ (its precise value is given in the beginning of Subsection 6.5.1).

Remark 23. *Despite the fact that the damage model (6.19) is non-local, the crack initiation is mesh-dependent. For instance, mesh-refinement at corners of the shape Ω makes the crack initiation easier at these corners. For this reason we rely on uniform meshes ($h_{\max}$ and $h_{\min}$ of the same order), so that the crack initiation is unbiased.*

The spaces $V, C_v, H^1_D(\Omega), Z$, defined by (6.2), (6.3), (6.4) are discretized by piecewise affine continuous (linear) finite elements and their discrete counterparts are denoted by $V^h, C^h_v, H^1_D(\Omega^h), Z^h$, respectively.

The time interval $[0, T]$ is discretized in N intervals of length $\delta t = T/N$. The time at the end of the n-th time interval is denoted $t_n,\ n = 1, 2, \cdots, N$.

6.4.1 Solving the fracture model

We still denote by $(\boldsymbol{u}_n, \alpha_n) \in Z^h$ the time-space discretized solution at time interval t_n (we do not write its dependence to h). The space discretized version of (6.44) is simply the same variational formulation with the finite dimensional spaces $V^h, H^1_D(\Omega^h), Z^h$ replacing their continuous counterparts $V, H^1_D(\Omega), Z$. Thus, the following: find $(\boldsymbol{u}_n, \alpha_n) \in Z^h$, $0 < n \leq N$ such that $(\boldsymbol{u}_0, \alpha_0) = (\boldsymbol{0}, \alpha_0)$, $\boldsymbol{u}_n = \bar{\boldsymbol{u}}(t_n)$

$$\int_\Omega \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v})\ dx = 0\ ds \qquad \forall \boldsymbol{v} \in V^h \tag{6.65a}$$

$$\text{and} \quad \int_\Omega \frac{G_c}{2c_w}\left(\ell \nabla\alpha_n \cdot \nabla\beta + \frac{w'(\alpha_n)}{2\ell}\right) dx + \int_\Omega \frac{1}{2}\mathbb{C}'(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n)\ dx$$
$$+ \int_\Omega \frac{G_c}{\epsilon}\left(M_\epsilon(\alpha_n - 1)\beta\ dx - \int_\Omega M_\epsilon\left(\frac{\alpha_{n-1} - \alpha_n}{\delta t}\right)\beta\right) dx = 0 \qquad \forall \beta \in \mathbb{P}^1(\Omega^h). \tag{6.65b}$$

For the resolution of the coupled problem (6.65), a simple Newton-Raphson might not converge to the solution desired $(\boldsymbol{u}_n, \alpha_n)$. Instead, one can follow the approach in [35], where the nonlinear variational formulation (6.65) is solved by a sequentially alternate algorithm: fixing α_n, solve (6.65a) for $\boldsymbol{u}_n$; fixing $\boldsymbol{u}_n$, solve (6.65b) for α_n. This algorithm exploits the fact that the total energy (6.11) is separately convex in $\boldsymbol{u}$ and α, but not with respect to the couple $(\boldsymbol{u}, \alpha)$, which may hinder the convergence of a standard Newton algorithm. This algorithm of [35] is precisely recalled in Algorithm 6.

Algorithm 6 Numerical resolution of (6.65)

Initialization: $(\boldsymbol{u}^0_n, \alpha^0_n) = (\boldsymbol{0}, 0)$ for $n = 0$ and $(\boldsymbol{u}^0_n, \alpha^0_n) = (\boldsymbol{u}^0_{n-1}, \alpha^0_{n-1})$ for $n > 0$.
For $i = 0, 1, \cdots,$ do

1. Substitute $\alpha_n = \alpha^i_n$ in (6.65a), and solve it for $\boldsymbol{u}_n = \boldsymbol{u}^{i+1}_n$
2. Substitute $\boldsymbol{u}_n = \boldsymbol{u}^{i+1}_n$ in (6.65b), and solve it for $\alpha_n = \alpha^{i+1}_n$ using a Newton-Raphson or a fixed point algorithm
3. If $\left\|\alpha^{i+1}_n - \alpha^i_n\right\|_{L^2(\Omega^h)} < tol\, \|1\|_{L^2(\Omega^h)}$, then exit the loop,
   else repeat

The tolerance tol in Algorithm 6 is chosen to be 10^{-6}. The resolution for α_n (step 2 in the algorithm 6) using Newton algorithm is easy given that there are no constraints on α_n, because of the penalization. Without penalization, one needs to apply a constrained Newton algorithm, the numerical implementation of which is not straight-forward. Algorithm 6 builds a sequence $(\boldsymbol{u}^i_n, \alpha^i_n) \longrightarrow (\boldsymbol{u}_n, \alpha_n)$ as $i \longrightarrow \infty$. Typically, its convergence is very fast when α_n is close to zero everywhere in Ω^h. Whereas the convergence is very slow when α_n approaches unity and there is crack formation in Ω^h.

Of course, the solution $(\boldsymbol{u}_n, \alpha_n)$ depends on the mesh size $h_{\max}$ and on the time step δt. Furthermore, each solution at time t_n depends on the whole time history before t_n. In practice, the solution depends on the initialization $(\boldsymbol{u}^0_n, \alpha^0_n)$ which is usually taken as the solution at the previous time step $(\boldsymbol{u}_{n-1}, \alpha_{n-1})$. Therefore, it is not clear that Algorithm 6 delivers an approximation of the global minimizer of the total energy (6.11) [44]. Rather, we may end up in local minima. Following again [35], we rather use the backtracking Algorithm 7 which escapes from local minima in practice. The idea of Algorithm 7 is that, if the solution found at time-step n using Algorithm 6 features a strong increase of damage, the solution at the previous time step $n-1$ is re-computed using the solution at time step n as initial guess. In doing so, we expect to find a new solution $(\boldsymbol{u}_{n-1}, \alpha_{n-1})$ which is a better minimizer of the total energy (6.11).

Now briefly discuss the influence of different κ values on the solution $(\boldsymbol{u}_n, \alpha_n)$. For that, we consider a wedge, shown in Fig.6.17. Let an imposed displacement $\boldsymbol{u}_i = (0, 2.33) \times 10^{-4} m, \quad t \in [0, 1]$ be applied, for

Algorithm 7 Numerical resolution of (6.44) with backtracking

Set $n = 1$ and $(\tilde{\boldsymbol{u}}, \tilde{\alpha}) = (\mathbf{0}, 0)$
While $n \leq N$ (time steps), do

1. Solve (6.44) for $(\boldsymbol{u}_n, \alpha_n)$ using algorithm 6 using an initial guess $(\boldsymbol{u}_n^0, \alpha_n^0) = (\tilde{\boldsymbol{u}}, \tilde{\alpha})$
2. If $\|\alpha_n - \tilde{\alpha}\|_{L^\infty(\Omega^h)} > 0.5$ and $n > 1$

   then $(\tilde{\boldsymbol{u}}, \tilde{\alpha}) = (\boldsymbol{u}_n, \alpha_n)$, and set $n = n - 1$,

   else $(\tilde{\boldsymbol{u}}, \tilde{\alpha}) = (\boldsymbol{u}_n, \alpha_n)$, and set $n = n + 1$.

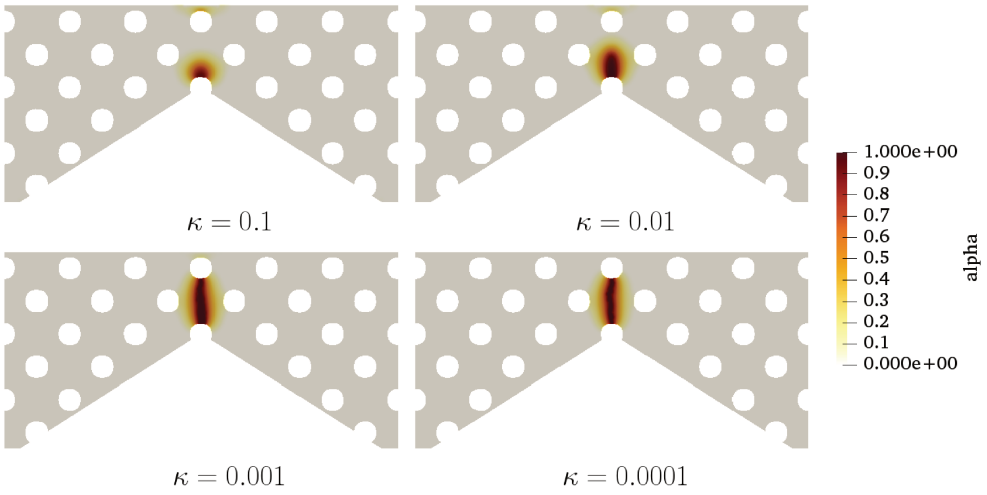


Figure 6.3: The damage variable α at final time corresponding to different κ in (6.66)

`DL` model (6.17), along with the degradation function

$$\mathbb{C}(\alpha) = (1 - \alpha)^2 \mathbb{C}_0 + \mathbb{C}_{res}, \tag{6.66}$$

where the residual stiffness is $\mathbb{C}_{res} = \kappa \mathbb{C}_0$. It corresponds to the lowest value of $|\mathbb{C}(\alpha)|$, i.e., value of the tensor $|\mathbb{C}(1)|$. This is done so that the stiffness matrix encountered while solving (6.65a) (in algorithm 6), does not have zero eigenvalues. We solve the penalized problem (6.65) on an unstructured mesh. For this case, it rarely happens that the value of α is exactly one. It instead takes values in the neighborhood of unity and the stiffness matrix does not have zero eigenvalues. Moreover, we use the iterative solver of `FreeFEM` (based on conjugate gradient) that is fairly robust. Hence, we $\kappa = 0$ for our numerical experiments in Section 6.5 and this does not pose any problem.

We investigate the solution when κ goes from 0.1 to 0.0001. The damage profile at the final time $t = 1$ is plotted in Fig.6.3. As seen in this figure, the damage profile at $\kappa = 0.1$ is very different from the one at $\kappa = 0.0001$. The solution for $\kappa = 0.001$ and $\kappa = 0.0001$ seem to be almost the same, implying convergence. At $\kappa = 0.1$, there is no crack unlike at $\kappa = 0.0001$. This is because when the residual stiffness is high, it takes more energy to create a crack. We shall use this fact to implement an "incremental approach" in Subsection 6.5.3, where κ shall be decreased with respect to shape optimization iterations (according to (6.77)), so that the intermediate shapes, that are not optimal do not undergo a crack.

6.4.2 Solving the adjoint problem

We still denote by $(\boldsymbol{v}_n, \beta_n) \in Z^h$ the time-space discretized adjoint solution at time interval t_n. As for the state solution in the previous subsection, the space discretized version of (6.46) is simply the same variational formulation with the finite dimensional spaces $V^h, H_D^1(\Omega^h), Z^h$ replacing their continuous counterparts

$V, H^1_D(\Omega), Z$. Hence the adjoint problem reads: for $\beta_N = 0$, find $(\boldsymbol{v}_n, \beta_n) \in Z^h$, $N-1 \geq n \geq 0$, such that

$$\int_\Omega (\partial_{\boldsymbol{u}_n} m(\boldsymbol{u}_n, \alpha_n)\boldsymbol{\psi} + \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{v}_n) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}) + \mathbb{C}'(\alpha_n)\beta_n\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{\psi}))\, dx + \int_{\Gamma_N} \partial_{\boldsymbol{u}_n} p(\boldsymbol{u}_n, \alpha_n)\boldsymbol{\psi}\, ds = 0 \qquad \forall \boldsymbol{\psi} \in V^h, \quad (6.67a)$$

$$\text{and } \int_\Omega \frac{G_c \ell}{2c_w} \nabla\varphi \cdot \nabla\beta_n\, dx + \int_\Omega \left(\mathbb{C}'(\alpha_n)\varphi\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v}_n) + \frac{1}{2}\mathbb{C}''(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n) + \frac{G_c}{4c_w}\frac{w''(\alpha_n)}{\ell} \right) \varphi\beta_n\, dx + \int_\Omega \frac{G_c}{\epsilon}\left(\mathcal{M}(\alpha_n - 1) + \frac{1}{\delta t}\mathcal{M}(\alpha_{n-1} - \alpha_n) \right)\varphi\beta_n\, dx + \int_\Omega \partial_{\alpha_n} m(\boldsymbol{u}_n, \alpha_n)\varphi\, dx + \int_{\Gamma_N} \partial_{\alpha_n} p(\boldsymbol{u}_n, \alpha_n)\boldsymbol{\psi}\, ds = \frac{1}{\epsilon\delta t}\int_\Omega \mathcal{M}(\alpha_n - \alpha_{n+1})\beta_{n+1}\varphi\, dx \qquad \forall \varphi \in H^1_D(\Omega^h). \quad (6.67b)$$

As usual the adjoint problem is solved backward in time, i.e., for decreasing indices $n = N-1, \cdots, 1, 0$. One ought to solve the state equation (6.44) (using Algorithm 7) until the last time step, store the solutions $(\boldsymbol{u}_n, \alpha_n)$ for every time-step and retrieve the solutions starting from the last time step. As explained in Section 6.3.3, the bilinear form in the variational formulation (6.46) is not coercive. Hence, for numerical implementations, one ought to use a direct solver or an iterative technique like `GMRES`, that is capable of resolving indefinite matrices.

Remark 24. *The numerical resolution of (6.46) using a direct solver or `GMRES` is slow. If the damage variable is small $\|\alpha_n\|_{L^\infty(\Omega^h)} \ll 1$, the adjoint problem is close to a simple linear elasticity problem and one can rather use an iterative solver meant for positive definite matrices, for instance `CG`, to save computational effort.*

Finally, the space-time discretized version of the shape derivative of (6.45) is

$$J'_N(\Omega^h)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta}\cdot\boldsymbol{n} \sum_{n=0}^{N} \delta t \left(m(\boldsymbol{u}_n, \alpha_n) + \mathbb{C}'(\alpha_n)\beta_n\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{u}_n) + \frac{G_c\ell}{2c_w}\nabla\alpha_n \cdot \nabla\beta_n + \frac{G_c}{4c_w}\frac{w'(\alpha_n)}{\ell}\beta_n + \mathbb{C}(\alpha_n)\boldsymbol{\varepsilon}(\boldsymbol{u}_n) : \boldsymbol{\varepsilon}(\boldsymbol{v}_n) + \frac{G_c}{\epsilon}\left(\mathcal{M}(\alpha_n - 1) - \mathcal{M}\left(\frac{\alpha_{n-1} - \alpha_n}{\delta t} \right) \right)\beta_n \right) ds. \quad (6.68)$$

6.4.3 Regularization and extension of the shape derivative

The shape derivative (6.68) is an approximation of the conical derivative of the gradient-damage equation (6.15), which is inherently non-smooth. This is coupled to the fact that the optimal shape we look for may not have a smooth boundary. Hence neither the shape derivative nor the normal to the boundary Γ may be smooth enough. This necessitates a regularization for the shape derivative (6.68) [40, 8], which ensures that the decent direction is still satisfied. One possibility is consider the H^1 scalar product (instead of the L^2 scalar product) by finding a function $dj(\Omega^h) \in \mathbb{P}^1(D) \subset H^1(D)$, such that

$$\int_D \left(\alpha^2 \nabla dj(\Omega^h) \cdot \nabla\varphi + dj(\Omega^h)\varphi \right) dx = \int_D \mathcal{H}(\phi)\mathcal{H}(h_{\max} - \phi) j(\Omega^h)\, \varphi\, dx \qquad \forall \varphi \in \mathbb{P}^1(D), \quad \text{where} \quad (6.69)$$

- $\alpha = h_{\max}$ is the fixed maximal mesh size
- $\mathcal{H}(\cdot)$ represents the Heaviside function
- ϕ is the distance function corresponding to Ω
- $j(\Omega^h)$ a function defined by formula (6.68) with

$$J'(\Omega^h)(\boldsymbol{\theta}) = \int_\Gamma \boldsymbol{\theta}\cdot\boldsymbol{n}\, j(\Omega^h)\, ds. \quad (6.70)$$

Usually, the right hand side in (6.69) involves an integral only over Γ. Here, since we have taken $\mathbb{P}^1$ functions for state and adjoint variables, their spatial derivative $j(\Omega^h) \in \mathbb{P}^0(\Omega^h)$ and hence is discontinuous on Γ. Thus, we prefer to regularize over a small band of thickness $h_{\max}$ around Γ to ensure that the spatial derivatives are well-captured.

6.4.4 Shape optimization algorithm

We consider the shape optimization problem (6.35) where the admissible shapes must satisfy a constraint on the target volume V_f. To do so, the following Lagrangian is introduced

$$\mathcal{L}(\boldsymbol{u}, \alpha, \boldsymbol{v}, \beta, \Omega^h, \lambda) = J(\Omega^h) + \frac{\lambda}{C_V}\left(V_{\Omega^h} - V_f\right), \tag{6.71}$$

where λ is the Lagrange multiplier for the volume constraint, V_{Ω^h} is the volume of the optimizable domain, and C_V is a normalization constant. The optimizable domain is D unless otherwise mentioned. Denoting by Ω_0^h the initial shape, the constant C_V is defined by (2.27).

We apply a standard gradient-based Uzawa-type algorithm to the Lagrangian (6.71). Let $\mathcal{N}$ be the maximal number of shape optimization iterations (typically $\mathcal{N} = 200$ for most of the test cases in the next section). The iteration number is denoted by i with $1 \leq i \leq \mathcal{N}$. At each iteration i, once the shape derivative $dj(\Omega_i^h)$ is evaluated by (6.69), a pseudo-time step (or descent step) τ is defined by (2.28) (where $h_{\min}$ is the minimal mesh size of the first iteration). The multiplier λ is also updated at each iteration using (2.29), ensuring that the volume will converge (slowly) to the target volume. Then, for the descent step τ, the transport equation (2.17) is solved with a velocity θ_i, given by (6.72)

$$\boldsymbol{\theta}_i = \left(dj(\Omega_i^h) + \frac{\lambda_{i+1}}{C_V}\right)\boldsymbol{n}, \tag{6.72}$$

where $\boldsymbol{n} = \nabla\phi_i$ is the normal to the level-set function associated to the shape Ω_i^h.

To improve the satisfaction of the volume constraint, we apply the following trick. As soon as the volume is close to the volume target, namely $|V_{\Omega_{i+1}} - V_f| \leq 10^{-1}V_f$, we apply a projection algorithm to satisfy the target volume exactly. More precisely, the level-set ϕ_{i+1} is iteratively updated by (2.31) until $|V_{\Omega_{i+1}} - V_f| \leq 10^{-4}V_f$. The newly obtained shape Ω_{i+1}^h is remeshed with `MMG` [61]. Eventually, the objective function $J(\Omega_{i+1}^h)$ is evaluated but is not compared to the previous value $J(\Omega_i^h)$. Summing up this sub-section, we basically implement Algorithm 8.

Remark 25. *If at iteration i the objective function does not decrease, compared to its value at the previous iteration $i-1$, we do not step back to the previous iteration with a reduced descent or pseudo-time step (in order to ensure a decrease in objective function). This is due to the non-smooth nature of the damage problem (6.44), where the onset of fracture is very sensitive to the loading and the geometry and yields large variations of the objective function. We shall explain this issue in greater details in Remark 26 in Sub-section 6.5.2.*

Algorithm 8 Shape optimization for the damage model

Initialize with a shape Ω_0^h and repeat over $i = 1, \cdots, \mathcal{N}$

1. Solve for the state $(\boldsymbol{u}, \alpha)$ in Ω_i^h marching in time from t_1 until t_N using Algorithm 7
2. Solve for the adjoint $(\boldsymbol{v}, \beta)$ in Ω_i^h backward in time from t_N up to t_1
3. Compute the shape derivative using (6.68) and regularize it with (6.69) to deduce $dj(\Omega_i^h)$
4. Update the Lagrange multiplier λ_{i+1} with (2.29)
5. Solve the transport equation (2.17) with the velocity given by (6.72) for the pseudo-time step τ given by (2.28) to obtain the new level-set function $\tilde{\phi}_{i+1}$
6. Re-initialize $\tilde{\phi}_{i+1}$ to the signed distance function ϕ_{i+1} (defining a new shape Ω_{i+1}^h)
7. Compute the volume $V_{\Omega_{i+1}}$. If it is close to the volume target, apply the projection algorithm (2.31) to satisfy exactly the volume constraint.
8. Remesh the box D using `MMG` [61] to obtain the body fitted mesh of the new shape Ω_{i+1}^h

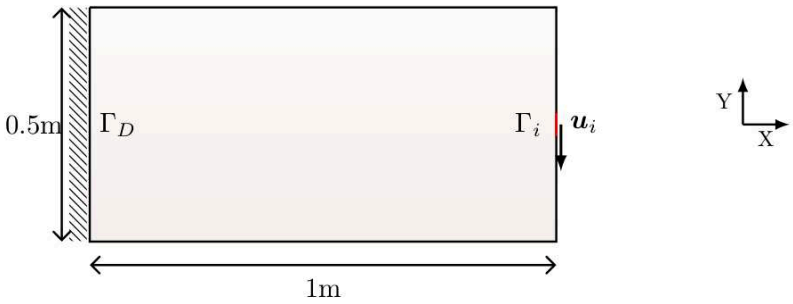


Figure 6.4: 2D Cantilever boundary conditions

6.5 Numerical Results

We now present 2D and 3D shape optimization results for the damage model (6.19), which prove the effectiveness of our algorithm to produce crack-free structures.

6.5.1 Setting and parameters

We choose to maximize rigidity, namely to minimize an objective function which is the total compliance. As already explained in the Remark 17 in Subsection 6.2.1, in case of surface and body forces, it is hard to determine a solution numerically. Hence, we only impose given displacements to the structure, and assume the forces $\boldsymbol{f} = \boldsymbol{0}$ and $\boldsymbol{g} = \boldsymbol{0}$. In such a case, there is a subtle definition of compliance, see e.g. [24], which takes into account the reaction force on the surface where the displacement is imposed. Since the work done by this reaction force is equal to the elastic energy, the total compliance is defined as

$$J(\Omega) = -\int_0^T \int_\Omega \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) dx\ dt. \tag{6.73}$$

The objective function (6.73) is minimized in all test cases, except otherwise mentioned.

The final time is $T = 1s$ and the time step is $\delta t = 0.15s$. The solution $(\boldsymbol{u}, \alpha)$ is computed by solving the penalized formulation (6.44) using Algorithm 7. The penalization parameter is chosen equal to $\epsilon = 10^{-5}$.

The isotropic degradation function (6.66) is considered for all test cases except the ones in Subsubsections 6.5.2.2 and 6.5.4.1. The residual stiffness $\mathbb{C}_{res}$ is taken to be zero (for reasons explained at the end of the Subsection 6.4.1). For all test cases in this section, only the dissipation function (6.17) of the `DL` model is considered, except the one in Subsubsection 6.5.2.1, where the dissipation function (6.16) of the `DQ` model shall be considered.

Since we are in a quasi-static evolution framework, the rate of increment of the imposed displacement has no effect on the solution $(\boldsymbol{u}, \alpha)$ at the final time $T = 1s$. But the rate does have a strong influence on the objective function (6.73) (since it contains a time-integral). In order to see a greater influence of the damage variable α on the optimized shape, we consider an imposed displacement $\bar{\boldsymbol{u}}(t)$ that grows from zero to a certain value and then remains constant for some period.

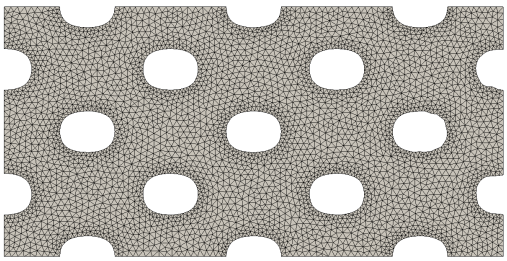
6.5.2 2D Cantilever

We study a 2D cantilever beam represented by a rectangle of dimensions $1m \times 0.5m$ as shown in Fig. 6.4. The cantilever beam is clamped all along its leftmost edge and subjected to an applied vertical displacement,

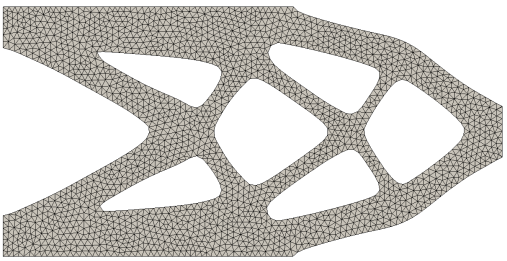
$$\bar{\boldsymbol{u}}(t) = (0, 4\min(1.2t, 1)) \times 10^{-4} m, \quad t \in [0, 1], \tag{6.74}$$

on a centered part of its rightmost edge $\Gamma_{\bar{\boldsymbol{u}}}$ of length $0.1m$. The above displacement (6.74) is chosen in such a way that the initial shape of Fig. 6.5a suffers from a crack as seen in Fig. 6.7a. A target volume $V_f = 0.25m^2$ is imposed for all the test cases in this subsection. The parameters of the remeshing tool `MMG` are: $h_{\min} = 0.0064m, h_{\max} = 0.0128m$.

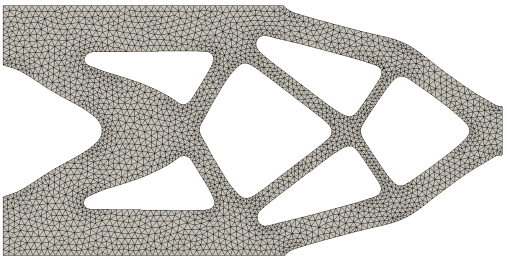
The initial shape is displayed in Fig. 6.5a. The shape obtained by minimizing (6.73) for linear elasticity (without any damage) is plotted in Fig. 6.5b. The shape obtained by minimizing (6.73) for the damage model is plotted in Fig. 6.5c. The convergence history is plotted in Fig. 6.6.



(a) Initialization

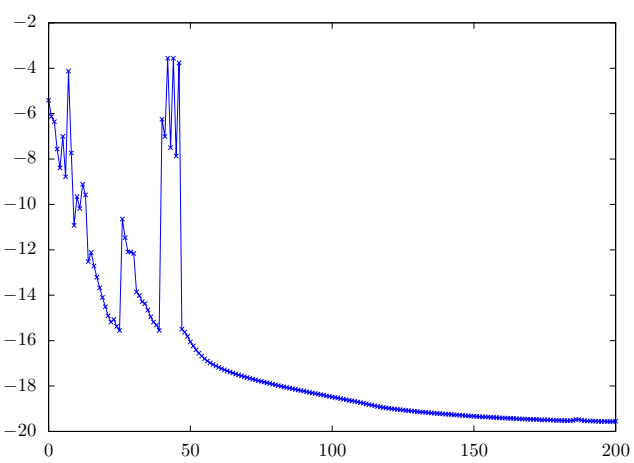


(b) Final shape obtained for linear elasticity

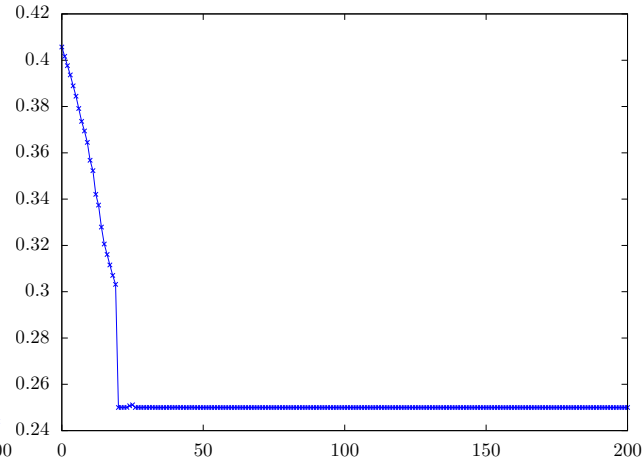


(c) Final shape for model (6.17), isotropic degradation (6.66)

Figure 6.5: Meshes for different wedge shapes for minimization of (6.73), imposed displacement (6.74)



(a) Total elastic energy (6.73) vs iterations

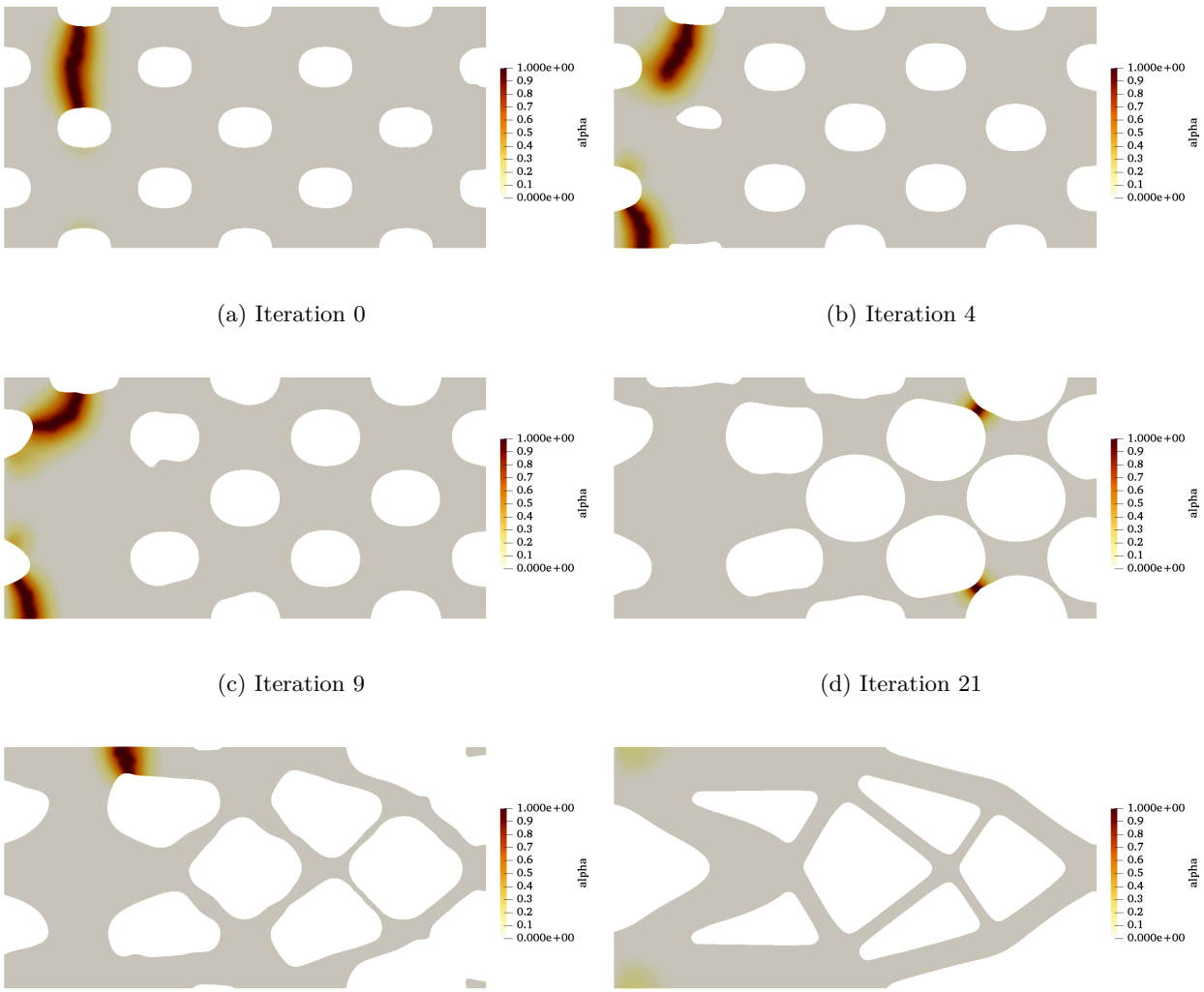


(b) Volume vs iterations

Figure 6.6: Convergence history for the shapes (6.5c)

As can be seen in Fig. 6.5, the optimized shapes, with or without damage in the mechanical model, are slightly different, but share the same topology. The shape in Fig. 6.5c do not undergo a crack, unlike the shape in Fig. 6.5b that does undergo a crack (see Fig. 6.8a). For the shape in Fig. 6.5c, the damaged region in the intermediate shapes is plotted in Fig. 6.7. The optimization algorithm indeed tries hard to remove every damaged or cracked region that appears. We observe that the crack appears in the cantilever at several locations, taking different configurations at each iteration and disappears finally after the 45-th iteration.

Remark 26. *One can see in Fig.6.6 that the objective function (6.73) features very strong oscillations during the first 50 iterations although the volume constraint is nicely satisfied after 20 iterations. These peaks in the convergence history occur whenever there is a transition from a shape without any crack to a shape with*



(a) Iteration 0

(b) Iteration 4

(c) Iteration 9

(d) Iteration 21

(e) Iteration 40

(f) Iteration 200

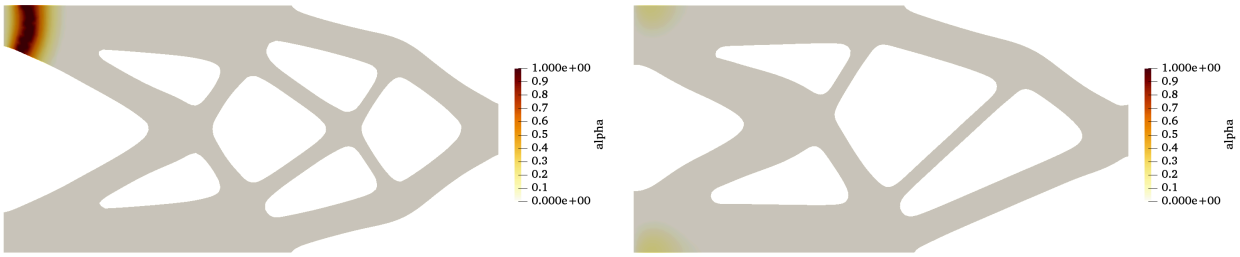
Figure 6.7: damage variable α at the final time plotted for several optimization iterations for the shape in Fig.6.5c

a crack, or when the crack changes from one position to a completely different one (see Fig. 6.7). In other words, small perturbations in the shape Ω can result in the appearance or disappearance of cracks, leading to abrupt changes in the objective function (6.73). Reducing the descent step τ would not help here because the onset of fracture is a discontinuous process with respect to load or geometry variations. Typically, the growth in time of a fracture can be discontinuous. Therefore, it is plausible that a small change in the geometry of the shape can induce a large change in the crack profile and thus in the value of the objective function (6.73). This non-smooth character of fracture or damage is well documented in the theoretical literature [29, 45, 46, 47, 148] but also in the numerical literature [35]. Note that our derivation of the shape derivative in Subsection 6.3.2 was performed under the assumption of a smooth solution of the damage model (6.19).

The regularization of the damage field α with the characteristic length ℓ or the penalization process of the damage irreversibility do not help at all on this matter. Our numerical experiments confirm this non-smoothness of the damage problem and the discontinuity of the objective function with respect to shape variations. Therefore, it is questionable to use a gradient-descent method to minimize the objective function (6.73). Nevertheless, the presented test cases show that, after some early oscillations, our gradient algorithm does converge to a crack-free optimal shape. There are two key ingredients for this relative success. First, although the descent step τ is adapted at each iteration by formula (2.28), we do not test if the objective

function decreases at each iteration and we never step back with a smaller descent step. Second, we rely on the backtracking Algorithm 7 (following [35]) which plays a pivotal role in ensuring a stable damage evolution. As a consequence, the shapes obtained in the final iterations are more stable in the sense that small geometric perturbations do not cause the appearance of a crack. Hence oscillations in the objective function are avoided and a smooth convergence of the objective function is attained. This is confirmed by our attempt to replace the backtracking Algorithm 7 by the simpler Algorithm 6 (without backtracking) in our shape optimization Algorithm 5. We noticed that the fluctuations in the objective function were more violent than the ones obtained with backtracking, which was hindering convergence.

One could think that changing the initialization could improve the convergence of the shape optimization for the damage model. For example, instead of starting from the periodically perforated initial shape in Fig. 6.5a, it is possible to initialize the damage model optimization with the optimal shape for linear elasticity in Fig. 6.5b. We perform this new test case and the result in Fig. 6.8b is quite deceiving. Indeed, it takes 600 iterations (3 times more, see Fig. 6.9) to converge to the shape of Fig. 6.8b, which is different from that previously obtained in Fig. 6.5c, slightly less optimal since the objective function for Fig. 6.5c is 0.7% better than for the shape of Fig. 6.8b. Note that both shapes of Fig. 6.8b and 6.5c feature no crack.



(a) Initial shape (6.5b) (b) Final shape

Figure 6.8: damage variable α at the final time for minimization of (6.73), imposed displacement (6.74), DL model (6.17), isotropic degradation (6.66)

6.5.2.1 Quadratic dissipation function DQ

Only in this subsubsection, we consider the DQ model in (6.16). In the DQ model, the fracture onsets from $t = 0$, resulting in a greater damage compared to the DL model (6.17). Thus, the imposed displacement is slightly reduced from (6.74) to

$$\bar{\boldsymbol{u}}(t) = (0, 3.125 \min(1.2t, 1)) \times 10^{-4} m, \quad t \in [0, 1].$$

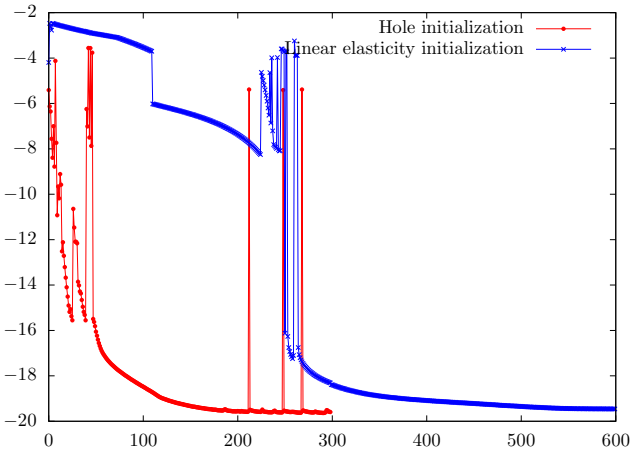


Figure 6.9: Objective (6.73) vs iterations for the shapes (6.5c) and (6.8b)

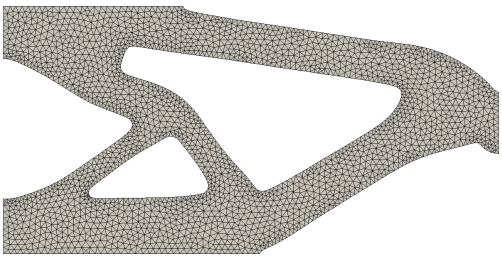


Figure 6.10: Final shape for model (6.16), anisotropic degradation (6.66)

The initial shape in Fig.6.5a converges to the shape in Fig.6.10. This shape is asymmetric, unlike in the case of DL model (shape in Fig.6.5c). This could be attributed to the higher tendency to damage in the structure, as seen in the intermediate shapes obtained for this optimization in Fig.6.11. Here, we see that the damage does not localize (or concentrate) as much as in the case of DL model (see Fig.6.7).

6.5.2.2 Traction-only degradation

For this test case only, we replace the isotropic degradation function (6.66) by the following traction-only degradation function [132]

$$\mathbb{C}(\boldsymbol{u},\alpha) = \Big(\mathcal{H}(div\boldsymbol{u})(-1+(1-\alpha)^2)+1\Big)\mathbb{C}_0 + \mathbb{C}_{res}, \tag{6.75}$$

where $\mathcal{H}$ denotes the Heaviside function. If $div\boldsymbol{u} \geq 0$ the material is said to be in traction, otherwise it is in compression. The degradation function (6.75) is constructed in such a way that damage occurs only under tension. In other words, when $div\boldsymbol{u} < 0$, whatever the value of α, one has $\mathbb{C}(\boldsymbol{u},\alpha) = \mathbb{C}_0$. Such a traction-only degradation function is more realistic since it can make a difference between an opening and a closing cracks (this idea was introduced in [21] with a slightly different degradation function). The Hooke's tensor $\mathbb{C}(\boldsymbol{u},\alpha)$ obviously depends on $\boldsymbol{u}$, and furthermore is not even differentiable with respect to $\boldsymbol{u}$. Nevertheless, for the numerical test here, we ignore this dependence and do not take it into account in the adjoint equation.

The same 2D cantilever beam, as in Subsection 6.5.2, is considered for the new degradation function (6.75). Starting from the initialization in Fig. 6.5a the algorithm converges to the shape of Fig. 6.12. The convergence history, in Fig. 6.13, is slightly smoother than in Fig. 6.6. The optimized shape in Fig. 6.12 is somehow intermediate between those in Fig. 6.5 for linear elasticity and the original isotropic degradation function (6.66).

6.5.3 2D Cantilever: some variants

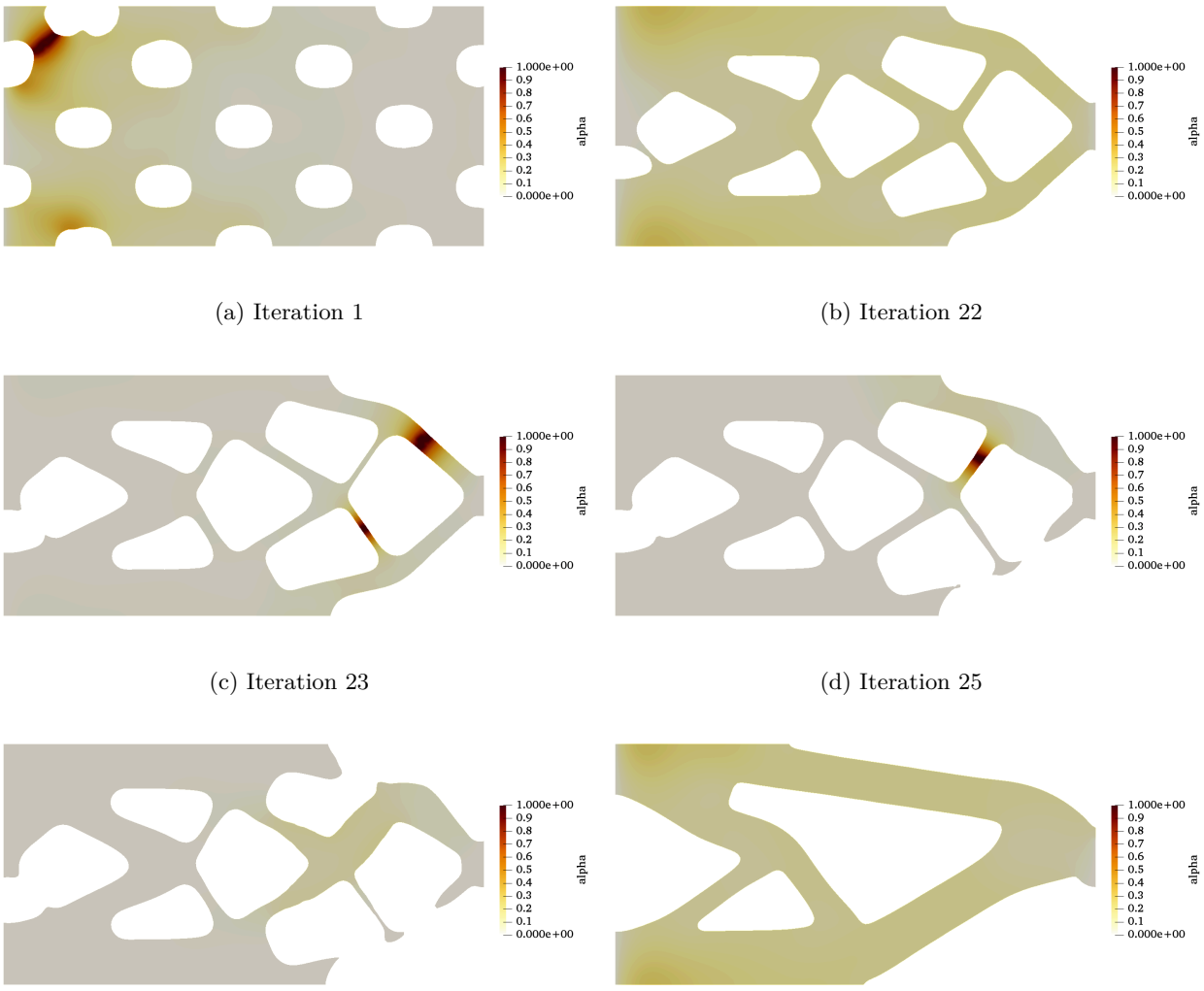
From the test cases of the previous subsections, we make two important observations:

1. the objective function converges in a highly non-smooth manner (see Fig. 6.6),
2. the final optimized shapes have zones that undergo slight damage (see Fig. 6.5 and 6.8b).

The first point is a consequence of two facts: (i), the damage model is very sensitive to all parameters (loading, geometry, material parameters), (ii), the initialization with a periodically perforated shape, which is very far from any possible optimal shape, implies that the optimization process will explore "wild" intermediate shapes. In order to ensure that the convergence of the objective function is smoother and that, if possible, the final shape features no damaged zones, we implemented the following three variants of our approach to improve the smoothness of the optimization process.

1. Incremental approach: the residual stiffness in the degradation function (6.66) is changed, taking a decreasing residual stiffness, as the iteration number increases

$$\mathbb{C}_{res} = \kappa_i \mathbb{C}_0 \tag{6.76}$$



(a) Iteration 1

(b) Iteration 22

(c) Iteration 23

(d) Iteration 25

(e) Iteration 28

(f) Iteration 200

Figure 6.11: damage variable α at the final time plotted for several optimization iterations for the shape in Fig.6.10

where κ_i is (heuristically) given by

$$\kappa_i = \frac{1}{10}\left(1 - \frac{i}{\mathcal{N}}\right)^p, \tag{6.77}$$

where i it the iteration number, p is an integer exponent and $\mathcal{N}$ is the total number of optimization iterations.

2. Symmetrization: the level-set function is symmetrized about the horizontal x-axis (given the observation that the optimal shapes of Fig. 6.5 have this symmetry which can be broken by the non-linear character of the damage model).

3. Weighted objective function: to give more importance to the final time, when the structure is more likely to endure fracture, replace the objective function (6.73) by its weighted or truncated version

$$J(\Omega) = -\int_{T_0}^{T}\int_{\Omega} \mathbb{C}(\alpha)\boldsymbol{\varepsilon}(\boldsymbol{u}) : \boldsymbol{\varepsilon}(\boldsymbol{u}) dx\, dt, \tag{6.78}$$

where $T = 1$ and $T_0 = 0.85 > 1/1.2$ (thus the imposed displacement (6.74) remains constant on this time interval).

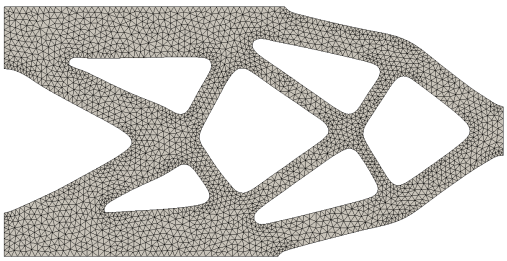
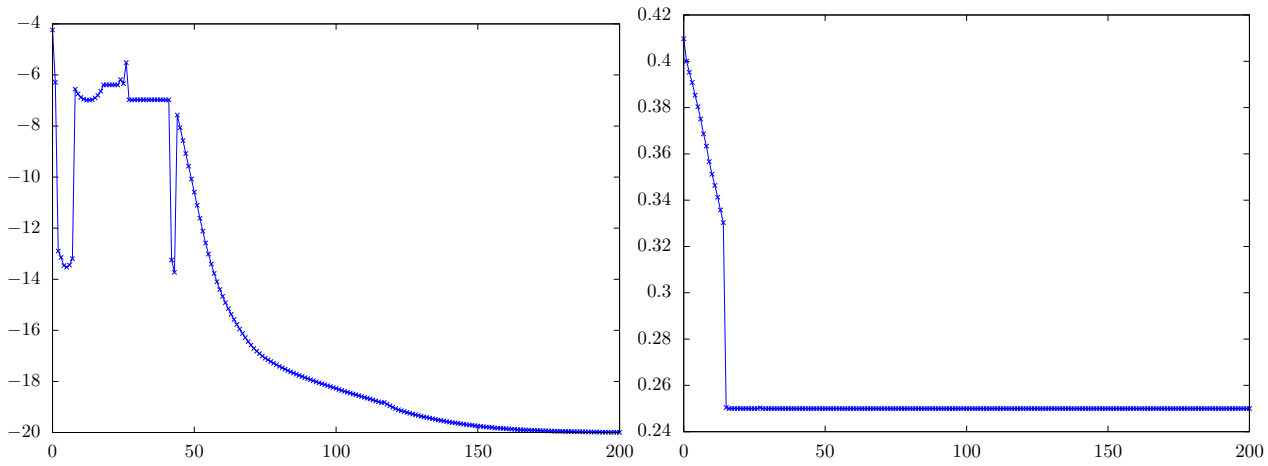


Figure 6.12: Final shape for model (6.17), anisotropic degradation (6.75)



(a) Objective function (6.73) vs iterations

(b) Volume vs iterations

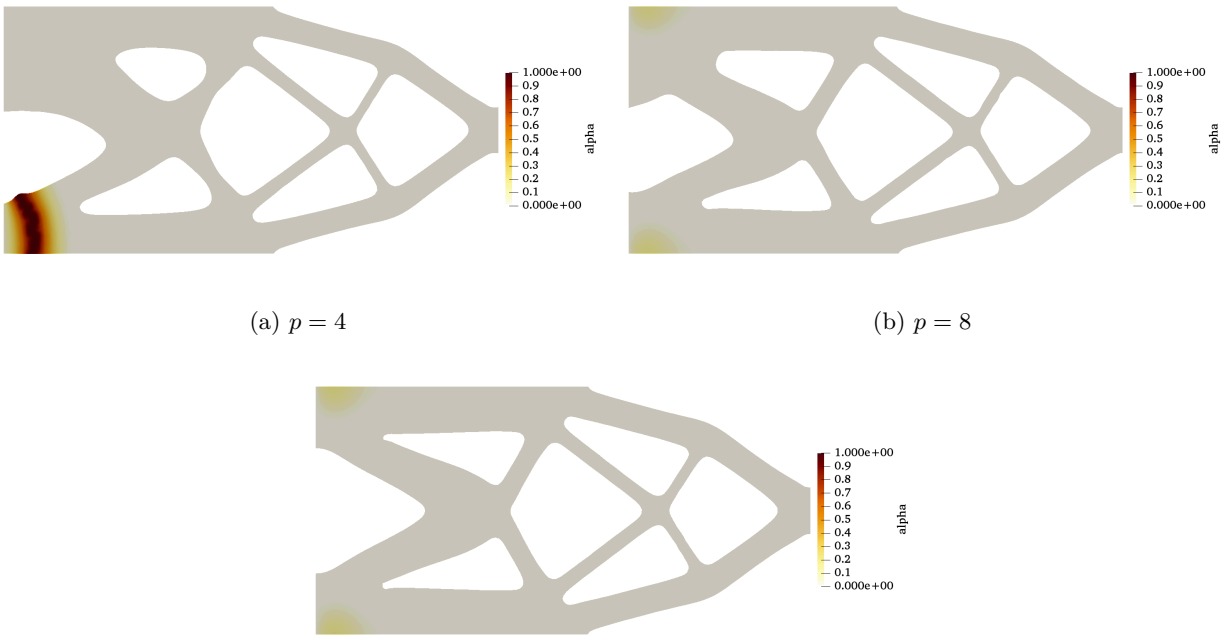
Figure 6.13: Convergence history for the shape of Fig. 6.12

These three variants are tested with Algorithm 5 and exactly 200 iterations.

In the first variant, the parameter κ_i of the residual stiffness in (6.76) is essentially the stiffness of the damaged phase. If κ_i is not small, the objective function does not increase significantly when the structure undergoes a crack. However, when κ_i is small, meaning that the residual stiffness approaches zero, the objective function increases dramatically when a crack appears. Given that the objective function oscillates in the first iterations (see Fig. 6.6), it is legitimate to try to reduce these oscillations by having a larger residual stiffness at the start and then gradually decrease it, as is the case with (6.77). Three different exponent are tested: $p = 4, 8$ and 10. The graphs of κ_i are plotted in Fig. 6.15a. The shapes obtained for the three exponents are displayed in Fig. 6.14 and the corresponding convergence histories are given in Fig. 6.15b. As can be seen in Fig. 6.14, the shapes obtained for $p = 4$ and $p = 8$ are not optimal, unlike the shape obtained for $p = 10$ (which resembles the one of Fig. 6.5c). In the convergence history of Fig. 6.15b, we observe that the fluctuations in the objective functions are high for $p = 4$, lesser for $p = 8$ and least for $p = 10$. Hence it seems preferable to decrease the residual stiffness in (6.77) with a large exponent p.

In the second variant, Algorithm 5 is slightly modified. After the re-initialization step 6, the level-set function ϕ_{i+1} is symmetrized around the horizontal central axis, which ensures that the shape remains symmetric. After 200 iterations of shape optimization, the shape in Fig. 6.16a is obtained, which is obviously not optimal as it undergoes a crack, unlike the shape in Fig. 6.12.

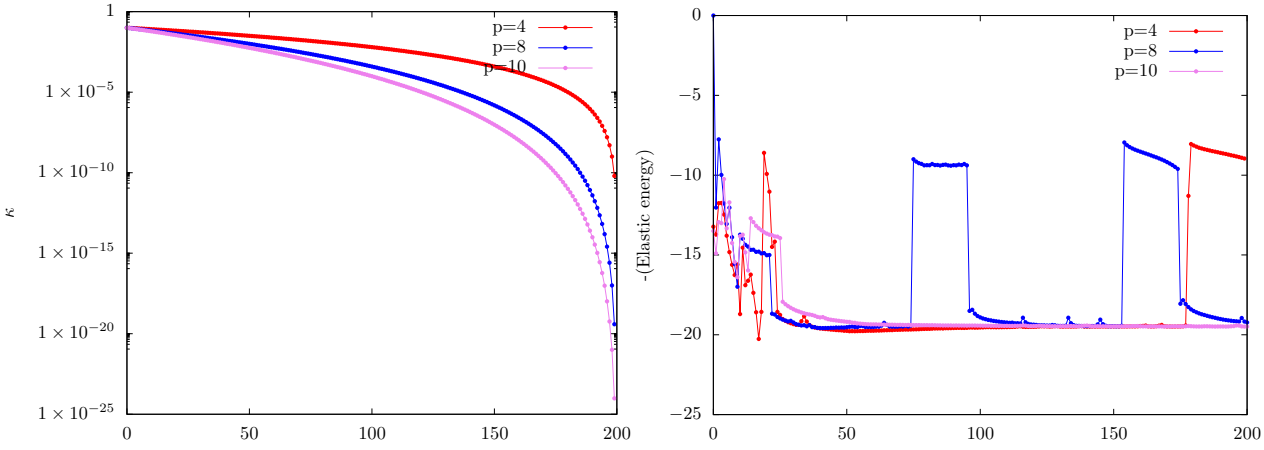
For the third variant, we obtain the shape of Fig. 6.16b, which is similar to the one obtained in Fig. 6.5c and undergoes only a very minor damage. This is quite deceiving as we weighted or truncated the objective function (6.73) in the hope of totally eliminating any damage zone. Thus, considering a weighted version of the objective function (6.73) is not effective in avoiding damaged zones.



(a) $p = 4$ (b) $p = 8$

(c) $p = 10$

Figure 6.14: damage variable α at the final time for minimization of (6.73) (200 iterations), imposed displacement (6.74), `DL` model (6.17), isotropic degradation (6.76) and residual stiffness κ decremented as in Fig.6.15a



(a) κ vs iterations (b) Objective (6.73) vs iterations

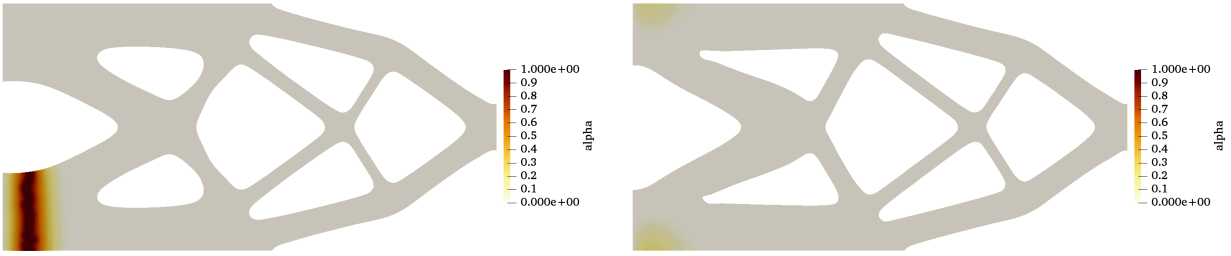
Figure 6.15: Convergence history for the shapes in Fig.6.14

6.5.4 2D wedge

We now study a 2D wedge as shown in Fig.6.17. The wedge is fixed on its leftmost support, with a zero vertical displacement on its rightmost support, and subjected to an imposed displacement,

$$\bar{\boldsymbol{u}}(t) = (0, 2.33\min(1.2t, 1)) \times 10^{-4} m, \quad t \in [0, 1], \tag{6.79}$$

on the center of its top edge as shown in Fig. 6.17. These prescribed boundary conditions and the sharp corner at the center ensure that there is a crack initiation at the notch in the center. The input parameters of the remesher `MMG` are $h_{\min} = 0.0064m$ and $h_{\max} = 0.0128m$. The target volume is taken to be $0.2m^2$. The initial shape, as shown in Fig. 6.18a, converges to the shape in Fig. 6.18b for linear elasticity and to



(a) Symmetrization, strategy (2) (b) Weighted objective (6.78), strategy (3)

Figure 6.16: α plotted at the final time in shape obtained after 200 iterations

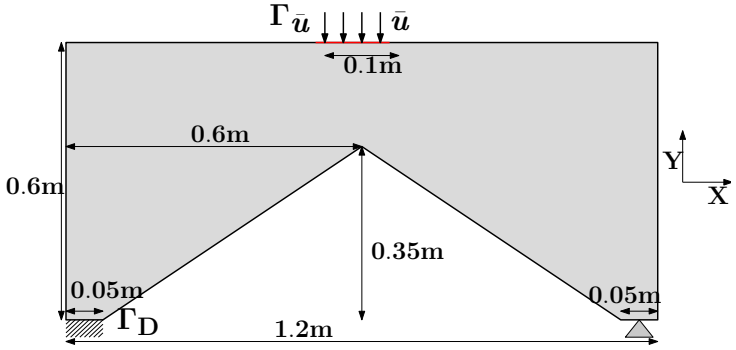


Figure 6.17: 2D wedge boundary conditions

the shape in Fig. 6.18c for the damage model. The convergence history is plotted in Fig. 6.19. Like in the previous 2D-cantilever case, the shapes obtained for linear elasticity or for the damage model have the same topology. In Fig. 6.20 are plotted the damaged regions for several intermediate shapes corresponding to the final shape of Fig.6.18c. A crack is clearly present during the first iterations although the notch is rounded by a hole. Nevertheless, after 21 iterations, the algorithm finds a crack-free shape, finally converging to the final shape at iteration 200, which features a very little amount of damage close to the notch.

6.5.4.1 Traction-only degradation

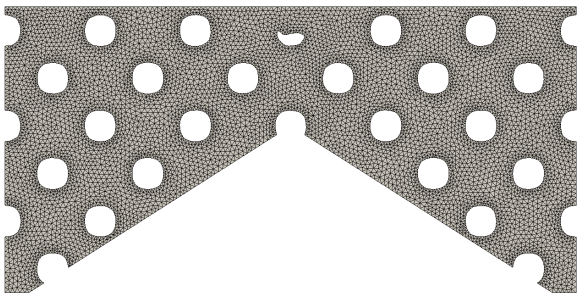
We now consider the traction-only degradation function (6.75) for the 2D wedge of this subsection. This function is treated in the same way as explained in Subsubsection 6.5.2.2. Starting from the initialization in Fig. 6.18a the algorithm converges to the shape of Fig. 6.21. The final shape obtained here in Fig. 6.21 seems to be very rigid near the notch at the center, compared to the shape in Fig.6.18c. This is because the notch is the only zone that can undergo damage and ultimately fracture (unlike the previous case where the support legs can as well undergo damage). The damaged regions in the intermediate shapes are plotted for the shape (6.21) in Fig.6.22. After the 41 iterations, the algorithm finds a crack-free shape, finally converging to the shape in iteration 200 that damages very little.

6.5.5 2D L-beam

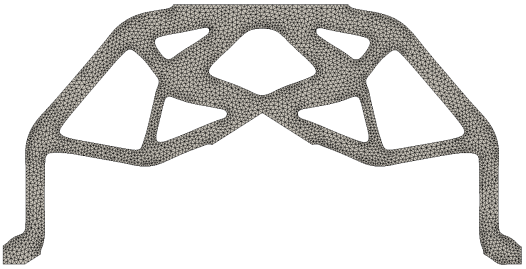
The last 2D example is an L-beam which is a meaningful test case because of its re-entrant right-angled corner, which is prone to crack initiation. As shown in Fig. 6.23, the L-beam is fixed on its topmost edge, subjected to an imposed displacement,

$$\bar{\boldsymbol{u}}(t) = (0, 1 \min(1.2t, 1)) \times 10^{-3} m, \quad t \in [0, 1], \tag{6.80}$$

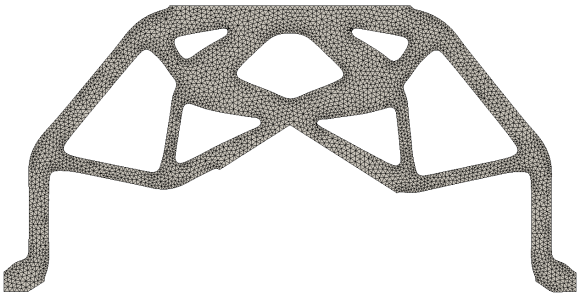
on a small part of its rightmost edge with a vertical force as shown in Fig. 6.23. A target volume $V_f = 2m^2$ is considered. The input parameters for `MMG` are $h_{\min} = 0.01m$ and $h_{\max} = 0.02m$. From the initial shape in Fig. 6.24a, the minimization of the objective function (6.73) for linear elasticity leads to the final shape



(a) Initialization

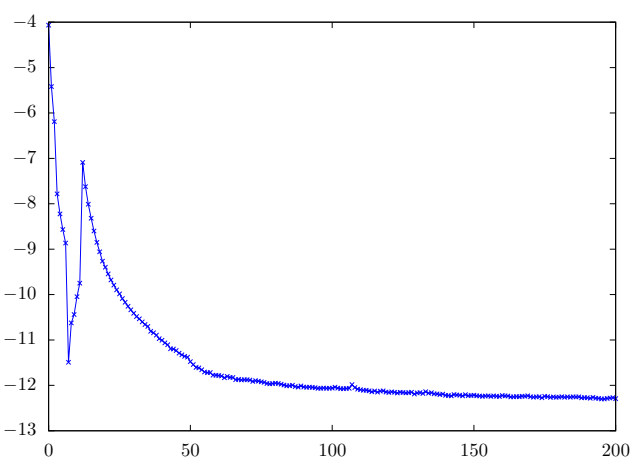


(b) Final shape obtained for Linear elasticity

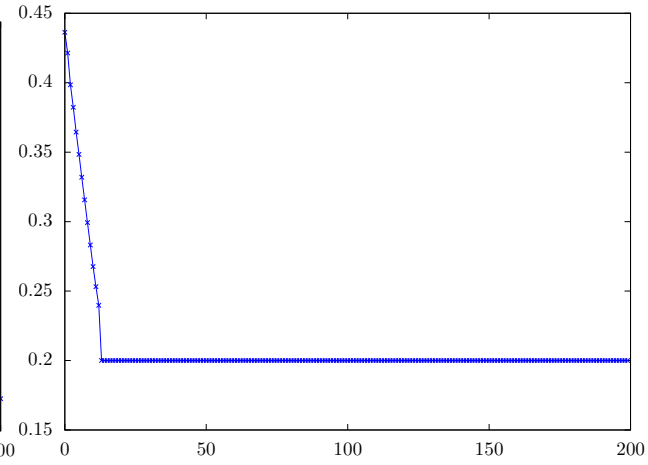


(c) Final shape for model (6.17), isotropic degradation (6.66)

Figure 6.18: Meshes for different wedge shapes for minimization of (6.73), imposed displacement (6.79)



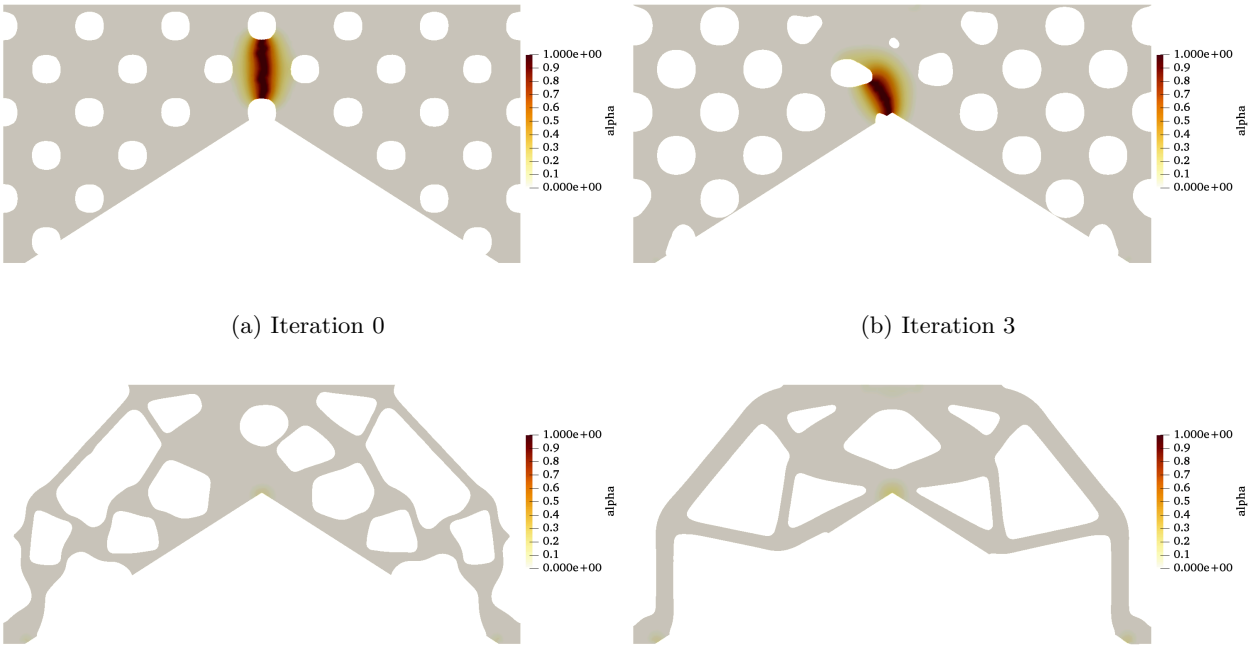
(a) Total elastic energy (6.73) vs iterations



(b) Volume vs iterations

Figure 6.19: Convergence history for the shapes (6.18c)-(6.21)

in Fig. 6.24b. Contrary to the previous test cases, we now initialize the optimization for the damage model with this shape of Fig. 6.24b, instead of the periodically perforated initialization of Fig. 6.24a. It yields the final shape of Fig. 6.24c. In order to compare the shapes in Fig. 6.24b and in Fig. 6.24c, we perform a damage computation for both shapes with $\bar{\boldsymbol{u}}(t)$, given by (6.80), and plot the damage variable α at the final time in Fig. 6.25: obviously, the shape in Fig. 6.24c does not undergo a crack, unlike the shape in Fig. 6.24b. The damage variable for the intermediate shapes corresponding to the optimization for the shape in Fig.6.24c are plotted in Fig.6.26.



(a) Iteration 0

(b) Iteration 3

(c) Iteration 21

(d) Iteration 200

Figure 6.20: Damage variable α at the final time T plotted for several optimization iterations for the shape in Fig.6.18c

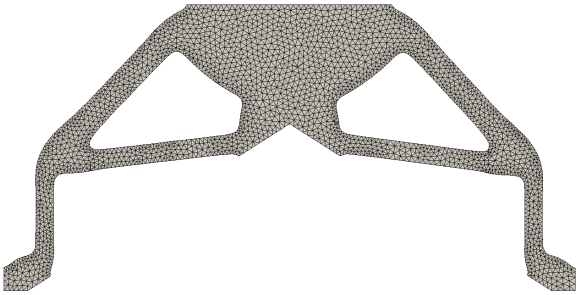


Figure 6.21: Final shape for model (6.17), anisotropic degradation (6.75)

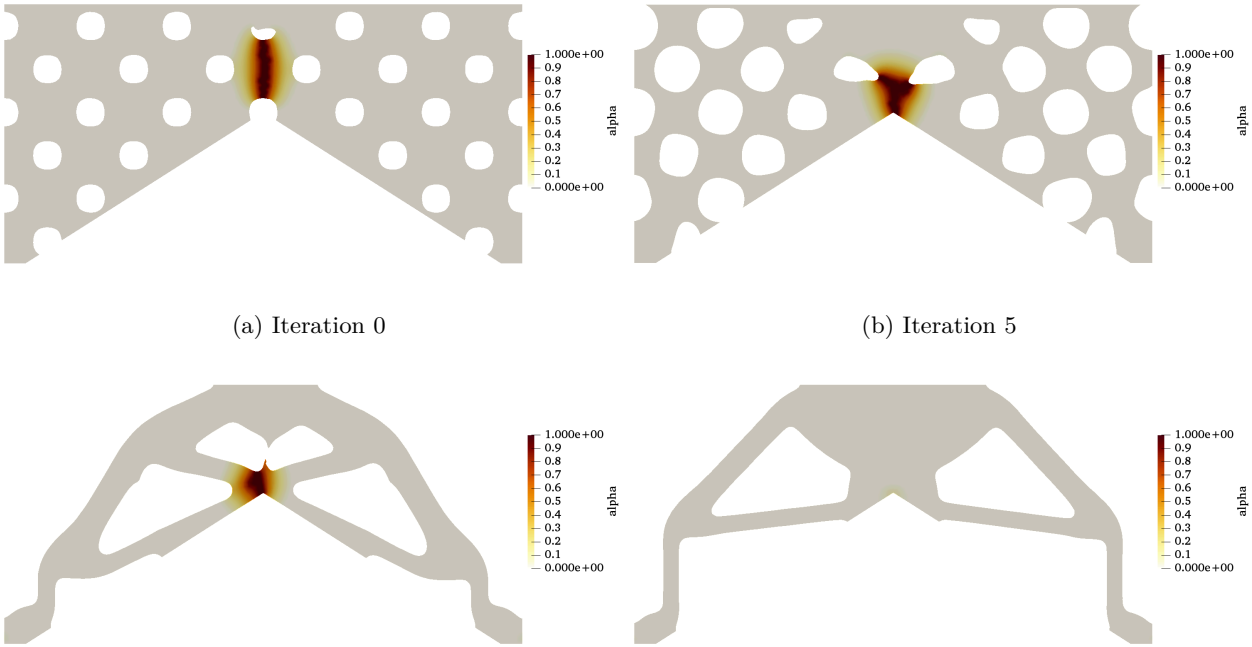
6.5.6 3D wedge

Continuing the study of test cases with sharp corners, we consider a 3D wedge as shown in Fig. 6.27. The wedge is supported on four square surfaces, each being $0.05m \times 0.05m$, at the bottom of the working domain, see Fig. 6.27. The wedge is clamped in all three directions on one surface and only along the vertical y-direction on the remaining three surfaces. The wedge is subjected to an imposed displacement $\bar{\boldsymbol{u}}(t)$ on a $0.1m \times 0.1m$ square surface at the top (in yellow in Fig. 6.27). The input parameters for `MMG` are $h_{\min} = 0.013m$ and $h_{\max} = 0.026m$. Here, we investigate the impact of increasing the magnitude of the imposed displacement on the optimized shape and hence consider two functions

$$\bar{\boldsymbol{u}}(t) = (0, 7\min(1.2t, 1)) \times 10^{-5} m, \quad t \in [0, 1], \quad \text{and} \tag{6.81}$$

$$\bar{\boldsymbol{u}}(t) = (0, 8.5\min(1.2t, 1)) \times 10^{-5} m, \quad t \in [0, 1]. \tag{6.82}$$

The target volume is chosen to be $0.7m^3$. The initial shape of Fig. 6.28a converges to the shape of Fig. 6.28b for linear elasticity. For the damage model, the same initial shape converges to the shape of Fig. 6.28c



(a) Iteration 0

(b) Iteration 5

(c) Iteration 41

(d) Iteration 200

Figure 6.22: damage variable α at the final time plotted for several optimization iterations for the shape in Fig.6.21

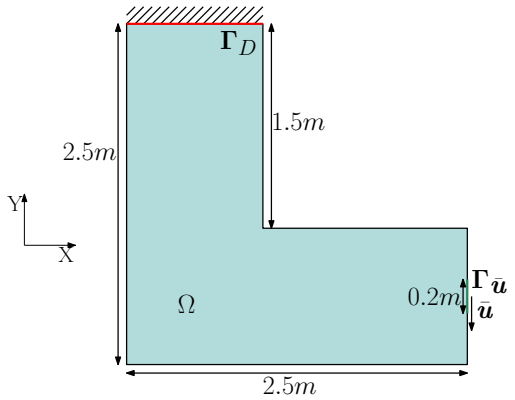


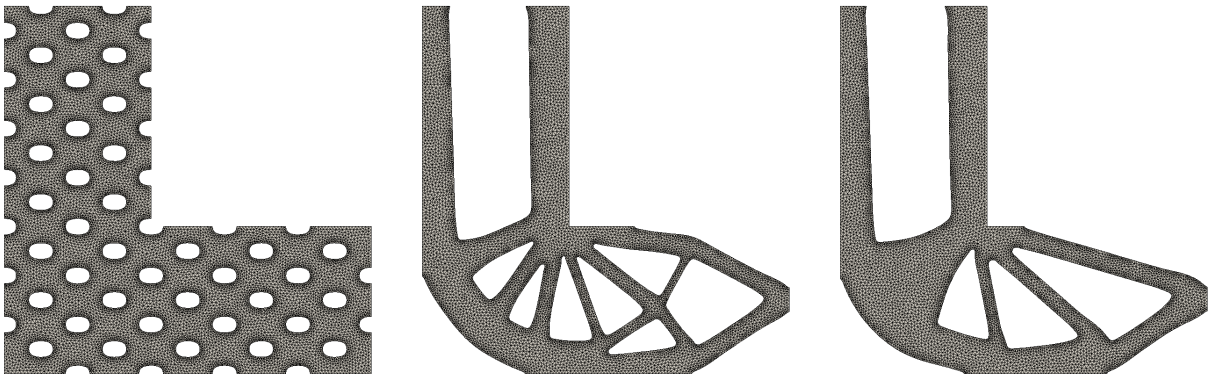
Figure 6.23: 2D L-beam boundary conditions

for the displacement (6.81) and to the shape of Fig. 6.28d for the displacement (6.82). The corresponding convergence histories are plotted in Fig. 6.29. It is remarkable that the shapes in Fig.6.28c and in Fig.6.28d are very different, although the imposed displacement (6.81) and (6.82) (for which they were optimized) are quite close. This illuminates the highly non-linear nature of the damage model (6.19).

The damage variable α for the intermediate shapes corresponding to the optimal shape in Fig. 6.28d are plotted in Fig.6.30. There, one can check that the intermediate shapes undergo crack, not only at the notch at the center, but also at other places. As expected, the resulting final shape in Fig. 6.82 is crack-free.

6.5.7 3D Column

This new test case is a 3D column (see Fig. 6.31a) which is $4m$ high, fixed at the bottom (in red) and subjected to an imposed displacement $\bar{\boldsymbol{u}}(t)$ on the top (in yellow). The precise geometrical definition of this column can be found in [78].



(a) Initialization mesh (b) Final shape for Linear elasticity (c) Fracture

Figure 6.24: Meshes for different L-beam shapes for minimization of (6.73), model (6.17), isotropic degradation (6.66)

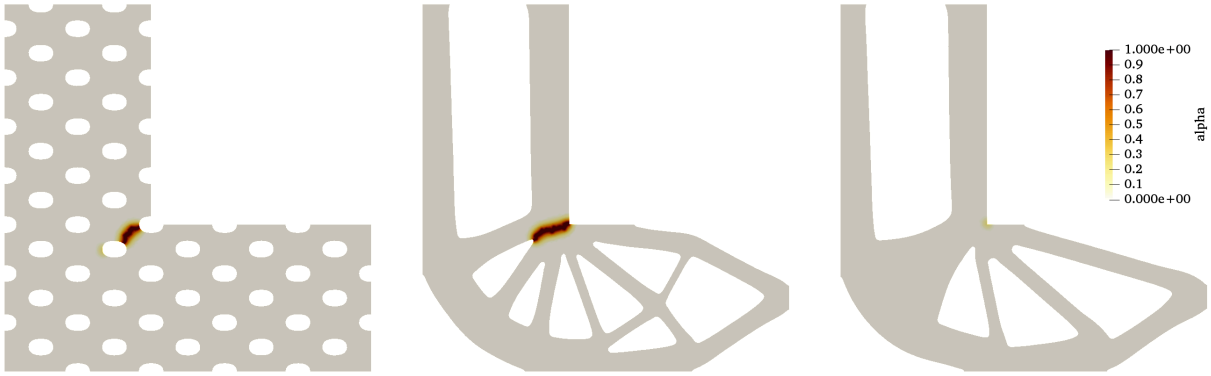


Figure 6.25: damage variable α at the final time for shapes in Fig.6.24, model (6.17), isotropic degradation (6.75)

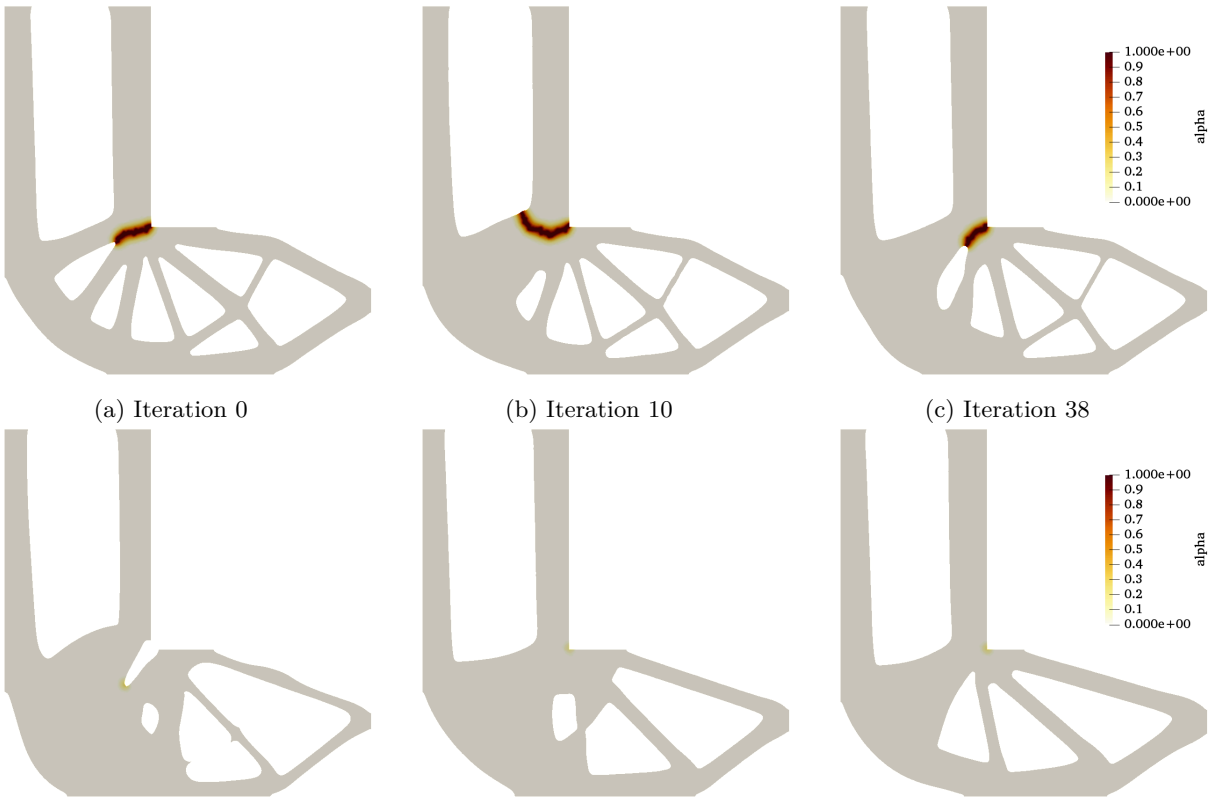
Here again, we investigate the impact of increasing the magnitude of the imposed displacement on the optimized shape and hence consider two functions

$$\bar{\boldsymbol{u}}(t) = (0, 2.88\min(1.2t, 1)) \times 10^{-4}m, \quad t \in [0,1]s, \quad \text{and} \tag{6.83}$$

$$\bar{\boldsymbol{u}}(t) = (0, 2.97\min(1.2t, 1)) \times 10^{-4}m, \quad t \in [0,1]s. \tag{6.84}$$

The input parameters of the remesher `MMG` are $h_{\min} = 0.02m$ and $h_{\max} = 0.04m$. Since the characteristic length (using the formula (6.62)) is $\ell = 0.0375m$, the condition (6.64) is violated. Instead of choosing a smaller $h_{\max}$, we increase the characteristic length to $\ell = 0.075m$ to ensure that the condition (6.64) is satisfied. The target volume is $2.5m^3$. The shape is initialized as shown in Fig. 6.31b. The final shape obtained for linear elasticity is plotted in Fig. 6.32a. For the damage model, it is plotted in Fig. 6.32b and 6.32c for the imposed displacement (6.83) and (6.84), respectively. Clearly again, we see that a slight increase in the imposed displacement results in a very different optimized shape for the damage model.

We compare the performance of the three shapes in Fig. 6.32 for the same damage model and for the same linear elasticity system with the imposed displacement (6.84) and plot the objective function (6.73) values in Table 6.1. As can be expected, the shapes optimized for damage have much better performances with the damage model. But, surprisingly, the performance of the three shapes for linearized elasticity are very similar (the optimal shape for linear elasticity is less than a fraction of percent better than the two other ones). The damage variable α for those three shapes is plotted in Fig. 6.35. As one can check in Fig. 6.35, the shapes in Fig. 6.32b and 6.32c do not undergo a crack whereas the shape in Fig.6.32a does.



(a) Iteration 0 (b) Iteration 10 (c) Iteration 38

(d) Iteration 49 (e) Iteration 80 (f) Iteration 200

Figure 6.26: damage variable α at the final time plotted for several optimization iterations for the shape in Fig.6.24c

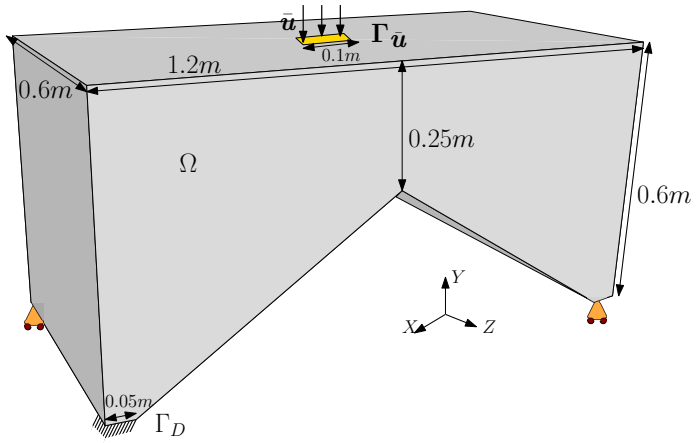
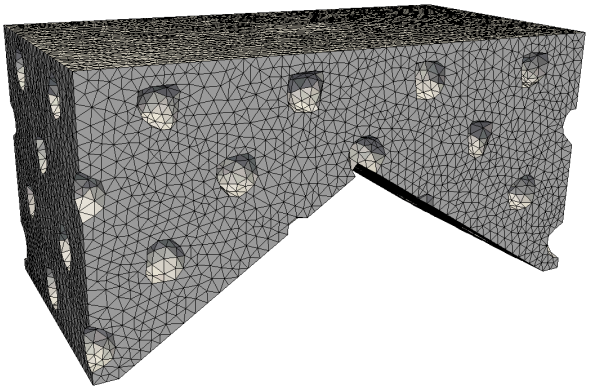


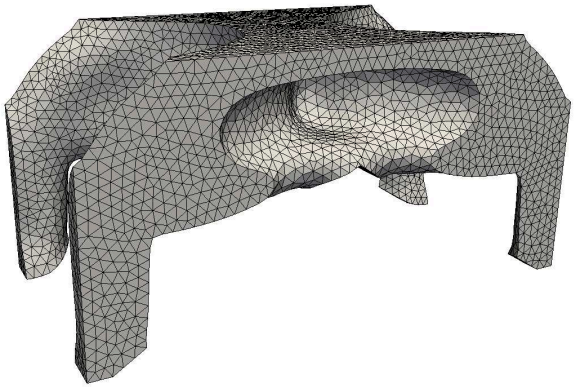
Figure 6.27: 3D wedge boundary conditions

	Fig. 6.32a	Fig. 6.32b	Fig. 6.32c
Objective (6.73) for linear elasticity	−183.47	−183.65	−186.83
Objective (6.73) for damage model	−179.75	−186.49	−186.13

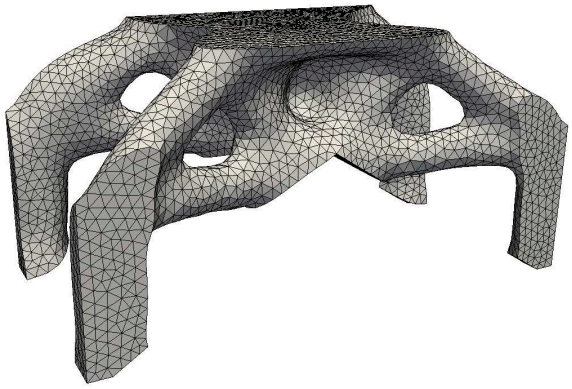
Table 6.1: Comparison between shapes in Fig. 6.32a and 6.32c for the imposed displacement (6.84)



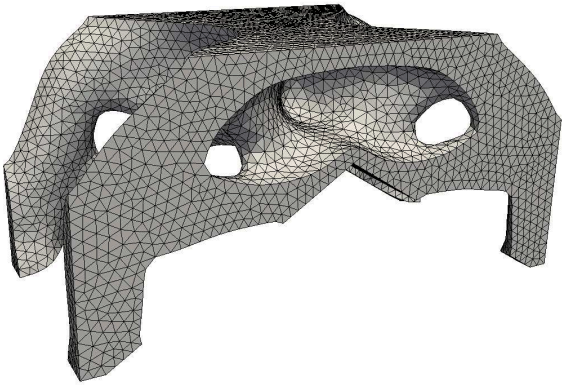
(a) Initialization



(b) Final shape for Linear elasticity

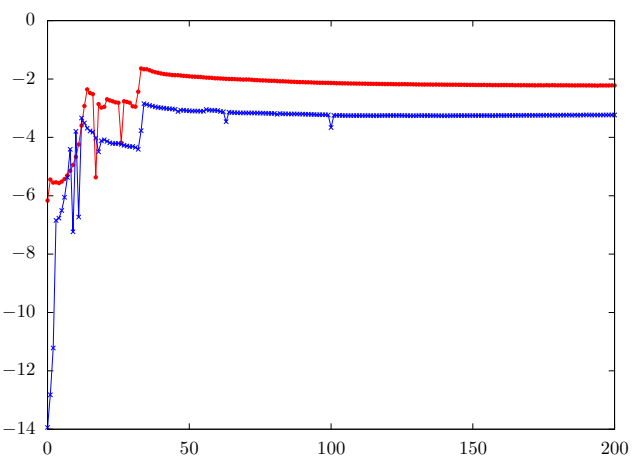


(c) Final shape for imposed displacement (6.81)

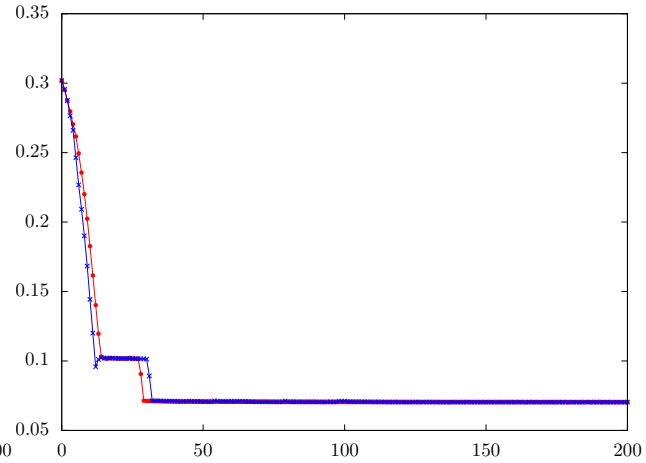


(d) Final shape for imposed displacement (6.82)

Figure 6.28: Meshes for wedge shapes obtained for minimization of (6.73), damage model (6.17), isotropic degradation (6.75)

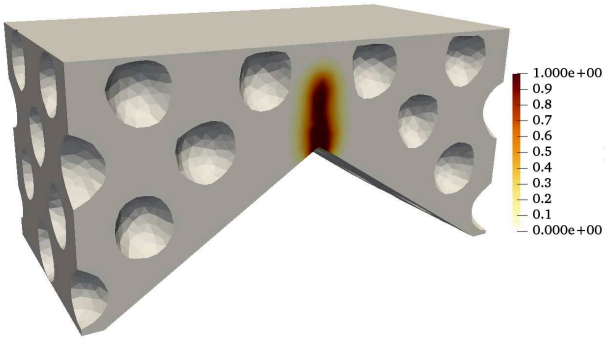


(a) Total elastic energy (6.73) vs iterations

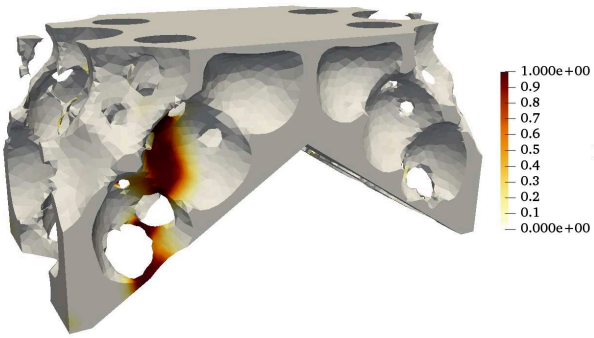


(b) Volume vs iterations

Figure 6.29: Convergence history for the shapes (6.28c)-(6.28d)

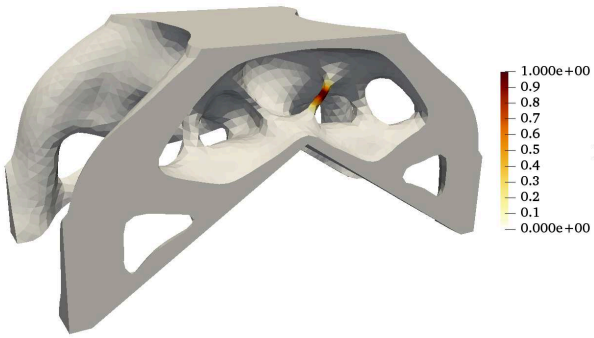


(a) Iteration 3

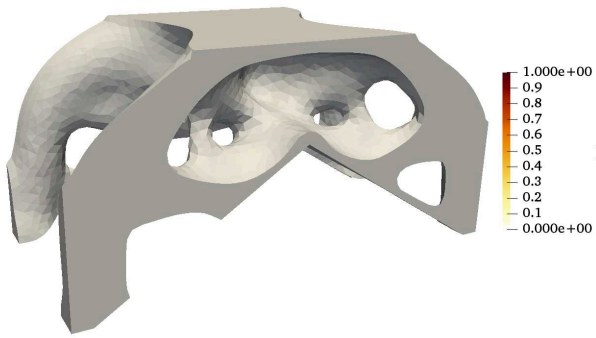


(b) Iteration 12

(c) Iteration 26



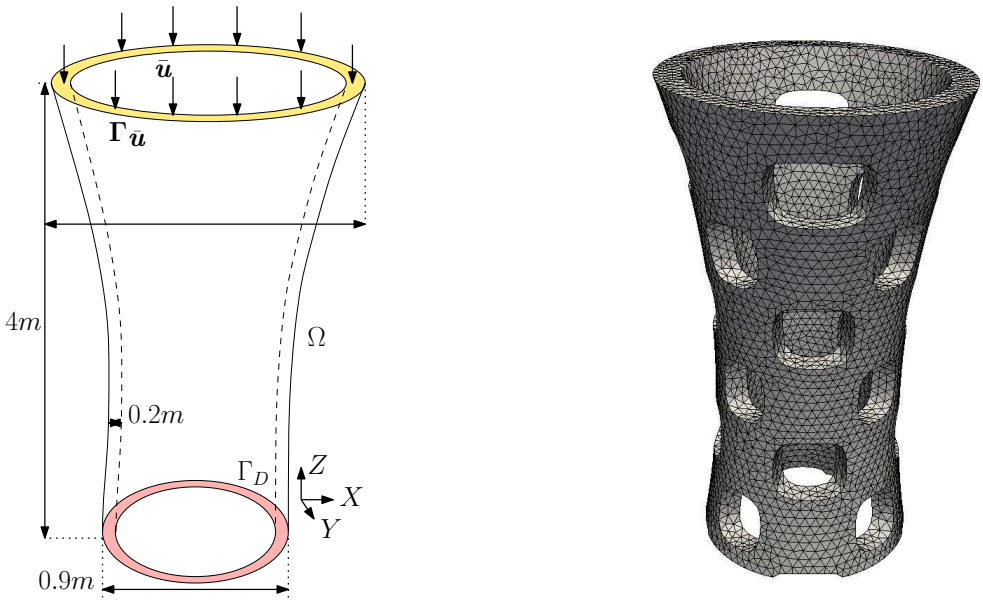
(d) Iteration 57



(e) Iteration 90

(f) Iteration 300

Figure 6.30: damage variable α at the final time plotted for several optimization iterations for the shape in Fig.6.32b

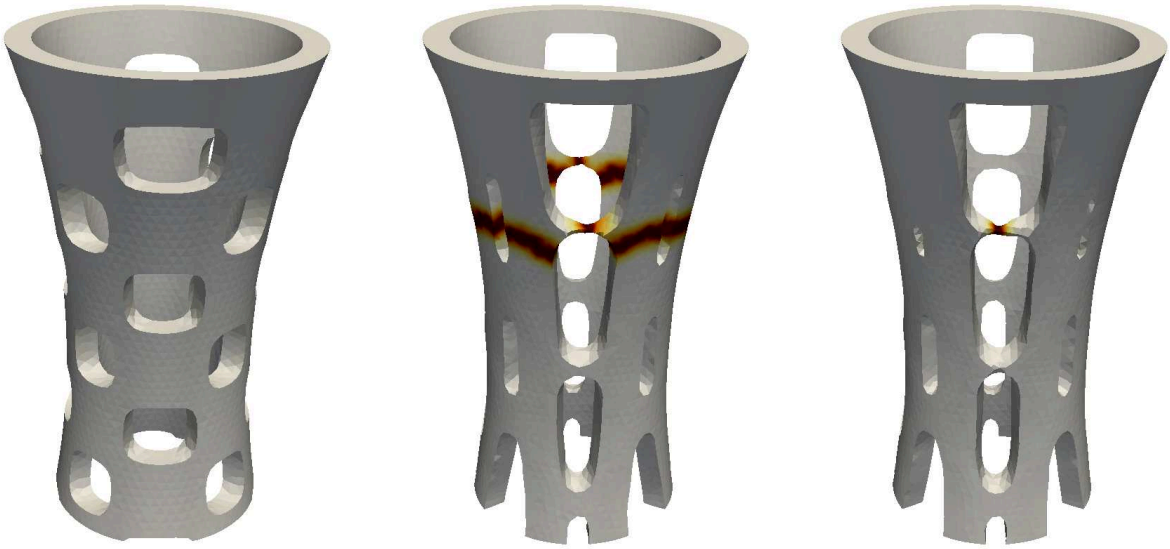


(a) (b)

Figure 6.31: 3D column: Boundary conditions (left), initial shape (right)

(a) Linear elasticity (b) Imposed displacement (6.83) (c) Imposed displacement (6.84)

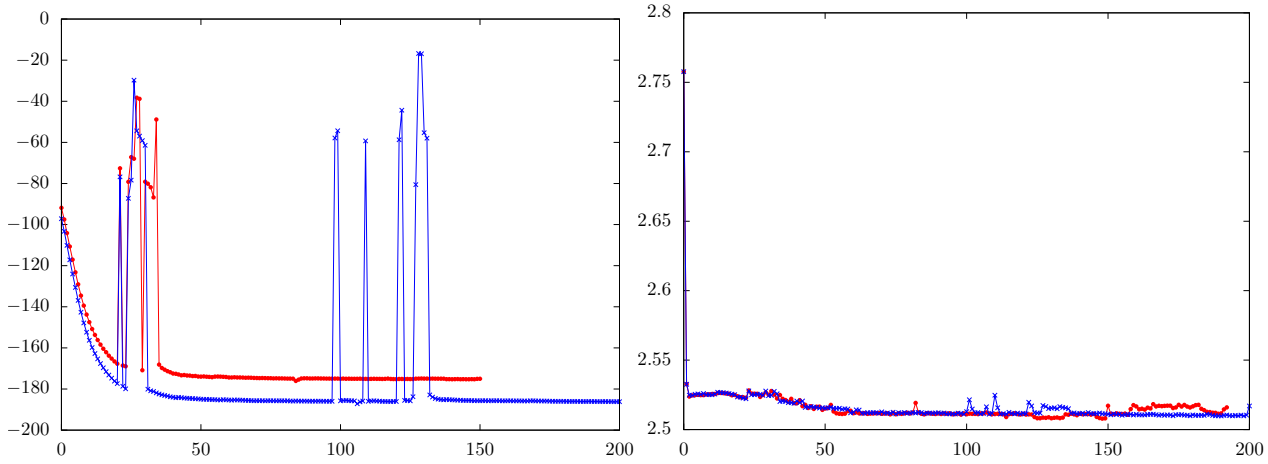
Figure 6.32: Final shape obtained for different cantilever shapes for minimization of (6.73), damage model (6.17) and isotropic degradation function (6.66), initialized from (6.31b)



(a) Iteration 0 (b) Iteration 21 (c) Iteration 23

(d) Iteration 25 (e) Iteration 60 (f) Iteration 139

Figure 6.33: damage variable α at the final time plotted for several optimization iterations for the shape in Fig.6.32b



(a) Total elastic energy (6.73) vs iterations

(b) Volume vs iterations

Figure 6.34: Convergence history for the shapes (6.32b)-(6.32c)

Figure 6.35: damage variable α at the final time computed for the optimized shapes in Fig.6.32, damage model (6.17), isotropic degradation (6.75), imposed displacement (6.84)

Chapter 7

High performance computing

Contents

This chapter is devoted to topology optimization for linear elasticity (algorithm 1) and for brittle fracture mechanics (algorithm 8) using parallel computing. This allows for a very large scale simulation, typically a structure (not considering the entire design space) meshed with $200,000$-$400,000$ vertices. We first outline the implementation details of the parallel computing in Section 7.1. Followed by Section 7.2, where we show several optimized 3D topologies. In this chapter, whenever we talk about matrices, we mean sparse matrices (arising out of the finite element discretization). Moreover, parallel computing is understood to be synonymous to high performance computing.

Using high performance computing, one can observe very smooth convergence of the objective function. When the mesh is coarse, the convergence of the objective function is slightly oscillatory, as can be seen in the previous chapters. In the case of linear elasticity this is because

S.1 The shape derivative (2.12) is computed for the continuous problem and not for the linear elasticity problem discretized using finite elements

S.2 The shape derivative (2.12) does not take into account remeshing

S.3 The descent step $d\tau$ may be chosen too big, and the first order derivative is not guaranteeing a descent direction

The above is true for brittle fracture as well, except there, additional oscillations arise from the inherent discontinuous nature of the fracture problem. When the mesh is sufficiently refined, the shape derivative for the discretized problem approaches the shape derivative (2.12) for the continuous problem. Moreover, the perturbations in the displacement vector $\boldsymbol{u}$ arising from remeshing are minimized. Along with this, if the time step is chosen small (i.e., $h_{\min}$ which is small since mesh is refined), then a very smooth convergence curve can be obtained. This is observed in Section 7.2. We do not quantify the extent of mesh refinement needed to go from an oscillatory convergence to a smooth convergence, which remains a subject to be studied. We first outline the implementation details in Section 7.1 and then present numerical results in Section 7.2 and Section 7.3. In all the sections, by the term "resolution of a matrix", we mean the resolution of the corresponding linear system using direct or iterative methods.

7.1 Implementation details

Broadly speaking, each shape optimization iteration performed in all the previous chapters encompass the following steps

L.1 Resolution of the state equation (the governing differential equation)

L.2 Resolution of the adjoint equation

L.3 Resolution of the regularization of the shape derivative

L.4 Updating the optimization parameters (Lagrange multipliers, objective function value, etc.)

L.5 Resolution of the transport equation (2.17) to evolve the shape

L.6 Remeshing the newly obtained shape

In the above, from a theoretical perspective, the computationally most expensive part is the resolution of the state and the adjoint equation. The state and the adjoint equation involves the resolution of a positive definite stiffness matrix using conjugate gradient, whose time complexity of computation is $\mathcal{O}(N^{1+1/d})$ [158] (where N is the degrees of freedom and d is the dimension of the problem). In linear elasticity, there is only one stiffness matrix to be resolved for every shape optimization iteration. Whereas for brittle fracture mechanics involves the resolution of at least a few hundred stiffness matrices (using algorithm 7) for every shape optimization iteration. Here, the adjoint fracture problem is relatively cheaper to compute, yet significantly more expensive than resolution of a single matrix. Hence, if the state and adjoint problem are implemented without parallel computing for large N (more a hundred thousand), the computation shall take several hours, and the complete shape optimization shall take a few weeks.

On the other hand, the time complexity of the regularization, advection and updating the optimization parameters is close to $\mathcal{O}(N)$. The regularization of the shape derivative is computationally cheap despite the fact that it involves the resolution of a matrix (the corresponding linear system of equations), whose size is proportional to the number of mesh vertices. This is because the regularization parameter (α in 2.23) is very small (see subsection 7.1.2 for more details). Since the remeshing algorithm in `MMG` is still not implemented using parallel computing, it remains slow.

In order to reduce the computation time, it is therefore worthwhile to implement the three steps **L.1**, **L.2** and **L.3** using parallel computing. Parallel computing can be brought about using a mutli-core CPU, a GPU or a combination of the two. A multi-core processor typically has 4 cores (8 with hyper-threading), with each core having a clock frequency 2.1GHz. In contrast, a GPU typically has 500-2000 cores; yet each core is relatively slower, clocked at a frequency of 500MHz. If one does a simple arithmetic to compare a CPU to a GPU, one GPU seems to be at least 30 times faster than a quad-core CPU. A matrix vector multiplication ought to be then about 30 times faster on a GPU than on a quad-core CPU. However, due to indirect and irregular matrix accesses, the speed-up factor of a matrix vector multiplication on a GPU is between 2 and 8 [25]. On a structured mesh (grid) however, the speedup factor is more than 10 [28]. The bottleneck of a typical parallel algorithm is communication between processors and the structure of accessing the data. The more there is communication, slower is the computation, and lesser is the scalability of the computation time with respect to the number of cores. We work with unstructured meshes, leading to irregular data access and a slower computation on GPU. Here, we do not perform any computation on GPU and instead stick to multi-processor CPU. However, we mention that GPU based topology optimization has been applied to density based methods [124], ESO (evolutionary structural optimization) [123] and to level-set based method [187].

In all the previous chapters, we had used the open-source software `FreeFEM` for all of our numerical experiments. For parallel-computing, we prefer not to re-code everything on a new software because `FreeFEM` comes with built-in OpenMPI and the open source package `PETSc` [1]. The `PETSc` library [1] has been integrated into `FreeFEM` by P. Jolivet[100, 68, 99]. This integration allows to use the `FreeFEM` tools, and finite element construction of `FreeFEM` in conjunction with PETSc solvers and its preconditioners. All of our numerical experiments in this chapter are performed using `FreeFEM 4.8` [92], installed on a workstation that has Intel(R) Xeon(R) Gold 6230 CPU and 40 processors, clocked at a frequency of 2.10GHz.

For the detail on the usage of the remeshing tool `MMG`, see Subsection 2.2.3. We modify our existing `FreeFEM` script meant for serial computing by simply modifying the part of our code corresponding to the three steps **L.1**, **L.2** and **L.3**. A `FreeFEM` script is typically launched on command-line. For the parallel

computation of the first three steps, we execute another FreeFEM script inside the existing FreeFEM script. The rough sketch of the modified code is given below.

Listing 7.1: Sketch of FreeFEM code for shape optimization

```
//shape optimization loop
for(int i=0; i<niter; i++)
{
  // Script to solve the state equation
  exec( "mpirun -np 4 FreeFem++-mpi  parallel_solver.edp" );

  //Script to solve the adjoint equation
  exec( "mpirun -np 4 FreeFem++-mpi parallel_adjoint.edp" );

  //Script to find and regularize the shape derivative
  exec( "mpirun -np 4 FreeFem++-mpi parallel_regularization.edp" );

  // Then update the optimization parameters
  ....
  // And finally performing advect
  exec( "advect -nocfl ..." );

  // Read the new levelset function and remesh using MMG
  exec( "mmg3d_O3 ..." );

  //To find the signed distance function of the new shape
  exec( "mshdist ..." );
}
```

In the next subsection, we give a few details about the numerical resolution of matrices using HPC (high performance computing) on FreeFEM.

In case of shape optimization for linear elasticity, using parallel computing for the state and adjoint equation, the remeshing step ends up being the slowest, consuming up to 90% of the total computation time (for example, see Table 7.3). A parallel version of MMG, called ParMMG is under development. Once ParMMG is released, the remeshing step **L.6** can be as fast as the other steps. In case of shape optimization for brittle fracture, the time taken for remeshing is much lesser compared to the resolution of the state and the adjoint equation, which are significantly more computationally expensive.

7.1.1 Resolution of the state and adjoint equation using parallel computing

Evidently, the resolution of state and adjoint equation in linear elasticity involves only one positive definite matrix. The resolution of the state equation for brittle fracture (using the algorithm 7) involves positive definite matrices, whereas the resolution of the adjoint equation for brittle fracture (using (4.107)) involves indefinite matrices.

Resolution of these matrices can be brought about using direct and iterative solvers. It is well-known that an iterative solver needs a preconditioner. Iterative solvers are powerful because they can be easily implemented using parallel computing. In contrast, direct solvers are by nature, not meant for parallel computing. There are however, multi-frontal approaches that facilitate a direct solver using a parallel computing (see MUMPS [19]).

In case of the ersatz approach (see Subsection 2.2.1), the presence of residual stiffness (measured using ϵ, see Def.(2.20)) contributes to ill-conditioning of the stiffness matrix of the state and the adjoint equation. This along with a higher degrees of freedom (compared to remeshing approach, due to the presence of ersatz material) results in a high computation for an iterative solver. In contrast, in body-fitted remeshing approach, the degrees of freedom are lower and the stiffness matrix is better conditioned (due to complete absence of ersatz). Preconditioning is still nevertheless needed as the condition number of the matrix is $\mathcal{O}(N^{2/d})$ [158].

PETSc has several direct and iterative solvers. Direct solvers involving factorizing approaches, like Cholesky, LU are available on PETSc. Incomplete-factorization approaches can be used as preconditioners. Iterative solvers like the Krylov subspace algorithms (like Conjugate gradient and GMRES) are available. In addition, PETSc has many preconditioners for iterative solvers like Jacobi, Additive Schwartz (ASM), Balancing Domain Decomposition (BDD), Geometric Algebraic Multigrid (GAMG), incomplete Cholesky

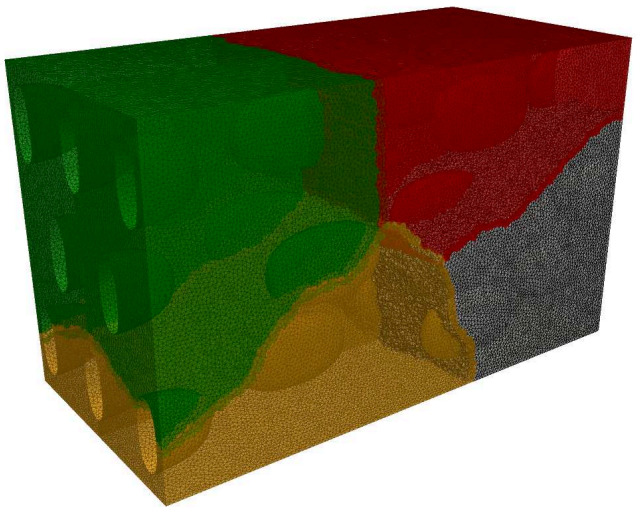


Figure 7.1: Mesh partitioning of a cantilever (having 3 million tetrahedra) using `METIS` on $np = 4$ processors

(ICC), incomplete LU to name a few [23]. For many preconditioners (like ASM, BDD and GAMG), it is imperative to perform a domain decomposition, namely, splitting the mesh in np subdomains, where np is the number of processors used. This can be brought about on `FreeFEM` by using SCOTCH or METIS [102]. Fig.7.1 illustrates an example of mesh decomposition into 4 subdomains using METIS. For all the matricial resolutions (the state and the adjoint) involving positive definite matrices, we stick to the conjugate gradient (CG) algorithm, and for the resolution (6.67) involving indefinite matrices, we use `GMRES`. In order to determine the best preconditioner (resulting in the fasting matrix resolution), we perform linear elasticity computation on the 3D cantilever-beam (shown in Fig.7.1 for boundary conditions (7.2a)) for different preconditioners and compare the computation time in table 7.1. As can be seen in 7.1, GAMG outperforms other preconditioners. Hence we choose GAMG preconditioner for CG and GMRES in our test cases. Now,

Preconditioner	Time
Jacobi	1206.51
ASM	734.23
ILU	61.6
GAMG	40.47

Table 7.1: Time taken in seconds with CG solver on 4 processors

we present a `FreeFEM` code in (7.2) that can solve linear elasticity using GAMG-CG solver on PETSc. We do so in order to demonstrate the ease with which one can write a parallel computing code on `FreeFEM`.

Listing 7.2: `FreeFEM` code using parallel computing

```
load "PETSc"   // PETSc plugin
macro dimension()2 //dimension of the problem
include "macro_ddm.idp"//domain decomposition tools

real lambda = .., mu = ..;//lame coefficients

mesh Th;
...// code to define the mesh Th

mesh Thglob = Th;//storing the global mesh

int[int] n2o;
macro ThN2O()n2o// EOM
buildDmesh(Th); //decompose and distribute mesh Th to each processor
// Th is now parallel!
```

```
// n2o is correspondence of local to global mesh Th

Mat A;//matrix of PETSc type, global matrix

macro def(u)[u, u#b]// displacement vector field definition
func Pk = [P1, P1];//finite element space of  displacement vector

macro init(i)[i, i])// EOM        // vector field initialization
createMat(Th, A, Pk);//A partitioned!
// A is now decomposed and sent to each processor

//parallel finite element space
fespace Vh(Th, Pk);

macro div(u) (dx(u) + dy(u#b))//EOM
macro espilon(u) [ dx(u), dy(u#b), 1/sqrt(2.)*( dy(u) + dx(u#B) )  ]//EOM
varf linelas(def(u), def(v)) = intV(Th)(lambda * div(u) * div(v)
      + 2.0 * mu * (epsilon(u)` * epsilon(v)))
      + intS(Th, neumann)( G`*def(v))
      + on(dirichlet, u = 0, u#B = 0);
A = linelas(Vh, Vh); // parallel matrix assembly
real[int] b = linelas(0, Sh); //right hand side

set(A, sparams = "-pc_type gamg  -ksp_max_it 1000 -ksp_rtol 1.e-10", bs = 2);

Vh def(u); //local displacement solution
u[] = A^-1 * b;//matrix resolution
```

As seen in the above code, the linear elasticity resolution is fairly simple to implement. The adjoint solver for linear elasticity can be implemented in a similar manner. For brittle fracture mechanics using algorithm 7, the code is however more complicated as it involves a transient scheme involving a Newton-Fixed point algorithm coupled to back-tracking.

7.1.2 Regularization implementation on FreeFEM

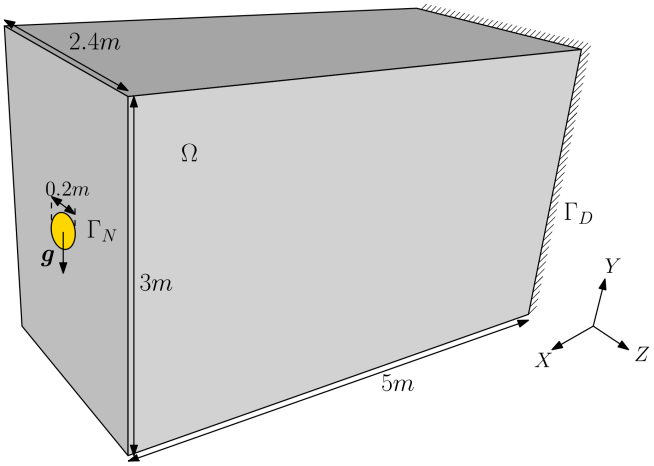
The previous section gives an idea about the resolution of the state and the adjoint equation. Once we have the solution to the two equations, we can determine the shape derivatives. The shape derivative in (2.12) and in (6.38) are defined only on the boundary $\partial\Omega$. In order to perform the regularization (2.23) and (6.69), the velocity $j'(\Omega)$ (along the shape derivative) ought to be extended by zero to the rest of the design space D.

Using the extended shape derivative, the regularization Equations (2.23) and (6.69) can be solved. Since the mesh is highly refined, it seems that the numerical solution of the regularization problem is time-consuming. This is however not the case. For the refined 3D cantilever beam (shown in Fig.7.1), we perform regularization using (2.23) for the compliance criterion. The solver is CG and there is no preconditioning. The number of iterations taken for different regularization parameter α is given in table 7.2. Surprizingly,

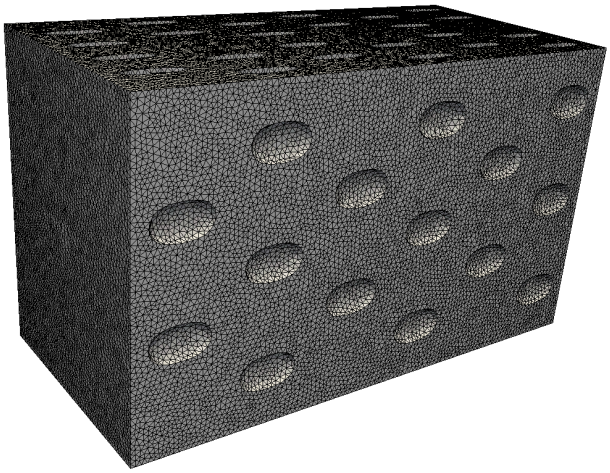
α	Iterations
$h_{\min}$	7
$2h_{\min}$	11
$4h_{\min}$	20
$8h_{\min}$	40

Table 7.2: Iterations taken with CG solver without preconditioning for the shape in Fig.2.12

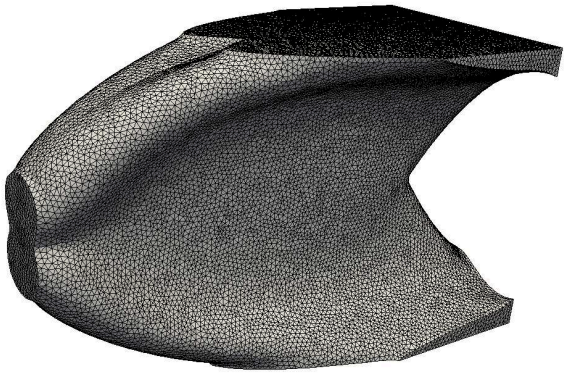
the convergence of CG takes place very fast, even without any preconditioner. As α increases, the number of iterations taken also increase. This is quite legitimate as α is measure of the diffusivity coefficient; smaller the value of α, lesser the diffusion of $dj(\Omega)$, faster the convergence of CG. Since we take $\alpha = h_{\min}$, the time -complexity of the regularization step is $\mathcal{O}(N)$, as stated earlier. We nevertheless implement this step using parallel computing.



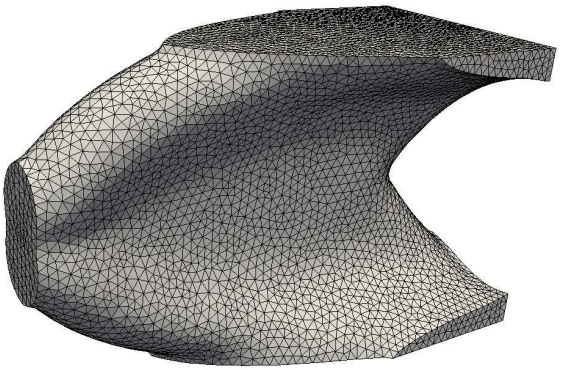
(a) Boundary conditions



(b) Initial shape mesh



(c) Final shape mesh



(d) Final shape mesh, without refinement

Figure 7.2: Cantilever optimization for compliance, $V_t = 12m^3$

7.2 3D Numerical Results: Linear Elasticity

For all test cases in this section, the compliance (2.3) is minimized. The material properties considered are $E = 210\text{GPa}$ and $\nu = 0.3$.

7.2.1 Cantilever beam

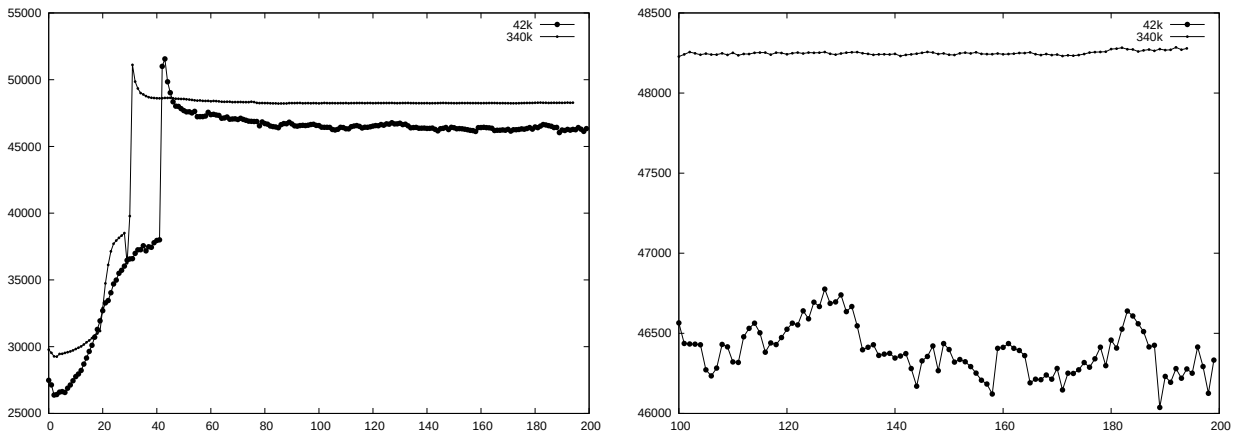
We consider a 3D cantilever beam of dimensions $5m \times 2.4m \times 3m$. As shown in the Fig.7.2a, the cantilever is fixed on the rightmost plane and forced on a circular part (zone in gold) on the right with a force $\boldsymbol{g} = [0, 10^3, 0]\text{MN/m}^2$. The target volume is taken to be $12m^3$. All the linear elasticity computations are performed for half of the domain (using symmetry). We now consider two cases

1. Coarse computation
   Remesher `MMG` considers $h_{\min} = 0.025m, h_{\max} = 0.05m$

2. Fine computation
   Remesher `MMG` considers $h_{\min} = 0.05m, h_{\max} = 0.1m$

For the fine computation, the initial shape in Fig.7.2b possessing $629,884$ vertices (or 1.8 million degrees of freedom) converges to the shape in Fig.7.2c. The final shape in Fig.7.2c has $676,068$ (or 2 million degrees of freedom). On an average, every shape in this shape optimization has about two million degrees of freedom. Despite the fact that the initial shape in Fig.7.2b has several holes, which facilitate the formation of a complex optimized topology, we find a final shape in Fig.7.2c that has very simple topology, a single plate at the center. From the mechanical standpoint, it is very well-known that the plates are better at dealing with shear stress compared to bars and rods. This test case verifies this heuristics.

For the coarse computation, the initial shape is taken to be a coarser version of the mesh in Fig.7.2b. The final shape obtained is plotted in Fig.7.2d and has $27,758$ vertices. As seen in this figure, the shape seems to be exactly the same as (7.2c) except for a few minor differences (in the thickness of the plate at the center). We then plot the convergence history for the two cases in Fig.7.3. The refined one is denoted by $340k$ and the coarse one is denoted by $42k$. As can be seen from the convergence history, the oscillations in the objective are far lesser for the refined case than for the coarse case. Towards convergence, the oscillations in the objective function arising from two factors **S.1** and **S.2** are minimized.



(a) Compliance vs iterations

(b) Compliance vs iterations, zoomed

Figure 7.3: Convergence history for the shapes (7.2c)-(7.2d)

For the fine computation, the average time taken for different operations for each shape optimization iteration is given in Table 7.3.

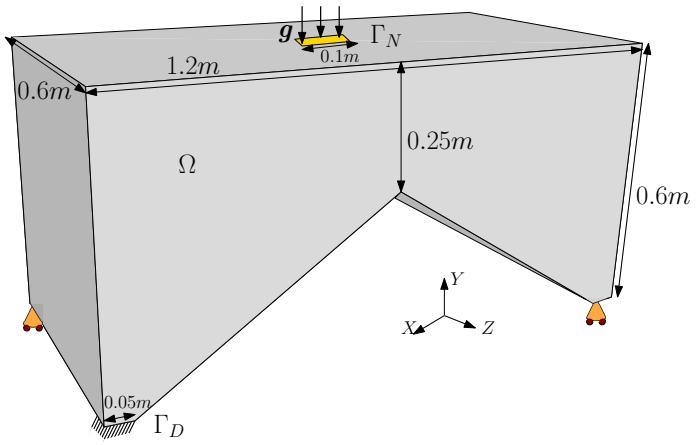
7.2.2 Wedge

We consider a 3D wedge as shown in Fig.7.4a. The wedge is supported on four square surfaces each being $0.05m \times 0.05m$, three of which can be seen in the Fig.7.4a. The wedge is clamped along all the three axes on one surface (visible in Fig.7.4a) and only along y-direction on the remaining three surfaces. The wedge

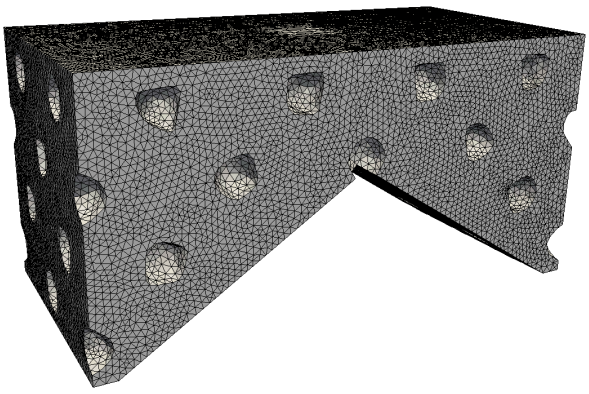
Step	Operation	Time (in seconds)
L.1	Matrix assembly	15
L.1	State equation resolution	10
L.2	Adjoint equation resolution	10
L.3	Regularization	5
L.4	Updating parameters	negligible
L.5	Transport equation	1
L.6	Remeshing	150
L.6	Redistancing	120

Table 7.3: Average time taken for different operations per iterations of the shape optimization for the cantilever beam in Fig.7.2c

is subjected to an external surface force $\boldsymbol{g} = [0, -500, 0]\text{MN/m}^2$. The remesher `MMG` considers the bounds $h_{\min} = 0.006m, h_{\max} = 0.016m$. The target volume is chosen to be $0.07m^3$. This results in the initial shape in Fig.7.4b possessing $547,194$ degrees of freedom. The compliance minimization starting from the shape in Fig.7.4b results in the shape in Fig.7.4c. The final shape in Fig.7.4c has $569,499$ degrees of freedom. Here again, we start the optimization with an initial shape in Fig.7.4c containing several holes, and end up with a final shape in Fig.7.4c with a fairly simple topology, consisting of four plates connected to the center. The convergence curve is plotted in Fig.7.5. As can be seen in the figure, the convergence of the objective is very smooth, as expected.



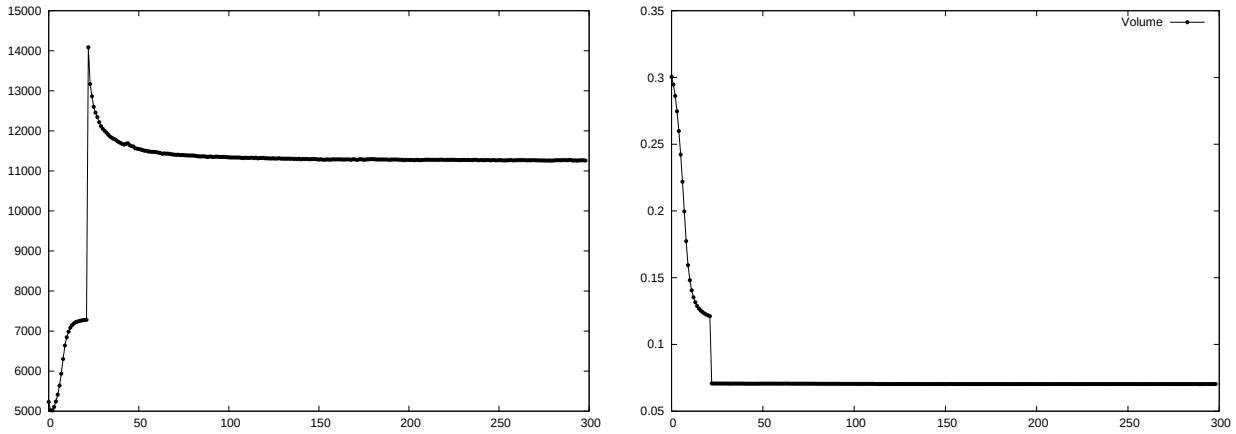
(a) Boundary conditions



(b) Initial shape mesh

(c) Final shape mesh

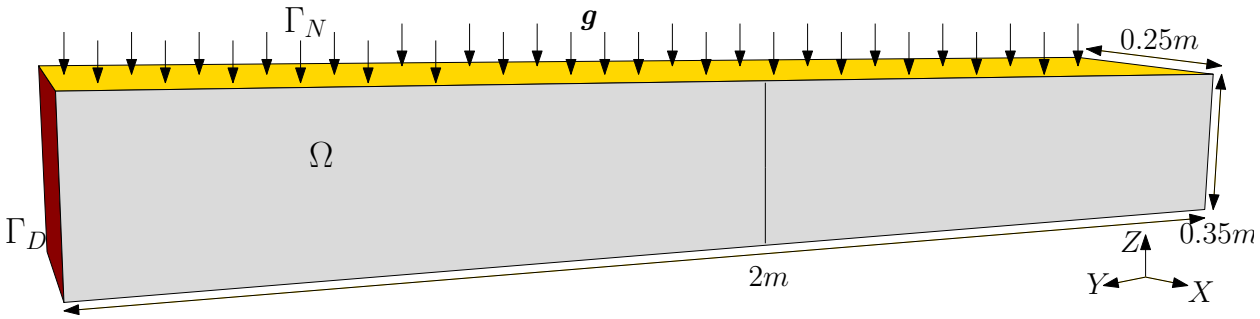
Figure 7.4: Wedge optimization for compliance



(a) Compliance vs iterations

(b) Volume vs iterations

Figure 7.5: Convergence history for the shape in Fig.7.4c



(a) Boundary conditions

7.2.3 Bridge

We now study a 3D bridge as shown in the Fig.7.4a. The bridge has dimensions $2m \times 0.35m \times 0.25m$. The bridge is fixed on the two ends (one surface seen in red in Fig.7.4a and the other cannot be seen) and forced on the topmost surface (in gold), with $\boldsymbol{g} = (0, 100, 0)\text{MN/m}^2$. All the linear elasticity computations are performed for half of the domain (using symmetry, partition can be seen at the center in Fig.7.4a). The target volume is taken to be $V_t = 0.053m^3$ (for each half). We consider two cases

1. Coarse computation
   Remesher `MMG` considers $h_{\min} = 0.02m, h_{\max} = 0.04m$

2. Fine computation
   Remesher `MMG` considers $h_{\min} = 0.04m, h_{\max} = 0.08m$

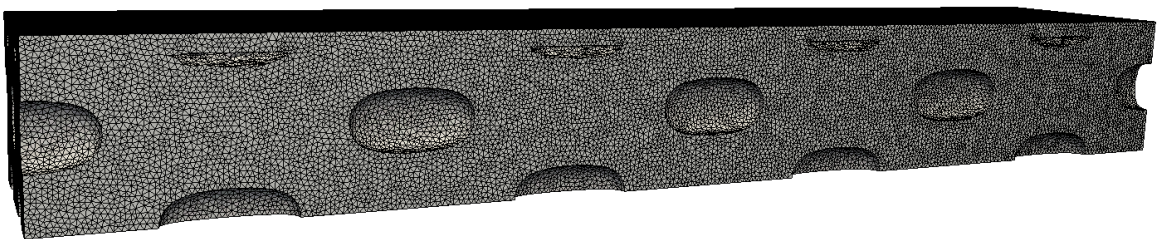
On the topmost surface Γ_N, we impose non-optimizable plate-like region of thickness $0.08m$. For the fine computation, the initial shape in Fig.7.7a converges to the shape in Fig.7.7b. The final shape in Fig.7.7b has 3.4 million degrees of freedom. For the coarse computation, we obtain the shape in Fig.7.7c.

7.2.4 Column

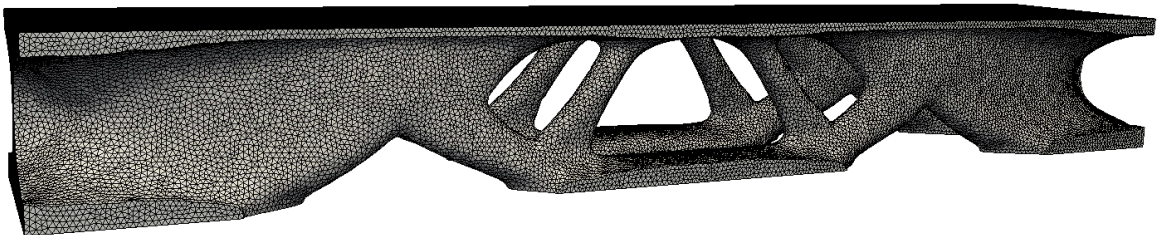
We revisit the same column, we considered in the Chapter 6, subsection 6.5.7, but with a highly refined mesh (which can be found in [78]). The target volume is chosen to be $2.5m^3$. The column is fixed on the bottom. For the external loading and objective function, we consider two cases:

C.1 External applied force $g = (0, 0, 100)\text{MN/m}^2$ and minimization for compliance; boundary conditions are shown in Fig.7.8a.

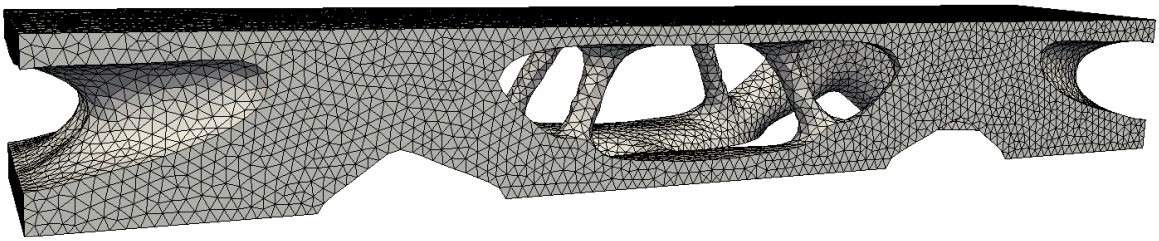
C.2 External imposed displacement $\bar{\boldsymbol{u}} = (0, 0, 10^{-3})\text{m}$ and elastic energy maximization; boundary conditions are shown in Fig.6.31a.



(a) Initial shape



(b) Final shape, fine mesh



(c) Final shape, coarse mesh

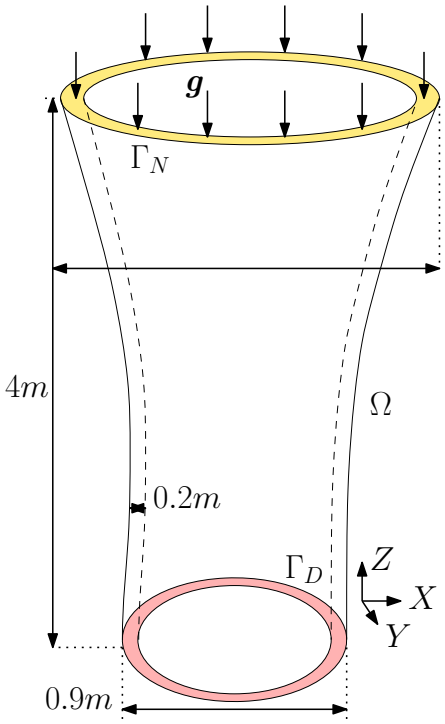
Figure 7.7: Minimization for compliance, fine mesh

As discussed in the beginning of Section 6.5, in case of an imposed displacement, the compliance can be defined using the total energy. Hence in the case **C.2**, we maximize the elastic energy (which is equivalent to minimizing compliance).

For the two cases, the shape is initialized as shown in Fig.7.8b and the final shapes obtained are plotted in Fig.7.8c-7.8d. The convergence curve for shape in Fig.7.8c is plotted in Fig.7.9. As can be seen in Fig.7.8c-7.8d, we make two important observations. First, there are slight differences in the thickness of the vertical bars. This is surprising as in the two cases, we minimize the compliance. Second, the column in Fig.7.8d is very different from the same column obtained with a coarse mesh (in Fig.7.8c). This is unlike in the case of the 3D cantilever beam (see subsection 7.2.1), where identical shapes were obtained with and without refinement. We suspect that this is because the 3D cantilever beam is a bulky structure, compared to the 3D column, which is a relatively thin structure. The displacement solution to a thin structure is more dependent on the mesh refinement, and thus refining the mesh plays an important role and leads to a much different (possibly better) optimal structure.

7.3 3D Numerical Results: Brittle Fracture

The goal of this example is to show that our optimization approach for fracture minimization is amenable to high performance computing (HPC) for the 3D column (treated in the previous subsection). This combined to the fact that we use true material parameters and true-to-scale dimensions, makes it the most important test case of this chapter. The column is subjected to the same imposed displacement (6.84) as shown in Fig.6.31a. The damage model `DL` (6.17) and isotropic degradation function (6.66) are considered. The input parameters of the remesher `MMG` are $h_{\min} = 0.015m$ and $h_{\max} = 0.03m$. It implies that the initial mesh has 243,641 vertices (1,359,805 tetrahedra) and the final mesh has 241,852 vertices (1,365,125 tetrahedra). The characteristic length, given by (6.62), is $\ell = 0.0375m$, satisfying condition (6.64). The target volume is again



(a) Boundary conditions

(b) Initial shape

(c) Final shape optimized for case **C.1**

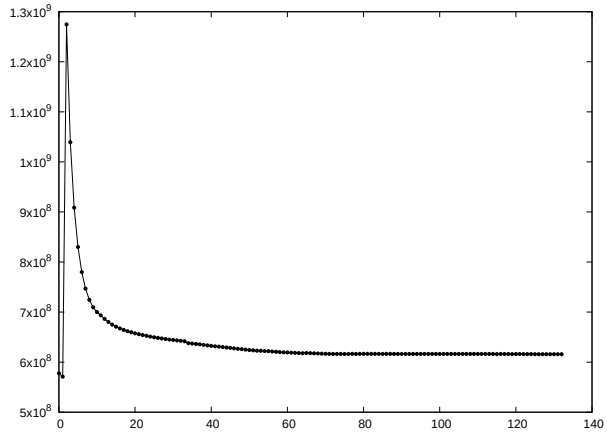
(d) Final shape optimized for case **C.2**

Figure 7.8: 3D column optimization

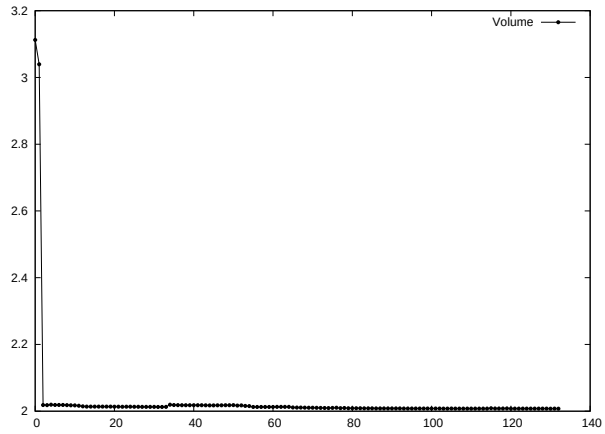
chosen to be $2.5m^3$. An incremental residual strategy is used, where the residual stiffness is defined with (6.76), and the residual stiffness parameter is defined as

$$\kappa = \frac{1}{2}\left(1 - \frac{i}{\mathcal{N}}\right)^8,$$

where i is the iteration number and $\mathcal{N}$ is the total number of iterations. This formula differs from (6.77) by a factor of 5, heuristically found to yield a smoother convergence. This incremental approach is preferred in



(a) Compliance vs iterations



(b) Compliance vs iterations, zoomed

Figure 7.9: Convergence history for the shape in Fig.7.8c

(a) Initialization

(b) Final shape obtained for damage model

Figure 7.10: Column shapes for minimization of (6.73), damage model (6.17), isotropic degradation function (6.66), imposed displacement (6.84)

order to ensure that the intermediate shapes do not undergo brittle fracture too often. The mesh is so fine that the time or descent step τ, given by (2.28), is very small and the convergence is too slow. Therefore, for this test case, we multiply it by a factor of 2, namely $\tau = h_{\min}/C_i$. The initial shape (with a rich topology) can be seen in Fig. 7.10a. The final shape obtained for linear elasticity is plotted in Fig. 7.8d and, for the damage model, is plotted in Fig. 7.10b (a snapshot of the mesh is given in Fig.7.12). The damage variable α is plotted for some intermediate shapes in Fig.7.11. Actually, there were only 3 intermediate shapes, which were cracked. Our incremental approach of the residual stiffness was thus not able to eliminate completely the appearance of cracks during the optimization process. The final shape obtained is very similar to the

one obtained for linear elasticity in Fig.7.8d. The total computational time for this optimization was 7 days.

(a) Iteration 15 (b) Iteration 17 (c) Iteration 18

(d) Iteration 43 (e) Iteration 83 (f) Iteration 200

Figure 7.11: Damage parameter α at the last time instant plotted for several optimization iterations for the shape in Fig.7.10b

Figure 7.12: Mesh of the shape (7.10b)

Bibliography

[1] S. Abhyankar, J. Brown, E. M. Constantinescu, D. Ghosh, B. F. Smith, and H. Zhang. Petsc/ts: A modern scalable ode/dae solver library. *arXiv preprint arXiv:1806.01437*, 2018.

[2] G. Allaire. Conception optimale de structures. *Mathematiques and Applications*, 58:280, 2007.

[3] G. Allaire. *Numerical analysis and optimization: an introduction to mathematical modelling and numerical simulation*. Oxford university press, 2007.

[4] G. Allaire. *Shape optimization by the homogenization method*, volume 146. Springer Science & Business Media, 2012.

[5] G. Allaire and C. Dapogny. A linearized approach to worst-case design in parametric and geometric shape optimization. *Mathematical Models and Methods in Applied Sciences*, 24(11):2199–2257, 2014.

[6] G. Allaire, C. Dapogny, and P. Frey. Topology and geometry optimization of elastic structures by exact deformation of simplicial mesh. *Comptes Rendus Mathematique*, 349(17-18):999–1003, 2011.

[7] G. Allaire, C. Dapogny, and P. Frey. Shape optimization with a level set based mesh evolution method. *Computer Methods in Applied Mechanics and Engineering*, 282:22–53, 2014.

[8] G. Allaire, C. Dapogny, and F. Jouve. Shape and topology optimization. *Geometric Partial Differential Equations, part II, A. Bonito and R. Nochetto eds., Handbook of Numerical Analysis*, 22:1–132, 2020.

[9] G. Allaire and G. Delgado. Stacking sequence and shape optimization of laminated composite plates via a level-set method. *Journal of the Mechanics and Physics of Solids*, 97:168–196, 2016.

[10] G. Allaire and A. Henrot. On some recent advances in shape optimization. *Comptes Rendus de l'Académie des Sciences-Series IIB-Mechanics*, 329(5):383–396, 2001.

[11] G. Allaire, F. Jouve, and G. Michailidis. Molding direction constraints in structural optimization via a level-set method. *Variational Analysis and Aerospace Engineering*, pages 1–39, 2016.

[12] G. Allaire, F. Jouve, and A.M. Toader. A level-set method for shape optimization. *Comptes Rendus Mathematique*, 334(12):1125–1130, 2002.

[13] G. Allaire, F. Jouve, and A.M. Toader. Structural optimization using sensitivity analysis and a level-set method. *Journal of Computational Physics*, 194(1):363–393, 2004.

[14] G. Allaire and O. Pantz. Structural optimization with F*reefem++*. *Structural and Multidisciplinary Optimization*, 32(3):173–181, 2006.

[15] S. Almi and U. Stefanelli. Topology optimization for quasistatic elastoplasticity, 2021.

[16] M. Ambati, T. Gerasimov, and L. De Lorenzis. A review on phase-field models of brittle fracture and a new fast hybrid formulation. *Computational Mechanics*, 55(2):383–405, 2015.

[17] L. Ambrosio and V.M. Tortorelli. Approximation of functional depending on jumps by elliptic functional via t-convergence. *Communications on Pure and Applied Mathematics*, 43(8):999–1036, 1990.

[18] L. Ambrosio and V.M. Tortorelli. On the approximation of free discontinuity problems. *Bollettino dell'unione matematica italiana*, page 105–123, 1992.

[19] P. R Amestoy, I. S Duff, J.Y. L'Excellent, and J. Koster. MUMPS: a general purpose distributed memory sparse solver. In *International Workshop on Applied Parallel Computing*, pages 121–130. Springer, 2000.

[20] O. Amir. A topology optimization procedure for reinforced concrete structures. *Computers & Structures*, 114:46–58, 2013.

[21] H. Amor, J.J. Marigo, and C. Maurini. Regularized formulation of the variational brittle fracture with unilateral contact: Numerical experiments. *Journal of the Mechanics and Physics of Solids*, 57:1209–1229, 2009.

[22] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, A. Dener, V. Eijkhout, W. Gropp, D. Karpeyev, D. Kaushik, M. Knepley, D. May, L. Curfman McInnes, R. Mills, T. Munson, K. Rupp, P. Sanan, B. Smith, S. Zampini, H. Zhang, and H. Zhang. PETSc users manual. Technical report, 2019.

[23] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith. Efficient management of parallelism in object oriented numerical software libraries. In E. Arge, A. M. Bruaset, and H. P. Langtangen, editors, *Modern Software Tools in Scientific Computing*, pages 163–202. Birkhauser Press, 1997.

[24] C. Barbarosie and S. Lopes. A generalized notion of compliance. *Comptes Rendus Mécanique*, 339(10):641–648, 2011.

[25] M. M. Baskaran and R. Bordawekar. Optimizing sparse matrix-vector multiplication on GPUs. *IBM Research Report RC24704*, (W0812–047), 2009.

[26] P. Bažant and G. Pijaudier-Cabot. Measurement of characteristic length of nonlocal continuum. *Journal of Engineering Mechanics*, 115(4):755–767, 1989.

[27] F. B. Belgacem. The mortar finite element method with lagrange multipliers. *Numerische Mathematik*, 84(2):173–197, 1999.

[28] N. Bell and M. Garland. Implementing sparse matrix-vector multiplication on throughput-oriented processors. In *Proceedings of the conference on high performance computing networking, storage and analysis*, pages 1–11, 2009.

[29] A. Benallal and J.J. Marigo. Bifurcation and stability issues in gradient theories with softening. *Modelling and Simulation in Materials Science and Engineering*, 15(1):S283, 2006.

[30] M. P. Bendsøe and N. Kikuchi. Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering*, 71(2):197–224, 1988.

[31] M. P. Bendsøe and O. Sigmund. Material interpolation schemes in topology optimization. *Archive of Applied Mechanics*, 69(9):635–654, 1999.

[32] M. P. Bendsoe and O. Sigmund. *Topology optimization: theory, methods, and applications*. Springer-Verlag Berlin Heidelberg, 2 edition, 2004.

[33] M. Bogomolny and O. Amir. Conceptual design of reinforced concrete structures using topology optimization with elastoplastic material modeling. *International Journal for Numerical Methods in Engineering*, 90(13):1578–1597, 2012.

[34] B. Bourdin. Numerical implementation of the variational formulation for quasi-static brittle fracture. *Interfaces and free boundaries*, 9(3):411–430, 2007.

[35] B. Bourdin, G.A. Francfort, and J.J. Marigo. The variational approach to fracture. *Journal of Elasticity*, 91(1-3):5–148, 2008.

[36] A. Braides. *Approximation of free-discontinuity problems*, volume 1964. Lecture notes in mathematics. Springer-Verlag.

[37] T. E. Bruns and D.A. Tortorelli. Topology optimization of non-linear elastic structures and compliant mechanisms. *Computer Methods in Applied Mechanics and Engineering*, 190(26-27):3443–3459, 2001.

[38] T. Buhl, C. Pedersen, and O. Sigmund. Stiffness design of geometrically nonlinear structures using topology optimization. *Structural and Multidisciplinary Optimization*, 19(2):93–104, 2000.

[39] C. Bui, C. Dapogny, and P. Frey. An accurate anisotropic adaptation method for solving the level set advection equation. *International Journal for Numerical Methods in Fluids*, 70(7):899–922, 2012.

[40] M. Burger. A framework for the construction of level set methods for shape optimization and reconstruction. *Interfaces and Free boundaries*, 5(3):301–329, 2003.

[41] E. Burman, S. Claus, P. Hansbo, M.G. Larson, and A. Massing. CutFEM: discretizing geometry and partial differential equations. *International Journal for Numerical Methods in Engineering*, 104(7):472–501, 2015.

[42] C. Carstensen, O. Scherf, and P. Wriggers. Adaptive finite elements for elastic bodies in contact. *SIAM Journal on Scientific Computing*, 20(5):1605–1626, 1999.

[43] J.L. Chaboche. Continuum damage mechanics: Part I—General concepts. *ASME*, 1988.

[44] A. Chambolle. An approximation result for special functions with bounded deformation. *Journal de mathématiques pures et appliquées*, 83(7):929–954, 2004.

[45] A. Chambolle, G.A. Francfort, and J.J. Marigo. When and how do cracks propagate? *Journal of the Mechanics and Physics of Solids*, 57(9):1614–1622, 2009.

[46] A. Chambolle, G.A. Francfort, and J.J. Marigo. Revisiting energy release rates in brittle fracture. *Journal of Nonlinear Science*, 20(4):395–424, 2010.

[47] A. Chambolle, A. Giacomini, and M. Ponsiglione. Crack initiation in brittle materials. *Arch. Ration. Mech. Anal.*, 188(2):309–349, 2008.

[48] D. Chenais. On the existence of a solution in a domain identification problem. *Journal of Mathematical Analysis and Applications*, 52(2):189–219, 1975.

[49] F. Chouly and P. Hild. A Nitsche-based method for unilateral contact problems: numerical analysis. *SIAM Journal on Numerical Analysis*, 51(2):1295–1307, 2013.

[50] B.D. Coleman and M.E. Gurtin. Thermodynamics with internal state variables. *The Journal of Chemical Physics*, 47(2):597–613, 1967.

[51] C. Comi and U. Perego. Fracture energy based bi-dissipative damage model for concrete. *International Journal of Solids and Structures*, 38(36-37):6427–6454, 2001.

[52] G. Costa, M. Montemurro, and J. Pailhès. NURBS hyper-surfaces for 3D topology optimization problems. *Mechanics of Advanced Materials and Structures*, pages 1–20, 2019.

[53] J. Céa. Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût. *ESAIM: Mathematical Modelling and Numerical Analysis*, 20(3):371–402, 1986.

[54] D. Da and J. Yvonnet. Topology optimization for maximizing the fracture resistance of periodic quasi-brittle composites structures. *Materials*, 13(15):3279, 2020.

[55] P. R. Dahl. A solid friction model. Technical report, Aerospace Corp El Segundo Ca, 1968.

[56] M. Dambrine, C. Dapogny, and H. Harbrecht. Shape optimization for quadratic functionals and states with random right-hand sides. *SIAM Journal on Control and Optimization*, 53(5):3081–3103, 2015.

[57] M. Dambrine and D. Kateb. On the ersatz material approximation in level-set methods. *ESAIM: Control, Optimisation and Calculus of Variations*, 16(3):618–634, 2010.

[58] M. Dambrine and J. Lamboley. Stability in shape optimization with second variation. *Journal of Differential Equations*, 267(5):3009–3045, 2019.

[59] M. Dambrine and A. Laurain. A first order approach for worst-case shape optimization of the compliance for a mixture in the low contrast regime. *Structural and Multidisciplinary Optimization*, 54(2):215–231, 2016.

[60] C. Dapogny. *Shape optimization, level set methods on unstructured meshes and mesh evolution*. PhD thesis, Paris 6, 2013.

[61] C. Dapogny, C. Dobrzynski, and P. Frey. Three-dimensional adaptive domain remeshing, implicit domain meshing, and applications to free and moving boundary problems. *Journal of Computational Physics*, 262:358–378, 2014.

[62] C. Dapogny and P. Frey. Computation of the signed distance function to a discrete contour on adapted triangulation. *Calcolo*, 49(3):193–219, 2012.

[63] E. De Giorgi. Sulla differenziabilita e l'analiticita delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino*, 3:25–43, 1957.

[64] F. De Gournay. Velocity extension for the level-set method and multiple eigenvalues in shape optimization. *SIAM Journal on Control and Optimization*, 45(1):343–367, 2006.

[65] F. De Gournay, G. Allaire, and F. Jouve. Shape and topology optimization of the robust compliance via the level set method. *ESAIM: Control, Optimisation and Calculus of Variations*, 14(1):43–70, 2008.

[66] J. Desai, G. Allaire, and F. Jouve. Topology optimization of structures undergoing brittle fracture. preprint on https://hal.archives-ouvertes.fr/hal-03277281/, 2021.

[67] J. Desai, G. Allaire, F. Jouve, and C. Mang. Topology optimization in quasi-static plasticity with hardening using a levet-set method. preprint on https://hal.archives-ouvertes.fr/hal-03259408/, 2021.

[68] V. Dolean, P. Jolivet, and F. Nataf. *An introduction to domain decomposition methods. Algorithms, theory, and parallel implementation*, volume 144. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2015.

[69] G. Duvant and J.L. Lions. *Inequalities in mechanics and physics*, volume 219. Springer Science & Business Media, 2012.

[70] P. Duysinx, L. Van Miegroet, T. Jacobs, and C. Fleury. Generalized shape optimization using x-fem and level set methods. In *IUTAM Symposium on Topological Design Optimization of Structures, Machines and Materials*, pages 23–32. Springer, 2006.

[71] C. Eck and J. Jarusek. Existence results for the static contact problem with Coulomb friction. *Mathematical Models and Methods in Applied Sciences*, 8(03):445–468, 1998.

[72] C. Eck, J. Jarusek, and M. Krbec. *Unilateral Contact Problems: Variational Methods and Existence Theorems*. Chapman & Hall/CRC Pure and Applied Mathematics. CRC Press, 2005.

[73] F. Feppon, G. Allaire, F. Bordeu, J. Cortial, and C. Dapogny. Shape optimization of a coupled thermal fluid–structure problem in a level set mesh evolution framework. *SeMA Journal*, 76(3):413–458, 2019.

[74] F. Feppon, G. Allaire, C. Dapogny, and P. Jolivet. Topology optimization of thermal fluid–structure systems using body-fitted meshes and parallel computing. *Journal of Computational Physics*, 417:109574, 2020.

[75] F. Fernandez, M. Puso, J. Solberg, and D.A. Tortorelli. Topology optimization of multiple deformable bodies in contact with large deformations. *Computer Methods in Applied Mechanics and Engineering*, 371:113288, 2020.

[76] G. Fichera. Sul problema elastostatico di Signorini con ambigue condizioni al contorno. *Atti Accad. Naz. Lincei, VIII. Ser., Rend., Cl. Sci. Fis. Mat. Nat*, 34:138–142, 1963.

[77] G.A. Francfort and J.J. Marigo. Revisiting brittle fracture as an energy minimization problem. *Journal of the Mechanics and Physics of Solids*, 46(8):1319–1342, 1998.

[78] N. Gaudillière, R. Duballet, C. Bouyssou, A. Mallet, Ph. Roux, M. Zakeri, and J. Dirrenberger. Chapter 3 - building applications using lost formworks obtained through large-scale additive manufacturing of ultra-high-performance concrete. In J.G. Sanjayan, A. Nazari, and B. Nematollahi, editors, *3D Concrete Printing Technology*, pages 37–58. Butterworth-Heinemann, 2019.

[79] A. Giacomini. Ambrosio-Tortorelli approximation of quasi-static evolution of brittle fractures. *Calculus of Variations and Partial Differential Equations*, 22(2):129–172, 2005.

[80] R. Glowinski. *Lectures on numerical methods for non-linear variational problems*. Springer Science & Business Media, 2008.

[81] H. M. Gomes. Truss optimization with dynamic constraints using a particle swarm algorithm. *Expert Systems with Applications*, 38(1):957–968, 2011.

[82] J. E. Gordon. *Structures: or why things don't fall down*. Da Capo Press, 2009.

[83] F.E. Graves. Nuts and bolts. *Scientific American*, 250(6):136–145, 1984.

[84] A.A. Griffith. VI. the phenomena of rupture and flow in solids. *Philosophical transactions of the royal society of london. Series A, containing papers of a mathematical or physical character*, 221(582-593):163–198, 1921.

[85] M. E. Gurtin. *The Linear Theory of Elasticity*, pages 1–295. Springer Berlin Heidelberg, Berlin, Heidelberg, 1973.

[86] W. Han and B.D. Reddy. *Plasticity: Mathematical Theory and Numerical Analysis*. Number XVI in 9. Springer-Verlag New York, 2 edition, 2013.

[87] W. Han, B.D. Reddy, and G.C. Schroeder. Qualitative and numerical analysis of quasi-static problems in elastoplasticity. *SIAM Journal on Numerical Analysis*, 34(1):143–177, 1997.

[88] J. Haslinger, P. Neittaanmäki, and T. Tiihonen. Shape optimization in contact problems based on penalization of the state inequality. *Aplikace matematiky*, 31(1):54–77, 1986.

[89] J. Haslinger and J.C. Nédélec. Approximation of the Signorini problem with friction, obeying the Coulomb law. *Mathematical Methods in the Applied Sciences*, 5(1):422–437, 1983.

[90] B. Hassani and E. Hinton. A review of homogenization and topology optimization iii—topology optimization using optimality criteria. *Computers & Structures*, 69(6):739–756, 1998.

[91] B. Hassani and E. Hinton. A review of homogenization and topology optimization i—homogenization theory for media with periodic structure. *Computers & Structures*, 69(6):707–717, 1998.

[92] F. Hecht. New development in FreeFEM++. *Journal of Numerical Mathematics*, 20(3-4):251–266, 2012.

[93] A. Henrot and M. Pierre. *Shape Variation and Optimization*. EMS Tracts in Mathematics Vol. 28, 2018.

[94] X. Huang and Y. M. Xie. Bi-directional evolutionary topology optimization of continuum structures with one or multiple materials. *Computational Mechanics*, 43(3):393–401, 2009.

[95] S. Hüeber and B.I. Wohlmuth. A primal–dual active set strategy for non-linear multibody contact problems. *Computer Methods in Applied Mechanics and Engineering*, 194(27-29):3147–3166, 2005.

[96] K. A. James and H. Waisman. Topology optimization of structures under variable loading using a damage superposition approach. *International Journal for Numerical Methods in Engineering*, 101(5):375–406, 2015.

[97] K.A. James and H. Waisman. Topology optimization of viscoelastic structures using a time-dependent adjoint method. *Computer Methods in Applied Mechanics and Engineering*, 285:166–187, 2015.

[98] M. Jansen, G. Lombaert, M. Schevenels, and O. Sigmund. Topology optimization of fail-safe structures using a simplified local damage model. *Structural and Multidisciplinary Optimization*, 49(4):657–666, 2014.

[99] P. Jolivet. FreeFEM tutorial. http://jolivet.perso.enseeiht.fr/FreeFem-tutorial/, 2020.

[100] P. Jolivet, F. Hecht, F. Nataf, and C. Prud'Homme. Scalable domain decomposition preconditioners for heterogeneous elliptic problems. *Scientific Programming*, 22(2):157–171, 2014.

[101] Z. Kang, P. Liu, and M. Li. Topology optimization considering fracture mechanics behaviors at specified locations. *Structural and Multidisciplinary Optimization*, 55(5):1847–1864, 2017.

[102] G. Karypis and V. Kumar. METIS: A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices. 1997.

[103] J. Kato, H. Hoshiba, S. Takase, K. Terada, and T. Kyoya. Analytical sensitivity in topology optimization for elastoplastic composites. *Structural and Multidisciplinary Optimization*, 52(3):507–526, 2015.

[104] N. Kikuchi and J. T. Oden. *Contact problems in elasticity: A study of variational inequalities and finite element methods*, volume 8. Philadelphia, PA: SIAM, 1988.

[105] N. Kikuchi and Y. J. Song. Penalty/finite-element approximations of a class of unilateral problems in linear elasticity. *Quarterly of Applied Mathematics*, 39(1):1–22, 1981.

[106] H. S. Kim. New extruded multi-cell aluminum profile for maximum crash energy absorption and weight efficiency. *Thin-Walled Structures*, 40(4):311–327, 2002.

[107] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications*, volume 31. SIAM, 1980.

[108] A Klarbring and J Haslinger. On almost constant contact stress distributions by shape optimization. *Structural Optimization*, 5(4):213–216, 1993.

[109] A. Klarbring and N. Strömberg. Topology optimization of hyperelastic bodies including non-zero prescribed displacements. *Structural and Multidisciplinary Optimization*, 41(3):37–48, 2013.

[110] R. Kohn and R. Temam. Dual spaces of stresses and strains, with applications to Hencky plasticity. *Applied Mathematics and Optimization*, 10(1):1–35, 1983.

[111] H. Kristiansen, K. Poulios, and N. Aage. Topology optimization for compliance and contact pressure distribution in structural problems with friction. *Computer Methods in Applied Mechanics and Engineering*, 364:112915, 2020.

[112] C. J. Larsen, C. Ortner, and E. Süli. Existence of solutions to a regularized model of dynamic fracture. *Mathematical Models and Methods in Applied Sciences*, 20(07):1021–1048, 2010.

[113] M. Lawry and K. Maute. Level set topology optimization of problems with sliding contact interfaces. *Structural and Multidisciplinary Optimization*, 52(6):1107–1119, 2015.

[114] M. Lawry and K. Maute. Level set shape and topology optimization of finite strain bilateral contact problems. *International Journal for Numerical Methods in Engineering*, 113(8):1340–1369, 2018.

[115] B. S. Lazarov and O. Sigmund. Filters in topology optimization based on Helmholtz-type differential equations. *International Journal for Numerical Methods in Engineering*, 86(6):765–781, 2011.

[116] L. Li, G. Zhang, and K. Khandelwal. Design of energy dissipating elastoplastic structures under cyclic loads using topology optimization. *Structural and Multidisciplinary Optimization*, 56(2):391–412, 2017.

[117] E. Lindgaard and J. Dahl. On compliance and buckling objective functions in topology optimization of snapthrough problems. *Structural and Multidisciplinary Optimization*, 47(3):409–421, 2013.

[118] P. Liu, Y. Luo, and Z. Kang. Multi-material topology optimization considering interface behavior via XFEM and level set method. *Computer Methods in Applied Mechanics and Engineering*, 308:113–133, 2016.

[119] D.G. Luenberger and Y. Ye. *Linear and nonlinear programming*, volume 228. New York, NY: Springer, 2016.

[120] N. D. Mankame and G.K. Ananthasuresh. Topology optimization for synthesis of contact-aided compliant mechanisms using regularized contact modeling. *Computers & structures*, 82(15-16):1267–1290, 2004.

[121] J.J. Marigo. Constitutive relations in plasticity, damage and fracture mechanics based on a work property. *Nuclear Engineering and Design*, 114(3):249–272, 1989.

[122] J.J. Marigo. From Clausius-Duhem and Drucker-Ilyushin inequalities to standard materials. *Continuum Thermomechanics*, pages 289–300, 2000.

[123] J. Martínez-Frutos and D. Herrero-Pérez. GPU acceleration for evolutionary topology optimization of continuum structures using isosurfaces. *Computers & Structures*, 182:119–136, 2017.

[124] J. Martínez-Frutos, P. J. Martínez-Castejón, and D. Herrero-Pérez. Efficient topology optimization using GPU computing with multilevel granularity. *Advances in Engineering Software*, 106:47–62, 2017.

[125] J. Martínez-Frutos, G. Allaire, C. Dapogny, and F. Periago. Structural optimization under internal porosity constraints using topological derivatives. *Computer Methods in Applied Mechanics and Engineering*, 345:1–25, 2019.

[126] J. Martínez-Frutos and D. Herrero-Pérez. Evolutionary topology optimization of continuum structures under uncertainty using sensitivity analysis and smooth boundary representation. *Computers & Structures*, 205:15–27, 2018.

[127] J. Martínez-Frutos, D. Herrero-Pérez, M. Kessler, and F. Periago. Robust shape optimization of continuous structures via the level set method. *Computer Methods in Applied Mechanics and Engineering*, 305(Supplement C):271 – 291, 2016.

[128] A. Maury. *Shape optimization for contact and plasticity problems thanks to the level set method*. Theses, Université Pierre et Marie Curie - Paris VI, December 2016.

[129] A. Maury, G. Allaire, and F. Jouve. Shape optimisation with the level set method for contact problems in linearised elasticity. *SMAI-Journal of Computational Mathematics*, (3):249–292, 2017.

[130] A. Maury, G. Allaire, and F. Jouve. Elasto-plastic shape optimization using the level set method. *SIAM Journal on Control and Optimization*, 56(1):556–581, 2018.

[131] K. Maute, S. Schwarz, and E. Ramm. Adaptive topology optimization of elastoplastic structures. *Structural Optimization*, 15(2):81–91, 1998.

[132] C. Miehe, F. Welschinger, and M. Hofacker. Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field FE implementations. *International Journal for Numerical Methods in Engineering*, 83(10):1273–1311, 2010.

[133] F. Mignot. Contrôle dans les inéquations variationelles elliptiques. *Journal of Functional Analysis*, 22(2):130–185, 1976.

[134] F. Mignot and J.P. Puel. Optimal control in some variational inequalities. *SIAM Journal on Control and Optimization*, 22(3):466–476, 1984.

[135] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. Oxford: Clarendon Press, 2001.

[136] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.*, 42(5):577–685, 1989.

[137] A. Myśliński. Level set method for shape and topology optimization of contact problems. In *IFIP Conference on System Modeling and Optimization*, pages 397–410. Springer, 2007.

[138] J.C. Nagtegaal. On the implementation of inelastic constitutive equations with special reference to large deformation problems. *Computer Methods in Applied Mechanics and Engineering*, 33(1-3):469–484, 1982.

[139] J.C. Nagtegaal and J.E. De Jong. Some aspects of non-isotropic work-hardening in finite strain plasticity. *Plasticity of metals at finite strain: theory, experiment and computation*, pages 65–102, 1982.

[140] P. Nakshatrala and D.A. Tortorelli. Topology optimization of multiscale elastoviscoplastic structures. *Journal for Numerical Methods in Engineering*, 106(6):430–453, 2016.

[141] NX Nastran. User's guide. sl: Siemens plm software inc. 2007.

[142] Q.S. Nguyen. Bifurcation and post-bifurcation analysis in plasticity and brittle fracture. *Journal of the Mechanics and Physics of Solids*, 35(3):303–324, 1987.

[143] S. Osher and R. Fedkiw. Signed distance functions. In *Level set methods and dynamic implicit surfaces*, pages 17–22. Springer, 2003.

[144] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153. Springer Science & Business Media, 2006.

[145] S. Osher and F. Santosa. Level set methods for optimization problems involving geometry and constraints: I. Frequencies of a two-density inhomogeneous drum. *Journal of Computational Physics*, 171(1):272–288, 2001.

[146] O. Pantz. A frictionless contact algorithm for deformable bodies. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45(2):235–254, 2011.

[147] P. Pedersen. On thickness and orientational design with orthotropic materials. *Structural Optimization*, 3(2):69–78, 1991.

[148] K. Pham, H. Amor, J.J. Marigo, and C. Maurini. Gradient damage models and their use to approximate brittle fracture. *International Journal of Damage Mechanics*, 20(4):618–652, 2011.

[149] W. Prager. A New Method of Analyzing Stresses and Strains in Work-Hardening Plastic Solids. *Journal of Applied Mechanics*, 23(4):493–496, 06 2021.

[150] L. Rakotondrainibe. *Optimisation topologique des liaisons dans les systèmes mécaniques*. PhD thesis, 2020. Mathématiques appliquées Institut polytechnique de Paris 2020.

[151] L. Rakotondrainibe, G. Allaire, and P. Orval. *Topology optimization of connections in mechanical systems*. Theses, Institut Polytechnique de Paris, December 2020.

[152] L. Rakotondrainibe, J. Desai, P. Orval, and G. Allaire. Coupled topology optimization of structure and connections for bolted mechanical systems. preprint on https://hal.archives-ouvertes.fr/hal-03273522/, 2021.

[153] B. D. Reddy and J. B. Martin. Internal Variable Formulations of Problems in Elastoplasticity: Constitutive and Algorithmic Aspects. *Applied Mechanics Reviews*, 47(9):429–456, 09 1994.

[154] M. Sauter. *Numerical Analysis of Algorithms for Infinitesimal Associated and Non-Associated Elasto-Plasticity*. PhD thesis, Karlsruher Institut für Technologie (KIT), 2010.

[155] J. Schröder, O. Klaas, E. Stein, and C Miehe. A physically nonlinear dual mixed finite element formulation. *Computer Methods in Applied Mechanics and Engineering*, 144(1-2):77–92, 1997.

[156] Y.D. Seo, H.J. Kim, and S.K. Youn. Isogeometric topology optimization using trimmed spline surfaces. *Computer Methods in Applied Mechanics and Engineering*, 199(49-52):3270–3296, 2010.

[157] J.A. Sethian and A. Wiegmann. Structural boundary design via level set and immersed interface methods. *Journal of Computational Physics*, 163(2):489–528, 2000.

[158] J. R. Shewchuk. An introduction to the conjugate gradient method without the agonizing pain, 1994.

[159] H.P. Shrivastava, Z. Mroz, and R.N. Dubey. Yield criterion and the hardening rule for a plastic solid. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 53(9):625–633, 1973.

[160] O. Sigmund. Morphology-based black and white filters for topology optimization. *Structural and Multidisciplinary Optimization*, 33(4-5):401–424, 2007.

[161] O. Sigmund and K. Maute. Topology optimization approaches. *Structural and Multidisciplinary Optimization*, 48(6):1031–1055, 2013.

[162] J.C. Simo and T.J.R. Hughes. *Computational inelasticity*, volume 7. Springer Science and Business Media, 2006.

[163] J.C. Simo and T.A. Laursen. An augmented lagrangian treatment of contact problems involving friction. *Computers & Structures*, 42(1):97–116, 1992.

[164] J.C. Simo and R. L. Taylor. Consistent tangent operators for rate-independent elastoplasticity. *Computer Methods in Applied Mechanics and Engineering*, 48(1):101–118, 1985.

[165] J.C. Simo and R.L. Taylor. A return mapping algorithm for plane stress elastoplasticity. *International Journal for Numerical Methods in Engineering*, 22(3):649–670, 1986.

[166] S.W. Sloan and J.R. Booker. Removal of singularities in Tresca and Mohr–Coulomb yield functions. *Communications in Applied Numerical Methods*, 2(2):173–179, 1986.

[167] M. Sofonea and A. Matei. *Variational inequalities with applications: a study of antiplane frictional contact problems*, volume 18. Springer Science & Business Media, 2009.

[168] M. Sofonea and A. Matei. *Mathematical models in contact mechanics*, volume 398. Cambridge University Press, 2012.

[169] J. Sokolowski and A. Zochowski. On the topological derivative in shape optimization. *SIAM Journal on Control and Optimization*, 37(4):1251–1272, 1999.

[170] J. Sokolowski and J. P. Zolésio. On shape sensitivity analysis for visco-elastic-plastic problems. In *Control of Distributed Parameter Systems 1989*, pages 225–228. Elsevier, 1990.

[171] J. Sokolowski and J. P. Zolésio. *Introduction to shape optimization*. Springer, 1992.

[172] G. Strang. *Computational Science and Engineering*. Wellesley-Cambridge Press, 2007.

[173] N. Strömberg. Topology optimization of orthotropic elastic design domains with mortar contact conditions. In *World Congress of Structural and Multidisciplinary Optimisation*, pages 1427–1438. Springer, 2017.

[174] N. Strömberg and A. Klarbring. Topology optimization of structures in unilateral contact. *Structural and Multidisciplinary Optimization*, 41(1):57–64, 2010.

[175] Dassault Systemes. Abaqus/standard user's manual, 2014.

[176] E. Tanné, T. Li, B. Bourdin, J.J. Marigo, and C. Maurini. Crack nucleation in variational phase-field models of brittle fracture. *Journal of the Mechanics and Physics of Solids*, 110:80–99, 2018.

[177] M. Tisza. Development of lightweight steels for automotive applications. In *Engineering Steels and High Entropy-Alloys*. IntechOpen, 2020.

[178] M. Tur, F.J. Fuenmayor, and P. Wriggers. A mortar-based frictional contact formulation for large deformations using lagrange multipliers. *Computer Methods in Applied Mechanics and Engineering*, 198(37-40):2860–2873, 2009.

[179] J.L. Vie. *Second-order derivatives for shape optimization with a level-set method*. PhD thesis, Paris Est, 2016.

[180] C. H. Villanueva and K. Maute. Density and level set-XFEM schemes for topology optimization of 3-D structures. *Computational Mechanics*, 54(1):133–150, 2014.

[181] M. Wallin, N. Ivarsson, and M. Ristinmaa. Large strain phase-field based multi-material topology optimization. *International Journal for Numerical Methods in Engineering*, 104(9):887–904, 2015.

[182] M. Wallin, V. Jönsson, and E. Wingren. Topology optimization based on finite strain plasticity. *Structural and Multidisciplinary Optimization*, 54(4):783–793, 2016.

[183] F. Wang, B. S. Lazarov, and O. Sigmund. On projection methods, convergence and robust formulations in topology optimization. *Structural and Multidisciplinary Optimization*, 43(6):767–784, 2011.

[184] M. Wang, X. Wang, and D. Guo. A level set method for structural topology optimization. *Computer Methods in Applied Mechanics and Engineering*, 192(1-2):227–246, 2003.

[185] M.L. Wilkins. Calculation of elastic-plastic flow. Technical report, California Univ Livermore Radiation Lab, 1963.

[186] C. Wu, J. Fang, S. Zhou, Z. Zhang, G. Sun, G. P. Steven, and Q. Li. Level-set topology optimization for maximizing fracture resistance of brittle materials using phase-field fracture model. *International Journal for Numerical Methods in Engineering*, 121(13):2929–2945, 2020.

[187] Z. Xia, Y. Wang, Q. Wang, and C. Mei. GPU parallel strategy for parameterized LSM-based topology optimization using isogeometric analysis. *Structural and Multidisciplinary Optimization*, 56(2):413–434, 2017.

[188] Y.M. Xie and G.P. Steven. A simple evolutionary procedure for structural optimization. *Computers & Structures*, 49(5):885–896, 1993.

[189] T. Yamada, K. Izui, S. Nishiwaki, and A. Takezawa. A topology optimization method based on the level set method incorporating a fictitious interface energy. *Computer Methods in Applied Mechanics and Engineering*, 199(45-48):2876–2891, 2010.

[190] M. Zhou and R. Fleury. Fail-safe topology optimization. *Structural and Multidisciplinary Optimization*, 54(5):1225–1243, 2016.

[191] H. Ziegler. A modification of Prager's hardening rule. *Quarterly of Applied mathematics*, 17(1):55–65, 1959.

[192] O.C. Zienkiewicz and I.C. Cormeau. Visco-plasticity—plasticity and creep in elastic solids—a unified numerical solution approach. *International Journal for Numerical Methods in Engineering*, 8(4):821–845, 1974.