Discrete element and time-integration methods for elasto-plasticity and dynamic cracking

Frédéric Marazzato

UNIVERSITÉ PARIS-EST

Thèse

présentée pour obtenir le grade de

Docteur de l'université Paris-Est

École doctorale MSTIC, mention Mathématiques Appliquées

par Frédéric Marazzato

Discrete element and time-integration methods for elasto-plasticity and dynamic cracking

Méthodes d'éléments discrets et d'intégration temporelle pour l'élasto-plasticité et la fissuration dynamique

---

Soutenue publiquement le 29 mai 2020 devant le jury composé de :

Rapporteurs	Pr. Yves Renard	INSA Lyon
	Pr. Eugenio Oñate	UPC
Examinateurs	Pr. Patrick Joly	Inria Saclay
	Pr. Ivan Iordanoff	Arts et Métiers ParisTech
	Dr. Flaviana Iurlano	LJLL
Directeur de thèse	Pr. Alexandre Ern	ENPC - Université Paris-Est
Co-directeur de thèse	Pr. Karam Sab	ENPC - Université Paris-Est
Encadrant industriel	Dr. Jean-Philippe Braeunig	CEA
Invité	Dr. Laurent Monasse	Inria Sophia-Antipolis

À ma famille, à mes amis, à mon amour,

"You are your worst enemy. It is your negative thoughts
that hold you back, nothing else."

Leon Brown

"To infinity and beyond !"

Light Buzz

Remerciements

Je commence ces remerciements par les membres de mon jury de thèse et tout d'abord les rapporteurs Yves Renard et Eugenio Oñate. Merci à vous d'avoir pris le temps de vous plonger dans mon manuscrit et d'avoir écrit des rapports aussi élogieux. Je remercie ensuite Patrick Joly qui m'a fait l'honneur de présider mon jury. Je le remercie également pour son intérêt concernant notre méthode d'intégration temporelle. Je remercie ensuite Flaviana Iurlano et Ivan Iordanoff d'avoir été des examinateurs aussi sympathiques et intéressés par mes travaux de thèse.

Je remercie également Alexandre Ern, Karam Sab et Laurent Monasse pour m'avoir encadré pendant ces quelques années, ce qui n'a pas dû être de tout repos. Merci à Alexandre pour sa rigueur, sa disponibilité et son investissement qui m'ont permis de réaliser ma thèse dans les meilleurs conditions et d'apprendre un nombre incalculable de concepts passionnants. Merci à Karam d'avoir toujours su être disponible pour discuter mécanique malgré les nombreuses contraintes que le poste de directeur du laboratoire Navier implique. Enfin, merci à Laurent d'avoir toujours été à mes côtés pendant ces longues années. Je le remercie pour son aide qui a été salvatrice concernant le développement des codes de calcul développés au cours de cette thèse ainsi que dans le débogage. Dieu sait que nous nous sommes arrachés les cheveux !

Je me dois également de remercier le CEA pour avoir financé cette thèse et m'avoir fait confiance pendant ces années périlleuses. Je remercie également CM puis JPB et LA pour avoir suivi avec beaucoup d'intérêt l'évolution de mes travaux de thèse.

Je tiens également à remercier à l'ENPC Jérémy Bleyer d'avoir toujours pu trouver un peu de temps pour discuter avec moi et m'aider à m'y retrouver dans la documentation aride de Fenics. Je remercie également Gabriel Stoltz et Frédéric Tayeb de m'avoir fait suffisamment confiance pour me confier des enseignements. J'ai beaucoup apprécié ces moments et le contact avec les étudiants. L'enseignement est une activité profondément enrichissante. De plus, l'enseignement ne nourrit pas seulement les connaissances des étudiants mais aussi celles de l'enseignant et ce phénomène est loin d'être négligeable.

Je remercie ensuite Isabelle Simunic qui a été une figure tutélaire pendant cette thèse et qui a toujours fait de son mieux pour que les doctorants puissent travailler dans les meilleurs conditions possibles au CERMICS. Je remercie ensuite Fatna puis Stéphanie qui ont mis en place tout le nécessaire pour que mes déplacements se passent sans encombre. Avec de tels soutiens, le doctorat ne pouvait que bien se passer. Je remercie également les doctorants du CERMICS avec qui nous avons partagé joies et désillusions. Je ne me risquerait pas à faire de liste nominative de peur d'oublier quelqu'un. Mais je suis convaincu que ceux auxquels je pense se reconnaîtront.

Je remercie également Christian, Jérémy et Alain de l'équipe quantic à l'Inria pour ces séances de courses qui m'ont permis de m'améliorer nettement et de ne pas flancher dans les moments difficiles de la rédaction. Je remercie également Vincent et Michiel pour de sympathiques discussions. Je salue ensuite l'équipe Serena de l'Inria Paris qui m'a accueilli à bras ouverts et a été d'un soutient sans faille dans les moments difficiles. Je remercie donc François, Michel, Géraldine, Florent, Ani, Patrick, Jad, Martin et Derya qui sait maintenir le cap et faire avancer tous les dossiers sans qu'ils soient oubliés dans des arcanes administratives sans fin et dont la porte est toujours ouverte si besoin. Merci à vous tous !

Je voudrais également remercier certains de mes professeurs qui ont su nourrir ma curiosité pour les sciences et la rigueur mathématique. Ce furent d'abord MM. Lefort, Mansart et Pot au lycée. J'ai des souvenirs très tendres de mes années de lycée et ces professeurs n'y sont pas étrangers. Suivirent Mme Chabane, MM. Teller et Alosy en classes préparatoires. Ce furent des années difficiles mais j'ai appris énormément à leurs côtés. Par la suite, je dois remercier Alain Ehrlacher qui fût bien plus qu'un professeur et un mentor. Son dévouement pour le département GMM (Génie Mécanique et Matériau) n'a d'égal que l'attention et le soutient qu'il sait apporter à chacun de ses étudiants et à ses projets. GMM ne sera plus le même sans toi Alain ! Merci pour tout et notamment de m'avoir encouragé lorsque j'ai voulu poursuivre avec le Master Mathématiques et Applications qui n'est as le débouché naturel de GMM. J'en profite pour remercier Eric Cancès (alors responsable du master pour l'ENPC) de m'avoir fait confiance et de m'avoir autorisé à suivre ce master qui s'est révélé très riche d'enseignements. Je n'ai aucun regret !

Je passe ensuite aux moments de détente qui sont au moins aussi importants que ceux de travail. Je remercie Régis, Mathieu, Laurent et tous les autres grimpeurs que j'ai pu connaître sur les murs d'Arkose ainsi que les membres du staff. Vous m'avez permis de me détendre et de faire respirer mon cerveau après des journées parfois complexes et stressantes. Vous avez aussi toujours su me challenger pour donner le meilleur de moi-même et tout cela dans une bonne humeur et une camaraderie incomparable. Merci à vous !

Je remercie ensuite ma famille qui m'a apporté un soutient sans faille durant toutes mes années d'études et sans laquelle je ne serai certainement pas la personne que je suis. Rassure-toi maman c'est (presque) fini !

Je remercie évidemment les vieux de la vieille pour ... tout. Il y aurait tellement de choses à citer après toutes ces années vécues ensemble que je préfère y englober toute mon expérience. Je ne serai pas arrivé là sans votre soutient, notamment dans les moments difficiles de la vie. Tous les remerciements que je pourrai écrire ici ne pourront jamais retranscrire toute ma gratitude et l'importance que j'attache à notre amitié.

Le meilleur pour la fin, je remercie celle qui a supporté mes angoisses et partagé mes joies au cours des montagnes russes que représente cette thèse.

Table des matières

Résumé

Cette thèse propose des contributions aux méthodes éléments discrets (MED) et à l'intégration temporelle explicite avec pour objectif applicatif les calculs de fissuration et de fragmentation pour des matériaux métalliques soumis à des chargements dynamiques. Les MED, qui sont traditionnellement utilisées pour simuler le comportement de matériaux granulaires, sont réinterprétées à la lumière des méthodes de discrétisation de gradient afin d'être appliquées à la simulation de matériaux continus. Les maillages utilisables avec la MED proposée ont été étendus des maillages de Voronoi à des maillages polyédriques généraux. Les comportements simulables par la méthode ont été étendus de l'élasto-dynamique à l'élasto-plasticité dynamique par l'ajout d'un degré de liberté tensoriel par cellule. De plus, la méthode est robuste par rapport à la limite incompressible et ses paramètres ne dépendent que des paramètres matériau. Une méthode d'intégration temporelle explicite conservant une pseudo-énergie, même pour des comportements non-linéaires et des pas de temps variables, a également été développée afin d'éviter une dissipation numérique de l'énergie disponible pour la dissipation plastique et la fissuration. Cette méthode a été couplée avec la MED précédente. Enfin, la propagation de fissures de Griffith à travers les facettes du maillage a été intégrée à la MED pour des comportements élastiques linéaires en deux dimensions d'espace. Le taux de restitution d'énergie est calculé pour chaque mode de fissuration à partir des facteurs d'intensités des contraintes qui sont approchés près de la fissure. Enfin, un critère de maximisation de la densité d'énergie élastique sur les facettes liée à la pointe de fissure permet de simuler l'orientation de la propagation.

Mot-clefs : Méthode éléments discrets, dynamique rapide, fissuration, elasto-plasticité, conservation d'énergie, intégration temporelle explicite

Abstract

The present Ph.D. dissertation proposes contributions to discrete element methods (DEM) and explicit time integration schemes with a view towards dynamic cracking for metallic materials under dynamic loading. DEM, which are usually used to simulate granular materials, are understood through the prism of gradient discretization methods in order to simulate continuous materials. The method has been extended from previous Voronoi meshes to support general polyhedral meshes. Material behaviours have been extended from elasto-dynamics to dynamic elasto-plasticity through the addition of a tensorial degree of freedom per mesh cell. The method is robust with respect to the incompressible limit and its parameters only depend on material parameters. Moreover, an explicit pseudo-energy conserving time integration method has been developed, even for nonlinear behaviours and variable time steps, so as to avoid the dissipation of energy available for plastic dissipation and cracking. The method has been coupled to the proposed DEM. Finally, Griffith crack propagation through the mesh facets has been adapted to the present DEM for linear elastic behaviours in two space dimensions. The energy release rate is computed for every cracking mode using the stress intensity factors approximated close to the crack. A criterion of maximization of elastic energy density is used to simulate kinking.

Keywords: Discrete element methods, fast dynamics, cracking, elasto-plasticity, energy conservation, explicit time integration

Chapter 1

Introduction

Cette thèse de doctorat a été préparée, pour sa partie académique au CERMICS, en collaboration avec le Laboratoire Navier. Le CERMICS est le laboratoire de mathématiques appliquées de l'École Nationale des Ponts et Chaussées (ENPC) et le laboratoire Navier est une unité mixte de recherche entre l'ENPC, l'IFSTTAR et le CNRS. Concernant sa partie applicative, cette thèse a été préparée au Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA).

Dans cette thèse sont proposées des contributions aux méthodes d'intégration temporelle pour les Équations aux Dérivées Ordinaires (EDO) ayant une structure hamiltonienne, aux discrétisations en espace de type éléments discrets pour la résolution des équations de l'élasto-plasticité et aux méthodes numériques de calculs de fissuration quasi-statique et dynamique.

Ce chapitre introductif est structuré en quatre sections. La première décrit le problème expérimental qui a motivé cette thèse ainsi que ses mécanismes physiques sous-jacents. La deuxième section présente une mise en équation de ces phénomènes physiques. La troisième section présente un état de l'art des méthodes numériques en lien avec les travaux effectués dans cette thèse. Enfin, la dernière section décrit les contributions qui sont contenues dans chaque chapitre du manuscrit.

1.1 Motivations

Nous commençons par décrire brièvement dans cette première section les missions du CEA ainsi que les enjeux scientifiques et industriels ayant mené le CEA à financer cette thèse. Puis, les phénomènes physiques étudiés dans le cadre de cette thèse sont présentés.

1.1.1 Contexte industriel

Le CEA est structuré en plusieurs directions : la direction de l'énergie nucléaire, la direction de la recherche fondamentale, la direction des sciences du vivant, la direction de la recherche technologique et enfin la direction des applications militaires, ci-après dénotée DAM, au sein de laquelle cette thèse de doctorat s'est inscrite. La mission historique de la DAM consiste à développer et maintenir la force de dissuasion nucléaire française. Des missions complémentaires ont été confiées au cours du temps à la DAM, notamment dans le cadre des traités internationaux de non-prolifération des armes nucléaires et d'interdiction des essais nucléaires. Ces missions consistent en plusieurs activités de surveillance de l'environnement, comme par exemple l'analyse d'événements sismiques potentiellement consécutifs à des essais nucléaires prohibés ou encore la détection de la présence de matériaux fissibles à l'échelle mondiale. Concernant les missions d'analyse de l'environnement géophysique, la DAM assure le développement et le maintien de codes de calculs permettant de simuler des phénomènes sismiques et dynamiques dans de nombreux matériaux. Ayant comme objectif des questions de sécurité et fort de cette expertise, le CEA a pour ambition de réaliser, notamment à l'aide du super

calculateur Tera1000, la simulation d'une expérience de dynamique rapide consistant en la fragmentation d'une coquille métallique qui sera décrite dans la sous-section suivante.

1.1.2 Problème phare : fragmentation d'une coquille métallique

Une expérience de fragmentation d'une coquille métallique est d'abord décrite. Les enjeux techniques que cette expérience recouvre pour le CEA sont ensuite présentés. Cette expérience ne sera pas simulée dans le cadre de cette thèse, mais les outils numériques qui y seront développés ont pour but de contribuer à cette simulation.

Description de l'expérience

Dans [Bolis et al., 2013] est décrit le dispositif expérimental ainsi que les résultats de l'essai de fragmentation intéressant le CEA. La figure 1.1 présente un schéma du dispositif expérimental. Une coquille en titane est posée, et non encastrée, sur un support horizontal plan. La coquille est représentée en noir et le support en bleu. Elle est libre de contraintes sur sa face externe et accueille en son sein une charge explosive parallélépipédique, repérée par la lettre C sur la figure 1.1, qui n'est pas en contact avec sa face interne. Lors de la détonation des

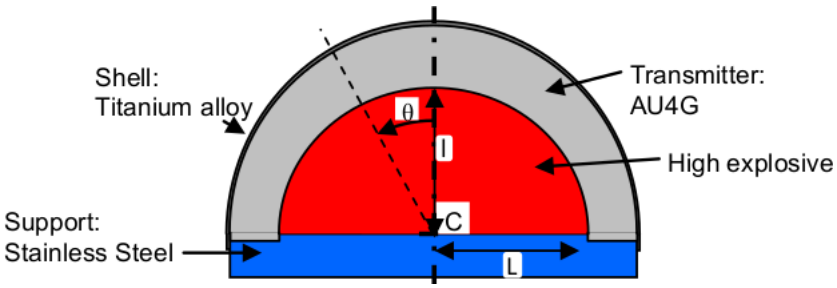


FIGURE 1.1 – Dispositif expérimental ([Bolis et al., 2013])

explosifs, les produits de l'explosion sont retenus par le transmetteur, représenté en gris sur la figure 1.1, afin de ne pas impacter la coquille. Ce dernier permet également de transmettre à la coquille l'onde de pression de l'air, représentée en rouge. Dans les conditions de l'expérience, une onde de compression atteint après 1μs la face interne de la coquille au niveau du point le plus haut et y engendre la propagation d'une onde élastique de compression. Dans la mesure où la forme des explosifs n'est pas isotrope, l'onde de compression qui atteint la coquille ne l'est pas non plus. Cette dernière est alors soumise à une compression qui n'est pas purement isotrope. Lors du passage de l'onde de compression dans la coquille, cette dernière plastifie mais ne se fissure pas. Une fois la face externe de la coquille atteinte, l'onde se réfléchit et génère des ondes de cisaillement et de traction qui engendrent alors l'amorçage de diverses fissures. Après de nombreux branchements et jonctions de fissures, de premiers fragments se détachent de la coquille aux alentours de 10μs et entrent alors en vol balistique. La coquille est entièrement fragmentée après approximativement 15μs. Des caméras à très grande résolution temporelle filment l'expérience tandis que des lasers, dont la résolution temporelle est encore supérieure, mesurent les déplacements de la coquille. Le nombre et la taille des fragments sont mesurés a posteriori à l'aide de techniques d'analyse d'images, comme présenté dans la figure 1.2. La taille moyenne des fragments est de l'ordre de 150mm^2.

Objectifs de la simulation de l'expérience

L'objectif de l'expérience est de mesurer le nombre et la distribution de taille des fragments et de les comparer, dans un premier temps, à des lois analytiques comme, par exemple, celle proposée dans [Grady and Olsen, 2003]. Dans un second temps, l'objectif est de simuler numériquement l'expérience afin de permettre la validation de codes industriels de dynamique rapide.

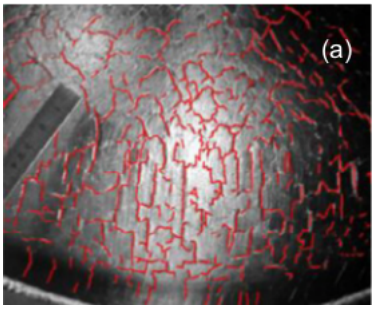


FIGURE 1.2 – Fragments détectés par analyse d'image [Bolis et al., 2013]

À plus long terme, l'utilisation de codes industriels validés par des expériences permettra leur utilisation pour dimensionner des structures. On peut citer comme exemple le dimensionnement d'enceintes de confinement soumises à des surpressions ou encore à des explosions. L'utilisation de la simulation numérique peut également permettre d'accéder à des données difficilement mesurables expérimentalement comme l'énergie cinétique des fragments expulsés.

1.1.3 Principaux phénomènes physiques

La description de la déformation macroscopique de la coquille est étudiée dans le cadre de la Mécanique des milieux continus (MMC) qui est une description continue de la matière et qui néglige la nature atomistique de cette dernière. La suite de cette section s'emploie à décrire les phénomènes physiques et ordres de grandeurs entrant dans le cadre de la MMC et intervenant dans la simulation de l'expérience décrite en section 1.1.2. Tous ces phénomènes ne sont pas pris en compte dans cette thèse ; l'objectif est ici d'en dresser un panorama relativement complet.

Dynamique rapide

La mécanique peut traditionnellement être étudiée dans le cadre de plusieurs régimes qui caractérisent la vitesse à laquelle une structure se déforme en fonction des efforts qui lui sont appliqués. Le régime « statique » consiste en le calcul de la déformation à l'équilibre d'une structure sous chargement donné sans aucune évolution temporelle. La structure est alors entièrement considérée à l'équilibre mécanique et les termes inertiels sont nuls. Le régime « quasi-statique » recouvre le calcul de l'évolution temporelle de la déformation d'une structure mais avec des temps caractéristiques permettant de négliger raisonnablement les termes inertiels. Dans les cas où cette approximation n'est pas pertinente, on parle alors de dynamique. La dynamique rapide est une des branches de la dynamique. Elle consiste notamment à calculer l'évolution temporelle complète de la structure en tenant compte des termes inertiels, et en particulier des phénomènes de propagation d'ondes, qui jouent un rôle prépondérant dans la physique étudiée. Les temps caractéristiques typiques des phénomènes physiques relevant de la dynamique rapide sont de l'ordre de la μs à la s ; on peut citer comme exemples typiques la propagation d'ondes de choc et la fragmentation.

Comportement asymétrique en dilatation et en contraction

Cette section présente un mécanisme physique entrant en compte dans la modélisation du comportement des métaux intervenant dans l'expérience de la section 1.1.2 mais qui n'est pas traité directement dans ce manuscrit. Le comportement en dilatation est considéré comme linéaire et réversible, ce qui est communément dénommé « élastique linéaire ». Cela signifie qu'après avoir mis en traction un échantillon, la force appliquée pour le maintenir déformé est proportionnelle à son niveau de déformation et que lorsque les contraintes sont relâchées, ce dernier retrouve son état originel sans déformation macroscopique rémanente. Concernant le

comportement en contraction, il a été observé que la matière fortement comprimée a tendance à raidir. Ce phénomène est mesuré puis ensuite décrit par une courbe appelée « courbe d'Hugoniot ». La figure 1.3 présente des exemples de courbes d'Hugoniot pour plusieurs matériaux. Cette courbe caractérise la relation entre l'inverse de la masse volumique d'un échantillon

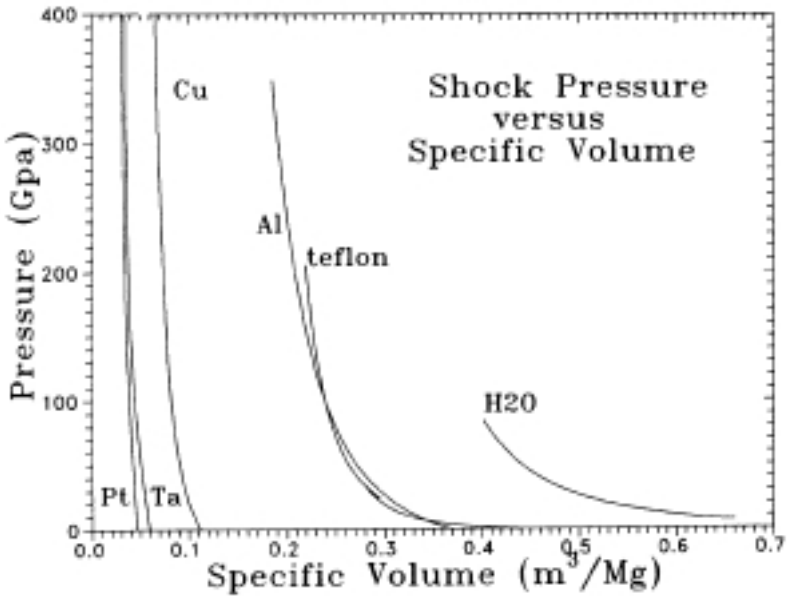


FIGURE 1.3 – Courbes d'Hugoniot [Asay and Shahinpoor, 2012, p. 79]

et la pression isotrope qui lui est appliquée. Les comportements considérés en compression pour l'étude de la coquille en expansion sont non-linéaires et réversibles. Le raidissement en compression mène potentiellement à la création de chocs au sein de l'échantillon étudié et leur propagation soulève des problèmes concernant leur traitement numérique qui ne seront pas traités dans ce manuscrit. Le comportement en dilatation et en contraction est considéré comme non-directement dépendant du temps au sens où le comportement d'un échantillon n'est impacté que par l'évolution temporelle de sa déformation mais pas directement par sa vitesse de déformation.

Plasticité

La déformation dite « plastique », contrairement à la déformation élastique, n'est pas réversible. Cela signifie qu'après relâchement des contraintes imposées à un échantillon, ce dernier présentera une déformation macroscopique rémanente. Un exemple de déformation plastique est présenté en figure 1.4. Dans le cas des matériaux cristallins, tels les métaux, cette déformation rémanente est due à la formation et à l'accumulation de dislocations au sein de l'échantillon étudié. Les dislocations sont des défauts microscopiques apparaissant dans les matériaux cristallins et brisant les symétries des structures cristallines. Au niveau macroscopique, la plasticité est généralement décrite par un critère portant sur l'état de contrainte local du matériau. Ce point de vue sera adopté dans la suite de ce manuscrit. Concernant l'expérience de la section 1.1.2, le comportement devrait être considéré comme visco-plastique à cause des très grandes vitesses de déformation observées. Le caractère « visqueux » d'un échantillon traduit que son comportement dépend directement de la vitesse de déformation dans l'échantillon étudié, contrairement au comportement considéré précédemment qui n'en dépend pas directement.

1. https://tobinelliott.wordpress.com/2013/03/16/spoon-feeding-the-boy/

FIGURE 1.4 – Déformation plastique d'une cuillère métallique [1]

Fissuration

La fissuration est un phénomène physique consistant en la propagation de discontinuités macroscopiques dans des échantillons précédemment considérés comme macroscopiquement sains, c'est-à-dire ne présentant pas de discontinuités macroscopiques. La sensibilité des matériaux à la fissuration peut se mesurer expérimentalement. La fissuration est un mécanisme de dissipation d'énergie qui permet à un échantillon de dissiper de l'énergie mécanique, qui est de nature volumique, en ouvrant des fissures en son sein, qui sont de nature surfacique. La figure 1.5 présente un exemple de fissure dans un milieu continu. La dissipation d'énergie lors de

FIGURE 1.5 – Assiette fissurée [2]

la fissuration est majoritairement due à la libération de phonons lors de la rupture de liens atomiques entre les lèvres de fissure et minoritairement due à la dissipation d'énergie directe provoquée par la rupture de ces mêmes liens atomiques. Des simulations de dynamique moléculaire visant à reproduire les comportements observés macroscopiquement ont été réalisées par exemple dans [Atrash et al., 2011]. La fissuration est généralement étudiée sur des échelles de temps longues, de l'ordre de grandeur de quelques secondes à quelques heures (au-delà, on parle plutôt de « fatigue »). La fissuration rentre généralement dans le cadre de l'étude quasi-statique des structures, décrite précédemment. Dans le cadre de la dynamique rapide, la fissuration, qui est alors dynamique, prend généralement la forme de la fragmentation, dont il est question dans la sous-section suivante.

Fragmentation

La fissuration dynamique, par opposition à la fissuration quasi-statique, est étudiée sur des temps beaucoup plus courts allant de la micro-seconde à quelques secondes. Dans la mesure où le

2. https://www.18h39.fr/articles/astuce-reparer-vaisselle-cassee-lait.html

temps caractéristique du chargement appliqué est beaucoup plus court que dans le cadre quasi-statique, il se rapproche ou est inférieur à celui de la propagation des ondes élastiques à travers l'échantillon étudié. Dans ce cas, l'échantillon n'est globalement plus à l'équilibre mécanique et des singularités peuvent alors apparaître indépendamment à différents endroits de l'échantillon. La fissuration dynamique se caractérise en général par la propagation de plusieurs fissures avec de nombreux branchements et jonctions de fissures, comme illustré en figure 1.6. À travers

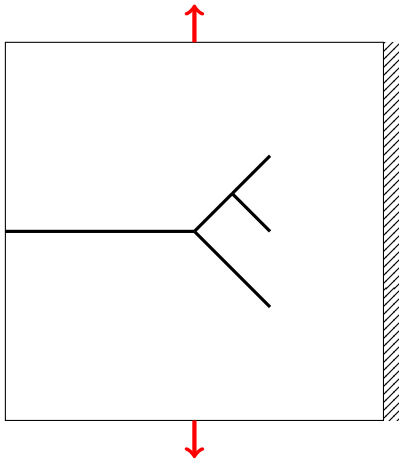


FIGURE 1.6 – Branchements d'une fissure lors d'un essai en traction d'un échantillon (schéma de principe).

ces phénomènes, des morceaux, également appelés « fragments », se forment et se détachent de l'échantillon originellement sain, comme présenté dans la figure 1.7. C'est ce phénomène, pouvant être accompagné d'une éjection des fragments avec une énergie cinétique importante, que l'on appelle fragmentation. Du point de vue de l'ingénierie, la prévision de la fragmentation

FIGURE 1.7 – Assiette fragmentée [3]

est critique dans le dimensionnement de structures pour des raisons de sécurité. Par exemple, des fragments éjectés à très haute vitesse sont potentiellement mortels pour les utilisateurs d'une voiture lors d'un crash.

1.2 Équations et lois de comportement

Cette deuxième section présente les Équations aux Dérivées Partielles (EDP) choisies pour modéliser les phénomènes physiques comme ceux intervenant dans l'expérience décrite à la

3. https://stock.adobe.com/fr/search?k=%22assiette+cass%C3%A9e%22&asset_id=209407

section 1.1.2. La première sous-section est consacrée aux équations de la dynamique des milieux continus. La deuxième sous-section traite le comportement des matériaux considérés dans l'expérience. La troisième sous-section est consacrée à la fissuration et à la fragmentation.

1.2.1 Dynamique des milieux continus

La simulation est effectuée sur un intervalle temporel fini noté $[0, T]$, où $T > 0$ est la durée de la simulation et le temps courant est noté $t \in [0, T]$. La masse volumique du matériau est notée $\rho > 0$ et considérée indépendante du temps. Le domaine spatial de la structure est noté $\Omega \subset \mathbb{R}^d$, où $d \in \{2, 3\}$. Le bord du domaine est noté $\partial\Omega$ et est partitionné en 2 parties : $\partial\Omega_N$ sur laquelle est imposé la contrainte normale $g(t)$, et $\partial\Omega_D$ sur laquelle est imposé le déplacement $u_D(t)$. La figure 1.8 illustre les grandeurs évoquées. Le déplacement est noté $u \equiv u(t, x)$ où

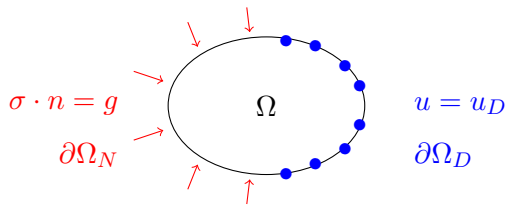


FIGURE 1.8 – Domaine et conditions de bord

$x \in \Omega$, la vitesse est notée $\dot{u}$ et les contraintes σ. Les conditions de bord sont décrites par

$$\begin{cases} u = u_D \text{ sur } \partial\Omega_D, \\ \sigma \cdot n = g \text{ sur } \partial\Omega_N, \end{cases} \quad \forall t \in [0, T], \tag{1.1}$$

où n est la normale sortante à $\partial\Omega$. Les conditions initiales prescrivent un déplacement initial u_0 et une vitesse initiale v_0 sous la forme :

$$\begin{cases} u(0) = u_0 \text{ dans } \Omega, \\ \dot{u}(0) = v_0 \text{ dans } \Omega. \end{cases} \tag{1.2}$$

L'équation suivante constitue l'équivalent, dans le cadre de la MMC, de la seconde loi de Newton qui régit l'évolution dynamique d'un point matériel :

$$\operatorname{div}(\sigma) + \rho f - \rho\ddot{u} = 0 \text{ dans } \Omega, \quad \forall t \in [0, T], \tag{1.3}$$

où f représente la densité volumique d'efforts appliqués et $\ddot{u}$ désigne l'accélération. Le système constitué des équations (1.1), (1.2) et (1.3) n'est pas fermé et nécessite l'adjonction d'une « loi de comportement » reliant les contraintes σ avec les déformations qui sont calculées à partir du gradient en espace des déplacements ∇u.

1.2.2 Lois de comportement

Les lois de comportement se doivent d'être représentatives de la réponse d'un matériau à différents types de sollicitation. Elles dépendent des propriétés mésoscopiques, microscopiques et, parfois même nanoscopiques, des matériaux considérés. Dans la mesure où la MMC se contente de décrire les échantillons étudiés macroscopiquement, elle prend pour hypothèse le comportement des matériaux et les suppose comme une donnée d'entrée. C'est alors le rôle de la science des matériaux et de divers autres domaines de la physique de définir les lois de comportement les plus adaptées pour chaque matériau et chaque gamme d'utilisation de ce matériau, ce qui comprend par exemple la température, la pression, le niveau de déformation, l'historique d'irradiation ou le changement de phase que peut subir un échantillon donné. Cette section aborde en premier lieu l'élasticité linéaire avant de continuer avec les comportements non-linéaires que sont la plasticité et l'hyper-élasticité.

Elasticité linéaire

Il existe plusieurs mesures des déformations d'un solide. Dans ce manuscrit, nous utiliserons le tenseur d'ordre 2 des déformations de Green-Lagrange noté $e \equiv e(u)$ et dont l'expression est donnée par

$$e(u) := \frac{1}{2}\left(\nabla u + \nabla u^T + \nabla u^T \cdot \nabla u\right). \tag{1.4}$$

Lorsque les déformations dans l'échantillon étudié sont faibles, $\|\nabla u\| \ll 1$, il est possible d'approcher le tenseur $e(u)$ par le tenseur d'ordre 2 des déformations linéarisées noté $\varepsilon \equiv \varepsilon(u)$ donné par

$$\varepsilon(u) := \frac{1}{2}\left(\nabla u + \nabla u^T\right). \tag{1.5}$$

De plus, lorsque les déplacements sont faibles, $|u| \ll \mathrm{diam}(\Omega)$ où diam désigne le diamètre d'un ouvert de $\mathbb{R}^d$, il est possible de confondre la géométrie de Ω en configuration initiale et en configuration déformée et donc de travailler avec le tenseur des contraintes σ en configuration initiale. La réunion des deux hypothèses précédentes permet de travailler dans l'hypothèse des petites perturbations ou « HPP ». En HPP, une loi de comportement se note généralement $\sigma(u) = \mathbb{C}(x) : \varepsilon(u)$ où $x \in \Omega$ et $\mathbb{C}$ est un tenseur d'ordre 4, appelé tenseur de rigidité. Dans le cadre des comportements étudiés dans cette thèse, le comportement est supposé homogène, ce qui signifie que $\mathbb{C}$ est constant en espace et en temps, et isotrope, ce qui implique que $\mathbb{C}$ appartient à un espace vectoriel de dimension deux. $\mathbb{C}$ dépend donc de deux paramètres qui peuvent être E (le module de Young) et ν (le coefficient de Poisson), ou bien λ et μ (les coefficients de Lamé). Les premiers s'expriment en fonction des seconds et réciproquement. Par exemple, on a

$$\lambda = \frac{\nu E}{(1-\nu)(1+2\nu)} \text{ et } \mu = \frac{E}{2(1+\nu)}. \tag{1.6}$$

On obtient alors la loi de comportement élastique linéaire homogène isotrope suivante, où δ est le tenseur d'ordre 2 unité, δ_4 est le tenseur d'ordre 4 unité et $\otimes$ le produit tensoriel :

$$\mathbb{C} = \lambda\delta \otimes \delta + 2\mu\delta_4. \tag{1.7}$$

Il est possible de réécrire la loi de comportement précédente afin de faire apparaître la partie sphérique (qui recouvre la dilation et la contraction mentionnées précédemment) et la partie déviatorique :

$$\mathbb{C} = (3\lambda + 2\mu)\frac{1}{3}\delta \otimes \delta + 2\mu(\delta_4 - \frac{1}{3}\delta \otimes \delta). \tag{1.8}$$

Elasto-plasticité HPP

La déformation plastique rémanente est notée ε_p. Contrairement à la déformation totale $\varepsilon(u)$, ε_p n'est pas issu d'une dérivation directe des déplacements. Le critère de plasticité est noté $\varphi \leq 0$, où φ dépend du tenseur des contraintes σ et d'autres variables telles la déformation plastique cumulée p et sa dérivée ou encore la température T. Lorsque l'inégalité $\varphi < 0$ est vérifiée, la déformation est purement élastique linéaire. Dans le cas contraire, on a $\varphi = 0$ et un taux de déformation plastique vient incrémenter ε_p. Concernant le seuil d'élasticité, un critère de type Von Mises est choisi car ce dernier ne tient compte que de la composante déviatorique des contraintes, ce qui est représentatif du comportement des métaux. La contrainte équivalente est alors notée σ_{eq} et se calcule sous la forme

$$\sigma_{eq} = \sqrt{\frac{3}{2}\mathrm{dev}(\sigma) : \mathrm{dev}(\sigma)}, \tag{1.9}$$

où dev est l'opérateur linéaire donnant la partie déviatorique d'un tenseur a d'ordre 2 sous la forme $\mathrm{dev}(a) = a - \frac{1}{3}\mathrm{tr}(a)\delta$. Une loi d'écoulement plastique dite « associée » est utilisée pour

caractériser le flot plastique, ce qui se traduit par

$$\begin{cases} \Lambda\varphi = 0,\ \Lambda \geq 0 \text{ et } \varphi \leq 0 \text{ dans } \Omega, \\ \dot{\varepsilon}_p = \Lambda \dfrac{\partial \varphi}{\partial \sigma} \text{ dans } \Omega, \end{cases} \qquad \forall t \in [0, T], \tag{1.10}$$

où $\Lambda \geq 0$ est le multiplicateur de Lagrange associé à la contrainte inégalité $\varphi \leq 0$. Les vitesses de déformation pouvant être très élevées, le matériau peut, en toute rigueur, être considéré comme élasto-visco-plastique. Une loi d'écoulement de type Johnson–Cook [Johnson and Cook, 1983] peut alors être considérée comme représentative de ce type de comportement. Cette loi prend la forme suivante :

$$\varphi(\sigma, p, \dot{p}, T) = \sigma_{eq} - (\sigma_0 + Bp^n)\left(1 + C\log\left(\frac{\dot{p}}{\dot{p}_0}\right)\right)\left(1 - \left(\frac{T - T_0}{T_f - T_0}\right)^m\right), \tag{1.11}$$

où σ_0 est la limite élastique de référence du matériau et B et C sont des paramètres matériau, T_0 est la température de référence de la loi, T la température du matériau et T_f la température de fusion du matériau. La déformation plastique cumulée p et sa dérivée $\dot{p}$ se calculent comme suit :

$$\dot{p}(t) = \sqrt{\frac{2}{3}\dot{\varepsilon}_p(t) : \dot{\varepsilon}_p(t)}, \tag{1.12}$$

$$p(t) = \int_0^t \dot{p}(\tau) d\tau. \tag{1.13}$$

Enfin, $\dot{p}_0$ est un paramètre de normalisation qui dépend des gammes de vitesses atteintes lors des essais permettant de caractériser les paramètres du matériau. Cependant, dans le cadre de cette thèse, par souci de simplicité, seuls des comportements élasto-plastiques linéaires et isothermes sont considérés. Dans ce cas, l'équation (1.11) se simplifie sous la forme suivante :

$$\varphi(\sigma, p) = \sigma_{eq} - \sigma_0 - Bp. \tag{1.14}$$

Enfin, la loi de comportement élasto-plastique HPP s'écrit :

$$\sigma(u) = \mathbb{C} : (\varepsilon(u) - \varepsilon_p)\,. \tag{1.15}$$

On notera que la loi de comportement élasto-plastique HPP n'agit que sur la partie déviatorique du comportement puisque l'on a

$$\mathrm{dev}\,(\sigma(u)) = 2\mu\ \mathrm{dev}(\varepsilon(u)). \tag{1.16}$$

La partie restante des contraintes est de la forme $P\delta$, où P est appelée pression, et vérifie

$$P = \frac{1}{3}\mathrm{tr}\,(\sigma(u))\,. \tag{1.17}$$

Hyper-élasticité en dilatation

Comme indiqué en sous-section 1.1.3, ce type de comportement n'est pas directement traité dans ce manuscrit de thèse mais il est brièvement évoqué pour compléter la présentation. La variation de volume, notée J, se calcule à partir du gradient de déplacements en utilisant l'équation :

$$J = \det(\nabla u + \delta), \tag{1.18}$$

où det est le déterminant de l'endomorphisme associé à $\nabla u + \delta$, qui est un tenseur d'ordre 2. La pression P est alors une fonction de J, ce qui est noté génériquement $P = \mathcal{F}(J)$. La fonction $\mathcal{F}$ est généralement une loi puissance avec un exposant α tel que $\alpha < -1$. Cette dernière inégalité traduit le fait qu'il faut augmenter plus que linéairement la pression dans un échantillon afin de continuer à réduire linéairement sa masse volumique, ce qui permet de modéliser le raidissement du matériau. On notera que le lien entre contrainte et déformation devient non-linéaire.

1.2.3 Fissuration et fragmentation

Cette sous-section présente d'abord une vision quasi-statique de la fissuration, comme introduite à l'origine par Griffith, puis poursuit par la vision dynamique développée par Freund. Puis, les fissures cohésives, qui généralisent les fissures de Griffith, seront abordées. Enfin, cette sous-section se termine par une approche statistique de la fragmentation. Ces deux derniers aspects de la modélisation ne seront pas traités dans cette thèse.

Dans ce chapitre introductif, la présentation de la fissuration est effectuée en trois dimensions ($d = 3$) même si les développements numériques réalisés dans cette thèse seront effectués, pour des raisons de simplicité, en deux dimensions ($d = 2$). $\Gamma(t)$ représente une fissure supposée être une variété $(d-1)$-dimensionnelle plongée dans un espace de dimension d. Par souci de simplicité, $\Gamma(t)$ est ici supposée connexe. On suppose en tout point $y \in \partial\Gamma(t)$ l'existence d'un plan tangent noté T_y. Il existe alors un vecteur unitaire τ, tangent à $\partial\Gamma(t)$ en y, et contenu dans T_y et un vecteur unitaire n, normal à $\partial\Gamma(t)$ en y, contenu dans T_y, et sortant de $\Gamma(t)$ comme représenté dans la figure 1.9. Enfin, le vecteur $N := \tau \wedge n$ est le vecteur normal à T_y et donc normal à $\Gamma(t)$ en y.

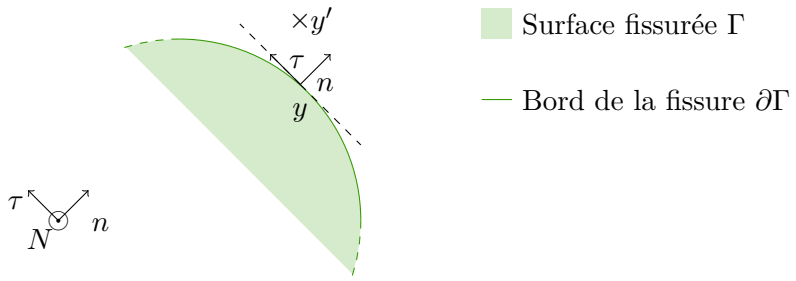


FIGURE 1.9 – Fissure en trois dimensions.

Energie de Griffith

Quasi-statique Le modèle de fissuration de Griffith a été introduit dans [Griffith, 1921]. Il est destiné à prévoir la propagation d'une fissure déjà pré-amorcée selon un tracé prédéfini. Il stipule que la dissipation d'énergie élastique δE_e lors de la création d'une fissure de mesure $\mathcal{H}^{d-1}(\delta\Gamma)$, où $\mathcal{H}^q$ est la mesure de Hausdorff q-dimensionnelle, est proportionnelle à cette nouvelle aire fissurée :

$$\delta E_e = G_c \mathcal{H}^{d-1}(\delta\Gamma). \tag{1.19}$$

Le facteur de proportionnalité noté G_c, et dénommé « taux de restitution d'énergie critique », est un paramètre matériau. On note $A(t) := \mathcal{H}^{d-1}(\Gamma(t))$ la mesure, en fonction du temps, de la surface de fissure. Afin de mesurer la restitution d'énergie potentielle lors de l'avancée de fissure, la fonction G, dénommée « taux de restitution d'énergie », est introduite :

$$G(A) := \frac{\partial E_e}{\partial A}(A). \tag{1.20}$$

Le critère énergétique de fissuration proposé par Griffith se résume alors par l'équation :

$$G(A) = G_c, \tag{1.21}$$

où G dépend du domaine Ω, de la surface de la fissure A et des déformations ε. L'énergie totale dissipée par le processus de fissuration jusqu'à l'instant t est $G_c A(t)$. La formule d'Irwin qui s'écrit

$$G(y) = \frac{1-\nu^2}{E}\left(K_1(y)^2 + K_2(y)^2\right) + \frac{E}{1+\nu}K_3(y)^2, \tag{1.22}$$

permet de relier le taux de restitution d'énergie aux facteurs d'intensité des contraintes (FIC) K_1, K_2 et K_3 qui caractérisent la singularité en bord de fissure. En effet, une fissure dans un domaine Ω crée une singularité géométrique en son bord car le domaine Ω n'est alors plus Lipschitizien en son voisinage. En utilisant les notations de la figure 1.9, la singularité des contraintes est en $(d(y', y))^{-1/2}$ où $d(\cdot, \cdot)$ est la distance euclidienne usuelle de $\mathbb{R}^d$. Les FIC sont alors définis en dimension trois par

$$\begin{cases} K_1(y) := \lim\limits_{y' \to y} \sigma_{NN}(y')\sqrt{2\pi d(y, y')}, \\ K_2(y) := \lim\limits_{y' \to y} \sigma_{Nn}(y')\sqrt{2\pi d(y, y')}, \\ K_3(y) := \lim\limits_{y' \to y} \sigma_{N\tau}(y')\sqrt{2\pi d(y, y')}. \end{cases} \tag{1.23}$$

La figure 1.10 présente les trois modes de fissuration correspondant aux trois facteurs d'intensité des contraintes.

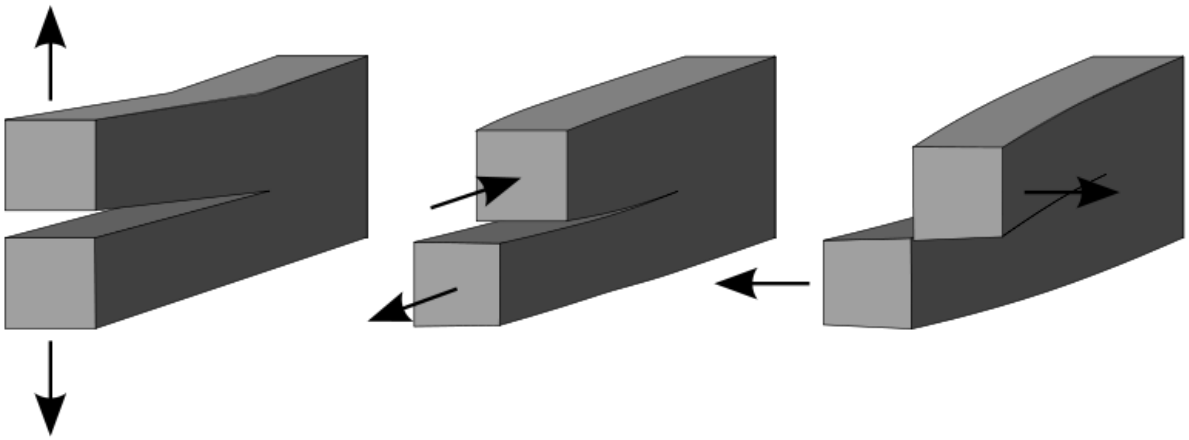


FIGURE 1.10 – Modes de fissuration. Gauche : mode 1 (ouverture). Milieu : mode 2 (cisaillement plan). Droite : mode 3 (cisaillement hors-plan).

Dynamique Freund a généralisé le critère quasi-statique de l'équation (1.21) afin de tenir compte de l'influence des termes inertiels dans le cadre de la fissuration dynamique. Il a alors proposé le critère généralisé :

$$G(A, S) = G_c. \tag{1.24}$$

La différence majeure de (1.24) par rapport à (1.21) consiste en la dépendance supplémentaire du critère de fissuration en la vitesse de fissuration S. L'ouvrage [Freund, 1998] présente un panorama très complet de la fissuration dynamique. Dans ce manuscrit toutefois, les critères tenant compte de la vitesse de fissuration ne seront pas considérés.

Energie cohésive

La fissuration au sens de Griffith ne considère que deux états possibles de la matière : fissurée ou saine. Barenblatt a alors proposé dans [Barenblatt, 1962] une généralisation du modèle de Griffith en introduisant les fissures dites « cohésives ». La généralisation de Barenblatt consiste à ajouter un troisième état possible appelé « endommagé ». Cet état suit généralement une loi d'endommagement. On note $\psi \equiv \psi(d)$ l'énergie surfacique dissipée correspondant à un endommagement $d \in [0, 1]$. Tant que $d = 0$, la surface est saine. Quand $0 < d < 1$, la surface est dite endommagée et quand $d = 1$, la surface est rompue. Une surface ne libère donc plus toute son énergie en rompant immédiatement mais la dissipe au fur et à mesure de son

endommagement. Les énergies cohésives sont une généralisation de l'énergie de Griffith au sens où le modèle de Griffith en est une borne supérieure. En effet, comme indiqué dans la figure 1.11, lorsque qu'une surface $\delta\Gamma$ rompt, elle dissipe $G_c\mathcal{H}^{d-1}(\delta\Gamma)$. Le modèle de Griffith correspond alors à prendre $\psi(d) = G_c$ si $d > 0$ et $\psi(0) = 0$. La supériorité des énergies

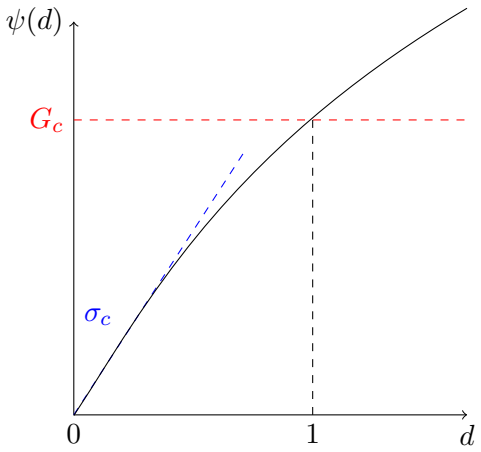


FIGURE 1.11 – Dissipation surfacique d'énergie lors de l'endommagement

cohésives, comparées à l'énergie de Griffith, réside dans leur capacité à décrire l'amorçage de fissure : celui-ci se produit lorsque, localement, on a

$$\sigma = \sigma_c, \tag{1.25}$$

où $\sigma_c = \psi'(0)$ est un paramètre matériau. En revanche, l'énergie de Griffith ne permet pas, dans un cadre conventionnel, de prédire l'amorçage de fissure en l'absence de singularités.

Fragmentation

Diverses approches fondées sur des études statistiques et physiques se proposent de prédire la distribution de taille et le nombre de fragments lors d'une expérience de fragmentation. Une première approche se fonde sur un choix de distribution statistique de défauts, comme par exemple la loi de Weibull dans [Rosin and Rammler, 1933]. Une seconde approche introduite, par exemple, dans [Mott, 1947] et développée dans [Grady and Olsen, 2003] se fonde sur une approche énergétique. Un bilan énergétique, liant énergie libérée par la fissuration et énergie mécanique, permet de calculer le nombre et la taille des fragments en fonction de la densité ρ, la vitesse c des ondes élastiques, du taux de déformation $\dot{\varepsilon}$ et du taux de libération d'énergie critique G_c. L'ouvrage [Asay and Shahinpoor, 2012] résume ces différentes approches et propose de nombreuses comparaisons avec des expériences de fragmentation comme celle de la sous-section 1.1.2. Dans le cadre de la simulation de cette expérience, le taux de déformation est $\dot{\varepsilon} \sim 10^6\text{s}^{-1}$. Or, pour des taux de déformation tels que $\dot{\varepsilon} \geq 10^4\text{s}^{-1}$, le régime énergétique de fragmentation domine le régime des défauts comme discuté dans [Asay and Shahinpoor, 2012, p. 300].

1.3 État de l'art pour la simulation

La simulation d'une expérience telle que celle présentée en sous-section 1.1.2 nécessite trois discrétisations principales. La description de quelques méthodes numériques associées à ces discrétisations compose les trois parties de cette section. La première discrétisation est spatiale et permet de simuler la déformation de la structure étudiée. La deuxième discrétisation est temporelle et s'attache à calculer la dynamique de la structure. Enfin la troisième discrétisation concerne la simulation de la fissuration.

1.3.1 Discrétisation en espace

Cette sous-section commence par présenter la méthode la plus communément utilisée et appelée méthode des éléments finis (MEF). Puis est présentée l'approche retenue dans le cadre de cette thèse et appelée méthode des éléments discrets (MED). Enfin, on présente les méthodes de Galerkine discontinues d'ordre minimal ainsi que la théorie des méthodes de discrétisation de gradients qui permet d'étudier, dans un cadre unifié, les discrétisations présentées.

Méthode des éléments finis

La méthode des éléments finis (MEF) a été développée à partir de la moitié du XX$^{\text{ème}}$ siècle et a été introduite par des ingénieurs mécaniciens afin de réaliser des calculs de structures. Elle a par la suite été formalisée et est devenue la méthode de référence en mécanique numérique des solides. La présentation suivante se place dans le cadre de la résolution de l'équation linéaire générique suivante, où $V \subset H^1(\Omega;\mathbb{R}^d)$ est un espace de Hilbert et un sous-espace vectoriel fermé de l'espace de Sobolev $H^1(\Omega;\mathbb{R}^d)$ des fonctions de carré sommable, ainsi que leurs dérivées partielles, sur un ouvert Ω :

$$\text{Trouver } u \in V, \text{ tel que : } \forall \tilde{u} \in V, \quad a(u,\tilde{u}) = b(\tilde{u}), \tag{1.26}$$

avec a est une forme bilinéaire symétrique définie positive sur $V \times V$ et b une forme linéaire continue sur V. Pour simplifier, on ne considère pas de fonctions dépendantes du temps. Le théorème de Lax-Milgram permet de prouver l'existence et l'unicité d'une solution à l'équation (1.26). Cette dernière étant rarement connue de manière analytique, il est nécessaire d'employer des méthodes numériques, comme la MEF, afin d'approcher cette solution. Le domaine Ω est alors approché par un polyèdre, lui-même formé par une collection de simplexes jointifs, appelé « maillage ». N degrés de liberté (ddl) sont placés à tous les nœuds du maillage pour les éléments finis de type « Lagrange ». u et $\tilde{u}$ se décomposent sous la forme :

$$\forall x \in \Omega, \quad u(x) = \sum_{k=1}^{N} u_k \xi_k(x), \tag{1.27}$$

où $(u_k)_{k=1,\dots,N}$ sont les ddl, qui sont des réels, et $(\xi_k)_{k=1,\dots,N}$ des fonctions dans V, appelées « fonctions de forme ». Les éléments de Lagrange les plus couramment utilisés sont les éléments appelés P^1 et P^2. Les fonctions de forme correspondantes sont illustrées en figure 1.12 dans le cas uni-dimensionnel et à valeurs scalaires. Le choix d'une méthode éléments finis revient, pour

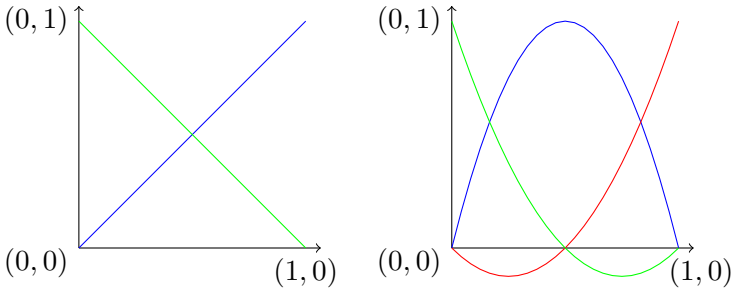


FIGURE 1.12 – Gauche : Fonctions de forme P^1. Droite : Fonctions de forme P^2.

simplifier, à choisir la position des degrés de liberté et les fonctions de forme. En appliquant la décomposition (1.27) à l'équation (1.26) et en notant U le vecteur collection des $(u_k)_{k=1,\dots,N}$, on obtient le système linéaire de $\mathbb{R}^N$:

$$AU = B, \tag{1.28}$$

pour lequel $A = (a(\xi_k, \xi_l))_{k,l=1,\dots,N}$ et $B = (b(\xi_l))_{l=1,\dots,N}$. A est alors une matrice symétrique définie positive et la valeur des ddl contenus dans U est calculée en inversant le système (1.28). Enfin, la solution u de l'équation (1.26) est aussi solution du problème équivalent de minimisation énergétique suivant :

$$\min_{v \in V} E_V(v), \tag{1.29}$$

où $E_V(v) := \frac{1}{2}a(v,v) - b(v)$. Cette propriété est reproduite au niveau discret puisque le vecteur U minimise dans $\mathbb{R}^N$ l'énergie :

$$E_{\mathbb{R}^N}(V) := \frac{1}{2}V^T A V - V^T B. \tag{1.30}$$

Notons que $E_{\mathbb{R}^N}(U) \geq E_V(u)$.

Méthodes des éléments discrets

Les MED ont été introduites dans [Cundall and Strack, 1979] pour simuler le comportement de matériaux granulaires comme par exemple les géomatériaux. De nombreux développements ont suivi et les MED sont aujourd'hui aussi bien utilisées pour simuler le comportement de milieux continus que le comportement d'écoulements granulaires. Des développements récents permettent de proposer deux classifications de MED : les « MED standard » et les « MED variationnelles ».

MED standard L'approche standard des MED consiste à simuler le comportement macroscopique d'un matériau en ayant une connaissance de son comportement mésoscopique. Le matériau est supposé être composé de grains, assimilés à des boules, qui interagissent entre eux à travers les lois de Coulomb du frottement. Le déplacement dans l'échantillon est alors représenté par le déplacement des grains. L'interpénétration des grains est pénalisée par des forces normales F_n et les frottements sont représentés par des forces tangentielles F_t, comme illustré en figure 1.13. Le calcul de la dynamique du système résulte alors de l'intégration du système de particules en utilisant la seconde loi de Newton. Dans la mesure où les grains peuvent être

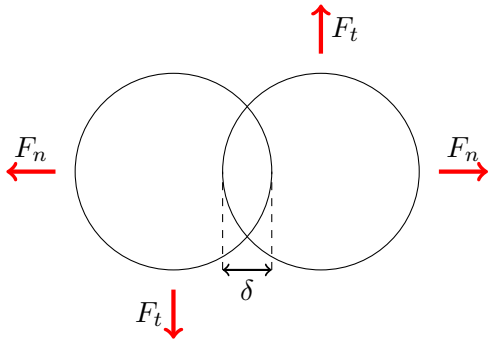


FIGURE 1.13 – Forces dues au contact entre deux éléments discrets

considérés comme des solides rigides, la rotation des éléments peut également être ajoutée comme variable. Les particules interagissent alors également à l'aide de moments purs et de moments résultant de l'application des forces précédentes. Bien que relativement efficaces pour reproduire des comportements de matériaux dont la structure est effectivement granulaire, ce type d'approche permet plus difficilement de simuler le comportement de milieux continus. En effet, les paramètres mésoscopiques doivent alors être ajustés comme dans [Jebahi et al., 2015] ou [André et al., 2013], afin que le comportement macroscopique calculé pour certains essais soit correct. Des études de convergence numérique et une mesure de l'écart-type des paramètres doivent ensuite être réalisées. L'écart-type ne converge cependant pas vers zéro avec la taille

des éléments. Les comportements continus avec des modules de Poisson $\nu \geq 0.3$ sont également difficilement simulables avec ce type d'approche, comme expliqué dans [Jebahi et al., 2015]. Une autre difficulté tient à la représentation du domaine étudié par le biais d'une collection de boules non-jointives comme illustré dans la figure 1.14. Ce genre de représentation ne remplit pas entièrement le volume étudié et la définition d'interface entre éléments n'est pas univoque. De plus, il est difficile de produire ce genre de maillage pour des domaines complexes. Enfin, la reconstruction de grandeurs macroscopiques telles que les déformations et les contraintes, comme dans [André et al., 2013], est délicate.

FIGURE 1.14 – Maillage d'éléments discrets constitués de boules non-jointives [4]

La déformation plastique est généralement prise en compte par l'intermédiaire de patins dissipant de l'énergie et placés entre deux éléments comme dans [Oñate and Rojek, 2004], par exemple. Reproduire les mécanismes de plasticité volumiques comme exprimés dans le cadre de la MMC avec cette approche reste toutefois délicat. L'introduction des MED variationnelles répond à la nécessité de remédier à ces problèmes.

MED variationnelles L'approche des MED variationnelles est, en quelque sorte, inverse à celle des MED standards. Dans le cas des MED variationnelles, le comportement mésoscopique entre grains est inconnu contrairement au comportement du matériau continu que l'on souhaite simuler. L'objectif est alors de dériver les interactions entre particules à partir uniquement des paramètres continus.

Une première formulation de MED discrétisant l'élasticité linéaire en 2d et utilisant des cellules de Voronoi afin d'occuper tout le volume de l'échantillon est proposée dans [Mariotti, 2007] et un exemple de maillage est présenté en figure 1.15. Les ddl en déplacements se trouvent au barycentre des cellules. Des ddl en rotation sont également utilisés et placés au barycentre des cellules. Par conséquent, les particules interagissent par le biais de forces et de moments. Les quantités mésoscopiques doivent être dérivées à partir des quantités macroscopiques. Cependant, la preuve de convergence formelle de la méthode n'est pas aisée. Une première tentative de formalisation de la convergence de ce type de méthodes est proposée dans [Monasse and Mariotti, 2012] mais reste incomplète.

D'autres approches visant à dériver une MED à partir d'une formulation variationnelle peuvent être citées comme [Notsu and Kimura, 2014]. Dans cette approche, la méthode dérive d'élements finis de Lagrange P^1. Une preuve de convergence est disponible. Cette approche souffre cependant d'une limitation en termes de module de Poisson. En effet, afin de conserver la positivité de l'énergie de chaque lien entre éléments, une restriction sur ν est imposée sous la forme $\lambda < \mu$ où λ et μ sont les coefficients de Lamé du matériau.

4. `https://hal.archives-ouvertes.fr/hal-01006009/document`

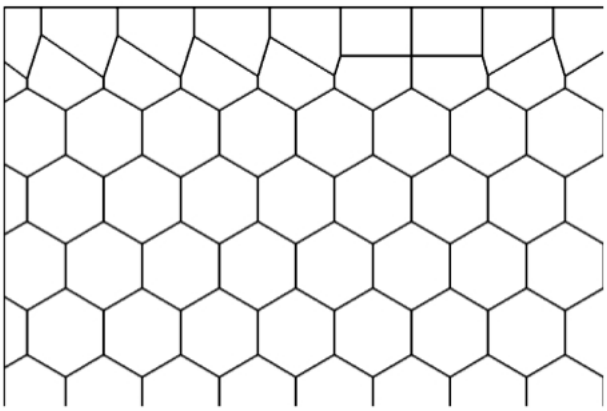


FIGURE 1.15 – Maillage de Voronoi [Mariotti, 2007]

Méthodes de discrétisation de gradient

Les méthodes de Galerkine discontinues d'ordre minimal étudiées dans [Di Pietro, 2012] ainsi que les méthodes de type volumes finis introduites dans [Eymard et al., 2004] et [Eymard et al., 2009], partagent avec les MED variationnelles le fait d'avoir des ddl au barycentre des cellules. La nouveauté, par rapport aux MED décrites ci-dessus, consiste dans l'utilisation de reconstructions affines pour calculer les déformations au sein de chaque maille. La stabilité, le caractère bien posé ainsi que la convergence de ces méthodes peuvent être étudiés dans le cadre de la théorie des méthodes de discrétisation de gradients introduite dans [Droniou et al., 2010] et [Droniou et al., 2013]. [Droniou et al., 2018] est un ouvrage de référence concentrant la plupart des résultats actuellement disponibles. Enfin, une application des méthodes de discrétisation de gradient à l'élasticité linéaire et à quelques modèles non-linéaires élastiques est développée dans [Droniou and Lamichhane, 2015].

1.3.2 Discrétisation en temps

Une fois la discrétisation en espace réalisée, la variable temporelle doit également être discrétisée. Cela correspond à discrétiser les EDO résultant de la discrétisation en espace. Les intégrateurs temporels utilisés dans le cadre de la dynamique rapide sont généralement explicites et une condition CFL s'applique alors sur le pas de temps. Il est également souhaitable d'utiliser des intégrateurs qui permettent de conserver l'énergie du système semi-discrétisé en espace.

Intégrateurs explicites

Les comportements des matériaux simulés dans le cadre de la dynamique rapide sont couramment non-linéaires, et les intégrateurs temporels utilisés sont explicites. Ces derniers sont préférés aux intégrateurs implicites dans la mesure où, comme les comportements simulés sont non-linéaires, les pas de temps utilisables avec des méthodes implicites sont du même ordre que ceux des méthodes explicites afin d'assurer la convergence des itérations de point fixe ou de Newton. Les intégrateurs temporels utilisés sont généralement des schémas symétriques réversibles d'ordre 2 à un pas. La méthode de Störmer –Verlet, encore appelée Leapfrog, est par exemple utilisée dans [Mariotti, 2007] et [Monasse and Mariotti, 2012]. Le schéma aux différences centrées, encore appelée Newmark explicite ($\beta = 0$ et $\gamma = \frac{1}{2}$), est, par exemple, utilisé dans [Noels and Radovitzky, 2008] et [Camacho and Ortiz, 1996] et est une reformulation de la méthode Leapfrog. Toute ces méthodes étant seulement conditionnellement stables, la gestion de la condition CFL nécessite une attention particulière. En effet, une CFL n'est connue que

dans le cas linéaire. Dans le cas non-linéaire, il est parfois nécessaire de diminuer substantiellement le pas de temps afin de réaliser un calcul stable.

Conservation de l'énergie semi-discrétisée en espace

On se replace dans le cadre de la sous-section 1.3.1 et on conserve les mêmes notations. On considère la version instationnaire du problème (1.26). Pour simplifier, supposons que la solution exacte u soit telle que

$$u \in W = \left\{ v \in L^2([0,T];V), \ddot{v} \in L^2([0,T];V) \right\}.$$

La version instationnaire du problème (1.26) est alors

$$\text{Trouver } u \in W, \text{ tel que dans } L^2([0,T]),\ \forall \tilde{u} \in V, \quad m(\ddot{u}(t),\tilde{u}) + a(u(t),\tilde{u}) = b(\tilde{u}), \tag{1.31}$$

où $m(\ddot{u}(t),\tilde{u}) := \int_\Omega \rho \ddot{u}(t) \cdot \tilde{u}$. Si l'on suppose que a et b ne dépendent pas du temps, alors u présente une propriété de conservation de l'énergie, qui s'écrit sous la forme suivante :

$$\frac{d}{dt} E_W(u(t)) = 0, \quad \forall t \in [0,T], \tag{1.32}$$

où E_W est définie par

$$E_W(u(t)) := \frac{1}{2} m(\dot{u}(t),\dot{u}(t)) + \frac{1}{2} a(u(t),u(t)) - b(u(t)), \quad \forall t \in [0,T]. \tag{1.33}$$

On en déduit alors que

$$E_W(u(t)) = E_{W,0}, \quad \forall t \in [0,T], \tag{1.34}$$

où $E_{W,0} := \frac{1}{2} m(v_0,v_0) + \frac{1}{2} a(u_0,u_0) - b(u_0)$ et u_0 et v_0 sont les déplacements et vitesses initiaux. Cette propriété est reproduite au niveau semi-discret en espace (le temps restant continu), puisque l'on a

$$E_{\mathbb{R}^N}(U(t)) = E_{\mathbb{R}^N,0}, \quad \forall t \in [0,T], \tag{1.35}$$

où

$$E_{\mathbb{R}^N}(U(t)) := \frac{1}{2} \dot{U}^T(t) M \dot{U}(t) + \frac{1}{2} U^T(t) A U(t) - U^T(t) B, \quad \forall t \in [0,T], \tag{1.36}$$

$E_{\mathbb{R}^N,0} := \frac{1}{2} V_0^T M V_0 + \frac{1}{2} U_0^T A U_0 - U_0^T B$, $M := (\int_\Omega \xi_k \cdot \xi_l)_{k,l}$ et U_0 et V_0 sont les vecteurs composantes de u_0 et v_0 dans l'espace des fonctions de forme.

Conservation de l'énergie discrétisée en espace-temps

La conservation de l'énergie mécanique totale est une propriété importante des systèmes continus et de leur semi-discrétisation en espace, comme présenté en équation (1.36). Il est alors nécessaire de conserver cette propriété au cours de l'intégration temporelle afin de prédire correctement l'évolution du système. Les schémas explicites mentionnés ci-dessus conservent tous une énergie discrète dans le cas linéaire, lorsqu'ils sont utilisés à pas de temps constant, comme décrit dans [Belytschko and Hughes, 1983]. Ce n'est cependant plus le cas pour un système non-linéaire, c'est-à-dire un système avec une énergie non-quadratique, ou pour un système linéaire ou non-linéaire intégré avec un pas de temps variable. Pour les systèmes raidissants, les difficultés se conjuguent car le raidissement impose une diminution du pas de temps. La stabilité du calcul est donc difficile à assurer et oblige à considérer des pas de temps très faibles.

Des méthodes implicites qui conservent l'énergie mécanique pour des systèmes non-linéaires ont été développées par exemple dans [Quispel and McLaren, 2008] dans le cadre des systèmes dynamiques, [Chabassier and Joly, 2010] dans le cadre des systèmes hyperboliques non-linéaires non-dégénérés et [Hager et al., 2012] dans le cadre de la mécanique du contact. La conservation de l'énergie mécanique est fondamentale dans le cadre des simulations de fragmentation car, comme indiqué dans la sous-section 1.2.3, la taille des fragments dépend en général de la quantité d'énergie mécanique disponible et du paramètre G_c. Étant implicites, ces méthodes sont plus coûteuses par pas de temps que les méthodes explicites mentionnées précédemment.

1.3.3 Fissuration et fragmentation

La fissuration et la fragmentation ont fait l'objet de nombreux traitements numériques. Cette section se limite à mentionner quelques méthodes qui possèdent des similarités avec les méthodes développées dans ce manuscrit.

Modèles à zones cohésives

Des modèles à zone cohésive ont été utilisés en complément d'éléments finis pour effectuer des calculs de fissuration dynamique, comme par exemple dans [Camacho and Ortiz, 1996] ou [Doyen et al., 2013]. Une approche similaire a été utilisée dans [Mariotti et al., 2009] dans le cas des MED, en utilisant la MED variationnelle de [Mariotti, 2007] pour la discrétisation de l'élasticité. Les modèles cohésifs de fissuration sont cependant encore mal compris, comme montré dans [Braides et al., 1999], et posent de nombreuses questions théoriques non résolues. Concernant les MED standards, le critère de fissuration adopté dans [André et al., 2013] et [André et al., 2019] correspond à celui d'amorçage présenté dans l'équation (1.25). Une loi cohésive n'est cependant pas utilisée pour incrémenter une variable d'endommagement et, par conséquent, un contrôle sur l'énergie dissipée dans la fissure n'est pas disponible.

Champ de phase

Dans [Francfort and Marigo, 1998], une reformulation variationnelle de la fissuration quasi-statique de Griffith, présentée en sous-section 1.2.3, est adoptée. De nombreux travaux, dont [Chambolle and Crismale, 2019], ont permis de caractériser les solutions de cette formulation afin d'obtenir une caractérisation mathématique précise des solutions des équations de la fissuration quasi-statique. Dans [Bourdin et al., 2000] est introduite la méthode à champ de phase pour calculer la propagation quasi-statique de fissures. Cette méthode permet également de calculer l'amorçage de fissures. La convergence formelle de cette méthode est le fruit de nombreux travaux, dont [Chambolle, 2004] qui traite le cas de l'élasticité linéaire 2d. Enfin, la méthode à champ de phase a été adaptée pour calculer la propagation de fissures dynamique notamment dans [Li et al., 2016] et [Li and Marigo, 2017]. Malgré d'excellents résultats, cette méthode a cependant l'inconvénient de nécessiter l'inversion d'un système global à chaque pas de temps de l'intégration explicite afin de faire évoluer le champ de phase. Ce coût de calcul significatif peut limiter fortement la taille en espace et en temps des systèmes simulables.

Méthodes sans maillage

Les MED standards sont généralement qualifiées, de façon discutable, de méthodes « sans maillage », car elles n'utilisent pas un maillage en tant que tel. Une autre méthode, qualifiée de « sans maillage », développée pour effectuer des calculs de fissuration dynamique, est la méthode des éléments libres, présentée par exemple dans [Belytschko et al., 1995]. Cette méthode estime les FIC et calcule ensuite le taux de restitution d'énergie (1.21) à travers l'utilisation de la formule d'Irwin (1.22). Afin de capturer correctement la singularité, des techniques de remaillage dynamique et d'analyse a posteriori peuvent être utilisées, comme dans [Stone and Babuška, 1998] par exemple. Leur usage a cependant un coût calculatoire important car cette méthode nécessite un recalcul des reconstructions des fonctions à chaque étape de l'intégration temporelle. Une autre difficulté apparaît dans la gestion des degrés de liberté qui ont été traversés par la fissure et qui doivent alors être dédoublés et placés sur le bord de la fissure.

1.4 Plan du manuscrit et contributions

Le reste de ce manuscrit est structuré en trois chapitres représentant les trois principales contributions de ce travail de thèse et un chapitre de conclusion et perspectives. Les trois

principales contributions sont exposées dans chacune des sous-sections suivantes.

1.4.1 Intégration temporelle et conservation de l'énergie

La conservation de l'énergie mécanique est une propriété fondamentale des systèmes continus et discrets étudiés. Par exemple, la répartition de taille des fragments générés dans la simulation de l'expérience de la sous-section 1.1.2 dépend directement de la quantité d'énergie disponible pour être dissipée dans des fissures. Or, comme évoqué dans la sous-section 1.3.2, la discrétisation temporelle est réduite à des intégrateurs explicites pour des raisons de coûts des calculs. Le comportement raidissant en compression interdit l'utilisation de la méthode Leap-frog et des différences centrées à cause de la variabilité du pas de temps qui retire leur caractère conservatif à ces méthodes. Une méthode explicite conservative en énergie est alors développée dans le chapitre 2 pour des systèmes Hamiltoniens. Cette dernière conserve une pseudo-énergie, y compris pour des comportements non-linéaires. Elle est également assortie d'un critère de stabilité a posteriori qui permet une gestion facilitée de la stabilité du schéma pour un système raidissant par rapport aux schémas décrits dans la sous-section 1.3.2. Ce chapitre ne traite pas directement le raidissement tel que décrit en sous-sections 1.1.3 et 1.2.2 mais traite tout de même un autre système non-linéaire raidissant issu de [Chabassier and Joly, 2010].

Cette contribution a fait l'objet d'une publication dans le journal Computer Methods in Applied Mechanics and Engineering [Marazzato et al., 2019a].

1.4.2 MED variationnelle pour l'élasto-plasticité

Une MED variationnelle discrétisant l'élasto-plasticité HPP est développée dans le chapitre 3. Cette méthode suit une approche similaire à [Mariotti, 2007] et [Monasse and Mariotti, 2012]. Les ddl en rotation des particules sont supprimés car ces derniers ne sont pas nécessaires pour la discrétisation d'un matériau de Cauchy. Une première contribution de la méthode développée au Chapitre 3 est de s'appliquer à des maillages polyédriques généraux et non plus de Voronoi comme dans les travaux cités précédemment. Cela permet notamment l'utilisation de maillages tétraédriques dont l'utilisation pour mailler des domaines complexes est plus souple. Cette extension est rendue possible grâce à l'utilisation de reconstructions de gradients issues des méthodes de type volumes finis, comme celle développée dans [Eymard et al., 2009] ou des méthodes de type Galerkine discontinues, comme celle développée dans [Di Pietro, 2012]. La deuxième contribution de ce chapitre est d'étendre la MED de [Mariotti, 2007] et [Monasse and Mariotti, 201 à la plasticité. L'idée est d'ajouter des ddl volumiques pour la plasticité et de les traiter de manière volumique comme cela est traditionnellement fait en MEF et non plus de façon surfacique, comme cela est traditionnellement fait en MED. Enfin l'intégrateur temporel développé dans le chapitre précédent est couplé à la MED développée dans ce chapitre afin d'être appliqué à des calculs d'élasto-plasticité dynamique.

Une première version de cette contribution a fait l'objet d'une publication dans les actes de Congrès Français de Mécanique [Marazzato et al., 2019c]. Une version augmentée a été soumise pour publication dans le journal International Journal for Numerical Methods in Engineering [Marazzato et al., 2019b] et est en cours de révision.

1.4.3 MED variationnelle pour la fissuration dynamique de Griffith

Une méthode de calcul de fissures de Griffith en deux dimensions d'espace a été développée dans le chapitre 4, fondée sur la MED variationnelle présentée dans le chapitre 3. Le choix d'une modélisation des fissures par une énergie de Griffith à la place d'une énergie cohésive, comme utilisée dans [Mariotti et al., 2009], est justifié par la meilleure maîtrise mathématique du modèle de Griffith avec les connaissances actuelles. L'approche utilisée permet de formaliser,

dans le cadre du modèle de Griffith, l'approche utilisée dans [Mariotti et al., 2009], qui consiste à laisser les fissures se propager à travers les facettes du maillage, et de conserver la localité du critère de fissuration. En effet, contrairement à la méthode à champ de phase développée dans [Li et al., 2016] et [Li and Marigo, 2017], le critère de fissuration considéré ici est purement local et ne nécessite donc pas l'inversion d'une matrice à chaque pas de temps de l'intégration explicite. Le taux de restitution d'énergie est calculé pour chaque mode de fissuration à partir des facteurs d'intensités des contraintes qui sont approchés près de la fissure. Par la suite, un critère de maximisation de la densité d'énergie élastique sur les facettes liées à la pointe de fissure permet de simuler l'orientation de la propagation. Enfin, l'approche développée dans le chapitre 4, bien que similaire à [Belytschko et al., 1995], conduit à un coût numérique plus contenu car le recalcul des reconstructions des déplacements n'est nécessaire que lorsque la fissure se propage.

Cette contribution fait l'objet d'un article soumis pour publication au journal Computer Methods in Applied Mechanics and Engineering.

Chapter 2

An explicit pseudo-energy conserving time-integration scheme for Hamiltonian dynamics

Contents

The material in this chapter has been published in [Marazzato et al., 2019a].

We propose a new explicit pseudo-energy and momentum conserving scheme for the time integration of Hamiltonian systems. The scheme, which is formally second-order accurate, is based on two key ideas: the integration during the time-steps of forces between free-flight particles and the use of momentum jumps at the discrete time nodes leading to a two-step formulation for the acceleration. The pseudo-energy conservation is established under exact force integration, whereas it is valid to second-order accuracy in the presence of quadrature errors. Moreover, we devise an asynchronous version of the scheme that can be used in the framework of slow-fast time-stepping strategies. The scheme is validated against classical benchmarks and on nonlinear or inhomogeneous wave propagation problems.

2.1 Introduction

Energy and momentum conservation is an important property of numerical schemes for a large number of physical problems. For instance, in statistical physics, accurately conserving first integrals constitutes a fundamental requirement to capture the correct behaviour of the system. In mechanics, conservation of the mechanical energy (together with momentum) is an important feature for systems such as the acoustics in a piano [Chabassier and Joly, 2010] or nonlinear contact dynamics [Hauret and Le Tallec, 2006, Fetecau et al., 2003]. In this work,

we consider Hamiltonian systems consisting of N particles in dimension d (typically, $d = 1$, 2 or 3) where $q_i, p_i \in \mathbb{R}^d$ are the position and momentum of the particle $i \in \{1, \dots, N\}$. We assume that the Hamiltonian has the following split form:

$$H(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{M}^{-1}\boldsymbol{p} + V(\boldsymbol{q}), \tag{2.1}$$

where $\boldsymbol{q} = (q_1, \dots, q_N) \in \mathbb{R}^{dN}$ is the position vector of the particles, $\boldsymbol{p} = (p_1, \dots, p_N) \in \mathbb{R}^{dN}$ is the momentum vector of the particles, $\boldsymbol{M}$ is the symmetric positive definite mass matrix and V is the potential energy. The system is thus driven by the equations

$$\dot{\boldsymbol{q}} = \boldsymbol{M}^{-1}\boldsymbol{p}, \qquad \dot{\boldsymbol{p}} = -\nabla V(\boldsymbol{q}). \tag{2.2}$$

Several approaches have been proposed to tackle the issue of conservation when integrating numerically (2.2). A first possibility consists in the use of symplectic schemes [Hairer et al., 2006], which integrate a modified (not explicitly exhibited, except in certain simple cases) Hamiltonian and thus preserve the first integrals of the dynamics over exponentially long times (with respect to the time-step), up to fluctuations whose amplitudes grow with the time-step. However, in the case of variable time-steps, symplectic schemes lose their conservation properties since the modified Hamiltonian changes with the time-step [Calvo and Sanz-Serna, 1993]. When the time-step size is driven by the shape of the Hamiltonian (e.g. in Kepler's problem with high eccentricity), a workaround consists in adding a perturbation accounting for the time-step variation in order for a rescaled dynamic to remain Hamiltonian [Hairer, 1997]. In practice, for mechanical problems, such a condition on the time-step can become impractical, since the time-step could be imposed due to coupling or stiffness phenomena not accounted for in the Hamiltonian part. For an extended review of variational integrators in mechanics, we refer the reader to [Marsden and West, 2001]. Another approach consists in imposing the exact conservation of energy and momentum at each step of the numerical scheme. Integrating on the constant energy manifold can be carried out using projection [Hughes et al., 1978] or Lie group integration [Iserles et al., 2000], but these methods are computationally expensive as soon as the manifold of constant energy and momentum has a complex shape. Another class of methods, energy-momentum conserving schemes, have been proposed in [Simo and Oliver, 1994, Gonzalez and Simo, 1996, Hauret and Le Tallec, 2006, Chabassier and Joly, 2010] for nonlinear mechanics, contact mechanics and nonlinear wave equations, among others. The general principle is to integrate the nonlinear forces at a special time during the time-step, which is determined through a nonlinear implicit procedure. A higher-order version of these implicit schemes has been derived for linear wave propagation in [Chabassier and Imperiale, 2013]. Variational integrators combining features of symplectic and energy-momentum schemes have been developed for variable time-step strategies [Kane et al., 1999a] and nonlinear mechanical problems in [Groß et al., 2005].

To the best of our knowledge, no explicit pseudo-energy conserving scheme has been proposed to date for nonlinear problems. With the motivation that explicit schemes often result in greater computational efficiency, the goal of the present work is to develop such an explicit scheme for nonlinear mechanics, where pseudo-energy conservation holds exactly for exact force integration and up to second-order accuracy in the presence of quadratures. The present scheme hinges on two key ideas. The first one, already considered in [Mariotti, 2016], is to approximate the dynamics of the particles by free-flight trajectories during each time-step. The second one is to use momentum jumps at the discrete time nodes to approximate the acceleration. In doing so, we circumvent the negative result on the existence of explicit schemes in [Chabassier and Joly, 2010, Lemma 3.3] through the use of a two-step strategy. This idea has some links with the implicit energy-conserving average vector field method [Quispel and McLaren, 2008] where the conservation of the Hamiltonian is formulated using an implicit integral of the forces derived from the potential V over the time-step. A high-order generalization of the average vector field method using collocation has been developed

in [Hairer, 2010]. The present numerical scheme shares with average vector field methods the salient feature of average force integration over each time-step. However, the two schemes differ on the discretisation of the acceleration, which is based here on momentum jumps.

A further development of the present work is to devise an asynchronous version of our scheme that lends itself to slow-fast decompositions as presented in [Hairer et al., 2006], with the goal to further reduce the computational cost of the simulation. In the case of mechanical systems with local stiffness, the conditional stability of an explicit time-integration scheme typically involves small time-steps for the whole system. A promising direction to mitigate this drawback consists in using a local time-stepping strategy. In the linear case, explicit high-order energy-momentum conserving methods with local time-stepping have been proposed in [Diaz and Grote, 2009]. In the nonlinear case, a modified Störmer–Verlet method for Hamiltonian systems containing slow and fast components is developed in [Hairer et al., 2006]. It is proved that this time-integrator remains symplectic, but the ratio of the fast and slow time-steps strongly influences the error on the total energy and, in general, a good balance has to be found experimentally. This phenomenon is called resonance since it is encountered for certain slow/fast ratios. Similarly, asynchronous variational integrators generally exhibit resonances when the local time-steps are close to certain rational ratios, so that ensuring stability requires adequate fitting of the local time-steps [Fong et al., 2008]. In contrast, the asynchronous version of the present scheme allows one to make slow-fast time-integration while conserving a pseudo-energy (in the absence of quadrature errors). Our numerical tests show that the asynchronous scheme still exhibits second-order accuracy; a mathematical proof of this property is postponed to future work.

This paper is organized as follows. In Section 2.2, we present the scheme for a Hamiltonian system of interacting particles with a synchronous time-integration and establish the main properties of the scheme including second-order accuracy, time-reversibility, linear stability under a CFL condition, and pseudo-energy conservation under exact force integration. In Section 2.3, we test the synchronous scheme on various benchmarks from the literature including a nonlinear wave propagation problem. In Section 2.4, we present the slow-fast time-stepping capabilities of the asynchronous version of the scheme, together with numerical results on model particle systems connected by springs and on an inhomogeneous wave equation. These results demonstrate the efficiency gains of the asynchronous scheme with respect to the synchronous scheme.

2.2 Synchronous scheme

In this section, we present our scheme in its synchronous version and establish its main properties.

2.2.1 Definition of the scheme

We consider a sequence of discrete time nodes t^n, $n = 0, 1, \ldots$, with time-steps $h_n = t^{n+1} - t^n$ and time intervals $I_n = [t^n, t^{n+1}]$. The scheme is written at step n as follows: knowing $\boldsymbol{p}^{n-1/2}$, $\boldsymbol{q}^n$, and $[\boldsymbol{p}]^n$, one computes

$$\boldsymbol{p}^{n+1/2} = \boldsymbol{p}^{n-1/2} + [\boldsymbol{p}]^n, \tag{2.3a}$$

$$\boldsymbol{q}^{n+1} = \boldsymbol{q}^n + h_n \boldsymbol{M}^{-1} \boldsymbol{p}^{n+1/2}, \tag{2.3b}$$

$$\frac{1}{2}\left([\boldsymbol{p}]^{n+1} + [\boldsymbol{p}]^n\right) = -\int_{I_n} \nabla V(\hat{\boldsymbol{q}}^n(t))dt, \tag{2.3c}$$

with the free-flight trajectory over the time interval I_n defined by

$$\hat{\boldsymbol{q}}^n(t) = \boldsymbol{q}^n + (t - t^n)\boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2} \quad \forall t \in I_n. \tag{2.4}$$

Here, $[\boldsymbol{p}]^n$ represents the jump of the momentum vector at time t^n, $\boldsymbol{q}^n$ the position vector at time t^n, and $\boldsymbol{p}^{n+1/2}$ is the momentum vector between t^n and t^{n+1}. We observe that $\boldsymbol{q}^{n+1} = \hat{\boldsymbol{q}}^n(t^{n+1})$. We initialize the scheme as follows:

$$\boldsymbol{p}^{-1/2} = \boldsymbol{p}(t^0), \quad \boldsymbol{q}^0 = \boldsymbol{q}(t^0), \quad [\boldsymbol{p}]^0 = \boldsymbol{0}, \tag{2.5}$$

where $\boldsymbol{q}(t^0), \boldsymbol{p}(t^0)$ are the given position and momentum vectors at the initial time t^0. The scheme (2.3) can alternatively be written as the following 2-step scheme without jumps: knowing $\boldsymbol{p}^{n-1/2}$, $\boldsymbol{q}^n$, and $\boldsymbol{p}^{n+1/2}$, one computes

$$\boldsymbol{q}^{n+1} = \boldsymbol{q}^n + h_n \boldsymbol{M}^{-1} \boldsymbol{p}^{n+1/2}, \tag{2.6a}$$

$$\frac{1}{2}\left(\boldsymbol{p}^{n+3/2} - \boldsymbol{p}^{n-1/2}\right) = -\int_{I_n} \nabla V(\hat{\boldsymbol{q}}^n(t))dt, \tag{2.6b}$$

with the free-flight trajectory defined by (2.4). The initialization of the scheme, equivalent to (2.5), is as follows:

$$\boldsymbol{p}^{-1/2} = \boldsymbol{p}(t^0), \quad \boldsymbol{q}^0 = \boldsymbol{q}(t^0), \quad \boldsymbol{p}^{1/2} = \boldsymbol{p}(t^0). \tag{2.7}$$

This initialization is tailored to achieve exact pseudo-energy conservation under exact force integration, as shown in Theorem 1 below. Other choices for the initialization are possible, for instance using a one-step method.

In the numerical implementation of the scheme, the integral in (2.3c) (or in (2.6b)) is usually not computed exactly but with a quadrature of the form

$$Q_n(f(t); t^n; t^{n+1}) = h_n \sum_{i=0}^{I} \omega_i f(\lambda_i t^n + (1-\lambda_i) t^{n+1}) \approx \int_{I_n} f(t)dt, \tag{2.8}$$

where the real numbers ω_i are the weights and the real numbers $\lambda_i \in [0,1]$ define the quadrature points. Applying the quadrature componentwise for the calculation of the forces and exploiting that the position of the particles varies linearly in time during the free flight, we obtain

$$Q_n\left(\nabla V(\hat{\boldsymbol{q}}^n(t)); t^n; t^{n+1}\right) = h_n \sum_{i=0}^{I} \omega_i \nabla V(\lambda_i \boldsymbol{q}^n + (1-\lambda_i)\boldsymbol{q}^{n+1})) \approx \int_{I_n} \nabla V(\hat{\boldsymbol{q}}^n(t))dt, \tag{2.9}$$

and we replace (2.3c) with

$$\frac{1}{2}\left([\boldsymbol{p}]^{n+1} + [\boldsymbol{p}]^n\right) = -Q_n\left(\nabla V(\hat{\boldsymbol{q}}^n(t)); t^n; t^{n+1}\right), \tag{2.10}$$

and a similar modification for (2.6b). In what follows, we assume that the quadrature is symmetric:

$$\forall i \in \{0, \dots, I\}, \quad \omega_i = \omega_{I-i} \quad \text{and} \quad \lambda_i = 1 - \lambda_{I-i}, \tag{2.11}$$

and at least of order two (i.e., that the quadrature integrates exactly affine polynomials). We also assume that V is of class C^2, i.e., $V \in C^2(\mathbb{R}^{dN}; \mathbb{R})$. This implies that

$$Q_n\left(\nabla V(\hat{\boldsymbol{q}}^n(t)); t^n; t^{n+1}\right) = \int_{I_n} \nabla V(\hat{\boldsymbol{q}}^n(t))dt + \mathcal{O}(h_n^3). \tag{2.12}$$

2.2.2 Properties of the scheme

We now establish various properties of the scheme: pseudo-energy conservation (in the absence of quadrature errors), symmetry (or time-reversibility), second-order accuracy, and linear stability (with constant time-step).

Theorem 1 (Pseudo-energy conservation)**.** *Assume that the quadrature is exact. Then, the scheme (2.3) exactly conserves the following pseudo-energy:*

$$\tilde{H}^n := V(\boldsymbol{q}^n) + \frac{1}{2}\left(\boldsymbol{p}^{n-1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2}. \tag{2.13}$$

Moreover, denoting $H^0 := H(\boldsymbol{q}(t^0), \boldsymbol{p}(t^0))$ the value of the exact Hamiltonian at the initial time time t^0, we have $\tilde{H}^n = H^0$ for all $n \geq 0$ if the scheme is initialized using (2.5).

Proof. Using (2.3.a), (2.4), and the chain rule, we obtain

$$\frac{d}{dt}\left(V(\hat{\boldsymbol{q}}^n(t))\right) = \nabla V\left(\hat{\boldsymbol{q}}^n(t)\right) \cdot (\hat{\boldsymbol{q}}^n)'(t) = \left(\boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2}\right)^{\mathrm{T}} \nabla V\left(\hat{\boldsymbol{q}}^n(t)\right).$$

Integrating in time and using (2.3c) and the symmetry of $\boldsymbol{M}$, we infer that

$$\begin{aligned} V(\boldsymbol{q}^{n+1}) - V(\boldsymbol{q}^n) &= \left(\boldsymbol{p}^{n+1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1} \int_{t^n}^{t^{n+1}} \nabla V(\hat{\boldsymbol{q}}^n(t))dt \\ &= -\left(\boldsymbol{p}^{n+1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1} \frac{1}{2}\left([\boldsymbol{p}]^{n+1} + [\boldsymbol{p}]^n\right) \\ &= -\left(\boldsymbol{p}^{n+1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1} \frac{1}{2}\left(\boldsymbol{p}^{n+3/2} - \boldsymbol{p}^{n-1/2}\right). \end{aligned}$$

This leads to

$$V(\boldsymbol{q}^{n+1}) + \frac{1}{2}\left(\boldsymbol{p}^{n+1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1}\boldsymbol{p}^{n+3/2} = V(\boldsymbol{q}^n) + \frac{1}{2}\left(\boldsymbol{p}^{n-1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2},$$

showing that $\tilde{H}^{n+1} = \tilde{H}^n$, thereby proving the first assertion. Finally, using the initialization (2.5), we obtain $\tilde{H}^0 = H^0$, and this concludes the proof. □

Remark 2.1 (Quadratures)**.** *In practice, the integral in Equation (2.3c) can be computed exactly only for polynomial potentials V. For instance, using the n-point Gauss–Lobatto (resp., Gauss–Legendre) quadrature, polynomials of degree up to $2n-3$ (resp., $2n-1$) are integrated exactly. The use of quadratures instead of exact integration for general nonlinear potentials entails only an approximate conservation of the pseudo-energy. Since the scheme is second-order accurate (see Theorem 2 below), we expect that pseudo-energy conservation is second-order accurate at best:*

$$\tilde{H}^n = \tilde{H}^0 + O(h^2). \tag{2.14}$$

where $h := \sup_n h_n$. The order of the quadrature has an influence on the multiplicative constant in $O(h^2)$, with the constant (swiftly) decreasing when increasing the quadrature order. Numerical results are presented on a nonlinear wave propagation problem in Section 2.3.3.

Remark 2.2 (Momentum conservation)**.** *Let $(1, ..., 1)$ be the vector of size dN filled with ones. Assume that the system is isolated, i.e., $(1, .., 1)^{\mathrm{T}} \cdot \nabla V(\boldsymbol{q}) = \boldsymbol{0}$ for all $\boldsymbol{q} \in \mathbb{R}^{dN}$. Then, the total momentum, defined as $P_{n+1/2} := (1, .., 1)^{\mathrm{T}} \cdot \boldsymbol{p}_{n+1/2}$ for all $n \geq 0$, is conserved. This follows by taking the product of $(1, ..., 1)^{\mathrm{T}}$ with Equation (2.6b) and using the null initialisation of the momentum jump which follows from Equation (2.5).*

Proposition 2.3 (Symmetry)**.** *If the quadrature (2.9) is exact or symmetric, then the scheme (2.6) is symmetric (or time-reversible).*

Proof. Let $\boldsymbol{Y}^n = (\boldsymbol{q}^n, \frac{\boldsymbol{p}^{n-1/2}+\boldsymbol{p}^{n+1/2}}{2}, \boldsymbol{p}^{n+1/2} - \boldsymbol{p}^{n-1/2})^{\mathrm{T}}$. Since we are going to consider positive and negative time-steps in this proof, we denote by $\mathrm{sign}(h_n)$ the sign of the time-step. The

numerical scheme can be written as $\boldsymbol{Y}^{n+1} = \boldsymbol{\Phi}_{h_n}(\boldsymbol{Y}^n)$, where for a generic column vector $\boldsymbol{Y} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3)^{\mathrm{T}}$, we have

$$\boldsymbol{\Phi}_{h_n}(\boldsymbol{Y}) = \begin{pmatrix} \boldsymbol{Y}_1 + h_n \boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 + \dfrac{\mathrm{sign}(h_n)}{2}\boldsymbol{Y}_3\right) \\ \boldsymbol{Y}_2 - Q_n\left(\nabla V\left(\boldsymbol{Y}_1 + t\boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 + \dfrac{\mathrm{sign}(h_n)}{2}\boldsymbol{Y}_3\right)\right); 0; h_n\right) \\ -\boldsymbol{Y}_3 - 2\mathrm{sign}(h_n) Q_n\left(\nabla V\left(\boldsymbol{Y}_1 + t\boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 + \dfrac{\mathrm{sign}(h_n)}{2}\boldsymbol{Y}_3\right)\right); 0; h_n\right) \end{pmatrix},$$

where we used the invariance by translation of the quadrature Q_n. Therefore, we have

$$\boldsymbol{\Phi}_{-h_n}(\boldsymbol{Y}) = \begin{pmatrix} \boldsymbol{Y}_1 - h_n \boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 - \dfrac{\mathrm{sign}(h_n)}{2}\boldsymbol{Y}_3\right) \\ \boldsymbol{Y}_2 - Q_n\left(\nabla V\left(\boldsymbol{Y}_1 + t\boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 - \dfrac{\mathrm{sign}(h_n)}{2}\boldsymbol{Y}_3\right)\right); 0; -h_n\right) \\ -\boldsymbol{Y}_3 + 2\mathrm{sign}(h_n) Q_n\left(\nabla V\left(\boldsymbol{Y}_1 + t\boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 - \dfrac{\mathrm{sign}(h_n)}{2}\boldsymbol{Y}_3\right)\right); 0; -h_n\right) \end{pmatrix}.$$

It remains to verify that $\boldsymbol{\Phi}_{h_n} \circ \boldsymbol{\Phi}_{-h_n}(\boldsymbol{Y}) = \boldsymbol{Y}$ using that the quadrature is symmetric or exact. To fix the ideas, we assume that $h_n > 0$. Let us write $\boldsymbol{Y}' = \boldsymbol{\Phi}_{-h_n}(\boldsymbol{Y})$ and $\boldsymbol{Y}'' = \boldsymbol{\Phi}_{h_n}(\boldsymbol{Y}')$. Since $\boldsymbol{Y}_1' = \boldsymbol{Y}_1 - h_n \boldsymbol{M}^{-1}(\boldsymbol{Y}_2 - \frac{1}{2}\boldsymbol{Y}_3)$ and $\boldsymbol{Y}_2' + \frac{1}{2}\boldsymbol{Y}_3' = \boldsymbol{Y}_2 - \frac{1}{2}\boldsymbol{Y}_3$, we infer that

$$\boldsymbol{Y}_1'' = \boldsymbol{Y}_1' + h_n \boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2' + \frac{1}{2}\boldsymbol{Y}_3'\right) = \boldsymbol{Y}_1. \tag{2.15}$$

For the second component, we obtain

$$\begin{aligned} \boldsymbol{Y}_2'' &= \boldsymbol{Y}_2' - Q_n\left(\nabla V\left(\boldsymbol{Y}_1 + (t - h_n)\boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 - \frac{1}{2}\boldsymbol{Y}_3\right)\right); 0; h_n\right) \\ &= \boldsymbol{Y}_2' - Q_n\left(\nabla V\left(\boldsymbol{Y}_1 + t\boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 - \frac{1}{2}\boldsymbol{Y}_3\right)\right); -h_n; 0\right) \\ &= \boldsymbol{Y}_2' + Q_n\left(\nabla V\left(\boldsymbol{Y}_1 + t\boldsymbol{M}^{-1}\left(\boldsymbol{Y}_2 - \frac{1}{2}\boldsymbol{Y}_3\right)\right); 0; -h_n\right) = \boldsymbol{Y}_2, \end{aligned}$$

where we used (2.15) in the first line and invariance by translation and symmetry of the quadrature Q_n in the second and third lines respectively. The proof that $\boldsymbol{Y}_3'' = \boldsymbol{Y}_3$ follows along similar lines. □

Theorem 2 (Consistency error)**.** *Assume that $V \in C^2(\mathbb{R}^{dN}; \mathbb{R})$. If the quadrature (2.9) is exact or at least of order two, the scheme (2.3) has second-order accuracy in time.*

Proof. Let $\boldsymbol{q}(t), \boldsymbol{p}(t)$ be the exact solution to (2.2). Let us consider the column vector $\boldsymbol{Y}(t^n) = (\boldsymbol{q}(t^n), \frac{\boldsymbol{p}(t^{n-1/2}) + \boldsymbol{p}(t^{n+1/2})}{2}, \boldsymbol{p}(t^{n+1/2}) - \boldsymbol{p}(t^{n-1/2}))^{\mathrm{T}}$. The consistency error is defined as

$$\boldsymbol{\eta}^{n+1} := \frac{\boldsymbol{Y}(t^{n+1}) - \boldsymbol{\Phi}_{h_n}(\boldsymbol{Y}(t^n))}{h_n},$$

where $\boldsymbol{\Phi}_{h_n}$ is defined in the previous proof. Let us denote by $\boldsymbol{\eta}_1^{n+1}, \boldsymbol{\eta}_2^{n+1}, \boldsymbol{\eta}_3^{n+1}$ the three components of the consistency error. We have

$$\begin{aligned} h_n \boldsymbol{\eta}_1^{n+1} &= \boldsymbol{q}(t^{n+1}) - \boldsymbol{q}(t^n) - h_n \boldsymbol{M}^{-1} \boldsymbol{p}(t^{n+1/2}), \\ h_n \boldsymbol{\eta}_2^{n+1} &= \frac{\boldsymbol{p}(t^{n+3/2}) - \boldsymbol{p}(t^{n-1/2})}{2} + Q_n\left(\nabla V(\breve{\boldsymbol{q}}^n(t)); t^n; t^{n+1}\right), \\ h_n \boldsymbol{\eta}_3^{n+1} &= 2 h_n \boldsymbol{\eta}_2^{n+1}, \end{aligned}$$

where $\check{\boldsymbol{q}}^n(t) = \boldsymbol{q}(t^n) + \boldsymbol{M}^{-1}\boldsymbol{p}(t^{n+1/2})(t-t^n)$. Using a Taylor expansion and the equation $\dot{\boldsymbol{q}}(t) = \boldsymbol{M}^{-1}\boldsymbol{p}(t)$, we infer that

$$h_n\boldsymbol{\eta}_1^{n+1} = h_n\dot{\boldsymbol{q}}(t^{n+1/2}) - h_n\boldsymbol{M}^{-1}\boldsymbol{p}(t^{n+1/2}) + \mathcal{O}(h_n^3) = \mathcal{O}(h_n^3).$$

Moreover, since the quadrature is of second-order (at least) so that it can be replaced by the mid-point quadrature up to $\mathcal{O}(h_n^3)$, and using the equations $\dot{\boldsymbol{q}}(t) = \boldsymbol{M}^{-1}\boldsymbol{p}(t)$ and $\dot{\boldsymbol{p}}(t) = -\nabla V(\boldsymbol{q})(t)$, we obtain

$$\begin{aligned} h_n\boldsymbol{\eta}_2^{n+1} &= h_n\dot{\boldsymbol{p}}(t^{n+1/2}) - h_n\nabla V(\check{\boldsymbol{q}}(t^{n+1/2})) + \mathcal{O}(h_n^3) \\ &= h_n\dot{\boldsymbol{p}}(t^{n+1/2}) - h_n\nabla V(\boldsymbol{q}(t^n) + \tfrac{1}{2}h_n\dot{\boldsymbol{q}}(t^{n+1/2})) + \mathcal{O}(h_n^3) \\ &= h_n\dot{\boldsymbol{p}}(t^{n+1/2}) - h_n\nabla V(\boldsymbol{q}(t^{n+1/2})) + \mathcal{O}(h_n^3) = \mathcal{O}(h_n^3). \end{aligned}$$

We conclude that $\boldsymbol{\eta}^{n+1} = \mathcal{O}(h_n^2)$, i.e., the scheme is second-order accurate. □

Proposition 2.4 (Linear stability)**.** *Assume that the potential V is quadratic with a positive definite Hessian $\boldsymbol{H} := D^2V$. Let λ be the largest eigenvalue of $\boldsymbol{H}$. Let $\mu > 0$ be the smallest eigenvalue of $\boldsymbol{M}$. Then the scheme (2.3) is conditionally stable for a constant time-step h under the following CFL condition:*

$$h \;<\; 2\sqrt{\frac{\mu}{\lambda}}. \tag{2.16}$$

Proof. Since the potential V is quadratic, the dynamical system (2.2) is linear. Let $\boldsymbol{Z}^n$ be the column vector such that $\boldsymbol{Z}^n = \left(\boldsymbol{q}^n, \boldsymbol{p}^{n-1/2}, \boldsymbol{p}^{n+1/2}\right)^{\mathrm{T}}$. Adding a linear functional to V does not change the nature of the Hamiltonian system. We thus consider $\nabla V(\boldsymbol{0}) = \boldsymbol{0}$ and $V(\boldsymbol{0}) = 0$. Since $\nabla V(\boldsymbol{q})$ is by assumption linear in $\boldsymbol{q}$, we have $\nabla V(\hat{\boldsymbol{q}}^n(t)) = \nabla V(\boldsymbol{q}^n) + (t-t^n)\boldsymbol{H}\boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2}$, so that

$$\int_{I_n} \nabla V(\hat{\boldsymbol{q}}^n(t))dt = h\boldsymbol{H}\boldsymbol{q}^n + \frac{1}{2}h^2\boldsymbol{H}\boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2}.$$

Therefore, the scheme (2.3) can be written as $\boldsymbol{Z}^{n+1} = \boldsymbol{A}\boldsymbol{Z}^n$ with

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_{dN} & \boldsymbol{0}_{dN} & h\boldsymbol{M}^{-1} \\ \boldsymbol{0}_{dN} & \boldsymbol{0}_{dN} & \boldsymbol{I}_{dN} \\ -2h\boldsymbol{H} & \boldsymbol{I}_{dN} & -h^2\boldsymbol{H}\boldsymbol{M}^{-1} \end{pmatrix}.$$

The matrix $\boldsymbol{M}$ being symmetric definite positive, its square root $\boldsymbol{M}^{1/2}$ is well-defined. We then observe that

$$\begin{aligned} \tilde{\boldsymbol{A}} &= \begin{pmatrix} \boldsymbol{M}^{1/2} & \boldsymbol{0}_{dN} & \boldsymbol{0}_{dN} \\ \boldsymbol{0}_{dN} & \boldsymbol{M}^{-1/2} & \boldsymbol{0}_{dN} \\ \boldsymbol{0}_{dN} & \boldsymbol{0}_{dN} & \boldsymbol{M}^{-1/2} \end{pmatrix} \boldsymbol{A} \begin{pmatrix} \boldsymbol{M}^{-1/2} & \boldsymbol{0}_{dN} & \boldsymbol{0}_{dN} \\ \boldsymbol{0}_{dN} & \boldsymbol{M}^{1/2} & \boldsymbol{0}_{dN} \\ \boldsymbol{0}_{dN} & \boldsymbol{0}_{dN} & \boldsymbol{M}^{1/2} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{I}_{dN} & \boldsymbol{0}_{dN} & h\boldsymbol{I}_{dN} \\ \boldsymbol{0}_{dN} & \boldsymbol{0}_{dN} & \boldsymbol{I}_{dN} \\ -2h\boldsymbol{S} & \boldsymbol{I}_{dN} & -h^2\boldsymbol{S} \end{pmatrix}, \end{aligned}$$

where we introduced the symmetric positive definite matrix $\boldsymbol{S} = \boldsymbol{M}^{-1/2}\boldsymbol{H}\boldsymbol{M}^{-1/2}$. Up to an adequate change of variable for each of the coordinates, it is possible to assume that $\boldsymbol{S}$ is diagonal. Denoting $(\sigma_i)_{1\le i\le dN}$ the eigenvalues of $\boldsymbol{S}$ and scaling the momenta in $\boldsymbol{Z}_n$ by the factors $(\sqrt{2\sigma_i})_{1\le i\le dN}$, $-\tilde{\boldsymbol{A}}$ is block diagonal in the following matrices of order 3, for all $i \in \{1, \dots, dN\}$:

$$a_i = \begin{pmatrix} -1 & 0 & -h\sqrt{2\sigma_i} \\ 0 & 0 & -1 \\ h\sqrt{2\sigma_i} & -1 & h^2\sigma_i \end{pmatrix}.$$

The characteristic polynomial χ_{a_i} of a_i is $\chi_{a_i}(X) = (X-1)(X^2 - X(h^2\sigma_i - 2) + 1)$, which shows that 1 is an eigenvalue of a_i. Moreover, the polynomial $X^2 - X(h^2\sigma_i - 2) + 1$ is positive as long as $h < \frac{2}{\sqrt{\sigma_i}}$, and the two complex conjugate eigenvalues, written b_i and $\bar{b}_i$, have a modulus equal to 1 and a nonzero imaginary part. Thus, the three eigenvalues are distinct, and the 3×3 matrix a_i can be diagonalized for all $i \in \{1, \dots, dN\}$. Let P_i the matrix such that $P_i^{-1} a_i P_i = \mathrm{Diag}(1, b_i, \bar{b}_i)$. Then, writing P^{-1} the block-diagonal matrix composed of the elementary matrices $(P_i^{-1})_{1\le i\le dN}$ and D the block-diagonal matrix composed of the diagonal matrices of eigenvalues of $(a_i)_{1\le i\le dN}$, we infer that, for all $n \in \mathbb{N}$:

$$\|\tilde{\boldsymbol{A}}^n\| \le \|P^{-1}\|\|P\|\|D\|^n \le \|P^{-1}\|\|P\|,$$

because the diagonal matrix D has diagonal entries of modulus 1, and thus $\|D\| = 1$. Hence, for all $n \in \mathbb{N}$, we obtain $\|\boldsymbol{Z}^n\| \le C\|\boldsymbol{Z}^0\|$, for a constant C independent of n. Since the eigenvalues σ_i of $\boldsymbol{S}$ are positive and smaller than $\frac{\lambda}{\mu}$, we conclude that linear stability holds true under the CFL condition (2.16).

□

Remark 2.5 (Comparison with Störmer–Verlet)**.** *A possible writing of the Störmer–Verlet method is the following:*

$$\boldsymbol{q}^{n+1} = \boldsymbol{q}^n + h_n \boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2},$$
$$\boldsymbol{p}^{n+3/2} = \boldsymbol{p}^{n+1/2} - h_{n+1}\nabla V(\boldsymbol{q}^{n+1}).$$

Both the present scheme and the Störmer–Verlet scheme are of leapfrog-type, have a similar CFL condition for linear stability, and are second-order accurate. The main difference is that, using a mid-point quadrature, the forces used to update the momenta at t^{n+1} are computed at $t^{n+1/2}$ with the present scheme, i.e., $\frac{1}{2}(\boldsymbol{p}^{n+3/2} - \boldsymbol{p}^{n-1/2}) = -h_n\nabla V(\hat{\boldsymbol{q}}(t^{n+1/2}))$, whereas the momentum update can be rewritten as

$$\frac{1}{2}(\boldsymbol{p}^{n+3/2} - \boldsymbol{p}^{n-1/2}) = -\frac{1}{2}\Big(h_{n+1}\nabla V(\boldsymbol{q}^{n+1}) + h_n\nabla V(\boldsymbol{q}^n)\Big),$$

in the Störmer–Verlet scheme. Moreover, the origin of energy conservation is different for the two schemes. The Störmer–Verlet scheme is energy-conserving only for constant time-steps due to its symplecticity. The present scheme enjoys an algebraic pseudo-energy preservation property for every time-step (constant or not) up to quadrature errors.

Remark 2.6 (Adaptive time-stepping for discrete energy control)**.** *The conservation of the pseudo-energy does not imply stability since $\left(\boldsymbol{p}^{n-1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1}\boldsymbol{p}^{n+1/2}$ does not have a sign a priori. However, defining the discrete energy*

$$H^n := V(\boldsymbol{q}^n) + \frac{1}{8}\left(\boldsymbol{p}^{n-1/2} + \boldsymbol{p}^{n+1/2}\right)^{\mathrm{T}} \boldsymbol{M}^{-1}\left(\boldsymbol{p}^{n-1/2} + \boldsymbol{p}^{n+1/2}\right), \tag{2.17}$$

a straightforward calculation shows that

$$H^n = \tilde{H}^n + \frac{1}{8}([\boldsymbol{p}]^n)^{\mathrm{T}} M^{-1}[\boldsymbol{p}]^n = \tilde{H}^n + \frac{1}{8}|M^{-1/2}[\boldsymbol{p}]^n|^2. \tag{2.18}$$

This implies that

$$0 \le H^n - \tilde{H}^n \le \mathcal{O}(h_n^2),$$

where we used the identity (2.18) for the lower bound and we invoked Theorem 2 for the upper bound. One can use the identity (2.18) during the computations for an on-the-fly monitoring of possible departures of the conserved pseudo-energy $\tilde{H}^n$ from the discrete energy H^n. The idea is to check whether $\frac{1}{8}|M^{-1/2}[\boldsymbol{p}]^n|^2 \le \epsilon_{fly}\tilde{H}^n$ after every momenta computation and to halve the time-step if this bound is not met (note that the momentum jumps converge to zero with the time-step). The benefits of such an adaptive time-stepping strategy are illustrated in Section 2.3.3.

2.3 Numerical results

In this section, we present numerical results for the scheme (2.3). We consider classical benchmarks from the literature and a nonlinear wave equation from [Chabassier and Joly, 2010].

2.3.1 Convergence study

We perform a convergence study with a single particle in dimension $d = 1$. The reference solution is $q(t) = \sin(t)^4 + 1$, and the corresponding potential energy is

$$V(q(t)) = 8\left((q(t) - 1)^{3/2} - (q(t) - 1)^2\right).$$

We apply the scheme (2.3) to this Hamiltonian system over 10^3 seconds using the mid-point quadrature as well as the three- and five-point Gauss–Lobatto quadratures of order 3 and 7, respectively, for the integration of the forces. We report the ℓ_1-error with respect to the reference solution (the sum of the errors at the discrete time nodes divided by the number of time-steps) in Figure 2.1 as a function of the number of force evaluations. We observe that for the three quadratures, the convergence is of second order as expected. The quadrature order does not impact the convergence rate but has an influence on the computational efficiency. We note that in this case, the mid-point quadrature is more efficient than the three- and five-point Gauss–Lobatto quadratures of order 3 and 7.

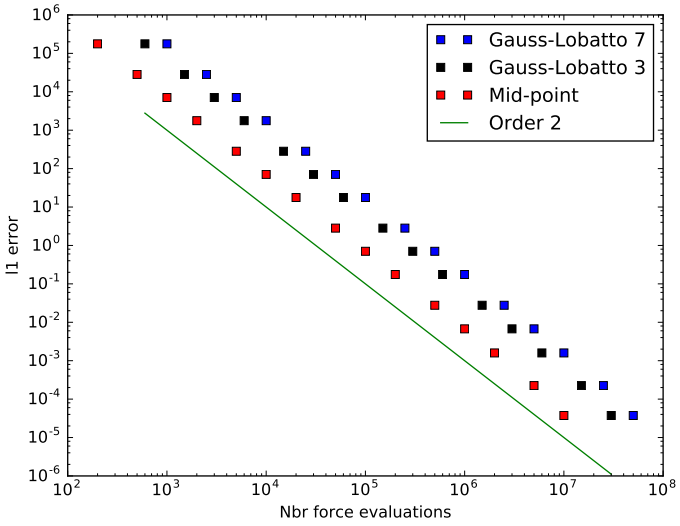


Figure 2.1 – Convergence test: ℓ_1-convergence for a single particle

2.3.2 Fermi–Pasta–Ulam

This test case was proposed in [Hairer et al., 2006, Chap. I.4]. It consists in having stiff linear springs linked to soft nonlinear springs in an alternating way, in dimension $d = 1$. Figure 2.2 illustrates the setting.

The Hamiltonian is

$$H(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}\sum_{i=1}^{m}(p_{2i-1}^2 + p_{2i}^2) + \frac{\omega^2}{4}\sum_{i=1}^{m}(q_{2i} - q_{2i-1})^2 + \sum_{i=0}^{m}(q_{2i+1} - q_{2i})^4,$$

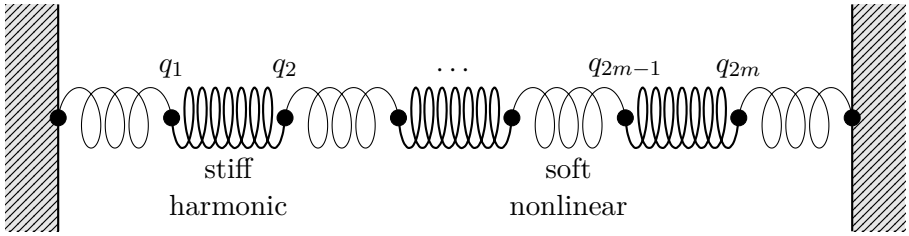


Figure 2.2 – Fermi–Pasta–Ulam test case

with typically $\omega \gg 1$. Introducing the variables

$$\begin{aligned} x_i &= (q_{2i} + q_{2i-1})/\sqrt{2}, & y_i &= (p_{2i} + p_{2i-1})/\sqrt{2}, \\ x_{m+i} &= (q_{2i} - q_{2i-1})/\sqrt{2}, & y_{m+i} &= (p_{2i} - p_{2i-1})/\sqrt{2}, \end{aligned}$$

the Hamiltonian can be rewritten as

$$\begin{aligned} H(\boldsymbol{x}, \boldsymbol{y}) = &\frac{1}{2}\sum_{i=1}^{2m} y_i^2 + \frac{\omega^2}{2}\sum_{i=1}^{m} x_{m+i}^2 \\ &+ \frac{1}{4}\left((x_1 - x_{m+1})^4 + \sum_{i=1}^{m-1} (x_{i+1} - x_{m+i+1} - x_i - x_{m+i})^4 + (x_m + x_{2m})^4 \right). \end{aligned} \tag{2.19}$$

As the system is Hamiltonian, the total energy of the system should be conserved by the numerical scheme. The Fermi–Pasta–Ulam system has yet another quasi-invariant. Letting $I_j(x_{m+j}, y_{m+j}) = \frac{1}{2}\left(y_{m+j}^2 + \omega^2 x_{m+j}^2\right)$ be the oscillatory energy of the jth stiff spring, the total oscillatory energy $I = I_1 + I_2 + \cdots + I_m$ is close to a constant value as proved in [Hairer et al., 2006, p.22]:

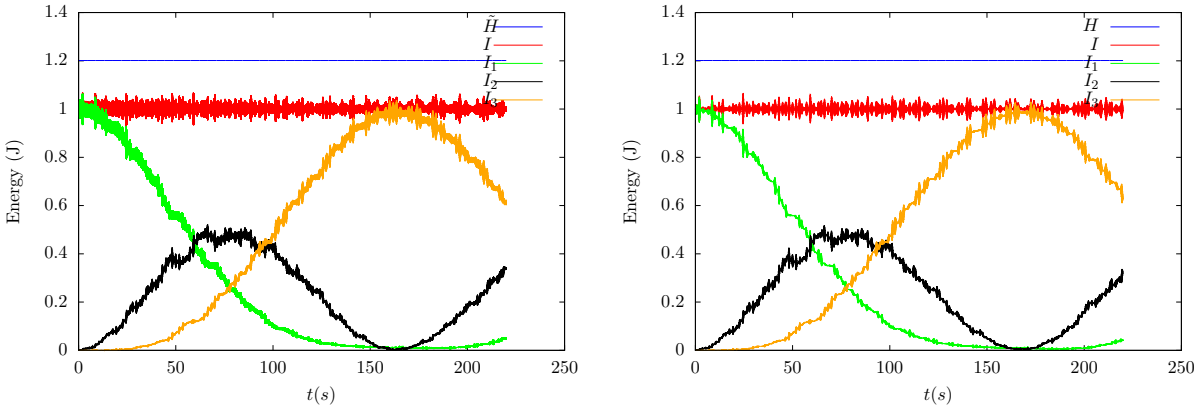
$$I(\boldsymbol{x}(t), \boldsymbol{y}(t)) = I(\boldsymbol{x}(0), \boldsymbol{y}(0)) + \mathcal{O}(\omega^{-1}).$$

In our numerical experiment, we set $m = 3$ and $\omega = 50$. Figure 2.32.1 (left panel) shows the variation of the oscillating energies and of the pesudo-energy $\tilde{H}^n$ over time for a constant time-step $h = 10^{-3}$. The energy exchange between the oscillatory modes is remarkably similar to the reference solution given in [Hairer et al., 2006, Chap. I.4] and represented in Figure 2.32.2 (right panel). The reference solution was computed with high accuracy using a Runge–Kutta 4 integrator with a time-step of $h = 10^{-4}$. In particular, the total oscillatory energy I displays fast oscillations around a fixed constant. The conservation of energy is verified up to machine precision, even with a mid-point quadrature. The results being already very satisfactory, the results computed with higher order quadratures are omitted for brevity. A more detailed study of the influence of the order of quadrature on pseudo-energy conservation is presented in the next section.

2.3.3 Nonlinear wave equation

The setting comes from [Chabassier and Joly, 2010]. The interval $\Omega = [0, 1]$ represents a one-dimensional string. Let $V : \mathbb{R}^d \to \mathbb{R}$ be the potential energy, with dimension $d = 2$. It is assumed that V verifies the following conditions:

— Smoothness: V is of class C^2;
— Convexity: V is strictly convex;
— Coercivity: $\exists K > 0$ so that $V(u) \geq K|u|^2$ for all $u \in \mathbb{R}^2$;
— Boundedness: $\exists M > 0$ so that $|\nabla V(u)|^2 \leq M \min(V(u), (1 + |u|^2))$ for all $u \in \mathbb{R}^2$.



(2.1) (2.2)

Figure 2.3 – Fermi–Pasta–Ulam test case: (2.1) Energy variation, present scheme, $h = 10^{-3}$; (2.2) reference RK4 solution, $h = 10^{-4}$

The problem of interest is to find $u : \Omega \times \mathbb{R}^+ \to \mathbb{R}^2$ such that

$$\begin{cases} \partial_{tt}^2 u - \partial_x(\nabla V(\partial_x u)) = 0, \\ u(0,t) = 0, \quad u(1,t) = 0, \\ u(x,0) = u^0(x), \quad \partial_t u(x,0) = v^0(x), \end{cases} \tag{2.20}$$

with given initial conditions $u^0 : \Omega \to \mathbb{R}^2$ and $v^0 : \Omega \to \mathbb{R}^2$. For a pair $(u_1, u_2) \in \mathbb{R}^2$, the functional V takes the following value:

$$V(u_1, u_2) = \frac{u_1^2 + u_2^2}{2} - \alpha \left(\sqrt{(1+u_1)^2 + u_2^2} - (1+u_1) \right),$$

where the parameter $\alpha \in [0, 1)$ is related to the tension of the string, such that the string behaviour is nonlinear when $\alpha > 0$ and the strength of the nonlinearity increases with α. The following variational formulation in $\mathcal{V} := H_0^1(\Omega; \mathbb{R}^2)$ is considered:

$$\frac{d^2}{dt^2} \left(\int_\Omega \boldsymbol{u} \cdot \boldsymbol{v} \right) + \int_\Omega \nabla V(\partial_x \boldsymbol{u}) \cdot \partial_x \boldsymbol{v} = 0, \quad \forall \boldsymbol{v} \in \mathcal{V}, \ \forall t > 0.$$

We use H^1-conforming $\mathbb{P}_1$ Lagrange finite elements for the space discretization. Let N be the number of nodes discretizing the string and $(\varphi_i)_{1 \le i \le 2N}$ be the nodal basis functions associated with the degrees of freedom of the string in the two directions. These basis functions span the finite-dimensional subspace $\mathcal{V}_N \subsetneq \mathcal{V}$. The space semi-discrete function approximating the exact solution is $\boldsymbol{u}_N(t) = \sum_{i=1}^{2N} \boldsymbol{q}_i(t)\varphi_i(x) \in \mathcal{V}_N$ and solves the following space semi-discrete problem:

$$\frac{d^2}{dt^2} \left(\int_\Omega \boldsymbol{u}_N \cdot \boldsymbol{v}_N \right) + \int_\Omega \nabla V(\partial_x \boldsymbol{u}_N) \cdot \partial_x \boldsymbol{v}_N = 0, \quad \forall \boldsymbol{v}_N \in \mathcal{V}_N, \ \forall t > 0.$$

Introducing the vector $\boldsymbol{q} = (q_1, \ldots, q_{2N}) \in \mathbb{R}^{2N}$, the following Hamiltonian system has to be integrated in time:

$$H(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2} \boldsymbol{p}^{\mathrm{T}} \boldsymbol{M}^{-1} \boldsymbol{p} + V(\boldsymbol{q}), \qquad V(\boldsymbol{q}) = \int_\Omega V \left(\sum_{i=1}^{2N} \boldsymbol{q}_i \partial_x \varphi_i \right),$$

where $\boldsymbol{M}$ is the classical $\mathbb{P}_1$ Lagrange finite element mass matrix. Assuming that all the components of $\boldsymbol{q}$ associated with the first direction are enumerated first and then those associated

with the second direction, the matrix $\boldsymbol{M}$ is a 2×2 block-diagonal matrix and each diagonal block is a tridiagonal matrix of size $N \times N$ equal to $\Delta x\,\mathrm{tridiag}(1/6, 2/3, 1/6)$.

In our numerical experiments, we consider the values $\alpha = 0$ (which corresponds to the linear case), $\alpha = 0.8$ (which corresponds to a mildly nonlinear behavior), and $\alpha = 0.99$ (which corresponds to a strongly nonlinear behavior). The space discretisation is such that $\Delta x = 0.01$ and thus $N = 99$ basis functions are used in each direction. The time-step is $\Delta t = 0.0033$. Using the same space discretisation, the greatest stable constant time-step has been found to be $\Delta t_{\max} = 0.0055$. Three numerical simulations are performed in every case by letting the amplitude of the initial condition u^0 at time t^0 be 0.01, 0.1, or 0.3. The initial velocity at time t^0 is always taken to be zero. The results are reported in Figure 2.4 where in all cases, a mid-point quadrature is used. Six snapshots of the $\mathbb{R}^2$-valued deformation vector $u_0(x) + u(x,t)$ of the string over one second are represented horizontally in various colors; specifically, at each snapshot in time, the deformation vector is plotted in the corresponding vertical plane. The role played by the nonlinearity can be observed in the fact that the amplitude of u^0 influences the vibration of the string. The tension which causes nonlinearity also changes the wave celerity. We observe an excellent agreement between the present results and the results reported in [Chabassier and Joly, 2010].

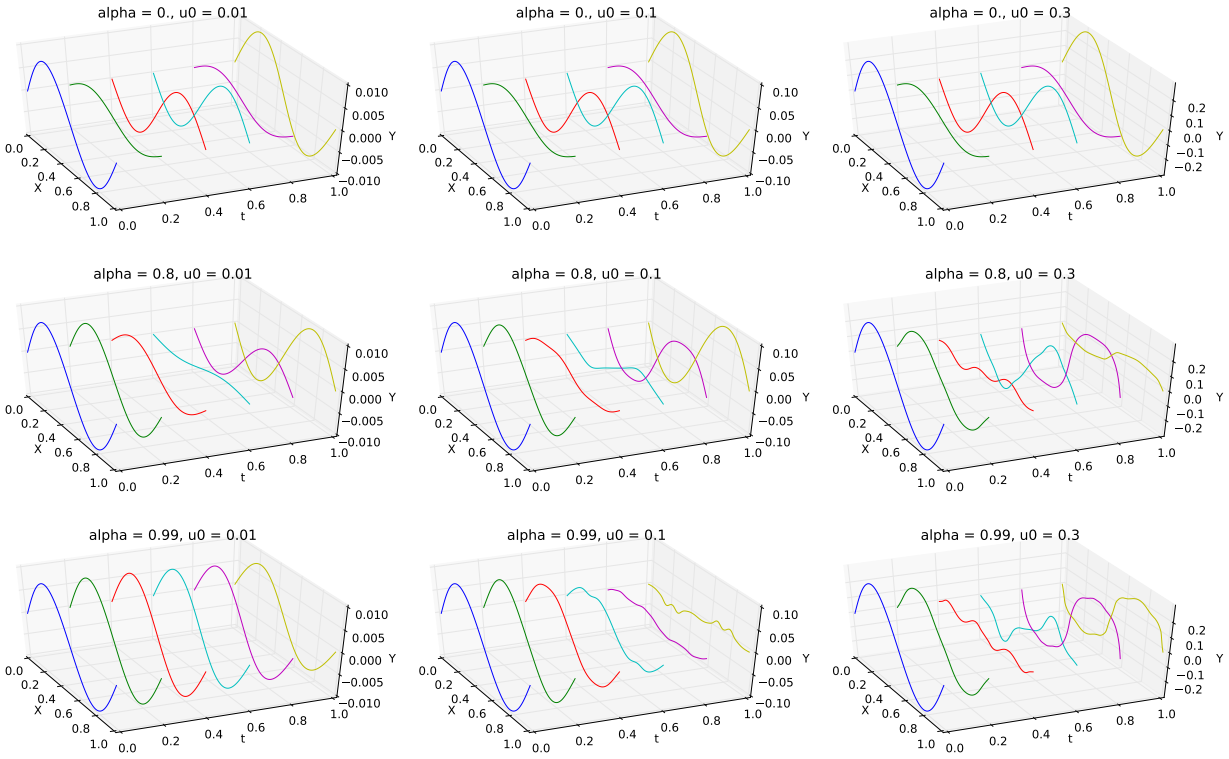
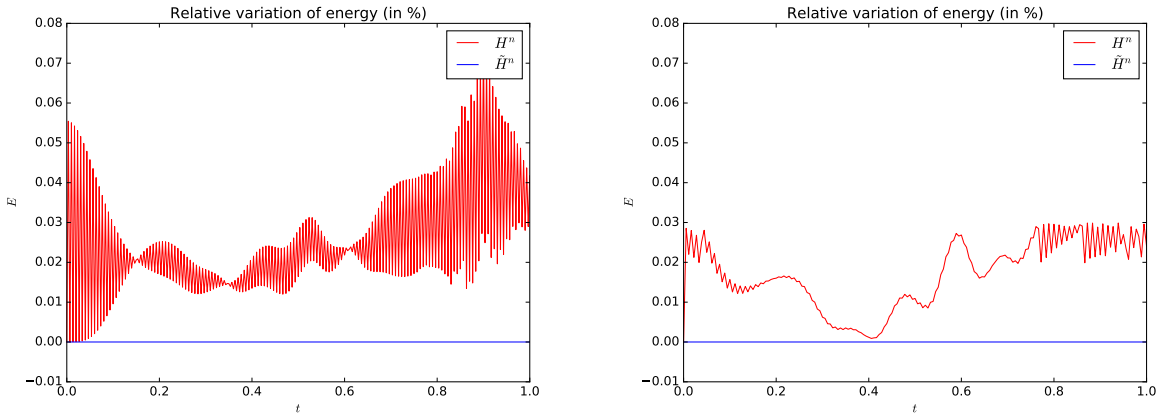


Figure 2.4 – Nonlinear wave equation: Deformations of the string over time with nonlinearity parameter $\alpha = 0$ (top), $\alpha = 0.8$ (middle), and $\alpha = 0.99$ (bottom); the amplitude of u^0 is $u_0 = 0.01$ (left), $u_0 = 0.1$ (middle), and $u_0 = 0.3$ (right)

The time-variation of the discrete pseudo-energy $\tilde{H}^n$ defined by (2.13) and of the discrete energy H^n defined by (2.17) are shown in Figure 2.52.1 in the most challenging case where $\alpha = 0.99$ and $u^0 = 0.3$. We first observe that the variations the discrete energy H^n are very moderate. In Figure 2.52.2, we illustrate the adaptive time-stepping strategy discussed in Remark 2.6, where we take $\epsilon_{\mathrm{fly}} = 0.03\%$ to control the departure of $\tilde{H}^n$ from H^n at each iteration. In this situation, the adaptive time steps take values in the range $[0.0024, 0.005]$. Concerning the discrete pseudo-energy $\tilde{H}^n$, we observe conservation up to machine precision when employing a five-point Gauss–Legendre quadrature of order 9.



(2.1) (2.2)

Figure 2.5 – Nonlinear wave equation: time-variation of the discrete energy H^n and pseudo-energy $\tilde{H}^n$ over a unit time interval for $\alpha = 0.99$ and an amplitude of 0.3 for u^0: (2.1) Fixed time-step $\Delta t = 0.0033$; (2.2) Adaptive time-step as in Remark 2.6

α	0			0.8			0.99		
u^0	0.01	0.1	0.3	0.01	0.1	0.3	0.01	0.1	0.3
MP	me	me	me	5.5e-08	5.6e-05	2.4e-05	1.5e-08	1.2e-05	1.1e-04
GL5	me	me	me	1.7e-13	1.2e-14	4.4e-14	3.3e-12	6.4e-13	1.8e-13
GL9	me	me	me	1.6e-13	1.2e-14	me	3.1e-12	me	me

Table 2.1 – Nonlinear wave equation: Maximal errors in the conservation of the pseudo-energy $\tilde{H}^n$ for the mid-point (MP), three-point Gauss–Legendre of order 5 (GL5) and five-point Gauss–Legendre of order 9 (GL9) quadratures; 'me' means machine error

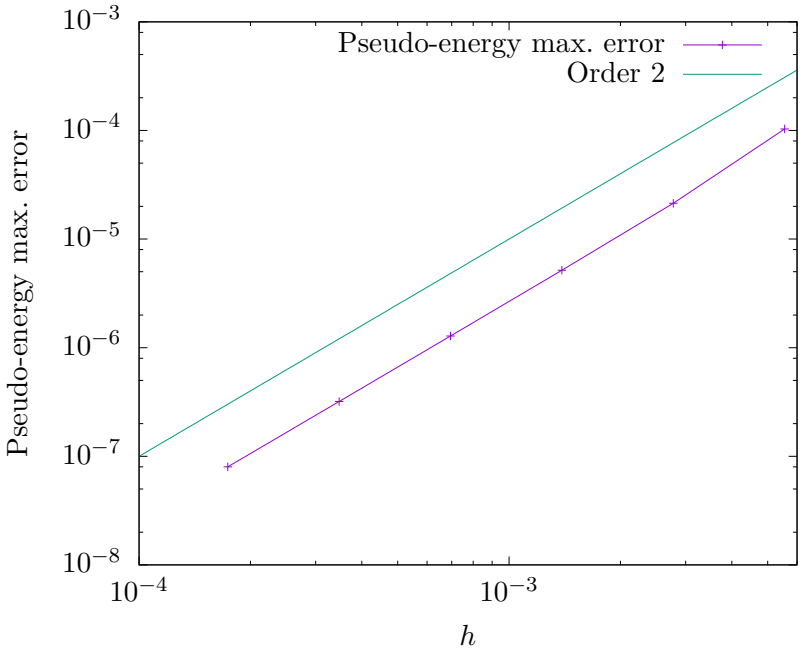


Figure 2.6 – Nonlinear wave equation: Maximal error on the conservation of the pseudo-energy $\tilde{H}^n$ as a function of the time-step in the case $\alpha = 0.99$, $u^0 = 0.3$, and a mid-point quadrature

To illustrate the impact of quadratures on pseudo-energy conservation, we perform two numerical experiments. First, Table 2.1 reports the maximal variation of the discrete pseudo-energy $\tilde{H}^n$ depending on the quadrature used with a constant time-step $\Delta t = 0.0033$. The mid-point quadrature is precise enough when used on the linear equation ($\alpha = 0$). The three-point Gauss–Legendre quadrature of order 5 is found to give very satisfactory results for the two nonlinear cases ($\alpha = 0.8$ and $\alpha = 0.99$). As announced in Remark 2.1, the maximal error on pseudo-energy conservation is observed to decrease when increasing the quadrature order. The total number of force evaluations is 300, 900, and 1500 when using the mid-point quadrature and the three- and five-point Gauss–Legendre quadratures of order 5 and 9, respectively. In the second experiment, we illustrate the second-order accuracy of pseudo-energy conservation when using the mid-point quadrature. We consider again the most challenging case where $\alpha = 0.99$ and $u^0 = 0.3$. Figure 2.6 shows the maximal variation of the discrete pseudo-energy $\tilde{H}^n$ with respect to the value of the (fixed) time-step used in the simulation, confirming the second-order accuracy.

To conclude this section, we present some comments on the relative costs of the present scheme with respect to an implicit scheme, e.g., the one devised in [Chabassier and Joly, 2010]. In the most challenging case where $\alpha = 0.99$ and $u^0 = 0.3$, the energy conservation in [Chabassier and Joly, 2010] is machine error. For a comparable error, we can consider the present scheme with the five-point Gauss–Legendre quadrature of order 9. The convergence criterion of the Newton's method in the implicit scheme can be estimated to require at least a couple of iterations per time-step and the same number of Hessian computations and global matrix inversions per time-step, whereas the explicit method with the five-point Gauss–Legendre quadrature of order 9 requires only 5 force evaluations per time-step. Thus, although no general conclusions can be drawn, the present explicit method stands good chances to be quite competitive with respect to an implicit method.

2.4 Asynchronous scheme

Owing to the CFL condition (2.16), the time-step can be required to be small in regions with stiff or nonsmooth dynamics. The overall efficiency of the computation would be compromised by the large number of integral calculations in the whole domain, while most of these would be redundant in smooth regions. We therefore propose an asynchronous version of the scheme which preserves the general properties of the synchronous version. In this section, we first present the idea behind slow-fast decomposition of the particles and we devise an asynchronous scheme for which we prove pseudo-energy conservation at the slow time nodes under exact force integration. Second-order accuracy is expected and is illustrated numerically on two test cases including an inhomogeneous wave equation.

2.4.1 Slow-fast splitting

In order to simplify the presentation of the asynchronous scheme, we limit ourselves here to the integration of a slow-fast dynamics, i.e., we consider a system with essentially two distinct time scales. The forces between the particles are supposed to be split into a "fast" set with an associated time-step h_F and a "slow" set with an associated time-step $h_S > h_F$. For example, the splitting can result from the relative stiffness of the forces in the system. Consequently, the particles are split into three sets: the slow particles are subjected only to slow forces, the fast particles are subjected only to fast forces, and the remaining particles, which are called mixed particles, are subjected to both slow and fast forces. This definition means that the slow particles do not interact with the fast particles directly, so that the potential V can be decomposed as follows:

$$V(\boldsymbol{q}) = V_S(\boldsymbol{q}_S) + V_M(\boldsymbol{q}_M, \boldsymbol{q}_S) + V_F(\boldsymbol{q}_F, \boldsymbol{q}_M),$$

where $\boldsymbol{q}_F$, $\boldsymbol{q}_S$, and $\boldsymbol{q}_M$ denote respectively the positions of the fast, slow, and mixed particles, the potential V_S describes the interactions between slow particles, V_M the interactions between slow and mixed particles, and V_F the interactions between mixed and fast particles (or between themselves). For instance, the purple particle in Figure 2.7 and the Particle 3 in Figure 2.8 are mixed particles. The mixed particle in Figure 2.7 is subjected to a "fast" force by the stiff spring on its right and a "slow" force by the soft spring on its left. The fast particle (in red) is only subjected to a "fast" force by the stiff spring. The slow particles (in blue) are only subjected to "slow" forces by the two soft springs. Finding a slow-fast decomposition is not possible for every Hamiltonian system. For example, in the case where all the particles interact with each other, no slow-fast splitting is available. The most favorable configuration is the one where the slow and the fast particles interact essentially among themselves and have very few interactions with mixed particles. This configuration is encountered in inhomogeneous problems where an interface separates two zones where the properties are different; the slow and the fast particles are then located in the two zones, whereas the mixed particles are located at the interface.

In what follows, we abuse the notation by denoting F, M and S the sets collecting the indices in $\{1, \dots, N\}$ of the fast, mixed and slow particles, respectively. For simplicity, we assume that the mass matrix $\boldsymbol{M}$ is diagonal and denote $\boldsymbol{M}_F$, $\boldsymbol{M}_M$ and $\boldsymbol{M}_S$ the restriction of $\boldsymbol{M}$ to the F, M and S particles respectively. Still for simplicity, we assume that both time-steps h_S and h_F are kept constant.

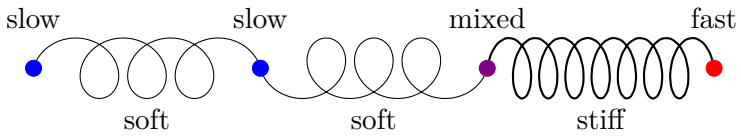


Figure 2.7 – Example of system of particles with a slow-fast splitting

2.4.2 Presentation of the asynchronous scheme

Without much loss of generality, we can suppose that the slow and fast time-steps are commensurate so that $h_S = K h_F$ with $K \in \mathbb{N}^*$. We then define the coarse time nodes $t^n = n h_S$ and the fine time nodes $t^{n,m} = t^n + m h_F$ for all $m \in \{0, \dots, K\}$. The asynchronous scheme consists in integrating K times the dynamics of the F and M particles with the "fast" forces computed at each time-step of length h_F and in updating the S particles with the "slow" forces computed once at the end of each time-step of length h_S. The general procedure is depicted in Figure 2.8 for four particles in the same configuration as in Figure 2.7. The efficiency of the asynchronous scheme hinges on the fact that each particle has a free-flight movement during each time-step, with the neighbouring particle forces acting only at the end of the time-step.

Let us now describe in more detail the asynchronous scheme over the coarse time interval $I_n = [t^n, t^{n+1}]$. At the beginning, we have at our disposal the triple $(p_i^{n-1/2}, q_i^n, p_i^{n+1/2})$ for the slow particles ($i \in S$) and the triple $(p_i^{n,-1/2} = p_i^{n-1,K-1/2}, q_i^{n,0}, p_i^{n,1/2})$ for the fast and the mixed particles ($i \in F \cup M$). The asynchronous scheme then proceeds as follows (we use here the two-step formulation which reduces to (2.6) in the synchronous case):

— For the fast particles ($i \in F$), one computes for all $m \in \{0, \dots, K-1\}$,

$$q_i^{n,m+1} = q_i^{n,m} + h_F \frac{1}{m_i} p_i^{n,m+1/2}, \tag{2.21a}$$

$$p_i^{n,m+3/2} = p_i^{n,m-1/2} - 2 \int_{t^{n,m}}^{t^{n,m+1}} \frac{\partial V_F}{\partial q_i}(\hat{\boldsymbol{q}}_F^{n,m}(t), \hat{\boldsymbol{q}}_M^{n,m}(t)) dt, \tag{2.21b}$$

with the free-flight trajectories for the fast and the mixed particles defined as

$$\hat{q}_j^{n,m}(t) = q_j^{n,m} + \frac{1}{m_j} p_j^{n,m+1/2}(t - t^{n,m}), \quad \forall t \in [t^{n,m}, t^{n,m+1}],\ \forall j \in F \cup M. \tag{2.22}$$

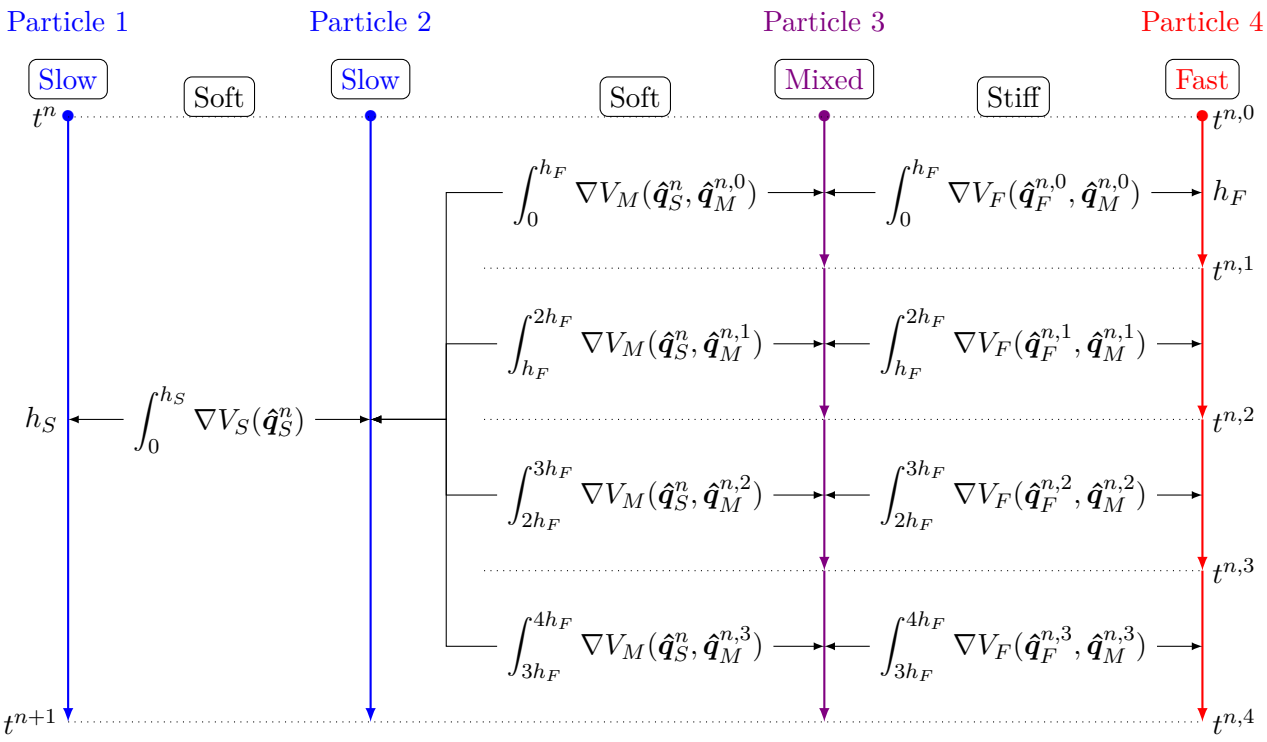


Figure 2.8 – Asynchronous integration of four particles with a slow-fast synamics, $h_S = 4h_F$

— For the mixed particles ($i \in M$), one computes for all $m \in \{0, \dots, K-1\}$, the position $q_i^{n,m+1}$ as in (2.21a), whereas equation (2.21b) is replaced by

$$p_i^{n,m+3/2} = p_i^{n,m-1/2} - 2\int_{t^{n,m}}^{t^{n,m+1}} \left(\frac{\partial V_F}{\partial q_i}(\hat{\boldsymbol{q}}_F^{n,m}(t), \hat{\boldsymbol{q}}_M^{n,m}(t)) + \frac{\partial V_M}{\partial q_i}(\hat{\boldsymbol{q}}_M^{n,m}(t), \hat{\boldsymbol{q}}_S^n(t)) \right) dt, \tag{2.23}$$

where the free-flight trajectories of the slow particles are computed over the coarse time interval as follows:

$$\hat{q}_j^n(t) = q_j^n + \frac{1}{m_j} p_j^{n+1/2}(t - t^n), \quad \forall t \in [t^n, t^{n+1}],\ \forall j \in S. \tag{2.24}$$

— For the slow particles ($i \in S$), one computes

$$q_i^{n+1} = q_i^n + h_S \frac{1}{m_i} p_i^{n+1/2}, \tag{2.25a}$$

$$p_i^{n+3/2} = p_i^{n-1/2} - 2\sum_{m=0}^{K-1} \int_{t^{n,m}}^{t^{n,m+1}} \frac{\partial V_M}{\partial q_i}(\hat{\boldsymbol{q}}_M^{n,m}(t), \hat{\boldsymbol{q}}_S^n(t))dt - 2\int_{t^n}^{t^{n+1}} \frac{\partial V_S}{\partial q_i}(\hat{\boldsymbol{q}}_S^n(t))dt, \tag{2.25b}$$

with the free-flight trajectories defined above.

Note that the slow forces between slow and mixed particles need to be evaluated at every fine time-step. In the worst case scenario, every slow force links a slow particle with a mixed particle, which results in the asynchronous scheme reverting to the synchronous scheme. Such a case typically occurs when the particles all interact or when the system alternates fast and slow forces. On the other hand, the efficiency of the asynchronous scheme compared to the synchronous scheme is maximal in the case where the mixed particles constitute a small fraction of the particles and their interaction is limited to a small fraction of the slow particles. A typical case is a nearest-neighbour interaction with slow and fast particles located in distinct regions, the mixed particles being confined in a lower dimensional delimiting interface. In the limit of a

large number of particles, the computational cost per large time-step h_S reduces to K integrals of the fast forces and one integral of the slow forces.

Proposition 2.7 (Synchronization of particles)**.** *Assume that the numerical integration is exact. Then the numerical scheme (2.21)–(2.25) exactly conserves the following pseudo-energy at the coarse time nodes t^n:*

$$\tilde{H}^n = V_S(\boldsymbol{q}_S^n) + V_M(\boldsymbol{q}_M^{n,0}, \boldsymbol{q}_S^n) + V_F(\boldsymbol{q}_F^{n,0}, \boldsymbol{q}_M^{n,0}) + \sum_{i\in S} \frac{1}{2m_i} \left(p_i^{n-1/2}\right)^{\mathrm{T}} p_i^{n+1/2} + \sum_{i\in F\cup M} \frac{1}{2m_i} \left(p_i^{n,-1/2}\right)^{\mathrm{T}} p_i^{n,1/2}. \tag{2.26}$$

Proof. Let us set

$$\begin{aligned}
\tilde{H}_S^n =& V_S(\boldsymbol{q}_S^n) + \sum_{i\in S} \frac{1}{2m_i} \left(p_i^{n-1/2}\right)^{\mathrm{T}} p_i^{n+1/2}, \\
\tilde{H}_{FM}^{n,m} =& V_F(\boldsymbol{q}_F^{n,m}, \boldsymbol{q}_M^{n,m}) + V_M(\boldsymbol{q}_M^{n,m}, \hat{\boldsymbol{q}}_S^n(t^{n,m})) + \sum_{i\in F\cup M} \frac{1}{2m_i} \left(p_i^{n,m-1/2}\right)^{\mathrm{T}} p_i^{n,m+1/2},
\end{aligned}$$

for all $m \in \{0, \dots, K\}$, so that $\tilde{H}^n = \tilde{H}_S^n + \tilde{H}_{FM}^{n,0}$. Following the same calculations as in the proof of Theorem 1 for equation (2.25), we infer that

$$\tilde{H}_S^{n+1} = \tilde{H}_S^n - \sum_{m=0}^{K-1} \int_{t^{n,m}}^{t^{n,m+1}} \frac{\partial V_M}{\partial \boldsymbol{q}_S}(\hat{\boldsymbol{q}}_M^{n,m}(t), \hat{\boldsymbol{q}}_S^n(t)) \cdot \left(\boldsymbol{M}_S^{-1} \boldsymbol{p}_S^{n+1/2}\right) dt.$$

Similarly, for all $m \in \{0, \dots, K-1\}$, using (2.21) and (2.23), we have

$$\tilde{H}_{FM}^{n,m+1} = \tilde{H}_{FM}^{n,m} + \int_{t^{n,m}}^{t^{n,m+1}} \frac{\partial V_M}{\partial \boldsymbol{q}_S}(\hat{\boldsymbol{q}}_M^{n,m}(t), \hat{\boldsymbol{q}}_S^n(t)) \cdot \left(\boldsymbol{M}_S^{-1} \boldsymbol{p}_S^{n+1/2}\right) dt,$$

and summing over m, we obtain

$$\tilde{H}_{FM}^{n+1,0} = \tilde{H}_{FM}^{n,K} = \tilde{H}_F^{n,0} + \sum_{m=0}^{K-1} \int_{t^{n,m}}^{t^{n,m+1}} \frac{\partial V_M}{\partial \boldsymbol{q}_S}(\hat{\boldsymbol{q}}_M^{n,m}(t), \hat{\boldsymbol{q}}_S^n(t)) \cdot \left(\boldsymbol{M}_S^{-1} \boldsymbol{p}_S^{n+1/2}\right) dt,$$

which gives the result. □

Remark 2.8 (Asynchronous pseudo-energy conservation)**.** *The pseudo-energy $\tilde{H}^n$ of Theorem 1 is not conserved after every integration over a fast time-step h_F in the asynchronous setting. This results from the fact that during a "slow" time-step h_S, the effect of forces has been taken into account for the "fast" particles but not for the "slow" particles.*

2.4.3 Numerical results

In this section, we present numerical results on the asynchronous scheme. We first consider a variant of the Fermi–Pasta–Ulam system with a slow-fast dynamics and then an inhomogeneous wave propagation problem.

Fermi–Pasta–Ulam system with slow-fast dynamics

We propose a slight variation of the Fermi–Pasta–Ulam test case in order to assess the efficiency of the asynchronous scheme. Contrary to the usual setting where stiff and soft springs alternate, we suppose here that the system is composed of one stiff region and one soft region, delimited by an interface in the middle of the domain. Figure 2.9 illustrates the setting. There are $(m-1)$ fast particles, 1 mixed particle, and m slow particles.

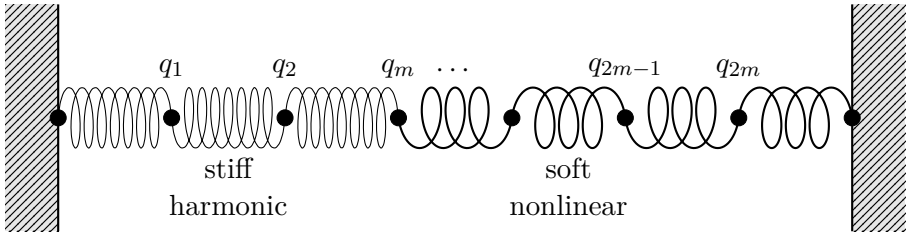


Figure 2.9 – Setting for the Fermi–Pasta–Ulam system with slow-fast dynamics

We consider a problem in dimension $d = 1$. The Hamiltonian is given by

$$H(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2} \sum_{i=1}^{2m} p_i^2 + \frac{\omega^2}{4} \sum_{i=1}^{m} (q_i - q_{i-1})^2 + \sum_{i=m}^{2m} (q_{i+1} - q_i)^4.$$

In the present experiment, we take $m = 3$ and $\omega^2 = 10$. The fast forces being generated by stiff linear springs, the fast time-step h_F should respect the CFL condition from Equation (2.16), which here leads to $h_F < 3 \cdot 10^{-3}$. The small time-step h_S being controlled by soft nonlinear springs, the CFL condition (2.16) is not applicable. A constant stable time-step has been found empirically to be $h_S \leq 10^{-1}$ using the five-point Gauss–Lobatto quadrature of order 7. The dynamics of the particles is presented in Figure 2.10 for $h_S = 0.01$ and $h_F = 2 \cdot 10^{-4}$, so that 50 iterations of the fine time-step are carried out for each iteration of the coarse time-step. Observe that, as expected, the fast particles ($1 \leq i \leq m$) exhibit oscillations with a typical frequency ω, whereas the slow particles ($m+1 \leq i \leq 2m$) have tame nonlinear oscillations with a frequency smaller than 1. Figure 2.11 shows that the conservation of the discrete pseudo-energy $\tilde{H}^n$ defined by (2.13) is as perfect for the asynchronous scheme as for the synchronous scheme with the five-point Gauss–Lobatto quadrature of order 7.

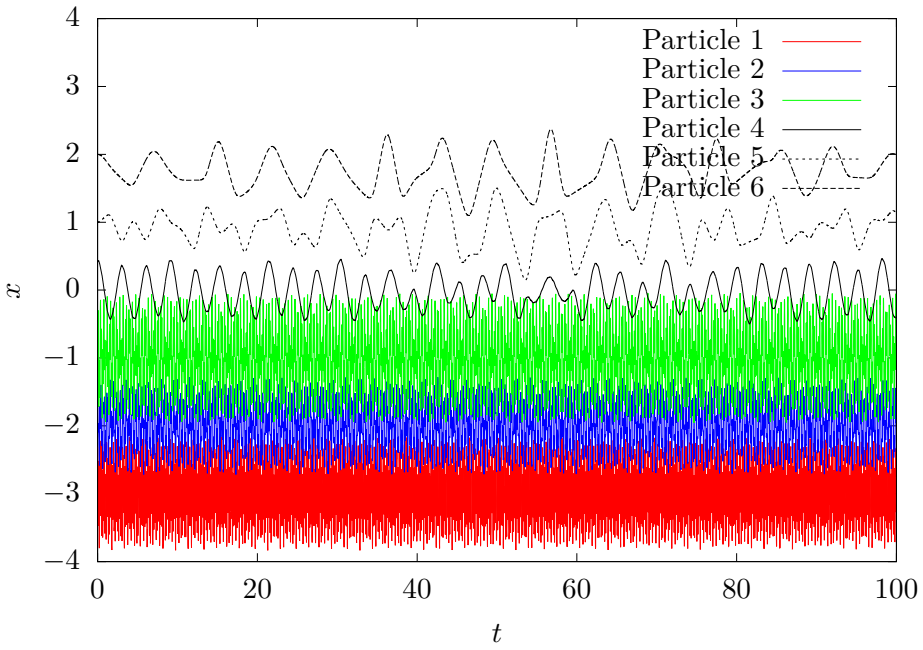


Figure 2.10 – Fermi–Pasta–Ulam system with slow-fast dynamics: Position dynamics for the asynchronous scheme ($h_S = 0.01$, $h_F = 2 \cdot 10^{-4}$)

The computational cost of the scheme is proportional to the number $\mathcal{N}$ of force evaluations. With an n-point quadrature and a total integration time T, the numbers of force evaluations

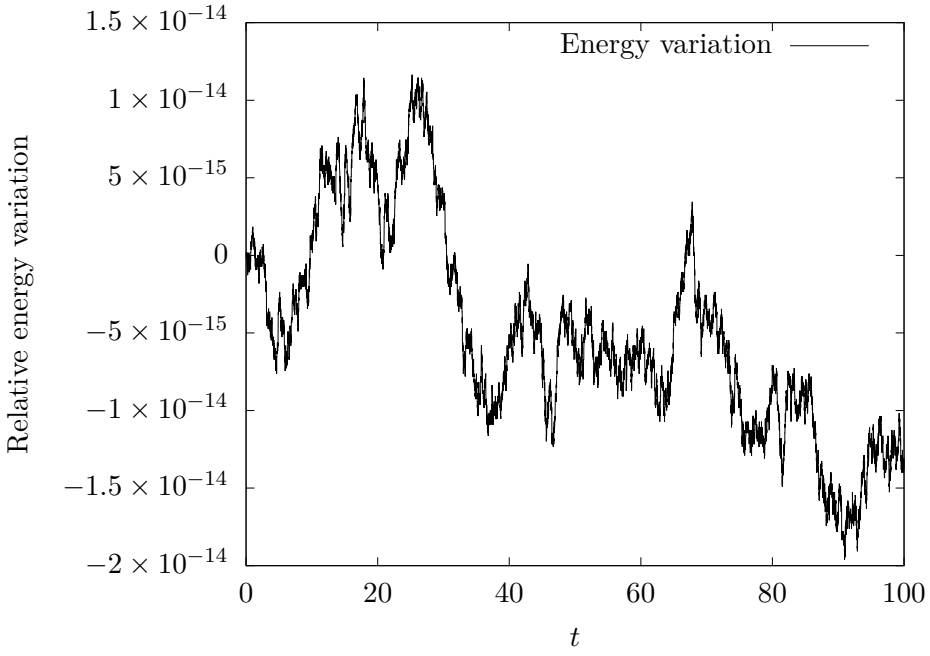


Figure 2.11 – Fermi–Pasta–Ulam system with slow-fast dynamics: Relative variation of the discrete pseudo-energy $\tilde{H}^n$ for the asynchronous scheme ($h_S = 0.01$, $h_F = 2 \cdot 10^{-4}$)

$\mathcal{N}_s$ and $\mathcal{N}_a$ for the synchronous and asynchronous schemes respectively on the present slow-fast problem are given by:

$$\mathcal{N}_s = (n-1)T\frac{2m+1}{h_F}, \qquad \mathcal{N}_a = (n-1)T\left(\frac{m+1}{h_F} + \frac{m}{h_S}\right).$$

Recalling that $K = \frac{h_S}{h_F} \geq 1$ is the number of fast steps per slow step, the cost reduction η of the asynchronous scheme with respect to the synchronous scheme is given by

$$\eta = \frac{\mathcal{N}_a}{\mathcal{N}_s} = \frac{1 + \frac{m}{(m+1)K}}{1 + \frac{m}{m+1}}.$$

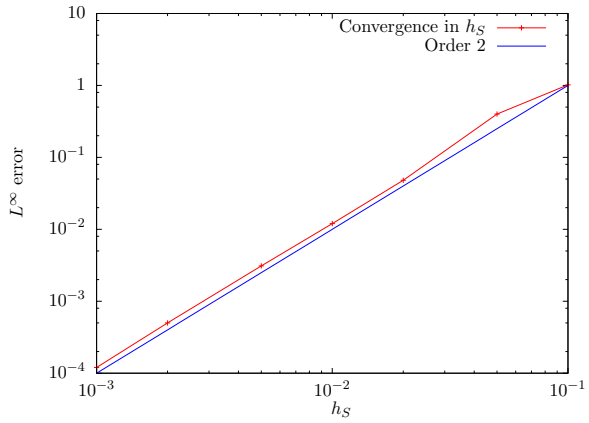
For $h_S = 0.01$, $h_F = 2 \cdot 10^{-4}$ and $T = 100$, $\mathcal{N}_a = 1.015 \cdot 10^8$, to be compared with $\mathcal{N}_s = 1.75 \cdot 10^8$. As m increases,

$$\eta \xrightarrow{m \to +\infty} \frac{1 + \frac{1}{K}}{2}.$$

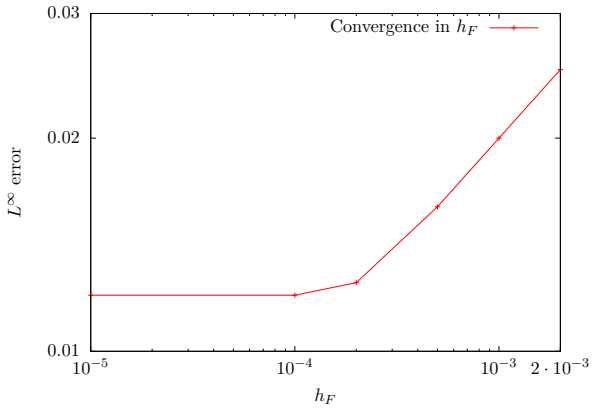
When the number of fast subiterations K increases, η tends to 0.5, which means that the computational cost reduction of the asynchronous scheme compared to the synchronous scheme approaches 50%. This is the best-case scenario, since the computational cost is concentrated on the fast dynamics where frequent evaluations are required, whereas the slow dynamics is almost costless.

In order to assess the accuracy of the asynchronous scheme, we consider the L^∞-error of the position of the asynchronous solution with respect to the synchronous solution using the small time-step h_F. Figure 2.122.1 (left panel) shows the evolution of the error as the coarse time-step h_S is refined, with fixed fine time-step $h_F = 10^{-4}$. We observe a second-order convergence of the error. Figure 2.122.2 (right panel) displays the evolution of the error as the fine time-step h_F is further refined, with fixed coarse time-step $h_S = 10^{-2}$. We observe that the error decreases until it reaches a plateau, which is due to the error on the slow particles. These observations confirm that reducing the fine time-step beyond $h_F = h_S/50$ does not significantly improve the error since the error is dominated by the error on the slow particles.

Conversely, the error reduction due to the coarse time-step reduction is not compromised by the asynchronous scheme.



(2.1)



(2.2)

Figure 2.12 – Fermi–Pasta–Ulam system with slow-fast dynamics: Convergence of the asynchronous scheme (2.1) with respect to the coarse time-step h_S, with fixed fine time-step $h_F = 10^{-4}$, and (2.2) with respect to the fine time-step h_F, with fixed coarse time-step $h_S = 10^{-2}$

Finally, a convergence test is carried out with a constant ratio $\frac{h_S}{h_F} = 25$ and using the five-point Gauss–Lobatto quadrature of order 7. The error is measured as previously by the L^∞-error on the positions between the synchronous and asynchronous schemes. The results are presented in Figure 2.13. The synchronous method is used with a constant time-step h_F. We observe second-order convergence as both time-steps are refined simultaneously.

Inhomogeneous wave propagation

As a physically relevant example of the slow-fast test case, we consider the propagation of a wave in a linear elastic material in dimension $d = 1$, with an inhomogeneous speed of sound. Denote the domain Ω, $u^0 : \Omega \to \mathbb{R}$ and $v^0 : \Omega \to \mathbb{R}$ initial conditions for displacement and velocity respectively, and $u : \Omega \times \mathbb{R}^+ \to \mathbb{R}$ the displacement, u follows the equations:

$$\begin{cases} \partial_{tt} u = \partial_x \Big(c(x)^2 \partial_x u \Big) & \text{in } \Omega, \\ u_{|\partial\Omega} = 0, \\ u(x,0) = u^0(x), \quad \partial_t u(x,0) = v^0(x). \end{cases} \tag{2.27}$$

We take $\Omega = (0,1)$ and we set

$$c(x) = \begin{cases} 10 & \text{if } x \leq 0.5, \\ 1 & \text{if } x > 0.5. \end{cases}$$

Setting $N \in \mathbb{N}^*$, $\Delta x = \frac{1}{N}$ and $x_i = i\Delta x$ for all $i \in \{0, \dots, N\}$, the partial differential equation (2.27) can be semi-discretized in space with the following centered finite difference scheme (which is equivalent to a discretization using H^1-conforming $\mathbb{P}_1$ Lagrange finite elements after

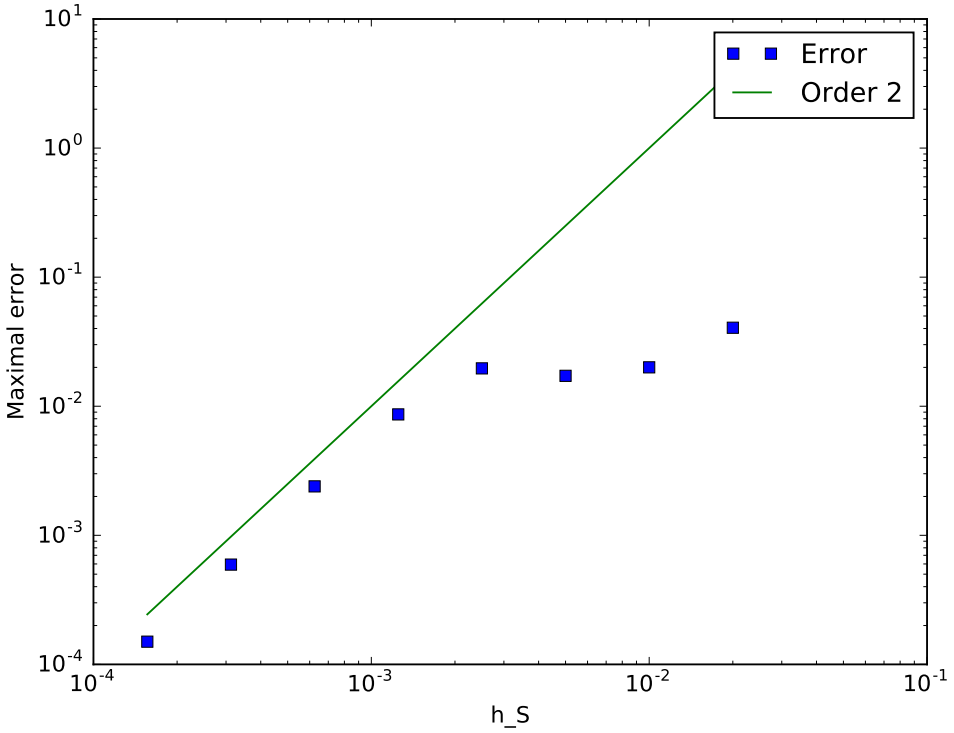


Figure 2.13 – Fermi–Pasta–Ulam system with slow-fast dynamics: L^∞-error on the positions between the synchronous scheme (with time-step h_F) and the asynchronous scheme (with time-steps h_S and h_F having fixed ratio $\frac{h_S}{h_F} = 25$).

lumping the mass matrix):

$$\begin{cases} \dfrac{d^2u_i}{dt^2} = \dfrac{1}{\Delta x^2}\left(c(x_{i-1/2})^2(u_{i-1}-u_i) - c(x_{i+1/2})^2(u_i-u_{i+1})\right) & \text{for } i \in \{1,\dots,N-1\}, \\ u_0 = u_N = 0, \quad \dfrac{du_0}{dt} = \dfrac{du_N}{dt} = 0, \\ u_i(0) = u^0(x_i), \quad \dfrac{du_i}{dt}(0) = v^0(x_i). \end{cases} \tag{2.28}$$

Setting $\boldsymbol{q} = (u_i)_{0\le i\le N}$, $\boldsymbol{p} = \left(\frac{du_i}{dt}\right)_{0\le i\le N}$ and $\omega_{i-1/2} = \frac{c(x_{i-1/2})}{\Delta x}$, the ordinary differential equation in (2.28) is derived from the following Hamiltonian:

$$H(\boldsymbol{p},\boldsymbol{q}) = \frac{1}{2}\sum_{i=1}^{N-1} p_i^2 + \frac{1}{2}\sum_{i=1}^{N} \omega_{i-1/2}^2 (q_i - q_{i-1})^2.$$

The CFL condition (2.16) becomes

$$h < 2\frac{\Delta x}{\omega_{i-1/2}}, \qquad \forall i \in \{1,\dots,N\}.$$

For the indices i such that $x_i \le 0.5$, one must then take $h \le 0.2\Delta x$, while for $x_i > 0.5$, it suffices that $h \le 2\Delta x$. In what follows, we therefore set the slow (resp. fast) particules as the elements i such that $x_i > 0.5$ (resp. $x_i < 0.5$) and define $h_S = \Delta x$ and $h_F = 0.1\Delta x$. The mixed particle is the particle at the interface between the fast and slow particles.

The numerical solution and the exact solution for the displacement and the velocity computed with $\Delta x = 5 \times 10^{-4}$ are presented in Figures 2.14 and 2.15 respectively. The system is initialized with the functions

$$u^0(x) = 10^{-2} e^{-(20(x-0.2))^2} \mathbb{1}_{(0,0.5)}(x), \quad v^0(x) = 8(x-0.2) e^{-(20(x-0.2))^2} \mathbb{1}_{(0,0.5)}(x).$$

The initial condition propagates to the right with the speed of sound $c_1 = 10$, until it reaches $x = 0.5$. At the boundary between the slow and fast domain, it is partly transmitted to the right with speed of sound $c_2 = 1$ and partly reflected with speed $-c_1$. The reflected wave reflects again on the left boundary $x = 0$ of the domain. Successive reflections and transmissions occur, which result in the final state in Figures 2.142.6 and 2.152.6. The exact solution can be expressed as follows for all $t > 0$:

$$\forall x \in (0, 0.5), \quad u(x,t) = \sum_{k \geq 0} \left(\frac{c_2 - c_1}{c_1 + c_2} \right)^k (u_0(x + k - c_1 t) - u_0(k - x - c_1 t)),$$

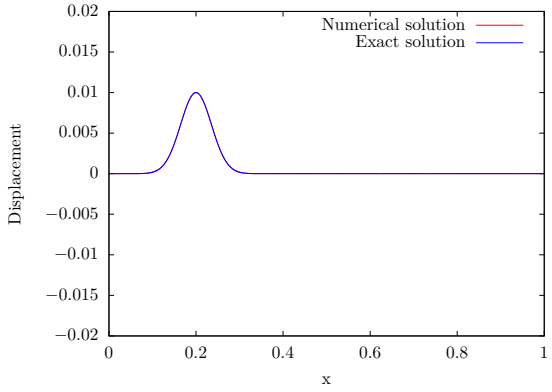
$$\forall x \in (0.5, 1), \quad u(x,t) = \frac{2c_1}{c_1 + c_2} \sum_{k \geq 0} \left(\frac{c_2 - c_1}{c_1 + c_2} \right)^k u_0 \left(\frac{c_1}{c_2}(x - 0.5) + k + 0.5 - c_1 t \right).$$

The numerical solution matches very well the exact solution. We can observe slight overshoots near the extrema and at the tail of the peaks, especially in the slow domain. This can be explained by the fact that the space-discretization (2.28) is slightly dispersive so that steep variations tend to generate oscillations (similar to the Gibbs phenomenon). Figure 2.16 presents the behavior of the error for the asynchronous and synchronous schemes with respect to the number of force evaluations. We consider the maximal error at the end of the simulation, i.e., at $t = 0.5$. The results are obtained by letting $h_S, h_F \to 0$ while keeping the ratio h_S/h_F fixed. The convergence is of order 1 with respect to the number of force evaluations. Since the number of force evaluations scales like $(\Delta x \cdot \Delta t)^{-1}$, this is compatible with a second-order convergence in space and in time after taking into account the CFL condition. In conclusion, the asynchronous scheme displays similar errors to the synchronous scheme, with roughly half of the number of force evaluations involved as noted in Section 2.4.3. This confirms the efficiency of the asynchronous scheme.

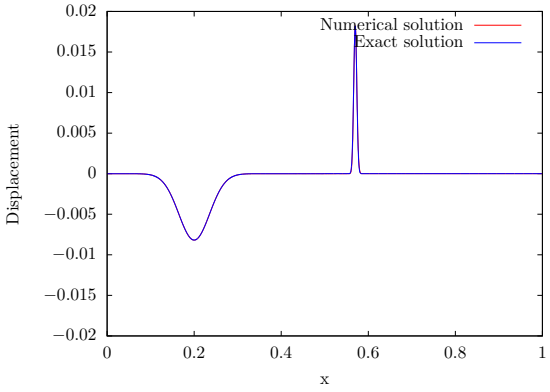
2.5 Conclusion

In this paper, a new explicit pseudo-energy conserving time-integration scheme has been proposed. It is capable of handling general nonlinear Hamiltonian systems and has been tested on classical numerical benchmarks and on a nonlinear wave propagation problem. The present scheme enables the use of local time-stepping strategies to circumvent stiff CFL conditions on the time-step and to enhance computational efficiency in the context of slow-fast dynamics.

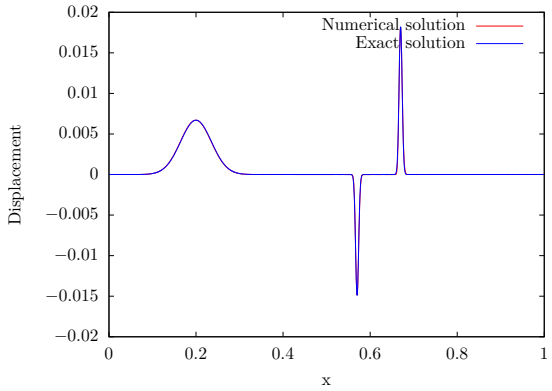
Various perspectives of the present work can be considered. We believe that the time-integration of dissipative systems should be a straightforward extension of the present scheme. Variational integrators have been proposed for dissipative systems and have proven to be able to accurately track the physical dissipation of energy [Kane et al., 2000]. Other possible developments lie in the adaptation of the scheme to constrained Hamiltonian systems [Leyendecker et al., 2008], such as mechanical contact problems [Kane et al., 1999b, Wohlmuth, 2011] and rigid body rotations [Krysl and Endres, 2005, Salomon et al., 2008, Mariotti, 2016]. Another perspective is the high-order extension of the present scheme.



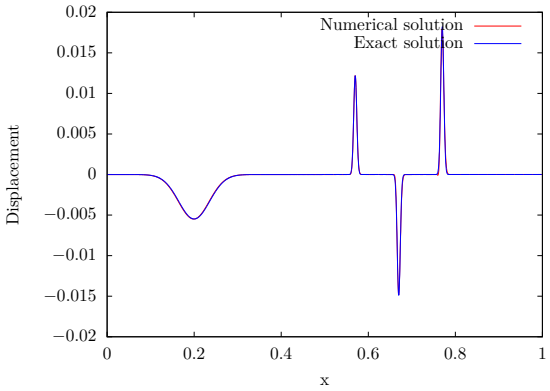
(2.1) $t = 0$



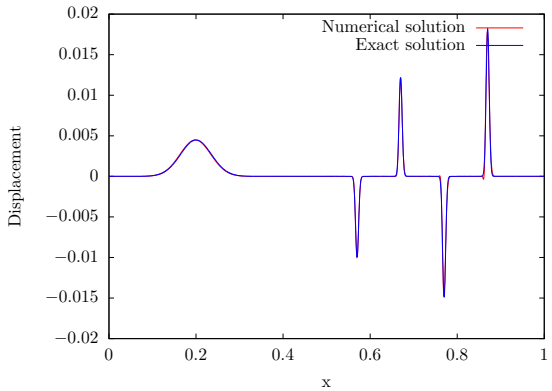
(2.4) $t = 0.1$



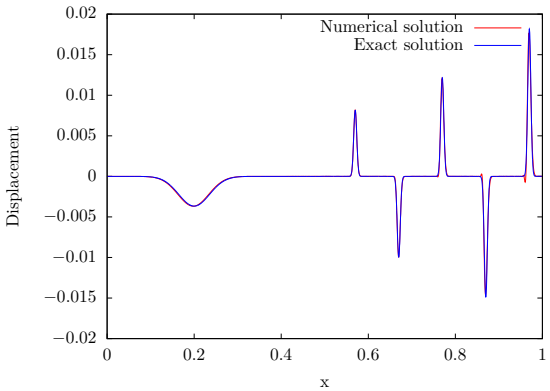
(2.2) $t = 0.2$



(2.5) $t = 0.3$

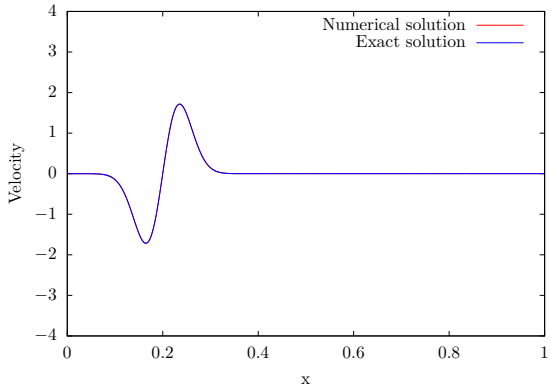


(2.3) $t = 0.4$

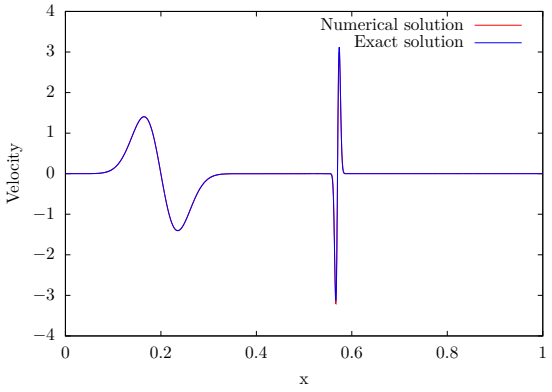


(2.6) $t = 0.5$

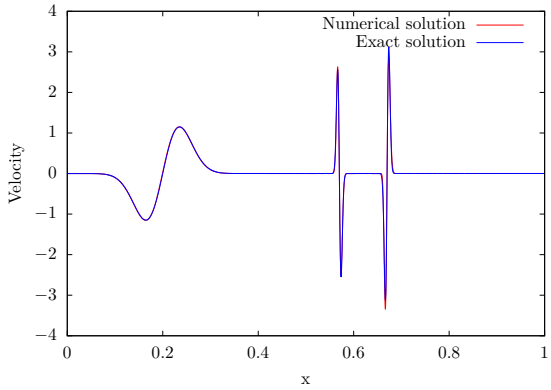
Figure 2.14 – Inhomogeneous wave propagation: Displacement u for $\Delta x = 5 \times 10^{-4}$ at times (2.1) $t = 0$, (2.4) $t = 0.1$, (2.2) $t = 0.2$, (2.5) $t = 0.3$, (2.3) $t = 0.4$, (2.6) $t = 0.5$



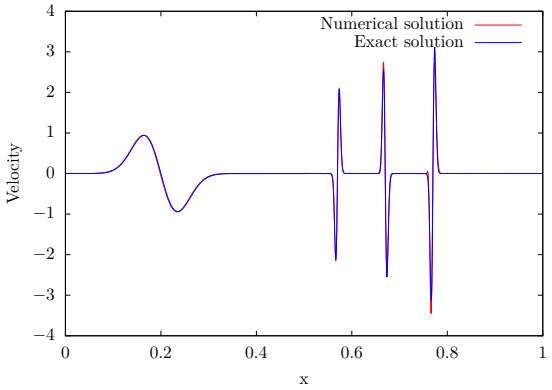
(2.1) $t = 0$



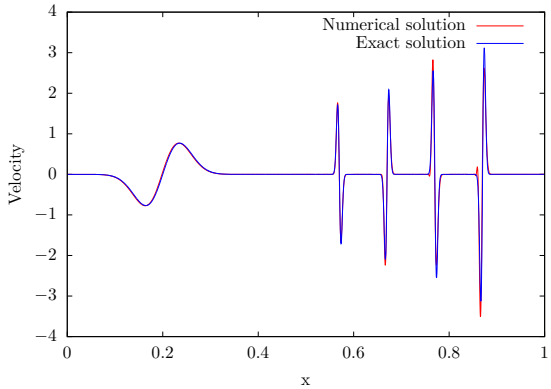
(2.4) $t = 0.1$



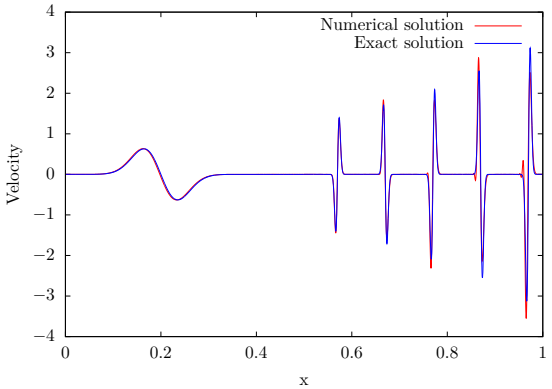
(2.2) $t = 0.2$



(2.5) $t = 0.3$



(2.3) $t = 0.4$



(2.6) $t = 0.5$

Figure 2.15 – Inhomogeneous wave propagation: Velocity $\frac{du}{dt}$ for $\Delta x = 5 \times 10^{-4}$ at times (2.1) $t = 0$, (2.4) $t = 0.1$, (2.2) $t = 0.2$, (2.5) $t = 0.3$, (2.3) $t = 0.4$, (2.6) $t = 0.5$

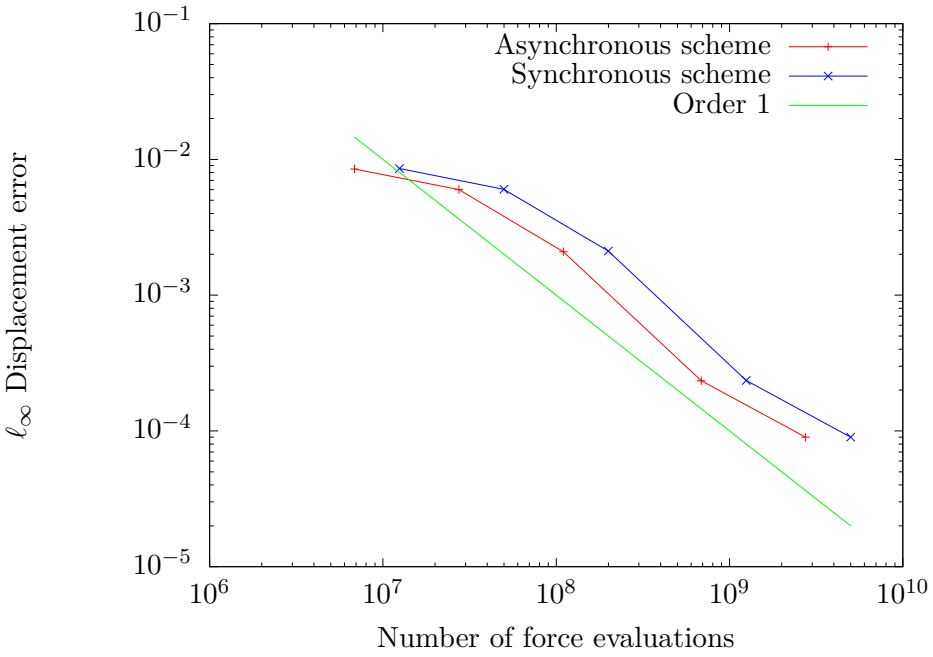


Figure 2.16 – Inhomogeneous wave propagation: Maximal displacement error at $t = 0.5$ against the number of force evaluations for the asynchronous and the synchronous schemes

Chapter 3

A variational discrete element method for quasi-static and dynamic elasto-plasticity

Contents

The material in this chapter has been submitted to IJNME [Marazzato et al., 2019b].

3.1 Introduction

Discrete element methods (DEM) constitute a large class of particle methods which have originally been used for crystalline materials [Hoover et al., 1974] and geotechnical problems [Cundall and Strack, 1979] and have found applications in granular materials, soil and rock mechanics. In their original formulation, DEM consisted in representing a domain by small spherical particles interacting by means of forces and torques. A wide range of models for the expression of these bonds has been developed depending on the material constitutive law. Computing the deformation of the domain then consists in computing the evolution of the particle system. Advantages of DEM are their ability to deal with discontinuous materials, such as fractured or porous materials, as well as the possibility to take advantage of GPU computations [Spellings et al., 2017]. Other, similar, particle methods have been derived in the context of Smooth Particles Hydrodynamics (SPH) methods, which require an

interaction kernel [Han and Meng, 2001]. The main difficulty in DEM consists in deriving a correct set of forces between elements to discretize the continuous equations (in the present case, dynamic elasto-plasticity). DEM originally used sphere packing to discretize the domain [Labra and Oñate, 2009] and were forced to fit parameters in order to obtain relevant values for the Young modulus E or the Poisson ratio ν [Jebahi et al., 2015, Celigueta et al., 2017]. Moreover simulating a material with a Poisson ratio ν larger than 0.3 met with difficulties [André et al., 2013]. Note also the possibility to use DEM only in a limited zone, where crack occurs for instance, in order to mitigate these issues. For example, a modified DEM (MDEM) has been coupled with a consistent virtual element method (VEM) for elasticity to discretize fracturing porous media [Nilsen et al., 2017].

A discrete element method was developed in [Monasse and Mariotti, 2012] and was formally proved to be consistent with Cauchy elasticity. A first attractive feature of this method was that the discrete force parameters were directly derived from the Young modulus and the Poisson ratio without the need for a fitting process. Moreover the method exhibited robustness in the incompressibility limit ($\nu \to 0.5$). Similar ideas have been used to handle brittle fracture [André et al., 2019]. However several limitations remain in this approach. First the evaluation of the forces between particles hinges on the use of a Voronoi mesh and does not adapt to general (not even tetrahedral) meshes. This is due to a nearest-neighbour evaluation of the gradient on a facet of the mesh (known in the Finite Volume community as the "two-point flux problem"). Secondly the expression of the forces for a Cauchy continuum cannot be readily extended to more general behaviour laws. Finally the convergence proof is mostly formal (on a Cartesian grid) and no convergence proof is given on general (Voronoi) meshes, apart from numerical evidence.

The main goal of the present contribution is to circumvent the above issues by extending the discrete element method of [Monasse and Mariotti, 2012] to general polyhedral meshes and elasto-plastic behaviour laws. The present scheme hinges on volumetric unknowns representing the (vector-valued) displacement and the (tensor-valued) plastic strain in each mesh cell. In addition (vector-valued) displacement unknowns at the boundary vertices are also considered. The key ingredient of the method is a piecewise constant gradient reconstruction in each mesh cell evaluated from local displacement reconstructions at the facets of the mesh. These reconstructions at the facets are, in turn, evaluated from the displacement unknowns attached to the mesh cells and the boundary vertices (the latter are only introduced to facilitate these reconstructions, as further discussed below). We devise the scheme for both quasi-static and dynamic elasto-plasticity, and in the latter situation we perform the time discretization using the explicit pseudo-energy conserving time-integration method developed in [Marazzato et al., 2019a]. Notice that a diagonal mass matrix is always employed.

The present scheme can be viewed as a variational DEM. As such, it shares a number of properties with finite volume [Eymard et al., 2004, Eymard et al., 2009] and lowest-order discontinuous Galerkin (dG) methods [Di Pietro, 2012]. The main difference with [Eymard et al., 2009] is in the mass distribution of the method. Our choice is motivated by the fact that, to use an explicit integration with a lumped mass matrix, every dof needs a mass to compute its velocity, even dofs on boundary facets. The function reconstruction of [Eymard et al., 2009] uses only cell dofs and not boundary facet dofs. Incidentally we point out that the convergence of the present variational DEM can be studied using the framework of gradient discretization methods (GDM) [Droniou et al., 2018]. GDM lead to a unified and powerful framework allowing one to prove convergence and error estimates for a wide range of numerical schemes. We also observe that the present variational DEM is not a mesh-free method (contrary to classical DEM) since it uses the mesh geometry to compute the above reconstructions. However, it can still be loosely viewed as a particle method owing to the use of a diagonal mass matrix and of mostly cell unknowns.

This paper is organised as follows. Section 3.2 briefly recalls the equations of dynamic elasto-plasticity in a Cauchy continuum. Section 3.3 introduces the proposed DEM and presents

the space discretization of the governing equations. Some numerical tests to verify the convergence of the space discretization in a steady setting are reported. Section 3.4 deals with the DEM discretization for quasi-static elasto-plasticity and presents test cases in two and three space dimensions. Section 3.5 addresses the time discretization of the dynamic elasto-plasticity problem using the explicit pseudo-energy conserving time-integrator developed in [Marazzato et al., 2019a]. This section also assesses the coupled DEM and time discretization on test cases in three space dimensions. Finally Section 3.6 draws some conclusions.

3.2 Governing equations for dynamic elasto-plasticity

We consider an elasto-plastic material occupying the domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, in the reference configuration and evolving dynamically on the finite time interval $(0,T)$, $T>0$, under the action of volumetric forces and boundary conditions. The strain regime is restricted to small strains so that the linearized strain tensor is $\varepsilon(u) := \frac{1}{2}(\nabla u + (\nabla u)^\mathsf{T})$, where u is the displacement field. The plastic constitutive law hinges on a von Mises criterion with nonlinear isotropic hardening. The material is supposed to be homogeneous, isotropic and rate-independent. The presented formalism can be extended to the case of anisotropic, inhomogeneous, rate-dependent, anisothermal materials as well as finite strains. The stress tensor $\sigma \in \mathbb{R}^{d\times d}_{\text{sym}}$ is such that

$$\sigma := \mathbb{C} : (\varepsilon(u) - \varepsilon_p), \tag{3.1}$$

where $\mathbb{C}$ is the fourth-order stiffness tensor, the subscript sym stands for symmetric tensors and $\varepsilon_p \in \mathbb{R}^{d\times d}_{\text{sym}}$ is the (trace-free) tensor of remanent plastic strain. The von Mises yield function φ is given by

$$\varphi(\sigma,p) := \sqrt{\frac{3}{2}}|\text{dev}(\sigma)| - (\sigma_0 + R(p)), \tag{3.2}$$

where $\text{dev}(\sigma)$ is the deviatoric part of σ and $|\tau| = (\sum_{i,j=1}^d \tau_{ij}^2)^{\frac{1}{2}}$ for a second-order tensor τ, p is the scalar cumulated plastic deformation, $R(p) := \frac{d\omega_p}{dp}$ where the function ω_p is the part of the Helmholtz free energy related to isotropic hardening, and σ_0 is the initial yield stress, so that the actual yield stress is $\sigma_0 + R(p)$. Admissible states are characterized by the inequality $\varphi(\sigma,p) \le 0$, the material is in the elastic domain if $\varphi(\sigma,p) < 0$ and in the plastic domain if $\varphi(\sigma,p) = 0$.

In strong form, the dynamic elasto-plasticity equations consist in searching for the displacement field $u : (0,T)\times\Omega \to \mathbb{R}^d$, the remanent plastic strain tensor $\varepsilon_p : (0,T)\times\Omega \to \mathbb{R}^{d\times d}_{\text{sym}}$, and the scalar cumulated plastic deformation $p : (0,T)\times\Omega \to \mathbb{R}$ such that the following equations hold in Ω for all $t \in (0,T)$:

$$\begin{cases} \rho\ddot{u} - \text{div}(\sigma) = f, \\ \varphi(\sigma,p) \le 0, \\ \lambda = \dot{p} \ge 0, \quad \lambda\varphi(\sigma,p) = 0, \quad \dot{\varepsilon}_p = \lambda\dfrac{\partial\varphi}{\partial\sigma}(\sigma), \end{cases} \tag{3.3}$$

where $\rho > 0$ is the density of the material, dots indicate time derivatives, f is the imposed volumetric force, and λ is the Lagrange multiplier associated with the constraint $\varphi(\sigma,p) \le 0$. Note that owing to (3.2), we have $\dot{\varepsilon}_p = \lambda\sqrt{\frac{3}{2}}\frac{\text{dev}(\sigma)}{|\text{dev}(\sigma)|}$, so that $\dot{p} = \lambda = \sqrt{\frac{2}{3}}|\dot{\varepsilon}_p|$.

Let $\partial\Omega = \partial\Omega_N \cup \partial\Omega_D$ be a partition of the boundary of Ω. By convention $\partial\Omega_D$ is a closed set and $\partial\Omega_N$ is a relatively open set in $\partial\Omega$. The boundary $\partial\Omega_D$ has an imposed displacement u_D, whereas a normal stress g is imposed on $\partial\Omega_N$, i.e. we enforce

$$u = u_D \ \text{ on } (0,T)\times\partial\Omega_D, \qquad \sigma\cdot n = g_N \ \text{ on } (0,T)\times\partial\Omega_N. \tag{3.4}$$

Note that u_D and g_N can be time-dependent. Finally the initial conditions prescribe that $u(0) = u_0$, $\dot{u}(0) = v_0$ and $p(0) = 0$ in Ω.

To formulate the governing equations (3.3) together with the Neumann boundary condition from (3.4) in weak form, we consider time-dependent functions with values in space-dependent functional spaces. Let us set

$$V_D := \left\{ v \in H^1(\Omega;\mathbb{R}^d) \mid v_{|\partial\Omega_D} = u_D \right\}, \qquad V_0 := \left\{ v \in H^1(\Omega;\mathbb{R}^d) \mid v_{|\partial\Omega_D} = 0 \right\}. \tag{3.5}$$

(Note that the space V_D is actually time-dependent if the Dirichlet data is time-dependent.) We also set

$$Q := L^2(\Omega;\mathbb{R}^{d\times d}_{\mathrm{sym}}), \qquad Q_0 := \{\eta_p \in Q \mid \mathrm{tr}(\eta_p) = 0\}, \tag{3.6}$$

where $P := L^2(\Omega)$. Here, for any vector space S, $L^2(\Omega;S)$ is the Hilbert space composed of S-valued square-integrable functions in Ω, and $H^1(\Omega;S)$ is the subspace of $L^2(\Omega;S)$ composed of those functions whose weak gradient is also square-integrable. All of the above functional spaces are equipped with their natural inner product. Then the weak solution is searched as a triple $(u,\varepsilon_p,p) : (0,T) \to V_D \times Q_0 \times P$. To alleviate the mathematical formalism, we do not specify here the regularity in time (see [Han and Reddy, 2012] and [Carstensen, 1999]). We introduce the mass bilinear form such that

$$m(a,\tilde{v}) := \langle \rho a, \tilde{v} \rangle_{V_0',V_0}, \quad \forall (a,\tilde{v}) \in V_0' \times V_0, \tag{3.7}$$

where V_0' denotes the dual space of V_0 and $\langle\cdot,\cdot\rangle_{V_0',V_0}$ the duality pairing, and the stiffness bilinear form parameterized by a member $\eta_p \in Q_0$ such that

$$a(\eta_p; v, \tilde{v}) := \int_\Omega (\mathbb{C} : (\varepsilon(v) - \eta_p)) : \varepsilon(\tilde{v}), \quad \forall (v,\tilde{v}) \in V_D \times V_0, \tag{3.8}$$

The governing equations (3.3) are rewritten as follows: Find $(u,\varepsilon_p,p) : (0,T) \to V_D \times Q_0 \times P$ such that, for all $t \in (0,T)$,

$$\begin{cases} m(\ddot{u}(t),\tilde{v}) + a(\varepsilon_p(t); u(t), \tilde{v}) = l(t;\tilde{v}), & \forall \tilde{v} \in V_0, \\ \varphi(\sigma,p) \le 0, & \text{in } \Omega, \\ \lambda = \dot{p} \ge 0, \quad \lambda\varphi(\sigma,p) = 0, \quad \dot{\varepsilon}_p = \lambda \dfrac{\partial \varphi}{\partial \sigma}(\sigma), & \text{in } \Omega, \end{cases} \tag{3.9}$$

where the time-dependency is left implicit in the second and third equations, and with the linear form $l(t;\cdot)$ acting on V_0 as follows:

$$l(t;\tilde{v}) := \int_\Omega f(t) \cdot \tilde{v} + \int_{\partial\Omega_N} g_N(t) \cdot \tilde{v}. \tag{3.10}$$

Note that the Dirichlet condition is enforced strongly, whereas the Neumann condition is enforced weakly. Define the elastic energy $E_{\mathrm{elas}}(t) := \frac{1}{2}\int_\Omega \sigma(t) : \mathbb{C}^{-1} : \sigma(t)$ with $\sigma(t) := \mathbb{C} : (\varepsilon(u(t)) - \varepsilon_p(t))$, the kinetic energy $E_{\mathrm{kin}}(t) := \frac{1}{2} m(\dot{u}(t),\dot{u}(t))$, the plastic dissipation $E_{\mathrm{plas}}(t) := \int_\Omega \sigma_0 p(t) + \omega_p(p(t))$, and the work of external loads $E_{\mathrm{ext}}(t) := \int_0^t l(s;\dot{u}(s))ds$. Then assuming for simplicity a homogeneous Dirichlet condition, we have the following energy equation:

$$E_{\mathrm{elas}}(t) + E_{\mathrm{kin}}(t) + E_{\mathrm{plas}}(t) = E_{\mathrm{ext}}(t), \tag{3.11}$$

showing that the total energy at time t is balanced with the work of external loads up to time t.

3.3 Space semi-discretization

In this section we present the DEM space semi-discretization of the weak formulation (3.9), and we present a few verification test cases for static linear elasticity.

3.3.1 Degrees of freedom

The domain Ω is discretized with a mesh $\mathcal{T}_h$ of size h made of polyhedra with planar facets in three space dimensions or polygons with straight edges in two space dimensions. We assume that Ω is itself a polyhedron or a polygon so that the mesh covers Ω exactly, and we also assume that the mesh is compatible with the partition of the boundary $\partial\Omega$ into the Dirichlet and Neumann parts.

Let $\mathcal{C}$ denote the set of mesh cells and $\mathcal{Z}^\partial$ the set of mesh vertices sitting on the boundary of Ω. Vector-valued volumetric degrees of freedom (dofs) for a generic displacement field $v_\mathcal{C} := (v_c)_{c\in\mathcal{C}} \in \mathbb{R}^{d\#(\mathcal{C})}$ are placed at the barycentre of every mesh cell $c \in \mathcal{C}$, where $\#(S)$ denotes the cardinality of any set S. Additional vector-valued boundary degrees of freedom $v_{\mathcal{Z}^\partial} := (v_z)_{z\in\mathcal{Z}^\partial} \in \mathbb{R}^{d\#(\mathcal{Z}^\partial)}$ for the displacement are added at every boundary vertex $z \in \mathcal{Z}^\partial$. The reason why we introduce boundary vertex dofs is motivated in Section 3.3.3. These dofs are also used to enforce the Dirichlet condition on $\partial\Omega_D$. We use the compact notation $v_h := (v_\mathcal{C}, v_{\mathcal{Z}^\partial})$ for the collection of all the cell dofs and all the boundary vertex dofs. Figure 3.1 illustrates the position of the displacement dofs. In addition a (trace-free) symmetric tensor-valued dof representing the internal plasticity variable $\eta_{p,c} \in \mathbb{R}^{d\times d}_{\text{sym}}$ is attached to every mesh cell $c \in \mathcal{C}$, as well as a scalar dof p_c representing the cumulated plastic deformation. We write $\eta_{p,\mathcal{C}} := (\eta_{p,c})_{c\in\mathcal{C}} \in Q_h := \left(\mathbb{R}^{d\times d}_{\text{sym}}\right)^{\#(\mathcal{C})}$ and $p_\mathcal{C} := (p_c)_{c\in\mathcal{C}} \in P_h := \mathbb{R}^{\#(\mathcal{C})}$.

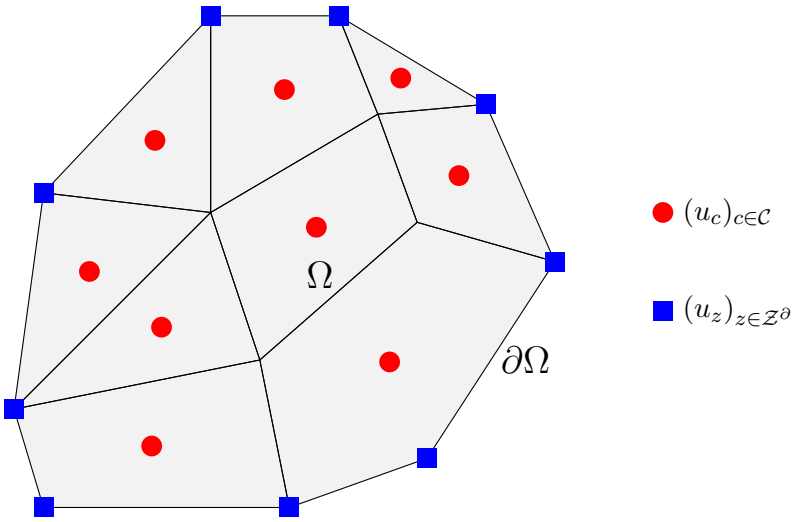


Figure 3.1 – Continuum Ω covered by a polyhedral mesh and vector-valued degrees of freedom for the displacement.

Let $\mathcal{F}$ denote the set of mesh facets. We partition this set as $\mathcal{F} = \mathcal{F}^i \cup \mathcal{F}^b$, where $\mathcal{F}^i$ is the collection of the internal facets shared by two mesh cells and $\mathcal{F}^b$ is the collection of the boundary facets sitting on the boundary $\partial\Omega$ (such facets belong to the boundary of only one mesh cell). Using the cell and boundary-vertex dofs introduced above, we reconstruct a collection of displacements $v_\mathcal{F} := (v_F)_{F\in\mathcal{F}} \in \mathbb{R}^{d\#(\mathcal{F})}$ on the mesh facets. The facet reconstruction operator is denoted $\mathcal{R}$ and we write

$$v_\mathcal{F} := \mathcal{R}(v_h). \tag{3.12}$$

The precise definition of the facet reconstruction operator is given in Section 3.3.3. Using the reconstructed facet displacements and a discrete Stokes formula, it is possible to devise a discrete $\mathbb{R}^{d\times d}$-valued piecewise-constant gradient field for the displacement that we write $G_\mathcal{C}(v_\mathcal{F}) := (G_c(v_\mathcal{F}))_{c\in\mathcal{C}} \in \mathbb{R}^{d^2\#(\mathcal{C})}$. Specifically we set in every mesh cell $c \in \mathcal{C}$,

$$G_c(v_\mathcal{F}) := \sum_{F\in\partial c} \frac{|F|}{|c|} v_F \otimes n_{F,c}, \qquad \forall v_\mathcal{F} \in \mathbb{R}^{d\#(\mathcal{F})}, \tag{3.13}$$

where the summation is over the facets F of c and $n_{F,c}$ is the outward normal to c on F. Note that for all $v_h \in V_h$, we have

$$G_c(\mathcal{R}(v_h)) = \sum_{F\in\partial c} \frac{|F|}{|c|}(\mathcal{R}(v_h)_F - v_c) \otimes n_{F,c}, \tag{3.14}$$

since $\sum_{F\in\partial c} |F| n_{F,c} = 0$. We define a constant linearized strain tensor in every mesh cell $c \in \mathcal{C}$ such that

$$\varepsilon_c(v_{\mathcal{F}}) := \frac{1}{2}(G_c(v_{\mathcal{F}}) + G_c(v_{\mathcal{F}})^{\mathsf{T}}) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{3.15}$$

Finally, we define two additional reconstructions. The first is a cellwise nonconforming P^1 reconstruction $\mathfrak{R}$ defined for all $c \in \mathcal{C}$ by

$$\mathfrak{R}(v_h)_c(\mathbf{x}) := v_c + G_c(\mathcal{R}(v_h)) \cdot (\mathbf{x} - \mathbf{x}_c), \tag{3.16}$$

where $\mathbf{x} \in c$ and $\mathbf{x}_c$ if the barycentre of the cell c. The second is a P^1 reconstruction on boundary facets, written $\mathfrak{R}^\partial(v_{\mathcal{Z}^\partial})$, and computed using the vertex dofs of the boundary facets. In case of simplicial facets, it reduces to classical P^1 Lagrange interpolation. For non-simplicial facets, generalised barycentric coordibnates need to be used, see [Budninskiy et al., 2016], for instance.

3.3.2 Discrete problem

Let us set $V_h := \mathbb{R}^{d\#(\mathcal{C})} \times \mathbb{R}^{d\#(\mathcal{Z}^\partial)}$ and (recall that $\partial\Omega_D$ is a closed set)

$$\begin{cases} V_{hD} := \{v_h \in V_h \mid v_z = u_D(z)\ \forall z \in \mathcal{Z}^\partial \cap \partial\Omega_D\}, \\ V_{h0} := \{v_h \in V_h \mid v_z = 0\ \forall z \in \mathcal{Z}^\partial \cap \partial\Omega_D\}, \\ W_{hD} := \{v_h \in V_h \mid v_z = \dot{u}_D(z)\ \forall z \in \mathcal{Z}^\partial \cap \partial\Omega_D\}. \end{cases} \tag{3.17}$$

(Note that the spaces V_{hD} and W_{hD} are actually time-dependent if the Dirichlet data is time-dependent.) The discrete stiffness bilinear form is parameterized by a member $\eta_{p,\mathcal{C}} \in Q_h$ and is such that, for all $(v_h, \tilde{v}_h) \in V_{hD} \times V_{h0}$ (compare with (3.8)),

$$a_h(\eta_{p,\mathcal{C}}; v_h, \tilde{v}_h) := \sum_{c\in\mathcal{C}} |c|(\mathbb{C} : (\varepsilon_c(\mathcal{R}(v_h)) - \eta_{p,c})) : \varepsilon_c(\mathcal{R}(\tilde{v}_h)) + s_h(v_h, \tilde{v}_h). \tag{3.18}$$

Here s_h is a weakly consistent stabilization bilinear form intended to render a_h coercive and which is defined on $V_h \times V_h$ as follows:

$$s_h(v_h, \tilde{v}_h) = \sum_{F\in\mathcal{F}} \frac{\eta}{h_F} |F| [\mathfrak{R}(v_h)]_F \cdot [\mathfrak{R}(\tilde{v}_h)]_F \tag{3.19}$$

where h_F is the diameter of the facet $F \in \mathcal{F}$. For an interior facet $F \in \mathcal{F}^i$, writing c_- and c_+ the two mesh cells sharing F, i.e., $F = \partial c_- \cap \partial c_+$, and orienting F by the unit normal vector n_F pointing from c_- to c_+, one has

$$[\mathfrak{R}(v_h)]_F := \mathfrak{R}(v_h)_{c_-}(\mathbf{x}_F) - \mathfrak{R}(v_h)_{c_+}(\mathbf{x}_F). \tag{3.20}$$

The sign of the jump is irrelevant in what follows. The role of the summation over the interior facets in (3.19) is to penalize the jumps of the cell reconstruction $\mathfrak{R}$ across the interior facets. For a boundary facet $F \in \mathcal{F}^b$, we denote c_- the unique mesh cell containing F, we orient F by the unit normal vector $n_F := n_{c_-}$ which points outward Ω, and we define

$$[\mathfrak{R}(v_h)]_F := \mathfrak{R}^\partial(v_{\mathcal{Z}^\partial})_F(\mathbf{x}_F) - \mathfrak{R}(v_h)_{c_-}(\mathbf{x}_F). \tag{3.21}$$

The role of the summation over the boundary facets in (3.19) is to penalize the jumps between the cell reconstruction $\mathfrak{R}$ and the boundary facet reconstruction $\mathfrak{R}^\partial$. The parameter $\eta > 0$

in (3.19) is user-defined with the only requirement that $\eta > 0$. The bilinear form s_h is classical in the context of discontinuous Galerkin methods (see [Arnold, 1982, Di Pietro and Ern, 2011] for instance, see also [Di Pietro, 2012] for cell-centred Galerkin methods). In practice, the penalty parameter η scales as $\eta = \beta\mu$ where μ is the second Lamé coefficient of the material and β is a dimensionless factor that remains user-dependent. Notice that this choice is robust with respect to $\nu \to 0.5$. We present a numerical test in Section 3.3.5 illustrating the moderate impact of β on the numerical computations.

We can next define a discrete mass bilinear form m_h similar to (3.7) and a discrete load linear form $l_h(t)$ similar to (3.10); details are given below. Then the space semi-discrete version of the evolution problem (3.9) amounts to seeking $(u_h, \varepsilon_{p,\mathcal{C}}, p_{\mathcal{C}}) : (0,T) \to V_{hD} \times Q_h \times P_h$ such that, for all $t \in (0,T)$,

$$\begin{cases} m_h(\ddot{u}_h(t), \tilde{v}_h) + a_h(\varepsilon_{p,\mathcal{C}}(t); u_h(t), \tilde{v}_h) = l_h(t, \tilde{v}_h), & \forall \tilde{v}_h \in V_{h0}, \\ \varphi(\Sigma_c(u_h), p_c) \le 0, & \forall c \in \mathcal{C}, \\ \lambda_c = \dot{p}_c \ge 0, \quad \lambda_c \varphi(\Sigma_c(u_h), p_c) = 0, \quad \dot{\varepsilon}_{p,c} = \lambda_c \dfrac{\partial \varphi}{\partial \sigma}(\Sigma_c(u_h)), & \forall c \in \mathcal{C}, \end{cases} \tag{3.22}$$

where the time-dependency is left implicit in the second and third equations and where we introduced in every mesh cell $c \in \mathcal{C}$ the local stress tensor

$$\Sigma_c(u_h) := \mathbb{C} : (\varepsilon_c(\mathcal{R}(u_h)) - \varepsilon_{p,c}) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}. \tag{3.23}$$

Note that the plasticity relations in (3.22) are enforced cellwise, i.e., a mesh cell $c \in \mathcal{C}$ is either in the elastic state or in the plastic state depending on the value of $\varphi(\Sigma_c(u_h), p_c)$. This is due to the fact that stresses are cell-wise constant and thus the second relation of (3.22) can only be enforced cell-wise.

The definition of the discrete mass bilinear form m_h hinges on subdomains to condense the mass associated with the dofs. Figure 3.2 represents our choice for the subdomains. For all the interior cells, the subdomain ω_c is chosen as the whole cell, i.e., $\omega_c = c$. For the boundary vertices and for the cells having a boundary face, a dual barycentric subdomain is constructed, leading to subdomains denoted by ω_z and ω_c, respectively (see Figure 3.2). For the discrete load linear form, we compute averages of the external loads f and g_N in the mesh cells and on the Neumann boundary facets, respectively. As a consequence, $m_h(\cdot,\cdot)$ and $l_h(t;\cdot)$ can be written as follows for all $(v_h, \tilde{v}_h) \in V_h \times V_h$ (compare with (3.7) and (3.10)):

$$m_h(v_h, \tilde{v}_h) := \sum_{z\in\mathcal{Z}^\partial} m_z v_z \cdot \tilde{v}_z + \sum_{c\in\mathcal{C}} m_c v_c \cdot \tilde{v}_c, \tag{3.24}$$

$$l_h(t, \tilde{w}_h) := \sum_{c\in\mathcal{C}} f_c(t) \cdot \tilde{w}_c + \sum_{F\in\mathcal{F}^b\cap\partial\Omega_N} g_F(t) \cdot \mathcal{R}(\tilde{v}_h)_F, \tag{3.25}$$

with $m_z := \int_{\omega_z} \rho$, $m_c := \int_{\omega_c} \rho$, $f_c(t) := \int_c f(t)$ and $g_F(t) := \int_F g_N(t)$.

3.3.3 Reconstruction operator on facets

The reconstruction operator $\mathcal{R}$ is constructed in the same way as in the finite volume methods studied in [Eymard et al., 2009, Sec. 2.2] and in the cell-centered Galerkin methods from [Di Pietro, 2012]. For a given facet $F \in \mathcal{F}$, we select neighbouring boundary vertices collected in a subset denoted $\mathcal{Z}_F^\partial$ and neighbouring cells collected in a subset denoted $\mathcal{C}_F$, as well as coefficients $(\alpha_F^z)_{z\in\mathcal{Z}_F^\partial}$ and $(\alpha_F^c)_{c\in\mathcal{C}_F}$, and we set

$$\mathcal{R}(v_h)_F := \sum_{z\in\mathcal{Z}_F^\partial} \alpha_F^z v_z + \sum_{c\in\mathcal{C}_F} \alpha_F^c v_c, \qquad \forall v_h \in V_h. \tag{3.26}$$

The neighbouring dofs should stay $\mathcal{O}(h)$ close to the facet F. An algorithm is detailed thereafter to explain the selection of the neighbouring dofs in $\mathcal{Z}_F^\partial$ and $\mathcal{C}_F$. The coefficients α_F^z and α_F^c are

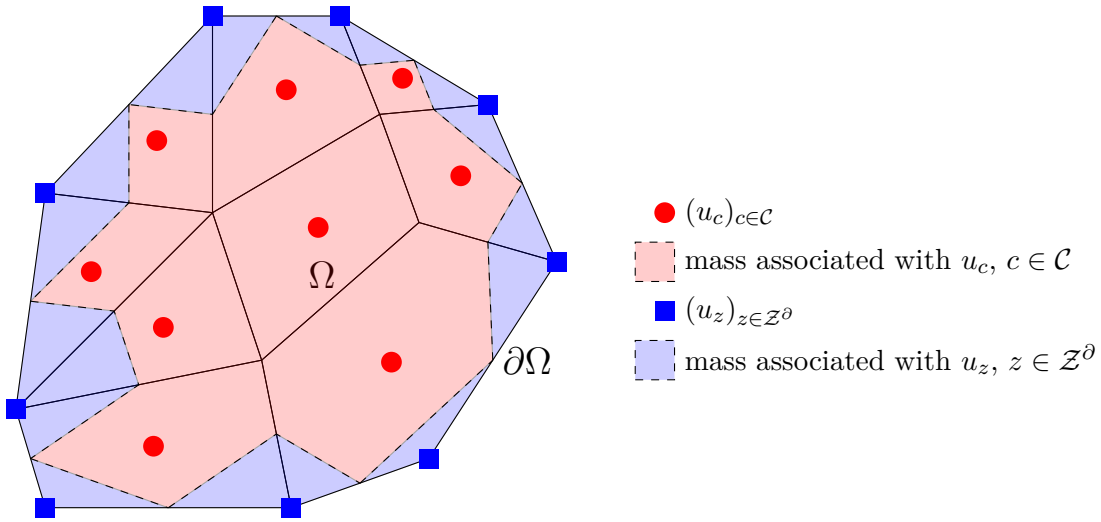


Figure 3.2 – Integration domains to determine the mass associated with the displacement dofs.

chosen as the barycentric coordinates of the facet barycentre $\mathbf{x}_F$ in terms of the location of the boundary vertices in $\mathcal{Z}_F^\partial$ and the barycentres of the cells in $\mathcal{C}_F$. For any interior facet $F \in \mathcal{F}^i$, the set $\mathcal{Z}_F^\partial \cup \mathcal{C}_F$ is constructed so as to contain exactly $(d+1)$ points forming the vertices of a non-degenerate simplex. Thus, the barycentric coefficients $\alpha_F^{\mathbf{z}}$ and α_F^c are computed by solving the linear system:

$$\begin{cases} \displaystyle\sum_{\mathbf{z}\in\mathcal{Z}_F^\partial} \alpha_F^{\mathbf{z}} + \sum_{c\in\mathcal{C}_F} \alpha_F^c = 1, & \forall F \in \mathcal{F}, \\ \displaystyle\sum_{\mathbf{z}\in\mathcal{Z}_F^\partial} \alpha_F^{\mathbf{z}}\mathbf{z} + \sum_{c\in\mathcal{C}_F} \alpha_F^c \mathbf{x}_c = \mathbf{x}_F, & \forall F \in \mathcal{F}, \end{cases} \tag{3.27}$$

where the vertex and its position are written $\mathbf{z}$ and $\mathbf{x}_c$ is the position of the barycentre of the cell c.

The main rationale for choosing the neighboring dofs is to ensure as much as possible that all the coefficients α_F^z or α_F^c lie in the interval $(0,1)$, so that the definition of $\mathcal{R}(v_h)_F$ in (3.26) is based on an interpolation formula (rather than an extrapolation formula if some coefficients lie outside the interval $(0,1)$.) For most internal facets $F \in \mathcal{F}^i$, far from the boundary $\partial\Omega$, it is possible to choose an interpolation-based reconstruction operator using only cell dofs, i.e., we usually have $\mathcal{Z}_F^\partial := \emptyset$. Figure 3.3 presents an example for an interior facet F using three neighbouring cell dofs located at the cell barycentres $\mathbf{x}^i$, $\mathbf{x}^j$ and $\mathbf{x}^k$. Close to the boundary $\partial\Omega$, the use of boundary vertex dofs helps to prevent extrapolation. In all the cases we considered, interpolation was always possible using the algorithm described below.

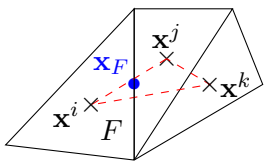


Figure 3.3 – Dofs associated with the interior facet F used for the reconstruction.

On the boundary facets, the reconstruction operator only uses interpolation from the boundary vertices of the facet, i.e., we always set $\mathcal{C}_F := \emptyset$ for all $F \in \mathcal{F}^b$. In three space dimensions, the facet can be polygonal and the barycentric coordinates are generalized barycentric coordinates. This is achieved using [Budninskiy et al., 2016] and the package 2D Triangulation of the geometric library CGAL. In the case of simplicial facets (triangles in three space dimensions

and segments in two space dimensions), we recover the classical barycentric interpolation as described above.

The advantage of using interpolation rather than extrapolation is relevant in the context of explicit time-marching schemes where the time step is restricted by a CFL condition depending on the largest eigenvalue $\lambda_{\max}$ of the stiffness matrix associated with the discrete bilinear form $a_h(0;\cdot,\cdot)$ (see, e.g., (3.48)). It turns out that using extrapolation can have an adverse effect on the maximal eigenvalue of the stiffness matrix, thereby placing a severe restriction on the time step, and this restriction is significantly alleviated if enough neighboring dofs are used in (3.26) to ensure that interpolation is being used. We refer the reader to Table 3.1 below for an illustration.

Let us now briefly outline the algorithm used for selecting the reconstruction dofs associated with a given mesh inner facet $F \in \mathcal{F}^i$. This algorithm has to be viewed more as a proof-of-concept than as an optimized algorithm, and we observe that this algorithm is only used in a pre-processing stage of the computations. Fix an integer $I \geq d+1$.

1. Compute a list of points $(\mathbf{x}_i)_{1\leq i\leq I}$ ordered by increasing distance to $\mathbf{x}_F$; each point can be either the barycentre of a mesh cell or a boundary vertex. To this purpose, we use the KDTree structure of the `scipy.spatial` module of Python.
2. Check if $\mathbf{x}_F$ lies in the convex hull of the set $(\mathbf{x}_i)_{1\leq i\leq I}$. To this purpose we use the ConvexHull structure of `scipy.spatial`. If that is not true, then extrapolation must be used. Otherwise find a subset of $(\mathbf{x}_i)_{1\leq i\leq I}$ containing $(d+1)$ points and use the barycentric coordinates of the resulting simplex to evaluate the interpolation coefficients to be used in (3.26).

Note that I must be large enough to enable the recovery of at least one simplex per inner facet that is not too flat, independently of the use of extrapolation or interpolation. In our computations, we generally took $I = 10$ if $d = 2$ and $I = 25$ if $d = 3$.

To illustrate the performance of the above algorithm in alleviating time step restrictions based on a CFL condition, we report in Table 3.1 the largest eigenvalue $\lambda_{\max}$ of the stiffness matrix and the percentage of the mesh facets where extrapolation has to be used as a function of the integer parameter I from the above algorithm. The results are obtained on the two three-dimensional meshes (called "coarse" and "fine") considered in Section 3.5.3 together with the DEM discretization for the dynamic elasto-plastic evolution of a beam undergoing flexion. Note that the minimal value is $I = 7$ for the coarse mesh and $I = 9$ for the fine mesh.

mesh	$I=7$		$I=9$		$I=12$		$I=15$		$I=18$	
coarse	$4{\cdot}10^{10}$	15%	$6{\cdot}10^{09}$	4%	$7{\cdot}10^{07}$	0.8%	$2{\cdot}10^{05}$	0%	$2{\cdot}10^{05}$	0%
fine	-	-	$7{\cdot}10^{09}$	1.6%	$3{\cdot}10^{07}$	0.1%	$1{\cdot}10^{07}$	0.01%	$8{\cdot}10^{05}$	0%

Table 3.1 – Largest eigenvalue of the stiffness matrix and percentage of inner facets with extrapolation for various values of the parameter I on the coarse and fine meshes used in the DEM discretizations reported in Section 3.5.3.

3.3.4 Interpretation as a Discrete Element Method

In this section we rewrite the first equation in (3.22) as a particle method by introducing the dofs of the discrete displacement $u_h(t) \in V_{hD}$ attached to the mesh cells and to the boundary vertices lying on the Neumann boundary, which we write $U_{\mathrm{DEM}} := (U_p(t))_{p\in\mathcal{P}}$ with $\mathcal{P} := \mathcal{C}\cup\mathcal{Z}_N^\partial$ and $\mathcal{Z}_N^\partial := \{z \in \mathcal{Z}^\partial \mid z \in \partial\Omega_N\}$. Here $\mathcal{P}$ can be viewed as the indexing set for the set of particles. Note that U_{DEM} is a collection of dof values forming a column vector, whereas u_h is a piecewise-constant function. Recalling the definition of the discrete mass bilinear form, we set $m_p := \int_{\omega_c}\rho$ if $p = c \in \mathcal{C}$ and $m_p := \int_{\omega_{\mathbf{z}}}\rho$ if $p = \mathbf{z} \in \mathcal{Z}_N^\partial$. Concerning the external loads, we set $F_{\mathrm{DEM}}(t) := (F_p(t))_{p\in\mathcal{P}}$ with $F_p(t) := f_c(t) = \int_c f(t)$ if $p = c$ and $F_p(t) := \sum_{F\in\mathcal{F}_{\mathbf{z}}}\alpha_F^{\mathbf{z}} g_F(t) =$

$\sum_{F\in\mathcal{F}_\mathbf{z}}\alpha_F^\mathbf{z}\int_F g_N(t)$ if $p=\mathbf{z}$, where $\mathcal{F}_\mathbf{z}\subset\mathcal{F}^b$ is the collection of the boundary facets to which $\mathbf{z}$ belongs and the coefficients $\alpha_F^\mathbf{z}$ are those used in (3.26) for the facet reconstruction. Since the Neumann boundary $\partial\Omega_N$ is relatively open in $\partial\Omega$, all the facets in $\mathcal{F}_\mathbf{z}$ belong to $\partial\Omega_N$ if $\mathbf{z}\in\mathcal{Z}_N^\partial$.

Recall that $\varepsilon_{p,\mathcal{C}}:(0,T)\to Q_h$ is the discrete tensor of remanent plastic strain. Let us use the shorthand notation $\Sigma_c(t):=\Sigma_c(u_h(t))$ as defined in (3.23), as well as $\Sigma_\mathcal{C}(t):=(\Sigma_c(t))_{c\in\mathcal{C}}$. For a piecewise-constant function defined on the mesh cells, say $w_\mathcal{C}=(w_c)_{c\in\mathcal{C}}$, we define the mean-value $\{w_\mathcal{C}\}_F:=\frac12(w_{c_-}+w_{c_+})$ for all $F=\partial c_-\cap\partial c_+\in\mathcal{F}^i$. Recall that the interior facet F is oriented by the unit normal vector n_F pointing from c_- to c_+ and that the jump across $F\in\mathcal{F}^i$ is defined such that $[w_\mathcal{C}]_F:=w_{c_-}-w_{c_+}$. Recall also that for a boundary facet $F\in\mathcal{F}^b$, c_- denotes the mesh cell to which F belongs and that n_F is the unit normal vector to F pointing outward Ω. Then a direct calculation shows that for all $\tilde v_h\in V_{h0}$,

$$\begin{aligned}-a_h(\varepsilon_{p,\mathcal{C}}(t);u_h(t),\tilde v_h)=&\sum_{F\in\mathcal{F}^i}|F|(\{\Sigma_\mathcal{C}(t)\}_F{\cdot}n_F)\cdot[\tilde v_\mathcal{C}]_F\\&+\sum_{F\in\mathcal{F}^i}|F|([\Sigma_\mathcal{C}(t)]_F{\cdot}n_F)\cdot(\{\tilde v_\mathcal{C}\}_F-\mathcal{R}(\tilde v_h)_F)\\&+\sum_{F\in\mathcal{F}^b}|F|(\Sigma_{c_-}(t){\cdot}n_F)\cdot(\tilde v_{c_-}-\mathcal{R}(\tilde v_h)_F)\\&-\sum_{F\in\mathcal{F}}\frac{\eta}{h_F}|F|[\mathfrak{R}(u_h(t))]_F\cdot[\mathfrak{R}(\tilde v_h)]_F.\end{aligned}\tag{3.28}$$

To simplify some expressions, we are going to neglect the second term on the above right-hand side since this term is of higher-order (it is essentially the product of two jumps). Recall that, by definition, the reconstruction operator $\mathcal{R}$ on a boundary facet $F\in\mathcal{F}^b$ makes use only of the vertex dofs of that facet. Then, letting $(\tilde V_p)_{p\in\mathcal{P}}$ be the collection of the dofs of the discrete test function $\tilde v_h$, we infer that

$$-a_h(\varepsilon_{p,\mathcal{C}}(t);u_h(t),\tilde v_h)\simeq\sum_{p\in\mathcal{P}}\Phi_p^{\rm ep}(t)\cdot\tilde V_p+\Phi_p^{\rm pen}(t)\cdot\tilde V_p,\tag{3.29}$$

where $\Phi_p^{\rm ep}(t)$ is the elasto-plastic force acting on the particle $p\in\mathcal{P}$ and $\Phi_p^{\rm pen}(t)$ is the force induced by the penalty and acting on the same particle. For all $c\in\mathcal{C}$, we have $\Phi_c^{\rm ep}(t):=\sum_{F\in\mathcal{F}_c^i\cup\mathcal{F}_c^N}\Phi_{c,F}^{\rm ep}(t)$ with $\mathcal{F}_c^i:=\{F\in\mathcal{F}^i\mid F\subset\partial c\}$, $\mathcal{F}_c^N:=\{F\in\mathcal{F}^b\mid F\subset\partial c\cap\partial\Omega_N\}$, and

$$\Phi_{c,F}^{\rm ep}(t):=\begin{cases}\iota_{c,F}|F|\{\Sigma_\mathcal{C}(t)\}_F{\cdot}n_F&\text{if }F\in\mathcal{F}_c^i,\\|F|\Sigma_{c_-}(t){\cdot}n_F&\text{if }F\in\mathcal{F}_c^N,\end{cases}\tag{3.30}$$

with $\iota_{c,F}:=n_c\cdot n_F$, and for all $z\in\mathcal{Z}_N^\partial$, we have

$$\Phi_z^{\rm ep}(t):=-\sum_{F\in\mathcal{F}_z}\alpha_F^z\Phi_{c_-,F}^{\rm ep}(t),\tag{3.31}$$

with $\Phi_{c_-,F}^{\rm ep}$ defined in (3.30) (recall that c_- denotes the unique mesh cell containing the boundary facet $F\in\mathcal{F}^b$). Note that the principle of action and reaction is encoded in the fact that $\iota_{c_-,F}+\iota_{c_+,F}=0$ for all $F=\partial c_-\cap\partial c_+\in\mathcal{F}^i$.

Remark 3.1 (Physical forces)**.** *The quantities of* (3.30) *are remarkable in the sense that they correspond to the physical force that one expects between two particles: the average of the stress in each particle multiplied by the surface shared by the two particles and contracted with the normal of the corresponding facet. Let us note that this force only depends on the macroscopic material parameters and does not depend on any added microscopic parameter.*

The penalty force is composed of two terms, that is, for all $p \in \mathcal{P}$, $\Phi_p^{\mathrm{pen}}(t) := \Phi_p^{\mathrm{pen},1}(t) + \Phi_p^{\mathrm{pen},2}(t)$. We define the first term for a cell $c \in \mathcal{C}$ and a facet $F \in \mathcal{F}$, such that $F \subset \partial c$, as

$$\Phi_{c,F}^{\mathrm{pen},1}(t) := -\iota_{c,F} \frac{\eta}{h_F} |F| [\mathfrak{R}(u_h(t))]_F, \tag{3.32}$$

and for all $z \in \mathcal{Z}_N^\partial$, as

$$\Phi_z^{\mathrm{pen},1}(t) := -\sum_{F \in \mathcal{F}_z} \alpha_F^z \Phi_{c_-,F}^{\mathrm{pen},1}(t). \tag{3.33}$$

The second term is more intricate since it is a byproduct of the extended stencil of the method. The set of facets using the dof of the particle $p \in \mathcal{P}$ (whether cell or boundary vertex) is defined as $\mathfrak{F}_p$. This means that if $F \in \mathfrak{F}_p$, then one has either $p \in \mathcal{C}_F$ or $p \in \mathcal{Z}_F^\partial$. Let us recall that for a facet $F \in \mathcal{F}$, the cell dofs used for the reconstruction $\mathcal{R}(u_h(t))_F$ are collected in $\mathcal{C}_F$ and the boundary vertex dofs in $\mathcal{Z}_F^\partial$. Then the second penalty term writes for $p \in \mathcal{P}$ and $F \in \mathfrak{F}_p$ as

$$\Phi_{p,F}^{\mathrm{pen},2}(t) := -\sum_{c \text{ s.t. } F \supset \partial c} \sum_{F' \subset \partial c} \iota_{c,F'} \frac{\eta}{h_{F'}} |F'| [\mathfrak{R}(u_h(t))]_{F'} \frac{|F|}{|c|} n_{F,c} \cdot (\mathbf{x}_{F'} - \mathbf{x}_c) \alpha_F^p, \tag{3.34}$$

where α_F^p is the barycentric coordinate associated with the particle $p \in \mathcal{P}$ in the reconstruction over the facet F. As a consequence, the total second penalty term writes:

$$\Phi_p^{\mathrm{pen},2}(t) := \sum_{F \in \mathfrak{F}_p} \Phi_{p,F}^{\mathrm{pen},2}(t) \tag{3.35}$$

Finally, putting everything together, we infer that the the first equation in (3.22) becomes

$$m_p \ddot{U}_p(t) \simeq \Phi_p^{\mathrm{ep}}(t) + \Phi_p^{\mathrm{pen}}(t) + F_p(t), \qquad \forall p \in \mathcal{P}. \tag{3.36}$$

Remark 3.2 (Matrix formulation)**.** *Let us briefly describe the matrix formulation of the space semi-discrete problem (3.22) in the case of elastodynamics, i.e., without plasticity. For simplicity we focus on the pure Neumann problem. A matrix* $\mathbf{R} \in \mathbb{R}^{d\#(\mathcal{F}) \times d\#(\mathcal{P})}$ *corresponding to the reconstruction operator* $\mathcal{R}$ *is first constructed. Its entries are the barycentric coefficients of the dofs used for the reconstruction on the face associated with the given line of* $\mathbf{R}$*. The lines of* $\mathbf{R}$ *associated with boundary facets have, by construction, non-zero entries only for boundary vertices dofs. The linearized strain matrix* $\mathbf{E} \in \left(\mathbb{R}_{\mathrm{sym}}^{d\times d}\right)^{\#(\mathcal{C})} \times \mathbb{R}^{d\#(\mathcal{F})}$ *is composed of the tensorial coefficients* $\frac{1}{2}\frac{|F|}{|c|}(\otimes n_F + n_F \otimes)$ *on the lines associated with the mesh cell* $c \in \mathcal{C}$ *and the columns associated with the facets* $F \subset \partial c$*. The linear elasticity matrix* $\mathbf{C} \in \left(\mathbb{R}_{\mathrm{sym}}^{d\times d}\right)^{\#(\mathcal{C})\times\#(\mathcal{C})}$ *can be written as the block-diagonal matrix where each block corresponds to the double contraction with the fourth-order elastic tensor* $\mathbb{C}$ *and multiplication by* $|c|$*. The jump matrix* $\mathbf{J} \in \mathbb{R}^{d\#(\mathcal{F})\times d\#(\mathcal{P})}$ *discretises the jumps* $[\mathfrak{R}(u_h)]_F$ *on a facet* F*. Its assembly is not detailed for simplicity. However, it is assembled using the connectivity matrix of adjacent cells, the gradient matrix (similar to* $\mathbf{E}$ *but non-symetrized and composed of the tensorial coefficients* $\frac{|F|}{|c|} \otimes n_F$*) and* $\mathbf{R}$*. Denoting* $\mathbf{D} \in \mathbb{R}^{d\#(\mathcal{F})\times d\#(\mathcal{F})}$ *the diagonal matrix with entry* $\frac{\eta}{h_F}|F|$ *for the facet* F*, the stabilization matrix* $\mathbf{S}$ *corresponding to the bilinear form* s_h *can be written* $\mathbf{S} := \mathbf{J}^\mathsf{T}\mathbf{D}\mathbf{J} \in \mathbb{R}^{d\#(\mathcal{P})\times d\#(\mathcal{P})}$*. Finally, denoting* $\mathbf{K} := \mathbf{R}^\mathsf{T}\mathbf{E}^\mathsf{T}\mathbf{C}\mathbf{E}\mathbf{R} + \mathbf{S} \in \mathbb{R}^{d\#(\mathcal{P})\times d\#(\mathcal{P})}$ *the stiffness matrix, the space semi-discrete system (3.22) in the case of elastodynamics reduces to* $\mathbf{M}\ddot{U}_{\mathrm{DEM}}(t) + \mathbf{K}U_{\mathrm{DEM}}(t) = F_{\mathrm{DEM}}(t)$*.*

3.3.5 Convergence tests for linear elasticity

The goal of this section is to briefly verify the correct implementation of the method in the case of static linear elasticity by comparing the numerical predictions using DEM with some

analytical solutions and reporting the orders of convergence on sequences of uniformly refined meshes. The model problem thus consists of finding $u \in V_D$ such that

$$\int_\Omega \varepsilon(\nabla u) : \mathbb{C} : \varepsilon(\nabla \tilde{u}) = \int_\Omega f \cdot \tilde{u}, \qquad \forall \tilde{u} \in V_0. \tag{3.37}$$

The L^2-error is computed as $\|u - \mathfrak{R}(u_h)\|_{L^2(\Omega)}$. The energy error is based on the reconstructed linearized strain of the discrete solution in each mesh cell is computed as $\|u-u_h\|_{\text{en}} := \frac{1}{2}a_h(0, u-u_h, u-u_h)$ where a_h is defined in (3.18). Note that this last term also contains the energy associated with the penalty bilinear form s_h. The convergence rates are approximated as

$$\text{order} = d \log\left(\frac{e_1}{e_2}\right)\left(\log\left(\frac{n_2}{n_1}\right)\right)^{-1}, \tag{3.38}$$

where e_1, e_2 denote the errors on the computations with mesh sizes h_1, h_2 and the number of dofs n_1, n_2.

Choice of penalty factor

In this section, we simulate the torsion of a cylinder with various values of $\eta = \beta\mu$ obtained by varying β. We compare the results to first-order penalised Crouzeix–Raviart finite elements (FE) [Hansbo and Larson, 2003]. A mesh of size $h = 29\text{mm}$ is chosen for both computations. The DEM computation contains $6,332$ vectorial displacement dofs and the Crouzeix–Raviart $11,760$. The geometry and boundary conditions are similar to Figure 3.7. The material is supposed to be isotropic elastic with $E = 70 \cdot 10^3\text{Pa}$ and $\nu = 0.3$. The imposed torsion angle is $\alpha = 2.1 \cdot 10^{-2}\text{rad}$. As the solution of this torsion test consists in pure shear stress, we compare in Figure 3.4 $\text{tr}(\sigma)$ and the Von Mises stress for DEM with $\beta = 1$, $\beta = 0.1$ and $\beta = 0.01$, and the reference penalised Crouzeix–Raviart computation with $\beta = 1$. We expect the Von Mises stress to be constant on the lateral side of the cylinder and the trace of the stress tensor to be zero.

We observe that the influence of the parameter β is rather light. Indeed, the trace of the trace tensor remains of order 10^1 and the Von Mises stress varies between similar values. We observe that the results with the penalised Crouzeix–Raviart method vary slightly less. We believe that this is due to the higher number of dofs in this computation. As a consequence, so as to have a penalty term of the same order as the elastic term, we choose $\beta = 1$ in our subsequent computations (unless explicitly mentioned).

Manufactured solution

Let us first consider an isotropic two-dimensional elasticity test case in the domain $\Omega = (0,1)^2\text{m}^2$ with the Young modulus $E = 70 \cdot 10^3\text{Pa}$ and the Poisson ratio $\nu = 0.3$. The reference solution is $u(x,y) = \frac{a}{2}(x^2+y^2)(e_x+e_y)$ with $a = 0.8\text{m}^{-1}$ and (e_x, e_y) is the Cartesian basis of $\mathbb{R}^2$. The load term f, which is computed accordingly, is $f(x,y) = -a(\lambda+3\mu)(e_x+e_y)$, where λ and μ are the Lamé coefficients. The corresponding non-homogeneous Dirichlet boundary condition is enforced on the whole boundary. Convergence results are reported in Table 3.2 showing that the energy error converges to first-order with the mesh size and the L^2-error to second-order.

3.4 Quasi-static elasto-plasticity

In this section we present the quasi-static elasto-plasticity problem, its DEM discretization, and we perform numerical tests to assess the methodology.

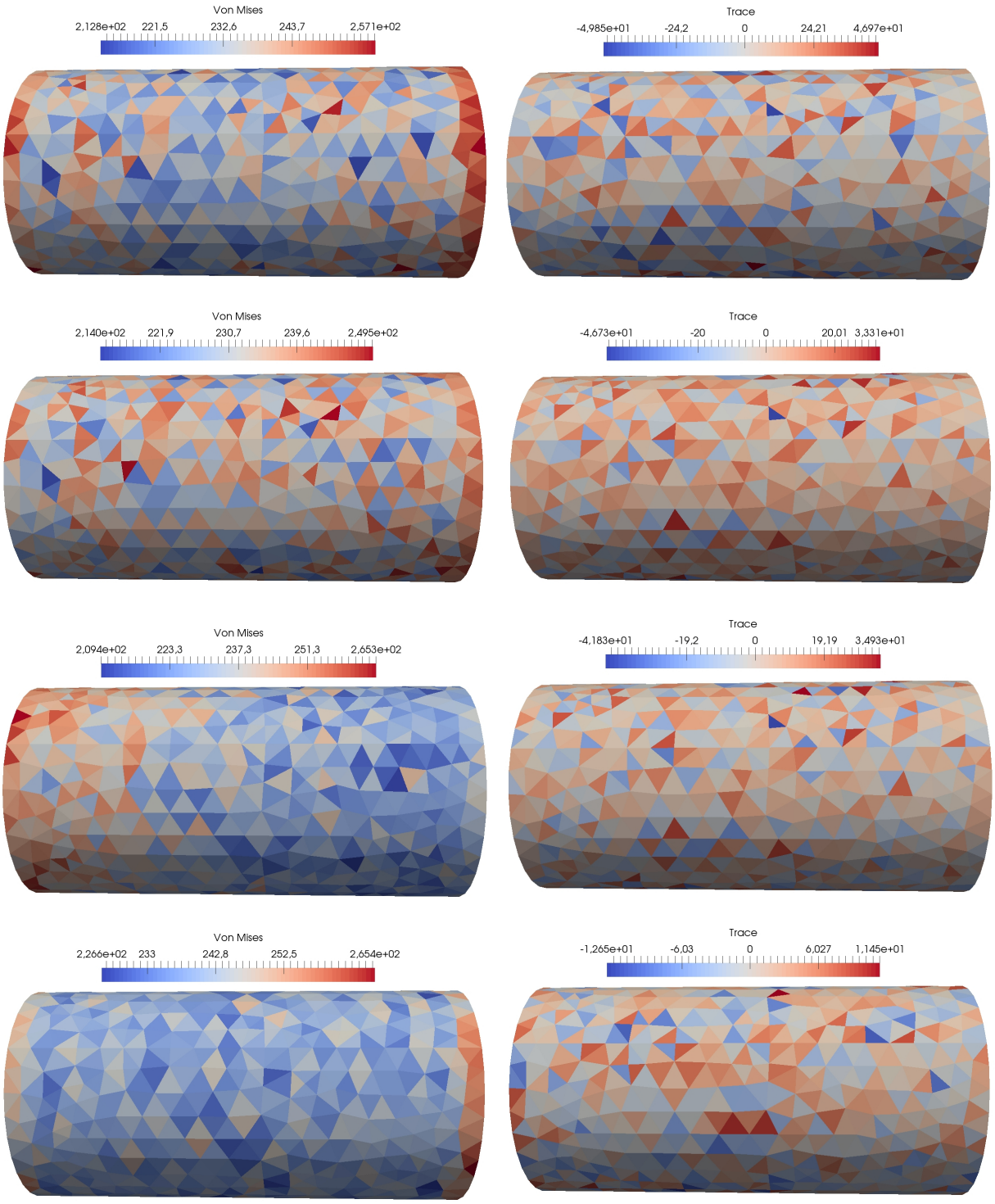


Figure 3.4 – Choice of penalty factor: Left: Von Mises equivalent stress. Right: Trace of the stress tensor. First row: DEM with $\beta = 1$. Second row: DEM with $\beta = 0.1$. Third row: DEM with $\beta = 0.01$. Fourth row: Penalised Crouzeix–Raviart FE with $\beta = 1$ (reference).

h	dofs	$\|u - \mathfrak{R}(u_h)\|_{L^2(\Omega)}$	order	$\|u - u_h\|_{\text{en}}$	order
0.0356	8,928	5.67942e-05	-	4.72e-03	-
0.0177	35,072	8.62031e-06	2.76	1.50e-03	1.60
0.00889	139,008	1.80278e-06	2.27	5.87e-04	1.55

Table 3.2 – Linear elasticity test case: size of the mesh, number of dofs, L^2-error and order of convergence, energy error and order of convergence.

3.4.1 Governing equations

The quasi-static elasto-plastic problem is a simplified formulation of (3.9) where the inertia term in the mass bilinear form is negligible and where the time derivatives are substituted by discrete increments. Thus we consider a sequence of loads $l^n \in V_0'$ for all $n = 1, \dots, N$, and we consider the following sequence of nonlinear problems where $(u^n, \varepsilon_p^n, p^n) \in V_D \times Q_0 \times P$ for all $n = 1, \dots, N$:

$$\begin{cases} a(\varepsilon_p^n; u^n, \tilde{v}) = l^n(\tilde{v}), & \forall \tilde{v} \in V_0, \\ (\varepsilon_p^n, p^n, \mathbb{C}_{\text{ep}}^n, \sigma^n) = \texttt{PLAS\_IMP}(\varepsilon_p^{n-1}, p^{n-1}, \varepsilon^{n-1}, \varepsilon^n), & \text{in } \Omega, \end{cases} \tag{3.39}$$

where $\varepsilon^{n-1} := \varepsilon(u^{n-1})$, $\varepsilon^n := \varepsilon(u^n)$, and where variables with a superscript $^{n-1}$ come from the solution of the quasi-static problem (3.39) at the previous load increment or from a prescribed initial condition if $n = 1$. Given a quadruple $(\varepsilon_p^{\text{old}}, p^{\text{old}}, \varepsilon^{\text{old}}, \varepsilon^{\text{new}})$, the procedure PLAS_IMP returns a quadruple $(\varepsilon_p^{\text{new}}, p^{\text{new}}, \mathbb{C}_{\text{ep}}^{\text{new}}, \sigma^{\text{new}})$ such that

$$\begin{cases} \varphi(\sigma^{\text{new}}, p^{\text{new}}) \leq 0, \\ \lambda^{\text{new}} := \delta p := p^{\text{new}} - p^{\text{old}} \geq 0, \quad \lambda^{\text{new}} \varphi(\sigma^{\text{new}}, p^{\text{new}}) = 0, \\ \delta\varepsilon_p := \varepsilon_p^{\text{new}} - \varepsilon_p^{\text{old}} = \lambda^{\text{new}} \dfrac{\partial \varphi}{\partial \sigma}(\sigma^{\text{new}}). \end{cases} \tag{3.40}$$

Moreover $\sigma^{\text{new}} = \mathbb{C} : (\varepsilon^{\text{new}} - \varepsilon_p^{\text{new}})$ is the new stress tensor, and $\mathbb{C}_{\text{ep}}^{\text{new}}$ is the consistent elasto-plastic modulus [Simo and Taylor, 1985] such that

$$\sigma^{\text{new}} = \sigma^{\text{old}} + \mathbb{C}_{\text{ep}}^{\text{new}} : \delta\varepsilon, \quad \delta\varepsilon := \varepsilon^{\text{new}} - \varepsilon^{\text{old}}, \quad \sigma^{\text{old}} := \mathbb{C} : (\varepsilon^{\text{old}} - \varepsilon_p^{\text{old}}). \tag{3.41}$$

The consistent elastoplastic modulus is instrumental to solve (3.39) iteratively using an implicit radial return mapping technique (close to Newton–Raphson iterations) [Son, 1977]: Starting from $k = 0$, we solve iteratively the linear problem in $u^{n,k} \in V_D$ such that

$$\begin{cases} (\mathbb{C}_{\text{ep}}^{n,k} : \varepsilon(u^{n,k+1}), \varepsilon(\tilde{v}))_Q = r^{n,k}(\tilde{v}) := l^n(\tilde{v}) - (\sigma^{n,k}, \varepsilon(\tilde{v}))_Q, & \forall \tilde{v} \in V_0, \\ (\varepsilon_p^{n,k}, p^{n,k}, \mathbb{C}_{\text{ep}}^{n,k}, \sigma^{n,k}) = \texttt{PLAS\_IMP}(\varepsilon_p^{n-1}, p^{n-1}, \varepsilon^{n-1}, \varepsilon^{n,k}), & \text{in } \Omega, \end{cases} \tag{3.42}$$

where the state for $k = 0$ comes from the previous loading step or the initial condition. Convergence of the iterative process in k is reached when the norm of the residual $r^{n,k}$ is small enough.

3.4.2 DEM space discretization

Using the DEM space discretization, the sequence of quasi-static problems (3.39) amounts to seeking a discrete triple $(u_h^n, \varepsilon_{p,\mathcal{C}}^n, p_{\mathcal{C}}^n) \in V_{hD} \times Q_h \times P_h$ for all $n = 1, \dots, N$, solving the following nonlinear problem:

$$\begin{cases} a_h(\varepsilon_{p,\mathcal{C}}^n; u_h^n, \tilde{v}_h) = l_h^n(\tilde{v}_h), & \forall \tilde{v}_h \in V_{h0}, \\ (\varepsilon_{p,c}^n, p_c^n, \mathbb{C}_{ep,c}^n, \sigma_c^n) = \texttt{PLAS\_IMP}(\varepsilon_{p,c}^{n-1}, p_c^{n-1}, \varepsilon_{p,c}^{n-1}, \varepsilon_c^n), & \forall c \in \mathcal{C}, \end{cases} \tag{3.43}$$

where l_h^n represents a suitable discretization of the load linear form l^n. Using the radial return mapping technique as in (3.42) and starting from $k = 0$, we solve iteratively the linear problem in $u_h^{n,k} \in V_{hD}$ such that

$$\begin{cases} \sum_{c\in\mathcal{C}} |c|(\mathbb{C}_{ep,c}^{n,k} : \varepsilon_c(\mathcal{R}(u_h^{n,k+1}))) : \varepsilon_c(\mathcal{R}(\tilde{v}_h)) = r_{\mathcal{C}}^{n,k}(\tilde{v}_h), & \forall \tilde{v}_h \in V_{h0}, \\ (\varepsilon_{p,c}^{n,k}, p_c^{n,k}, \mathbb{C}_{ep,c}^{n,k}, \sigma_c^{n,k}) = \texttt{PLAS\_IMP}(\varepsilon_{p,c}^{n-1}, p_c^{n-1}, \varepsilon_c^{n-1}, \varepsilon_c^{n,k}), & \forall c \in \mathcal{C}, \end{cases} \tag{3.44}$$

with the residual $r_{\mathcal{C}}^{n,k}(\tilde{v}_h) := l^n(\tilde{v}_h) - \sum_{c\in\mathcal{C}} |c|\sigma_c^{n,k} \cdot \varepsilon_c(\mathcal{R}(\tilde{v}_h))$, and where the discrete state for $k = 0$ comes from the previous loading step or by interpolating the values of the initial condition at the cell barycentres and the boundary vertices. Convergence of the iterative process in k is reached when the norm of the residual $r_{\mathcal{C}}^{n,k}$ is small enough (we use a scaled Euclidean norm).

3.4.3 Numerical tests

This section contains two three-dimensional tests, a beam in quasi-static flexion and a beam in quasi-static torsion, and a two-dimensional test case on the swelling of an infinite cylinder with internal pressure.

Beam in quasi-static traction

A beam of square section $S = 0.016\text{m}^2$ and $L = 1\text{m}$ is stretched by a displacement $u_D(t)$ imposed at its right extremity, whereas the normal displacement and the tangential component of the normal stress are null at the left extremity. An homogeneous Neumann condition ($\sigma \cdot n = 0$) is enforced on the four remaining sides of the beam. Figure 3.5 shows a sketch of the problem setup. The Young modulus is $E = 70 \cdot 10^3\text{Pa}$ and the Poisson ratio $\nu = 0.3$. The yield stress is $\sigma_0 = 250\text{Pa}$, and the material is supposed to be elasto-plastic with linear kinematic hardening. Specifically the tangent plastic modulus is set to $E_t = \frac{1}{5}E$, so that we have $R(p) = Hp$ with $H = \frac{EE_t}{E-E_t}$. The imposed displacement is linearly increased in 20 loading steps from 0 to $2\delta_y$, where $\delta_y = \frac{\sigma_0}{E}L$ is the yield displacement. For this test case the analytical solution is available. Figure 3.6 shows the stress-strain response curve, showing perfect agreement with

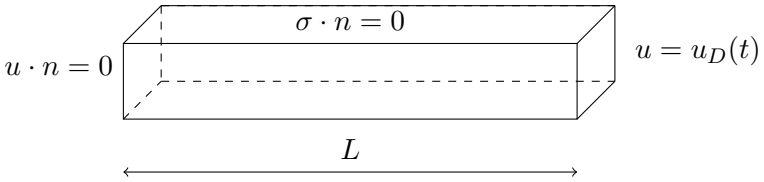


Figure 3.5 – Beam in quasi-static traction: problem setup.

the analytical solution using a mesh of size $h = 0.2\text{m}$. Note that in this test case, the stress tensor is constant in the beam.

Beam in quasi-static torsion

A beam of length $L = 0.2\text{m}$ with a circular section of radius $R = 0.05\text{m}$ is subjected to torsion at one of its extremities. The Young modulus is $E = 70 \cdot 10^3\text{Pa}$ and the Poisson ratio $\nu = 0.3$. The yield stress is $\sigma_0 = 250\text{Pa}$, and the material is supposed to be perfectly plastic so that $R(p) = 0$. The beam is clamped at one of its extremities, a torsion angle $\alpha(t)$ is imposed at the other extremity, and the rest of the boundary of the beam is stress free ($\sigma \cdot n = 0$). Figure 3.7 presents the problem setup. The torsion angle $\alpha(t)$ is increased linearly in 20 loading steps from 0 to $\alpha_{\max} = 2\alpha_y$, where $\alpha_y = \frac{\sigma_0 L}{\mu R\sqrt{3}}$ is the yield angle and μ is the second Lamé coefficient.

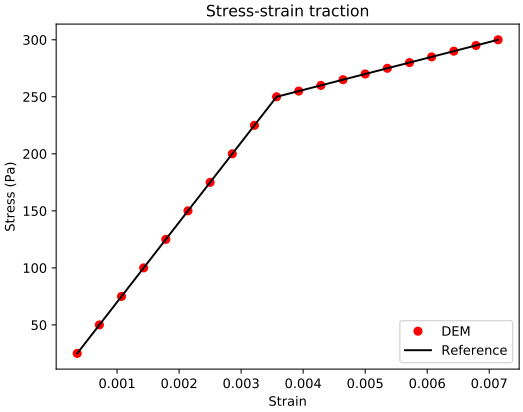


Figure 3.6 – Beam in quasi-static traction: stress-strain response curve for the analytical solution and the DEM solution.

An analytical solution is available in the cylindrical frame (e_r, e_θ, e_z): the displacement field is $u(r,z,t) = \alpha(t) r \frac{z}{L} e_\theta$, and the stress field is $\sigma(r,t) = \tau(r,t)\,(e_\theta \otimes e_z + e_z \otimes e_\theta)$, where

$$\tau(r,t) := \begin{cases} \mu\alpha(t)\dfrac{r}{L}, & 0 \le r \le R\dfrac{\alpha_y}{\alpha(t)}, \\ \sigma_0, & R\dfrac{\alpha_y}{\alpha(t)} \le r \le R. \end{cases}$$

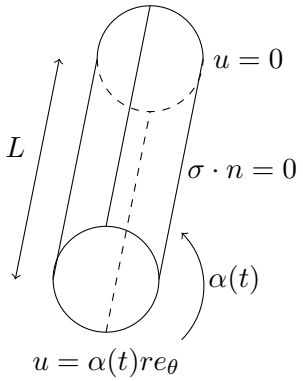


Figure 3.7 – Beam in quasi-static torsion: problem setup.

Table 3.3 reports the maximum L^2-error on the displacement (evaluated as in Section 3.3.5) over the simulation interval and the energy error (including the energy of the penalty terms). The errors are evaluated as described in Section 3.3.5. First-order convergence in the energy norm is observed, as expected. However, full second-order convergence in L^2 norm does not seem satisfied (although the convergence order is still above 1.77). That can be due to the fact that in perfect plasticity $u \notin H^2(\Omega)$ which is a usual hypothesis to obtain full second-order convergence. Figure 3.8 presents the torque-angle response curve for the reference solution and the DEM solution on various meshes, showing good agreement and the convergence of the DEM predictions as the mesh is refined.

Inner swelling of an infinite cylinder

This test case consists in the inner swelling of an infinite cylinder. The inner radius is $R_i = 1\text{m}$ and the outer radius is $R_o = 1.3\text{m}$. Owing to the symmetries, the computation is

h	dofs	$\|u - \mathfrak{R}(u_h)\|_{L^2(\Omega)}$	order	$\|u - u_h\|_{\text{en}}$	order
0.05263	3,753	2.62e-06	-	5.87e-04	-
0.03294	12,726	1.02e-06	2.32	1.97e-04	2.68
0.02871	18,996	7.75e-07	2.04	1.53e-04	1.89
0.01965	47,670	4.50e-07	1.77	8.23e-05	2.02
0.01418	160,146	2.14e-07	1.84	5.36e-05	1.06

Table 3.3 – Beam in quasi-static torsion: size of the mesh, number of dofs, L^2-error and order of convergence, energy error and order of convergence.

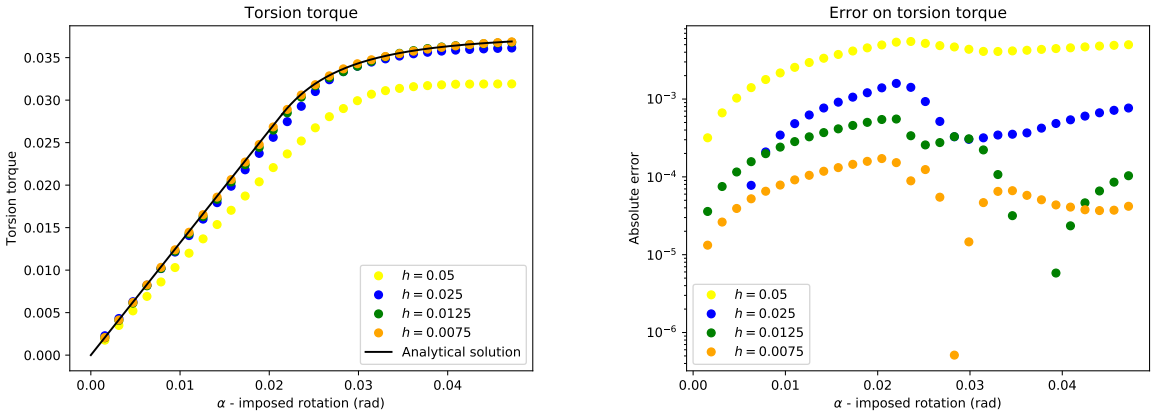


Figure 3.8 – Beam in quasi-static torsion. Left: torque-angle response curve for the analytical solution and the DEM solution on various meshes. Right: difference between the analytical solution and the DEM solution on various meshes.

carried out on a quarter of a planar section of the cylinder with a plane strain formulation. A sketch of the problem setup is presented in Figure 3.9. On the lateral sides of the quarter of cylinder, a null normal displacement and a null tangential component of the normal stress are enforced. The outer side of the cylinder is stress free ($\sigma \cdot n = 0$), and the inner pressure ϖ imposed on the inner side is linearly increased from 0 to $p_{\max} = \frac{2}{\sqrt{3}}\sigma_0 \ln\left(\frac{R_o}{R_i}\right)$, where $\sigma_0 = 250\text{N.m}^{-2}$ is the initial yield stress. The Young modulus and the tangent plastic modulus are set to $E = 70 \cdot 10^3\text{Pa}$ and $E_t = \frac{1}{100}E$.

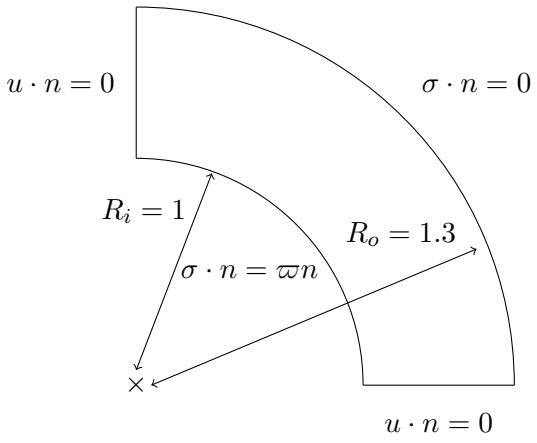


Figure 3.9 – Inner swelling of an infinite cylinder: problem setup.

Table 3.4 reports the L^2-error on the displacement (evaluated as in Section 3.3.5) and on the cumulated plastic strain. The reference solution is computed on the finest mesh and is based on P^2-Lagrange finite elements (FE) using the implementation available in [Bleyer, 2018]. The results in Table 3.4 show that the method converges at order two in the L^2-norm and at order one in the energy norm.

h	dofs	$\|u - \mathfrak{R}(u_h)\|_{L^2(\Omega)}$	order	$\|u - u_h\|_{\text{en}}$	order
0.07735	992	5.15e-4	-	3.05e-03	-
0.04217	3412	1.69e-4	1.80	8.38e-04	2.09
0.02879	7588	7.26e-5	2.12	3.76e-04	2.01
0.02172	13380	3.85e-5	2.23	2.10e-04	2.05

Table 3.4 – Inner swelling of an infinite cylinder: size of the mesh, number of dofs, L^2-error and order of convergence, energy error and order of convergence.

3.5 Fully space-time discrete elasto-plasticity

In this section we consider the dynamic elasto-plasticity equations from Section 3.2. The time discretization is performed by means of an explicit, pseudo-energy conserving, time-integration scheme recently introduced in [Marazzato et al., 2019a]. The space discretization is achieved by means of the DEM scheme discussed in Section 3.3. The main difference with Section 3.4.3 is that no iterative procedure is necessary in this section. Three-dimensional test cases are presented to assess the proposed methodology.

3.5.1 Time semi-discretization of dynamic elasto-plasticity

For simplicity we consider in this section only the time semi-discretization of the dynamic elasto-plasticity equations (3.9). We deal with the fully space-time discrete setting in the next section. The time-integration scheme [Marazzato et al., 2019a] is a two-step method of order two which ensures a discrete pseudo-energy conservation, if the integration of forces is exact, even for nonlinear systems. Symmetric Gaussian quadratures of the forces can be used in practice as long as they are of order at least two. The time interval $(0, T)$ is discretized using the time nodes $0 = t_0 < \ldots < t_n < \ldots < t_N = T$, and for simplicity we consider a constant time step Δt. We define the half-time nodes $t_{n+\frac{1}{2}} = \frac{1}{2}(t_n + t_{n+1})$ for all $n = 0, \ldots, N$. The time step is limited by a CFL condition which we specify in the fully discrete setting in the next section.

The key idea in the scheme [Marazzato et al., 2019a] is to approximate the displacement field at the time nodes by means of functions u^n, for all $n = 0, \ldots, N$, with u^0 specified by the initial condition on the displacement, and the velocity field at the half-time nodes by means of functions $v^{n+\frac{1}{2}}$, for all $n = 0, \ldots, N$, with $v^{\frac{1}{2}}$ specified by the initial condition on the velocity. For all $n = 0, \ldots, N$, given u^n and $v^{n+\frac{1}{2}}$, the scheme performs two substeps: (i) A time-dependent displacement field is predicted on the time interval $[t_n, t_{n+1}]$ using the free-flight expression $\tilde{u}(t) = u^n + (t - t_n)v^{n+\frac{1}{2}}$ for all $t \in [t_n, t_{n+1}]$; (ii) The velocity field $v^{n+\frac{3}{2}}$ is predicted by means of a quadrature on the time-integration of the forces in the time interval $[t_n, t_{n+1}]$. Let $\{t_{n,k}\}_{k\in\mathcal{K}}$ and $\{\omega_{n,k}\}_{k\in\mathcal{K}}$ be the nodes and the weights for the quadrature in the time interval $[t_n, t_{n+1}]$. We then set

$$\begin{cases} u^{n,k} = u^n + (t_{n,k} - t_n)v^{n+\frac{1}{2}}, & \forall k \in \mathcal{K}, \\ (\varepsilon_p^{n,k}, p^{n,k}) = \texttt{PLAS\_EXP}(\varepsilon_p^{n,k-1}, p^{n,k-1}, \varepsilon^{n,k}), & \forall k \in \mathcal{K}, \\ \dfrac{1}{2}m(v^{n+\frac{3}{2}} - v^{n-\frac{1}{2}}, \tilde{v}) = \displaystyle\sum_{k\in\mathcal{K}} \omega_{n,k}\Big(l(t_{n,k}, \tilde{v}) - a(\varepsilon_p^{n,k}; u^{n,k}, \tilde{v})\Big), & \forall \tilde{v} \in V_0, \end{cases} \tag{3.45}$$

where $\varepsilon^{n,k} := \varepsilon(u^{n,k})$ is known from the free-flight displacement prediction and where the state for the first Gauss node $k = 1$ comes from the previous time step or the initial condition. Given a triple $(\varepsilon_p^{\text{old}}, p^{\text{old}}, \varepsilon^{\text{new}})$, the procedure `PLAS_EXP` returns a pair $(\varepsilon_p^{\text{new}}, p^{\text{new}})$ such that, letting $\sigma^{\text{new}} := \mathbb{C} : (\varepsilon^{\text{new}} - \varepsilon_p^{\text{new}})$, we have

$$\begin{cases} \varphi(\sigma^{\text{new}}, p^{\text{new}}) \le 0, \\ \lambda^{\text{new}} := \delta p := p^{\text{new}} - p^{\text{old}} \ge 0, \quad \lambda^{\text{new}} \varphi(\sigma^{\text{new}}, p^{\text{new}}) = 0, \\ \delta\varepsilon_p := \varepsilon_p^{\text{new}} - \varepsilon_p^{\text{old}} = \lambda^{\text{new}} \dfrac{\partial \varphi}{\partial \sigma}(\mathbb{C} : (\varepsilon^{\text{new}} - \varepsilon_p^{\text{old}})). \end{cases} \tag{3.46}$$

The main difference with respect to the procedure `PLAS_IMP` described in (3.40) is on the increment of the tensor of remanent plastic strain.

3.5.2 Fully space-time discrete scheme

Full space-time discretization is achieved by combining the time-integration scheme [Marazzato et al., 201 described in the previous section with the DEM space discretization scheme from Section 3.3. For all $n = 1, \dots, N$, we compute a discrete displacement field $u_h^n \in V_{hD}$ and a discrete velocity field $v_h^{n+\frac{1}{2}} \in W_{hD}(t_{n+\frac{1}{2}})$ (recall that these spaces depend on n if the prescribed Dirichlet condition on the displacement is time-dependent). Moreover, we compute a (trace-free) tensor of remanent plastic strain $\varepsilon_{p,c}^{n,k}$ and a scalar cumulated plastic deformation $p_c^{n,k}$ for every mesh cell $c \in \mathcal{C}$ and every Gauss time-node $k \in \mathcal{K}$. We set $\varepsilon_{p,\mathcal{C}}^{n,k} := (\varepsilon_{p,c}^{n,k})_{c\in\mathcal{C}}$ and $p_{\mathcal{C}}^{n,k} := (p_c^{n,k})_{c\in\mathcal{C}}$. The fully discrete scheme reads as follows: For all $n = 1, \dots, N$, given u_h^n, $v_h^{n-\frac{1}{2}}$ and $v_h^{n+\frac{1}{2}}$, compute $\{u_h^{n,k}\}_{k\in\mathcal{K}}$, $v_h^{n+\frac{3}{2}}$, $\{\varepsilon_{p,\mathcal{C}}^{n,k}\}_{k\in\mathcal{K}}$, and $\{p_{\mathcal{C}}\}_{k\in\mathcal{K}}$ such that

$$\begin{cases} u_h^{n,k} = u_h^n + (t_{n,k} - t_n) v_h^{n+\frac{1}{2}}, & \forall k \in \mathcal{K}, \\ (\varepsilon_{p,c}^{n,k}, p_c^{n,k}) = \texttt{PLAS\_EXP}(\varepsilon_{p,c}^{n,k-1}, p_c^{n,k-1}, \varepsilon_c^{n,k}), & \forall k \in \mathcal{K},\ \forall c \in \mathcal{C}, \\ \dfrac{1}{2} m_h(v_h^{n+\frac{3}{2}} - v_h^{n-\frac{1}{2}}, \tilde{v}_h) = \displaystyle\sum_{k\in\mathcal{K}} \omega_{n,k} \Big(l_h(t_{n,k}, \tilde{v}_h) - a_h(\varepsilon_{p,\mathcal{C}}^{n,k}; u_h^{n,k}, \tilde{v}_h) \Big), & \forall \tilde{v}_h \in V_{h0}, \end{cases} \tag{3.47}$$

where $\varepsilon_c^{n,k} := \varepsilon_c(\mathcal{R}(u_h^{n,k}))$. Moreover, m_h and l_h are, respectively, the discrete mass bilinear form and the discrete load linear form. For the first Gauss node $k = 1$, the first two arguments in `PLAS_EXP` come from the previous time step or the initial condition. The initial displacement u_h^0 and the initial velocity $v_h^{\frac{1}{2}}$ are evaluated by using the values of the prescribed initial displacement u_0 and the prescribed initial velocity v_0 at the cell barycentres and the boundary vertices.

The time step is restricted by the following CFL stability condition:

$$\Delta t < 2\sqrt{\frac{\mu_{\min}}{\lambda_{\max}}}, \tag{3.48}$$

where $\mu_{\min}$ is the smallest entry of the diagonal mass matrix associated with the discrete mass bilinear form $m_h(\cdot,\cdot)$ and $\lambda_{\max}$ is the largest eigenvalue of the stiffness matrix associated the discrete stiffness bilinear form $a_h(0;\cdot,\cdot)$ (i.e., this maximal eigenvalue is computed in the worst-case scenario when there is no plasticity). The CFL condition (3.48) guarantees the stability of the time-integration scheme in the linear case [Marazzato et al., 2019a], i.e., when there is no plasticity. We expect that this condition is still reasonable in the nonlinear case since plasticity does not increase the stiffness of the material.

Finally, let us write the discrete equivalent of the energy conservation property (3.11). Define the discrete elastic energy at t_{n+1} as

$$E_{\text{elas},h}^{n+1} := \frac{1}{2} \sum_{c\in\mathcal{C}} |c| \sigma_c^{n+1} : \mathbb{C}^{-1} : \sigma_c^{n+1},$$

with $\sigma_c^{n+1} := \mathbb{C} : (\varepsilon_c(\mathcal{R}(u_h^{n+1})) - \varepsilon_{p,c}^{n+1})$, for all $c \in \mathcal{C}$. The discrete kinetic energy at t_{n+1} is defined by

$$E_{\mathrm{kin},h}^{n+1} := \frac{1}{2} m_h(v_h^{n+3/2}, v_h^{n+1/2}),$$

the discrete plastic dissipation at t_{n+1} is defined by

$$E_{\mathrm{plas},h}^{n+1} := \sum_{c\in\mathcal{C}} |c| \left(\sigma_0 p_c^{n+1} + \omega_p(p_c^{n+1})\right),$$

and finally the work of external loads at t_{n+1} is defined by

$$E_{\mathrm{ext},h}^{n+1} := \sum_{m=0}^{n} l_h(t_{m+1}; v_h^{m+1/2}) \Delta t.$$

Then assuming a homogeneous Dirichlet condition for simplicity, we have the following energy balance equation:

$$E_{\mathrm{elas},h}^{n+1} + E_{\mathrm{kin},h}^{n+1} + E_{\mathrm{plas},h}^{n+1} = E_{\mathrm{ext},h}^{n+1} + E_{\mathrm{elas},h}^{0} + E_{\mathrm{kin},h}^{0} + E_{\mathrm{plas},h}^{0} + \mathcal{O}(\Delta t^2), \tag{3.49}$$

where the term $\mathcal{O}(\Delta t^2)$ results from the use of quadratures to compute the integral of forces. The interested reader is referred to [Marazzato et al., 2019a] for further details on the effect of quadratures on the conservation properties of the integration method.

3.5.3 Numerical tests

This section contains two three-dimensional test cases: a beam in dynamic flexion and a beam in dynamic torsion. We use the midpoint quadrature for the integration of the forces in each time step within the time-integration scheme. We refer the reader to [Marazzato et al., 2019a] for a study of the influence of the quadrature on the scheme accuracy for various nonlinear problems with Hamiltonian dynamics.

Although the material parameters are indicated below using physical units, they are best interpreted in terms of characteristic times. For instance, considering a one-dimensional domain of length L, the characteristic time of the numerical experiments is $T_{\mathrm{ref}} := L\sqrt{\frac{\rho}{E}}$. The CFL condition (3.48) restricts the time-step as $\Delta t < 2\frac{\mu_{\min}}{\lambda_{\max}}$. Thus, one has

$$\frac{\Delta t}{T_{\mathrm{ref}}} < \frac{2}{L}\sqrt{\frac{\mu_{\min} E}{\lambda_{\max}\rho}}.$$

Since the ratio $\frac{E}{\lambda_{\max}}$ is independent of E, the same conclusion holds for $\frac{\Delta t}{T_{\mathrm{ref}}}$. Therefore we will report this time ratio in the computations.

Beam in dynamic flexion

This test case consists in computing the oscillations of an elastic and linearly isotropic plastic beam of length $L = 1\mathrm{m}$ with a rectangular section of $0.04 \times 0.1\mathrm{m}^2$. The simulation time is $T = 2.5\mathrm{s}$. The beam is clamped at one end, it is loaded by a uniform vertical traction $g(t)$ at the other end, and the four remaining lateral faces are stress free ($\sigma \cdot n = 0$). The load term $g(t)$ is defined as

$$g(t) := \begin{cases} -\frac{5t}{4} e_x & \text{for } 0 \le t \le \frac{4}{5}, \\ 0 & \text{for } \frac{4}{5} \le t \le T. \end{cases} \tag{3.50}$$

Figure 3.10 displays the problem setup. The material parameters are $E = 10^3\mathrm{Pa}$ for the Young modulus, $\nu = 0.3$ for the Poisson ratio, $\rho = 1\mathrm{kg{\cdot}m^{-3}}$ for the density, $\sigma_0 = 25\mathrm{Pa}$ for the yield stress, and $E_t = \frac{1}{100}E$ for the tangent plastic modulus. The present three-dimensional

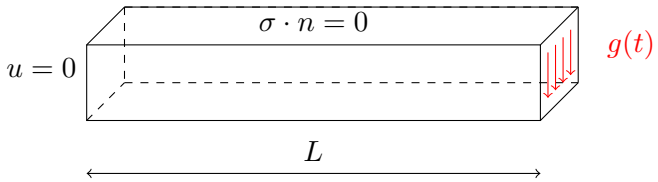


Figure 3.10 – Beam in dynamic flexion: problem setup.

implementation used as a starting point [Bleyer, 2018], where P^1-Lagrange FE and an implicit time-integration scheme are considered for a purely elastic material.

The proposed DEM is compared to penalised Crouzeix–Raviart FE [Hansbo and Larson, 2003]. This method is chosen as reference since it is known to be robust in the incompressible limit as $\nu \to 0.5$. The penalty parameter is chosen as $\eta = \beta\mu$ with $\beta = 0.5$. The reference computation is performed using $14,376$ vector-valued dofs and a time-step $\Delta t = 20\mu$s. Two computations are performed with the proposed DEM. The first uses a coarse mesh containing $4,668$ vector-valued dofs and a time-step $\Delta t = 1.4\mu$s, which is stable for the explicit integration. The second uses a fine mesh containg $13,302$ vector-valued dofs and a time-step $\Delta t = 1.1\mu$s, also stable for the explicit integration. Thus one has: $4.4 \cdot 10^{-7} \leq \frac{\Delta t}{T_{\text{ref}}} \leq 8.0 \cdot 10^{-6}$. The penalty parameter for both computations is similar to the reference computation with $\beta = 0.5$. As already mentioned, we used a midpoint quadrature of the forces. Higher-order symmetric quadratures have been found to give overlapping results with respect to the midpoint quadrature. In all the computations, the time-discretization error is negligible with respect to the space-discretization error, but larger time-steps cannot be considered owing to the CFL restriction.

The displacement and the velocity at the center of the loaded tip of the beam are compared in Figure 3.11. We notice the excellent agreement between the DEM prediction on the fine mesh and the reference computation. Figure 3.12 shows the balance of energies for the

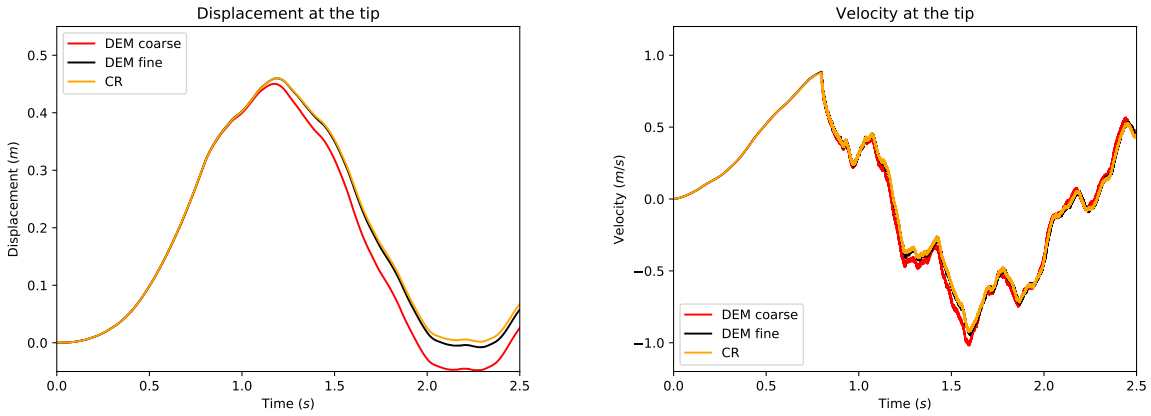


Figure 3.11 – Beam in dynamic flexion: comparison between the proposed scheme (DEM) on a coarse and a fine mesh and the reference solution (CR). Left: Displacement at the loaded tip of the beam. Right: Velocity at the same point.

reference computation and the fine DEM computation. One can first notice that the total energy for both DEM and Crouzeix–Raviart space semi-discretizations is well preserved by the time-integrator [Marazzato et al., 2019a] since the total mechanical energy (kinetic energy, elastic energy and plastic dissipation) and the work of the external load are perfectly balanced at all times. We also notice that the amount of plastic dissipation is rather significant at the end of the simulation. Figure 3.13 presents some further results of the DEM computations on the fine mesh so as to visualize at three different times during the simulation the spatial

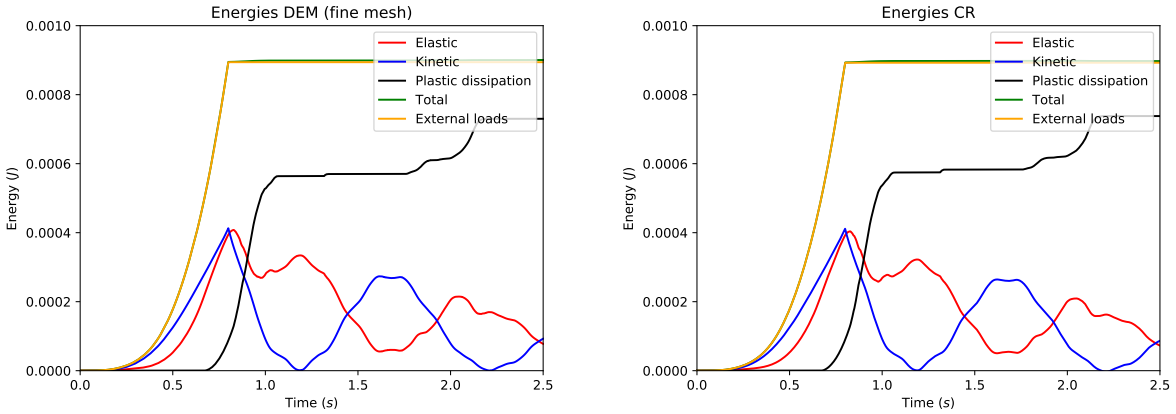


Figure 3.12 – Beam in dynamic flexion: energies during the simulation. Left: DEM (fine mesh). Right: reference solution (CR).

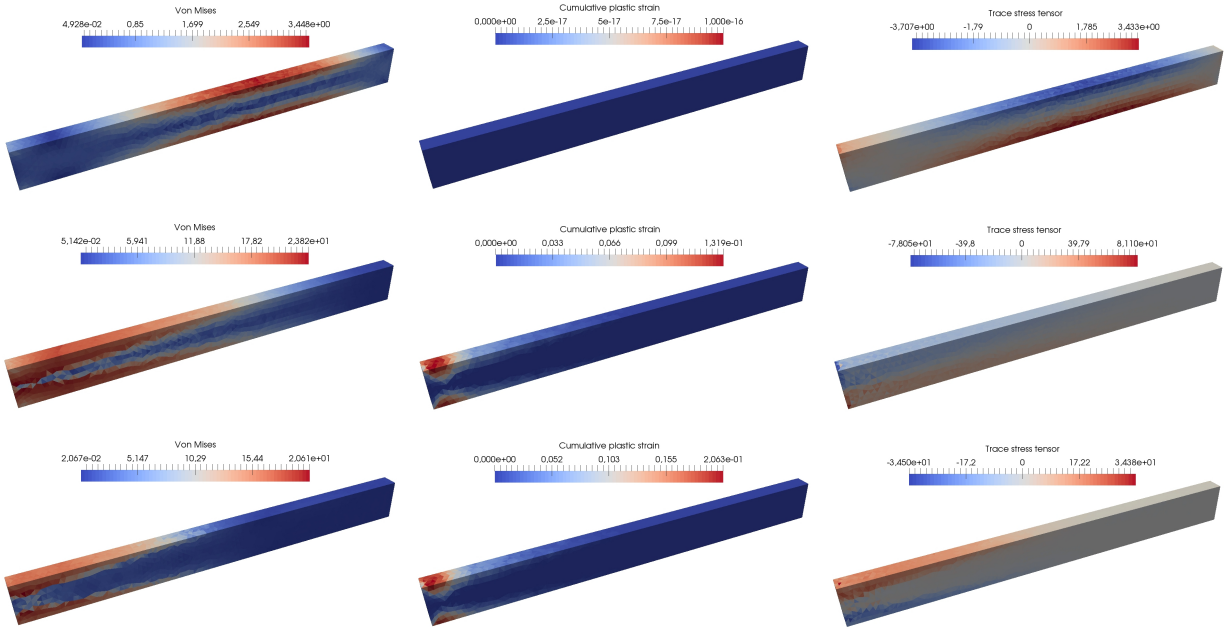


Figure 3.13 – Beam in dynamic flexion: DEM on the fine mesh. Von Mises equivalent stress (left column), p (middle column) and $\mathrm{tr}(\sigma)$ (right column) at $t = \frac{1}{10}T$ (top line), $t = \frac{1}{2}T$ (middle line) and $t = T$ (bottom line).

localization of the von Mises equivalent stress, the cumulated plastic strain, and the trace of the stress tensor. One can see that the plastic strain is concentrated close to the clamped tip of the beam, where the material undergoes the greatest stresses. The method does not exhibit any locking due to plastic incompressibility as indicated by the smooth behavior of the trace of the stress tensor.

Beam in dynamic torsion

The setting is similar to the quasi-static torsion test case presented in Section 3.4.3. The two differences are the material parameters and the plastic law which are similar to Section 3.5.3, and the boundary conditions on one side of the beam. Figure 3.14 shows the problem setup. On one of its extremities the beam is clamped, and on the other extremity the following normal stress is imposed:

$$g(t) = \mu\alpha(t)\frac{r}{L}e_\theta, \tag{3.51}$$

where r and e_θ are defined in Section 3.4.3. The angle $\alpha(t)$ is increased from 0 at $t = 0$ to $5\alpha_y$ at $t = T = 0.5$s, where α_y is the yield angle defined in Section 3.4.3. The plastic parameters are the same as those in Section 3.5.3.

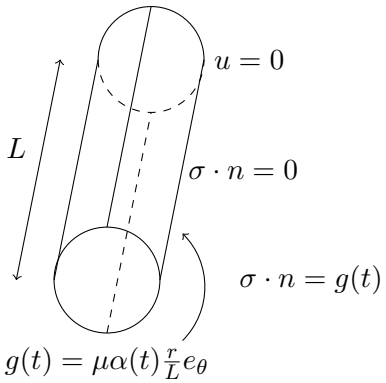


Figure 3.14 – Beam in dynamic torsion: problem setup.

Three different space discretizations are considered for this test case: P^1-Lagrange FE, penalised Crouzeix–Raviart FE and the proposed DEM. The Crouzeix–Raviart computations are used as reference since the method is known to be robust with respect to the incompressible limit. The Lagrange FE computations are used to illustrate their inability to deal with large plastic (deviatoric) strains. The goal of this test case is to show the ability of the proposed DEM to deal with deviatoric plasticity. The computations are not performed on the same meshes but rather with meshes leading to a similar number of dofs so as to give comparable results for DEM and Lagrange FE, whereas the meshes used with penalised Crouzeix–Raviart FE lead to twice as many dofs since they are employed to obtain a reference solution. The number of dofs and the time-steps used in the computations are presented in Table 3.5. One thus has: $8.4 \cdot 10^{-7} \leq$

Method	DEM		Lagrange FE		penalised CR FE	
Computation	coarse	fine	coarse	fine	fine	very fine
Vectorial dofs	6,978	14,438	6,584	12,853	13,052	27,711
Δt (μs)	4.1	1.3	3.9	2.6	2.3	0.42

Table 3.5 – Beam in dynamic torsion: number of vectorial dofs and time-step for all the computations.

$\frac{\Delta t}{T_{\text{ref}}} \leq 8.2 \cdot 10^{-6}$. All the reported time-steps are compatible with the CFL restriction. Also, for

all computations, we have verified that the time discretization error is negligible with respect to the space discretization error. The time-integration scheme [Marazzato et al., 2019a] is used with a midpoint quadrature of the forces for the coarse computations and with a Gauss-Legendre quadrature of the forces of order 5 for the fine computations. For details on the effect of the quadrature rule see [Marazzato et al., 2019a].

The comparison between the methods is performed by considering the displacement and the velocity of the point of coordinates $(0.9R, 0, \frac{1}{6}L)$ over the simulation time T. The results are reported in Figure 3.15. In the angular velocity plot, we can see for times $t \leq 0.2$s elastic waves

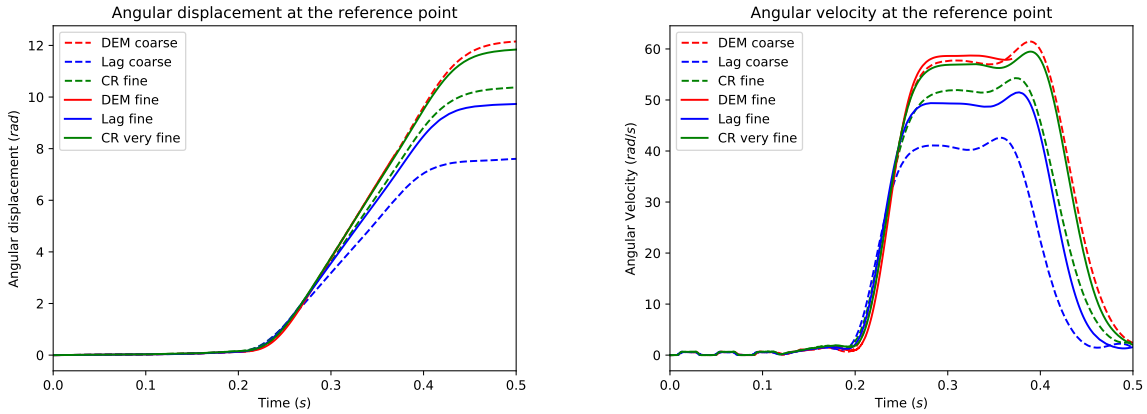


Figure 3.15 – Beam in dynamic torsion: comparison between DEM and FEM. Left: Displacement at the chosen point. Right: Velocity at the same point.

with a small magnitude travelling through the beam. The wave of larger amplitude arriving afterwards is a plastic wave whose velocity is ten times smaller than the elastic waves since $E_t = \frac{E}{100}$. We notice that the value for the simulation time is too long for the simulation to remain physically relevant within the small strain assumption owing to the large value reached by the angular displacement of the reference point. However this setting allows us to reach substantial amounts of plastic dissipation and thereby to probe the robustness of the space semi-discretization methods with respect to incompressibility. Recall that the remanent plastic strain tensor is trace-free, so that the stress tensor is nearly deviatoric in the entire beam at the end of the simulation. Such a situation is challenging for the P^1-Lagrange FEM since this method is known to lock in the incompressible limit. To highlight the volumetric locking incurred by Lagrange FE, Figure 3.16 displays at the time $t = \frac{1}{2}T$ the trace of the stress tensor predicted by Lagrange FE (fine mesh) and penalised Crouzeix–Raviart FE (coarse mesh), for a similar number of dofs. For Lagrange FE, significant oscillations are visible in the whole beam (the amplitude of these oscillations is about ten times the maximal value of the von Mises equivalent stress). Also, the amplitudes of the oscilliations of the trace tensor are about four times larger than for penalised Crouzeix–Raviart FE. Figure 3.17 reports the energies on the fine meshes for the DEM, Lagrange FE and penalised Crouzeix–Raviart FE. First, we notice as in the previous test case the prefect balance of the work of external loads with the different components of the mechanical energy. The orders of magnitude of the energies and plastic dissipations are similar for the three methods. However, a significant discrepancy in the plastic dissipation can be observed for Lagrange FE with respect to penalised Crouzeix–Raviart FE and DEM which both give a plastic dissipation similar to the reference computation.

Figure 3.18 presents some further results of the DEM computations on the fine mesh so as to visualize at three different times during the simulation the spatial localization of the von Mises equivalent stress, the cumulated plastic strain, and the trace of the stress tensor. We can see in the first row that the von Mises stress is nonzero in the entire beam and thus the elastic waves have already travelled through the entire beam whereas the cumulative plastic strain is

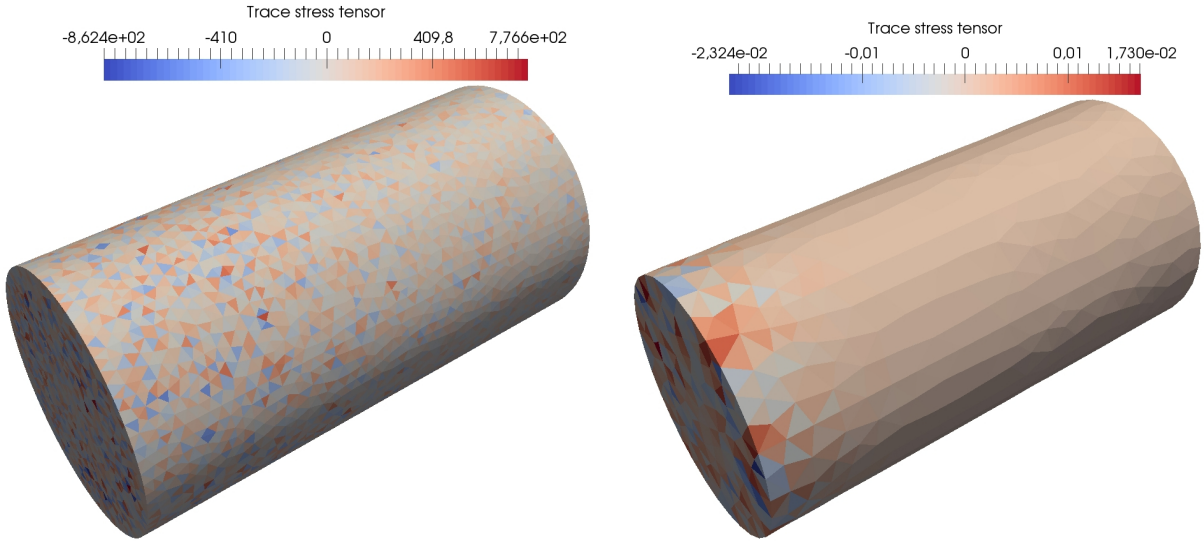


Figure 3.16 – Beam in dynamic torsion: $\mathrm{tr}(\sigma)$ at $t = \frac{1}{2}T$. Left: Lagrange FE (fine mesh). Right: Penalised Crouzeix–Raviart FE (fine mesh).

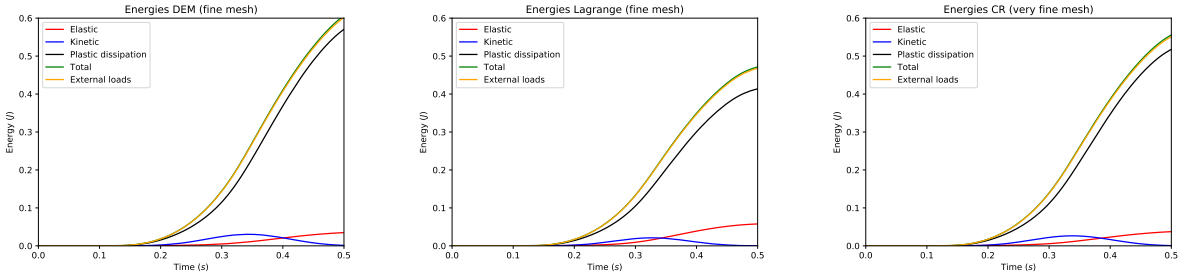


Figure 3.17 – Beam in dynamic torsion: energies during the simulation on the fine meshes. Left: DEM. Middle: Lagrange FE. Right: penalised Crouzeix–Raviart FE.

still zero and thus no plastic flow has occurred. In the second row, we can see that a plastic wave has started to propagate from one end of the beam. In the last row, we see that the plastic wave has reached the other side of the beam at the end of the simulation. Also, regarding the robustness with respect to $\nu \to 0.5$, the magnitude of the oscillations of the trace of the stress tensor is similar to penalised Crouzeix–Raviart FE. Indeed for the DEM, the extreme values of the trace of the stress tensor are $-3 \cdot 10^2$Pa and $3 \cdot 10^2$Pa at $t = \frac{T}{2}$ and $-9 \cdot 10^2$Pa and $7 \cdot 10^2$Pa at $t = T$ whereas for penalised Crouzeix–Raviart FE the extermes values are $-2 \cdot 10^2$Pa and $2 \cdot 10^2$Pa at $t = \frac{T}{2}$ and $-6 \cdot 10^2$Pa and $6 \cdot 10^2$Pa at $t = T$. Comparatively, for the P^1-Lagrange FE computations, the extreme values of the trace of the stress tensor are $-8 \cdot 10^2$Pa and $9 \cdot 10^2$Pa for $t = \frac{T}{2}$ and $-2 \cdot 10^3$Pa and $2 \cdot 10^3$Pa for $t = T$.

3.6 Conclusion

We have presented a new Discrete Element Method which is a consistent discretization of a Cauchy continuum and which only requires continuum macroscopic parameters as the Young modulus and the Poisson ratio for its implementation. The displacement degrees of freedom are attached to the barycentres of the mesh cells and to the boundary vertices. The key idea is to reconstruct displacements on the mesh facets and then to use a discrete Stokes formula to devise a piecewise constant gradient and linearized strain reconstructions. A simple geometric pre-processing has been devised to ensure that for almost all the mesh facets, the

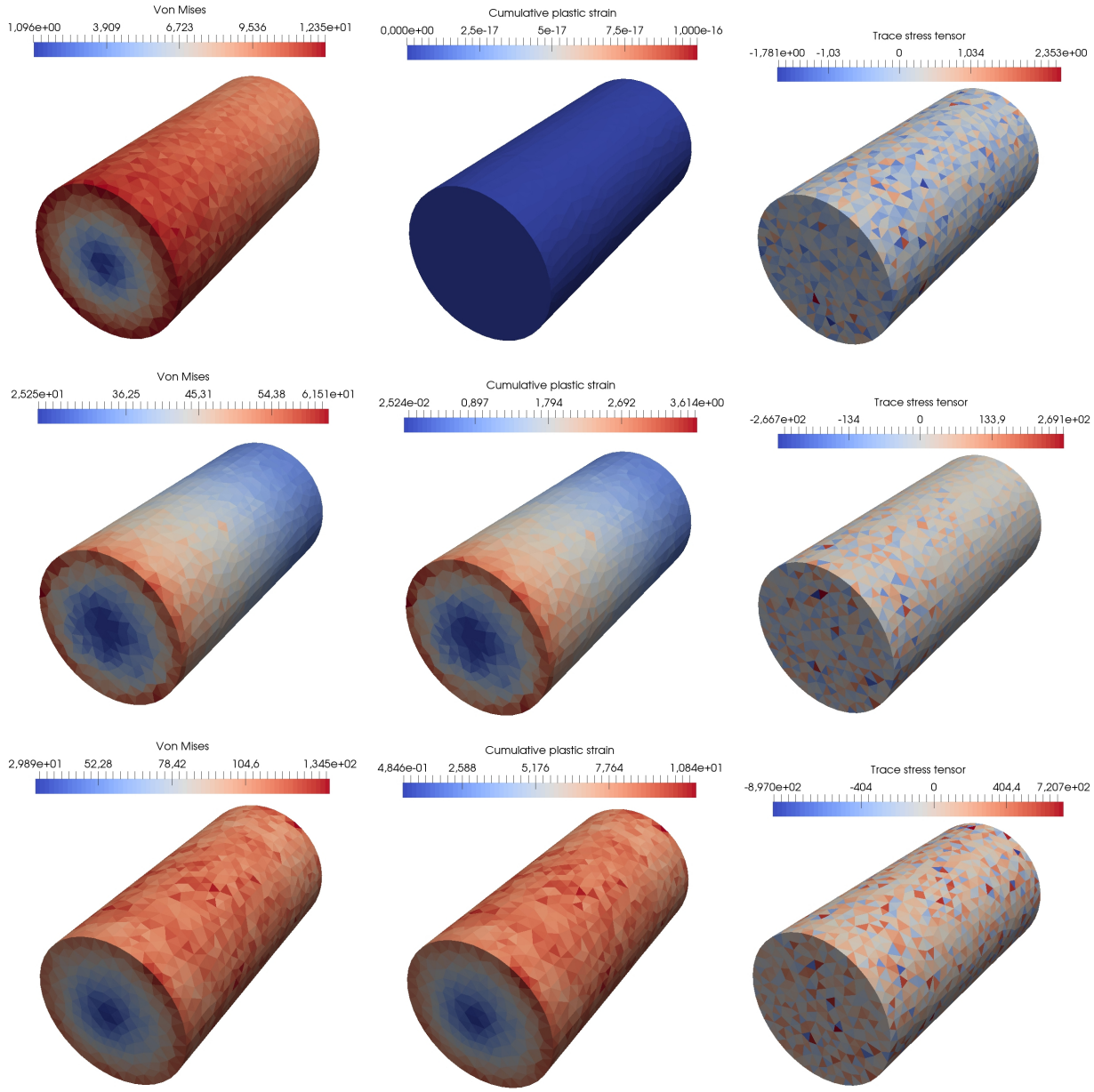


Figure 3.18 – Beam in dynamic torsion: DEM on the fine mesh. Von Mises equivalent stress (left column), p (middle column) and $\mathrm{tr}(\sigma)$ (right column) at $t = \frac{1}{10}T$ (top line), $t = \frac{1}{2}T$ (middle line) and $t = T$ (bottom line).

reconstruction is based on an interpolation (rather than extrapolation) formula and we have shown by numerical experiments that this choice can produce highly beneficial effects in terms of the largest eigenvalue of the stiffness matrix, and thus on the time step restriction within explicit time-marching schemes. Moreover, in the case of elasto-plastic behavior, the internal variables for plasticity are piecewise-constant in the mesh cells. The scheme has been tested on quasi-static and dynamic test cases using a second-order, explicit, energy-conserving time-marching scheme. Future work can include adapting the present framework to dynamic cracking and fragmentation as well as to Cosserat continua. An extension to large strain solid dynamics could also be performed by working in the reference configuration.

Chapter 4

Crack propagation with a Griffith criterion using a discrete element method

Contents

4.1 Introduction

Discrete element methods (DEM) are popular in the modelling of granular materials, soil and rock mechanics. DEM generally use sphere packing to discretize the domain as small spheres interacting through forces and torques [Labra and Oñate, 2009], but the main difficulty is to derive a suitable set of parameter values for those interactions so as to reproduce a given Young modulus E and Poisson ratio ν at the macroscopic level [Jebahi et al., 2015, Celigueta et al., 2017]. Advantages of DEM are their ability to deal with discontinuous materials, such as fractured or porous materials, as well as the possibility to take advantage of GPU computations [Spellings et al., 2017]. A first DEM parametrized only by E and ν has been proposed in [Monasse and Mariotti, 2012] for elasto-dynamic computations on Voronoi meshes.

In a consecutive work [Marazzato et al., 2019b], a variational DEM has been proposed for dynamic elasto-plasticity computations on polyhedral meshes using cell-wise reconstructions of the strains.

DEM for cracking have been developed in [André et al., 2013] and [André et al., 2019] with cracks propagating through the facets of the (Voronoi) mesh and using a critical stress criterion (initiation criterion). A similar approach, but using a different reconstruction of strains based on moving least-squares interpolations, can be traced back to [Belytschko et al., 1994] (2d), [Sukumar et al., 1997] (3d) and [Belytschko et al., 1995] (dynamic cracking). Crack propagation instead can be based on the Griffith criterion which relies on the computation of the stress intensity factors (SIF) at the crack tip. Virtual element methods (VEM) have been recently applied to crack propagation [Hussein et al., 2018]. Cracks were allowed to cut through the polyhedral mesh cells as in the extended finite element method (XFEM) which is based on an extended space of basis functions [Chahine et al., 2008] and a level-set description of the crack [Moës and Belytschko, 2002]. Phase-field methods instead smooth the crack and have been developed among others in [Bourdin et al., 2000] (quasi-static evolution) and [Li and Marigo, 2017] and [Bleyer et al., 2017] (dynamic evolution). Phase-field methods are not based on SIF computations but rather on a variational formulation of cracking [Francfort and Marigo, 1998]. Furthermore DEM using cohesive laws have been developed for fragmentation computations [Mariotti et al., 2009] with a view towards uniting initiation and propagation. These methods allow one to devise an initiation criterion and also to control the energy dissipation as with a Griffith criterion. The cracks still go through the mesh facets. This is also the case for similar methods of higher-order such as discontinuous Galerkin methods [Hansbo and Salomonsson, 2015]. Regarding time-integration, the above methods typically use an explicit time-integration scheme (see also [Doyen et al., 2013] on a quasi-explicit time-integration scheme for cohesive models).

The main goal of the present work is to develop a variational DEM on polyhedral meshes using a Griffith criterion to compute dynamic crack propagation through the mesh facets. The proposed method is close to [Marazzato et al., 2019b] (where there is no cracking) but the degrees of freedom (dofs) are different since we use here displacement dofs attached to the boundary and crack facets rather than dofs attached to the boundary vertices, in addition to the usual dofs attached to the mesh cells. The reason for replacing the dofs at boundary vertices by dofs at the barycenter of the crack facets is to avoid the singularity at the crack tip. The cracking algorithm hinges on two main ingredients. The first ingredient is an approximation of the SIF at the crack tip. To this purpose the SIF are evaluated using a reconstruction of the stresses on facets ahead of the crack tip. The energetic propagation criterion is based on the Irwin formula and the continuous energy release rate which is a material parameter. The second ingredient is a kinking criterion used to determine the next breaking facet and thus the crack path. The kinking criterion, in the spirit of [Sih, 1974], consists in selecting for the crack path the inner facet of the mesh that maximizes a quantity representing the local density of elastic energy.

The present chapter is organized as follows. Section 4.2 briefly recalls the equations of elasto-dynamics and cracking in a Cauchy continuum. Section 4.3 introduces the proposed DEM and presents the space semi-discretization of the governing equations. Moreover some numerical tests are reported to assess the convergence of the space discretization in the presence of a singularity. Section 4.4 and Section 4.5 address the full discretization of the quasi-static and dynamic cracking problems, respectively. Section 4.6 contains numerical results on quasi-static and dynamic crack propagation problems in two space dimensions. Finally, Section 4.7 draws some conclusions.

4.2 Governing equations for dynamic cracking

We consider an elastic fragile material occupying the domain $\Omega \subset \mathbb{R}^2$ in the reference configuration and evolving dynamically over the finite time interval $[0,T]$, $T>0$, under the action of a volumetric force f and boundary conditions. The strain regime is restricted to small strains so that we use the linearized strain tensor $\varepsilon(u) := \frac{1}{2}(\nabla u + (\nabla u)^{\mathsf{T}}) \in \mathrm{R}^{3\times 3}$, where u is the $\mathbb{R}^3$–valued displacement field. The material is supposed to be homogeneous and isotropic. The stress tensor $\sigma(u) \in \mathbb{R}^{3\times 3}$ is such that

$$\sigma(u) := \mathbb{C} : \varepsilon(u), \tag{4.1}$$

where $\mathbb{C}$ is the fourth-order stiffness tensor. The elastic material is characterized by the Young modulus E and the Poisson ratio ν or equivalently by the Lamé coefficients λ and μ. The boundary of Ω is partitioned as $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$, a Dirichlet condition is prescribed on $\partial\Omega_D$, and a Neumann condition on $\partial\Omega_N$, so that we enforce for all $t \in [0,T]$,

$$u = u_D \ \text{ on } \partial\Omega_D, \qquad \sigma(u)\cdot n = g_N \ \text{ on } \partial\Omega_N. \tag{4.2}$$

Since cracking can occur, we denote $\Gamma(t)$ the crack at any time $t \in [0,T]$ and the actual domain at time t is

$$\Omega(t) := \Omega \setminus \Gamma(t). \tag{4.3}$$

This implies that $\partial\Omega(t) = \partial\Omega_D \cup \partial\Omega_N \cup \Gamma(t)$. We enforce a homogeneous Neumann condition on $\Gamma(t)$ for all $t \in [0,T]$, i.e.,

$$\sigma(u)\cdot n = 0 \ \text{ on } \Gamma(t). \tag{4.4}$$

Since we are interested in crack propagation, we assume that $\Omega(0)$ already contains a crack, i.e., $\Gamma(0) \neq \emptyset$. The crack $\Gamma(t)$ is supposed to be a countably rectifiable 1–manifold for all $t \in [0,T]$ (see [Dal Maso, 2013]). This hypothesis ensures the almost everywhere (a.e.) existence of a normal vector n and a tangent vector τ to $\Gamma(t)$ at any point $y \in \Gamma(t)$ [Simon, 1983]. Figure 4.1 illustrates these quantities.

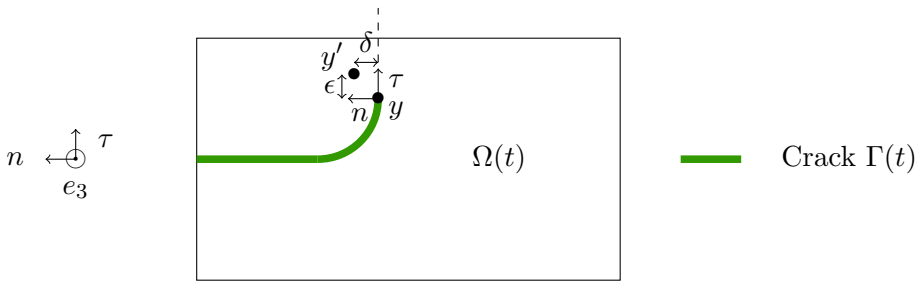


Figure 4.1 – Sketch of a crack in the two-dimensional domain $\Omega(t)$.

The stress intensity factors (SIF) at any point $y \in \Gamma(t)$ are usually defined for a purely elastic material as

$$\begin{cases} K_1(y) := \lim\limits_{\delta\downarrow 0,\epsilon\to 0} \sigma_{nn}(y')\sqrt{2\pi d(y,y')}, \\ K_2(y) := \lim\limits_{\delta\downarrow 0,\epsilon\to 0} \sigma_{n\tau}(y')\sqrt{2\pi d(y,y')}, \\ K_3(y) := \lim\limits_{\delta\downarrow 0,\epsilon\to 0} \sigma_{n3}(y')\sqrt{2\pi d(y,y')}, \end{cases} \tag{4.5}$$

where $d(\cdot,\cdot)$ is the Euclidean distance in $\mathbb{R}^3$ and $y' := y + \delta n + \epsilon\tau$ (see Figure 4.1). If the stresses remain bounded in the vicinity of $y \in \Gamma(t)$, then the SIF are null. Using the Irwin formula, one can define the energy release rate $\mathcal{G}(y)$ as

$$\mathcal{G}(y) := \frac{1-\nu^2}{E}\left(K_1(y)^2 + K_2(y)^2\right) + \frac{1+\nu}{E}K_3(y)^2. \tag{4.6}$$

Admissible states are characterized by the inequality

$$\mathcal{G}(y) \leq \mathcal{G}_c, \quad \forall y \in \Gamma(t), \tag{4.7}$$

where $\mathcal{G}_c$ is a material property associated with the capacity of the material to sustain loads without locally failing and thus opening cracks. The material remains healthy at the point $y \in \Gamma(t)$ if $\mathcal{G}(y) < \mathcal{G}_c$ and breaks if $\mathcal{G}(y) = \mathcal{G}_c$. The material parameter $\mathcal{G}_c$ is assumed to be homogeneous for simplicity.

To formulate the governing equations for dynamic cracking, we consider time-dependent functions with values in space-dependent functional spaces. Let us set

$$V_D(t) := \left\{ v \in H^1(\Omega(t); \mathbb{R}^d) \mid v_{|\partial\Omega_D} = u_D \right\}, \qquad V_0(t) := \left\{ v \in H^1(\Omega(t); \mathbb{R}^d) \mid v_{|\partial\Omega_D} = 0 \right\}, \tag{4.8}$$

where standard notation is used for the Hilbert Sobolev spaces. The weak solution is searched as a pair (u, Γ) such that for all $t \in [0, T]$, $u(t) \in V_D(t)$, $\Gamma(t) \subset \Omega$ is a 1–manifold satisfying the above assumptions, and

$$\begin{cases} m(t; \ddot{u}(t), \tilde{v}) + a(t; u(t), \tilde{v}) = l(t; \tilde{v}), & \forall \tilde{v} \in V_0(t), \\ \mathcal{G}(y) \leq \mathcal{G}_c, & \forall y \in \Gamma(t). \end{cases} \tag{4.9}$$

Here we introduced the mass and stiffness bilinear forms such that for all $(v, \tilde{v}) \in V_D(t) \times V_0(t)$,

$$m(t; v, \tilde{v}) := \int_{\Omega(t)} \rho v \cdot \tilde{v}, \tag{4.10}$$

where $\rho > 0$ is the density of the material, dots indicate time derivatives, and

$$a(t; v, \tilde{v}) := \int_{\Omega(t)} \varepsilon(v) : \mathbb{C} : \varepsilon(\tilde{v}), \tag{4.11}$$

and the linear form acting on $V_0(t)$ as follows:

$$l(t; \tilde{v}) := \int_{\Omega(t)} f(t) \cdot \tilde{v} + \int_{\partial\Omega_N} g_N(t) \cdot \tilde{v}. \tag{4.12}$$

Note that the Dirichlet condition on $\partial\Omega_D$ is enforced strongly, whereas the Neumann condition on $\partial\Omega_N \cup \Gamma(t)$ is enforced weakly. Finally the initial conditions for (4.9) prescribe $u(0) = u_0$, $\dot{u}(0) = v_0$, and an initial crack $\Gamma(0)$ in Ω.

4.3 Space semi-discretization

In this section we present the space semi-discretization of (4.9) using a DEM.

4.3.1 Discrete sets and degrees of freedom

The domain Ω is discretized with a mesh $\mathcal{T}_h$ of size h made of polyhedra with planar facets in three space dimensions or polygons with straight edges in two space dimensions. We assume that Ω is itself a polyhedron or a polygon so that the mesh covers Ω exactly. We also assume that the mesh is compatible with the initial crack position $\Gamma(0)$ and with the partition of the boundary into the Dirichlet and Neumann parts.

Let $\mathcal{C}$ denote the set composed of the mesh cells and, for all $t \in [0, T]$, let $\mathcal{F}(t)$ denote the set composed of the mesh facets. This set is time-dependent since a facet $F \in \mathcal{F}(t)$ is replaced, after cracking, by two boundary facets $F_-, F_+ \in \mathcal{F}(t)$ (F_-, F_+ are the same geometric object, but are different objects regarding the data structure since each one belongs to the boundary of a different mesh cell). We partition the set of mesh facets as $\mathcal{F}(t) = \mathcal{F}^i(t) \cup \mathcal{F}^b(t)$, where $\mathcal{F}^i(t)$ is composed of the internal facets shared by two mesh cells and $\mathcal{F}^b(t)$ is the collection of

the boundary facets sitting on the boundary $\partial\Omega(t) = \partial\Omega_D \cup \partial\Omega_N \cup \Gamma(t)$ (notice that such facets belong to the boundary of only one mesh cell). The subsets $\mathcal{F}^i(t)$ and $\mathcal{F}^b(t)$ are time-dependent since, as the facet $F \in \mathcal{F}^i(t)$ cracks, it is replaced by the facets $F_+, F_- \in \mathcal{F}^b(t)$. At any time $t \in [0, T]$, the discrete crack is denoted $\Gamma_h(t)$ and is composed of facets belonging to a subset of $\mathcal{F}^b(t)$. This subset is denoted $\mathcal{F}^\Gamma(t) \subset \mathcal{F}^b(t)$.

Vector-valued volumetric degrees of freedom (dofs) for a generic displacement field $v_\mathcal{C} := (v_c)_{c\in\mathcal{C}} \in \mathbb{R}^{d\#(\mathcal{C})}$ are placed at the barycentre of every mesh cell $c \in \mathcal{C}$. Additional displacement boundary dofs $v_{\mathcal{F}^b} := (v_F)_{F\in\mathcal{F}^b(t)} \in \mathbb{R}^{d\#(\mathcal{F}^b(t))}$ are added at the barycentre of every boundary facet $F \in \mathcal{F}^b(t)$. Note that the number of such boundary dofs is by construction time-dependent since it depends on the number of boundary facets and thus on the crack propagation. We use the compact notation $v_h := (v_\mathcal{C}, v_{\mathcal{F}^b})$ for the collection of all the cell dofs and all the boundary facet dofs. Figure 4.2 illustrates the position of the displacement dofs. The barycentre of a mesh cell $c \in \mathcal{C}$ is denoted by $\mathbf{x}_c$ and the barycentre of a mesh facet $F \in \mathcal{F}$ is denoted by $\mathbf{x}_F$.

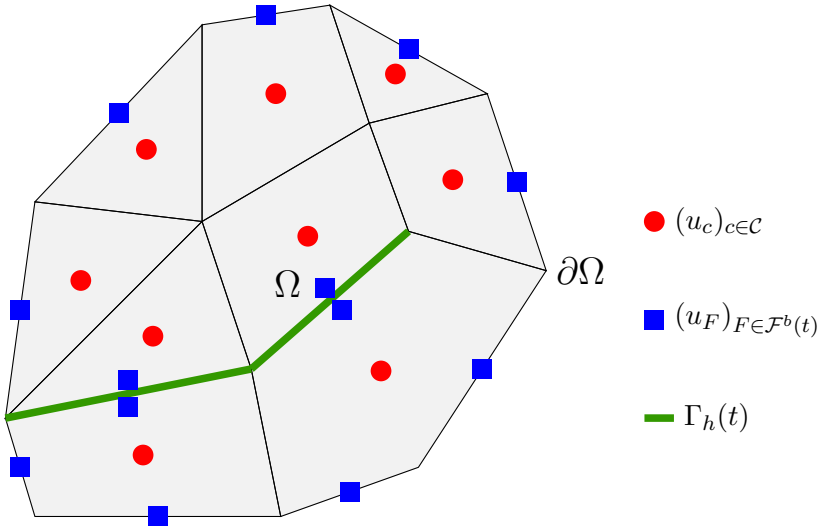


Figure 4.2 – Domain Ω covered by a polyhedral mesh and vector-valued degrees of freedom for the displacement.

4.3.2 Discrete bilinear and linear forms

The discrete stiffness bilinear form hinges on a reconstruction operator that provides a displacement value at every mesh facet by an interpolation formula (see below) from neighbouring cell (and possibly boundary facet) dofs. In other words, using the cell and boundary facet dofs, we reconstruct a collection of displacements $v_\mathcal{F} := (v_F)_{F\in\mathcal{F}(t)} \in \mathbb{R}^{d\#(\mathcal{F}(t))}$ on all the mesh facets. The reconstruction operator is denoted $\mathcal{R}(t;\cdot)$ and we write

$$v_\mathcal{F} := \mathcal{R}(t; v_h) \in \mathbb{R}^{d\#(\mathcal{F}(t))}. \tag{4.13}$$

The reconstruction operator depends on t because of the connectivity modifications due to the crack propagation.

On all the boundary facets, the reconstruction operator returns the actual value, so that we only need to specify the reconstruction on the inner facets. Let $F \in \mathcal{F}^i(t)$. The main idea to define v_F is to use a barycentric combination of the cell (and possibly boundary facet) dofs close to F. A similar idea has been considered for finite volume methods in [Eymard et al., 2009, Sec. 2.2] and for cell-centered Galerkin methods in [Di Pietro, 2012]. We thus select a subset of neighboring dofs of F, say $\mathcal{I}_F \subset \mathcal{C} \cup \mathcal{F}^b(t)$, and set

$$v_F := \sum_{i\in\mathcal{I}_F} \alpha_i(\mathbf{x}_F) v_i, \tag{4.14}$$

where the v_i's are the dofs of v_h and the coefficients $\alpha_i(\mathbf{x}_F)$ are the barycentric coordinates of the facet barycenter $\mathbf{x}_F$ in terms of the selected positions of the dofs. For this construction to be meaningful, the points associated with the selected dofs must not lie on the same hyperplane, so that, in particular, the cardinality of $\mathcal{I}_F$ is at least $(d+1)$.

Here, we use a reconstruction in which the two cells sharing the inner facet F play symmetric roles. We refer to this construction as symmetric reconstruction. Specifically, let c_+ and c_- be the two cells sharing the inner facet $F \in \mathcal{F}^i(t)$. Then, we select $\mathcal{I}_-$ (resp. $\mathcal{I}_+$) as being composed of the cell c_+ (resp. c_-) and of all the other cells sharing an inner facet with c_- (resp. c_+), and we set

$$v_F := \frac{1}{2} \sum_{i \in \mathcal{I}_- \cup \mathcal{I}_+} \alpha_i(\mathbf{x}_F) v_i, \tag{4.15}$$

so that, in the case of a simplicial mesh, $2(d+1)$ dofs are used for the reconstruction (always including c_- and c_+). Note that $\sum_{i\in\mathcal{I}_-} \alpha_i(\mathbf{x}_F) = \sum_{i\in\mathcal{I}_+} \alpha_i(\mathbf{x}_F) = 1$ here. For non-simplicial meshes, the barycentric coordinates $\alpha_i(\mathbf{x}_F)$ are generalized barycentric coordiantes and thus are not uniquely defined. They can be computed using, for instance, [Budninskiy et al., 2016].

Figure 4.3 presents an example where $c_- = c_i$, $c_+ = c_j$, $\mathcal{I}_- = \{j,k,l\}$ and $\mathcal{I}_+ = \{i,m,n\}$. If the facet F touches the boundary, one or some cell dofs are replaced by the corresponding facet dof (see the right panel of Figure 4.10 for an illustration).

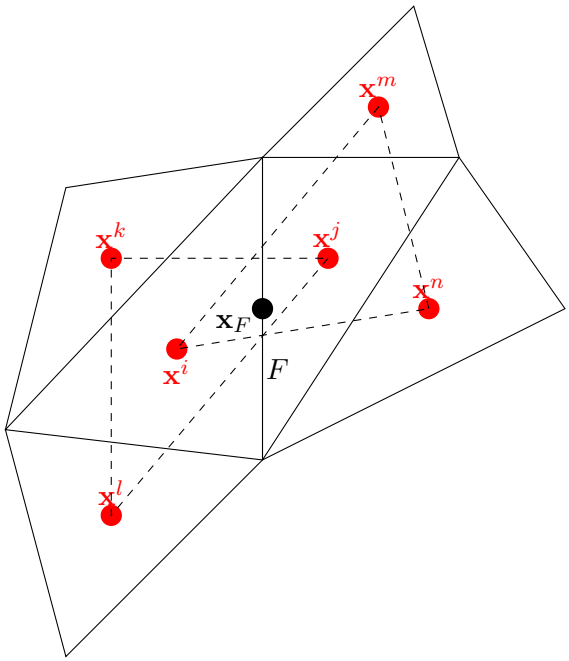


Figure 4.3 – Dofs associated with the interior facet F used in the reconstruction.

Having defined the reconstructed facet displacements, it is now possible to devise a discrete $\mathbb{R}^{d\times d}$-valued piecewise-constant gradient field for the displacement that we write $G_{\mathcal{C}}(v_{\mathcal{F}}) := (G_c(v_{\mathcal{F}}))_{c\in\mathcal{C}} \in \mathbb{R}^{d^2\#(\mathcal{C})}$. Specifically we set in every mesh cell $c \in \mathcal{C}$,

$$G_c(v_{\mathcal{F}}) := \sum_{F\in\partial c} \frac{|F|}{|c|} v_F \otimes n_{F,c}, \tag{4.16}$$

where the summation is over the facets F of c and $n_{F,c}$ is the outward normal to c on F. Note that (4.16) is motivated by a Stokes formula and that for all $v_h \in V_h$, we have

$$G_c(\mathcal{R}(t;v_h)) = \sum_{F\in\partial c} \frac{|F|}{|c|} (\mathcal{R}(t;v_h)_F - v_c) \otimes n_{F,c}, \tag{4.17}$$

since $\sum_{F\in\partial c}|F|n_{F,c}=0$. We define a constant linearized strain tensor in every mesh cell $c\in\mathcal{C}$ such that

$$\varepsilon_c(v_{\mathcal{F}}) := \frac{1}{2}(G_c(v_{\mathcal{F}}) + G_c(v_{\mathcal{F}})^{\mathsf{T}}) \in \mathbb{R}^{d\times d}, \tag{4.18}$$

and a constant stress tensor in every mesh cell $c\in\mathcal{C}$ such that

$$\Sigma_c(v_{\mathcal{F}}) := \mathbb{C} : \varepsilon_c(v_{\mathcal{F}}) \in \mathbb{R}^{d\times d}. \tag{4.19}$$

Finally, we define an additional reconstruction. It is a cellwise nonconforming P^1 reconstruction $\mathfrak{R}$ defined for all $c\in\mathcal{C}$ by

$$\mathfrak{R}(t;v_h)_c(\mathbf{x}) := v_c + G_c(\mathcal{R}(t;v_h))\cdot(\mathbf{x}-\mathbf{x}_c), \tag{4.20}$$

where $\mathbf{x}\in c$ and $\mathbf{x}_c$ if the barycentre of the cell c.

4.3.3 Discrete problem

Let us set $V_h(t) := \mathbb{R}^{d\#(\mathcal{C})}\times\mathbb{R}^{d\#(\mathcal{F}^b(t))}$ for all $t\in[0,T]$ and

$$\begin{cases} V_{hD}(t) := \{v_h\in V_h(t)\mid v_F = u_D(\mathbf{x}_F),\ \forall F\subset\partial\Omega_D\},\\ V_{h0}(t) := \{v_h\in V_h(t)\mid v_F = 0,\ \forall F\subset\partial\Omega_D\}. \end{cases} \tag{4.21}$$

The discrete stiffness bilinear form is such that for all $(v_h,\tilde{v}_h)\in V_{hD}(t)\times V_{h0}(t)$ (compare with (4.11))

$$a_h(t;v_h,\tilde{v}_h) := \sum_{c\in\mathcal{C}}|c|\varepsilon_c(\mathcal{R}(t;v_h)) : \mathbb{C} : \varepsilon_c(\mathcal{R}(t;\tilde{v}_h)) + s_h(t;v_h,\tilde{v}_h), \tag{4.22}$$

where the stabilization bilinear form s_h is intended to render a_h coercive and is defined as

$$s_h(t;v_h,\tilde{v}_h) = \sum_{F\in\mathcal{F}(t)}\frac{\eta}{h_F}|F|[\mathfrak{R}(t;v_h)]_F\cdot[\mathfrak{R}(t;\tilde{v}_h)]_F \tag{4.23}$$

where h_F is the diameter of the facet $F\in\mathcal{F}(t)$. For an interior facet $F\in\mathcal{F}^i(t)$, writing c_- and c_+ the two mesh cells sharing F, i.e., $F=\partial c_-\cap\partial c_+$, and orienting F by the unit normal vector n_F pointing from c_- to c_+, one has

$$[\mathfrak{R}(t;v_h)]_F := \mathfrak{R}(t;v_h)_{c_-}(\mathbf{x}_F) - \mathfrak{R}(t;v_h)_{c_+}(\mathbf{x}_F). \tag{4.24}$$

The sign of the jump is irrelevant in what follows. The role of the summation over the interior facets in (4.23) is to penalize the jumps of the cell reconstruction $\mathfrak{R}$ across the interior facets. For a boundary facet $F\in\mathcal{F}^b(t)$, we denote c_- the unique mesh cell containing F, we orient F by the unit normal vector $n_F := n_{c_-}$ which points outward Ω, and we define

$$[\mathfrak{R}(t;v_h)]_F := v_F - \mathfrak{R}(v_h)_{c_-}(\mathbf{x}_F). \tag{4.25}$$

The role of the summation over the boundary facets in (4.23) is to penalize the jumps between the cell reconstruction $\mathfrak{R}$ and the constant values in the boundary facets. Finally, the parameter $\eta>0$ in (4.23) is user-defined with the only requirement that $\eta>0$. The bilinear form s_h is classical in the context of discontinuous Galerkin methods (see [Arnold, 1982, Di Pietro and Ern, 2011] for instance, see also [Di Pietro, 2012] for cell-centred Galerkin methods). In practice, the penalty parameter η scales as $\eta=\beta\mu$ where μ is the second Lamé coefficient of the material and β is a dimensionless factor that remains user-dependent. In what follows, the value $\beta=1$ will be used consistently with the numerical experiments performed in Chapter 3.

The definition of the discrete mass bilinear form m_h hinges on subdomains to condense the mass associated with the dofs, as illustrated in Figure 4.4. For all the interior cells (having no boundary facet), the subdomain ω_c is chosen as the whole cell, i.e., $\omega_c := c$. For the boundary

facets and for the cells having a boundary facet, a dual barycentric subdomain is constructed, leading to subsets denoted by ω_F and ω_c, respectively. Owing to crack propagation, $\omega_c \equiv \omega_c(t)$ is time-dependent. The discrete mass bilinear form and the discrete load linear form are written as follows for all $(v_h, \tilde{v}_h) \in V_{hD}(t) \times V_{h0}(t)$ (compare with (4.10) and (4.12)):

$$m_h(t; v_h, \tilde{v}_h) := \sum_{c\in\mathcal{C}} m_c(t) v_c \cdot \tilde{v}_c + \sum_{F\in\mathcal{F}^b(t)} m_F v_F \cdot \tilde{v}_F, \tag{4.26}$$

$$l_h(t; \tilde{w}_h) := \sum_{c\in\mathcal{C}} f_c(t) \cdot \tilde{w}_c + \sum_{F\subset\partial\Omega_N} g_F(t) \cdot \tilde{v}_F, \tag{4.27}$$

with $m_c(t) := \int_{\omega_c(t)} \rho$, $m_F := \int_{\omega_F} \rho$, $f_c(t) := \int_c f(t)$ and $g_F(t) := \int_F g_N(t)$.

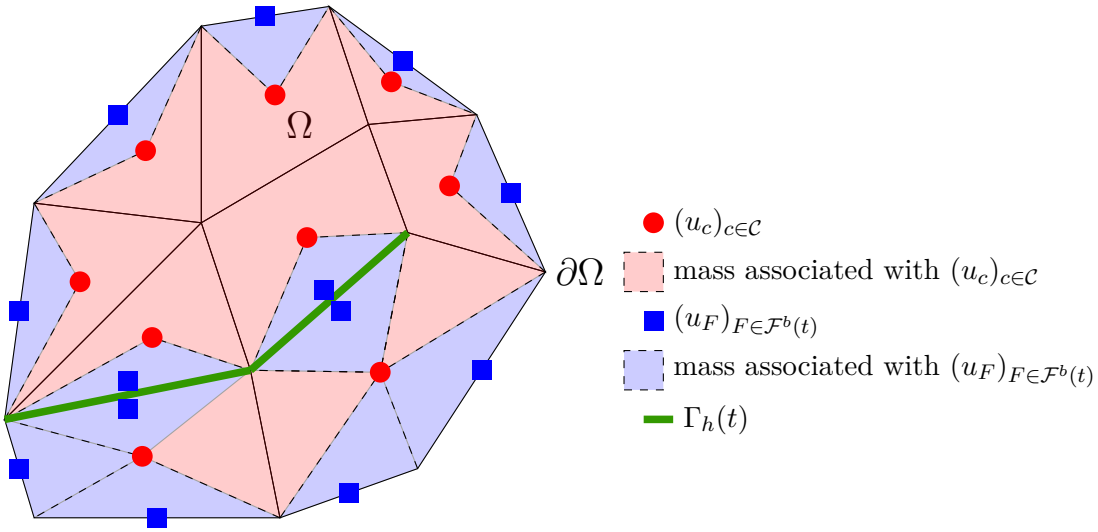


Figure 4.4 – Integration domains to determine the mass associated with the displacement dofs.

4.3.4 Verification test case

This section presents a verification test case related to the convergence rate with a singularity at the crack tip. The crack does not propagate, i.e., we consider a steady setting using the above discrete stiffness bilinear form and load linear form.

The convergence rate of the method in the presence of a singularity is tested in the case of an infinite plate under mode 3 loading at infinity as presented in Figure 4.5. A convergence

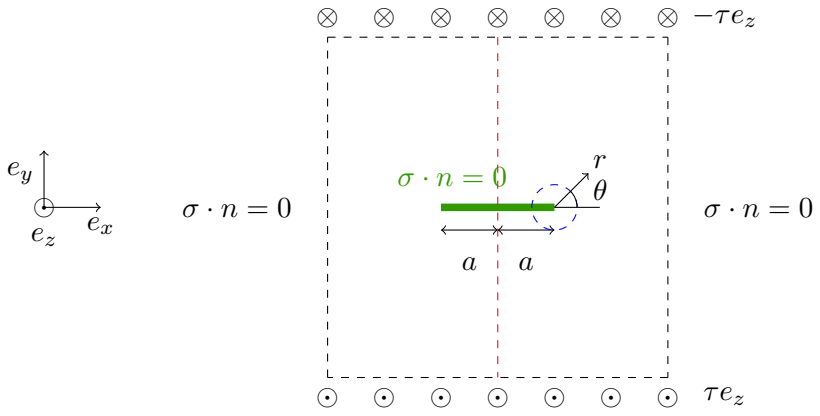


Figure 4.5 – Sketch of the antiplane shear experiment in an infinite plate.

rate of $O(\sqrt{h})$, similar to that obtained with Lagrange P^1 finite elements, is expected. The reference solution, close to the crack tip ($\frac{r}{a} \ll 1$), reads in polar coordinates [Kuna, 2013, p. 28]:

$$u(r,\theta) = \frac{2\tau}{\mu}\sqrt{\frac{ar}{2}}\sin\left(\frac{\theta}{2}\right)e_z, \tag{4.28}$$

where τ is the modulus of the antiplane shear stress imposed at infinity. The displacement defined in (4.28) verifies the statics equation in a strong form since $\mathrm{div}(u) = 0$. The stresses are

$$\sigma(r,\theta) = \tau\sqrt{\frac{a}{2r}}\left[\sin\left(\frac{\theta}{2}\right)e_r - \cos\left(\frac{\theta}{2}\right)e_\theta\right]\otimes e_z. \tag{4.29}$$

Figure 4.5 being symmetric with respect to the red dashed line, only the right part of the domain is considered. As the analytical solution (4.28) is only valid close to the crack tip, a small ball around the crack tip, which corresponds to the green dashed circle in Figure 4.5, is meshed. The setting is presented in Figure 4.6. The convergence in energy norm towards the analytical solution is checked on the meshed ball with the reference solution imposed as Dirichlet boundary condition over the whole ball boundary. The results of the computation

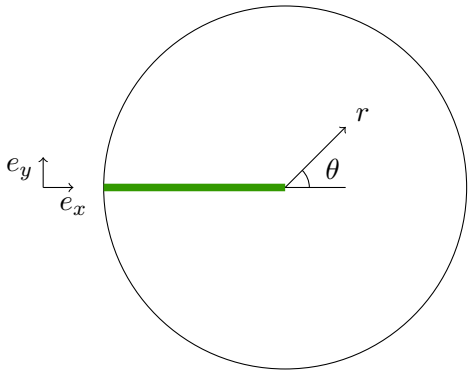


Figure 4.6 – Sketch of the meshed ball around the crack tip.

are reported in Table 4.1 which corroborates an $O(\sqrt{h})$ convergence rate on the elastic energy, as expected. The convergence rates are approximated as

nb dofs	$\|u - \mathfrak{R}(u_h)\|_{L^2}$	Convergence rate	$\|\nabla u - G_h(u_h)\|_{L^2}$	Convergence rate
182	2.90e-04	-	1.74e-01	-
632	9.82e-05	1.74	1.12e-01	0.70
2,534	4.06e-05	1.27	6.47e-02	0.80
9,744	1.69e-05	1.30	4.09e-02	0.68
37,580	8.29e-06	1.05	2.84e-02	0.54
148,218	3.86e-06	1.13	1.94e-02	0.56

Table 4.1 – Number of dofs, L^2-error and convergence rate, L^2-error on the gradient and convergence rate.

$$\text{order} = d\log\left(\frac{e_1}{e_2}\right)\left(\log\left(\frac{n_2}{n_1}\right)\right)^{-1}, \tag{4.30}$$

where e_1, e_2 denote the errors on the computations with mesh sizes h_1, h_2 and the number of dofs n_1, n_2.

4.4 Quasi-static crack propagation

In this section we formulate the discrete problem for quasi-static crack propagation. The space discretization is achieved by means of the DEM scheme presented in the previous section. The time interval $[0, T]$ is discretized by means of $(N+1)$ discrete time nodes $(t_n)_{n\in\{0,\dots,N\}}$ with $t_0 := 0$ and $t_N := T$. At every discrete time node t_n, we proceed in two steps. First, given the partition of the mesh facets into the subsets $\mathcal{F}^i(t_n)$ and $\mathcal{F}^b(t_n)$ and a description of the discrete crack $\Gamma_h(t_n)$ by means of the facets in $\mathcal{F}^\Gamma(t_n)$, we find the discrete displacement $u_h(t_n) \in V_{hD}(t_n)$ solving the quasi-static problem $a_h(t_n; u_h(t_n), \tilde{v}_h) = l_h(t_n; \tilde{v}_h)$ for all $\tilde{v}_h \in V_{h0}(t_n)$, where the discrete spaces $V_{hD}(t_n)$ and $V_{h0}(t_n)$ are defined in (4.21). Then we use the newly computed displacement field $u_h(t_n)$ to determine whether crack propagation occurs and update accordingly the subsets $\mathcal{F}^i(t_n)$, $\mathcal{F}^b(t_n)$, and $\mathcal{F}^\Gamma(t_n)$. The discrete quasi-static crack propagation scheme can thus be summarized as follows: Given $\mathcal{F}^i(t_0)$, $\mathcal{F}^b(t_0)$, and $\mathcal{F}^\Gamma(t_0)$, solve for all $n \in \{0, \dots, N\}$,

$$\begin{cases} \text{(i)} \;\; u_h(t_n) \in V_{hD}(t_n) \text{ s.t. } a_h(t_n; u_h(t_n), \tilde{v}_h) = l_h(t_n; \tilde{v}_h), \;\; \forall \tilde{v}_h \in V_{h0}(t_n), \\ \text{(ii)} \;\; (\mathcal{F}^\Gamma(t_{n+1}), \mathcal{F}^b(t_{n+1}), \mathcal{F}^i(t_{n+1})) = \texttt{CRACK\_QS}(\mathcal{F}^\Gamma(t_n), \mathcal{F}^b(t_n), \mathcal{F}^i(t_n), u_h(t_n)). \end{cases} \tag{4.31}$$

The rest of this section is devoted to the description of the procedure `CRACK_QS`. This procedure consists in three consecutive steps outlined in Figure 4.7. The first step involves the procedure `ESTIMATE` which considers the collection of vertices of the discrete crack $\Gamma_h(t_n)$ and computes for each of these vertices an approximate energy release rate. The second step involves the procedure `MARK` which flags among all the facets sharing a vertex with an energy release rate larger than the maximum value $\mathcal{G}_c$ those facets that will indeed break. The selection is made by using a discrete kinking criterion. The last step uses the procedure `UPDATE_QS` and simply consists in updating the data structure according to the crack propagation.

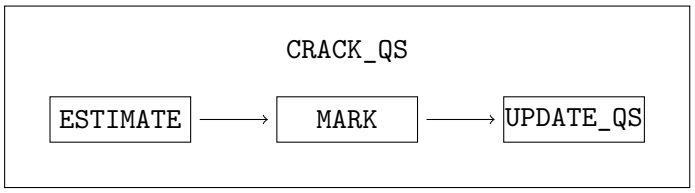


Figure 4.7 – Details of the procedure `CRACK_QS`.

4.4.1 Procedure `ESTIMATE`

Let $\mathcal{V}^\Gamma(t_n)$ denote the collection of the mesh vertices lying in $\Gamma_h(t_n)$. The procedure `ESTIMATE` computes an approximate energy release rate $\mathcal{G}_h(\mathbf{z})$ for all $\mathbf{z} \in \mathcal{V}^\Gamma(t_n)$. The computation of $\mathcal{G}_h(\mathbf{z})$ is based on the approximations of the SIF defined in (4.5) followed by the Irwin formula (4.6). Let $\mathcal{C}_\mathbf{z}$ be the set of the mesh cells containing the vertex $\mathbf{z}$. Let $F \in \mathcal{F}^\Gamma(t_n)$ be a facet in the crack to which $\mathbf{z}$ belongs. There are two such facets if $\mathbf{z}$ lies inside the crack and only one if $\mathbf{z}$ is positioned at a crack tip. Let n_F and τ_F be the normal and tangent vector to F. The geometric setting is illustrated in Figure 4.8. For all the mesh cells $c \in \mathcal{C}_\mathbf{z}$ and all the facets $F \in \mathcal{F}^\Gamma(t_n)$ containing the vertex $\mathbf{z}$, the approximate SIF associated with $\mathbf{z}$ are then evaluated as

$$\begin{cases} K_{1h}(\mathbf{z}, c, F) := (n_F \cdot \Sigma_c(t_n) \cdot n_F)\sqrt{2\pi d(\mathbf{x}_c, \mathbf{z})}, \\ K_{2h}(\mathbf{z}, c, F) := (n_F \cdot \Sigma_c(t_n) \cdot \tau_F)\sqrt{2\pi d(\mathbf{x}_c, \mathbf{z})}, \\ K_{3h}(\mathbf{z}, c, F) := (n_F \cdot \Sigma_c(t_n) \cdot e_3)\sqrt{2\pi d(\mathbf{x}_c, \mathbf{z})}, \end{cases} \tag{4.32}$$

where Σ_c is the piecewise-constant stress tensor evaluated at the mesh cell c (reconstructed using $u_h(t_n)$). The approximate energy release rate for the cell $c \in \mathcal{C}_\mathbf{v}$ and the facet $F \in \mathcal{F}^\Gamma(t_n)$,

$\mathbf{z} \in \partial F$, is evaluated as

$$\mathcal{G}_h(\mathbf{z}, c, F) := \frac{1-\nu^2}{E}\left(K_{1h}(\mathbf{z}, c, F)^2 + K_{2h}(\mathbf{z}, c, F)^2\right) + \frac{1+\nu}{E} K_{3h}(\mathbf{z}, c, F)^2. \tag{4.33}$$

Finally, the discrete energy release rate $\mathcal{G}_h(\mathbf{z})$ is evaluated as

$$\mathcal{G}_h(\mathbf{z}) := \max_{c \in \mathcal{C}_\mathbf{z}} \max_{F \in \mathcal{F}^\Gamma(t_n), \mathbf{z} \in \partial F} \mathcal{G}_h(\mathbf{z}, c, F). \tag{4.34}$$

To sum up, the output of the procedure `ESTIMATE` is the collection of approximate energy release rates $\{\mathcal{G}_h(\mathbf{z})\}_{\mathbf{z} \in \mathcal{V}^\Gamma(t_n)}$.

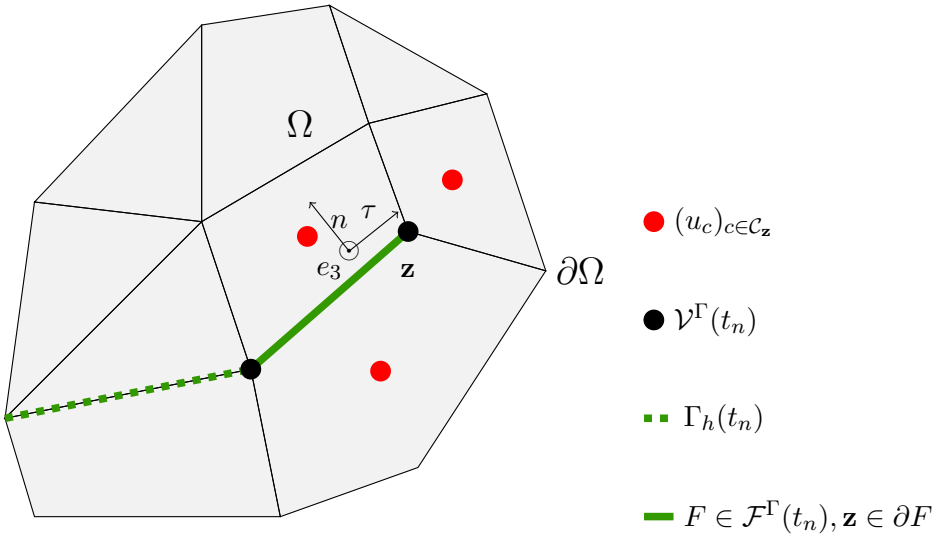


Figure 4.8 – Estimation of the SIF at the crack tip.

4.4.2 Procedure `MARK`

The goal of the procedure `MARK` is to identify those inner facets in $\mathcal{F}^i(t_n)$ through which the crack will propagate. Let us first mark the vertices in $\mathcal{V}^\Gamma(t_n)$ whose approximate energy release rate is larger than the material parameter $\mathcal{G}_c$:

$$\mathcal{V}^{\Gamma *}(t_n) := \{\mathbf{z} \in \mathcal{V}^\Gamma(t_n), \mathcal{G}_h(\mathbf{z}) \geq \mathcal{G}_c\}. \tag{4.35}$$

Then, for every marked vertex $\mathbf{z} \in \mathcal{V}^{\Gamma *}(t_n)$, we want to select an inner facet in

$$\mathcal{F}^i_\mathbf{z}(t_n) := \{F \in \mathcal{F}^i(t_n), \mathbf{z} \in \partial F\} \tag{4.36}$$

through which the crack will propagate. The setting is illustrated in Figure 4.9. The selection of the inner facet in $\mathcal{F}^i_\mathbf{z}(t_n)$ is made through a kinking criterion which, in the spirit of [Hussain et al., 1974], is based on the maximum of the density of elastic energy. This leads us to define the elastic energy density associated with an inner facet $F \in \mathcal{F}^i_\mathbf{z}(t_n)$ as

$$E_{F,\mathbf{z}}(t_n) := \frac{|\mathbf{x}_F - \mathbf{z}|}{2} \{\Sigma_\mathcal{C}(t_n)\}_F : \{\varepsilon_\mathcal{C}(t_n)\}_F, \tag{4.37}$$

where $\{\cdot\}_F$ denotes the arithmetic mean value at F of a piecewise constant quantity defined on the mesh cells, and where we used the shorthand notation $\varepsilon_\mathcal{C}(t_n) := (\varepsilon_c(\mathcal{R}(t_n; u_h(t_n))))_{c \in \mathcal{C}}$ and $\Sigma_\mathcal{C}(t_n) := (\mathbb{C} : \varepsilon_c(t_n))_{c \in \mathcal{C}}$. The presence of the factor $|\mathbf{x}_F - \mathbf{z}|$ in (4.37) is motivated by the fact that considering the singularity in $\frac{1}{\sqrt{r}}$ at the crack tip, the quantity in (4.37) is expected

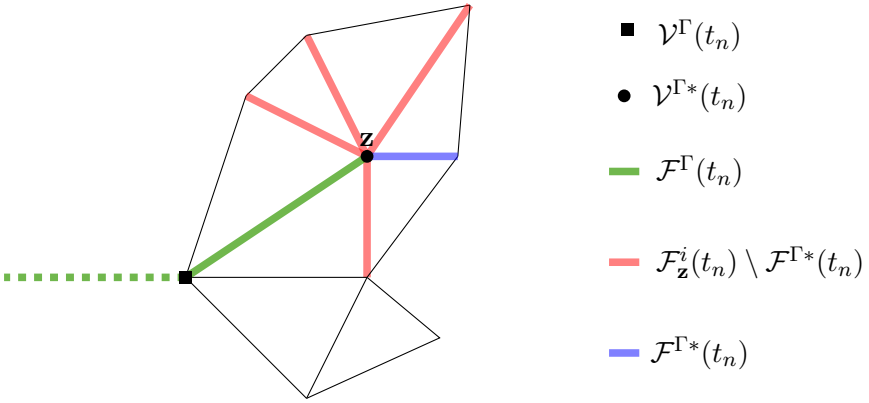


Figure 4.9 – Sketch of the discrete sets considered in the procedure `MARK`.

to depend essentially on θ and thus on the direction of the facet F. The kinking criterion then consists in selecting for all $\mathbf{z} \in \mathcal{V}^{\Gamma*}(t)$ the inner facet such that

$$F_{\mathbf{z}} := \underset{F \in \mathcal{F}^{i}_{\mathbf{z}}(t_n)}{\operatorname{argmax}} E_{F,\mathbf{z}}(t_n). \tag{4.38}$$

Finally, the output of the procedure `MARK` is the following collection of inner facets through which the crack will propagate:

$$\mathcal{F}^{\Gamma*}(t_n) := \bigcup_{\mathbf{z} \in \mathcal{V}^{\Gamma*}(t_n)} \{F_{\mathbf{z}}\}. \tag{4.39}$$

Notice that in most situations, the set $\mathcal{V}^{\Gamma*}(t_n)$ reduces to one vertex located at the crack tip and the set $\mathcal{F}^{\Gamma*}(t_n)$ contains only one mesh facet.

4.4.3 Procedure `UPDATE_QS`

The subsets $\mathcal{F}^{\Gamma}(t_{n+1})$, $\mathcal{F}^{i}(t_{n+1})$, and $\mathcal{F}^{b}(t_{n+1})$ can now be updated as follows:

$$\begin{cases} \mathcal{F}^{\Gamma}(t_{n+1}) := \mathcal{F}^{\Gamma}(t_n) \cup \displaystyle\bigcup_{F \in \mathcal{F}^{\Gamma*}(t_n)} \{F\}, \\ \mathcal{F}^{i}(t_{n+1}) := \mathcal{F}^{i}(t_n) \setminus \displaystyle\bigcup_{F \in \mathcal{F}^{\Gamma*}(t_n)} \{F\}, \\ \mathcal{F}^{b}(t_{n+1}) := \mathcal{F}^{b}(t_n) \cup \displaystyle\bigcup_{F \in \mathcal{F}^{\Gamma*}(t_n)} \{F_-, F_+\}, \end{cases} \tag{4.40}$$

where we recall that F_- and F_+ are the same geometric object as the inner facet F, but are now each one on the boundary of a single mesh cell. Note that $u_h(t_{n+1})$ has consequently $2d\#(\mathcal{F}^{\Gamma*}(t_n))$ more entries than $u_h(t_n)$.

Remark 4.1 (Update of a_h)**.** *The updates in (4.40) affect the reconstruction operator used to evaluate the discrete stiffness bilinear form. Figure 4.10 presents a sketch of an inner facet whose reconstruction has to be recomputed after a neighbouring inner facet broke. The purpose of recomputing the reconstruction on certain inner facets is to avoid using dof values on both sides of the crack in the same reconstruction. Incidentally, we notice that the recomputation of the stencil of the reconstruction operator can lead to slight variations in the elastic energy. Having a reconstruction stencil as small as possible helps to minimize this phenomenon.*

Remark 4.2 (Update of m_h)**.** *Assuming that the broken inner facet was shared by two cells with no facets on the boundary of the domain, the mass associated with the two newly created boundary facets is sketched in blue in Figure 4.11. The mass attributed to the two boundary facets is consistently removed from the mass of the corresponding cells.*

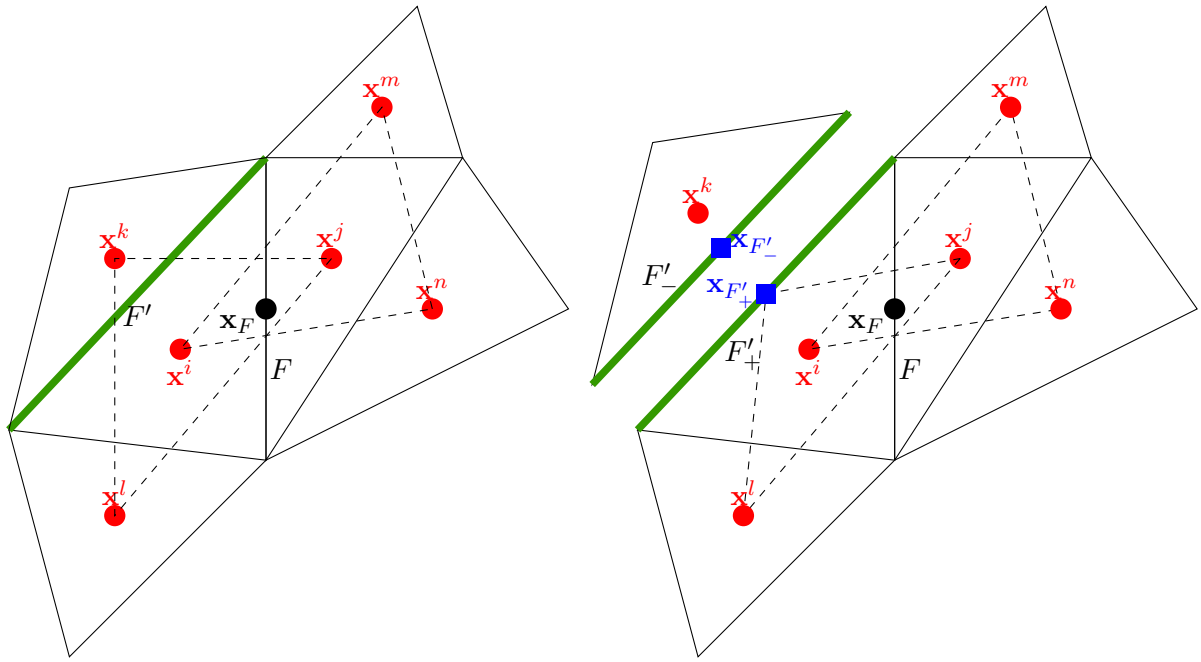


Figure 4.10 – Recomputation of the reconstruction stencil associated with the inner facet F after the breaking of the neighbouring inner facet F'. Left: reconstruction before cracking. Right: reconstruction after cracking. (The two cells separated by the crack are drawn slightly apart.)

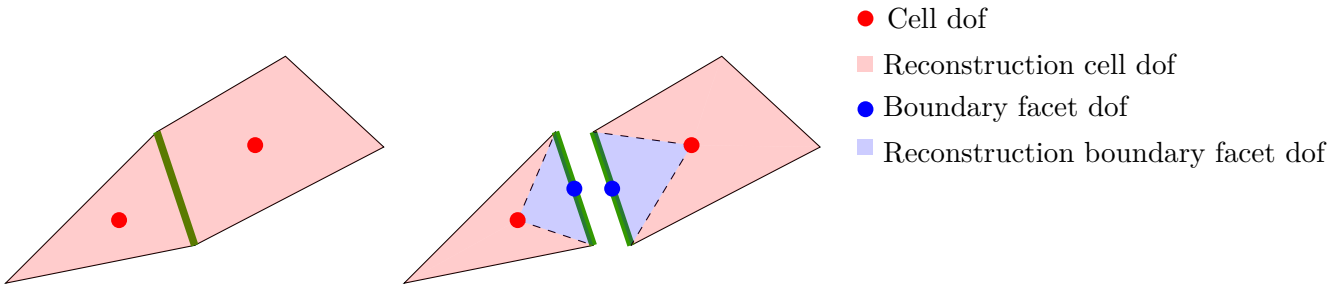


Figure 4.11 – Update of the discrete mass bilinear form after the breaking of an inner facet. (The two cells separated by the crack are drawn slightly apart.)

4.5 Fully discrete dynamic crack propagation

In this section we formulate the fully discrete problem for dynamic crack propagation. As in the quasi-static case, the space discretization is achieved by means of the DEM scheme discussed in Section 4.3, and the time interval $[0, T]$ is discretized by means of the $(N + 1)$ discrete time nodes $(t_n)_{n\in\{0,\dots,N\}}$ with $t_0 := 0$ and $t_N := T$.

4.5.1 The explicit time-integration scheme

The time discretization is performed by means of a Leapfrog method. The time-step is limited by a CFL condition which is discussed below. The key idea in the Leapfrog scheme is to approximate the displacement field at the time nodes by means of functions $u_h^n \in V_{hD}(t_n)$ for all $n \in \{0, \dots, N\}$, and the velocity field at the half-time nodes $t_{n+\frac{1}{2}} = \frac{1}{2}(t_n + t_{n+1})$ by means of functions $v_h^{n+\frac{1}{2}} \in W_{hD}(t_{n+\frac{1}{2}})$, where the discrete velocity space W_{hD} is defined as the discrete displacement space but using the time-derivative of the Dirichlet data on $\partial\Omega_D$. For

all $n \in \{0, \dots, N-1\}$, given u_h^n and $v_h^{n+\frac{1}{2}-}$, the Leapfrog scheme performs the following three substeps:

$$\begin{cases} \text{(i)} \quad u_h^{n+1-} = u_h^n + (t_{n+1} - t_n) v_h^{n+\frac{1}{2}-}, \\ \text{(ii)} \quad (\mathcal{F}^\Gamma(t_{n+1}), \mathcal{F}^b(t_{n+1}), \mathcal{F}^i(t_{n+1}), u_h^{n+1}, v_h^{n+\frac{1}{2}}) \\ \qquad = \texttt{CRACK\_DYN}(\mathcal{F}^\Gamma(t_n), \mathcal{F}^b(t_n), \mathcal{F}^i(t_n), u_h^{n+1-}, v_h^{n+\frac{1}{2}-}), \\ \text{(iii)} \quad m_h(t_{n+1}; v_h^{n+\frac{3}{2}-} - v_h^{n+\frac{1}{2}}, \tilde{v}_h) = (t_{n+3/2} - t_{n+1/2}) \Big(l_h(t_{n+1}; \tilde{v}_h) - a_h(t_{n+1}; u_h^{n+1}, \tilde{v}_h) \Big), \end{cases} \tag{4.41}$$

for all $\tilde{v}_h \in V_{h0}(t_{n+1})$. The substep (i) is performed using the crack $\Gamma_h(t_n)$ and the corresponding data structure associated with the facet subsets $\mathcal{F}^\Gamma(t_n)$, $\mathcal{F}^b(t_n)$, and $\mathcal{F}^i(t_n)$. This is why a superscript with a minus sign is used for u_h^{n+1-} and $v_h^{n+\frac{1}{2}-}$. This substep predicts the displacement field at t_{n+1} using a free-flight expression based on the current velocity field $v_h^{n+\frac{1}{2}-}$. The substep (ii), which is further commented below, treats the crack propagation from t_n to t_{n+1}, thereby obtaining the new data structure described by the facet subsets $\mathcal{F}^\Gamma(t_{n+1})$, $\mathcal{F}^b(t_{n+1})$, and $\mathcal{F}^i(t_{n+1})$, as well as the new displacement u_h^{n+1} and the new velocity $v_h^{n+\frac{1}{2}}$ accounting for the crack propagation. Finally the substep (iii) predicts the velocity field $v_h^{n+\frac{3}{2}-}$ by means of an evaluation of the forces at t_{n+1}.

The procedure `CRACK_DYN` invoked in the substep (ii) is similar to the procedure `CRACK_QS` invoked in the quasi-static setting. First, using u_h^{n+1-}, the energy release rate is estimated for all the vertices along the crack $\Gamma_h(t_n)$ using the procedure `ESTIMATE` described in Section 4.4.1. Then the inner facets through which the crack propagates are marked by means of the procedure `MARK` described in Section 4.4.2. Finally the facet subsets $\mathcal{F}^\Gamma(t_{n+1})$, $\mathcal{F}^b(t_{n+1})$, $\mathcal{F}^i(t_{n+1})$ are updated as in the quasi-static setting, but we need also to update the values of the displacement and the velocity at the newly created boundary facet dofs which now lie along the crack. Recall that for every newly cracked facet $F \in \mathcal{F}^{\Gamma *}(t_n)$, F_- and F_+ denote the boundary facets added to $\mathcal{F}^b(t_{n+1})$. Then we set

$$v_{F_-}^{n+\frac{1}{2}} := v_{c_{F_-}}^{n+\frac{1}{2}-}, \qquad v_{F_+}^{n+\frac{1}{2}} := v_{c_{F_+}}^{n+\frac{1}{2}-}, \tag{4.42}$$

and

$$u_{F_-}^{n+1} := \mathcal{R}(t_{n+1}; u_h^{n+1-})_F, \qquad u_{F_+}^{n+1} := \mathcal{R}(t_{n+1}; u_h^{n+1-})_F. \tag{4.43}$$

Instead, for every dof i that already existed at t_n, we simply set $u_i^{n+1} := u_i^{n+1-}$ and $v_i^{n+\frac{1}{2}} := v_i^{n+\frac{1}{2}-}$. (4.43) is chosen to enforce the continuity of displacements. Moreover, (4.42) is chosen to verify the conservation of linear momentum and of kinetic energy since these relations imply

Proposition 4.3 (Conservation on Kinetic energy and linear momentum)**.** *The kinetic energy is conserved through the cracking process. Assuming facets crack at t_{n+1}, one has*

$$\frac{1}{2} m_h(t_n; v_h^{n+\frac{1}{2}-}, v_h^{n+\frac{1}{2}-}) = \frac{1}{2} m_h(t_{n+1}; v_h^{n+\frac{1}{2}}, v_h^{n+\frac{1}{2}}). \tag{4.44}$$

The linear momentum is conserved through the cracking process. Assuming facets crack at t_{n+1}, one has

$$\sum_{c \in \mathcal{C}} m_c(t_n) v_c^{n+\frac{1}{2}-} + \sum_{F \in \mathcal{F}^b(t_n)} m_F v_F^{n+\frac{1}{2}-} = \sum_{c \in \mathcal{C}} m_c(t_{n+1}) v_c^{n+\frac{1}{2}} + \sum_{F \in \mathcal{F}^b(t_{n+1})} m_F v_F^{n+\frac{1}{2}}. \tag{4.45}$$

Finally, the scheme (4.41) is initialized by prescribing the displacement u_h^0 and the velocities $v_h^{\frac{1}{2}}$ from the values of the prescribed initial displacement u_0 and the prescribed initial velocity v_0 at all the initial dofs.

4.5.2 The CFL condition

A constant time-step Δt is used for simplicity. To ensure the initial stability of the method, the time step is restricted by the CFL stability condition

$$\Delta t < 2\sqrt{\frac{\mu^0_{\min}}{\lambda^0_{\max}}}, \tag{4.46}$$

where $\Delta t := t^{n+1} - t^n$, $\mu^0_{\min}$ is the smallest eigenvalue of the diagonal mass matrix associated with the discrete mass bilinear form $m_h(t_0;\cdot,\cdot)$, and $\lambda^0_{\max}$ is the largest eigenvalue of the stiffness matrix associated with the discrete stiffness bilinear form $a_h(t_0;\cdot,\cdot)$. As the time step is constant, a discrete total energy is conserved. The CFL condition (4.46) guarantees the stability of the Leapfrog time-integration scheme with a constant time step, as long as cracking does not occur. However, the updates of m_h and a_h, due to cracking, and performed by `UPDATE_QS` as described in Section 4.4.3, can potentially lead to a change in $\lambda^n_{\max}$ and $\mu^n_{\min}$, which denote respectively the largest and smallest eigenvalue of $a_h(t_n;\cdot,\cdot)$ and $m_h(t_n;\cdot,\cdot)$. As a consequence, the new CFL condition ensuring stability becomes

$$\Delta t < 2\sqrt{\frac{\mu^n_{\min}}{\lambda^n_{\max}}}. \tag{4.47}$$

Thus, the Δt chosen a priori using (4.46) should keep verifying (4.47) at all stages of the computation while remaining constant. In the test cases reported in Section 4.6, the condition (4.47) is fulfilled during all the computations.

4.6 Numerical experiments

In this section, we present numerical experiments for two-dimensional crack propagation in quasi-static and dynamic settings.

4.6.1 Verification test cases

Convergence of the discrete energy release rate

The domain is a rectangular plate of dimensions $L = 5\text{m}$ and $H = 2\text{m}$ as sketched in Figure 4.12. The plate contains a straight crack of length $a = 1\text{m}$. The material parameter

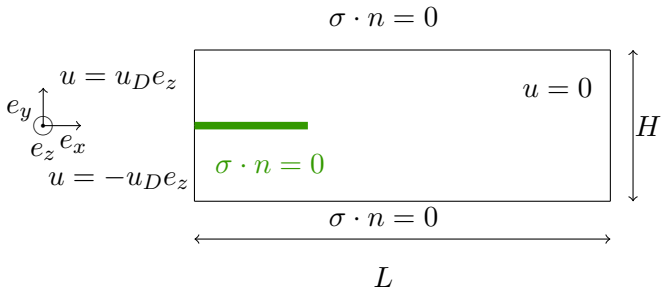


Figure 4.12 – Sketch of the 2d antiplane shear experiment.

is $\mu = 0.2\text{Pa}$. A constant opposite boundary displacement $u_D = 0.3\text{m}$ is imposed in the e_z direction at the two extremities of the plate as sketched in Figure 4.12. A sequence of refined structured triangular meshes is used to check the convergence of the discrete energy release rate in mode 3. The value $\mathcal{G}_h = 1.3483 \cdot 10^{-2}\text{J}\cdot\text{m}^{-2}$ computed on the finest mesh with $h = 3.1 \cdot 10^{-3}\text{m}$ and $4,100,480$ dofs is used as reference to compute the errors which determine the order of convergence in Table 4.2. A convergence rate of $\mathcal{O}(h)$ is observed. The values of the energy release rate are consistent with the quasi-static computation of [Li et al., 2016] for which $u_D \approx 0.3\text{m}$ and the propagation starts for $\mathcal{G}_c = 0.01\text{J}\cdot\text{m}^{-2}$.

Mesh size	Nb dofs	energy release rate	error	order
0.2	1,070	1.1663e-02	1.82e-03	-
0.1	4,140	1.2565e-02	9.18e-04	1.01
0.05	16,280	1.3043e-02	4.40e-04	1.08
0.025	64,560	1.3289e-02	1.93e-04	1.19
0.0125	1,026,240	1.3412e-02	7.11e-05	1.45
6.25e-02	4,100,480	1.3464e-02	1.85e-05	1.95

Table 4.2 – Number of scalar dofs, discrete energy release rate, absolute error, and order of convergence.

Test of the kinking criterion

This test case consists in simulating the propagation in mode 3 of an already initiated crack as already sketched in Figure 4.12 above. The evolution is quasi-static and is realized by enforcing increments on the imposed displacement $u_D(t)$. These increments are computed with a loading speed $k = 10^{-3}\text{m}\cdot\text{s}^{-1}$ over the time-interval $[0, T]$ where $T = 10^3\text{s}$. The second Lamé coefficient is chosen as $\mu = 0.2\text{Pa}$ and the stresses are defined as $\sigma = \mu\nabla u$ where u is the scalar out-of-plane displacement. The penalty parameter is chosen as $\eta = \mu \cdot 10^{-3}$ and the critical energy release rate as $\mathcal{G}_c = 0.01\text{J}\cdot\text{m}^{-2}$.

A first set of computations is performed on refined structured meshes. Three meshes with sizes $h = 5\text{cm}$, $h = 2.5\text{cm}$ and $h = 1.25\text{cm}$ are used. The corresponding constant time-steps are $\Delta t = 22\text{s}$, $\Delta t = 11\text{s}$ and $\Delta t = 5.6\text{s}$. The crack path at the final time is reported in Figure

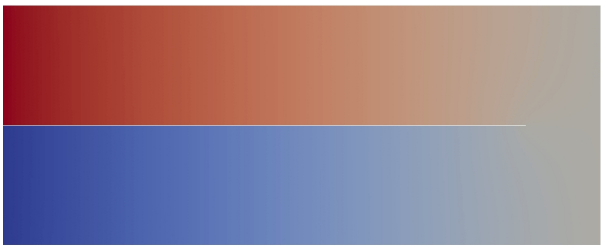


Figure 4.13 – Quasi-static kinking experiment: crack path at final time for mesh size $h = 1.25\text{cm}$.

4.13 for the finest mesh. Similar results are obtained with the coarsest meshes. The colors in Figure 4.13 represent the out-of-plane displacement u_z. One can see that the crack propagates straight, as expected.

A second set of computations is performed using a coarse unstructured mesh containing 10,862 scalar dofs of size $h = 7.4\text{cm}$ and a fine unstructured mesh containing 42,972 scalar dofs of size $h = 3.9\text{cm}$. The time-step used is $\Delta = 17\text{s}$ with the coarse mesh and $\Delta t = 8.6\text{s}$ with the fine mesh. Figure 4.14 shows the computed crack paths with the two meshes at the final time. With the unstructured meshes, the crack deviates from the straight path obtained in Figure 4.13. Similar results have been reported for phase-field methods in [Bourdin et al., 2008, p. 129–131] and seem to derive from the loss of symmetry of the discrete problem, with respect to the continuous problem, due to the loss of symmetry of the mesh. Mode 3 crack propagation is known to be metastable and thus these results seem reasonable. A noteworthy result is that the crack seems to go rather straight in both computations before kinking at 90 degrees and finishing their course at similar abscissas.

Finally, a last computation is performed on a fine mesh with 43,224 scalar dofs and mesh size $h = 3.8\text{cm}$ with a time step $\Delta t = 8.5\text{s}$. The straight propagation direction is present in the mesh so as to have similarities with the structured meshes but the mesh itself is not

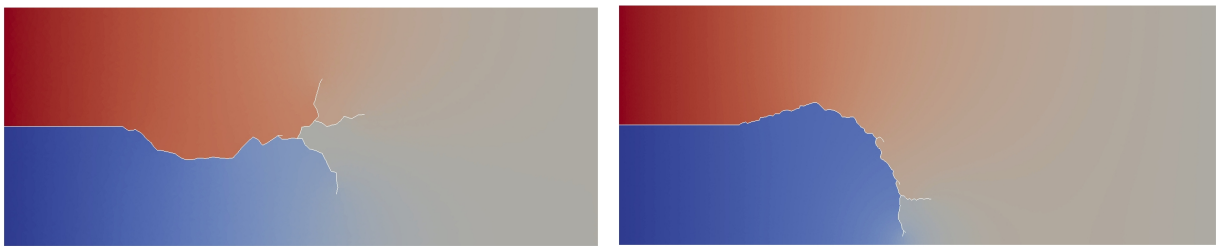


Figure 4.14 – Crack path computed with the kinking criterion (4.38) with an unstructured mesh.

symmetric as was the case with unstructured meshes. Figure 4.15 reports the crack path at the final time. The resulting crack path looks very similar to those obtained with the unstructured

Figure 4.15 – Crack path computed with the kinking criterion (4.38) with an unstructured mesh containing a straight line.

meshes thus confirming the importance of the symmetry of the mesh to obtain a straight crack evolution. An interesting feature of this crack path is that is ends at a similar abscissa as the crack paths obtained with the unstructured meshes. Brutal crack growth has been reported for this test case, using phase-field methods, in [Bourdin et al., 2008, p. 128], in the sense of an unstable crack reaching the end of the sample in a single time-step, when the crack reaches the abscissa $\frac{L}{2}$. In these computations we rather see the crack path stopping at the middle of the plate.

4.6.2 Test case with prescribed crack path

We consider a test case taken from [Li et al., 2016] in which the authors develop a dynamic phase-field model to compute dynamic cracking. The test case consists of an already cracked plate under antiplane shear loading. The difference with Section 4.6.1 is that, as in [Li et al., 2016], the crack is now forced to propagate along a straight line represented by the dashed line in Figure 4.16. The goal of this test case is to study the crack propagation velocity. We consider both quasi-static and dynamic loadings.

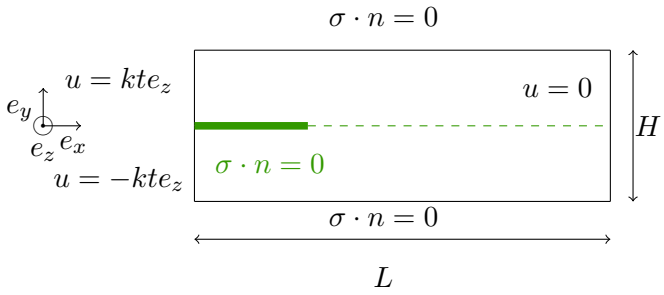


Figure 4.16 – Sketch of the 2d antiplane shear experiment with prescribed crack path.

Quasi-static evolution

The evolution is first assumed to be quasi-static and is realized by enforcing increments on the imposed displacement values. These increments are computed with a loading speed $k = 10^{-3}\text{m}\cdot\text{s}^{-1}$ and a time step Δt that we choose so as to obtain a load increment $\Delta u_D = k\Delta t$ from one quasi-static solution to the next. We are interested in the length of the crack with respect to the cumulated boundary loading displacement $u_D := kt$ where $t \in [0., 1/k]$, i.e. the final displacement load is $u_D = 1\text{m}$. In the plots of crack lengths presented below, every point corresponds to a facet cracking with the current crack length at the current u_D. The reference solution for the crack speed S with respect to the loading speed, taken from [Li et al., 2016], is $\frac{S}{k} = \sqrt{\frac{\mu H}{\mathcal{G}_c}} \approx 4.47$. As this solution is only valid when $L \to \infty$, we checked that doubling the length L of the strip did not lead to any significant change in the crack speeds. The computations are performed with structured 2d meshes of triangles with characteristic size h. The key observation from the numerical experiment is that the non-dimensional number $\chi := \frac{h}{\Delta u_D}$ has a strong influence on the computed crack speed. Two numerical experiments have been carried out to study its influence.

A first set of computations is performed with a fixed mesh size $h = 5\text{cm}$ and various constant load increments Δu_D to check their influence on the crack evolution. A reference value $\Delta u_{D_{\text{ref}}} = 11\text{mm}$ is chosen. Figure 4.17 reports the crack length as a function of the cumulated loading displacement u_D. One can see that the results for the two smallest boundary load

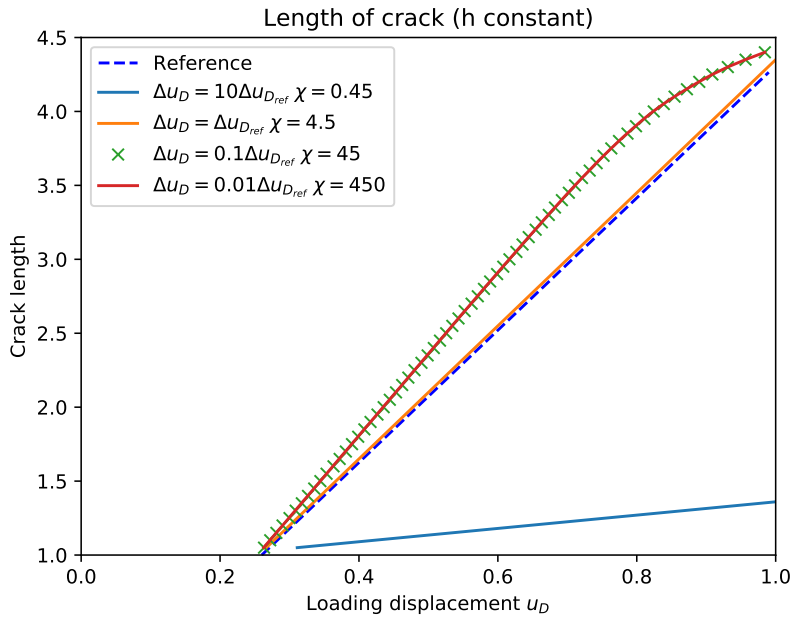


Figure 4.17 – Anti-plane shear experiment (quasi-static): Crack length as a function of cumulated loading displacement u_D.

increments (values of χ larger than 45) seem to overlap and thus one could believe that this reflects a convergence over the load increments. Unfortunately, they differ from the analytical solution. Instead, the results for $\chi = 4.5$ show an excellent agreement with the analytical solution. Finally, if χ takes lower values ($\chi = 0.45$), the discrepancy with the analytical solution is substantial.

A second set of computations is performed to verify the convergence when refining the mesh with a constant Δu_D. The load increment is fixed as $\Delta u_D = 5.6\text{mm}$. The results are reported in Figure 4.18. The evolution with $h = 0.025\text{m}$ and thus with $\chi = 4.5$ overlaps with the reference solution. The evolution with $\chi \leq 2.2$ underestimates the crack speed, whereas the evolutions with $\chi \geq 8.9$ overestimate the crack speed.

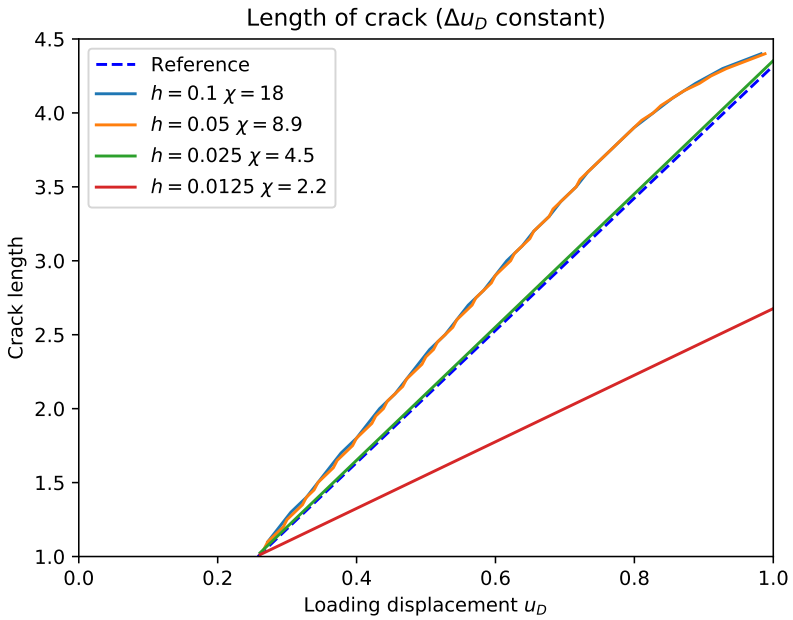


Figure 4.18 – Anti-plane shear experiment (quasi-static): Crack length as a function of cumulated loading displacement u_D.

Finally, a third set of computations is performed to verify the convergence when refining the mesh with a constant χ. Using the results of the previous computations, we set $\chi := 4.5$. We use three structured refined meshes of triangles with sizes $h = 0.1$m, $h = 0.05$m and $h = 0.025$m which correspond respectively to $4,140$ dofs, $16,280$ dofs and $64,560$ dofs. The results are reported in Figure 4.19. One can see a convergence when refining the mesh as the

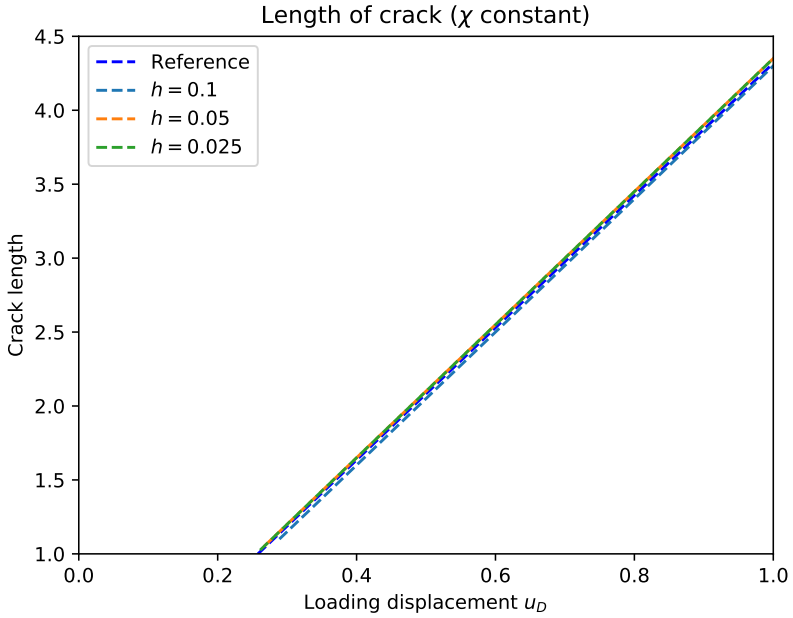


Figure 4.19 – Antiplane shear experiment (quasi-static): Crack length as a function of loading displacement u_D for mesh size and load increment fixed at the given rate $\chi = 4.5$.

two lines computed with the finest meshes overlap with the reference solution.

We understand the necessity to have a balance between h and Δu_D as follows. Assuming

h to be constant, if Δu_D is chosen too large (i.e. χ is too small), only one facet of the mesh is allowed to crack during the load increment Δu_D whereas several facets should have been broken. This is a limitation of the present algorithm that could be lifted with additional developments. Conversely, if Δu_D is chosen too small (i.e. χ is too large), a facet could break too soon thus displacing the singularity to the next facet which might break. Such a chain reaction would lead to overestimated crack speeds.

Dynamic evolution

The mass density is chosen as $\rho = 1\text{kg}\cdot\text{m}^{-3}$ and the shear wave speed is $c_s = \sqrt{\frac{\mu}{\rho}} \approx 0.45\text{m}\cdot\text{s}^{-1}$. A first computation is performed using the Leapfrog integration described above for three loading speeds: $k = 0.1\text{m}\cdot\text{s}^{-1}$, $k = 0.2\text{m}\cdot\text{s}^{-1}$ and $k = 0.3\text{m}\cdot\text{s}^{-1}$. Results are reported in Figure 4.20. The results for $h = 2.5\text{cm}$ are not reported as they are similar to those for $h = 5\text{cm}$. The time steps are $\Delta t = 2.7 \cdot 10^{-2}\text{s}$ for $h = 10\text{cm}$ and $\Delta t = 1.7 \cdot 10^{-2}\text{s}$ for $h = 5\text{cm}$ and $\Delta t = 6.7 \cdot 10^{-3}\text{s}$ for $h = 2.5\text{cm}$ so as to verify the CFL condition (4.46). The simulation

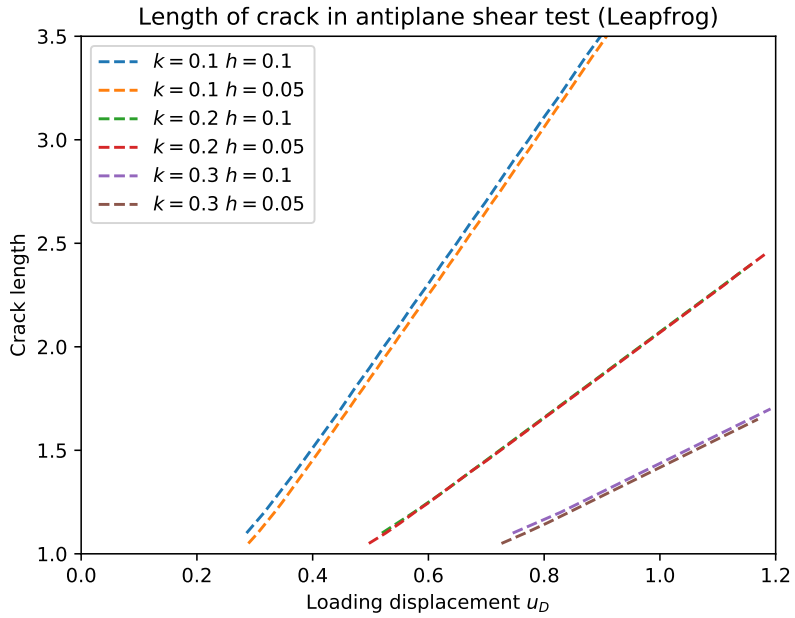


Figure 4.20 – Antiplane shear experiment (dynamic): Crack length as a function of cumulated loading displacement u_D.

period is $[0, 1.2/k]$ so that the final cumulated loading displacement is 1.2m. The initial times of propagation seem to coincide with [Li et al., 2016, Fig. 7]. However, the slope of the curves for $k = 0.1\text{m}\cdot\text{s}^{-1}$ is too steep. Having fixed Δt and h, one has for $k = 0.1\text{m}\cdot\text{s}^{-1}$, $\chi = 29$, for $k = 0.2\text{m}\cdot\text{s}^{-1}$, $\chi = 15$ and for $k = 0.3\text{m}\cdot\text{s}^{-1}$, $\chi = 9.8$. Referring to the results given above, the value $\chi = 29$ could be too high and thus lead to the overestimated crack speed. In the case $k = 0.1\text{m}\cdot\text{s}^{-1}$, having fixed h, one would want to increase Δt so as to decrease χ. However, one is limited by the CFL condition (4.46) in the increase of Δt.

A second set of computations is performed to assess the possibilities of the Leapfrog integration over a large interval of loading speeds. The reference crack speed from [Li et al., 2016] is

$$S = \sqrt{\frac{\mu H}{\mathcal{G}_c + \rho H k^2}}. \tag{4.48}$$

For all the computations, the mesh is fixed at a size $h = 5\text{cm}$ and the time-step is taken as $\Delta t = 1.7 \cdot 10^{-2}\text{s}$ so as to verify the CFL condition (4.46). The penalty parameter is fixed to

$\eta = \mu \cdot 10^{-3}$. The final time of computation goes from $T = \frac{1}{k}$ to $T = \frac{4}{k}$ so as to have enough points for the linear regression giving the crack speed. Figure 4.21 presents the results with respect to the reference solution (4.48). The results in Figure 4.21 seem to fit satisfactorily

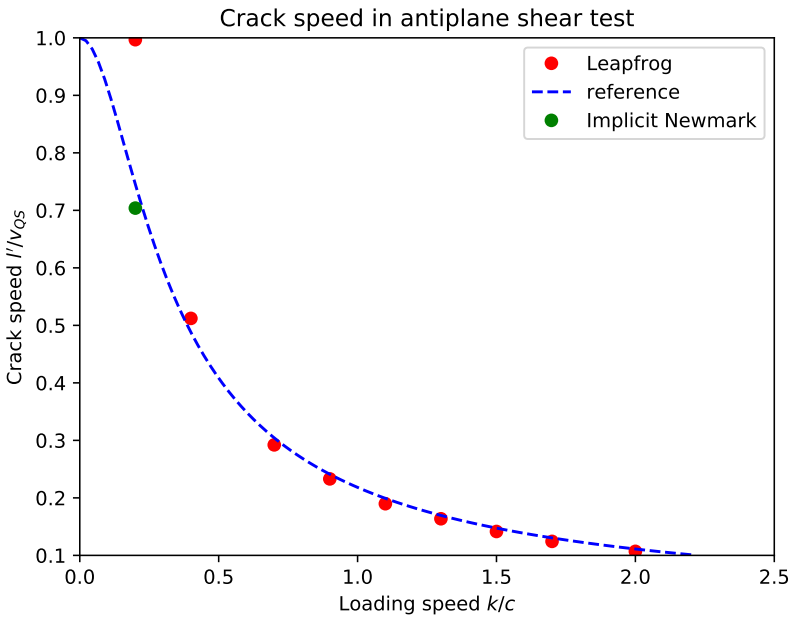


Figure 4.21 – Antiplane shear experiment (dynamics): crack speed as a function of loading speed.

the reference solution from Equation (4.48) apart for the point at $k = 0.1\text{m}\cdot\text{s}^{-1} \simeq 0.2c_s$. To improve on this value, an implicit Newmark-β scheme can be used, as in [Li et al., 2016], with $\beta = \frac{1}{4}$. With a similar mesh and a time step $\Delta t = 0.14\text{s}$ and thus about eight times larger than the time-step of the Leapfrog integration scheme, the result is much closer to the reference value.

The implicit Newmark–β scheme used to solve the cracking problem writes as follows. Let w_h^n be the approximate acceleration at the discrete time t^n. Assuming that the quantities u_h^n, v_h^n and w_h^n are known, one computes u_h^{n+1-}, v_h^{n+1-} and w_h^n by solving the following system:

$$\text{(i)} \begin{cases} u_h^{n+1-} = u_h^n + (t_{n+1} - t_n) v_h^{n+1-}, \\ v_h^{n+1-} = v_h^n + (t_{n+1} - t_n) \dfrac{w_h^n + w_h^{n+1-}}{2}, \\ m_h(t_{n+1}; w_h^{n+1-}, \tilde{v}_h) = l_h(t_{n+1}; \tilde{v}_h) - a_h(t_{n+1}; u_h^{n+1-}, \tilde{v}_h), \quad \forall \tilde{v}_h \in V_{h0}(t_n). \end{cases} \tag{4.49}$$

Then, the quantities u_h^{n+1}, v_h^{n+1} and w_h^{n+1} are computed as follows

$$\text{(ii)} \begin{cases} (\mathcal{F}^\Gamma(t_{n+1}), \mathcal{F}^b(t_{n+1}), \mathcal{F}^i(t_{n+1}), u_h^{n+1}, v_h^{n+1}), \\ = \texttt{CRACK\_DYN}(\mathcal{F}^\Gamma(t_n), \mathcal{F}^b(t_n), \mathcal{F}^i(t_n), u_h^{n+1-}, v_h^{n+1-}), \\ \forall F \in \mathcal{F}^{\Gamma *}(t_{n+1}), \quad w_{F_-}^{n+1} := 0, \quad w_{F_+}^{n+1} := 0, \end{cases} \tag{4.50}$$

i.e. the acceleration is initialized to zero at the newly created facets after the crack propagates.

4.6.3 Quasi-static test case in mode 1

This test case aims at confirming the effectiveness of the kinking criterion (4.38). The setting is similar to Figure 4.12 except that the computation is quasi-static and thus $\Delta U \equiv \Delta U(t)$. The

computations are performed on two unstructured meshes of sizes $h = 0.74$mm and $h = 0.15$mm corresponding respectively to $11,172$ dofs and $271,660$ dofs.

The finer mesh is not a refinement of the coarser one. The computation is performed with a Dirichlet boundary condition on the normal component of the displacement to assess the behaviour of cracking in mode 1. Figure 4.22 shows the crack paths obtained on the two meshes. The crack paths obtained in mode 1 are satisfactory as the propagation is rather

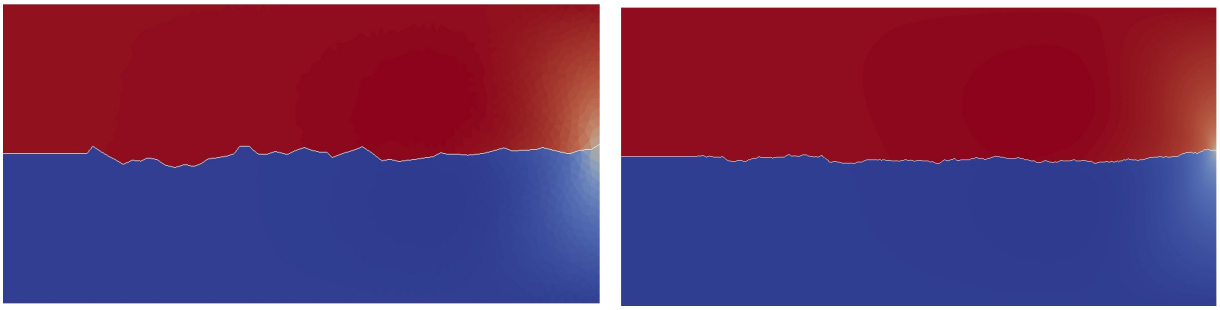


Figure 4.22 – Mode 1 (quasi-static): crack path on unstructured meshes.

straight and the results on the two meshes are quite similar.

4.7 Conclusion

We have presented a variational Discrete Element Method (DEM) to compute Griffith crack propagation. The crack propagates through the facets of the mesh and thus between discrete elements. The DEM is a consistent discretization of a Cauchy continuum and only requires continuum macroscopic parameters such as the Young modulus and the Poisson ratio for its implementation. The displacement degrees of freedom are attached to the barycentre of the mesh cells and to the barycentre of the boundary facets. A discrete Stokes formula is used to devise a piecewise constant gradient and linearized strain reconstructions. An approximation of the energy release rate similar to a Gmax criterion has been devised in the procedure `ESTIMATE` as well as a kinking criterion based on the density of elastic energy in the spirit of [Hussain et al., 1974] in the procedure `MARK`. Finally, two procedures `UPDATE_QS` and `UPDATE_DYN` have been devised to update the necessary discrete quantities after the facets that have been marked have been broken. Convergence tests in antiplane shear have confirmed the efficiency of the `ESTIMATE` procedure as well as the $\mathcal{O}(\sqrt{h})$ convergence rate in energy norm. The procedure `MARK` has been assessed in antiplane shear and in mixed modes (1 and 2) on structured and unstructured meshes. Finally, the subtle link between the time and space discretization has been explored to achieve correct crack speeds.

Several aspects of this work can be improved. The first is to adapt the present methodology to three-dimensional problems with two-dimensional cracks. The second aspect is to develop an algorithm that allows several contiguous facets to break in a single iteration. A third aspect that can be improved is the regularity of the crack surface. Indeed, in the spirit of [Francfort and Marigo, 1998], a crack should be a minimal surface. To achieve this goal, the DEM could be coupled to gradient flows used for surface lifting, as in [Romon, 2013], with the goal to move the vertices of the crack surface. One would then have to verify the convergence of the discrete crack area with tools similar to [Hildebrandt et al., 2006]. A fourth aspect that could be improved, is to approximate cohesive cracking laws instead of a Griffith cracking law so as to enable the possibility to simulate crack initiation as well as crack propagation. Inspiration can be found in [Mariotti et al., 2009] which is DEM with a linear cohesive law. Finally, a last aspect that could be improved is the convergence order of the method (which is only $\mathcal{O}(\sqrt{h})$). A similar enrichment as in [Chahine et al., 2008] could be used close to the crack tip to obtain a convergence order of $\mathcal{O}(h)$. Similar enrichments can be envisaged for

HHO methods [Di Pietro and Ern, 2015] so to have dofs on the facets but not at the nodes of the mesh.

Chapitre 5

Conclusions et perspectives

Les travaux menés dans cette thèse ont permis de proposer une méthode éléments discrets variationnelle qui, couplée à une méthode d'intégration temporelle explicite conservant une pseudo-énergie discrète, permet de traiter les problèmes de fragmentation dans des matériaux ayant des comportements non-linéaires comme ceux intéressant le CEA.

Dans le chapitre 2, une méthode d'intégration temporelle explicite pour les systèmes hamiltoniens, qui permet de conserver une pseudo-énergie, y compris dans des cas non-linéaires et avec un pas de temps variable, a été développée. La preuve de stabilité a été établie à pas de temps constant et dans le cas linéaire. En outre, un critère de stabilité énergétique a posteriori a été développé pour les cas non-linéaires et pour les cas à pas de temps variable. Ce critère permet, à moindre coût, d'adapter le pas de temps tout en s'assurant de la stabilité de la méthode. L'efficacité de la méthode a été évaluée en particulier dans le cas d'un système hyperbolique non-linéaire dégénéré présenté dans [Chabassier and Joly, 2010]. La possibilité d'utiliser le schéma avec des pas de temps locaux a également été testée numériquement avec succès.

Dans le chapitre 3, afin de pouvoir mener des calculs rigoureux avec des éléments discrets dans le cas de comportements non-linéaires, une méthode élements discrets variationnelle a été développée. Cette méthode est consistante et permet de mener des calculs d'élasto-plasticité quasi-statiques et dynamiques pour une loi d'écrouissage plastique durcissante. Les degrés de liberté (ddl) sont placés au barycentre des cellules du maillage et aux nœuds du bord. Les ddl des nœuds sur le bord permettent de n'avoir recours qu'à de l'interpolation pour reconstruire les valeurs au barycentre des facettes proches du bord, ce qui tempère la dégradation de la condition CFL provenant de la matrice de rigidité dont les valeurs propres peuvent croître fortement si de l'extrapolation est utilisée. La méthode développée permet d'appliquer les outils d'analyse numérique développés dans le cadre des méthodes aux éléments finis, en les étendant aux éléments discrets et de s'affranchir de la contrainte des maillages Voronoi en permettant l'utilisation de maillages comme les tétraèdres, hexaèdres ou prismes, par example. Enfin, la méthode a été validée sur des cas test quasi-statiques pour lesquels des solutions analytiques sont connues. La MED a également été couplée avec l'intégrateur temporel développé au chapitre précédent et comparée à des éléments finis de Lagrange P^1 en trois dimensions d'espace dans le cas d'une poutre en flexion dynamique et d'un cylindre en torsion dynamique. Les résultats indiquent divers avantages pour l'approche MED.

Dans le chapitre 4, la méthode développée au chapitre 3 a été légèrement adaptée afin de pouvoir mener des calculs de fissuration dynamique en deux dimensions d'espace. Les ddl sur le bord du domaine ne se trouvent plus aux nœuds mais au barycentre des facettes de bord. Ce choix permet de simplifier la création de ddl sur les lèvres de la fissure lors de sa propagation. Une reconstruction dite « symétrique » des déplacements dans les facettes a été mise en place afin de ne pas introduire d'anisotropie locale dans la reconstruction des contraintes. Un critère de propagation de fissure a été développé en se fondant sur une approximation des facteurs d'intensité des contraintes à proximité de la fissure. De plus, un critère de choix de direction

de propagation a été développé afin de déterminer le chemin de fissure. Ce dernier a été validé sur des cas de chargements pour lesquels l'orientation de fissure théorique est connue. Enfin, une étude de l'influence de la discrétisation en espace et en temps sur la vitesse de propagation de fissure a été menée aussi bien pour des évolutions quasi-statiques que dynamiques.

Plusieurs développements complémentaires ont été identifiés pour faire suite à ce travail de thèse.

Concernant la méthode d'intégration temporelle du chapitre 2, une preuve de la stabilité en pas de temps variable et pour des énergies non-quadratiques présenterait un intérêt notable. En effet, des preuves de stabilité dans ces cas pour des méthodes explicites ne sont pas, à notre connaissance, disponibles.

En vue de simuler l'expérience de fragmentation de coquille présentée dans le chapitre 1, il serait nécessaire d'inclure la physique manquante. Tout d'abord, l'ajout d'une gestion du contact avec frottement traité de façon consistante en utilisant des méthodes issues de [Chouly et al., 2017], par exemple, couplé avec la MED du chapitre 4 permettrait de rendre compte de l'ouverture et de la fermeture des pointes de fissure dans les calculs de fissuration dynamique. Ensuite, la prise en compte du raidissement des matériaux fortement comprimés demanderait également des développements supplémentaires à cause de la potentielle formation de chocs au sein des échantillons étudiés. En effet, les gradients centrés, comme la reconstruction du gradient utilisée dans les chapitres 3 et 4, peuvent engendrer des oscillations associées à un phénomène de Gibbs au voisinage des chocs, comme suggéré dans [Blanc and Josien, 2017]. À cause de la surestimation des contraintes due au phénomène de Gibbs, une dissipation plastique trop importante risquerait de fausser les résultats de la simulation numérique. Des possibilités pour atténuer ce phénomène, comme l'ajout d'une dissipation ou d'une pseudo-dissipation, sont envisageables. La modification du problème étudié afin d'ajouter de la dispersion, dans une approche similaire à [Pu et al., 2018], pourrait également contribuer à écarter le problème sans perte d'énergie due à la dissipation artificielle ajoutée. Enfin des méthodes permettant de préserver les symétries du problème et ayant prouvé leur efficacité dans le cas d'un problème de Burgers comme dans [Chhay and Hamdouni, 2010] peuvent également être envisagées.

Concernant la MED du chapitre 4, le premier développement complémentaire serait le passage à la fissuration en trois dimensions, qui présente de nombreux défis. Il serait également intéressant de modifier pendant un calcul les maillages utilisés à l'aide de flots de gradients géométriques, comme présentés dans [Romon, 2013], car cela permettrait de minimiser la courbure des surfaces de fissure. L'étude serait d'abord menée dans le cadre d'évolutions quasi-statiques afin d'essayer de prouver des Γ-convergences vers le modèle de [Francfort and Marigo, 1998], dans l'esprit de [Francfort et al., 2009]. En conséquence, cela permettrait de diminuer la dépendance du calcul au maillage en prouvant la convergence à la limite vers une solution du problème de fissuration. En complément, des estimateurs a posteriori, construits sur le modèle de [Hannukainen et al., 2012], pourraient être mis en place afin d'adapter dynamiquement le maillage de la MED du chapitre 4 aux singularités induites par l'amorçage et la propagation de fissures. Enfin, le développement d'une version cohésive de la MED variationnelle adaptée à la fissuration du chapitre 4, dans l'esprit de [Mariotti et al., 2009], permettrait de faire naturellement le lien entre le critère d'amorçage et celui de propagation.

Bibliographie

[André et al., 2019] André, D., Girardot, J., and Hubert, C. (2019). A novel DEM approach for modeling brittle elastic media based on distinct lattice spring model. *Computer Methods in Applied Mechanics and Engineering*, 350 :100–122.

[André et al., 2013] André, D., Jebahi, M., Iordanoff, I., Charles, J.-L., and Néauport, J. (2013). Using the discrete element method to simulate brittle fracture in the indentation of a silica glass with a blunt indenter. *Computer Methods in Applied Mechanics and Engineering*, 265 :136–147.

[Arnold, 1982] Arnold, D. (1982). An interior penalty finite element method with discontinuous elements. *SIAM journal on numerical analysis*, 19(4) :742–760.

[Asay and Shahinpoor, 2012] Asay, J. and Shahinpoor, M. (2012). *High-pressure shock compression of solids*. Springer Science & Business Media.

[Atrash et al., 2011] Atrash, F., Hashibon, A., Gumbsch, P., and Sherman, D. (2011). Phonon emission induced dynamic fracture phenomena. *Physical review letters*, 106(8) :085502.

[Barenblatt, 1962] Barenblatt, G. I. (1962). The mathematical theory of equilibrium cracks in brittle fracture. In *Advances in applied mechanics*, volume 7, pages 55–129. Elsevier.

[Belytschko and Hughes, 1983] Belytschko, T. and Hughes, T. J. R. (1983). *Computational methods for transient analysis*, volume 1. Amsterdam, North-Holland(Computational Methods in Mechanics.

[Belytschko et al., 1994] Belytschko, T., Lu, Y. Y., and Gu, L. (1994). Element-free Galerkin methods. *International journal for numerical methods in engineering*, 37(2) :229–256.

[Belytschko et al., 1995] Belytschko, T., Lu, Y. Y., Gu, L., and Tabbara, M. (1995). Element-free Galerkin methods for static and dynamic fracture. *International Journal of Solids and Structures*, 32(17-18) :2547–2570.

[Blanc and Josien, 2017] Blanc, X. and Josien, M. (2017). From the newton equation to the wave equation : the case of shock waves. *Applied Mathematics Research eXpress*, 2017(2) :338–385.

[Bleyer, 2018] Bleyer, J. (2018). Numerical tours of computational mechanics with fenics.

[Bleyer et al., 2017] Bleyer, J., Roux-Langlois, C., and Molinari, J.-F. (2017). Dynamic crack propagation with a variational phase-field model : limiting speed, crack branching and velocity-toughening mechanisms. *International Journal of Fracture*, 204(1) :79–100.

[Bolis et al., 2013] Bolis, C., Counilh, D., Lagrange, J., and Frugier, P. (2013). Fragmentation of a titanium alloy shell in expansion : From experiment to simulation. *Procedia Engineering*, 58 :672–677.

[Bourdin et al., 2000] Bourdin, B., Francfort, G. A., and Marigo, J.-J. (2000). Numerical experiments in revisited brittle fracture. *Journal of the Mechanics and Physics of Solids*, 48(4) :797–826.

[Bourdin et al., 2008] Bourdin, B., Francfort, G. A., and Marigo, J.-J. (2008). The variational approach to fracture. *Journal of elasticity*, 91(1-3) :5–148.

[Braides et al., 1999] Braides, A., Dal Maso, G., and Garroni, A. (1999). Variational formulation of softening phenomena in fracture mechanics : The one-dimensional case. *Archive for Rational Mechanics and Analysis*, 146(1) :23–58.

[Budninskiy et al., 2016] Budninskiy, M., Liu, B., Tong, Y., and Desbrun, M. (2016). Power coordinates : a geometric construction of barycentric coordinates on convex polytopes. *ACM Transactions on Graphics (TOG)*, 35(6) :241.

[Calvo and Sanz-Serna, 1993] Calvo, M. P. and Sanz-Serna, J. M. (1993). The development of variable-step symplectic integrators, with application to the two-body problem. *SIAM Journal on Scientific Computing*, 14 :936–952.

[Camacho and Ortiz, 1996] Camacho, G. T. and Ortiz, M. (1996). Computational modelling of impact damage in brittle materials. *International Journal of solids and structures*, 33(20-22) :2899–2938.

[Carstensen, 1999] Carstensen, C. (1999). Numerical analysis of the primal problem of elastoplasticity with hardening. *Numerische Mathematik*, 82(4) :577–597.

[Celigueta et al., 2017] Celigueta, M. A., Latorre, S., Arrufat, F., and Oñate, E. (2017). Accurate modelling of the elastic behavior of a continuum with the discrete element method. *Computational Mechanics*, 60(6) :997–1010.

[Chabassier and Imperiale, 2013] Chabassier, J. and Imperiale, S. (2013). Introduction and study of fourth order theta schemes for linear wave equations. *Journal of Computational and Applied Mathematics*, 245 :194–212.

[Chabassier and Joly, 2010] Chabassier, J. and Joly, P. (2010). Energy preserving schemes for nonlinear Hamiltonian systems of wave equations : Application to the vibrating piano string. *Computer Methods in Applied Mechanics and Engineering*, 199(45-48) :2779–2795.

[Chahine et al., 2008] Chahine, E., Laborde, P., and Renard, Y. (2008). Crack tip enrichment in the XFEM using a cutoff function. *International journal for numerical methods in engineering*, 75(6) :629–646.

[Chambolle, 2004] Chambolle, A. (2004). An approximation result for special functions with bounded deformation. *Journal de mathématiques pures et appliquées*, 83(7) :929–954.

[Chambolle and Crismale, 2019] Chambolle, A. and Crismale, V. (2019). A density result in $GSBD^p$ with applications to the approximation of brittle fracture energies. *Archive for Rational Mechanics and Analysis*, 232(3) :1329–1378.

[Chhay and Hamdouni, 2010] Chhay, M. and Hamdouni, A. (2010). Lie symmetry preservation by finite difference schemes for the Burgers equation. *Symmetry*, 2(2) :868–883.

[Chouly et al., 2017] Chouly, F., Hild, P., Lleras, V., and Renard, Y. (2017). Nitsche-based finite element method for contact with Coulomb friction. In *European Conference on Numerical Mathematics and Advanced Applications*, pages 839–847. Springer.

[Cundall and Strack, 1979] Cundall, P. and Strack, O. (1979). A discrete numerical model for granular assemblies. *geotechnique*, 29(1) :47–65.

[Dal Maso, 2013] Dal Maso, G. (2013). Generalised functions of bounded deformation. *Journal of the European Mathematical Society*, 15(5) :1943–1997.

[Di Pietro, 2012] Di Pietro, D. A. (2012). Cell centered Galerkin methods for diffusive problems. *ESAIM : Mathematical Modelling and Numerical Analysis*, 46(1) :111–144.

[Di Pietro and Ern, 2011] Di Pietro, D. A. and Ern, A. (2011). *Mathematical aspects of discontinuous Galerkin methods*, volume 69. Springer Science & Business Media.

[Di Pietro and Ern, 2015] Di Pietro, D. A. and Ern, A. (2015). A hybrid high-order locking-free method for linear elasticity on general meshes. *Computer Methods in Applied Mechanics and Engineering*, 283 :1–21.

[Diaz and Grote, 2009] Diaz, J. and Grote, M. J. (2009). Energy conserving explicit local time stepping for second-order wave equations. *SIAM Journal on Scientific Computing*, 31(3) :1985–2014.

[Doyen et al., 2013] Doyen, D., Ern, A., and Piperno, S. (2013). Quasi-explicit time-integration schemes for dynamic fracture with set-valued cohesive zone models. *Computational Mechanics*, 52(2) :401–416.

[Droniou et al., 2018] Droniou, J., Eymard, R., Gallouët, T., Guichard, C., and Herbin, R. (2018). *The gradient discretisation method*, volume 82. Springer.

[Droniou et al., 2010] Droniou, J., Eymard, R., Gallouët, T., and Herbin, R. (2010). A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Mathematical Models and Methods in Applied Sciences*, 20(02) :265–295.

[Droniou et al., 2013] Droniou, J., Eymard, R., Gallouet, T., and Herbin, R. (2013). Gradient schemes : a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic equations. *Mathematical Models and Methods in Applied Sciences*, 23(13) :2395–2432.

[Droniou and Lamichhane, 2015] Droniou, J. and Lamichhane, B. P. (2015). Gradient schemes for linear and non-linear elasticity equations. *Numerische Mathematik*, 129(2) :251–277.

[Eymard et al., 2004] Eymard, R., Gallouët, T., and Herbin, R. (2004). A finite volume scheme for anisotropic diffusion problems. *Comptes Rendus Mathématique*, 339(4) :299–302.

[Eymard et al., 2009] Eymard, R., Gallouët, T., and Herbin, R. (2009). Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI : a scheme using stabilization and hybrid interfaces. *IMA Journal of Numerical Analysis*, 30(4) :1009–1043.

[Fetecau et al., 2003] Fetecau, R. C., Marsden, J. E., Ortiz, M., and West, M. (2003). Nonsmooth Lagrangian mechanics and variational collision integrators. *SIAM Journal on Applied Dynamical Systems*, 2(3) :381–416.

[Fong et al., 2008] Fong, W., Darve, E., and Lew, A. (2008). Stability of asynchronous variational integrators. *Journal of Computational Physics*, 227(18) :8367–8394.

[Francfort et al., 2009] Francfort, G. A., Le, N. Q., and Serfaty, S. (2009). Critical points of Ambrosio-Tortorelli converge to critical points of Mumford-Shah in the one-dimensional Dirichlet case. *ESAIM : Control, Optimisation and Calculus of Variations*, 15(3) :576–598.

[Francfort and Marigo, 1998] Francfort, G. A. and Marigo, J.-J. (1998). Revisiting brittle fracture as an energy minimization problem. *Journal of the Mechanics and Physics of Solids*, 46(8) :1319–1342.

[Freund, 1998] Freund, L. B. (1998). *Dynamic fracture mechanics*. Cambridge university press.

[Gonzalez and Simo, 1996] Gonzalez, O. and Simo, J. C. (1996). On the stability of symplectic and energy-momentum algorithms for non-linear Hamiltonian systems with symmetry. *Computer Methods in Applied Mechanics and Engineering*, 134(3-4) :197–222.

[Grady and Olsen, 2003] Grady, D. and Olsen, M. (2003). A statistics and energy based theory of dynamic fragmentation. *International Journal of Impact Engineering*, 29(1-10) :293–306.

[Griffith, 1921] Griffith, A. A. (1921). 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.

[Groß et al., 2005] Groß, M., Betsch, P., and Steinmann, P. (2005). Conservation properties of a time FE method. part IV : Higher order energy and momentum conserving schemes. *International Journal for Numerical Methods in Engineering*, 63(13) :1849–1897.

[Hager et al., 2012] Hager, C., Hauret, P., Le Tallec, P., and Wohlmuth, B. I. (2012). Solving dynamic contact problems with local refinement in space and time. *Computer Methods in Applied Mechanics and Engineering*, 201 :25–41.

[Hairer, 1997] Hairer, E. (1997). Variable time step integration with symplectic methods. *Applied Numerical Mathematics*, 25(2-3) :219–227.

[Hairer, 2010] Hairer, E. (2010). Energy-preserving variant of collocation methods. *Journal of Numerical Analysis, Industrial and Applied Mathematics*, 5 :73–84.

[Hairer et al., 2006] Hairer, E., Lubich, C., and Wanner, G. (2006). *Geometric numerical integration : structure-preserving algorithms for ordinary differential equations*, volume 31. Springer.

[Han and Meng, 2001] Han, W. and Meng, X. (2001). Error analysis of the reproducing kernel particle method. *Computer Methods in Applied Mechanics and Engineering*, 190(46-47) :6157–6181.

[Han and Reddy, 2012] Han, W. and Reddy, B. D. (2012). *Plasticity : mathematical theory and numerical analysis*, volume 9. Springer Science & Business Media.

[Hannukainen et al., 2012] Hannukainen, A., Stenberg, R., and Vohralík, M. (2012). A unified framework for a posteriori error estimation for the Stokes problem. *Numerische Mathematik*, 122(4) :725–769.

[Hansbo and Larson, 2003] Hansbo, P. and Larson, M. (2003). Discontinuous galerkin and the crouzeix–raviart element : application to elasticity. *ESAIM : Mathematical Modelling and Numerical Analysis*, 37(1) :63–72.

[Hansbo and Salomonsson, 2015] Hansbo, P. and Salomonsson, K. (2015). A discontinuous Galerkin method for cohesive zone modelling. *Finite Elements in Analysis and Design*, 102 :1–6.

[Hauret and Le Tallec, 2006] Hauret, P. and Le Tallec, P. (2006). Energy-controlling time integration methods for nonlinear elastodynamics and low-velocity impact. *Comput. Methods Appl. Mech. Eng.*, 195(37) :4890–4916.

[Hildebrandt et al., 2006] Hildebrandt, K., Polthier, K., and Wardetzky, M. (2006). On the convergence of metric and geometric properties of polyhedral surfaces. *Geometriae Dedicata*, 123(1) :89–112.

[Hoover et al., 1974] Hoover, W., Ashurst, W., and Olness, R. (1974). Two-dimensional computer studies of crystal stability and fluid viscosity. *The Journal of Chemical Physics*, 60(10) :4043–4047.

[Hughes et al., 1978] Hughes, T. J. R., Liu, W. K., and Caughy, P. (1978). Transient finite element formulations that preserve energy. *Journal of Applied Mechanics*, 45 :366–370.

[Hussain et al., 1974] Hussain, M. A., Pu, S. L., and Underwood, J. (1974). Strain energy release rate for a crack under combined mode I and mode II. In *Fracture analysis : Proceedings of the 1973 national symposium on fracture mechanics, part II*. ASTM International.

[Hussein et al., 2018] Hussein, A., Hudobivnik, B., Aldakheel, F., Wriggers, P., Guidault, P.-A., and Allix, O. (2018). A virtual element method for crack propagation. *PAMM*, 18(1) :e201800104.

[Iserles et al., 2000] Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P., and Zanna, A. (2000). Lie-group methods. *Acta Numerica*, 9 :215–365.

[Jebahi et al., 2015] Jebahi, M., André, D., Terreros, I., and Iordanoff, I. (2015). *Discrete element method to model 3D continuous materials*. John Wiley & Sons.

[Johnson and Cook, 1983] Johnson, G. and Cook, W. (1983). A constitutive model and data for metals subjected to large strains, strain rates, and high pressures. In *Proceedings of the 7th International Symposium On Ballistics*.

[Kane et al., 1999a] Kane, C., Marsden, J. E., and Ortiz, M. (1999a). Symplectic-energy-momentum preserving variational integrators. *Journal of Mathematical Physics*, 40(7) :3353–3371.

[Kane et al., 2000] Kane, C., Marsden, J. E., Ortiz, M., and West, M. (2000). Variational integrators and the newmark algorithm for conservative and dissipative mechanical systems. *International Journal for Numerical Methods in Engineering*, 49(10) :1295–1325.

[Kane et al., 1999b] Kane, C., Repetto, E. A., Ortiz, M., and Marsden, J. E. (1999b). Finite element analysis of nonsmooth contact. *Computer Methods in Applied Mechanics and Engineering*, 180(1) :1–26.

[Krysl and Endres, 2005] Krysl, P. and Endres, L. (2005). Explicit Newmark/Verlet algorithm for time integration of the rotational dynamics of rigid bodies. *International Journal for Numerical Methods in Engineering*, 62(15) :2154–2177.

[Kuna, 2013] Kuna, M. (2013). *Finite elements in fracture mechanics*. Springer.

[Labra and Oñate, 2009] Labra, C. and Oñate, E. (2009). High-density sphere packing for discrete element method simulations. *Communications in Numerical Methods in Engineering with Biomedical Applications*, 25(7) :837–849.

[Leyendecker et al., 2008] Leyendecker, S., Marsden, J. E., and Ortiz, M. (2008). Variational integrators for constrained dynamical systems. *Journal of applied Mathematics and Mechanics*, 88(9) :677–708.

[Li and Marigo, 2017] Li, T. and Marigo, J.-J. (2017). Crack tip equation of motion in dynamic gradient damage models. *Journal of Elasticity*, 127(1) :25–57.

[Li et al., 2016] Li, T., Marigo, J.-J., Guilbaud, D., and Potapov, S. (2016). Numerical investigation of dynamic brittle fracture via gradient damage models. *Advanced Modeling and Simulation in Engineering Sciences*, 3(1) :26.

[Marazzato et al., 2019a] Marazzato, F., Ern, A., Mariotti, C., and Monasse, L. (2019a). An explicit pseudo-energy conserving time-integration scheme for Hamiltonian dynamics. *Computer Methods in Applied Mechanics and Engineering*, 347 :906 – 927.

[Marazzato et al., 2019b] Marazzato, F., Ern, A., and Monasse, L. (2019b). A consistent discrete element method for quasi-static and dynamic elasto-plasticity. `https://hal.archives-ouvertes.fr/hal-02343280`.

[Marazzato et al., 2019c] Marazzato, F., Ern, A., Monasse, L., and Sab, K. (2019c). A discrete element method recast as a lowest-order discontinuous Galerkin method and applications to elasto-plastic solid dynamics. In *CFM 2019 24e Congrès Français de Mécanique*.

[Mariotti, 2007] Mariotti, C. (2007). Lamb's problem with the lattice model Mka3D. *Geophysical Journal International*, 171(2) :857–864.

[Mariotti, 2016] Mariotti, C. (2016). A new leapfrog scheme for rotational motion in 3D. *International Journal for Numerical Methods in Engineering*, 107(4) :273–289.

[Mariotti et al., 2009] Mariotti, C., Michaut, V., and Molinari, J.-F. (2009). Modeling of the fragmentation by discrete element method. In *DYMAT 2009 9th Int. Conf. Mechanical and Physical Behaviour of Materials under Dynamic Loading*, pages 1523–1528.

[Marsden and West, 2001] Marsden, J. E. and West, M. (2001). Discrete mechanics and variational integrators. *Acta Numerica*, 10 :357–514.

[Moës and Belytschko, 2002] Moës, N. and Belytschko, T. (2002). X-FEM, de nouvelles frontières pour les éléments finis. *Revue européenne des Eléments*, 11(2-4) :305–318.

[Monasse and Mariotti, 2012] Monasse, L. and Mariotti, C. (2012). An energy-preserving discrete element method for elastodynamics. *ESAIM : Mathematical Modelling and Numerical Analysis*, 46 :1527–1553.

[Mott, 1947] Mott, N. F. (1947). Fragmentation of shell cases. *Proceedings of the Royal Society of London. Series A. Mathematical and physical sciences*, 189(1018) :300–308.

[Nilsen et al., 2017] Nilsen, H., Larsen, I., and Raynaud, X. (2017). Combining the modified discrete element method with the virtual element method for fracturing of porous media. *Computational Geosciences*, 21(5-6) :1059–1073.

[Noels and Radovitzky, 2008] Noels, L. and Radovitzky, R. (2008). An explicit discontinuous Galerkin method for non-linear solid dynamics : Formulation, parallel implementation and scalability properties. *International Journal for Numerical Methods in Engineering*, 74(9) :1393–1420.

[Notsu and Kimura, 2014] Notsu, H. and Kimura, M. (2014). Symmetry and positive definiteness of the tensor-valued spring constant derived from P1-FEM for the equations of linear elasticity. *Networks & Heterogeneous Media*, 9(4).

[Oñate and Rojek, 2004] Oñate, E. and Rojek, J. (2004). Combination of discrete element and finite element methods for dynamic analysis of geomechanics problems. *Computer methods in applied mechanics and engineering*, 193(27-29) :3087–3128.

[Pu et al., 2018] Pu, Y., Pego, R., Dutykh, D., and Clamond, D. (2018). Weakly singular shock profiles for a non-dispersive regularization of shallow-water equations. *arXiv preprint arXiv :1805.06842*.

[Quispel and McLaren, 2008] Quispel, G. and McLaren, D. I. (2008). A new class of energy-preserving numerical integration methods. *Journal of Physics A : Mathematical and Theoretical*, 41(4) :045206.

[Romon, 2013] Romon, P. (2013). *Introduction à la géométrie différentielle discrète*. Ellipses.

[Rosin and Rammler, 1933] Rosin, P. and Rammler, E. (1933). Regularities in the distribution of cement particles. *Journal of the Institute of Fuel*, 7 :29–33.

[Salomon et al., 2008] Salomon, J., Weiss, A. A., and Wohlmuth, B. (2008). Energy-conserving algorithms for a corotational formulation. *SIAM Journal on Numerical Analysis*, 46(4) :1842–1866.

[Sih, 1974] Sih, G. C. (1974). Strain-energy-density factor applied to mixed mode crack problems. *International Journal of fracture*, 10(3) :305–321.

[Simo and Oliver, 1994] Simo, J. and Oliver, J. (1994). A new approach to the analysis and simulation of strain softening in solids. *Fracture Damage Quasibrittle Struct.*, pages 25–39.

[Simo and Taylor, 1985] Simo, J. C. and Taylor, R. L. (1985). Consistent tangent operators for rate-independent elastoplasticity. *Computer methods in applied mechanics and engineering*, 48(1) :101–118.

[Simon, 1983] Simon, L. (1983). Lectures on geometric measure theory. In *Proceedings of the Centre for Mathematical Analysis, Australian National University*, volume 3. Australian National University Centre for Mathematical Analysis, Canberra.

[Son, 1977] Son, N. Q. (1977). On the elastic plastic initial-boundary value problem and its numerical integration. *International Journal for Numerical Methods in Engineering*, 11(5) :817–832.

[Spellings et al., 2017] Spellings, M., Marson, R. L., Anderson, J. A., and Glotzer, S. C. (2017). GPU accelerated discrete element method (DEM) molecular dynamics for conservative, faceted particle simulations. *Journal of Computational Physics*, 334 :460–467.

[Stone and Babuška, 1998] Stone, T. J. and Babuška, I. (1998). A numerical method with a posteriori error estimation for determining the path taken by a propagating crack. *Computer methods in applied mechanics and engineering*, 160(3-4) :245–271.

[Sukumar et al., 1997] Sukumar, N., Moran, B., Black, T., and Belytschko, T. (1997). An element-free Galerkin method for three-dimensional fracture mechanics. *Computational Mechanics*, 20(1-2) :170–175.

[Wohlmuth, 2011] Wohlmuth, B. (2011). Variationally consistent discretization schemes and numerical algorithms for contact problems. *Acta Numerica*, 20 :569–734.