



**HAL**  
open science

# Méthode des Éléments Virtuels pour le calcul de la déformation mécanique couplée aux écoulements en milieux poreux

Julien Coulet

► **To cite this version:**

Julien Coulet. Méthode des Éléments Virtuels pour le calcul de la déformation mécanique couplée aux écoulements en milieux poreux. Analyse numérique [math.NA]. Sorbonne-Université, 2019. Français. NNT: . tel-02416511v1

**HAL Id: tel-02416511**

**<https://theses.hal.science/tel-02416511v1>**

Submitted on 17 Dec 2019 (v1), last revised 11 Mar 2020 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SORBONNE UNIVERSITÉ

École Doctorale 386 – Sciences Mathématiques de Paris Centre  
Laboratoire Jacques-Louis Lions

---

# Méthode des Éléments Virtuels pour le calcul de la déformation mécanique couplée aux écoulements en milieux poreux

---

Thèse présentée par

**Julien COULET**

En vue de l'obtention du grade de

**Docteur de Sorbonne Université**

Discipline Mathématiques Appliquées

Soutenue publiquement le 13 novembre 2019

Devant un jury composé de :

---

Rapporteur	Grégoire ALLAIRE	Professeur	École Polytechnique
Rapporteur	Jérôme DRONIOU	Professeur	Monash University
Examineur	Frédéric HECHT	Professeur	Sorbonne Université
Directrice de thèse	Vivette GIRAULT	Professeure	Sorbonne Université
Directeur de thèse	Frédéric NATAF	Directeur de recherche	Sorbonne Université
Encadrante	Isabelle FAILLE	Ingénieure de recherche	IFPEN
Encadrant	Nicolas GUY	Ingénieur de recherche	IFPEN

---



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>7</b>
1.1	Contexte et applications . . . . .	8
1.2	Limites de l'approche actuelle et objectifs de la thèse . . . . .	9
1.3	Aperçu de l'existant et directions choisies . . . . .	11
1.4	Travaux effectués et plan du manuscrit . . . . .	13
<b>2</b>	<b>Modélisation : du système d'élasticité aux équations de Biot</b>	<b>15</b>
2.1	Mécanique du solide élastique . . . . .	16
2.1.1	Déformations, contraintes et loi de comportement . . . . .	16
2.1.2	Cas particulier des problèmes plans . . . . .	18
2.1.3	Système d'élasticité linéaire . . . . .	19
2.2	Mécanique des matériaux poroélastiques . . . . .	20
2.2.1	Système d'équations de Biot . . . . .	20
2.2.2	Stabilité de la solution du problème poroélastique . . . . .	22
2.2.3	Quelques problèmes de poroélasticité . . . . .	24
<b>3</b>	<b>Méthode des éléments virtuels pour le problème d'élasticité</b>	<b>29</b>
3.1	Construction de la méthode . . . . .	30
3.1.1	Espace d'approximation . . . . .	30
3.1.2	Formulation variationnelle approchée . . . . .	33
3.1.3	Explicitation de la projection . . . . .	37
3.2	Implémentation . . . . .	39
3.3	Paramétrage et validation numérique . . . . .	42
3.3.1	Choix du coefficient de stabilisation . . . . .	43
3.3.2	Convergence en espace . . . . .	46
3.4	Observations supplémentaires sur la méthode . . . . .	46
3.4.1	Traitement de maillages particuliers . . . . .	48
3.4.2	Comparaison aux éléments finis . . . . .	49
3.4.3	Comportement à la limite incompressible . . . . .	51
<b>4</b>	<b>Discrétisation VEM-TPFA, un premier schéma couplé pour le système poroélastique</b>	<b>57</b>
4.1	Construction du schéma numérique . . . . .	58
4.1.1	Discrétisation du système poroélastique . . . . .	58
4.1.2	Cas du schéma volumes finis 2 points . . . . .	60
4.2	Étude du schéma numérique VEM-TPFA . . . . .	61

4.2.1	Existence et unicité d'une solution discrète . . . . .	61
4.2.2	Majoration de l'erreur d'approximation . . . . .	63
4.2.3	Majoration des erreurs d'interpolation . . . . .	71
4.3	Implémentation et validation numérique . . . . .	73
4.3.1	Implémentation du schéma . . . . .	73
4.3.2	Patch test et limites de l'approximation à deux points . . . . .	76
4.3.3	Validation numérique de l'estimation d'erreur . . . . .	78
4.4	Calculs réalistes de problèmes physiques . . . . .	79
4.4.1	Cas de compaction . . . . .	79
4.4.2	Cas Mandel . . . . .	82
<b>5</b>	<b>Vers des maillages plus généraux avec une discrétisation VEM-MPFA</b>	<b>87</b>
5.1	Motivations . . . . .	88
5.1.1	$\bar{\kappa}$ -orthogonalité . . . . .	88
5.1.2	Maillages géosciences . . . . .	90
5.2	Approximation multipoints pour le schéma volumes finis . . . . .	91
5.2.1	Approximation MPFA pour les flux numériques . . . . .	91
5.2.2	Implémentation et couplage avec les éléments virtuels . . . . .	93
5.2.3	Illustrations numériques . . . . .	94
5.3	Modification des VEMs pour les faces non planes . . . . .	96
5.3.1	Transformation en maillage polyédrique . . . . .	96
5.3.2	Un espace d'approximation modifié pour les éléments virtuels . . . . .	99
5.3.3	Un mot sur les volumes finis . . . . .	101
5.4	Illustration sur une application géosciences . . . . .	102
5.4.1	Paramètres du cas . . . . .	102
5.4.2	Résultats numériques . . . . .	103
<b>6</b>	<b>Au-delà de la discrétisation : stratégies de résolution du système poroélastique</b>	<b>107</b>
6.1	Panorama des différentes méthodes . . . . .	108
6.1.1	Méthode directe ou méthode itérative . . . . .	108
6.1.2	Résolution monolithique ou résolution séquentielle . . . . .	109
6.1.3	Choix de la stratégie . . . . .	110
6.2	Stratégie séquentielle fixed-stress . . . . .	112
6.2.1	Formulation de l'algorithme fixed-stress . . . . .	112
6.2.2	Interprétation de la stratégie fixed-stress comme une méthode de point fixe . . . . .	114
6.2.3	Comparaison des différentes variantes . . . . .	116
6.3	Accélération de la stratégie fixed-stress . . . . .	117
6.3.1	Interprétation de la stratégie fixed-stress comme une méthode de Krylov . . . . .	118
6.3.2	Comparaison numérique des algorithmes . . . . .	120
6.4	Recherche de sous-solveurs efficaces pour les sous-problèmes . . . . .	120
6.4.1	Préconditionneurs ILU et AMG . . . . .	120
6.4.2	Méthode de décomposition de domaine pour l'élasticité . . . . .	123
6.5	Utilisation des sous-solveurs spécialisés parallèles dans le cadre de la stratégie fixed-stress	126
6.5.1	Utilisation de la librairie hpddm pour la décomposition de domaine . . . . .	126
6.5.2	Comportement sur le cas synthétique hétérogène . . . . .	129
6.5.3	Application au problème tridimensionnel réaliste . . . . .	131

<b>7 Conclusion et perspectives</b>	<b>133</b>
7.1 Travaux effectués . . . . .	133
7.2 Perspectives . . . . .	134
7.3 Réalisations et publications . . . . .	135
<b>Annexe : Poster</b>	<b>137</b>
<b>Bibliographie</b>	<b>139</b>



# 1

## Introduction

À l'horizon 2020, résoudre un problème de physique à des fins de compréhension du monde, de prédiction ou d'optimisation requiert une démarche pluridisciplinaire dont une proposition de décomposition pourrait être la suivante.

Une *théorie*, en physique, est la première marche vers la compréhension de phénomènes naturels. Ensemble d'hypothèses fondamentales et de propriétés déduites de ces hypothèses, elle permet de décrire et de systématiser un domaine physique. En physique moderne, les fondements d'une théorie proviennent généralement d'expériences de pensée, et donc de l'imagination humaine. Une théorie fournit une lecture de la nature, valide sous certaines conditions et précise jusqu'à un certain point, qui se doit d'être cohérente avec les observations expérimentales sous peine d'être invalidée. Selon le cadre d'application, des hypothèses simplificatrices y sont adjointes : nombre et dimension des variables étudiées, relations entre ces variables ou valeurs de grandeurs caractéristiques, ce qui conduit à la définition d'un *modèle*. Second niveau vers la description du réel, le modèle se veut plus spécifique et plus simplifié que la théorie. Il est intimement lié à l'expérience qui permet d'en calibrer les paramètres. Hormis quelques situations simples, les équations posées par le modèle ne sont généralement pas solubles par une méthode analytique. Au mieux, quelques indications peuvent être obtenues sur l'existence d'une solution, son unicité ou encore son comportement asymptotique. Une troisième étape consiste alors à discrétiser le problème, c'est à dire à trouver une *méthode numérique* permettant de passer du problème continu, dont la solution continue est de dimension infinie, à un problème approché discret dont la solution discrète est de taille finie. La méthode numérique doit formuler le problème approché et assurer l'existence et la calculabilité d'une solution discrète à ce problème. Il est évidemment souhaitable que cette solution discrète soit semblable à la solution continue, et donc que l'on dispose d'une estimation de l'écart entre ces deux solutions. Le *calcul* de la solution discrète s'effectue au moyen d'algorithmes qui transcrivent la méthode numérique en opérations concrètes. Il va de soi que ces algorithmes sont aujourd'hui exécutés par des machines (ordinateurs, calculateurs), dont la puissance potentielle n'a cessé de croître ces dernières décennies. Cette quatrième étape nécessaire à la résolution d'un problème physique correspond donc à l'adaptation des algorithmes aux ressources disponibles tout en assurant précision, rapidité et efficacité énergétique des calculs.

Physique théorique, modélisation, méthodes numériques et calcul scientifique font aujourd'hui encore l'objet de nombreux travaux de recherche. Communautés liées mais distinctes, les acteurs d'une thématique n'interviennent qu'occasionnellement dans une autre, et ce d'autant plus que la décomposition ci-dessus occulte quantité de sous-catégories dans chacune des disciplines. Ces travaux, initialement orga-



nisés autour d’une méthode numérique, ont par opportunisme et par goût inclus certaines problématiques rattachées au calcul scientifique. Sans prétendre en maîtriser l’une ou l’autre, on exposera donc dans ce manuscrit des recherches effectuées dans les deux disciplines. De manière plus large, souhaitons que, par les précisions apportées dans chacune des quatre catégories, ce document demeure accessible à toute personne intéressée par l’univers de la simulation numérique.

## Sommaire

---

<b>1.1</b>	<b>Contexte et applications</b> . . . . .	<b>8</b>
<b>1.2</b>	<b>Limites de l’approche actuelle et objectifs de la thèse</b> . . . . .	<b>9</b>
<b>1.3</b>	<b>Aperçu de l’existant et directions choisies</b> . . . . .	<b>11</b>
<b>1.4</b>	<b>Travaux effectués et plan du manuscrit</b> . . . . .	<b>13</b>

---

### 1.1 Contexte et applications

En modélisation mathématique, on désigne par géomécanique la théorie mathématique décrivant le comportement mécanique des sous-sols. Couplée aux données ou modèles issus de la géologie et de la géophysique, elle est indispensable dans les domaines techniques où une représentation fine des sous-sols est requise : prospection d’hydrocarbures, exploitation de l’énergie géothermique ou stockage géologique du dioxyde de carbone en sont quelques exemples. Bien que les échelles de temps et d’espace utilisées soit sans commune mesure selon l’application considérée – retracer l’évolution naturelle d’un bassin sédimentaire implique de s’intéresser à plusieurs centaines de kilomètres carrés pendant des dizaines ou centaines de millions d’années, tandis qu’étudier l’impact d’une installation humaine comme un puits extracteur nécessite de se focaliser sur quelques kilomètres carrés et des durées de l’ordre de quelques dizaines d’années – les modèles mathématiques rendent compte de ces diverses situations. Du point de vue de la géomécanique, un sous-sol est un milieu poreux constitué d’une structure solide et d’un ou plusieurs fluides pouvant se déplacer plus ou moins rapidement au sein de celle-ci. Originellement et par soucis de simplification, les déformations mécaniques de la structure et les écoulements de fluides ont longtemps été considérés comme deux processus distincts et indépendants. Si dans certains cas cette hypothèse est sans conséquences sur la qualité du modèle obtenu, elle constitue dans d’autres une approximation trop grossière qui entraîne une mauvaise représentation de la réalité [Zob07]. L’introduction de nouveaux modèles dits poromécaniques, d’abord par l’approche de Terzaghi [Ter43] puis par celle de Biot [Bio41], a permis de tenir compte du caractère couplé de ces deux problèmes et ainsi d’améliorer la qualité de la modélisation, par exemple en tenant compte de l’effet des contraintes tectoniques pour la modélisation de bassins ou l’effet de subsidence pour la modélisation de réservoirs.

Simultanément, le changement de modèle a ouvert la question du choix de la méthode numérique à adopter. Le modèle faisant intervenir deux équations, chacune d’elle peut être discrétisée par une méthode numérique différente. Plusieurs combinaisons ont ainsi été proposées. Historiquement, il semble qu’une des premières combinaisons consistait à utiliser pour chaque dimension physique du problème couplé la méthode numérique la plus répandue pour la physique correspondante. Ainsi, des méthodes de types éléments finis de Galerkin continu (CG) servaient à discrétiser la partie mécanique tandis que des méthodes de type volumes finis (FV) étaient utilisées pour discrétiser la partie écoulement [SM94]. Parmi les autres possibilités explorées, certaines placent également la partie écoulement dans un contexte éléments finis : c’est le cas des approches (CG-CG) [ML94], de leur analogue non conformes [HRGZ17] ou de celles utilisant des méthodes de Galerkin discontinues (DG-DG) [RTT17]. Une variante conservative de l’approxi-

mation CG, nommée Galerkin enrichie, est utilisée dans [CL18] pour une discrétisation (CG-EG). Toujours dans un contexte éléments finis, il existe également de nombreuses approches utilisant une formulation à trois champs où une variable supplémentaire est discrétisée dans l'équation fluide, ce qui permet d'utiliser des éléments finis mixtes dans des combinaisons (CG-Mixed) [PW07; BBKT17], voire des formulations à quatre champs pour des combinaisons (Mixed-Mixed) [Lee16]. À l'autre extrémité, plusieurs combinaisons peuvent être obtenues en s'inspirant des méthodes volumes finis pour discrétiser les équations issues de la mécanique : c'est par exemple le cas de l'approche (MPSA-FV) [Nor14], qui adapte un schéma de type volumes finis multipoints à l'équation d'élasticité (schéma multipoints d'approximation des contraintes, MPSA), ou de l'approche (HMM-HFV) [Lem13] qui discrétise la mécanique par une méthode de type hybride mimétique mixte (HMM) inspirée des volumes finis hybrides (HFV).

Si ces différentes possibilités coexistent à l'état de recherche, le choix est bien plus restreint dans un contexte industriel. À ce jour, il n'existe pas de code commercial de référence traitant le problème de poromécanique couplé dans le domaine des géosciences. Les simulateurs d'écoulement embarquent généralement un module de mécanique simplifiée pour les cas simples ; pour les situations plus complexes, les ingénieur-e-s font appel à deux codes de simulations distincts, chacun spécialisé dans sa physique et ses méthodes numériques. À titre d'exemple à IFPEN, la partie mécanique repose sur le logiciel commercial Abaqus ou sur le code open-source Code\_Aster tandis que la partie écoulement fait appel au simulateur PumaFlow développé en interne. C'est donc en arrière plan la discrétisation historique (CG-FV) qui est utilisée. Cette approche permet d'utiliser deux logiciels très riches dans leur modélisation, et est relativement économe en coût de développement. Toutefois, elle souffre de deux inconvénients majeurs.

## 1.2 Limites de l'approche actuelle et objectifs de la thèse

### Limites

La première contrainte réside en l'utilisation de deux modules de calcul différents. En toute logique, l'usage de deux codes de simulation distincts ne permet pas de rassembler efficacement les inconnues en pression et en déplacement dans un unique système linéaire et de le résoudre pour obtenir la solution, puisque chaque module a été conçu pour résoudre son propre problème uniquement. On utilise donc des techniques dites séquentielles [DGSM06], qui seront développées dans un prochain chapitre, et qui permettent de découpler les deux sous-problèmes pour les résoudre successivement jusqu'à convergence de la solution vers celle du système couplé. Or, utiliser une telle approche plutôt qu'une résolution entièrement couplée devrait être un choix stratégique et réfléchi, et non résulter d'une contrainte technologique. De fait, dans cette dernière situation, le choix de la méthode de résolution itérative et de ses paramètres relève d'avantage de l'expérience et l'intuition de l'ingénieur-e que des travaux mathématiques sur les méthodes. Enfin, utiliser deux codes indépendants peut entraîner ou une duplication des données, qui doivent être échangées entre les deux logiciels, ou une perte d'informations lorsque l'un des deux codes ne gère pas toutes les grandeurs physiques de l'autre. Cette chaîne de simulation est donc moins souple et moins performante qu'un unique code qui effectuerait une résolution séquentielle en gérant lui-même les deux sous-problèmes, et qui pourrait de plus proposer l'alternative de la résolution entièrement couplée.

La seconde difficulté, indépendante de la précédente mais en réalité la plus problématique, réside dans les contraintes qu'imposent les méthodes de discrétisation sur les maillages géométriques. L'utilisation de deux méthodes différentes requiert ou bien un maillage compatible avec chacune des deux méthodes, ou bien deux maillages distincts complétés d'opérateurs d'interpolation entre ceux-ci. Or, dans le cas plus spécifique des problématiques géomécaniques, c'est-à-dire dans le cadre applicatif de ces travaux, il est

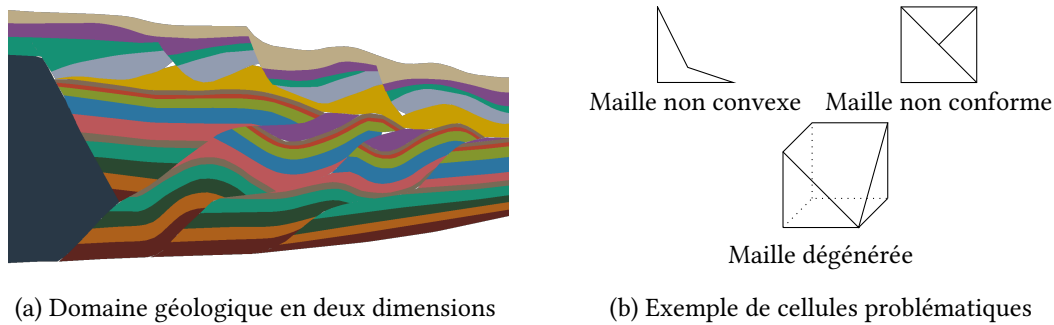


FIG. 1.1 : **Difficultés géométriques** : représentation d'un domaine géologique présentant des discontinuités (a) et schématisation de mailles incompatibles avec la méthode des éléments finis (b).

particulièrement ardu de générer des maillages compatibles avec la discrétisation éléments finis CG tout en conservant un nombre d'éléments raisonnable. En effet, les domaines spatiaux considérés se caractérisent par des discontinuités géométriques telles que des hétérogénéités, des failles géologiques ou encore des rapports entre les différentes dimensions défavorables, voir figure 1.1a. Les outils de maillage utilisés aujourd'hui, conçus à l'origine pour des problèmes d'écoulement et basés sur des hexaèdres, tendent à produire des cellules non convexes, non conformes ou dégénérées (c'est-à-dire dont plusieurs sommets sont confondus) représentés sur la figure 1.1b. Dans le cas tridimensionnel, il est parfois impossible d'obtenir des maillages basés sur des hexaèdres même en s'autorisant ce type de cellules, en particulier lorsque des failles sont présentes dans le domaine, d'où l'intérêt pour des maillages encore plus généraux. Du côté des méthodes volumes finis appliquées aux problèmes d'écoulement, des progrès significatifs sur les schémas ont été réalisés ces dernières années et permettent le traitement de tels types de maillages. La limitation est donc en majeure partie due à l'usage de la discrétisation par éléments finis continus de l'équation d'équilibre mécanique.

## Objectifs

Les équipes de recherche d'IFPEN ont démarré l'élaboration d'une nouvelle génération de codes de simulation développés dans une plateforme logicielle appelée Arcane, co-développée avec le CEA-DAM. Cette nouvelle génération intègre déjà certains schémas volumes finis capables de discrétiser l'équation d'écoulement sur des maillages généraux, telles que la classe des schémas multipoints (MPFA, [Aavo2]) ou celle des schémas hybrides (HFV, [EGHo9]). En revanche, la plateforme ne propose à ce jour aucune méthode de discrétisation ni pour l'équation de mécanique, ni *a fortiori* pour le système poromécanique qui soit utilisé dans un contexte industriel. Ceci constitue l'occasion de créer en interne un module de calcul des déformations mécaniques dans cette plateforme de manière à lever les deux limites soulevées au paragraphe précédent. L'utilisation d'une plateforme unifiée permet en effet une grande liberté dans le choix de la méthode de résolution des deux équations (entièrement couplée ou séquentielle). De plus, repartir de zéro pour les calculs de déformations mécaniques permet d'étudier et utiliser une méthode numérique adaptée aux maillages généraux. Concrètement, le produit attendu est un module de calcul pour la plateforme Arcane utilisant une nouvelle méthode de discrétisation pour l'équation de mécanique, couplée à une discrétisation de type volumes finis existante pour l'équation d'écoulement. Idéalement, ce simulateur doit permettre d'effectuer des calculs performants (précis, robustes et rapides) sur des cas représentatifs des applications visées par les ingénieurs.

## 1.3 Aperçu de l'existant et directions choisies

### Méthodes polyédriques

Concernant l'équation de mécanique et les contraintes liées aux formes des éléments, il faut choisir une méthode de discrétisation fonctionnant sur des maillages plus généraux. Il existe justement un certain nombre de méthodes, basées sur l'approximation de Galerkin, pour le traitement de problèmes elliptiques sur des maillages polyédriques. Une liste de ces méthodes peut par exemple être trouvée dans les préfaces [BBM14 ; BE16] d'éditions spéciales consacrées à ce sujet. Sans entrer dans les détails, on mentionne l'extension des fonctions de bases des éléments finis aux maillages polyédriques par le moyen de fonctions de base rationnelles, [Wac10], barycentriques [SM06] ou harmoniques [Bis14]. Les méthodes Galerkin discontinue [DE12], Galerkin discontinue hybride [CGL09] ou encore *Hybrid High Order* [DEL16] (certaines connections existant entre ces différentes méthodes) constituent d'autres techniques d'approximation sur les éléments généraux. Pourrait également être compatibles la méthode de Galerkin faible [WY14] ou celle des éléments finis *smoothed* [LDN07]. Un autre point de vue est à l'origine des méthodes appelées différences finies mimétiques [BLM14] qui, interprétées dans un cadre variationnel, ont donné naissance aux méthodes des éléments virtuels.

### Méthode des éléments virtuels

Bien que toutes les méthodes listées précédemment auraient pu convenir, et qu'il n'existe *a priori* pas de méthode miraculeuse, c'est la catégorie des éléments virtuels qui a été retenue. Plusieurs arguments en leur faveur, sans forcément leur être exclusifs, sont à l'origine de ce choix.

En premier lieu, le paradigme sur lequel elles reposent est naturel et séduisant : les fonctions de base additionnelles nécessaires sur les éléments non simpliciaux ne sont jamais calculées explicitement, la connaissance des degrés de liberté suffisant à elle-seule pour définir la méthode. Cette caractéristique contraste avec les fonctions non polynomiales, généralement lourdes à définir et à manipuler, utilisées pour les méthodes éléments finis polyédriques. D'impressionnants exemples de maillages en deux [PG15] ou trois [CBP17] dimensions compatibles avec ces méthodes peuvent être trouvés dans la littérature. De plus, l'utilisation d'un espace d'approximation continu a pour conséquence de travailler avec un nombre de degrés de liberté inférieur à celui des méthodes de type Galerkin discontinues, et présage donc un coût de calcul moindre à précision équivalente. Pour une première lecture sur l'implémentation effective de la méthode pour l'équation de Poisson, voir [BBMR14 ; Sut17].

En second lieu, les méthodes des éléments virtuels ont hérité d'un cadre théorique particulièrement avancé de la part des différences finies mimétiques. Introduites en 2013, elle ont ensuite été étendues à des espaces d'approximation non conformes [ALM16] ou à des espaces conformes plus généraux [BBMR16a], tout en utilisant des stratégies pour réduire le nombre de degrés de libertés [BBMR16b]. Plusieurs travaux ont été menés dans le but d'accroître la précision (l'ordre) de ces méthodes [BDR17] ou leur stabilité [CMRS15 ; BB17a], ou bien dans celui de les adapter à des maillages courbes [BRV19]. Par ailleurs, on trouve plusieurs analyses de stabilité et d'erreur *a priori* telles que [BLR17 ; BQL17], certaines même valides en présence de petites arêtes ou faces [BS18]. De manière plus récente, des estimations d'erreur *a posteriori* ont également vu le jour [CGPS17 ; BB17b]. Celles-ci peuvent éventuellement se baser sur de nouvelles méthodes de reconstruction du gradient, plus précises que la formulation originale [CBP19 ; AMLP19]. On trouve également des formulations de ces méthodes dans le cadre d'une approche multi-échelle [RBLE19 ; LJ19].

En troisième lieu, le domaine d'application de ces méthodes est en plein essor et les résultats récents

semblent très prometteurs au sujet de la robustesse vis-à-vis des déformations du maillage. Initialement illustrées sur le problème de Laplace [Bei+13], elles ont ensuite été étendues à une grande variété de problèmes, parmi lesquels l'élasticité linéaire [BBM13; GTP14; ABL17a; ZF18] et non linéaire [BLM15; ABL17b; HAW19] ou les déformations finies [CBP17]. Notons également leur usage dans le cadre de l'optimisation topologique [GPDM15], la géomécanique [ANR17], les problèmes de contacts [WRR16], d'endommagement [BWHZ18], de propagation de fractures [Hus+19], de torsion de plaques [BM13; Chi18], ou encore dans le cadre de l'étude de matériaux composites [Art18; AMS18]. Cette liste, dont les éléments ont été choisis pour leur lien avec la mécanique, est loin d'être exhaustive. Pour conclure sur l'engouement pour ces méthodes dans le cadre de la mécanique, citons leur implémentation récente dans le code de simulation Abaqus [DN19].

### Résolution efficace du système poroélastique

Par rapport au cas où deux codes de calculs distincts sont utilisés, le développement d'un schéma numérique couplé fonctionnant au sein d'un unique simulateur permet *a minima* d'économiser de coûteux échanges de données entre les programmes, ce qui simplifie et accélère la simulation. Le choix de la méthode de résolution du système linéaire obtenu après discrétisation reste toutefois capital, puisque plus de 80% du temps de calcul est consacré à celle-ci. De plus, depuis la dernière décennie, la fréquence de base des processeurs tend à stagner (en raison de problèmes de dissipation thermique), si bien que l'un des seuls moyens d'accélérer un algorithme consiste à l'adapter pour l'exécuter en parallèle sur plusieurs processeurs. Or, le système issu de la discrétisation du système poroélastique possède une structure le rendant particulièrement mal conditionné [FGT01], c'est à dire numériquement difficile à résoudre par des méthodes itératives, en particulier en parallèle. Plusieurs travaux ont visé à construire une méthode efficace de résolution pour ce système par des méthodes de Krylov (voir par exemple [CPL01; GPF03; CTP06; HOL12; BM12]) ou multigrilles [LRGO15]. Récemment, des connexions ont été établies [WCT16] entre certaines de ces méthodes et la méthode de résolution séquentielle fixed-stress split, qui est alors indirectement utilisée pour construire un préconditionneur pour le système couplé. À l'origine, la méthode fixed-stress split est l'une des quatre méthodes séquentielles (avec les méthodes *fixed-strain split*, *undrained split* et *drained split*, cf [MW13]) permettant de découpler les deux sous-problèmes, introduites dans le cadre de la simulation de réservoir [SM94] et largement étudiées depuis. La méthode fixed-stress, du fait qu'elle soit inconditionnellement stable contrairement aux méthodes fixed-strain split et drained split [KTJ11a; KTJ11b] tout en ayant une convergence potentiellement plus rapide que la méthode undrained split [MW13], a fait l'objet de nombreuses études. En plus des travaux sur la convergence et la stabilité de cette méthode, on peut citer par exemple son extension à l'élasticité non linéaire [BRKN18], celle à une approche multi-échelle [DGW18] ou encore celle au cas où le pas de temps diffère selon l'équation résolue [Alm+16]. La méthode fixed-stress semble donc être un moyen efficace d'améliorer la résolution du système linéaire, qu'elle soit appliquée directement ou en tant que préconditionneur du système couplé. Dans les deux cas, la difficulté est en réalité reportée sur la résolution des sous-problèmes. Or, les matrices associées aux problèmes découplés de déformation mécanique et d'écoulement ont un meilleur conditionnement que la matrice couplée, et devraient donc être plus faciles à résoudre. On peut de plus de cette manière capitaliser sur des solveurs efficaces et robustes développés pour ces problèmes dans un contexte plus large que celui de la poromécanique. Pour une revue de ces solveurs, voir par exemple [Ben02]. Dans le cadre de la résolution de systèmes linéaires et pour rebondir sur le développement actif des méthodes des éléments virtuels, on relève l'usage récent pour celles-ci de méthodes de décomposition de domaine (telles que celles de Schwarz [Cal19] ou bien l'approche *finite element tearing and interconnecting* [BPP17]), ainsi que l'investigation de méthodes multigrilles [AMV18].

## 1.4 Travaux effectués et plan du manuscrit

Deux grandes thématiques, de volume certes différent, se dessinent au fil de ce manuscrit de thèse. Elles ont en commun d'étudier, sous un angle ou un autre, le système d'équations modélisant le comportement mécanique de milieux poreux. La première de ces deux parties, constituée des chapitres 3 à 5, aborde le problème par le prisme des méthodes numériques, qui permettent de passer d'équations continues à un système linéaire discret. La seconde de ces deux parties, située en aval dans la chaîne de simulation et retranscrite dans le chapitre 6, s'intéresse à l'amélioration de la résolution de ce système et est donc plus proche des thématiques liées au calcul scientifique.

**Chapitre 2** Le chapitre 2 pose les bases du problème mécanique modélisé. Une première partie, dédiée à la mécanique du solide, décrit l'origine et les hypothèses simplificatrices qui mènent au système d'élasticité linéaire. C'est ce modèle de base de la mécanique des solides déformables qui sera utilisé au long de ces travaux pour décrire le comportement des solides. Dans la seconde partie, ce système est modifié de manière à tenir compte des effets de couplage entre le solide et le fluide : c'est le modèle poroélastique. Après avoir décrit les différentes quantités physiques intervenant dans ce système, et rappelé que celui-ci admet une unique solution continue, on démontre la stabilité de cette solution. La fin de cette partie expose deux problèmes analytiques de poroélasticité, dont la solution exacte est connue, et qui seront utilisés à des fins d'illustration et de validation numérique des codes développés.

**Chapitre 3** La méthode des éléments virtuels est introduite dans le chapitre 3, où seul le système d'élasticité est considéré. Ceci permet de construire pour ce problème la discrétisation par la méthode des éléments virtuels, à l'ordre d'approximation polynomial un, pour des maillages constitués de polygones (en dimension deux) ou de polyèdres (en dimension trois) quelconques. Ce chapitre peut être vu comme une reformulation synthétique des premiers travaux sur la méthode des éléments virtuels appliquée au problème d'élasticité. Après avoir détaillé l'espace d'approximation, la forme bilinéaire approchée ou encore le calcul des projections qui sont spécifiques à cette méthode, des détails fournis sur l'implémentation permettent de développer ce schéma. Les premiers tests numériques démontrent la convergence en espace de cette méthode à l'ordre attendu par la théorie, et illustrent leur capacité à traiter des maillages particulièrement allongés (ici en deux dimensions), sous réserve d'un certain choix pour le coefficient de stabilisation de la méthode, ou à traiter des maillages non conformes. De même, quelques résultats numériques permettent d'espérer un bon comportement sur les matériaux à la limite du domaine compressible, qui posent généralement un problème aux méthodes numériques.

**Chapitre 4** Les trois chapitres suivants constituent les contributions originales de ce manuscrit. Le chapitre 4 propose de réutiliser la méthode des éléments virtuels pour l'élasticité dans le cadre du problème poroélastique. L'équation de conservation de la masse du fluide  $y$  est alors discrétisée par la méthode volumes finis la plus aboutie théoriquement, c'est-à-dire celle utilisant une approximation à deux points pour les flux numériques. Après avoir expliqué comment combiner ces deux méthodes numériques pour créer le schéma couplé VEM-TPFA, une étude mathématique de celui-ci est proposée. L'existence et l'unicité d'une solution discrète est d'abord démontrée, avant de proposer une estimation de l'erreur par rapport à la solution continue du système. En particulier, on montre que le schéma converge à l'ordre 1 en espace et 1 en temps lorsqu'une certaine norme est utilisée pour mesurer la différence. Pour illustrer ce résultat théorique, et après avoir détaillé l'implémentation du schéma, un cas synthétique bidimensionnel permet de retrouver numériquement ces ordres de convergence. Ces résultats théoriques et numériques ont fait l'ob-

jet d'une publication [Cou+19]. À des fins de validation, mais également dans le but d'illustrer l'influence de certains paramètres tels que le choix des conditions initiales, les problèmes physiques réalistes décrits dans le chapitre 2 sont simulés numériquement.

**Chapitre 5** Le chapitre 5 est une extension du chapitre précédent à des maillages encore plus généraux. En effet, si la méthode des volumes finis employée dans le chapitre 4 est maîtrisée théoriquement, la condition d'orthogonalité sur les maillages qui l'accompagne reste trop restrictive pour une utilisation sur des grilles réelles. On cherche donc à remplacer celle-ci par une méthode de volumes finis plus générale. Pour rester proche du cadre de travail déjà mis en place, c'est la classe des méthodes à approximation multipoints du flux numérique qui est retenue. Après avoir détaillé la construction de celles-ci et leur couplage avec la méthode des éléments virtuels, qui donne lieu au schéma VEM-MPFA, on montre numériquement sur le cas synthétique précédemment utilisé que ce choix permet la convergence en espace et en temps à l'ordre 1 sur un échantillon de maillages plus large que pour la méthode VEM-TPFA. Malgré cette amélioration, la méthode n'est pas encore capable de traiter les maillages typiques des géosciences : ceux-ci contiennent effectivement des éléments qui ne sont pas même polyédriques. Une modification est donc apportée à la méthode des éléments virtuels pour gérer ces situations. Ce second schéma couplé pour le système de poroélasticité constitue également une contribution de ces travaux. Ce schéma couplé est illustré sur un maillage réellement issu d'un bassin sédimentaire, en utilisant des données inspirées d'une application réaliste : l'extraction d'un fluide piégé dans un milieu poreux stratifié. Les résultats obtenus sont cohérents avec de la physique du problème.

**Chapitre 6** Dans le chapitre 6, la question de la discrétisation est mise de côté pour se focaliser sur la résolution du système linéaire. Cette étape, qui concentre la majorité du temps de calcul, devient rapidement bloquante lorsque le nombre d'inconnues augmente et se doit d'être optimisée pour conserver un processus global performant. Partant du constat (connu et illustré ici numériquement) que les méthodes les plus classiques sont insuffisamment performantes (pour les méthodes directes) ou insuffisamment robustes (pour les méthodes itératives) sur le système couplé, on s'intéresse à la stratégie séquentielle fixed-stress split qui permet de découpler les deux sous-problèmes et donc de résoudre à la place deux sous-systèmes. Deux axes sont alors étudiés pour rendre cette méthode compétitive. D'une part, on montre que réinterpréter cet algorithme comme un problème d'annulation d'un résidu permet d'en accélérer la convergence à l'aide de méthodes de Krylov. D'autre part, disposant à présent de deux sous-problèmes physiques découplés, on cherche à qualifier pour chacun d'entre eux un sous-solveur robuste. Parmi les méthodes testées, on retient le préconditionneur standard AMG pour la résolution du problème d'écoulement, et une méthode de décomposition de domaine à deux niveaux pour le problème d'élasticité. Cette dernière approche étant moins classique, la construction algébrique de la méthode ainsi que son utilisation au travers de la bibliothèque hpddm sont détaillées. À notre connaissance, une telle combinaison de sous-solveurs est une originalité par rapport aux approches classiques. Ces deux axes sont finalement recoupsés pour la résolution du problème créé dans le chapitre 5 sur le maillage réellement issu d'un bassin sédimentaire, et permettent d'observer la robustesse de la résolution.

# 2

## Modélisation : du système d'élasticité aux équations de Biot

Dans les ouvrages d'analyse des équations aux dérivées partielles, les équations de l'élasticité linéaire sont souvent utilisées à titre d'illustration sans autre forme d'introduction. Inversement, il y a matière dans les œuvres consacrées à la mécanique des milieux continus à dédier un chapitre entier à la description du tenseur des contraintes. Ce chapitre ne remplacera la lecture ni de l'une ni de l'autre de ces théories, mais vise à fournir quelques bases de mécanique des solides de manière à comprendre les modèles manipulés dans la suite de ces travaux. Sa première section, consacrée à l'élasticité, introduit les notions de déformation et de contrainte pour établir le système d'élasticité linéaire. Sa seconde section y ajoute la présence d'un fluide et explique comment celui-ci interagit avec les déformations mécaniques. Après avoir explicité les équations de Biot, qui modélisent cette interaction, elle présente une courte étude mathématique de ce système afin d'introduire au niveau continu les techniques utilisées au niveau discret dans les chapitres suivants. Ce chapitre s'achève par quelques exemples de problèmes poroélastiques dont la solution analytique est connue, qui permettront d'illustrer et de valider les méthodes numériques étudiées.

Dans tout ce qui suit,  $d$  désigne la dimension du problème considéré ( $d = 2$  ou  $d = 3$ ). Le triplet  $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ , ou le doublet  $(\mathbf{e}_x, \mathbf{e}_y)$  le cas échéant, représente une base orthonormée de l'espace  $\mathbb{R}^d$ . Les quantités vectorielles sont notées en gras :  $\mathbf{v}$ , et les matrices sont surlignées d'un trait :  $\overline{M}$ . Les autres notations sont précisées au fil du document.

### Sommaire

---

<b>2.1</b>	<b>Mécanique du solide élastique</b> . . . . .	<b>16</b>
2.1.1	Déformations, contraintes et loi de comportement . . . . .	16
2.1.2	Cas particulier des problèmes plans. . . . .	18
2.1.3	Système d'élasticité linéaire . . . . .	19
<b>2.2</b>	<b>Mécanique des matériaux poroélastiques</b> . . . . .	<b>20</b>
2.2.1	Système d'équations de Biot. . . . .	20
2.2.2	Stabilité de la solution du problème poroélastique. . . . .	22
2.2.3	Quelques problèmes de poroélasticité . . . . .	24

---



## 2.1 Mécanique du solide élastique

Cette première section introduit les quantités d'intérêt et les hypothèses simplificatrices permettant de se ramener de la mécanique des milieux continus au domaine plus restreint de l'élasticité, puis à celui plus restreint encore de l'élasticité linéaire. Ceci permet d'établir le système d'équations de l'élasticité linéaire, dans sa formulation forte issue de la physique et dans sa formulation faible utilisée pour la discrétisation numérique. Les développements sont effectués dans le cas où  $d = 3$ . Le traitement du cas particulier  $d = 2$  fait l'objet d'une remarque spécifique en fin de section 2.1.1.

### 2.1.1 Déformations, contraintes et loi de comportement

#### Déformations et contraintes

La cinématique est l'étude des mouvements indépendamment des causes qui les produisent. Elle fournit plusieurs outils descriptifs permettant de représenter les déplacements de points matériels constituant un solide. À ces fins est tout d'abord défini le champ vectoriel de déplacement  $\mathbf{u}$  qui associe à tout point matériel (dans la configuration de départ) la différence entre sa nouvelle et son ancienne position. Le gradient (par rapport à la configuration initiale) de ce champ de déplacement permet de définir le tenseur  $\bar{E}$  d'ordre 2 selon la relation

$$2\bar{E} = \overline{\nabla\mathbf{u}}^T + \overline{\nabla\mathbf{u}} + \overline{\nabla\mathbf{u}}^T\overline{\nabla\mathbf{u}}.$$

Certains mouvements créent des déformations, c'est-à-dire des modifications d'angles ou de distances entre les points matériels du solide. Pour ces mouvements, les composantes de  $\bar{E}$  donnent facilement accès à certaines informations comme, par exemple, l'extension dans une direction parallèle aux vecteurs de base  $\mathbf{e}_i$ . Au contraire, d'autres mouvements, combinaisons de translations et de rotations, n'entraînent aucune déformation du solide. Ces mouvements sont appelés *mouvements de corps rigide* et le tenseur  $\bar{E}$  qui leur est associé est nul. Pour cette raison, le tenseur  $\bar{E}$  est qualifié de tenseur des déformations, aussi dénommé tenseur de Green-Lagrange. Lorsque l'amplitude des déformations est suffisamment faible – on parle alors de déformations infinitésimales – les termes quadratiques deviennent négligeables. Le tenseur des déformations infinitésimales est alors noté  $\bar{\epsilon}$  et satisfait

$$2\bar{\epsilon} = \overline{\nabla\mathbf{u}}^T + \overline{\nabla\mathbf{u}} \quad (2.1)$$

avec  $\epsilon_{ij} \ll 1$ . Certaines caractéristiques du déplacement peuvent être directement déduites de  $\bar{\epsilon}$ . En particulier, la trace du tenseur des déformations  $\text{tr}(\bar{\epsilon})$  est égale à la divergence du champ de déplacements  $\text{div}(\mathbf{u})$  et s'interprète comme la déformation volumique du solide. Enfin, le tenseur  $\bar{\epsilon}$  étant clairement symétrique, il est fréquent [DVS98] d'adopter la notation de Kelvin  $\hat{\epsilon}$  pour stocker ses composantes indépendantes sous la forme d'un vecteur de  $\mathbb{R}^6$  :  $\hat{\epsilon} = (\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, \sqrt{2}\epsilon_{yz}, \sqrt{2}\epsilon_{xz}, \sqrt{2}\epsilon_{xy})^T$ .

Un autre domaine de la physique consiste à traduire l'action de forces appliquées sur un solide continu, sans s'intéresser aux effets qu'engendrent ces forces. Étant donnée une action mécanique  $\mathbf{f}$  appliquée sur une surface  $dS$  de normale sortante  $\mathbf{n}$  autour d'un point  $\mathbf{M}$ , on définit le vecteur contrainte  $\mathbf{T}$  par la relation  $\mathbf{T}(\mathbf{M}, \mathbf{n}) = \frac{d\mathbf{f}}{dS}$ . On peut montrer que  $\mathbf{T}$  dépend linéairement de  $\mathbf{n}$  et donc définir une application linéaire, appelée tenseur des contraintes de Cauchy, satisfaisant  $\mathbf{T}(\mathbf{M}, \mathbf{n}) = \overline{\sigma^e}(\mathbf{M})\mathbf{n}$ . Ce tenseur représente l'état de contrainte en un point : la composante  $\sigma_{ij}^e$  du tenseur des contraintes correspond à la  $i^e$  composante du vecteur contrainte de direction  $\mathbf{e}_j$ .

À partir du principe fondamental de la dynamique et dans le cadre d'un problème statique, c'est-à-dire en l'absence d'accélération, on démontre [LC90] d'une part l'équation d'équilibre satisfaite par  $\overline{\sigma^e}$  sous

l'effet de forces extérieures  $\mathbf{f}$

$$\mathbf{div}(\overline{\boldsymbol{\sigma}}^e) + \mathbf{f} = \mathbf{0} \quad (2.2)$$

et d'autre part la symétrie du tenseur des contraintes. En utilisant à nouveau la notation de Kelvin, on note  $\hat{\boldsymbol{\sigma}}^e = (\sigma_{xx}^e, \sigma_{yy}^e, \sigma_{zz}^e, \sqrt{2}\sigma_{yz}^e, \sqrt{2}\sigma_{xz}^e, \sqrt{2}\sigma_{xy}^e)^T$ .

### Indicateurs scalaires dérivés des contraintes et déformations

Déformations et contraintes permettent de calculer certaines grandeurs scalaires pouvant présenter un intérêt en mécanique. Du point de vue du développement et de la validation de méthodes, ces grandeurs sont également plus facile à représenter et à utiliser pour des comparaisons que les champs vectoriels. On introduit ainsi la contrainte équivalente de Von Mises qui se calcule à partir du tenseur des contraintes selon la relation

$$\sigma_{\text{VM}}^e = \frac{1}{\sqrt{2}} \sqrt{(\sigma_{xx}^e - \sigma_{yy}^e)^2 + (\sigma_{yy}^e - \sigma_{zz}^e)^2 + (\sigma_{zz}^e - \sigma_{xx}^e)^2 + 6(\sigma_{xy}^e{}^2 + \sigma_{yz}^e{}^2 + \sigma_{zx}^e{}^2)}. \quad (2.3)$$

Cette contrainte équivalente représente l'énergie de distorsion par unité de volume. Pour certains types de matériaux, la limite du domaine élastique est atteinte lorsque  $\sigma_{\text{VM}}^e$  dépasse un seuil critique : le matériau commence alors à se déformer de manière plastique, et ne retrouve pas son état initial lorsque les sollicitations mécaniques cessent. En combinant les tenseurs des déformations et des contraintes, on définit également l'énergie de déformation élastique

$$W^e = \frac{1}{2} \int_{\Omega} \overline{\boldsymbol{\sigma}}^e : \bar{\boldsymbol{\epsilon}}, \quad (2.4)$$

réel positif associé au travail des forces qui ont généré la déformation du solide. L'énergie élastique est une grandeur physique essentielle sur laquelle sont bâtis plusieurs théorèmes (dits énergétiques) qui sortent du cadre de ces travaux. Elle est également utilisée pour construire les lois de comportement non linéaires.

### Loi de comportement

Les deux notions introduites précédemment, déformation et contrainte, peuvent être mises en relation à travers une loi de comportement. Cette loi dépend du matériau considéré, et diffère à la fois par la nature du comportement (élastique, plastique, etc.) et par les constantes mécaniques du matériau. La loi de comportement élastique est la plus simple à modéliser. Elle convient à des solides retrouvant de façon instantanée leur état initial lorsque les efforts extérieurs disparaissent. De plus, d'autres lois de comportement peuvent être approchées par une suite de problèmes élastiques, ce qui justifie l'étude de celle-ci. On parle d'élasticité linéaire lorsque les contraintes sont proportionnelles aux déformations. Cette relation s'écrit

$$\overline{\boldsymbol{\sigma}}^e_{ij} = \overline{\overline{\overline{C}}}_{ijkl} \bar{\boldsymbol{\epsilon}}_{kl} \quad (2.5)$$

où  $\overline{\overline{\overline{C}}}$ , tenseur d'ordre 4, est appelé tenseur des rigidités élastiques. La symétrie de  $\bar{\boldsymbol{\epsilon}}$  et de  $\overline{\boldsymbol{\sigma}}^e$  implique que  $\overline{\overline{\overline{C}}}_{ijkl} = \overline{\overline{\overline{C}}}_{jikl} = \overline{\overline{\overline{C}}}_{jilk}$ . Le tenseur d'ordre 4  $\overline{\overline{\overline{C}}}$  peut alors être réécrit sous la forme d'une matrice de taille  $6 \times 6$  notée  $\hat{C}$ , définie par l'équation (2.6) ci-après. De plus, les lois de la thermodynamique impliquent [LC90] que  $\overline{\overline{\overline{C}}}_{ijkl} = \overline{\overline{\overline{C}}}_{lkji}$  pour une transformation élastique isotherme uniforme, ou, de manière équivalente, que la

matrice  $\hat{C}$  est symétrique, de sorte que

$$\hat{C} = \begin{pmatrix} C_{xxxx} & C_{xxyy} & C_{xxzz} & \sqrt{2}C_{xxyz} & \sqrt{2}C_{xxxz} & \sqrt{2}C_{xxxxy} \\ & C_{yyxy} & C_{yyzz} & \sqrt{2}C_{yyyz} & \sqrt{2}C_{yxxz} & \sqrt{2}C_{yyxy} \\ & & C_{zzzz} & \sqrt{2}C_{zzyz} & \sqrt{2}C_{zzxz} & \sqrt{2}C_{zzxy} \\ & & & 2C_{yzyz} & 2C_{yzxz} & 2C_{yzxy} \\ & & & & 2C_{zxxz} & 2C_{zxxxy} \\ & & & & & 2C_{xyxy} \end{pmatrix}. \quad (2.6)$$

Les coefficients adoptés dans la notation de Kelvin permettent la correspondance entre le produit doublement contracté sur  $\mathbb{R}^{3 \times 3}$  et le produit scalaire sur  $\mathbb{R}^6$  et donc de réécrire :

$$\hat{\sigma}^e = \hat{C}\hat{\epsilon} \quad \text{et} \quad \bar{\sigma}^e : \bar{\epsilon} = \hat{\sigma}^e \cdot \hat{\epsilon}.$$

De plus, le matériau peut présenter certaines relations de symétrie matérielle qui réduisent encore le nombre de composantes indépendantes de  $\bar{C}$ . Le cas le plus dégénéré est celui d'un matériau dit *isotrope* dans lequel tous les axes sont matériellement identiques. Pour ce type de matériaux,  $\hat{C}$  ne dépend pas du repère choisi et ne possède que deux composantes indépendantes. On peut donc exprimer  $\hat{C}$  en fonction du module de Young  $E$  et du coefficient de Poisson  $\nu$  selon

$$\hat{C} = \frac{E}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & \nu & & & \\ \nu & 1-\nu & \nu & & & \\ \nu & \nu & 1-\nu & & & \\ & & & 1-2\nu & & \\ & & & & 1-2\nu & \\ & & & & & 1-2\nu \end{pmatrix} \quad (2.7)$$

ou bien, de manière équivalente, utiliser les coefficients de Lamé  $\lambda$  et  $G$  liés aux deux coefficients précédents par

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}, \quad G = \frac{E}{2(1+\nu)}. \quad (2.8)$$

### 2.1.2 Cas particulier des problèmes plans

Si les problèmes élastiques concernent tous des solides physiques de  $\mathbb{R}^3$ , certains d'entre eux présentent des invariances ou des propriétés géométriques qui permettent de se ramener à un problème plan. Les conséquences sur les tenseurs des déformations, des contraintes et des rigidités dépendent de la configuration considérée et sont explicités ici pour un matériau supposé isotrope.

**Déformations planes** Lorsque tous les plans perpendiculaires à la direction  $\mathbf{e}_z$  sont dans le même état cinématique (en général, c'est le cas d'un domaine de taille infinie en  $z$ ), on parle d'état plan de déformations :  $\epsilon_{xz} = \epsilon_{yz} = \epsilon_{zz} = 0$  et les composantes non nulles de  $\bar{\epsilon}$  ne dépendent pas de  $z$ . Ce cas de figure est le plus simple car il suffit alors de développer le produit matriciel  $\hat{\sigma}^e = \hat{C}\hat{\epsilon}$  en tenant compte des composantes nulles du vecteur  $\hat{\epsilon}$  pour se rendre compte que  $\sigma_{xz}^e = \sigma_{yz}^e = 0$  et réécrire

$$\begin{pmatrix} \sigma_{xx}^e \\ \sigma_{yy}^e \\ \sqrt{2}\sigma_{xy}^e \end{pmatrix} = \frac{E}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & 1-2\nu \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \sqrt{2}\epsilon_{xy} \end{pmatrix} \quad \text{et} \quad \sigma_{zz}^e = \frac{E\nu}{(1+\nu)(1-2\nu)}(\epsilon_{xx} + \epsilon_{yy}). \quad (2.9)$$

**Contraintes planes** Dans d'autres configurations, typiquement lorsque l'objet a une faible dimension dans la direction  $z$ , le solide se trouve en état plan de contraintes :  $\sigma_{xz}^e = \sigma_{yz}^e = \sigma_{zz}^e = 0$  et les composantes non nulles de  $\bar{\sigma}^e$  ne dépendent pas de  $z$ . Dans ce cas, en utilisant à nouveau la loi de comportement (2.5), on montre que  $\epsilon_{xy} = \epsilon_{yz} = 0$  et que

$$\begin{pmatrix} \sigma_{xx}^e \\ \sigma_{yy}^e \\ \sqrt{2}\sigma_{xy}^e \end{pmatrix} = \frac{E}{1-\nu^2} \begin{pmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & 1-\nu \end{pmatrix} \begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \sqrt{2}\epsilon_{xy} \end{pmatrix} \quad \text{et} \quad \epsilon_{zz} = \frac{-\nu}{E}(\sigma_{xx}^e + \sigma_{yy}^e) = \frac{\nu}{\nu-1}(\epsilon_{xx} + \epsilon_{yy}). \quad (2.10)$$

### 2.1.3 Système d'élasticité linéaire

**Formulation forte** Soit  $\Omega$  un domaine ouvert de  $\mathbb{R}^d$  de frontière  $\partial\Omega$ . Le système d'équations de l'élasticité linéaire est obtenu en complétant l'équation d'équilibre local (2.2) et la loi de comportement (2.5) par des conditions aux limites sur la frontière  $\partial\Omega$  de normale sortante  $\mathbf{n}$ . On suppose que cette frontière est séparée en deux sous-ensembles  $\Gamma_{D_d}$  et  $\Gamma_{N_d}$ , de sorte que  $\partial\Omega = \Gamma_{D_d} \cup \Gamma_{N_d}$  et  $\Gamma_{D_d} \cap \Gamma_{N_d} = \emptyset$ , correspondant respectivement aux conditions aux limites de Dirichlet (déplacement imposé) et de Neumann (contrainte imposée). Le problème d'élasticité linéaire s'écrit alors

$$-\operatorname{div}(\bar{\bar{C}}\bar{\epsilon}_u) = \mathbf{f} \quad \text{dans } \Omega, \quad (2.11a)$$

$$\mathbf{u} = \mathbf{g}_D \quad \text{sur } \Gamma_{D_d}, \quad (2.11b)$$

$$\bar{\sigma}^e \mathbf{n} = \mathbf{t}_N \quad \text{sur } \Gamma_{N_d}. \quad (2.11c)$$

**Formulation faible** On introduit à ce stade les notations  $\|\cdot\|_{L^2(\omega)}$  et  $\|\cdot\|_{H^1(\omega)}$  se référant aux normes classiques de Sobolev sur un domaine  $\omega$ . Ainsi, pour  $v \in L^2(\omega)$ ,  $\mathbf{v} \in (L^2(\omega))^d$  et  $\bar{\mathbf{v}} \in (L^2(\omega))^{d \times d}$ ,

$$\|v\|_{L^2(\omega)} = \left( \int_{\omega} v^2 \right)^{\frac{1}{2}}, \quad \|\mathbf{v}\|_{L^2(\omega)} = \left( \int_{\omega} \mathbf{v} \cdot \mathbf{v} \right)^{\frac{1}{2}} \quad \text{et} \quad \|\bar{\mathbf{v}}\|_{L^2(\omega)} = \left( \int_{\omega} \mathbf{v} : \mathbf{v} \right)^{\frac{1}{2}}.$$

Si de plus  $v \in H^1(\omega)$  ou  $\mathbf{v} \in (H^1(\omega))^d$ ,

$$\|v\|_{H^1(\omega)} = \left( \int_{\omega} v^2 + \nabla v \cdot \nabla v \right)^{\frac{1}{2}} \quad \text{et} \quad \|\mathbf{v}\|_{H^1(\omega)} = \left( \int_{\omega} \mathbf{v} \cdot \mathbf{v} + \nabla \mathbf{v} : \nabla \mathbf{v} \right)^{\frac{1}{2}}.$$

Par souci de lisibilité, on pourra noter simplement  $\|\cdot\|_{L^2}$  ou  $\|\cdot\|_{H^1}$  lorsque  $\omega \equiv \Omega$ . La formulation faible en déplacements nécessite l'introduction des espaces

$$V_g = \left\{ \mathbf{v} \in (H^1(\Omega))^d, \mathbf{v} = \mathbf{g}_D \text{ sur } \Gamma_{D_d} \right\}, \quad V_0 = \left\{ \mathbf{v} \in (H^1(\Omega))^d, \mathbf{v} = \mathbf{0} \text{ sur } \Gamma_{D_d} \right\} \quad (2.12)$$

et s'écrit :

$$\text{Trouver } \mathbf{u} \in V_g \text{ tel que } \forall \mathbf{v} \in V_0, \quad a^\Omega(\mathbf{u}, \mathbf{v}) := \int_{\Omega} \bar{\bar{C}}\bar{\epsilon}_u : \bar{\epsilon}_v = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_{N_d}} \mathbf{t}_N \cdot \mathbf{v} \quad (2.13)$$

La forme bilinéaire  $a^\Omega$  est continue ; elle est de plus coercive dès lors que  $\Gamma_{D_d}$  est de mesure non nulle. Ainsi, il existe deux constantes réelles  $M > 0$  et  $\mu > 0$  telles que pour tout  $\mathbf{u}, \mathbf{v}$  dans  $H^1(\Omega)$ ,

$$a^\Omega(\mathbf{u}, \mathbf{v}) \leq M \|\mathbf{u}\|_{H^1(\Omega)} \|\mathbf{v}\|_{H^1(\Omega)} \quad \text{et} \quad \mu \|\mathbf{u}\|_{H^1(\Omega)}^2 \leq a^\Omega(\mathbf{u}, \mathbf{u}).$$

La coercivité de  $a^\Omega$  s'établit grâce à l'inégalité de Korn (2.14) et à la relation  $\|\bar{\epsilon}_v\|_{L^2(\Omega)} \geq \tilde{C}\|\mathbf{v}\|_{H^1(\Omega)}$  pour tout  $\mathbf{v}$  avec  $\tilde{C}$  constante positive.

**Lemme 2.1 (Inégalité de Korn, [All12])** *Il existe une constante  $C > 0$  telle que, pour toute fonction  $\mathbf{v}$  de  $(H^1(\Omega))^d$ ,*

$$\|\mathbf{v}\|_{H^1(\Omega)} \leq C \left( \|\mathbf{v}\|_{L^2(\Omega)}^2 + \|\bar{\epsilon}_v\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}}. \quad (2.14)$$

Dans le cas  $\mathbf{g}_D = \mathbf{0}$ , on montre (voir par exemple [All12, thm 5.3.4]) que pour  $\mathbf{f} \in (L^2(\Omega))^d$  et  $\mathbf{t}_N \in (L^2(\Gamma_{N_d}))^d$ , le problème (2.11) admet une unique solution faible  $\mathbf{u}$  dans l'espace  $V_0$ . Un relèvement de solution permet d'étendre ce résultat au cas général  $\mathbf{g}_D \neq \mathbf{0}$ .

## 2.2 Mécanique des matériaux poroélastiques

Les notions d'élasticité linéaire introduites dans la section précédente servent de base à la description de matériaux plus complexes, dont le comportement élastique est modifié par l'action d'un fluide : les milieux poreux.

### 2.2.1 Système d'équations de Biot

Il existe une certaine structuration de la matière dans laquelle interagissent deux constituants principaux : un squelette solide déformable et un fluide occupant un réseau de *pores*, cavités du solide. À l'échelle macroscopique, un matériau présentant une telle organisation est dénommé milieu poreux. La modélisation de ces milieux poreux requiert la superposition des lois physiques régissant le comportement des deux constituants primaires, de manière à décrire le comportement de l'ensemble par des équations dans lesquelles interviennent des paramètres relatifs à chacun des deux constituants et à la nature de leur organisation. Cette branche particulière de la mécanique est appelée la *poromécanique*, et on doit à Biot [Bio41] les fondations de cette discipline. Dans le cas particulier où le squelette solide est sollicité de manière élastique, on emploie le terme plus spécifique de *poroélasticité*.

Le modèle poroélastique présenté présuppose également que le fluide est *monophasique* (un seul fluide, non réactif, saturant le milieu étudié) et que les déformations sont *quasi statiques* (les effets d'inertie sur la structure solide sont négligeables). Les paramètres physiques décrits dans le tableau 2.1 caractérisent les propriétés du milieu poreux étudié.

Paramètre	Dimension	Unité	Nom
$\phi$	-	-	Porosité
$\alpha$	-	-	Coefficient de Biot
$c_0$	$M^{-1}LT^2$	$\text{Pa}^{-1}$	Coefficient de stockage spécifique
$\rho_f$	$ML^{-3}$	$\text{kg}\cdot\text{m}^{-3}$	Masse volumique du fluide
$\bar{\kappa}$	$M^{-1}L^3T$	$\text{m}^2(\text{Pa}\cdot\text{s})^{-1}$	Matrice de mobilité
$\bar{\bar{C}}$	$ML^{-1}T^{-2}$	Pa	Tenseur des rigidités du solide

TAB. 2.1 : Paramètres physiques décrivant un milieu poroélastique

Dans le cadre du modèle poroélastique, l'interaction entre le squelette solide et le fluide se traduit par l'introduction d'une nouvelle définition de la contrainte. La contrainte totale subie par le matériau est la

somme de la contrainte dite effective  $\overline{\sigma}^e$  présentée précédemment, due aux efforts mécaniques, et d'un terme additionnel lié à la pression  $p$  du fluide présent dans les pores de sorte que

$$\overline{\sigma}(\mathbf{u}, p) = \overline{\sigma}^e(\mathbf{u}) - \alpha p \overline{I}_d.$$

Le coefficient de Biot peut donc être vu comme un quantificateur de la variation de contrainte lorsque la pression du fluide augmente à volume de pores constant, ou, symétriquement, comme un quantificateur de la variation du volume de pores lorsque la pression augmente à contrainte constante. Ce coefficient peut être déterminé à partir de la mesure expérimentale de la compressibilité de la phase solide  $C_s$  et de celle du squelette poreux  $C$  suivant la relation  $\alpha = 1 - \frac{C_s}{C}$ . Toutefois, l'approximation  $\alpha = 1$  est couramment réalisée pour des roches consolidées dans la mesure où leur squelette solide est bien plus compressible que les grains qui le composent. Comme dans le cas élastique simple, la contrainte effective est liée aux déformations par la loi de comportement

$$\overline{\sigma}^e(\mathbf{u}) = \overline{\overline{C}} \overline{\epsilon}_u \text{ avec } \overline{\epsilon}_u = \frac{1}{2} (\overline{\nabla \mathbf{u}} + \overline{\nabla \mathbf{u}}^T)$$

où, dans le cas de l'élasticité isotrope, le tenseur des rigidités  $\overline{\overline{C}}$  se déduit de deux coefficients indépendants selon la forme explicité par l'équation (2.7). L'équation d'équilibre mécanique reste vérifiée avec cette nouvelle définition de la contrainte totale et se réécrit donc

$$-\mathbf{div}(\overline{\sigma}(\mathbf{u}, p)) = \mathbf{f}. \quad (2.15)$$

Par ailleurs, le fluide présent dans les pores a la capacité de se déplacer à une vitesse proportionnelle au déséquilibre hydrostatique selon la loi de Darcy

$$\mathbf{v}_f = -\overline{\kappa} (\nabla p - \rho_f \mathbf{g})$$

où  $\mathbf{g}$  désigne la gravité et  $\rho_f$  la masse volumique du fluide. Le coefficient de proportionnalité est appelé matrice de mobilité, définie comme le quotient de la matrice de perméabilité  $\overline{\overline{K}}$ , qui mesure l'aptitude du milieu poreux à se laisser traverser par un fluide, et de la viscosité dynamique du fluide  $\mu_f$  qui mesure la résistance du fluide à un écoulement laminaire. La seconde équation du système poroélastique s'obtient à partir de l'équation de conservation de la masse du fluide. L'obtention du modèle n'est pas directe (voir par exemple [Mer16] pour les détails) et revient finalement à considérer une équation de conservation dans laquelle la quantité de fluide est donnée par

$$c_0 p + \alpha \mathbf{div}(\mathbf{u}).$$

Du point de vue de cette équation, le coefficient de Biot peut être vu comme un quantificateur de la variation de la quantité de fluide lorsque le volume des pores augmente à pression constante. Le coefficient de stockage spécifique  $c_0 \geq 0$  représente la quantité de fluide pouvant être ajouté dans le milieu lors d'une augmentation de pression à volume solide constant. Ce coefficient dépend des compressibilités du solide  $C_s$  et du fluide  $C_f$  ainsi que de la porosité de la structure  $\phi$ , laquelle caractérise le ratio entre le volume des pores et le volume total du milieu, à travers la relation

$$c_0 = \phi C_f + (\alpha - \phi) C_s.$$

En supposant que le fluide est faiblement compressible, l'équation de conservation de la masse se ramène à l'équation de conservation de volume

$$\partial_t (c_0 p + \alpha \mathbf{div}(\mathbf{u})) + \mathbf{div}(-\overline{\kappa} (\nabla p - \rho_f \mathbf{g})) = q, \quad (2.16)$$

$q$  représentant un terme source volumique.

La combinaison des équations (2.15) et (2.16) définit le système poroélastique linéaire. Soit un milieu poreux élastique linéaire représenté par  $\Omega$  ouvert de  $\mathbb{R}^d$  et un temps  $T > 0$ . Si ce milieu est sujet à une force volumique  $\mathbf{f} : \Omega \times ]0, T] \rightarrow \mathbb{R}^d$  et à un terme source  $q : \Omega \times ]0, T] \rightarrow \mathbb{R}$ , alors le champ de déplacement  $\mathbf{u} : \Omega \times ]0, T] \rightarrow \mathbb{R}^d$  et le champ de pression  $p : \Omega \times ]0, T] \rightarrow \mathbb{R}$  vérifient

$$-\mathbf{div}(\bar{\mathbf{C}}\bar{\boldsymbol{\epsilon}}_u - \alpha p\bar{\mathbf{l}}_d) = \mathbf{f} \quad \text{dans } \Omega \times ]0, T], \quad (2.17a)$$

$$\partial_t(c_0 p + \alpha \operatorname{div}(\mathbf{u})) + \operatorname{div}(-\bar{\kappa}(\nabla p - \rho_f \mathbf{g})) = q \quad \text{dans } \Omega \times ]0, T], \quad (2.17b)$$

complété par des conditions aux limites (avec  $\partial\Omega = \Gamma_{D_d} \cup \Gamma_{N_d}$ ,  $\partial\Omega = \Gamma_{D_p} \cup \Gamma_{N_p}$  et  $\mathbf{n}$  normale sortante)

$$\mathbf{u} = \mathbf{0} \quad \text{sur } \Gamma_{D_d} \times ]0, T], \quad (2.17c)$$

$$(\bar{\boldsymbol{\sigma}}^e - \alpha p\bar{\mathbf{l}}_d) \cdot \mathbf{n} = \mathbf{t}_N \quad \text{sur } \Gamma_{N_d} \times ]0, T], \quad (2.17d)$$

$$p = 0 \quad \text{sur } \Gamma_{D_p} \times ]0, T], \quad (2.17e)$$

$$-\bar{\kappa}(\nabla p - \rho_f \mathbf{g}) \cdot \mathbf{n} = \phi_N \quad \text{sur } \Gamma_{N_p} \times ]0, T], \quad (2.17f)$$

et des conditions initiales

$$p(\mathbf{x}, t = 0) = p^0(\mathbf{x}) \quad \text{dans } \Omega, \quad (2.17g)$$

$$\mathbf{u}(\mathbf{x}, t = 0) \text{ solution de (2.17a) avec } p = p^0 \quad \text{dans } \Omega. \quad (2.17h)$$

### 2.2.2 Stabilité de la solution du problème poroélastique

Des résultats d'existence, d'unicité et de régularité sont développés dans [Shoou]. La démonstration traite le problème poroélastique comme un problème d'évolution dans un espace de Hilbert en faisant appel à des outils d'analyse issus de la théorie des semi-groupes. En supposant que le domaine est suffisamment régulier et que le terme source  $q$  est continu au sens de Hölder, soit  $q \in \mathcal{C}^\alpha([0, T]; L^2)$  avec  $\alpha \in ]0, 1]$ , l'existence et l'unicité d'une solution est assurée.

Afin d'introduire les outils utilisés dans l'étude numérique du schéma section 4.2, la stabilité de la solution du système continu est étudiée formellement ci-dessous. Pour alléger les calculs, les conditions de Neumann et le terme de gravité sont supposés nuls dans cette section ( $\mathbf{t}_N = \mathbf{0}$ ,  $\phi_N = 0$ ,  $\mathbf{g} = \mathbf{0}$ ). Aux côtés des normes en espace introduites dans la section 2.1.3, on utilise pour les normes en temps sur l'intervalle  $[0, T]$  les notations

$$\|v\|_{L^2(0, T)} = \left( \int_0^T v(t)^2 dt \right)^{\frac{1}{2}} \quad \text{et} \quad \|v\|_{L^\infty(0, T)} = \max_{t \in [0, T]} |v(t)|$$

pour  $v \in L^\infty(0, T)$ . Si  $v$  dépend du temps et de l'espace, on notera pour  $v \in L^\infty(0, T; H^1)$

$$\|v\|_{L^2(0, T; H^1)} = \left( \int_0^T \|v(t)\|_{H^1(\Omega)}^2 dt \right)^{\frac{1}{2}} \quad \text{et} \quad \|v\|_{L^\infty(0, T; H^1)} = \max_{t \in [0, T]} \|v(t)\|_{H^1(\Omega)}.$$

On définit de la même manière les normes  $\|v\|_{L^2(0, T; L^2)}$  et  $\|v\|_{L^\infty(0, T; L^2)}$  pour  $v \in L^2(]0, T[ \times \Omega)$ . Enfin, on étend ces notations dans le cas où  $\mathbf{v}$  est à valeurs vectorielles. On démontre de manière formelle le résultat de stabilité suivant sur la solution  $(\mathbf{u}, p)$  :

**Proposition 2.2** *Pour tout  $t \in [0, T]$ , la solution  $(\mathbf{u}, p)$  satisfait*

$$\|\mathbf{u}(t)\|_{H^1(\Omega)}^2 + \|p(t)\|_{L^2(\Omega)}^2 + \int_0^t \int_{\Omega} \kappa \nabla p \cdot \nabla p \leq B \exp(t)$$

où  $B$  est une constante dépendant des données initiales, des seconds membres, de  $c_0$  et de la constante de coercivité de  $a^\Omega$  notée  $\mu$ .

En multipliant l'équation d'élasticité par  $\partial_t \mathbf{u}$ , l'équation de conservation par  $p$  puis en intégrant sur  $\Omega$ ,

$$\begin{aligned} - \int_{\Omega} \mathbf{div} (\overline{\sigma^e} - \alpha p \overline{l_d}) \partial_t \mathbf{u} &= \int_{\Omega} \mathbf{f} \cdot \partial_t \mathbf{u} \\ c_0 \int_{\Omega} p \partial_t p + \int_{\Omega} \alpha \partial_t (\operatorname{div}(\mathbf{u})) p + \int_{\Omega} \kappa \nabla p \cdot \nabla p &= \int_{\Omega} p q, \end{aligned}$$

soit en sommant les deux égalités

$$a^\Omega(\mathbf{u}, \partial_t \mathbf{u}) + c_0 \int_{\Omega} p \partial_t p + \int_{\Omega} \kappa \nabla p \cdot \nabla p = \int_{\Omega} \mathbf{f} \cdot \partial_t \mathbf{u} + \int_{\Omega} p q$$

puisque  $-\int_{\Omega} \mathbf{div} (\overline{\sigma^e} - \alpha p \overline{l_d}) \partial_t \mathbf{u} = a^\Omega(\mathbf{u}, \partial_t \mathbf{u}) - \int_{\Omega} \alpha p \operatorname{div}(\partial_t \mathbf{u})$  d'après la formule de Green, où  $a^\Omega(\cdot, \cdot)$  désigne la forme bilinéaire associée au problème élastique. Pour les deux formes bilinéaires que sont  $a^\Omega$  et le produit scalaire  $L^2$ , on a  $\langle \cdot, \partial_t \cdot \rangle = \frac{1}{2} \partial_t \|\cdot\|^2$  d'où, en intégrant entre 0 et  $t$  pour  $t$  donné dans  $[0, T]$  :

$$\|\mathbf{u}(t)\|_a^2 - \|\mathbf{u}^0\|_a^2 + c_0 \|p(t)\|_{L^2(\Omega)}^2 - c_0 \|p^0\|_{L^2(\Omega)}^2 + 2 \int_0^t \int_{\Omega} \kappa \nabla p \cdot \nabla p = 2 \int_0^t \int_{\Omega} \mathbf{f} \cdot \partial_t \mathbf{u} + 2 \int_0^t \int_{\Omega} p q.$$

Ci-dessus, la notation  $\|\cdot\|_a$  correspond à la norme induite par  $a^\Omega : \|\mathbf{u}\|_a^2 = a^\Omega(\mathbf{u}, \mathbf{u})$ . En notant  $\mu$  la constante de coercivité de  $a^\Omega$  ( $a^\Omega$  étant coercive d'après les inégalités de Korn et de Poincaré),

$$\mu \|\mathbf{u}(t)\|_{H^1(\Omega)}^2 + c_0 \|p(t)\|_{L^2(\Omega)}^2 + 2 \int_0^t \int_{\Omega} \kappa \nabla p \cdot \nabla p \leq 2 \int_0^t \int_{\Omega} \mathbf{f} \cdot \partial_t \mathbf{u} + 2 \int_0^t \int_{\Omega} p q + \mu \|\mathbf{u}^0\|_{H^1(\Omega)}^2 + c_0 \|p^0\|_{L^2(\Omega)}^2.$$

On majore les termes de droite en utilisant les inégalités de Cauchy-Schwarz et de Young : par intégration par parties

$$\int_0^t \int_{\Omega} \mathbf{f}(\mathbf{x}, s) \cdot \partial_t \mathbf{u}(\mathbf{x}, s) dx ds = - \int_0^t \int_{\Omega} \partial_t \mathbf{f}(\mathbf{x}, s) \cdot \mathbf{u}(\mathbf{x}, s) dx ds + \int_{\Omega} \mathbf{f}(\mathbf{x}, t) \cdot \mathbf{u}(\mathbf{x}, t) dx - \int_{\Omega} \mathbf{f}^0(\mathbf{x}) \cdot \mathbf{u}^0(\mathbf{x}) dx$$

d'où

$$\begin{aligned} 2 \left| \int_0^t \int_{\Omega} \mathbf{f} \cdot \partial_t \mathbf{u} \right| &\leq 2 \left| \int_0^t \int_{\Omega} \partial_t \mathbf{f} \cdot \mathbf{u} \right| + 2 \left| \int_{\Omega} \mathbf{f}^0 \cdot \mathbf{u}^0 \right| + 2 \left| \int_{\Omega} \mathbf{f}(t) \cdot \mathbf{u}(t) \right| \\ &\leq 2 \int_0^t \|\hat{\mathbf{f}}\|_{L^2(\Omega)} \|\mathbf{u}\|_{L^2(\Omega)} + 2 \|\mathbf{f}^0\|_{L^2(\Omega)} \|\mathbf{u}^0\|_{L^2(\Omega)} + 2 \|\mathbf{f}(t)\|_{L^2(\Omega)} \|\mathbf{u}(t)\|_{L^2(\Omega)} \\ &\leq 2 \left( \int_0^t \|\hat{\mathbf{f}}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \left( \int_0^t \|\mathbf{u}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} + 2 \|\mathbf{f}\|_{L^\infty(0, T; L^2)} \|\mathbf{u}^0\|_{L^2(\Omega)} + 2 \|\mathbf{f}\|_{L^\infty(0, T; L^2)} \|\mathbf{u}(t)\|_{L^2(\Omega)} \\ &\leq \frac{1}{\epsilon_2} \|\hat{\mathbf{f}}\|_{L^2(0, T; L^2)}^2 + \epsilon_2 \int_0^t \|\mathbf{u}\|_{H^1(\Omega)}^2 + \frac{1 + \epsilon_1}{\epsilon_1} \|\mathbf{f}\|_{L^\infty(0, T; L^2)}^2 + \epsilon_1 \|\mathbf{u}(t)\|_{H^1(\Omega)}^2 + \|\mathbf{u}^0\|_{H^1(\Omega)}^2 \end{aligned}$$



avec  $\epsilon_1 > 0$  et  $\epsilon_2 > 0$  à fixer, et

$$2 \int_0^t \int_{\Omega} pq \leq 2 \int_0^t \|p\|_{L^2} \|q\|_{L^2} \leq 2 \left( \int_0^t \|p\|_{L^2}^2 \right)^{\frac{1}{2}} \left( \int_0^t \|q\|_{L^2}^2 \right)^{\frac{1}{2}} \leq \frac{1}{\epsilon_3} \|q\|_{L^2(0,T;L^2)}^2 + \epsilon_3 \int_0^t \|p\|_{L^2(\Omega)}^2$$

avec  $\epsilon_3 > 0$  à fixer. Ainsi,

$$(\mu - \epsilon_1) \|\mathbf{u}(t)\|_{H^1(\Omega)}^2 + c_0 \|p(t)\|_{L^2(\Omega)}^2 + 2 \int_0^t \int_{\Omega} \kappa \nabla p \cdot \nabla p \leq \int_0^t \left( \epsilon_2 \|\mathbf{u}\|_{H^1(\Omega)}^2 + \epsilon_3 \|p\|_{L^2(\Omega)}^2 \right) + \tilde{B}$$

où l'on a noté

$$\tilde{B} = (1 + \mu) \|\mathbf{u}^0\|_{H^1(\Omega)}^2 + c_0 \|p^0\|_{L^2(\Omega)}^2 + \frac{1 + \epsilon_1}{\epsilon_1} \|\mathbf{f}\|_{L^\infty(0,T;L^2)}^2 + \frac{1}{\epsilon_2} \|\dot{\mathbf{f}}\|_{L^2(0,T;L^2)}^2 + \frac{1}{\epsilon_3} \|q\|_{L^2(0,T;L^2)}^2.$$

Enfin, en fixant  $\epsilon_1 < \mu$ , puis en divisant par  $\chi = \min\{\mu - \epsilon_1, c_0\}$  et en prenant  $\epsilon_2 = \epsilon_3 = \chi$ ,

$$\|\mathbf{u}(t)\|_{H^1(\Omega)}^2 + \|p(t)\|_{L^2(\Omega)}^2 + \int_0^t \int_{\Omega} \kappa \nabla p \cdot \nabla p \leq \int_0^t \left( \|\mathbf{u}\|_{H^1(\Omega)}^2 + \|p\|_{L^2(\Omega)}^2 \right) + B$$

avec

$$B = \frac{1 + \mu}{\chi} \|\mathbf{u}^0\|_{H^1(\Omega)}^2 + \frac{c_0}{\chi} \|p^0\|_{L^2(\Omega)}^2 + \frac{1 + \epsilon_1}{\epsilon_1 \chi} \|\mathbf{f}\|_{L^\infty(0,T;L^2)}^2 + \frac{1}{\chi^2} \|\dot{\mathbf{f}}\|_{L^2(0,T;L^2)}^2 + \frac{1}{\chi^2} \|q\|_{L^2(0,T;L^2)}^2.$$

Le lemme de Grönwall est utilisé pour conclure :

**Lemme 2.3** Soit  $z$  une fonction continue de  $[0, T]$  dans  $\mathbb{R}^+$  et  $a, b \in \mathbb{R}^+$  de sorte que

$$z(t) \leq b + a \int_0^t z(s) ds.$$

Alors

$$z(t) \leq b \exp(at).$$

En l'appliquant avec  $z(t) = \|\mathbf{u}(t)\|_{H^1(\Omega)}^2 + \|p(t)\|_{L^2(\Omega)}^2$ , on obtient la conclusion

$$\|\mathbf{u}(t)\|_{H^1(\Omega)}^2 + \|p(t)\|_{L^2(\Omega)}^2 + \int_0^t \int_{\Omega} \kappa \nabla p \cdot \nabla p \leq B \exp(t).$$

### 2.2.3 Quelques problèmes de poroélasticité

Il existe assez peu de problèmes de poroélasticité dont la solution analytique est connue. Proche des applications géomécaniques envisagées, on trouve plusieurs problèmes de type *consolidation*, phénomène pouvant être défini comme le tassement de sols saturés en eau sous l'application d'un chargement mécanique.

Le premier cas présenté ici est extrait des travaux [OGP17] et représente un bassin sur lequel des sédiments s'accumulent par dépôt au cours du temps. La masse de cette couche de sédiment génère une contrainte de compression sur la surface supérieure du bassin. Cette contrainte est une fonction linéaire du temps, de coefficient constant  $\sigma_a$ , ce qui est à la fois peu orthodoxe mais aussi réaliste puisque l'épaisseur de sédiments augmente au cours du temps. Le matériau est supposé homogène isotrope de sorte que  $\bar{\kappa} = \kappa \bar{I}_2$  et que  $\kappa, E$  et  $\nu$  sont constants. Les valeurs de ces coefficients mécaniques, ainsi que les conditions aux limites,

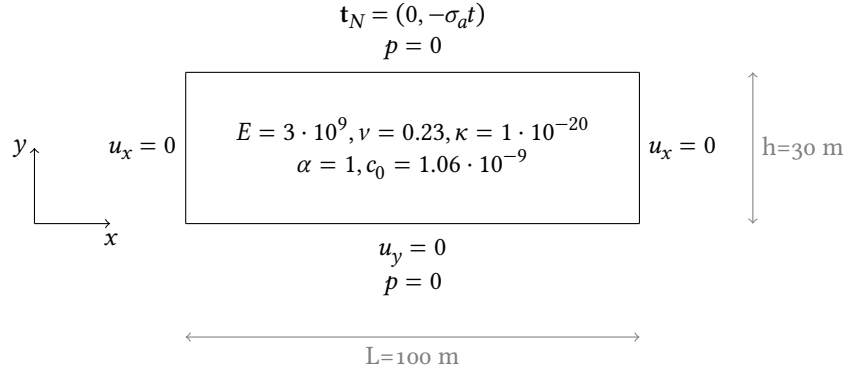


FIG. 2.1 : **Cas de compaction**, constante du matériau et conditions aux limites. L'absence d'indication correspond à une condition de contrainte nulle ou de flux nul.

sont représentées sur la figure 2.1. Enfin, la pression est supposée nulle à l'instant initial. Ces conditions aux limites correspondent à ce que l'on appelle un contexte *œdométrique* : les déplacements verticaux de la base ainsi que les déplacements latéraux des deux côtés sont bloqués. Dans ce contexte, il existe des conditions de symétrie qui permettent de relier certaines composantes du tenseur des contraintes et, après quelques calculs détaillés dans [OGP17], d'aboutir à une équation parabolique mono-dimensionnelle sur la pression. L'analogie entre cette équation et le problème de la chaleur 1D permet d'utiliser des résultats de thermique [CJ59, chap 3.3] pour déterminer l'expression de la pression analytique

$$p(\mathbf{x}, t) = \frac{A_0}{2} \left( (yh - y^2) - 8h^2 \sum_{n=0}^{\infty} \frac{\sin(\frac{\omega_n y}{h})}{\omega_n^3} \exp(-\frac{\omega_n^2}{h^2} \tau_s t) \right) \quad (2.18)$$

avec  $A_0 = \frac{\alpha(1+\nu)\sigma_a}{3\kappa K_d(1-\nu)}$ ,  $\tau_s = \frac{\kappa}{c_0 + \frac{3(1+\nu)\alpha^2}{(1-\nu)K_d}}$  et  $\omega_n = (2n+1)\pi$

où  $K_d$  est le module de compressibilité du solide qui s'exprime en fonction des constantes d'élasticité  $E$  et  $\nu$  selon la relation  $3(1-2\nu)K_d = E$ . Le fait que le problème soit mono-dimensionnel en espace se traduit par le fait que  $p$  ne dépend en réalité que de  $y$  (et du temps). Plus précisément, la pression est constituée d'un terme indépendant du temps  $yh - y^2$  maximal sur l'horizontale médiane et d'un terme d'atténuation sinusoidal qui tend à s'annuler lorsque  $t$  devient grand. Dans le contexte œdométrique, la relation entre la contrainte et les déplacements verticaux se réécrit

$$\epsilon_{yy} = E_{oe}(\sigma_{yy} + \alpha p) \quad \text{avec} \quad E_{oe} = \frac{(1+\nu)(1-2\nu)}{E(1-\nu)}.$$

De plus, la contrainte verticale est liée au taux de sédimentation constant selon la relation  $\sigma_{vv} = -\sigma_a t$ . D'après la relation  $\partial_y u_y = \epsilon_{yy}$ , on peut recalculer le déplacement en intégrant par rapport à  $y$  la quantité  $\epsilon_{yy} = E_{oe}(-\sigma_a t + \alpha p)$  et en utilisant la condition aux limites  $u_y(y=0, t) = 0$  pour déterminer la constante d'intégration. On obtient alors

$$u_y(\mathbf{x}, t) = -E_{eo}\sigma_a t y + \frac{E_{eo}A_0}{2} \left( \frac{y^2 h}{2} - \frac{y^3}{3} + 8h^3 \sum_{n=0}^{\infty} \frac{\cos(\frac{\omega_n y}{h}) - 1}{\omega_n^4} \exp(-\frac{\omega_n^2}{h^2} \tau_s t) \right). \quad (2.19)$$

Les déplacements dans la direction  $\mathbf{e}_x$  sont nuls.

Le second cas proposé, un peu plus courant dans la validation de codes de calcul pour la poroélasticité, est celui dit de Mandel. Nommé ainsi d'après la première personne à avoir proposé une solution analytique en pression, il a depuis été généralisé au cas anisotrope transverse pour lequel on dispose de l'expression analytique de la pression, des déplacements et des contraintes [Abo+96]. On utilise pour ce cas les mêmes paramètres mécaniques que pour le précédent : seuls les chargements et conditions aux limites sont différents. Le problème consiste en un milieu poreux de taille  $2L \times 2h$  situé entre deux plaques rigides. À l'instant initial, les déplacements et la pression sont nuls. A l'instant  $t = 0^+$ , une force surfacique d'intensité  $2F$  est appliquée verticalement sur la plaque supérieure tandis qu'une force de même intensité mais de signe opposé est appliquée verticalement sur la plaque inférieure, comme représenté sur la figure 2.2. En raison de sa symétrie, ce problème est souvent résolu uniquement dans le quart supérieur droit du domaine en utilisant les conditions aux limites détaillées sur la figure 2.2. Sur le bord supérieur, le déplacement vertical

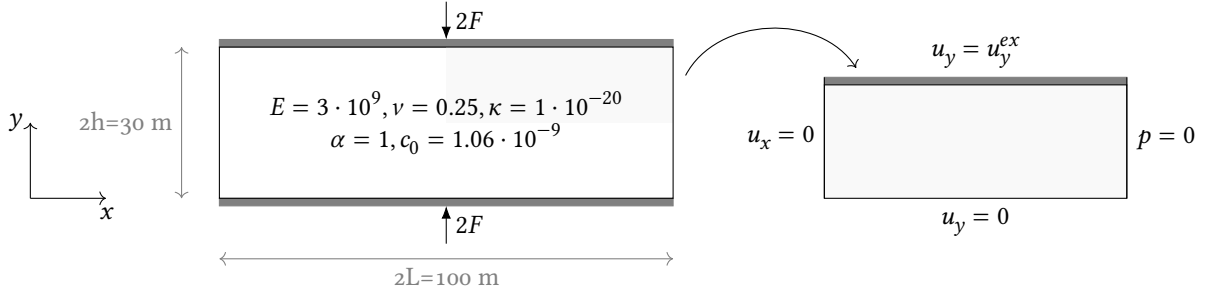


FIG. 2.2 : **Cas de Mandel**, constante du matériau et conditions aux limites. L'absence d'indication correspond à une condition de contrainte nulle ou de flux nul. En raison de la symétrie, les simulations sont effectués sur un quart du domaine.

est fixé égal au déplacement vertical exact pour plus de simplicité. En effet, imposer une condition de type contrainte pour représenter la force de compression nécessiterait d'inclure également dans la formulation le fait que le déplacement vertical reste uniforme [PW07].

La solution analytique fait intervenir quelques coefficients supplémentaires pouvant être calculés à partir de ceux définis dans la section 2.2 :

- le coefficient de diffusivité du fluide

$$c_f = \kappa \frac{\lambda + 2G}{\alpha^2 + (\lambda + 2G)c_0};$$

il s'agit du coefficient que l'on obtiendrait en reformulant l'équation de conservation de la masse en une équation de diffusion sur la quantité de fluide de la forme  $\partial_t \eta = c_f \Delta \eta$  avec  $\eta = c_0 p + \alpha \text{div}(\mathbf{u})$ ;

- le coefficient de Skempton

$$B = \frac{\alpha}{c_0 K_d + \alpha^2}$$

qui quantifie l'incrément de la pression interstitielle d'un matériau non drainé subissant un confinement ;

- le coefficient de Poisson non drainé

$$\nu_u = \frac{3\nu + \alpha B(1 - 2\nu)}{3 - \alpha B(1 - 2\nu)}$$

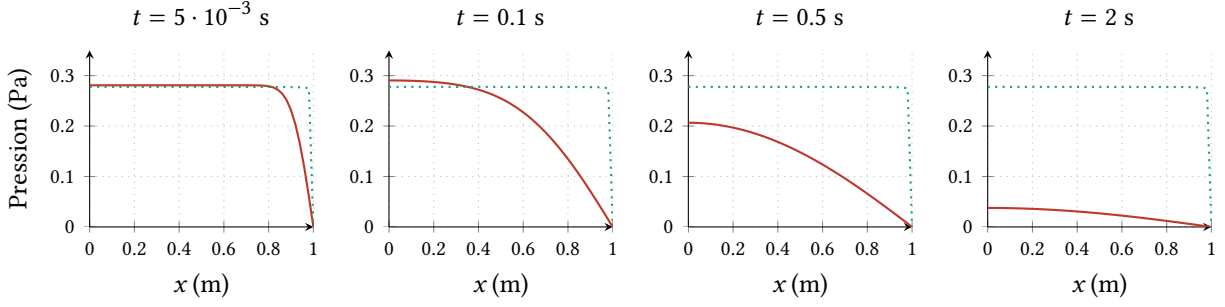


FIG. 2.3 : **Évolution de la pression analytique pour différents instants.** L'effet Mandel-Creyer s'observe sur le second graphe où la pression dépasse sa valeur initiale, représentée en pointillés, autour des  $x$  faibles. Pour simplifier les dimensions manipulées, cette figure est obtenue en fixant  $E, \kappa, c_0$  et  $L$  à 1 au lieu des valeurs prescrites par la figure 2.2.

qui est l'un des coefficients élastiques équivalents du matériau dans des conditions non drainées ( $\eta = 0$ ).

Enfin, on note  $\theta_n$  les solutions positives de l'équation non linéaire  $\tan(\theta_n) = \frac{1-\nu}{\nu_u-\nu}\theta_n$  qui doit être résolue numériquement. Ces coefficients permettent de calculer les déplacements et la pression analytiques d'après

$$p(\mathbf{x}, t) = \frac{2FB(1+\nu_u)}{3L} \sum_{n=1}^{\infty} \frac{\sin \theta_n \cos\left(\frac{\theta_n x}{L}\right) - \sin \theta_n \cos \theta_n}{\theta_n - \sin \theta_n \cos \theta_n} \exp\left(-\frac{\theta_n^2 c_f t}{L^2}\right), \quad (2.20a)$$

$$u_x(\mathbf{x}, t) = \left[ \frac{F\nu}{2GL} - \frac{F\nu_u}{GL} \sum_{n=1}^{\infty} \frac{\sin \theta_n \cos \theta_n}{\theta_n - \sin \theta_n \cos \theta_n} \exp\left(-\frac{\theta_n^2 c_f t}{L^2}\right) \right] x + \frac{F}{G} \sum_{n=1}^{\infty} \frac{\sin\left(\frac{\theta_n x}{L}\right) \cos \theta_n}{\theta_n - \sin \theta_n \cos \theta_n} \exp\left(-\frac{\theta_n^2 c_f t}{L^2}\right) \quad (2.20b)$$

$$u_y(\mathbf{x}, t) = \left[ \frac{-F(1-\nu)}{2GL} + \frac{F(1-\nu_u)}{GL} \sum_{n=1}^{\infty} \frac{\sin \theta_n \cos \theta_n}{\theta_n - \sin \theta_n \cos \theta_n} \exp\left(-\frac{\theta_n^2 c_f t}{L^2}\right) \right] y. \quad (2.20c)$$

Comme dans le cas précédent, les déplacements et la pression ne dépendent que d'une seule direction. En revanche, le champ de déplacement a cette fois ses deux composantes non nulles. Ce cas est numériquement délicat car la pression passe de 0 à l'instant  $t = 0$  à la valeur  $\frac{FB(1+\nu_u)}{3L}$  à l'instant  $t = 0^+$ . De même, les valeurs asymptotiques des champs  $u_x$  et  $u_y$  valent respectivement

$$u_x(x = L, t \rightarrow 0^+) = \frac{F\nu_u}{2G} \quad \text{et} \quad u_y(y = h, t \rightarrow 0^+) = \frac{-F(1-\nu_u)\alpha}{2GL}$$

et sont donc discontinues par rapport à la condition initiale  $\mathbf{u} = \mathbf{0}$ . De plus, ce cas permet d'illustrer un comportement dû à l'interaction fluide - solide. Il s'agit ici de l'effet de *Mandel-Creyer*, par lequel la déformation du solide fait apparaître un incrément de pression à proximité de l'axe  $x = 0$ , et joue donc le rôle de terme source pour l'équation de conservation. Cet effet est illustré sur la figure 2.3 qui représente l'évolution de la pression analytique pour différentes valeurs de  $t$ . Après l'instant initial, la valeur de  $p$  dépasse la valeur asymptotique (à  $t = 0.1$  s) aux alentours de  $x = 0$  avant de redescendre vers 0.



# 3

## Méthode des éléments virtuels pour le problème d'élasticité

Avant de réaliser le couplage entre les deux méthodes de discrétisation pour les deux équations du système poroélastique, il est indispensable de maîtriser les méthodes choisies sur les sous problèmes associés. Concernant la partie écoulement, les schémas volumes finis considérés sont relativement standards et ont pu être directement utilisés. En revanche, concernant la partie mécanique, l'usage de la méthode des éléments virtuels n'est pas une approche fréquente. Ce chapitre est donc consacré à l'application de cette méthode à l'équation de la mécanique. Par conséquent, on ne considère que le système d'élasticité linéaire défini par (2.11) sans tenir compte du fluide. La première section du chapitre détaille la construction de la méthode des éléments virtuels d'ordre 1 pour ce système. Les différents ingrédients tels que l'espace d'approximation, la forme bilinéaire discrète et les opérateurs de projection spécifiques aux méthodes éléments virtuels y sont présentés. Même si la lecture de cette première section est suffisante pour implémenter la méthode dans un code de calcul, quelques détails supplémentaires sur ce point sont exposés dans la section 3.2. Cette implémentation, réalisée dans la plateforme Arcane, est utilisée afin de proposer quelques illustrations dans la section 3.3 qui permettent d'ajuster certains paramètres numériques et de valider la méthode. Enfin, des exemples d'application un peu plus particuliers sont présentés dans la section 3.4.

### Sommaire

---

<b>3.1</b>	<b>Construction de la méthode</b>	<b>30</b>
3.1.1	Espace d'approximation	30
3.1.2	Formulation variationnelle approchée	33
3.1.3	Explicitation de la projection	37
<b>3.2</b>	<b>Implémentation</b>	<b>39</b>
<b>3.3</b>	<b>Paramétrage et validation numérique</b>	<b>42</b>
3.3.1	Choix du coefficient de stabilisation	43
3.3.2	Convergence en espace	46
<b>3.4</b>	<b>Observations supplémentaires sur la méthode</b>	<b>46</b>
3.4.1	Traitement de maillages particuliers	48
3.4.2	Comparaison aux éléments finis	49

### 3.1 Construction de la méthode

Historiquement, les méthodes des éléments virtuels tirent leurs origines des différences finies mimétiques, réinterprétées dans un cadre variationnel. Ce changement de point de vue a considérablement simplifié la définition des méthodes et a ainsi permis leur application à des problèmes physiques plus variés, la montée en ordre polynomial, le choix d'espaces d'approximation plus complexes ou encore une analyse mathématique des méthodes dans un cadre unifié. Si la logique sous-jacente à ces méthodes devait être résumée, on pourrait la formuler de la manière suivante :

1. Sur chaque élément, la dimension de l'espace d'approximation doit être égale au nombre de degrés de liberté. L'espace vectoriel des fonctions polynomiales (de degré  $k$ ) n'étant pas assez large, on y ajoute un jeu de fonctions de bases définies *virtuellement* par un ensemble de conditions.
2. Ne pouvant pas évaluer la forme bilinéaire exacte sans la connaissance explicite de ces fonctions, on la remplace par une forme *approchée* calculable à partir des informations à disposition.
3. Puisque les informations à disposition concernent majoritairement les fonctions polynomiales, des opérateurs de *projection* sont définis et utilisés pour passer d'une fonction virtuelle à sa projection dans l'espace des polynômes.

Si ces ingrédients sont communs à toutes les méthodes éléments virtuels, leur choix peut varier en fonction du problème étudié, de l'espace d'approximation choisi et de l'ordre d'approximation souhaité. Bien qu'on puisse trouver des définitions générales de la méthode pour le problème d'élasticité, on se focalise par soucis de simplification sur la construction d'une méthode d'ordre polynomial un pour le système d'élasticité linéaire.

#### 3.1.1 Espace d'approximation

##### Maillage

Soit  $\tau_h$  une famille de *polytopes* – polygones en deux dimensions ou polyèdres en trois dimensions – Lipschitz  $(K_i)_{i=1 \dots \#\tau_h}$ , non nécessairement convexes mais ne se recouvrant pas, formant une partition de  $\Omega$  :

$$\overline{\Omega} = \cup_{i=1}^{\#\tau_h} K_i \text{ et } \overset{\circ}{K}_i \cap \overset{\circ}{K}_j = \emptyset \text{ pour } i \neq j.$$

Pour chaque polytope,  $h_K$ ,  $|K|$  et  $\mathbf{x}^K$  représentent respectivement le diamètre, la mesure et le barycentre de  $K$ . On note  $M_K$  le nombre de sommets de  $K$  et  $\mathcal{M}_K$  l'ensemble des indices de ces sommets. Dans le cas tridimensionnel, l'indice ou l'exposant  $f$  est utilisé pour définir de manière similaire  $h_f$ ,  $|f|$ ,  $\mathbf{x}^f$ ,  $M_f$  et  $\mathcal{M}_f$  pour chaque face  $f$  de  $K$ . Dans le cas bidimensionnel, on conserve la notation  $K$  pour désigner les polygones. Au niveau global, on note  $M$  le nombre total de sommets du maillage et  $\mathcal{M}$  l'ensemble des indices de ces sommets. Cet ensemble se décompose en deux sous-ensembles  $\mathcal{M}_{D_d}$  et  $\mathcal{M}_{\overline{D_d}} = \mathcal{M} \setminus \mathcal{M}_{D_d}$  qui représentent respectivement les indices des sommets appartenants à la frontière  $\Gamma_{D_d}$  et les indices des autres sommets du maillage.

### Espace d'approximation local

Indépendamment de la dimension, l'espace d'approximation  $V_h^K \subset H^1(K)$  est construit sur chaque élément  $K$  de manière à satisfaire les assertions suivantes :

- Les polynômes de degré inférieur ou égal à 1 sur  $K$  sont inclus dans  $V_h^K$ .
- Les fonctions de  $V_h^K$  sont bien définies sur les sommets de  $K$ , et chaque fonction de  $V_h^K$  peut être déterminée de manière unique par ses valeurs sur les  $M_K$  sommets de  $K$ . En d'autres termes, les valeurs aux sommets de  $K$  constituent les degrés de liberté des fonctions de bases et la dimension de  $V_h^K$  est égale à  $M_K$ .
- En trois (resp. deux) dimensions, la trace de toute fonction de  $V_h^K$  sur n'importe quelle face  $f$  (resp. n'importe quelle arête  $e$ ) de la frontière  $\partial K$  de  $K$  ne dépend que des valeurs de la fonction sur les sommets de  $f$  (resp. de  $e$ ). Comme dans la théorie éléments finis classique, cette propriété garantie que le raccordement des espaces  $V_h^K$  donne une approximation conforme de  $H^1(\Omega)$ .

Dans un premier temps, on construit un espace de fonctions scalaires  $V_{h,scal}^K$  satisfaisant ces propriétés et dont l'expression dépend de la dimension considérée. Le passage à un espace d'approximation vectoriel réutilisera simplement cet espace scalaire.

En deux dimensions, la notation  $K$  désigne un polygone. L'espace  $V_{h,scal}^K$  est défini d'après [Bei+13] par l'équation (3.1) et les auteurs montrent que les valeurs aux nœuds constituent les degrés de liberté de l'espace local.

$$V_{h,scal}^K = \{v_h \in H^1(K) \text{ t.q. : pour toute arête } e \subset \partial K, v_{h|e} \in \mathbb{P}_1(e); v_{h|\partial K} \in \mathcal{C}^0(\partial K); \Delta v_h = 0 \text{ dans } K\}. \quad (3.1)$$

Autrement dit,  $V_{h,scal}^K$  est constitué de fonctions harmoniques linéaires sur chaque arête. En définissant la fonction de base locale  $\varphi_i \in V_{h,scal}^K$  associée au nœud  $i$ ,

$$\Delta \varphi_i = 0 \text{ dans } K, \quad \varphi_i = \delta_i \text{ sur } \partial K \text{ avec } \delta_i \text{ vérifiant } \begin{cases} \delta_i(V_j) = 1 \text{ si } i = j, 0 \text{ sinon,} \\ \delta_i \text{ linéaire sur chaque arête } e, \\ \delta_i \text{ continue sur } \partial K. \end{cases}$$

Il est clair que les polynômes de degré 1 vérifient les trois conditions de la définition 3.1, ce qui implique  $\mathbb{P}_1(K) \subset V_{h,scal}^K$ . On note également que dans le cas où  $K$  est un triangle, seuls les polynômes de degré 1 vérifient les conditions. Dans ce cas particulier, l'espace d'approximation coïncide avec l'espace des éléments finis  $\mathbb{P}_1$  de Lagrange :  $V_{h,scal}^K = \mathbb{P}_1(K)$ . Dans le cas général, l'espace d'approximation inclut d'autres fonctions *virtuellement* définies par la condition  $\Delta v_h = 0$ . La méthode des éléments virtuels est en effet construite de sorte que ces fonctions, *a priori* non polynomiales, n'aient jamais à être connues explicitement.

En trois dimensions, la définition de  $V_{h,scal}^K$  est moins directe. La notation  $K$  désigne à présent un polyèdre dont les faces  $f$  sont des polygones. Dans un premier temps, on définit sur chacune de ces faces un opérateur de projection  $\pi^{f,0} : \{v \in H^1(f) \text{ t.q. } v \in \mathcal{C}^0(\partial f)\} \rightarrow \mathbb{P}_1$  par

$$\pi^{f,0} v(\mathbf{x}) = \langle \nabla v \rangle (\mathbf{x} - \bar{\mathbf{x}}_f) + \bar{v}_f \quad \text{où} \quad \begin{cases} \langle \nabla v \rangle = \frac{1}{|f|} \int_f \nabla v \\ \bar{\zeta}_f = \frac{1}{M_f} \sum_{i \in \mathcal{M}_f} \zeta(V_i) \text{ avec } \zeta \equiv \mathbf{x} \text{ ou } \zeta \equiv v. \end{cases} \quad (3.2)$$



Ceci permet de construire pour chaque face un espace dans lequel  $\pi^{f,0}$  est une projection  $L^2$  :

$$V_{h,scal}^f = \left\{ v_h \in H^1(f) \text{ t.q. : } \forall e \in \partial f, v_{h|e} \in \mathbb{P}_1(e); v_{h|\partial f} \in \mathcal{C}^0(\partial f); \Delta v_h \in \mathbb{P}_1; \int_f (\pi^{f,0} v_h) q = \int_f v_h q \quad \forall q \in \mathbb{P}_1 \right\}$$

puis d'utiliser ces espaces sur les faces pour définir l'espace d'approximation  $V_{h,scal}^K$  sur la cellule par

$$V_{h,scal}^K = \left\{ v_h \in H^1(K) \text{ tel que } v_{h|f} \in V_{h,scal}^f \quad \forall f \in \partial K; \quad \Delta v_h = 0 \text{ dans } K \right\}. \quad (3.3)$$

À nouveau, on note que  $\mathbb{P}_1(K) \subset V_{h,scal}^K$  avec égalité et donc coïncidence avec les éléments finis  $\mathbb{P}_1$  de Lagrange dans le cas où  $K$  est un tétraèdre. Les degrés de liberté de  $V_{h,scal}^K$  sont les valeurs aux sommets du polyèdre, ce qui implique que la dimension de l'espace est égale à  $M_K$ .

*Remarque.* L'extension naturelle aurait été de prendre la définition du cas bidimensionnel (3.1) en remplaçant  $K$  par  $f$  pour les espaces polygonaux  $V_{h,scal}^f$ . Malheureusement, ce choix conduirait à un espace d'approximation qui, dans l'étape suivante, ne permettrait pas d'évaluer la forme bilinéaire approchée. On verra effectivement plus loin que dans le cas tridimensionnel, le calcul d'intégrales de la forme  $\int_f v_h$  est requis. Or, l'espace construit par (3.1) ne définit pas les fonctions  $v_h$  à l'intérieur des faces. La technique utilisée est classique des éléments virtuels : l'information manquante est ajoutée à l'espace d'approximation (via la condition  $\int_f (\pi^{f,0} v_h) q = \int_f v_h q \quad \forall q \in \mathbb{P}_1$ ), mais ce dernier reste de même dimension puisque la condition  $\Delta v_h = 0$  est remplacée par  $\Delta v_h \in \mathbb{P}_1$ . Les deux définitions de  $V_{h,scal}^f$  conduisent finalement à deux espaces équivalents en terme de dimension et de degrés de liberté, comme démontré dans [Ahm+13] : un jeu de valeurs aux sommets permet d'identifier de manière unique une fonction dans chacun de ces espaces, qui sont distinctes dans le cas général. Le choix effectué pour  $V_{h,scal}^f$  est purement guidé par le besoin d'avoir une forme bilinéaire calculable pour les fonctions de l'espace d'approximation.

### Espace d'approximation global

Quelle que soit la dimension, l'espace d'approximation global  $V_h$  est construit à partir de l'assemblage conforme des espaces locaux :

$$V_{h,scal} = \left\{ v_h \in H^1(\Omega) \text{ tel que } v_{h|K} \in V_{h,scal}^K \text{ pour tout } K \in \tau_h \right\}. \quad (3.4)$$

Les degrés de liberté globaux sont les valeurs de  $v_h$  sur chacun des sommets du maillage : à chacun de ces sommets est associé la fonction de base scalaire  $\varphi_i$  satisfaisant  $\varphi_i(V_j) = \delta_{ij} \quad \forall j \in \mathcal{M}$ . Toute fonction scalaire  $v_h$  de  $V_{h,scal}$  se décompose donc selon l'identité d'interpolation lagrangienne

$$v_h(\mathbf{x}) = \sum_{i \in \mathcal{M}} v_h(V_i) \varphi_i(\mathbf{x}).$$

Pour traiter le problème d'élasticité, qui est vectoriel, les espaces scalaires sont dupliqués pour créer les espaces d'approximation vectoriels  $V_h^K = (V_{h,scal}^K)^d$  et  $V_h = (V_{h,scal})^d$ . De la même manière, chaque fonction de base scalaire permet de créer  $d$  fonctions de base vectorielles selon  $\boldsymbol{\varphi}_i^k = \varphi_i \mathbf{e}_k$  ( $k = x, y, z$ ). Finalement, les fonctions vectorielles  $\mathbf{v}_h$  de  $V_h$  s'expriment dans la base  $\{\boldsymbol{\varphi}_i^k\}$  selon

$$\mathbf{v}_h(\mathbf{x}) = \sum_{i \in \mathcal{M}} (v_h(V_i))^x \boldsymbol{\varphi}_i^x(\mathbf{x}) + (v_h(V_i))^y \boldsymbol{\varphi}_i^y(\mathbf{x}) + (v_h(V_i))^z \boldsymbol{\varphi}_i^z(\mathbf{x}). \quad (3.5)$$

A titre d'illustration, on représente sur la figure 3.1 une fonction de base en un sommet choisi pour diverses mailles en dimension deux. Dans ce contexte uniquement, les fonctions sont calculées numériquement en résolvant l'équation  $\Delta\varphi_i = 0$  : les cellules sont discrétisées en sous éléments et les conditions aux limites à appliquer sont déduites de la définition de l'espace d'approximation. Dans l'utilisation normale des éléments virtuels, cette manipulation n'est ni souhaitable ni nécessaire.

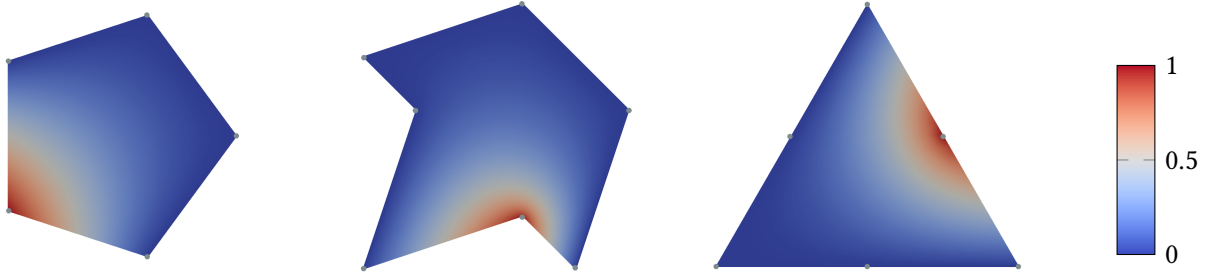


FIG. 3.1 : **Valeur de la fonction de base scalaire** associée à un sommet sur chacune de ces trois cellules. Contrairement aux apparences, la cellule de droite est définie comme un hexagone à six arêtes et non pas comme un triangle.

### 3.1.2 Formulation variationnelle approchée

L'espace d'approximation défini précédemment peut être utilisé dans des méthodes polygonales de type éléments finis, appelées méthodes des éléments finis harmoniques, voir par exemple [Bis14] pour le problème d'élasticité. Dans ces méthodes, l'évaluation de la forme bilinéaire  $a^\Omega$  issue de la formulation faible (2.13) nécessite d'approcher les fonctions de base, par exemple en partitionnant chaque cellule en simplexes et en résolvant par éléments finis le problème de Laplace, ce qui rend ces méthodes coûteuses. La philosophie des méthodes des éléments virtuels est différente : plutôt que de déterminer les fonctions de base, le problème variationnel est modifié de sorte que les termes soient calculables directement depuis les degrés de liberté. Pour aboutir à ce problème modifié, il faut approcher la forme bilinéaire et le second membre du système (2.13).

#### Forme bilinéaire discrète

Pour approcher la forme bilinéaire,  $a^\Omega$  est tout d'abord décomposée sur les éléments

$$a^\Omega(\mathbf{u}_h, \mathbf{v}_h) = \sum_{K \in \tau_h} a^K(\mathbf{u}_h, \mathbf{v}_h) \text{ avec } a^K(\mathbf{u}_h, \mathbf{v}_h) = \int_K \overline{\overline{C}} \epsilon_{u_h} : \overline{\overline{\epsilon_{v_h}}}$$

puis  $a^K$  est approchée sur chaque élément par une forme bilinéaire  $a_h^K$ . Ces formes approchées élémentaires sont ensuite regroupées pour définir la forme bilinéaire approchée globale  $a_h$  :

$$a_h(\mathbf{u}_h, \mathbf{v}_h) = \sum_{K \in \tau_h} a_h^K(\mathbf{u}_h, \mathbf{v}_h) \text{ avec } a_h^K(\mathbf{u}_h, \mathbf{v}_h) \approx a^K(\mathbf{u}_h, \mathbf{v}_h).$$

Malgré ce *crime variationnel*, la convergence de la méthode est assurée si les deux hypothèses suivantes sont respectées pour tout  $K \in \tau_h$  [Bei+13] :

- Consistance : pour tout  $\mathbf{p} \in \mathbb{P}_1^d$  et pour tout  $\mathbf{v}_h \in V_h^K$

$$a_h^K(\mathbf{v}_h, \mathbf{p}) = a^K(\mathbf{v}_h, \mathbf{p}) \quad (3.6)$$

- Stabilité : il existe deux constantes positives  $\alpha_*$  et  $\alpha^*$  indépendantes de  $K$  et de  $h$  telles que

$$\forall \mathbf{v}_h \in V_h^K, \quad \alpha_* a^K(\mathbf{v}_h, \mathbf{v}_h) \leq a_h^K(\mathbf{v}_h, \mathbf{v}_h) \leq \alpha^* a^K(\mathbf{v}_h, \mathbf{v}_h). \quad (3.7)$$

La propriété de consistance assure la précision de la méthode. En particulier, elle garantit la satisfaction du *patch test*, c'est-à-dire le fait que la méthode est exacte dans le cas où la solution exacte est linéaire. La propriété de stabilité permet de conserver la continuité et la coercivité de la forme bilinéaire approchée. Sans cette propriété, la matrice associée à la forme bilinéaire serait de rang déficient ce qui poserait de gros problèmes de stabilité.

Puisque  $V_h^K$  contient des éléments polynomiaux et non polynomiaux, et que la propriété de consistance porte sur les polynômes, il est naturel de définir sur chaque élément  $K$  un opérateur de projection  $\pi^K$  de l'espace  $V_h^K$  vers l'espace des polynômes  $\mathbb{P}_1^d$ , de sorte que pour tout élément  $\mathbf{v}_h$  dans  $V_h^K$ ,

$$\mathbf{v}_h = \pi^K(\mathbf{v}_h) + (\bar{I}_d - \pi^K)(\mathbf{v}_h). \quad (3.8)$$

Plus précisément, si on arrive à définir un tel opérateur de projection satisfaisant de plus la relation

$$a^K(\pi^K \mathbf{v}_h, \mathbf{q}) = a^K(\mathbf{v}_h, \mathbf{q}) \text{ pour tout } \mathbf{q} \in \mathbb{P}_1^d, \quad (3.9)$$

on montre facilement que la forme bilinéaire approchée localement définie par

$$\tilde{a}_h^K(\mathbf{u}, \mathbf{v}) := a^K(\pi^K \mathbf{u}, \pi^K \mathbf{v})$$

est consistante au sens de (3.6) : soit  $\mathbf{v}_h \in V_h^K$  et  $\mathbf{p} \in \mathbb{P}_1^d$ , alors

$$\begin{aligned} \tilde{a}_h^K(\mathbf{v}_h, \mathbf{p}) &= a^K(\pi^K \mathbf{v}_h, \pi^K \mathbf{p}) = a^K(\pi^K \mathbf{v}_h, \mathbf{p}) \text{ car } \pi^K \mathbf{p} = \mathbf{p} \text{ pour } \mathbf{p} \in \mathbb{P}_1^d \\ &= a^K(\mathbf{v}_h, \mathbf{p}) \text{ d'après (3.9).} \end{aligned}$$

Cependant, cette forme bilinéaire ne satisfait pas la propriété de stabilité, ce qui nécessite l'introduction d'une autre forme bilinéaire  $\tilde{s}_h^K$  qui aura pour rôle d'assurer cette propriété. Procédant comme dans [BLM15], on pose

$$\tilde{s}_h^K(\mathbf{u}, \mathbf{v}) := s_K \sum_{i \in \mathcal{M}_K} (\mathbf{u} - \pi^K \mathbf{u})(V_i) \cdot (\mathbf{v} - \pi^K \mathbf{v})(V_i) \quad (3.10)$$

où  $s_K$  désigne un coefficient réel positif dont le choix sera discuté par la suite. Puisque  $\tilde{s}_h^K$  s'annule lorsque l'un de ces deux arguments est un polynôme de degré un, elle n'altère pas la propriété de consistance. En revanche, elle permet d'assurer la propriété de stabilité (3.7) : sous certaines hypothèses sur la forme des mailles et sur le choix de  $s_K$ , il existe  $c_0, c_1$  constantes indépendantes de  $K$  et de  $h$  telles que

$$\forall \mathbf{v}_h \in V_h^K \text{ tel que } \pi^K \mathbf{v}_h = \mathbf{0}, \quad c_0 a^K(\mathbf{v}_h, \mathbf{v}_h) \leq \tilde{s}_h^K(\mathbf{v}_h, \mathbf{v}_h) \leq c_1 a^K(\mathbf{v}_h, \mathbf{v}_h).$$

Finalement, la somme de  $\tilde{a}_h^K$  et  $\tilde{s}_h^K$  définit la forme bilinéaire approchée consistante et stable

$$a_h^K(\mathbf{u}, \mathbf{v}) := \tilde{a}_h^K(\mathbf{u}, \mathbf{v}) + \tilde{s}_h^K(\mathbf{u}, \mathbf{v}) = a^K(\pi^K \mathbf{u}, \pi^K \mathbf{v}) + s_K \sum_{i \in \mathcal{M}_K} (\mathbf{u} - \pi^K \mathbf{u})(V_i) \cdot (\mathbf{v} - \pi^K \mathbf{v})(V_i). \quad (3.11)$$

Il reste à construire l'opérateur de projection  $\pi^K$  satisfaisant (3.9), ce qui sera l'objet de la section 3.1.3.

## Second membre

Dans le cadre de la méthode des éléments virtuels, le second membre local

$$l^K(\mathbf{v}_h) = \int_K \mathbf{f} \cdot \mathbf{v}_h + \int_{\partial K \cap \Gamma_{N_d}} \mathbf{t}_N \cdot \mathbf{v}_h$$

doit également être approché par une forme linéaire approchée  $l_h^K(\mathbf{v}_h)$  de manière à être calculable à partir des degrés de liberté. En suivant l'approche développée dans [GTP14], on utilise une formule de quadrature basée sur les nœuds en prenant

$$l_h^K(\mathbf{v}_h) := \sum_{i \in \mathcal{M}_K} \omega_i^K \mathbf{f}(V_i) \cdot \mathbf{v}_h(V_i) + \sum_{i \in \mathcal{M}_{K \cap \Gamma_{N_d}}} \omega_i^{\partial K} \mathbf{t}_N(V_i) \cdot \mathbf{v}_h(V_i) \quad (3.12)$$

où les poids d'intégration  $\omega_i^K$  et  $\omega_i^{\partial K}$  dépendent de la dimension considérée et sont représentés sur la figure 3.2. La force volumique s'applique sur les cellules (polyèdres en trois dimensions ou polygones en deux dimensions) tandis que la traction de Neumann s'applique sur les bords (faces polygonales pour les problèmes tridimensionnels ou arêtes pour les problèmes bidimensionnels). On montre que ces poids

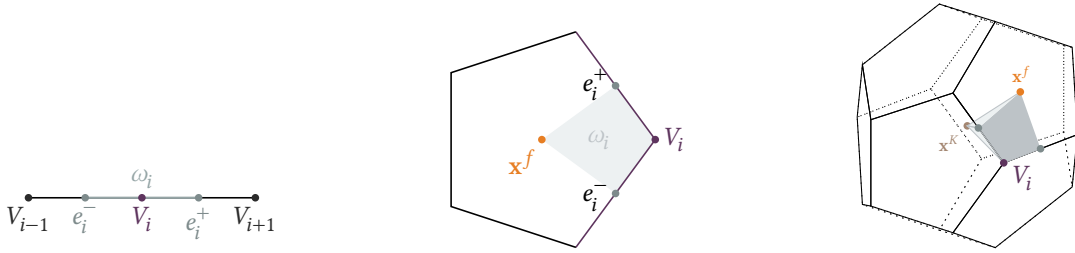


FIG. 3.2 : **Poids d'intégration** pour la forme de quadrature basée sur les nœuds. En dimension un,  $\omega_i$  est la longueur du segment formé par les milieux des arêtes  $e_i^-$ ,  $e_i^+$ . En dimension deux,  $\omega_i$  est l'aire du quadrangle formé par le sommet  $V_i$ , le barycentre  $\mathbf{x}^f$  et les milieux des arêtes  $e_i^-$ ,  $e_i^+$ . En dimension trois,  $\omega_i$  est la somme du volume des pyramides de base  $V_i$ ,  $\mathbf{x}^f$ , milieux des arêtes  $e_i$  et de sommet  $\mathbf{x}^K$  (barycentre de  $K$ ). Pour plus de lisibilité, seule une des trois pyramides est représentée sur la figure.

permettent d'intégrer exactement les fonctions linéaires selon  $\int_K p = \sum_{i \in \mathcal{M}_K} \omega_i^K p(V_i)$  : par exemple en dimension deux,

$$\begin{aligned} 3 \int_f p &= \sum_{i \in \mathcal{M}_f} |T_{V_i, V_{i+1}, \mathbf{x}^f}| (p(V_i) + p(V_{i+1}) + p(\mathbf{x}^f)) \text{ en triangulant par rapport au barycentre } \mathbf{x}^f \\ &= |f| p(\mathbf{x}^f) + 2 \sum_{i \in \mathcal{M}_f} \omega_i^f p(V_i) \text{ car } \omega_i^f = \frac{1}{2} (|T_{V_{i-1}, V_i, \mathbf{x}^f}| + |T_{V_i, V_{i+1}, \mathbf{x}^f}|) \\ &= \int_f p + 2 \sum_{i \in \mathcal{M}_f} \omega_i^f p(V_i) \text{ par définition du barycentre.} \end{aligned}$$

De cette manière, si les données  $\mathbf{f}$  et  $\mathbf{t}_N$  sont suffisamment régulières, l'erreur due à la formule de quadrature est comparable à l'erreur du schéma numérique et ne modifie pas l'ordre de convergence. D'autres méthodes d'approximation de la forme linéaire peuvent être utilisées, voir [BLM15 ; ANR17]. Comme pour la forme bilinéaire, la somme sur les cellules de ces seconds membres approchés définit le second membre approché global selon  $l_h = \sum_{K \in \mathcal{T}_h} l_h^K$ .

### Formulation faible approchée

Les formes bilinéaires et linéaires approchées étant définies, on peut à présent poser le problème variationnel approché résolu par la méthode des éléments virtuels. Pour tenir compte des conditions aux limites de Dirichlet  $\mathbf{u} = \mathbf{g}_D$  sur  $\Gamma_{D_d}$ , on introduit les espaces

$$V_{h,g} = \{\mathbf{v}_h \in V_h \text{ tel que } \mathbf{v}_h(V_i) = \mathbf{g}_D(V_i) \text{ pour } i \in \mathcal{M}_{D_d}\} \quad (3.13)$$

et

$$V_{h,0} = \{\mathbf{v}_h \in V_h \text{ tel que } \mathbf{v}_h(V_i) = \mathbf{0} \text{ pour } i \in \mathcal{M}_{D_d}\} \quad (3.14)$$

de sorte qu'un élément  $\mathbf{v}_h \in V_{h,g}$  se réécrit dans la base  $\{\boldsymbol{\varphi}_i^k\}$  selon

$$\mathbf{v}_h = \sum_{i \in \mathcal{M}_{\overline{D_d}}} (\mathbf{v}_h(V_i))^x \boldsymbol{\varphi}_i^x + (\mathbf{v}_h(V_i))^y \boldsymbol{\varphi}_i^y + (\mathbf{v}_h(V_i))^z \boldsymbol{\varphi}_i^z + \sum_{i \in \mathcal{M}_{D_d}} (\mathbf{g}_D(V_i))^x \boldsymbol{\varphi}_i^x + (\mathbf{g}_D(V_i))^y \boldsymbol{\varphi}_i^y + (\mathbf{g}_D(V_i))^z \boldsymbol{\varphi}_i^z.$$

La formulation variationnelle approchée s'écrit alors

#### Formulation variationnelle approchée

Trouver  $\mathbf{u}_h \in V_{h,g}$  tel que pour tout  $\mathbf{v}_h \in V_{h,0}$ ,

$$a_h(\mathbf{u}_h, \mathbf{v}_h) = l_h(\mathbf{v}_h) \quad (3.15a)$$

ou, de manière équivalente, trouver  $\mathcal{U} = (u_1^x, u_1^y, \dots, u_{\#\mathcal{M}_{\overline{D_d}}}^z)$  solution du système

$$\mathcal{A}\mathcal{U} = \mathcal{L}^u \quad \text{avec} \quad \mathcal{A}_{i_k j_l} = a_h(\boldsymbol{\varphi}_i^k, \boldsymbol{\varphi}_j^l) \quad \text{et} \quad \mathcal{L}_{i_k}^u = l_h(\boldsymbol{\varphi}_i^k) - \sum_{\substack{j \in \mathcal{M}_{D_d} \\ l=x,y,z}} (\mathbf{g}_D(V_j))^l a_h(\boldsymbol{\varphi}_i^k, \boldsymbol{\varphi}_j^l). \quad (3.15b)$$

De la même manière que dans la méthode des éléments finis, l'assemblage de la matrice  $\mathcal{A}$  et du second membre  $\mathcal{L}^u$  est effectué en calculant les matrices élémentaires  $\mathcal{A}^K$  et les seconds membres élémentaires  $\mathcal{L}^{u,K}$  sur chaque maille.

*Remarque.* En pratique, les conditions aux limites de Dirichlet sont définies en termes de degrés de liberté plutôt qu'en termes de sommets, comme le laissent entendre les notations précédentes utilisées pour plus de simplicité. Étant donné un nœud  $V_j$  situé sur un bord du domaine, on peut tout à fait mixer les conditions aux limites en imposant par exemple un déplacement fixé (condition de Dirichlet) pour la composante  $x$  du vecteur et une contrainte de traction (condition de Neumann) pour les composantes  $y$  et  $z$  du vecteur. Ce type de conditions aux limites est utile pour de nombreux problèmes de mécanique où les déplacements peuvent n'être bloqués que dans une seule direction, et sont d'ailleurs utilisées dans la section 2.2.3. La formulation variationnelle reste valable en plaçant dans ce cas  $j_x$  dans  $\mathcal{M}_{D_d}$  et  $j_y, j_z$  dans  $\mathcal{M}_{\overline{D_d}}$ .

La littérature sur les éléments virtuels [BBM13] fournit le résultat suivant :

**Lemme 3.1** *On suppose que chaque maille  $K$  est étoilée par rapport à une boule de rayon  $rh_K$ , où  $r$  est une constante inférieure au rapport de la plus petite arête de  $K$  et de son diamètre. Si la forme bilinéaire approchée  $a_h$  vérifie les hypothèses de consistance (3.6) et de stabilité (3.7), alors le problème approché (3.15a) admet une*

unique solution  $\mathbf{u}_h$ . De plus, pour toute approximation  $\mathbf{u}_I$  de  $\mathbf{u}$  dans  $V_h$  et pour toute approximation  $\mathbf{u}_\pi$  de  $\mathbf{u}$  dans  $\mathbb{P}_1$  pour chaque maille,

$$\|\mathbf{u}_h - \mathbf{u}\|_{H^1} \leq C \left( \|\mathbf{u} - \mathbf{u}_I\|_{H^1} + \left( \sum_{K \in \mathcal{T}_h} \|\mathbf{u} - \mathbf{u}_\pi\|_{H^1(K)}^2 \right)^{\frac{1}{2}} + \sup_{\mathbf{v}_h \in V_{h,0}} \frac{l(\mathbf{v}_h) - l_h(\mathbf{v}_h)}{\|\mathbf{v}_h\|_{H^1}} \right)$$

où  $C$  est une constante dépendant de  $\Omega, \bar{C}, \alpha^*, \alpha_*$  et de la constante de coercivité de  $a^\Omega$ .

### 3.1.3 Explicitation de la projection

Il ressort du paragraphe précédent que le cœur de la méthode des éléments virtuels consiste à trouver un projecteur  $\pi^K$  satisfaisant (3.9) et calculable depuis les degrés de liberté de l'espace  $V_h^K$ .

Dans le cas général, la définition de cette projection est constructive [BBMR14] : en réécrivant l'égalité (3.9)  $a^K(\pi^K \mathbf{v}_h, \mathbf{q}) = a^K(\mathbf{v}_h, \mathbf{q})$  pour chaque monôme polynomial, on aboutit à un système linéaire de taille  $\dim(\mathbb{P}_1^d)$  dont les inconnues sont les coefficients de  $\pi^K(\mathbf{v}_h)$ . Ce faisant, on observe alors que le choix des degrés de liberté et la définition de  $V_h^K$  permettent d'assembler ce système linéaire. En revanche, puisque la condition (3.9) est satisfaite à une constante près, les monômes d'ordre zéro  $\mathbf{q} = \mathbb{1}_K \mathbf{e}_k$  donnent l'égalité  $0 = 0$  et conduisent à un système de rang déficient. Une dernière étape consiste donc à ajouter une contrainte supplémentaire, typiquement

$$\sum_{i \in \mathcal{M}_K} \pi^K \mathbf{v}_h(V_i) = \sum_{i \in \mathcal{M}_K} \mathbf{v}_h(V_i).$$

La résolution de ce système linéaire pour chaque fonction de base permet d'expliciter la représentation matricielle de la projection  $\pi^K$ .

Dans notre cas particulier où l'ordre d'approximation polynomial n'est que de un, l'expression de  $\pi^K$  qui est finalement assez simple est directement donnée par l'équation (3.16).

#### Projecteur des éléments virtuels à l'ordre 1

$$\pi^K \mathbf{v}(\mathbf{x}) = \langle \nabla \mathbf{v} \rangle (\mathbf{x} - \bar{\mathbf{x}}_K) + \bar{\mathbf{v}} \quad (3.16)$$

où l'on a défini les notations suivantes assimilables à des moyennes :

- $\langle \mathbf{v} \rangle = \frac{1}{|K|} \int_K \mathbf{v}$  donc  $\langle \nabla \mathbf{v} \rangle$  représente le gradient moyen par maille,
- $\bar{\mathbf{v}} = \frac{1}{M_K} \sum_{i \in \mathcal{M}_K} \mathbf{v}(V_i)$  et, en particulier,  $\bar{\mathbf{x}}_K = \frac{1}{M_K} \sum_{i \in \mathcal{M}_K} \mathbf{x}(V_i)$ .

En effet, cette expression étant donnée, on peut effectuer le raisonnement précédemment décrit "à l'envers" pour vérifier que l'équation de consistance (3.9) est bien satisfaite, et que le projecteur est calculable pour chaque fonction de base en utilisant uniquement les informations de l'espace d'approximation  $V_h^K$  et les degrés de liberté.

Soit  $\mathbf{v}_h$  dans  $V_h^K$  et  $\mathbf{p}$  appartenant à  $\mathbb{P}_1^d$ . Par définition de  $a^K$ ,

$$\begin{aligned} a^K(\pi^K \mathbf{v}_h - \mathbf{v}_h, \mathbf{p}) &= \int_K \bar{C} \bar{\varepsilon}(\pi^K \mathbf{v}_h - \mathbf{v}_h) : \bar{\varepsilon}(\mathbf{p}) \\ &= \bar{C} \bar{\varepsilon}(\mathbf{p}) : \int_K \bar{\varepsilon}(\pi^K \mathbf{v}_h - \mathbf{v}_h) \text{ car } \bar{C} \text{ et } \bar{\varepsilon}(\mathbf{p}) \text{ sont constants dans } K. \end{aligned}$$

En utilisant les définitions de  $\bar{\epsilon}$  (voir (2.1)) et de  $\pi^K$ ,

$$\bar{\epsilon}(\pi^K \mathbf{v}_h - \mathbf{v}_h) = \frac{1}{2} (\langle \nabla \mathbf{v}_h \rangle + \langle \nabla \mathbf{v}_h^T \rangle - \nabla \mathbf{v}_h - \nabla \mathbf{v}_h^T)$$

et, dans la mesure où

$$\forall \mathbf{v}, \int_K \langle \nabla \mathbf{v} \rangle - \nabla \mathbf{v} = |K| \langle \nabla \mathbf{v} \rangle - \int_K \nabla \mathbf{v} = \mathbf{0} \text{ par définition de } \langle \nabla \mathbf{v} \rangle,$$

on a

$$\int_K \bar{\epsilon}(\pi^K \mathbf{v}_h - \mathbf{v}_h) = \bar{0} \text{ d'où } a^K(\pi^K \mathbf{v}_h, \mathbf{p}) = a^K(\mathbf{v}_h, \mathbf{p}).$$

La propriété (3.9) est donc bien satisfaite.

Il reste à montrer que  $\pi^K$  est calculable pour toute fonction de base de  $V_h^K$  à partir des degrés de liberté de cet espace, qui sont dans le cas considéré les valeurs aux nœuds. La moyenne aux sommets  $\bar{v}$  étant calculable sans difficultés, le seul terme à détailler est le gradient moyen des fonctions de bases  $\langle \nabla \varphi_i^k \rangle$  pour  $k = x, y, z$ . On remarque en premier lieu que ce calcul se ramène à celui du gradient moyen des fonctions de base scalaires selon la relation  $\nabla \varphi_i^k = \mathbf{e}_k \otimes \nabla \varphi_i$ . La suite diffère selon la dimension considérée.

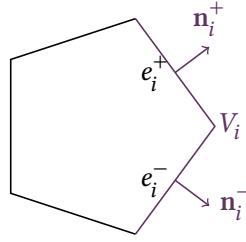


FIG. 3.3 : **Élément polygonal en deux dimensions.** Les sommets et arêtes sont numérotés dans le sens trigonométrique, avec la convention  $V_{K+1} \equiv V_1$ .

En deux dimensions, les sommets sont numérotés dans le sens trigonométrique et on désigne respectivement par  $e_i^-$  et  $e_i^+$  les arêtes connectant  $V_i$  à  $V_{i-1}$  et à  $V_{i+1}$ , et par  $\mathbf{n}_i^-$  et  $\mathbf{n}_i^+$  les normales sortantes à ces arêtes, comme illustré sur la figure 3.3. La formule de Stokes permet de transformer l'intégrale du gradient sur la maille en intégrale de la fonction de base sur les arêtes. Or, les fonctions de bases sont linéaires sur les arêtes et leurs valeurs aux sommets sont connues ( $\varphi_i(V_j) = \delta_{ij}$ ) ce qui permet d'intégrer exactement selon

$$\frac{1}{|K|} \int_K \nabla \varphi_i = \frac{1}{|K|} \int_{\partial K} \varphi_i \mathbf{n} = \frac{1}{2|K|} (|e_i^-| \mathbf{n}_i^- + |e_i^+| \mathbf{n}_i^+),$$

d'où l'expression du projecteur en deux dimensions :

$$\pi^K \varphi_i^k = \frac{1}{M_K} \mathbf{e}_k + \left( \mathbf{e}_k \otimes \frac{1}{2|K|} (|e_i^-| \mathbf{n}_i^- + |e_i^+| \mathbf{n}_i^+) \right) (\mathbf{x} - \bar{\mathbf{x}}_K). \quad (3.17)$$

En trois dimensions, la formule de Stokes est également utilisée pour transformer l'intégrale du gradient en intégrale sur les faces  $f$  de normales sortantes  $\mathbf{n}_f$  selon

$$\frac{1}{|K|} \int_K \nabla \varphi_i = \frac{1}{|K|} \sum_f \int_f \varphi_i \mathbf{n}_f$$

La suite n'est pas directe : si l'on sait que  $\varphi_i$  est linéaire sur chaque arête du polyèdre, on manque en revanche d'informations sur le comportement de  $\varphi_i$  à l'intérieur des faces, et on est donc *a priori* incapable de calculer cette intégrale. C'est à ce stade que le choix effectué pour la définition de l'espace  $V_h^K$  en trois dimensions dans la section 3.1.1 fait sens. Puisque cet espace d'approximation inclut la propriété

$$\int_f \pi^{f,0} v_h = \int_f v_h,$$

on peut remplacer le calcul de l'intégrale de la fonction par le calcul de l'intégrale de sa projection, en espérant que celle-ci soit calculable depuis les degrés de liberté. Dans notre cas, c'est évidemment possible puisque la définition de la projection  $L^2$   $\pi^{f,0}$  donnée par l'équation (3.2) coïncide avec la définition de la projection des éléments virtuels bidimensionnelle donnée par l'équation (3.16). On réutilise donc l'expression obtenue au paragraphe précédent pour écrire

$$\begin{aligned} \int_f \varphi_i &= \int_f \pi^{f,0} \varphi_i = \int_f \left( \frac{1}{M_f} + \frac{1}{2|f|} \left( |e_{f,i}^-| \mathbf{n}_{f,i}^- + |e_{f,i}^+| \mathbf{n}_{f,i}^+ \right) \cdot (\mathbf{x} - \bar{\mathbf{x}}_f) \right) \\ &= \frac{|f|}{M_f} + \frac{1}{2} \left( |e_{f,i}^-| \mathbf{n}_{f,i}^- + |e_{f,i}^+| \mathbf{n}_{f,i}^+ \right) \cdot (\mathbf{x}^f - \bar{\mathbf{x}}_f). \end{aligned}$$

Les notations de la figure 3.3 ont été réutilisées en ajoutant l'indice  $f$  pour souligner le fait que  $e_{f,i}^-$  et  $e_{f,i}^+$  (les longueurs des deux arêtes de la face  $f$  connectées au sommet  $V_i$ ) ainsi que  $\mathbf{n}_{f,i}^-$  et  $\mathbf{n}_{f,i}^+$  (les normales sortantes unitaires à ces arêtes, dans le plan de la face  $f$ ) dépendent de la face considérée. Pour obtenir cette dernière égalité, on a simplement intégré la quantité affine en utilisant la définition du barycentre des faces  $|f| \mathbf{x}^f = \int_f \mathbf{x}$ . Notons que pour des faces simples, par exemple triangulaires ou carrées, le barycentre  $\mathbf{x}^f$  se confond avec le centre moyen  $\bar{\mathbf{x}}_f$ , l'intégrale se réduisant alors à  $\int_f \varphi_i = \frac{|f|}{M_f}$ . Dans le cas général, ces deux points sont distincts. L'intégrale du gradient de la  $i$ -ème fonction de base fait donc intervenir  $F_{i,K}$ , l'ensemble des faces de  $K$  dont  $V_i$  est l'un des sommets, et s'écrit

$$\frac{1}{|K|} \int_K \nabla \varphi_i = \frac{1}{|K|} \sum_{f \in F_{i,K}} \left( \frac{|f|}{M_f} + \frac{1}{2} \left( |e_{f,i}^-| \mathbf{n}_{f,i}^- + |e_{f,i}^+| \mathbf{n}_{f,i}^+ \right) \cdot (\mathbf{x}^f - \bar{\mathbf{x}}_f) \right) \mathbf{n}_f. \quad (3.18)$$

Finalement, en injectant ce gradient moyen dans (3.16), on obtient l'expression du projecteur en trois dimensions pour chaque fonction de base :

$$\pi^K \varphi_i^k = \frac{1}{M_K} \mathbf{e}_k + (\mathbf{x} - \bar{\mathbf{x}}_K) \cdot \left( \mathbf{e}_k \otimes \frac{1}{|K|} \sum_{f \in F_{i,K}} \left( \frac{|f|}{M_f} + \frac{1}{2} \left( |e_{f,i}^-| \mathbf{n}_{f,i}^- + |e_{f,i}^+| \mathbf{n}_{f,i}^+ \right) \cdot (\mathbf{x}^f - \bar{\mathbf{x}}_f) \right) \mathbf{n}_f \right). \quad (3.19)$$

## 3.2 Implémentation

La méthode des éléments virtuels consistant à résoudre un problème variationnel (certes approché, tel que défini par l'équation (3.15a)), la procédure d'assemblage est identique à celle des méthodes éléments finis et est décrite par l'algorithme 1. Les paragraphes suivants détaillent les différentes étapes de cet algorithme.



---

**Algorithm 1** Implémentation des éléments virtuels pour l'élasticité
 

---

```

1 Read mesh and compute each cell  $\diamond$  geometry
2 for  $\diamond$  in mesh do
3   Compute local matrix  $\mathcal{A}^{\diamond}$ 
4   Compute local rhs  $\mathcal{L}^{u,\diamond}$  ▷  $\mathcal{L}^u$  includes forcing term and Neumann Bc
5   for non Dirichlet dof  $\bullet$  in  $\diamond$  do
6      $\mathcal{L}^{u,\bullet} += \mathcal{L}^{u,\diamond}(\bullet)$ 
7   for non Dirichlet  $\bullet$  in  $\diamond$  do
8      $\mathcal{A}(\bullet, \bullet) += \mathcal{A}^{\diamond}(\bullet, \bullet)$ 
9   for Dirichlet  $\bullet$  in  $\diamond$  do
10     $\mathcal{L}^{u,\bullet} -= \mathcal{A}^{\diamond}(\bullet, \bullet)g_D(\bullet)$ 
11 Solve  $d \cdot \text{card}(\mathcal{M}_{D_d}^-)$  linear system  $\mathcal{A}\mathcal{U} = \mathcal{L}^u$ 
12 Complete  $\mathcal{U}$  with  $d \cdot \text{card}(\mathcal{M}_{D_d})$  Dirichlet  $\bullet$  values
13 for  $\diamond$  in mesh do
14   Compute stress vector
15 Compute errors
    
```

---

**Calcul des matrices VEM locales**

Cette section explicite le calcul des matrices élémentaires de la méthode des éléments virtuels, se basant sur les développements de [ANR17]. La contribution de chaque élément  $K$  consiste à calculer la matrice de rigidité locale  $\mathcal{A}^K$  explicitée par l'équation (3.11) pour chacune des  $d \times M_K$  fonctions de base de l'élément.

On développe dans un premier temps le calcul de la partie consistante  $a^K(\pi^K \boldsymbol{\varphi}_i^k, \pi^K \boldsymbol{\varphi}_j^k)$ . Comme expliqué dans le chapitre 2.1, le tenseur des déformations a la propriété d'être nul lorsqu'il est associé à un mouvement de translation ou de rotation. Plus précisément, un calcul direct montre que  $\bar{\boldsymbol{\epsilon}}(\mathbf{u})$  est nul lorsque  $\mathbf{u}$  est de la forme  $\mathbf{u}(\mathbf{x}) = \bar{\boldsymbol{\Omega}}(\mathbf{x} - \mathbf{x}_0) + \mathbf{T}$  avec  $\bar{\boldsymbol{\Omega}}$  matrice antisymétrique de  $\mathbb{R}^{d \times d}$ ,  $\mathbf{x}_0$  et  $\mathbf{T}$  vecteurs de  $\mathbb{R}^d$ . Il est donc possible d'alléger l'évaluation du terme  $a^K(\pi^K \boldsymbol{\varphi}_i^k, \pi^K \boldsymbol{\varphi}_j^k)$  en séparant  $\pi^K \boldsymbol{\varphi}_i^k$  en deux termes, dont l'un s'annule à travers  $a^K$ . Cette procédure, qui est facultative, permet néanmoins d'économiser des calculs mais aussi de travailler sur des quantités symétriques ou anti-symétriques et donc d'utiliser la notation de Kelvin pour condenser les objets utilisés. On réécrit donc (3.16) sous la forme

$$\pi^K = \pi_c^K + \pi_r^K \quad \text{avec} \quad \begin{cases} \pi_c^K \mathbf{v}(\mathbf{x}) = \frac{1}{2} \langle \nabla \mathbf{v} + \nabla \mathbf{v}^T \rangle (\mathbf{x} - \bar{\mathbf{x}}_K) \\ \pi_r^K \mathbf{v}(\mathbf{x}) = \frac{1}{2} \langle \nabla \mathbf{v} - \nabla \mathbf{v}^T \rangle (\mathbf{x} - \bar{\mathbf{x}}_K) + \bar{\mathbf{v}} \end{cases} \quad (3.20)$$

et il est clair que la matrice  $\langle \nabla \mathbf{v} - \nabla \mathbf{v}^T \rangle$  est antisymétrique, de sorte que  $\bar{\boldsymbol{\epsilon}}(\pi_r^K \mathbf{v}) = \bar{\mathbf{0}}$  et, par bilinéarité,  $a^K(\pi_c^K \mathbf{u}, \pi_r^K \mathbf{v}) = a^K(\pi_c^K \mathbf{u}, \pi_c^K \mathbf{v})$ .

Concernant la partie non nulle  $\pi_c^K \mathbf{v}$ , puisque la matrice  $\frac{1}{2} \langle \nabla \mathbf{v} + \nabla \mathbf{v}^T \rangle$  est symétrique, on peut la représenter par un vecteur en adoptant la notation de Kelvin : on note donc  $W_{ci}^k$  le vecteur de  $\mathbb{R}^6$  (ou de  $\mathbb{R}^3$  en deux dimensions) associé à la quantité  $\frac{1}{2} \langle \nabla \boldsymbol{\varphi}_i^k + \nabla \boldsymbol{\varphi}_i^{kT} \rangle$ . Les  $d \times M_K$  vecteurs  $W_{ci}^k$  sont rassemblés dans une matrice  $\mathcal{W}_c = [W_{c1}^x, W_{c1}^y, \dots, W_{cM_K}^z]$  de taille  $6 \times 3M_K$  (ou  $3 \times 2M_K$  en deux dimensions) dont le bloc de colonnes  $(\mathcal{W}_c)_i = [W_{ci}^x, W_{ci}^y, W_{ci}^z]$  est explicité par l'équation (3.21). Pour repasser à l'expression de la composante  $l$  du projecteur  $\pi_c^K(\boldsymbol{\varphi}_i^k)$  évalué au sommet  $V_j$ , le vecteur  $W_{ci}^k$  doit être multiplié à gauche par la

quantité  $(\mathbf{x}_j - \overline{\mathbf{x}_K})^T$  stockée dans un vecteur ligne  $N_{c_j}^l$  de sorte que

$$(\pi_c^K \boldsymbol{\varphi}_i^k(\mathbf{x}_j)) \cdot \mathbf{e}_l = N_{c_j}^l W_{ci}^k.$$

À nouveau, les  $d \times M_K$  vecteurs  $N_c$  sont rassemblés dans une matrice  $\mathcal{N}_c = [N_{c1}^x, N_{c1}^y, \dots, N_{cM_K}^z]$  de taille  $3M_K \times 6$  (ou  $2M_K \times 3$  en deux dimensions) dont le bloc de lignes  $(\mathcal{N}_c)_i$  est également explicité par l'équation (3.21)

$$(\mathcal{W}_c)_i = \begin{pmatrix} q_i^x & 0 & 0 \\ 0 & q_i^y & 0 \\ 0 & 0 & q_i^z \\ 0 & \frac{1}{\sqrt{2}} q_i^z & \frac{1}{\sqrt{2}} q_i^y \\ \frac{1}{\sqrt{2}} q_i^z & 0 & \frac{1}{\sqrt{2}} q_i^x \\ \frac{1}{\sqrt{2}} q_i^y & \frac{1}{\sqrt{2}} q_i^x & 0 \end{pmatrix} \quad \text{et} \quad (\mathcal{N}_c)_i = \begin{pmatrix} r_i^x & 0 & 0 & 0 & \frac{r_i^z}{\sqrt{2}} & \frac{r_i^y}{\sqrt{2}} \\ 0 & r_i^y & 0 & \frac{r_i^z}{\sqrt{2}} & 0 & \frac{r_i^x}{\sqrt{2}} \\ 0 & 0 & r_i^z & \frac{r_i^y}{\sqrt{2}} & \frac{r_i^x}{\sqrt{2}} & 0 \end{pmatrix} \quad (3.21)$$

où l'on a noté

$$\mathbf{q}_i = \langle \nabla \varphi_i \rangle = \frac{1}{K} \int_K \nabla \varphi_i \quad \text{et} \quad \mathbf{r}_i = \mathbf{x}_i - \overline{\mathbf{x}_K}. \quad (3.22)$$

En toute rigueur, les quantités  $\mathcal{W}_c$ ,  $\mathcal{N}_c$ ,  $\mathbf{q}_i$  et  $\mathbf{r}_i$  sont locales à chaque cellule, mais on a omis l'exposant  $K$  pour plus de lisibilité. En utilisant cette décomposition, on a dans la notation de Kelvin  $\hat{\varepsilon}(\pi_c^K \boldsymbol{\varphi}_i^k) = W_{ci}^k$  et donc, pour deux fonctions de base  $\boldsymbol{\varphi}_i^k$  et  $\boldsymbol{\varphi}_j^l$  données,  $a^K(\pi_c^K \boldsymbol{\varphi}_i^k, \pi_c^K \boldsymbol{\varphi}_j^l) = |K| (W_{ci}^k)^T \hat{C} W_{cj}^l$ . Autrement dit, la première partie de la matrice de rigidité élémentaire associée à  $a_h^K$  se calcule par la relation

$$\tilde{\mathcal{A}}^K = |K| \mathcal{W}_c^T \hat{C} \mathcal{W}_c.$$

Il reste à expliciter la matrice associée à  $\pi_r^K$  pour calculer la seconde partie de la matrice de rigidité élémentaire  $\tilde{s}_h^K$ . En effet, cette partie fait intervenir la projection  $\pi^K$  dans son ensemble. On peut utiliser une décomposition similaire et définir les matrices  $\mathcal{W}_r$  et  $\mathcal{N}_r$  construites à partir des blocs  $(\mathcal{W}_r)_i$  et  $(\mathcal{N}_r)_i$  définis par

$$(\mathcal{W}_r)_i = \begin{pmatrix} \frac{1}{M_K} & 0 & 0 \\ 0 & \frac{1}{M_K} & 0 \\ 0 & 0 & \frac{1}{M_K} \\ 0 & \frac{1}{\sqrt{2}} q_i^z & \frac{-1}{\sqrt{2}} q_i^y \\ \frac{-1}{\sqrt{2}} q_i^z & 0 & \frac{1}{\sqrt{2}} q_i^x \\ \frac{1}{\sqrt{2}} q_i^y & \frac{-1}{\sqrt{2}} q_i^x & 0 \end{pmatrix} \quad \text{et} \quad (\mathcal{N}_r)_i = \begin{pmatrix} 1 & 0 & 0 & 0 & \frac{-r_i^z}{\sqrt{2}} & \frac{r_i^y}{\sqrt{2}} \\ 0 & 1 & 0 & \frac{r_i^z}{\sqrt{2}} & 0 & \frac{-r_i^x}{\sqrt{2}} \\ 0 & 0 & 1 & \frac{-r_i^y}{\sqrt{2}} & \frac{r_i^x}{\sqrt{2}} & 0 \end{pmatrix}. \quad (3.23)$$

À nouveau, on a

$$(\pi_r^K \boldsymbol{\varphi}_i^k(\mathbf{x}_j)) \cdot \mathbf{e}_l = N_{r_j}^l W_{ri}^k.$$

De cette manière, la projection  $\pi^K$  s'écrit sous forme matricielle  $\mathcal{P} = \mathcal{N}_r \mathcal{W}_r + \mathcal{N}_c \mathcal{W}_c$  et  $\mathcal{P}_{i_k j_l}$  représente la projection de la  $i_k^{\text{ème}}$  fonction de base évaluée au  $j_l^{\text{ème}}$  degré de liberté. La seconde partie de la matrice de rigidité évaluée sur les fonctions de base  $\boldsymbol{\varphi}_i^k$  et  $\boldsymbol{\varphi}_j^l$  vaut donc  $\tilde{s}_h^K(\boldsymbol{\varphi}_i^k, \boldsymbol{\varphi}_j^l) = s_K \sum_{m \in \mathcal{M}_K} (\delta_{i_k m_q} - \mathcal{P}_{i_k m_q}) (\delta_{j_l m_q} - \mathcal{P}_{j_l m_q})$

car les fonctions de base ne sont non nulles que sur le degré de liberté associé, et la matrice élémentaire de stabilisation s'écrit donc

$$\tilde{\mathcal{S}}_h^K = s_K (\bar{I}_d - \mathcal{P})^T (\bar{I}_d - \mathcal{P}).$$

Finalement, la matrice élémentaire  $\mathcal{A}^K$  s'exprime sur chaque élément  $K$  selon

$$\mathcal{A}^K = |K| \mathcal{W}_c^T \hat{\mathcal{C}} \mathcal{W}_c + s_K (\bar{I}_d - \mathcal{P})^T (\bar{I}_d - \mathcal{P}). \quad (3.24)$$

Enfin, le second membre élémentaire se calcule simplement sur chaque maille en appliquant la formule (3.12) aux fonctions de base de sorte que

$$\mathcal{L}_k^{u,K} = \begin{cases} \omega_i^K \mathbf{f}(V_i) \cdot \mathbf{e}_k + \omega_i^{\partial K} \mathbf{t}_N(V_i) \cdot \mathbf{e}_k & \text{si } V_i \text{ est un nœud de Neumann,} \\ \omega_i^K \mathbf{f}(V_i) \cdot \mathbf{e}_k & \text{sinon,} \end{cases} \quad (3.25)$$

où les coefficients  $\omega_i^K$  et  $\omega_i^{\partial K}$  ont été définis sur la figure 3.2.

### Calcul des contraintes et des écarts

Une fois le système linéaire résolu, les valeurs de  $\mathbf{u}_h$  sur les degrés de liberté sont connues. Sans expression pour les fonctions de base, il n'est cependant pas possible de reconstruire  $\mathbf{u}_h$  dans la cellule pour calculer la contrainte par la loi  $\hat{\sigma}_{\mathbf{u}_h}^e = \hat{\mathcal{C}} \hat{\epsilon}_{\mathbf{u}_h}$ . La meilleure information que l'on puisse obtenir est la déformation de la projection de la solution  $\hat{\epsilon}(\pi^K \mathbf{u}_h) = \mathcal{W}_c^K \mathcal{U}^K$  où la matrice  $\mathcal{W}_c^K$  est définie par (3.21) et  $\mathcal{U}^K$  est le vecteur stockant les valeurs calculées de  $\mathbf{u}_h$  sur la maille  $K$ . Cette approximation du tenseur des déformations peut être interprétée comme en étant une valeur moyenne sur la maille. On peut alors en déduire une valeur moyenne de la contrainte sur la maille selon  $\hat{\sigma}^e(\pi^K \mathbf{u}_h) = \hat{\mathcal{C}} \hat{\epsilon}(\pi^K \mathbf{u}_h)$ . Les déformations et contraintes approchées sont donc constantes par maille.

À partir de ces deux tenseurs, on peut calculer les grandeurs scalaires représentatives définies dans le chapitre 2 qui sont plus facile à représenter et à comparer : la contrainte équivalente de Von Mises, définie par (2.3) et constante sur chaque maille, ainsi que l'énergie de déformation élastique définie par (2.4) qui est une grandeur définie globalement sur le domaine. En pratique, cette énergie est calculée sur chaque maille puis sommée sur l'ensemble des mailles.

Afin de comparer solutions exactes et solutions numériques, on définit les écarts absolus et relatifs en déplacement  $E_{\text{abs}}^{\mathbf{u}}$  et  $E_{\text{rel}}^{\mathbf{u}}$  – comparables à des écarts dans l'espace  $L^2$  – ainsi que les écarts absolus et relatifs en contrainte  $E_{\text{abs}}^{\sigma}$  et  $E_{\text{rel}}^{\sigma}$  – comparables à des écarts dans l'espace  $H^1$  – par les expressions

$$E_{\text{abs}}^{\mathbf{u}} = \|\mathbf{u}_h - \mathbf{u}\|_{L^2}, \quad E_{\text{rel}}^{\mathbf{u}} = \frac{\|\mathbf{u}_h - \mathbf{u}\|_{L^2}}{\|\mathbf{u}\|_{L^2}}, \quad E_{\text{abs}}^{\sigma} = \|\bar{\sigma}^e \pi^K(\mathbf{u}_h) - \bar{\sigma}^e_{\mathbf{u}}\|_{L^2} \quad \text{et} \quad E_{\text{rel}}^{\sigma} = \frac{\|\bar{\sigma}^e \pi^K(\mathbf{u}_h) - \bar{\sigma}^e_{\mathbf{u}}\|_{L^2}}{\|\bar{\sigma}^e_{\mathbf{u}}\|_{L^2}}. \quad (3.26)$$

Pour évaluer les intégrales dans les produits scalaires  $L^2$ , la formule d'intégration numérique décrite sur la figure 3.2 est utilisée. Dans le cas de l'erreur en contrainte, la projection de la solution calculée a été utilisée conformément au commentaire du paragraphe précédent.

### 3.3 Paramétrage et validation numérique

Dans cette section, on achève de définir la méthode numérique en choisissant le paramètre de stabilisation  $s_K$ . On effectue également quelques tests simples qui permettent de valider l'implémentation, et de retrouver numériquement l'ordre de convergence du schéma. Dans ce cadre, on utilise différentes familles de maillages représentées sur la figure 3.4, chacune d'entre elles comprenant des maillages plus ou moins fins.

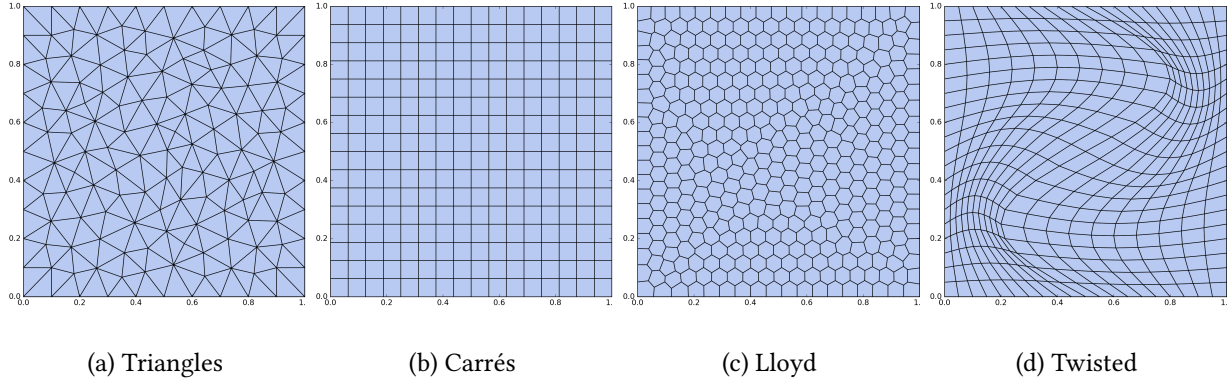


FIG. 3.4 : Familles de maillages du carré unité 2D utilisées dans les tests numériques.

**Dégénérescence en éléments finis** Avant toute chose, un moyen relativement simple de contrôler partiellement l'implémentation de la méthode consiste à utiliser un maillage triangulaire. Sur ce type d'éléments, l'espace d'approximation  $V_h^K$  est réduit à  $\mathbb{P}_1^d$ , et les fonctions de  $V_h^K$  sont donc égales à leur projection par  $\pi^K$ . Par conséquent, la forme bilinéaire de stabilisation est nulle. Dans ce cas particulier, la matrice des éléments virtuels est *rigoureusement* identique à celle qui aurait été obtenue par la méthode des éléments finis P1. Il est alors possible de vérifier le calcul des coefficients de la matrice de rigidité en utilisant un code élément finis tiers. Des comparaisons ont été effectuées avec le logiciel FreeFem++ [Hec12], et, en dehors de faibles écarts pouvant s'expliquer par un traitement différent des conditions aux limites, les résultats obtenus avec le code éléments virtuels concordent avec ceux du code élément finis.

### 3.3.1 Choix du coefficient de stabilisation

Dans la méthode des éléments virtuels, le choix de la stabilisation est à la fois important et secondaire. Secondaire, parce que la consistance de la méthode est assurée par la première partie de la forme bilinéaire approchée : la précision reste assurée tant que la stabilisation est du même ordre de grandeur que la partie consistante (un coefficient de stabilisation démesurément grand ou petit conduit à des instabilités dans la solution). Important, car un choix réfléchi permet d'assurer certaines propriétés supplémentaires. Par exemple, un choix différent de celui effectué pour  $\tilde{s}_h^K$  dans l'équation (3.10) peut améliorer la méthode dans le cas où l'ordre polynomial d'approximation est élevé [BDR17]. À l'ordre un, il semblerait que l'amélioration soit trop faible et injustifiée au regard de la difficulté à implémenter une formulation différente pour  $\tilde{s}_h^K$ , par rapport à l'expression de base qui est elle simple à calculer [Mas18]. Tout en restant sur le choix standard pour  $\tilde{s}_h^K$ , il reste la possibilité d'ajuster le coefficient réel  $s_K$  pour modifier le comportement du terme de stabilisation. En particulier, on cherche ici à conserver la stabilité de la méthode sur des maillages *d'aspect ratio* défavorable, c'est à dire dont les mailles sont particulièrement allongées. Plusieurs coefficients de stabilisation ont été proposés dans la littérature, reposant sur l'idée heuristique que celui-ci doit correspondre à l'ordre de grandeur des coefficients de  $\hat{C}$  et/ou des caractéristiques du maillage. On propose de comparer ici les trois approches suivantes utilisées pour la mécanique :

1. Le coefficient est choisi dans [GTP14] de manière à ce que, pour des déformations uniformes, l'énergie exacte calculée par la forme  $a^K$  et l'énergie approchée calculée par la forme  $\tilde{s}_h^K$  soient comparables pour les fonctions dont la projection  $\pi^K \mathbf{v}$  est nulle. Pour satisfaire cette condition, le coefficient  $s_K$

est défini par

$$s_K = \frac{|K|\text{tr}(\hat{C})}{\text{tr}(\mathcal{N}_c^T \mathcal{N}_c)}.$$

2. Les auteurs de [ANR17] remarquent que le coefficient précédent est instable vis-à-vis de la distorsion : l'écart entre la solution exacte et la solution calculée se détériore lorsque le ratio d'aspect des mailles augmente, c'est-à-dire lorsque celles-ci « s'aplatissent ». Pour corriger cette mauvaise propriété, les auteurs proposent l'expression

$$s_K = \frac{|K|\text{tr}(\hat{C})\text{tr}((\mathcal{N}_c^T \mathcal{N}_c)^{-1})}{9}.$$

3. En parallèle, les auteurs de [BLM15], dans le cadre plus général de problèmes non élastiques, proposent un coefficient dont la simplification au régime élastique linéaire donne

$$s_K = h_K^{d-2} \max |C_{ijkl}|$$

Afin de tester la stabilité de ces coefficients vis-à-vis de la distorsion, on considère trois maillages de base en deux dimensions de forme carrée, Lloyd ou twisted d'environ 400 éléments chacun. Les maillages triangulaires ne sont pas considérés puisque la forme de stabilisation est nulle sur les mailles triangulaires. Ces maillages sont allongés en multipliant l'ensemble de leurs coordonnées horizontales  $x$  par un facteur  $m_x$  variant entre 1 et 50, le nombre de mailles restant inchangé, voir par exemple la figure 3.5 pour une illustration sur le maillage de Lloyd.

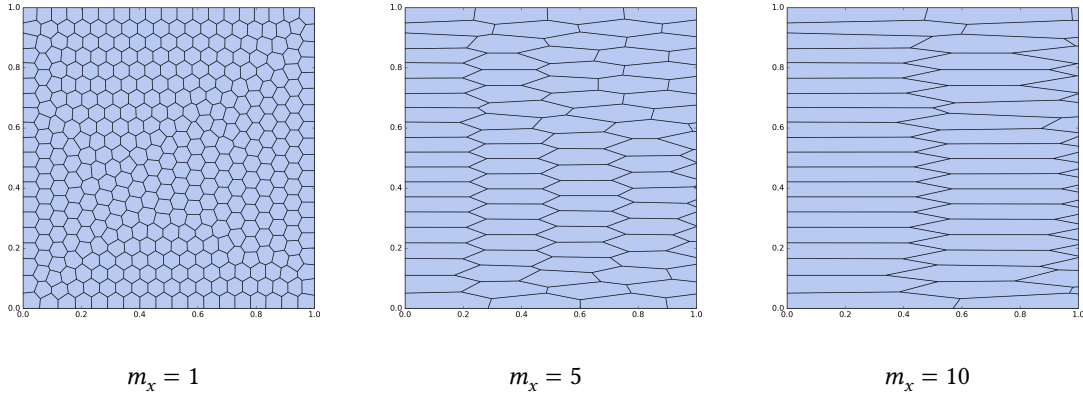


FIG. 3.5 : **Maillages de Lloyd distordus** en multipliant la coordonnée  $x$  des sommets par un facteur  $m_x$ . Pour que les cellules restent visibles sur les maillages distordus, seul le rectangle  $[0, 1] \times [0, 1]$  du domaine complet  $[0, m_x] \times [0, 1]$  est représenté.

Le cas mécanique considéré ici est fréquemment utilisé pour la validation de codes éléments finis [FV05]. On considère un domaine carré de côtés de longueur unitaire représenté sur la figure 3.6, constitué d'un matériau isotrope homogène. Dans le cadre des déformations planes, le tenseur des rigidités  $\hat{C}$  est donné par l'équation (2.9). Les coefficients élastiques sont fixés à  $E = 5$  GPa et  $\nu = 0.3$ . Pour ce problème, on suppose que le champ de déplacements est donné par  $\mathbf{u} = (0.01xy, -0.02xy)$  ce qui permet de calculer analytiquement la force constante appliquée  $\mathbf{f} = (0.02(G + \lambda), -0.01(G + \lambda))$  et les conditions aux limites à imposer, représentées sur la figure 3.6.

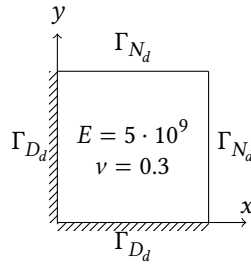


FIG. 3.6 : Cas de calcul élastique en dimension 2 avec représentation des conditions aux limites et des paramètres mécaniques.

L'erreur en déplacement  $E_{rel}^u$  en fonction du coefficient de distorsion  $m_x$  est représentée sur la figure 3.7. Le second coefficient, introduit pour améliorer le premier sur de tels maillages, est effectivement plus stable sauf sur le maillage carré, ce problème ayant été relevé par les auteurs et relié à des effets de symétries. Sur ce maillage carré, l'erreur d'ailleurs est d'un ordre de grandeur plus faible que sur les deux autres, pour chacun des trois coefficients. Le troisième choix de coefficient permet d'obtenir de bons résultats sur les trois familles de maillages et sera donc celui utilisé, sauf mention spécifique, dans les autres cas numériques.

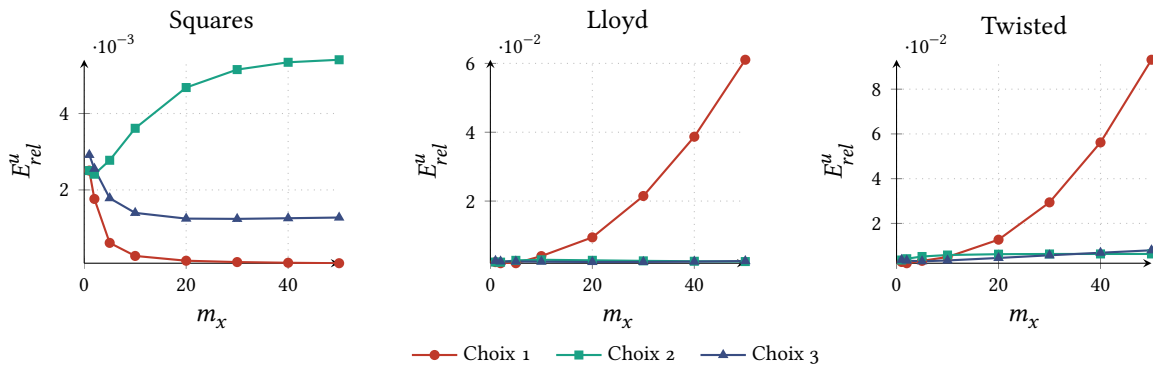


FIG. 3.7 : Écart L2 en déplacement en fonction de la distorsion du maillage pour les trois coefficients de stabilisation  $s_K$ . Le coefficient 3 semble être plus robuste que les deux autres.

**Patch test** La forme bilinéaire approchée étant à présent totalement déterminée, une autre propriété de la méthode peut être utilisée pour contrôler l'implémentation. Lorsque le champ de déplacements  $\mathbf{u}$  exact est composé de fonctions  $\mathbb{P}_1^d$ , l'erreur numérique du schéma doit être nulle (aux erreurs machines près), et ce quelque soit le maillage considéré. Numériquement, cette propriété est bien respectée sur les différents maillages bidimensionnels représentés plus haut. De la même manière, le patch test est satisfait sur des maillages tridimensionnels de différentes nature, voir par exemple la figure 3.8 pour une illustration sur un maillage Voronoï de la solution linéaire  $\mathbf{u} = 10^{-2}z\mathbf{e}_z$ . Même si ceci ne permet pas d'affirmer que le code est exempt de problèmes, une erreur dans l'assemblage, dans les calculs géométriques ou dans ceux des projections par exemple aurait été directement détectée et aurait stoppé la poursuite des tests numériques avant sa résolution.

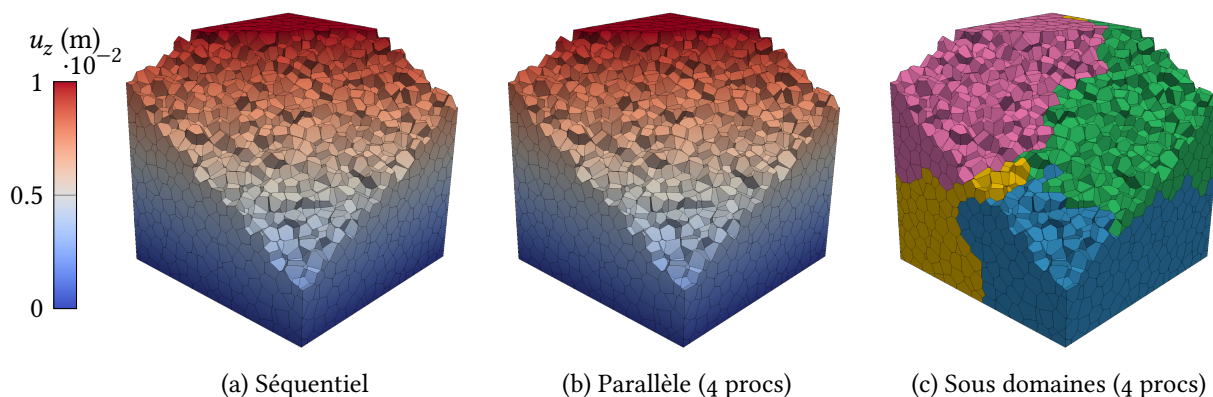


FIG. 3.8 : **Patch test** sur un maillage de Voronoï 3D. La solution analytique, linéaire, vaut  $\mathbf{u} = 10^{-2}z\mathbf{e}_z$  et les coefficients élastiques sont  $E = 1$ ,  $\nu = 0.25$ . Les figures (a) et (b) représentent le champ de déplacement  $u_z$ . La figure (c) représente le partitionnement utilisé pour la simulation parallèle. En séquentiel comme en parallèle, l'écart  $E_{\text{abs}}^{\mathbf{u}}$  entre la solution calculée et la solution exacte est de l'ordre de  $10^{-16}$ .

### 3.3.2 Convergence en espace

Toujours sur les familles de maillages bidimensionnels et avec le cas mécanique défini précédemment (figure 3.6), on représente pour chaque famille de maillage les écarts relatifs en déplacement et en contrainte en fonction du diamètre maximal des éléments du maillage sur la figure 3.9. En échelle logarithmique, les courbes sont des droites dont le coefficient directeur est lié à l'ordre de convergence de la méthode. On constate que les coefficients directeurs de chaque droite sont identiques et égaux à la valeur attendue : deux en déplacement et un en contrainte [BBM13].

Pour effectuer la même étude en trois dimensions, les familles de maillages représentées sur la figure 3.10 sont considérées. Une fois encore, différents niveaux de raffinement sont disponibles pour chacune de ces familles. Le cas d'élasticité homogène est étendu de la même manière que dans [Mar12] en prenant  $\mathbf{u} = (0.01xyz, -0.02xyz, -0.01xyz)$ . Le second membre, cette fois non constant, a alors pour expression  $\mathbf{f} = 0.01(G + \lambda)(y + 2z, x - z, 2x - y)$ . Les conditions aux limites du cas bidimensionnel sont étendues aux faces, et on adopte pour la troisième dimension une condition de Dirichlet sur la face  $z = 0$  et une condition de Neumann sur la face  $z = 1$  : on impose donc un déplacement fixé  $\mathbf{u} = \mathbf{0}$  sur les faces  $x = 0$ ,  $y = 0$  et  $z = 0$  et une contrainte fixée (calculée à partir de la solution exacte) sur les plans  $x = 1$ ,  $y = 1$  et  $z = 1$ . Les courbes d'erreur en fonction du pas du maillage pour les maillages tridimensionnels décrits sur la figure 3.10 sont tracées sur la figure 3.11. Les valeurs des coefficients directeurs sont à nouveau celles attendues.

## 3.4 Observations supplémentaires sur la méthode

Dans la section précédente, quelques cas numériques de base ont permis d'achever le paramétrage de la méthode et de retrouver numériquement l'ordre de convergence de celle-ci. On cherche à présent à illustrer quelques propriétés moins classiques, mais néanmoins désirables. Ces propriétés sont présentées plutôt dans un esprit qualitatif que dans un cadre formel.

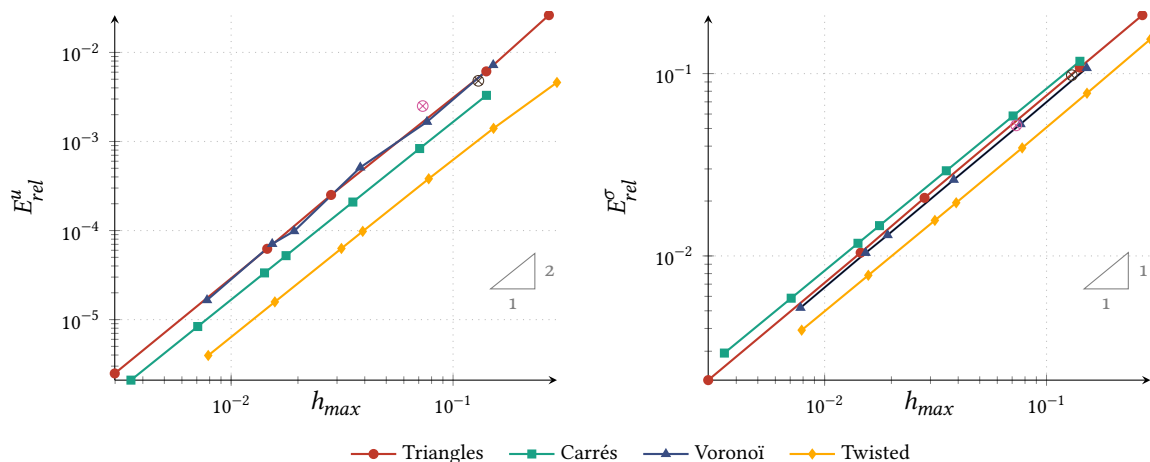


FIG. 3.9 : **Courbes de convergence** sur différents maillages pour le cas du carré homogène. Toutes les familles de maillages présentent les ordres attendus. Pour les deux points solitaires, voir section 3.4.1.

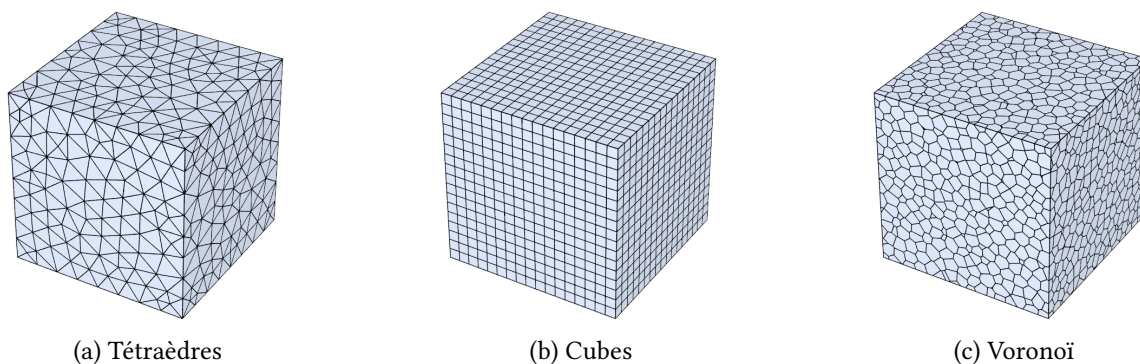


FIG. 3.10 : **Familles de maillages** du cube unité 3D utilisées dans les tests numériques. Les maillages de Voronoï sont générés à l'aide de l'outil Neper [QDB11].

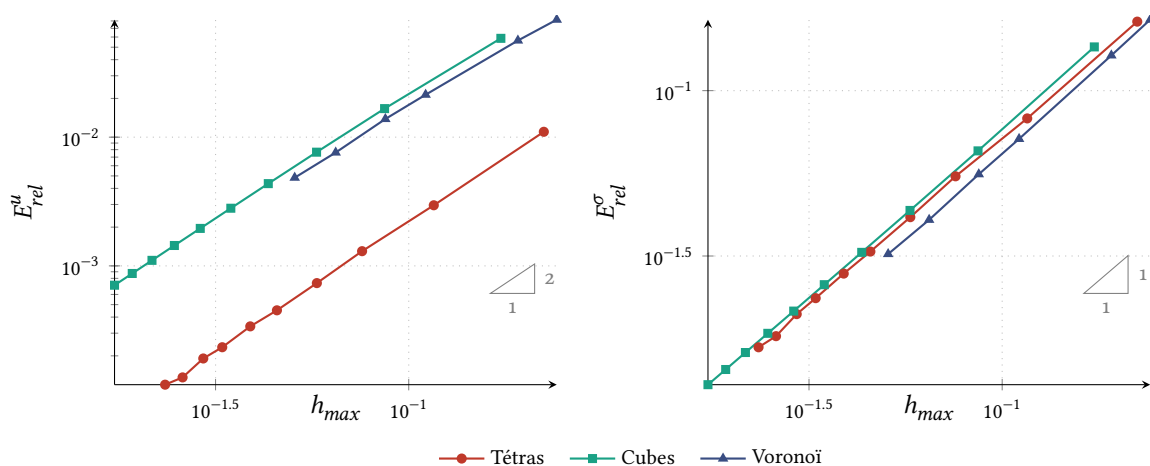


FIG. 3.11 : **Courbes de convergence** sur différents maillages pour le cas du cube homogène. Toutes les familles de maillages présentent les ordres attendus.



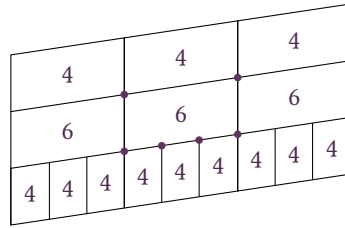


FIG. 3.12 : **Zoom sur la non conformité du maillage.** Chaque maille est naturellement traitée comme un polygone à 4 ou 6 sommets.

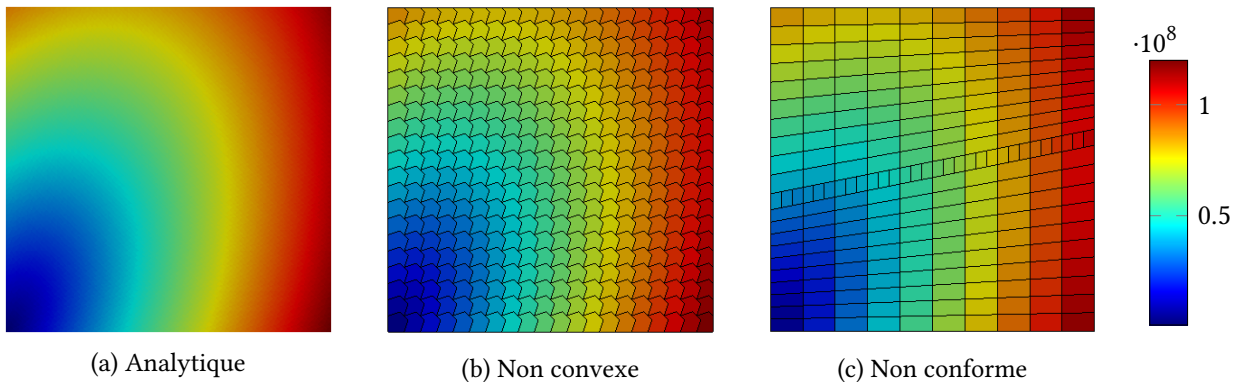


FIG. 3.13 : **Contrainte de Von Mises** calculée sur des maillages non convexes et non conforme. Les écarts par rapport à la solution exacte sont reportées sur la figure 3.9.

### 3.4.1 Traitement de maillages particuliers

Toujours dans le cas du problème homogène bidimensionnel, on considère des maillages moins classiques : un maillage non convexe et un maillage non conforme. On note que le maillage non conforme s'utilise de manière naturelle en traitant les mailles voisines de la bande centrale comme des polygones à 6 sommets (voir figure 3.12) : il n'y a pas de traitement particulier des *hanging nodes* comme en éléments finis classiques, ce qui laisse la possibilité d'effectuer facilement du raffinement local.

Après résolution du problème sur ces maillages, on représente la contrainte équivalente de Von Mises définie par l'équation (2.3) sur les figures 3.13b et 3.13c. La figure 3.13a représente à titre de comparaison la contrainte de Von Mises calculée à partir de la solution analytique. La valeur des champs calculés est qualitativement correcte sur les deux maillages. À défaut d'effectuer une véritable étude de convergence en espace, les couples  $(h_{max}, E_{rel}^u)$  et  $(h_{max}, E_{rel}^\sigma)$  calculés sur ces exemples permettent de situer les points sur le graphique 3.9, en rose pour le maillage non convexe et en marron pour le maillage non conforme. On observe alors que pour chacun des graphiques, ces deux points se situent quasiment sur les courbes obtenues avec les autres types de maillages. L'ordre de grandeur de l'erreur n'est donc pas fortement pénalisé par la forme singulière de ces maillages. Notons également que ces maillages sont relativement sobres par rapport à ceux usuellement utilisés pour illustrer la grande tolérance des VEM, que l'on trouvera par exemple dans [PG15].

### 3.4.2 Comparaison aux éléments finis

Pour positionner la méthode par rapport à celle des éléments finis, on considère un cas bidimensionnel de type poutre, ce qui est courant dans la communauté des mécaniciens, schématisé sur la figure 3.14. Fixée par son extrémité gauche, elle subit une flexion parabolique appliquée sur son extrémité droite. Pour

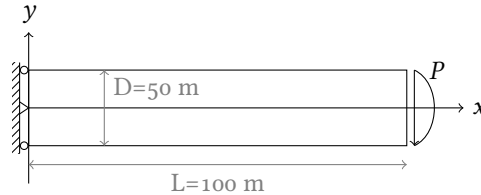


FIG. 3.14 : **Schématisme** d'un cas poutre plane, avec chargement parabolique. Ce type de problèmes admet une distribution plane de contraintes.

ce type de problèmes, une approche en contraintes planes peut être adoptée. Le tenseur des rigidités du matériau, homogène de constantes mécaniques  $E = 210$  GPa et  $\nu = 0.3$ , est donc donné par l'équation (2.10). Les caractéristiques géométriques sont  $L = 500$  m et  $D = 50$  m, et permettent de calculer le moment d'inertie selon la relation  $I = \frac{D^3}{12}$ . Le chargement  $P$  parabolique, maximal de valeur  $p = -10^{-4}$  en  $y = 0$  et nul en  $y = \pm \frac{D}{2}$ , s'écrit

$$P = \frac{-F}{2I} \left( \frac{D^2}{4} - y^2 \right) \quad \text{où} \quad F = -\frac{2}{3} p D = \frac{10^6}{3}$$

et est appliqué dans la direction  $\mathbf{e}_y$ . La solution analytique de ce problème [TG51] est donnée par

$$u_x = \frac{Fy}{6EI} \left( (6L - 3x)x + (2 + \nu)(y^2 - \frac{D^2}{4}) \right), \quad u_y = \frac{-F}{6EI} \left( 3\nu y^2(L - x) + (4 + 5\nu) \frac{D^2 x}{4} + (3L - x)x^2 \right),$$

et permet de calculer les déformations et les contraintes associées à cette configuration. Pour comparer la méthode à celle des éléments finis, la poutre est maillée avec des éléments rectangulaires ou triangulaires. Dans les deux cas, deux familles de maillages sont générées : une famille de maillages peu allongés, de ratio (1 : 2), et une famille de maillages très allongés, de ratio (1 : 1250), représentés sur la figure 3.15.

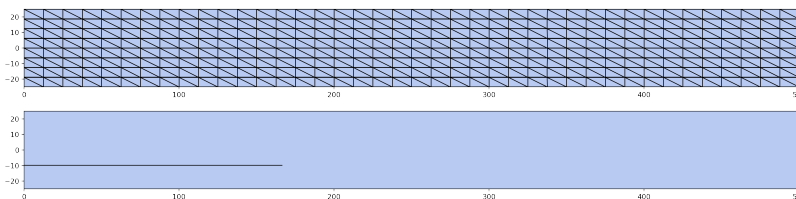


FIG. 3.15 : **Exemples de maillages au ratio 1 : 2 (haut) et au ratio 1 : 1250 (bas)**. Le maillage allongé est constitué de  $3 \times 375$  rectangles, mais seul l'un d'entre eux est représenté pour obtenir une illustration visible.

Le déplacement vertical de la fibre neutre (l'axe  $y = 0$ ) de la poutre, aussi appelé flèche, est représenté pour quelques maillages rectangulaires sur la figure 3.16. La légende indique le nombre d'éléments dans la direction  $\mathbf{e}_x$ . Les deux méthodes convergent vers la solution exacte lorsque la taille des mailles diminue. Sur

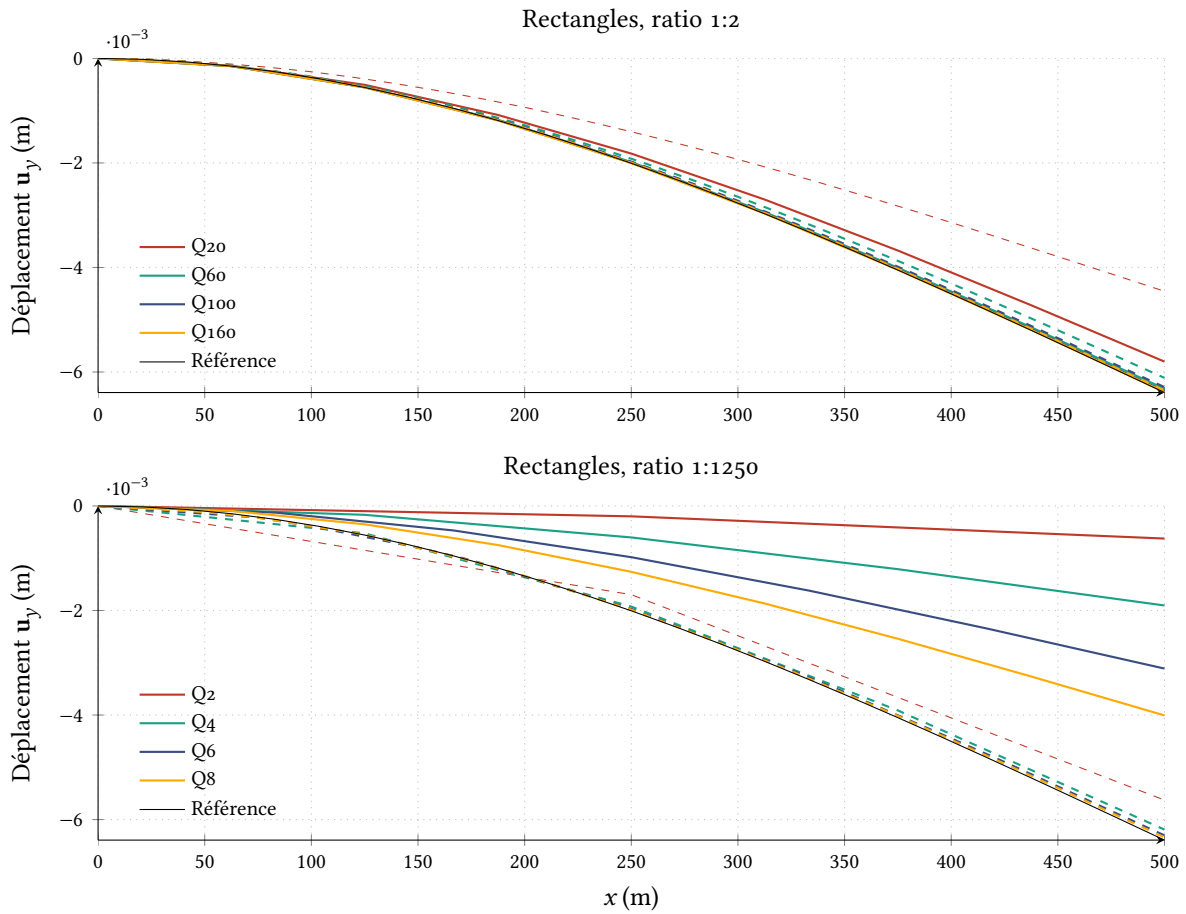


FIG. 3.16 : **Comparaison des éléments finis  $P_1$  (traits pleins) et des éléments virtuels (tirets)** sur les maillages quadrangulaires (Q), la légende indiquant le nombre de mailles dans la direction  $e_x$ . Sur les maillages allongés, les éléments virtuels convergent bien plus rapidement.

la famille de maillages peu allongés, les résultats sont quasiment identiques pour les deux méthodes, excepté sur les maillages les plus grossiers où les éléments virtuels donnent une moins bonne approximation. En revanche, sur les maillages fortement allongés, le résultat des éléments finis est largement plus éloigné de la solution exacte, et cette méthode converge lentement. La méthode des éléments virtuels, elle, donne un bien meilleur résultat même si le faible nombre d'éléments en  $x$  entraîne une mauvaise interpolation visible au niveau des  $x$  faibles. Cette figure illustre une seconde fois l'importance du terme de stabilisation, et confirme que le choix effectué dans la section 3.3.1 permet d'avoir une méthode robuste vis-à-vis de la déformation du maillage.

Sur les maillages triangulaires, comme expliqué précédemment, les éléments virtuels dégénèrent en éléments finis ce qui rend la comparaison inutile. En revanche, on propose ici de tirer à nouveau parti de la flexibilité autorisée dans les maillages pour améliorer la solution. Le maillage de base, constitué de triangles, est transformé en introduisant des nœuds supplémentaires au centre des arêtes comme schématisé sur la figure 3.17. Ainsi, les éléments ne sont plus vus comme des triangles mais comme des hexagones. En conséquences, l'espace d'approximation n'est plus réduit à  $P_1^d$  et le terme de stabilisation n'est plus identiquement nul. Cela permet d'obtenir des résultats différents des éléments finis, représentés sur la figure



FIG. 3.17 : **Transformation du type des cellules.** En ajoutant des sommets supplémentaires, chaque élément est considéré comme un hexagone à six côtés et non comme un triangle à 3 côtés.

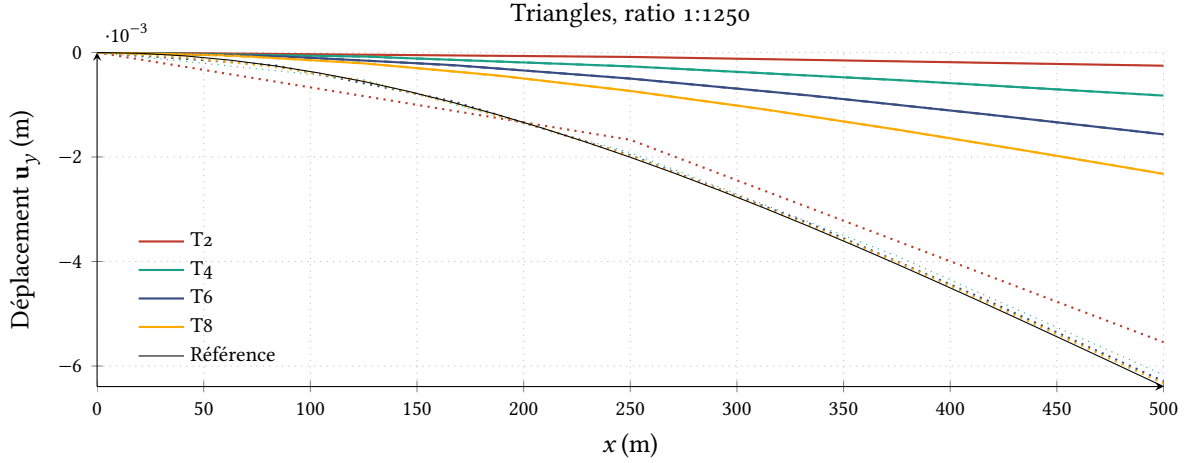


FIG. 3.18 : **Comparaison des éléments finis P1 (traits pleins), des éléments virtuels (tirets), et des éléments virtuels avec hexagones (pointillés)** sur les maillages triangles (T), la légende indiquant le nombre de mailles dans la direction  $\mathbf{e}_x$ . Sans modification du maillage, traits pleins et tirets sont confondus. La modification en hexagones permet d'améliorer les résultats.

3.18 pour les maillages triangles allongés. Évidemment, cette manipulation augmente le nombre de degrés de liberté, mais elle permet d'obtenir une solution numérique de bien meilleure qualité, tout en conservant un ordre d'approximation polynômial égal à un. On aurait d'ailleurs pu se contenter de n'ajouter qu'un nœud par triangle, par exemple sur les arêtes diagonales. Avec les éléments finis, un moyen d'améliorer le résultat aurait été de considérer des éléments P2, ce qui aurait également augmenté le nombre de degrés de liberté. Cet exemple illustre donc à nouveau une utilisation possible de la liberté sur les maillages offerte par la méthode.

### 3.4.3 Comportement à la limite incompressible

Une autre propriété intéressante des éléments virtuels est leur comportement lorsque le matériau est à la limite du domaine incompressible. Les matériaux presque incompressibles se caractérisent par un coefficient de Poisson  $\nu$  tendant vers  $\frac{1}{2}$ , ou, de manière équivalente, par un coefficient  $\lambda$  tendant vers  $+\infty$ . Pour comprendre ce que cela implique, la forme bilinéaire  $a^\Omega(\mathbf{u}, \mathbf{v}) = \int_\Omega \overline{\overline{\mathbf{C}}} \overline{\overline{\boldsymbol{\varepsilon}}}_u : \overline{\overline{\boldsymbol{\varepsilon}}}_v$  du problème élastique est réécrite sous sa forme équivalente

$$a^\Omega(\mathbf{u}, \mathbf{v}) = 2G \int_\Omega \overline{\overline{\boldsymbol{\varepsilon}}}_u : \overline{\overline{\boldsymbol{\varepsilon}}}_v + \lambda \int_\Omega \operatorname{div}(\mathbf{u}) \operatorname{div}(\mathbf{v}). \quad (3.27)$$

Pour que l'énergie reste bornée lorsque  $\lambda \rightarrow +\infty$ , il est nécessaire que  $\operatorname{div}(\mathbf{u}) \rightarrow 0$ . Cela est problématique en éléments finis classiques sur des triangles ou des rectangles car les seuls éléments appartenant

à l'espace d'approximation *et* dont la divergence est nulle sont des constantes, voir [BS08, chap. 11] ou [PW09]. À cause de cette dégénérescence de l'espace d'approximation, les résultats numériques obtenus sont fortement détériorés, et on emploie le terme *locking* pour qualifier ce phénomène. Une autre manière de voir les choses consiste à examiner les estimations d'erreurs habituellement établies pour les méthodes numériques, telle que celle formulée pour les éléments virtuels dans le lemme 3.1 : la constante contrôlant l'erreur explose lorsque les paramètres physiques du problème deviennent trop grands, et l'erreur entre la solution exacte et la solution numérique n'est plus contrôlée. Par ailleurs, des phénomènes de type locking peuvent également apparaître pour la même raison lors de la résolution de problèmes poroélastiques, selon les valeurs prises par le coefficient  $c_0$ , la matrice de mobilité  $\bar{\kappa}$  et le pas de temps. C'est pourquoi il est nécessaire d'être averti de l'existence de cette artéfact numérique et des éventuelles solutions à mettre en place.

Il a été remarqué qu'avec la méthode des éléments virtuels et dans le cas du problème élastique, deux paramètres peuvent être ajustés de manière à atténuer ce problème. Sans fournir d'explication rigoureuse, la suite de cette section expose ces observations. Le cas illustratif est un problème de cavité entraînée issu de la thèse [Lem13]. Dans un domaine carré de taille unitaire, un déplacement est imposé sur le bord supérieur  $y = 1$  par la condition de Dirichlet  $\mathbf{g}_D = (1, 0)$  tandis que les trois autres bords sont fixés ( $\mathbf{g}_D = \mathbf{0}$ ). Le second membre  $\mathbf{f}$  est nul, et les coefficients élastiques ont pour valeur  $E = 1000$  Pa et  $\nu = 0.4999$ . La solution de ce problème n'est analytiquement pas connue. Une solution de référence est donc calculée sur un maillage triangulaire très fin, de plus de cent mille sommets. En dehors de ce maillage de référence, les discrétisations utilisées sont décrites dans le tableau 3.1. Le nombre d'inconnues est volontairement faible

	Triangles	Carrés	Voronoi
# sommets	481	441	402
# cellules	896	400	200
$h_{max}$	$6.25 \cdot 10^{-2}$	$7.07 \cdot 10^{-2}$	$9.17 \cdot 10^{-2}$

TAB. 3.1 : Maillages utilisés pour la mise en évidence du locking sur le cas issu de [Lem13].

pour que la résolution du maillage ne fasse pas disparaître l'effet de locking. On tente également d'avoir un diamètre  $h$  comparable sur les différents maillages. Les maillages sont comparés en représentant l'une ou l'autre des variables calculées suivantes :

- la composante  $u_x$  du déplacement en fonction de  $y$  sur la ligne verticale  $x = \frac{1}{2}$  ;
- la composante  $u_y$  du déplacement en fonction de  $x$  sur la ligne verticale  $y = \frac{1}{2}$ .

**Augmentation du nombre de degrés de liberté par maille** Formellement, il suffit pour éviter le locking d'utiliser un espace d'approximation local contenant des fonctions à divergence nulle et non constantes. En éléments virtuels, plus l'élément considéré a de sommets, plus la taille de l'espace d'approximation local est grande. Il semble que dans ce cas, des fonctions de base possédant les propriétés requises existent dans l'espace d'approximation. La figure 3.19 expose le champ de déplacement calculé sur les maillages du tableau 3.1. Sur les deux graphes, on remarque que la solution calculée sur le maillage triangulaire ou sur le maillage carré est de mauvaise qualité, même si leur comportement semble différent : dans le premier cas, la solution a tendance à osciller, tandis que dans le second cas, la solution a tendance à rester constante. Pour la composante  $u_x$  du déplacement, le maillage de Voronoï permet d'obtenir une solution numérique nettement plus proche de la référence. C'est plus discutable pour la composante  $u_y$

du champ de déplacement. Le lien entre le nombre de degrés de liberté et le locking est approfondi dans [BLM14, chap. 8] pour la méthode des différences finies mimétiques qui est à l'origine de celle des éléments virtuels. Dans cet ouvrage, les auteurs recommandent l'ajout d'un degré de liberté sur les arêtes lorsqu'elles sont plus de trois à partager un même sommet, sous peine de pâtir du locking. Ici, les sommets des maillages carrés et triangles sont quasiment tous le point de rencontre de plus de trois arêtes (quatre dans le cas des carrés, jusqu'à six dans le cas des triangles). En revanche, pour les maillages de Voronoï, le nombre d'arêtes connectant chaque sommet est en général de trois. L'étude effectuée dans l'ouvrage sus-cité pourrait donc expliquer que le locking soit moins visible sur ce maillage.

**Changement du terme de stabilisation** Le second paramètre permettant de réduire le phénomène de locking est le coefficient  $s_K$  utilisé en facteur de la forme de stabilisation et explicité dans la partie 3.3.1. Jusqu'alors, le choix retenu en deux dimensions était  $s_K = \max |C_{ijkl}|$ , ce qui équivaut pour le problème d'élasticité isotrope à  $s_K = \lambda + 2G$ . Lorsque  $\lambda \rightarrow +\infty$ , le terme de stabilisation devient trop important par rapport au terme de consistance. Une stratégie peut alors consister à borner artificiellement le terme de stabilisation, voire à ne pas tenir compte de la partie en  $\lambda$  : on prend alors  $s_K = 2G$  et on représente à nouveau le déplacement  $\mathbf{u}_y$  sur la figure 3.20. Puisque le coefficient ne s'applique pas aux maillages triangulaires, seuls les maillages carrés et Voronoï sont testés ici. Pour ces deux maillages, on observe une amélioration significative du résultat lorsque cette technique est utilisée.

Il semblerait qu'on puisse établir un lien entre cette technique et certaines formulations mixtes du problème d'élasticité linéaire utilisées en éléments finis, connues pour éviter le problème de locking. En effet, ne stabiliser que la partie en  $G$  revient à écrire sur chaque maille

$$a^K(\mathbf{u}_h, \mathbf{v}_h) = 2G \int_K \bar{\epsilon}(\pi^K \mathbf{u}_h) : \bar{\epsilon}(\pi^K \mathbf{v}_h) + 2G \sum_{i \in \mathcal{M}_K} (\mathbf{u}_h - \pi^K \mathbf{u}_h)(V_i) \cdot (\mathbf{v}_h - \pi^K \mathbf{v}_h)(V_i) + \lambda \int_K \operatorname{div}(\pi^K \mathbf{u}_h) \operatorname{div}(\pi^K \mathbf{v}_h).$$

D'un autre coté, en utilisant une formulation mixte, le problème faible (3.27) peut se réécrire sous la forme

Trouver  $(\mathbf{u}, p) \in (V_g \times L^2(\Omega))$  tels que pour tout  $(\mathbf{v}, q) \in (V_0 \times L^2(\Omega))$ ,

$$\begin{aligned} \int_{\Omega} 2G \bar{\epsilon}_u : \bar{\epsilon}_v + \int_{\Omega} p \operatorname{div}(\mathbf{v}) &= \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \\ \int_{\Omega} (p - \lambda \operatorname{div}(\mathbf{u})) q &= 0. \end{aligned}$$

Pour le discrétiser, on cherche  $p_h$  dans un espace de dimension inférieure à celui dans lequel on cherche  $\mathbf{u}_h$ . Typiquement, on peut chercher  $\mathbf{u}_h$  dans  $V_{h,g}$  et  $p_h$  dans l'espace des constantes. Dans ce cas, en utilisant la projection  $L^2$  orthogonale  $\pi^0$  qui vérifie  $\int_K p_h = \int_K \pi^0 p_h$ , la seconde ligne permet d'obtenir

$$\int_K \lambda \operatorname{div}(\mathbf{u}_h) = \int_K p_h = \int_K \pi^0 p_h = \int_K \lambda \pi^0(\operatorname{div}(\mathbf{u}_h)).$$

De plus, en réutilisant l'égalité  $\operatorname{div}(\mathbf{u}_h) = \operatorname{tr}(\nabla \mathbf{u}_h)$ , l'expression du projecteur  $\pi^K$  donnée par l'équation (3.16), et la ligne précédente, on montre que

$$\forall \mathbf{v}_h \in V_{h,g}, \quad \int_K \operatorname{div}(\pi^K \mathbf{v}_h) = \int_K \operatorname{div}(\mathbf{v}_h) = \int_K \pi^0(\operatorname{div}(\mathbf{v}_h)).$$

En injectant ce résultat dans la première ligne du système mixte, on obtient finalement

$$\sum_K \int_K 2G \overline{\epsilon_{u_h}} : \overline{\epsilon_{v_h}} + \int_K \lambda \pi^0(\operatorname{div}(\mathbf{u}_h)) \pi^0(\operatorname{div}(\mathbf{v}_h)) = \sum_K \int_K \mathbf{f} \cdot \mathbf{v}_h$$

ou de manière équivalente

$$\sum_K \int_K 2G \overline{\epsilon_{u_h}} : \overline{\epsilon_{v_h}} + \int_K \lambda \operatorname{div}(\pi^K \mathbf{u}_h) \operatorname{div}(\pi^K \mathbf{v}_h) = \sum_K \int_K \mathbf{f} \cdot \mathbf{v}_h.$$

En approchant la forme bilinéaire approchée  $\int_K 2G \overline{\epsilon_{u_h}} : \overline{\epsilon_{v_h}}$  sur chaque maille par la procédure des éléments virtuels, on aboutit finalement à la même formulation que celle des éléments virtuels stabilisés par  $s^K = 2G$ . Pour plus de détails sur le lien entre ces deux approches, voir par exemple [BBM13; NNR18]. Cette méthode rappelle également les techniques de sous intégration utilisées en éléments finis pour réduire le phénomène de locking.

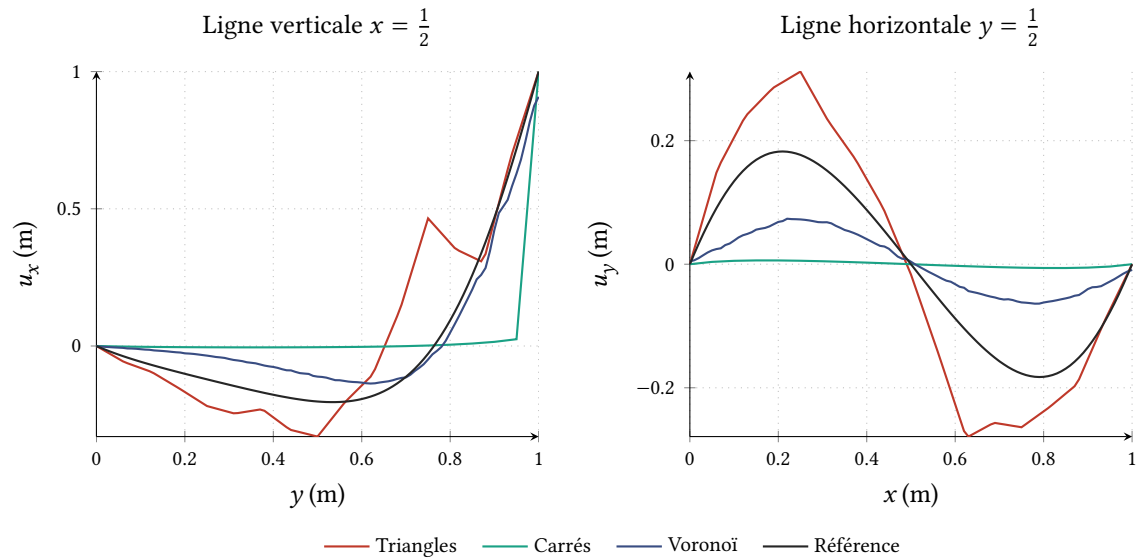


FIG. 3.19 : **Comparaison du déplacement calculé** avec différents maillages pour le problème de cavité entraîné.

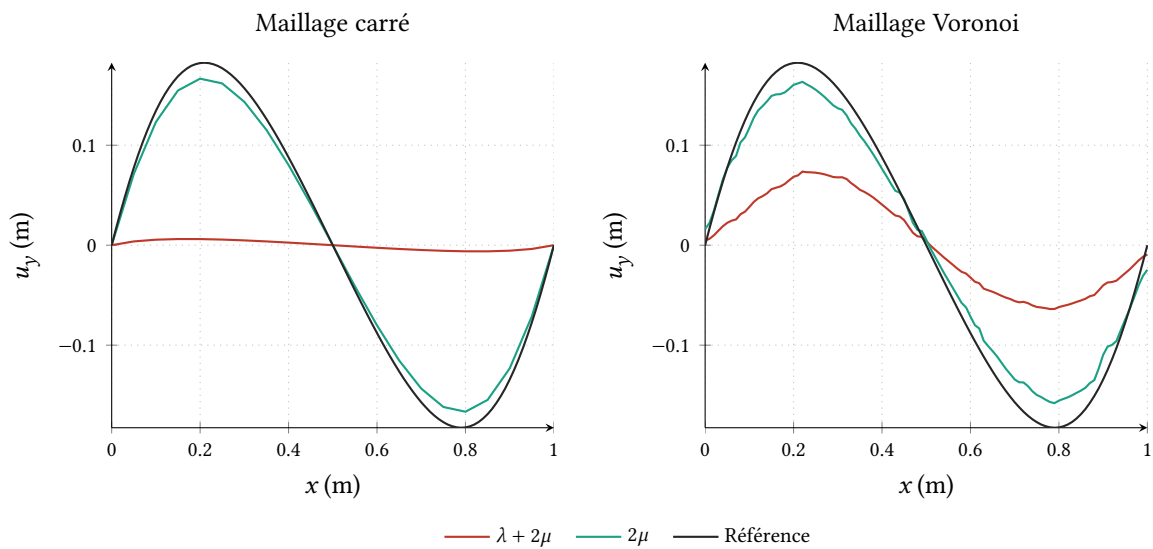


FIG. 3.20 : **Comparaison du déplacement calculé** avec différentes valeurs du coefficient de stabilisation pour le problème de cavité entraîné.





# 4

## Discrétisation VEM-TPFA, un premier schéma couplé pour le système poroélastique

Dans le chapitre précédent, seul le système d'élasticité était considéré de manière à introduire et paramétrer la discrétisation par éléments virtuels de ce problème. Cette étape était préliminaire à l'objectif visé dans ce chapitre : construire et étudier une méthode de discrétisation couplée pour le système poroélastique, dans laquelle l'équation de mécanique est traitée par les éléments virtuels. Pour ce premier schéma couplé, la méthode des volumes finis à deux points (*Two-Points Flux Approximation ou TPFA*) est utilisée pour discrétiser la partie écoulement du problème. La première partie de ce chapitre décrit la construction de ce schéma couplé. Se plaçant d'abord dans le cadre général des volumes finis, le schéma est par la suite précisé en utilisant l'approximation TPFA. L'analyse numérique de ce schéma est proposée dans la seconde partie, où l'on obtient un résultat d'existence et d'unicité de la solution discrète ainsi qu'une estimation de l'erreur d'approximation. À l'instar du chapitre 3, la troisième section de ce chapitre complète la première en ajoutant quelques détails sur l'implémentation. Elle propose également une illustration des résultats d'analyse numérique. Enfin, le chapitre s'achève sur l'usage du schéma dans le cadre de problèmes physiques plus réalistes, tels que ceux présentés à la fin du chapitre 2.

### Sommaire

---

<b>4.1</b>	<b>Construction du schéma numérique</b>	<b>58</b>
4.1.1	Discrétisation du système poroélastique	58
4.1.2	Cas du schéma volumes finis 2 points	60
<b>4.2</b>	<b>Étude du schéma numérique VEM-TPFA</b>	<b>61</b>
4.2.1	Existence et unicité d'une solution discrète	61
4.2.2	Majoration de l'erreur d'approximation	63
4.2.3	Majoration des erreurs d'interpolation	71
<b>4.3</b>	<b>Implémentation et validation numérique</b>	<b>73</b>
4.3.1	Implémentation du schéma	73
4.3.2	Patch test et limites de l'approximation à deux points	76

4.3.3	Validation numérique de l'estimation d'erreur . . . . .	78
<b>4.4</b>	<b>Calculs réalistes de problèmes physiques . . . . .</b>	<b>79</b>
4.4.1	Cas de compaction . . . . .	79
4.4.2	Cas Mandel . . . . .	82

---

## 4.1 Construction du schéma numérique

Le schéma couplé utilise une méthode éléments virtuels pour l'équation mécanique et une méthode volumes finis pour l'équation de conservation du fluide. Les développements de la première partie de cette section sont valides pour n'importe lesquelles de ces méthodes. À partir de la seconde partie et jusqu'à la fin de ce chapitre, on se focalise sur la méthode des éléments virtuels d'ordre 1 présentée précédemment et la méthode des volumes finis à deux points. Pour plus de simplicité, le terme de gravité dans (2.17b) n'est pas pris en compte.

### 4.1.1 Discrétisation du système poroélastique

#### Formulation intégrale du problème

La première étape consiste à réécrire les équations du système (2.17) dans un cadre variationnel ou intégral. Pour ce faire, l'espace fonctionnel  $V_0$ , introduit pour les déplacements dans la partie 2.1.3 est complété par l'espace  $Q$  de sorte que

$$V_0 = \left\{ \mathbf{v} \in (H^1(\Omega))^d, \quad \mathbf{v}|_{\Gamma_{D_d}} = \mathbf{0} \right\}, \quad (4.1a)$$

$$Q = \left\{ p \in H^1(\Omega), \quad p|_{\Gamma_{D_p}} = 0 \right\}. \quad (4.1b)$$

Étant donnée  $\mathbf{f} \in (L^2(]0, T[ \times \Omega))^d$ , une solution  $(\mathbf{u}, p)$  dans  $(C^1[0, T[; V_0) \times (C^1[0, T[; Q)$  du problème continu (2.17) satisfait d'une part pour tout  $\mathbf{v} \in V_0$

$$-\int_{\Omega} \mathbf{div}(\overline{\sigma_u}) \cdot \mathbf{v} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \quad \text{pp dans } ]0, T[$$

soit, en intégrant par parties,

$$\int_{\Omega} \overline{\sigma_u} : \overline{\epsilon_v} = \int_{\Omega} \overline{\overline{C}} \overline{\epsilon_u} : \overline{\epsilon_v} - \int_{\Omega} \alpha p \operatorname{div}(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_{N_d}} \mathbf{t}_N \cdot \mathbf{v} \quad (4.2)$$

et d'autre part, en réutilisant la décomposition de l'espace  $\Omega$  en mailles  $K$  introduite dans la section 3.1.1, en intégrant (2.17b) sur chaque maille et en utilisant la formule de Stokes,

$$\forall K, \quad \int_K \partial_t(c_0 p + \alpha \operatorname{div}(\mathbf{u})) + \int_{\partial K \setminus \Gamma_{N_p}} -\overline{\kappa} \nabla p \cdot \mathbf{n} = \int_K q - \int_{\partial K \cap \Gamma_{N_p}} \phi_N. \quad (4.3)$$

### Discrétisation en temps

L'intervalle  $[0, T]$  est discrétisé en une suite de pas de temps ( $t^n$ ) tels que l'écart entre deux instants  $\Delta t = t^{n+1} - t^n$  soit constant. Toute fonction  $\zeta(\mathbf{x}, t)$  évaluée au temps  $t^n$  est désignée par  $\zeta^n(\mathbf{x}) = \zeta(\mathbf{x}, t^n)$  et les dérivées temporelles au premier ordre sont approchées par la méthode d'Euler implicite :

$$\partial_t \zeta^n(\mathbf{x}) \approx D^- \zeta^n(\mathbf{x}) := \frac{\zeta^n(\mathbf{x}) - \zeta^{n-1}(\mathbf{x})}{\Delta t}. \quad (4.4)$$

### Discrétisation en espace

L'équation de la mécanique (4.2) diffère de celle obtenue en élasticité seule (2.13) uniquement par le terme de couplage  $\int_{\Omega} \alpha p \operatorname{div}(\mathbf{v})$ . En combinant le formalisme des méthodes variationnelles avec la discrétisation temporelle, on obtient le problème approché au temps  $t^n$  :

Trouver  $\mathbf{u}_h^n \in V_{h,0}$  tel que  $\forall \mathbf{v}_h \in V_{h,0}$ ,

$$a_h(\mathbf{u}_h^n, \mathbf{v}_h) - \sum_K \alpha \int_K \operatorname{div}(\mathbf{v}_h) p_h^n = \int_{\Omega} \mathbf{f}^n \cdot \mathbf{v}_h + \int_{\Gamma_{N_d}} \mathbf{t}_N^n \cdot \mathbf{v}_h \quad (4.5)$$

où  $V_{h,0}$  et  $a_h(\cdot, \cdot)$  désignent respectivement l'espace d'approximation des déplacements et la forme bilinéaire associée à la discrétisation. Cette formulation reste générale au sens où il est possible d'utiliser n'importe quelle méthode variationnelle, par exemple celle des éléments finis de Lagrange, pour définir  $V_{h,0}$  et  $a_h(\cdot, \cdot)$ . Évidemment, on réutilise ici la méthode des éléments virtuels construite dans le chapitre 3, de sorte que l'espace  $V_{h,0}$  est défini par l'équation (3.14) et la forme bilinéaire  $a_h(\cdot, \cdot)$  correspond à la forme approchée définie par (3.11).

D'autre part, la solution discrète en pression  $p_h$  est à chercher dans un espace discret  $Q_h$  qui dépend du choix de la méthode d'approximation pour l'équation d'écoulement et sera donc précisé ultérieurement (paragraphe 4.1.2). L'équation de conservation (4.3) est justement discrétisée en suivant le formalisme des volumes finis, décrit en détails dans [EGH00]. La caractéristique des méthodes volumes finis est d'approcher le flux  $\int_{f \in \partial K} -\bar{\kappa} \nabla p \cdot \mathbf{n}$  sortant de la maille  $K$  à travers la face  $f$  par un flux numérique noté  $\mathcal{F}_{Kf}$ . Plusieurs choix sont possibles pour la construction du flux numérique  $\mathcal{F}_{Kf}$ , mais ces schémas doivent respecter une propriété de conservativité : si la face  $f$  sépare les mailles  $K$  et  $L$ , alors l'égalité  $\mathcal{F}_{Kf} = -\mathcal{F}_{Lf}$  doit être respectée. En utilisant la discrétisation en temps introduite plus haut, le schéma volumes finis s'écrit sous la forme

$$\forall K \in \tau_h, \forall n \in [1, N], \quad \int_K (c_0 p_h^n + \alpha \operatorname{div}(\mathbf{u}_h^n)) + \Delta t \sum_{\substack{f \subset \partial K \\ f \notin \Gamma_{N_p}}} \mathcal{F}_{Kf}^n = \int_K (\Delta t q^n + c_0 p_h^{n-1} + \alpha \operatorname{div}(\mathbf{u}_h^{n-1})) - \Delta t \int_{\partial K \cap \Gamma_{N_p}} \phi_N^n. \quad (4.6)$$

La combinaison de ces deux discrétisations conduit au schéma couplé : pour chaque pas de temps  $t^n$  avec  $n \geq 1$ , trouver une solution  $(\mathbf{u}_h^n, p_h^n) \in (V_{h,0} \times Q_h)$  au système (4.7) :

**Discrétisation VEM-FV pour le système poroélastique**

 Trouver  $(\mathbf{u}_h^n, p_h^n) \in (V_{h,0} \times Q_h)$  tel que

$$\begin{aligned} \forall \mathbf{v}_h \in V_{h,0}, \quad a_h(\mathbf{u}_h^n, \mathbf{v}_h) - \sum_K \alpha \int_K \operatorname{div}(\mathbf{v}_h) p_h^n &= \int_{\Omega} \mathbf{f}^n \cdot \mathbf{v}_h + \int_{\Gamma_{Nd}} \mathbf{t}_N^n \cdot \mathbf{v}_h \\ \forall K \in \tau_h, \int_K (c_0 p_h^n + \alpha \operatorname{div}(\mathbf{u}_h^n)) + \Delta t \sum_{\substack{f \subset \partial K \\ f \notin \Gamma_{Np}}} \mathcal{F}_{Kf}^n &= \int_K (\Delta t q^n + c_0 p_h^{n-1} + \alpha \operatorname{div}(\mathbf{u}_h^{n-1})) - \Delta t \int_{\partial K \cap \Gamma_{Np}} \phi_N^n \end{aligned} \quad (4.7)$$

 avec à l'instant initial  $p_h^0$  projection de  $p^0 = p(\mathbf{x}, t^0)$  dans l'espace  $Q_h$  et  $\mathbf{u}_h^0 \in V_{h,0}$  solution du problème

$$\forall \mathbf{v}_h \in V_{h,0}, \quad a_h(\mathbf{u}_h^0, \mathbf{v}_h) = \sum_K \alpha \int_K \operatorname{div}(\mathbf{v}_h) p_h^0 + \int_{\Omega} \mathbf{f}^0 \cdot \mathbf{v}_h + \int_{\Gamma_{Nd}} \mathbf{t}_N^0 \cdot \mathbf{v}_h. \quad (4.8)$$

**4.1.2 Cas du schéma volumes finis 2 points**

Cette formulation générale est à préciser en choisissant l'espace d'approximation  $Q_h$  et la méthode d'approximation du flux numérique. Différentes alternatives existent (voir par exemple [Dro14] pour une revue des différentes méthodes), plus ou moins sophistiquées mais également plus ou moins coûteuses, offrant plusieurs possibilités pour le couplage avec l'équation de la mécanique. Dans un premier temps, on s'intéresse à la méthode dite à deux points qui fait intervenir une inconnue en pression par maille. Elle appartient à l'ensemble des méthodes dites *cell centered* où les inconnues sont localisées dans les mailles, qui sont compatibles avec la formulation (4.3). En conséquence, l'espace d'approximation  $Q_h$  est l'espace des fonctions constantes dans chaque cellule.

$$Q_h = \{p_h \in L^2(\Omega) \text{ tel que } p_{h_K} \in \mathbb{P}_0(K)\}. \quad (4.9)$$

On note que  $Q_h$  est inclus dans  $L^2$  sans être un sous espace de  $Q$  :  $Q_h \not\subset Q$ . Dans chaque élément  $K$ , la pression est calculée en utilisant un point de calcul  $\mathbf{x}_K$  qui peut possiblement différer du barycentre de la cellule, propriété qui sera illustrée dans la section 4.3. Le fait que la pression discrète soit constante par maille facilite grandement l'évaluation des intégrales puisque  $\int_K p_h^n d\mathbf{x} = p_{h_K}^n \int_K d\mathbf{x}$ . En particulier, le traitement du terme de couplage  $\int_K \alpha \operatorname{div}(\mathbf{v}_h) p_h^n$  se réduit au calcul de  $\alpha p_{h_K}^n \int_K \operatorname{div}(\mathbf{v}_h)$ , et donc à l'intégrale de la divergence des fonctions de base. Or, cette intégrale a déjà été calculée dans la méthode des éléments virtuels.

La méthode à deux points, approximation de base des volumes finis, souffre de certains inconvénients, le principal étant la contrainte dite d'orthogonalité. Dans le cas d'un tenseur  $\bar{\kappa}$  isotrope, si une face  $f$  sépare les mailles  $K$  et  $L$ , dont la pression est approchée aux points  $\mathbf{x}_K$  et  $\mathbf{x}_L$ , alors le schéma n'est consistant que si le segment formé par ces points de calcul est orthogonal à la face : la propriété  $(\mathbf{x}_K \mathbf{x}_L) \perp f$  doit être vérifiée. Dans le cas où le tenseur  $\bar{\kappa}$  est anisotrope, la consistance est encore plus compliquée à obtenir puisque le maillage doit dans ce cas vérifier la contrainte dite de  $\bar{\kappa}$ -orthogonalité

$$\bar{\kappa}^{-1}(\mathbf{x}_K \mathbf{x}_L) \perp f. \quad (4.10)$$

Dans la suite de ce chapitre, on supposera donc que  $\bar{\kappa} = \kappa \bar{I}_d$ , avec  $\kappa$  scalaire réel.

En contrepartie de cette limitation, la méthode à deux points est relativement simple à mettre en place et peu coûteuse. En effet, peu d'inconnues interviennent dans le calcul des flux numériques, ce qui conduit à un stencil étroit pour la matrice du système linéaire associé à la discrétisation. Cette méthode constitue donc un bon choix pour mettre en place un premier schéma couplé et en effectuer l'analyse mathématique. Avec cette méthode, le flux numérique est défini par

$$\mathcal{F}_{Kf}^n = \begin{cases} T_f(p_{h_K}^n - p_{h_L}^n) & \text{avec } T_f = \frac{|f|\kappa^{eq}}{d_{KL}} \text{ si } f \subset \Omega, \\ T_f p_{h_K}^n & \text{avec } T_f = \frac{|f|\kappa}{d_{Kf}} \text{ si } f \subset \Gamma_{D_p}, \end{cases} \quad (4.11)$$

où  $T_f$  est appelé *transmissivité* à travers la face  $f$  et fait intervenir pour les arêtes internes la perméabilité équivalente  $\kappa^{eq}$  calculée de manière harmonique grâce à la formule

$$\frac{d_{KL}}{\kappa^{eq}} = \frac{d_{Lf}}{\kappa_L} + \frac{d_{Kf}}{\kappa_K}. \quad (4.12)$$

La quantité  $d_{KL}$  (resp.  $d_{Kf}$ ) correspond à la distance euclidienne entre les points  $\mathbf{x}_K$  et  $\mathbf{x}_L$  (resp. entre le point  $\mathbf{x}_K$  et la face  $f$ ). Ce flux définit également la semi-norme  $H^1$  discrète (qui est en réalité une norme  $H^1$  discrète dans le cas où  $\Gamma_{D_p} \neq \emptyset$ )

$$|p_h|_{1,\tau} = \left( \sum_{f \in \overline{\Omega} \setminus \Gamma_{N_p}} T_f D_f(p_h)^2 \right)^{\frac{1}{2}} \text{ avec } D_f(p_h) = \begin{cases} |p_{h_K} - p_{h_L}| & \text{si } f \subset \Omega, \\ |p_{h_K}| & \text{si } f \subset \Gamma_{D_p}. \end{cases} \quad (4.13)$$

L'usage de l'espace  $Q_h$  et de l'approximation à deux points pour calculer le flux dans l'équation (4.7) définit donc le schéma couplé VEM-TPFA.

## 4.2 Étude du schéma numérique VEM-TPFA

Dans cette section, on travaille sur l'étude numérique du schéma construit dans la section précédente. Pour simplifier les démonstrations, on travaille avec des conditions de Neumann nulles ( $\mathbf{t}_N = \mathbf{0}, \phi_N = 0$ ). On montre dans la section 4.2.1 l'existence et l'unicité d'une solution discrète au schéma numérique, et on travaille sur une estimation de l'erreur numérique dans les sections 4.2.2 et 4.2.3.

### 4.2.1 Existence et unicité d'une solution discrète

On démontre dans cette partie le théorème suivant :

**Theorème 4.1** *Le schéma (4.7) avec les flux de la méthode à deux points (4.11) admet pour tout temps  $t^n$  une unique solution discrète  $(\mathbf{u}_h^n, p_h^n)$ .*

Soit  $\mathcal{O}$  l'opérateur associé au système d'équations (4.7). On cherche à chaque temps  $t^n$  une solution satisfaisant  $\mathcal{O}(\mathbf{u}_h^n, p_h^n) = b$  où  $b$  dépend des données et de la solution au temps précédent. Remarquant que :

- l'opérateur  $\mathcal{O}$  est linéaire (car  $a_h$  est bilinéaire et que les flux  $\mathcal{F}_{Kf}^n$  ainsi que les termes de couplage sont linéaires);

- le système est carré car composé de  $d$  équations ( $d = 2$  ou  $3$ ) par nœuds "non Dirichlet" pour sa première ligne, par définition de  $V_{h,0}$ , plus une équation par maille pour sa seconde ligne ce qui est égal au nombre d'inconnues ( $d$  par nœuds en déplacement et une par maille en pression) ;
- à l'instant initial, la solution  $(\mathbf{u}_h^0, p_h^0)$ , est uniquement définie par les données initiales ; plus précisément,  $p_h^0$  est la projection de  $p^0$  dans l'espace discret et  $\mathbf{u}_h^0$  est solution du problème élastique discret (4.8), ce problème admettant une unique solution d'après le lemme 3.1 ;

il suffit de montrer qu'à chaque instant  $t^n$ , l'unique solution du système homogène

$$\begin{aligned} \forall \mathbf{v}_h \in V_{h,0}, \quad a_h(\mathbf{u}_h^n, \mathbf{v}_h) - \sum_K \alpha \int_K \operatorname{div}(\mathbf{v}_h) p_h^n &= 0 \\ \forall K \in \tau_h, \quad \int_K (c_0 p_h^n + \alpha \operatorname{div}(\mathbf{u}_h^n)) + \Delta t \sum_{\substack{f \subset \partial K \\ f \notin \Gamma_{N_p}}} \mathcal{F}_{Kf}^n &= 0 \end{aligned} \quad (4.14)$$

est  $(0, 0)$ . La première égalité est évaluée en  $\mathbf{v}_h = \mathbf{u}_h^n$  (car  $\mathbf{u}_h^n \in V_{h,0}$ ) tandis que la seconde est multipliée par la valeur constante  $p_{h_{|K}}^n$ , puis sommée sur l'ensemble des cellules :

$$\begin{aligned} a_h(\mathbf{u}_h^n, \mathbf{u}_h^n) - \sum_K \alpha \int_K \operatorname{div}(\mathbf{u}_h^n) p_h^n &= 0 \\ \sum_K \int_K (c_0 p_h^n p_{h_{|K}}^n + \alpha \operatorname{div}(\mathbf{u}_h^n) p_{h_{|K}}^n) + \Delta t \sum_K p_{h_{|K}}^n \sum_{\substack{f \subset \partial K \\ f \notin \Gamma_{N_p}}} \mathcal{F}_{Kf}^n &= 0. \end{aligned}$$

En sommant ces deux égalités, il y a simplification des termes  $\int_K \alpha \operatorname{div}(\mathbf{u}_h^n) p_h^n$  d'où

$$a_h(\mathbf{u}_h^n, \mathbf{u}_h^n) + \int_{\Omega} c_0 p_h^n p_h^n + \Delta t \sum_K p_{h_{|K}}^n \sum_{\substack{f \subset \partial K \\ f \notin \Gamma_{N_p}}} \mathcal{F}_{Kf}^n = 0. \quad (4.15)$$

Les deux premiers termes sont positifs par positivité de  $a_h$  et du produit scalaire. Pour le troisième terme, en réorganisant la somme sur les faces,

$$\begin{aligned} \sum_K p_{h_{|K}}^n \sum_{\substack{f \subset \partial K \\ f \notin \Gamma_{N_p}}} \mathcal{F}_{Kf}^n &= \sum_{f \subset \Omega \setminus \partial \Omega} T_f \left( (p_{h_{|K}}^n - p_{h_{|L}}^n) p_{h_{|K}}^n + (p_{h_{|L}}^n - p_{h_{|K}}^n) p_{h_{|L}}^n \right) + \sum_{f \subset \Gamma_{D_p}} T_f (p_{h_{|K}}^n)^2 \\ &= \sum_{f \subset \Omega \setminus \partial \Omega} T_f (p_{h_{|K}}^n - p_{h_{|L}}^n)^2 + \sum_{f \subset \Gamma_{D_p}} T_f (p_{h_{|K}}^n)^2 \geq 0. \end{aligned} \quad (4.16)$$

Chacun des trois termes dans la somme (4.15) est donc nul. La forme  $a_h$  étant définie positive, on a directement  $\mathbf{u}_h^n = \mathbf{0}$ . Pour  $p_h^n$ , on a la nullité si  $c_0 > 0$  grâce au produit scalaire. Dans le cas contraire, le fait que la somme (4.16) soit nulle implique que  $p_{h_{|K}}^n = 0$  lorsque  $f \subset \Gamma_{D_p}$ . En raisonnant de proche en proche, on a  $p_{h_{|L}}^n = 0$  pour toutes les autres mailles et finalement  $p_h^n = 0$ .

Ainsi, on a montré que l'unique solution du système homogène est  $(0, 0)$  et donc que l'opérateur  $\mathcal{O}$  associé est inversible. Pour tout second membre, le système non homogène  $\mathcal{O}(\mathbf{u}_h^n, p_h^n) = b$  admet une unique solution d'après les arguments énoncés en début de section.

### 4.2.2 Majoration de l'erreur d'approximation

Pour quantifier l'écart entre la solution exacte et la solution calculée, on introduit des opérateurs d'interpolation permettant d'obtenir le représentant de la solution exacte dans les espaces d'approximation discrets. L'opérateur permettant de passer d'une fonction suffisamment régulière de  $Q$  à son interpolée dans  $Q_h$  est noté  $P$  et est défini par

$$\forall K \in \tau_h, \forall \mathbf{x} \in K, \quad Pp(\mathbf{x}) = p(\mathbf{x}_K).$$

L'opérateur continu permettant de passer d'une fonction suffisamment régulière de  $V_0$  à son interpolée dans  $V_{h,0}$  est noté  $\tilde{P}$  et est défini à partir des degrés de liberté (*ie* les valeurs aux sommets) :

$$\forall i \in \mathcal{M}, \quad (\tilde{P}\mathbf{u})(V_i) = \mathbf{u}(V_i).$$

L'erreur entre la solution réelle et la solution approchée est décomposée en erreur d'interpolation  $E^I$ , qui dépend de la qualité des espaces d'approximation, et erreur du schéma  $E^A$  qui dépend de la qualité du schéma numérique :

$$\begin{aligned} \mathbf{u} - \mathbf{u}_h &= \mathbf{u} - \tilde{P}\mathbf{u} + \tilde{P}\mathbf{u} - \mathbf{u}_h = E_{\mathbf{u}}^I + E_{\mathbf{u}}^A = E_{\mathbf{u}} \\ p - p_h &= p - Pp + Pp - p_h = E_p^I + E_p^A = E_p. \end{aligned} \quad (4.17)$$

Pour les besoins de l'analyse d'erreur, une projection depuis  $V_0$  vers l'espace des fonctions affines par mailles  $\mathbb{P}_1(K)^d$  est nécessaire. Dès lors que  $\mathbf{u} \in \mathcal{C}^0(\bar{\Omega}) \cap H^1(\Omega)$ , cette projection coïncide sur chaque maille avec celle définie par l'équation (3.16). L'erreur de projection associée à cet opérateur est notée

$$E_{\mathbf{u}}^{\pi} = \mathbf{u}|_K - \pi^K \mathbf{u}|_K. \quad (4.18)$$

Puisque cette quantité n'est pas globalement dans  $H^1(\Omega)$ , car discontinue aux interfaces, on définit la norme  $H^1$  brisée

$$\|\cdot\|_{H^1(\tau)} = \left( \sum_{K \in \tau_h} \|\cdot\|_{H^1(K)}^2 \right)^{\frac{1}{2}}.$$

Le travail de cette section vise à démontrer l'estimation suivante sur les erreurs d'approximation.

**Proposition 4.2** *Sous réserve que la solution exacte soit suffisamment régulière et que  $\Delta t \leq \frac{c_0}{2}$ , l'erreur d'approximation du schéma satisfait*

$$\|E_{\mathbf{u}}^{A,N}\|_{H^1(\Omega)}^2 + \|E_p^{A,N}\|_{L^2(\Omega)}^2 + \sum_{n=1}^N \Delta t |E_p^{A,n}|_{1,\tau}^2 \leq \exp(t^N) (B\Delta t^2 + B' \|E_{\mathbf{u}}^{A,0}\|_{H^1(\Omega)}^2 + E^I) \quad (4.19)$$

où  $E_{\mathbf{u}}^{A,n}$  et  $E_p^{A,n}$  sont les quantités définies par (4.17) évaluées au temps  $t^n$ ,  $B$  et  $B'$  sont deux constantes positives dépendant des paramètres physiques  $c_0, \alpha, d, |\Omega|$ , des constantes d'ellipticité de  $a$  et  $a_n$ , des dérivées temporelles de la solution exacte  $\partial_{tt}\mathbf{u}$  et  $\partial_{tt}p$  et où le terme  $E^I$  défini par (4.33) regroupe les différentes erreurs d'interpolation et de projection.

La démonstration est séparée en trois étapes. En premier lieu, une égalité faisant intervenir les erreurs d'approximation est obtenue. En second lieu, le membre de droite est borné par une quantité indépendante des erreurs d'approximation. Enfin, une version discrète du lemme de Grönwall est utilisée pour conclure.



**Première partie – obtention d’une égalité**

Dans cette partie, toutes les grandeurs sont considérées à un temps donné  $t^n$ . L’exposant  $n$  des inconnues discrètes est donc omis pour une meilleure lisibilité. En utilisant la bilinéarité de  $a$  et  $a_h$ , des termes additionnels sont introduits dans les formulations variationnelles exactes (4.2) et approchées (4.7) :

$$\forall \mathbf{v}_h \in V_{h,0},$$

$$\begin{aligned} \sum_{K \in \tau_h} a(\mathbf{u}|_K - \pi^K \mathbf{u}|_K, \mathbf{v}_h) + \sum_{K \in \tau_h} a(\pi^K \mathbf{u}|_K, \mathbf{v}_h) - \alpha \int_{\Omega} p \operatorname{div}(\mathbf{v}_h) &= \int_{\Omega} \mathbf{f} \cdot \mathbf{v}_h \\ -a_h(\tilde{P}\mathbf{u} - \mathbf{u}_h, \mathbf{v}_h) + a_h(\tilde{P}\mathbf{u} - \sum_{K \in \tau_h} \pi^K \mathbf{u}|_K, \mathbf{v}_h) + \sum_{K \in \tau_h} a_h(\pi^K \mathbf{u}|_K, \mathbf{v}_h) - \alpha \int_{\Omega} p_h \operatorname{div}(\mathbf{v}_h) &= \int_{\Omega} \mathbf{f}_h \cdot \mathbf{v}_h. \end{aligned}$$

La soustraction de ces deux égalités et le fait que  $a_h(\pi^K \mathbf{u}|_K, \mathbf{v}_h) = a(\pi^K \mathbf{u}|_K, \mathbf{v}_h)$  pour tout  $K$  car  $\pi^K \mathbf{u}|_K \in P_1(K)^d$  conduit à

$$a_h(E_{\mathbf{u}}^A, \mathbf{v}_h) - \alpha \int_{\Omega} E_p \operatorname{div}(\mathbf{v}_h) = a \left( \sum_{K \in \tau_h} \pi^K \mathbf{u}|_K - \mathbf{u}|_K, \mathbf{v}_h \right) + a_h \left( \tilde{P}\mathbf{u} - \sum_{K \in \tau_h} \pi^K \mathbf{u}|_K, \mathbf{v}_h \right) + \int_{\Omega} (\mathbf{f} - \mathbf{f}_h) \cdot \mathbf{v}_h.$$

Puisque  $D^- E_{\mathbf{u}}^A \in V_{h,0}$  (l’opérateur  $D^-$  étant défini par (4.4)), on a en choisissant  $\mathbf{v}_h = D^- E_{\mathbf{u}}^A$

$$a_h(E_{\mathbf{u}}^A, D^- E_{\mathbf{u}}^A) - \alpha \int_{\Omega} E_p \operatorname{div}(D^- E_{\mathbf{u}}^A) = a(-E_{\mathbf{u}}^{\pi}, D^- E_{\mathbf{u}}^A) + a_h(E_{\mathbf{u}}^{\pi} - E_{\mathbf{u}}^I, D^- E_{\mathbf{u}}^A) + \int_{\Omega} (\mathbf{f} - \mathbf{f}_h) \cdot D^- E_{\mathbf{u}}^A. \quad (4.20)$$

Pour l’équation de conservation, on soustrait de la même manière les deux équations ci-dessous satisfaites par chacune des deux solutions

$$\begin{aligned} \int_K c_0 (D^- - D^- + \partial_t)(p(\mathbf{x}, t^n)) + \alpha \operatorname{div}((D^- - D^- + \partial_t)(\mathbf{u}(\mathbf{x}, t^n))) + \int_{\partial K} -\kappa \nabla p(\mathbf{x}, t^n) \cdot \mathbf{n} &= \int_K q(\mathbf{x}, t^n) \\ \int_K c_0 D^- p_{h|K} + \alpha \operatorname{div}(D^- \mathbf{u}_h) + \sum_{f \subset \partial K} \mathcal{F}_{Kf} &= \int_K q(\mathbf{x}, t^n) \end{aligned}$$

puis on multiplie par la grandeur constante par maille  $E_p^A$  et on somme sur l’ensemble des mailles pour obtenir

$$\begin{aligned} \int_{\Omega} c_0 D^- (E_p) E_p^A + \alpha \int_{\Omega} \operatorname{div}(D^- E_{\mathbf{u}}) E_p^A + \sum_K \sum_{f \subset \partial K} (\overline{F_{Kf}} - F_{Kf}^n) E_p^A \\ = \int_{\Omega} c_0 (D^- p - \partial_t p) E_p^A + \int_{\Omega} \alpha \operatorname{div}((D^- \mathbf{u} - \partial_t \mathbf{u})) E_p^A \end{aligned} \quad (4.21)$$

où l’on a noté  $\overline{F_{Kf}}$  le flux exact défini au temps  $t^n$  par  $\overline{F_{Kf}} = \int_{f \cap \partial K} -\kappa \nabla p(\mathbf{x}, t^n) \cdot \mathbf{n}$ . La différence entre le flux numérique  $\mathcal{F}$  et le flux exact  $\overline{F}$  doit être explicitée afin de séparer l’erreur due au schéma volumes finis et l’erreur due à l’interpolation dans l’espace discret  $Q_h$  : ainsi, on note  $F_{Kf}^*$  le flux numérique évalué en la projection  $Pp$  de la solution exacte de sorte que

$$\sum_K \sum_{f \subset \partial K} (\overline{F_{Kf}} - \mathcal{F}_{Kf}) E_p^A = \sum_K \sum_{f \subset \partial K} (\overline{F_{Kf}} - F_{Kf}^*) E_p^A + \sum_K \sum_{f \subset \partial K} (F_{Kf}^* - \mathcal{F}_{Kf}) E_p^A.$$

On remarque directement que la différence  $(F_{Kf}^* - \mathcal{F}_{Kf})$  est égale au flux numérique  $\mathcal{F}$  évalué en  $E_p^A$  et non en  $p_h$ . Ainsi, en procédant à une réorganisation de la somme comme pour l'équation (4.16), on a

$$\sum_K \sum_{f \in \partial K} (\overline{F_{Kf}} - \mathcal{F}_{Kf}) E_p^A = \sum_K \sum_{f \in \partial K} (\overline{F_{Kf}} - F_{Kf}^*) E_p^A + |E_p^A|_{1,\tau}^2.$$

La seconde double somme s'assimile à une erreur d'interpolation et sera traitée prochainement. La première partie s'achève en sommant les équations (4.20) et (4.21) en remarquant la simplification partielle des termes de la forme  $\int_{\Omega} p \operatorname{div}(\mathbf{u})$  (cette simplification était totale lors de l'étude de stabilité du schéma continu) pour obtenir

$$a_h(E_{\mathbf{u}}^A, D^- E_{\mathbf{u}}^A) + \int_{\Omega} c_0 D^- E_p^A E_p^A + |E_p^A|_{1,\tau}^2 = B_1 + B_2 + B_3 + B_4 + B_5 + B_6$$

avec

$$\begin{aligned} B_1 &= -\alpha \int_{\Omega} \operatorname{div}(D^- E_{\mathbf{u}}^I) E_p^A + \alpha \int_{\Omega} \operatorname{div}(D^- E_{\mathbf{u}}^A) E_p^I, & B_2 &= a(-E_{\mathbf{u}}^{\pi}, D^- E_{\mathbf{u}}^A) + a_h(E_{\mathbf{u}}^{\pi} - E_{\mathbf{u}}^I, D^- E_{\mathbf{u}}^A), \\ B_3 &= \int_{\Omega} c_0 (D^- p - \partial_t p) E_p^A + \int_{\Omega} \alpha \operatorname{div}(D^- \mathbf{u} - \partial_t \mathbf{u}) E_p^A, & B_4 &= \int_{\Omega} (\mathbf{f} - \mathbf{f}_h) \cdot D^- E_{\mathbf{u}}^A, \\ B_5 &= -\sum_K \sum_{f \in \partial K} (\overline{F_{Kf}} - F_{Kf}^*) E_p^A, & B_6 &= -c_0 \int_{\Omega} E_p^A D^- E_p^I. \end{aligned}$$

## Deuxième partie – majoration

À ce stade, l'exposant temporel  $n$  est réintroduit et l'équation précédente est sommée sur  $n$  entre 1 et  $N$ . En utilisant l'égalité  $2\mathcal{B}(u - v, u) = \mathcal{B}(u, u) + \mathcal{B}(u - v, u - v) - \mathcal{B}(v, v)$  valable pour toute forme bilinéaire symétrique, on a en notant  $\|\mathbf{u}\|_{a_h}^2 = a_h(\mathbf{u}, \mathbf{u})$  la norme associée à la forme bilinéaire  $a_h$

$$\begin{aligned} \int_{\Omega} c_0 D^- E_p^{A,n} E_p^{A,n} &\geq \frac{c_0}{2\Delta t} (\|E_p^{A,n}\|_{L^2}^2 - \|E_p^{A,n-1}\|_{L^2}^2); \\ a_h(E_{\mathbf{u}}^{A,n}, D^- E_{\mathbf{u}}^{A,n}) &\geq \frac{1}{2\Delta t} (\|E_{\mathbf{u}}^{A,n}\|_{a_h}^2 - \|E_{\mathbf{u}}^{A,n-1}\|_{a_h}^2). \end{aligned}$$

La somme est donc télescopique et, en remarquant de plus que  $E_p^{A,0} = 0$  et en utilisant la constante de coercivité de  $a_h$  notée  $\tilde{\mu}$ , celle-ci devient

$$\tilde{\mu} \|E_{\mathbf{u}}^{A,N}\|_{H^1}^2 + c_0 \|E_p^{A,N}\|_{L^2}^2 + 2 \sum_{n=0}^N \Delta t |E_p^{A,n}|_{1,\tau}^2 \leq \tilde{\mu} \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2 + 2\Delta t \sum_{n=1}^N (B_1^n + B_2^n + B_3^n + B_4^n + B_5^n + B_6^n).$$

Il reste à majorer chacune des sommes en temps du second membre. Pour ce faire, les outils utilisés sont :

1. l'inégalité de Young

$$2ab \leq \epsilon a^2 + \frac{b^2}{\epsilon} \text{ avec } \epsilon \geq 0, \quad (4.22)$$

souvent utilisée avec un  $\epsilon$  différent sur le dernier pas de temps des sommes  $B_i^n$ ;

2. la relation

$$\forall \mathbf{v} \in (H^1(\Omega))^d, \|\operatorname{div}(\mathbf{v})\|_{L^2} \leq \sqrt{d} |\mathbf{v}|_{H^1}; \quad (4.23)$$

3. les développements de Taylor de fonctions  $\mathcal{C}^2$  en temps

$$D^-v(t^n) = \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} \partial_t v(s) ds, \quad (4.24a)$$

$$D^-v(t^n) - \partial_t v(t^n) = \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} \partial_{tt} v(s) (t^{n-1} - s) ds; \quad (4.24b)$$

4. la formule de sommation par parties discrète, valide pour toute forme bilinéaire  $\mathcal{B}$

$$\sum_{k=1}^N \mathcal{B}(a_k, b_k - b_{k-1}) = - \sum_{k=1}^{N-1} \mathcal{B}(b_k, a_{k+1} - a_k) + \mathcal{B}(a_N, b_N) - \mathcal{B}(a_1, b_0); \quad (4.25)$$

5. la double identité de Cauchy-Schwarz : pour deux variables discrétisées scalaires ou vectorielles  $a, b$  et  $n \in [1, N]$ , et en notant  $b^n = b(\mathbf{x}, t^n)$ ,

$$\left| \int_{\Omega} (b^n - b^{n-1}) a^n \right| \leq \|a^n\|_{L^2} \left( \Delta t \int_{t^{n-1}}^{t^n} \|\partial_t b(s)\|_{L^2}^2 ds \right)^{\frac{1}{2}}. \quad (4.26)$$

Dans ce qui suit,  $\epsilon_r$  et  $\epsilon_l$  sont deux réels strictement positifs quelconques et  $\tilde{\epsilon}_r$  est défini par  $\frac{1}{\tilde{\epsilon}_r} = 1 + \frac{1}{\epsilon_r}$ .

• Majoration de  $\Delta t B_1^n = -\alpha \Delta t \int_{\Omega} \operatorname{div} (D^- E_{\mathbf{u}}^{I,n}) E_p^{A,n} + \alpha \Delta t \int_{\Omega} \operatorname{div} (D^- E_{\mathbf{u}}^{A,n}) E_p^{I,n}$ .

Les deux intégrales sont traitées séparément. Pour le premier terme, les outils 5 et 2 sont successivement utilisés,

$$\left| \Delta t \int_{\Omega} E_p^{A,n} \operatorname{div} (D^- E_{\mathbf{u}}^{I,n}) \right| = \sqrt{\Delta t} \|E_p^{A,n}\|_{L^2} \left( \int_{t^{n-1}}^{t^n} \|\operatorname{div} (\partial_t E_{\mathbf{u}}^I)\|_{L^2}^2 \right)^{\frac{1}{2}} \leq \sqrt{d \Delta t} \|E_p^{A,n}\|_{L^2} \left( \int_{t^{n-1}}^{t^n} \|\partial_t E_{\mathbf{u}}^I\|_{H^1}^2 \right)^{\frac{1}{2}}$$

puis l'inégalité de Young est appliquée avec  $\epsilon = \frac{1}{4}$  pour le dernier terme de la somme et  $\epsilon = \frac{\epsilon_r}{4}$  pour les autres ce qui permet d'écrire

$$\begin{aligned} \left| 2 \sum_{n=1}^N \alpha \Delta t \int_{\Omega} E_p^{A,n} \operatorname{div} (D^- E_{\mathbf{u}}^{I,n}) \right| &\leq \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_p^{A,n}\|_{L^2}^2 + \frac{\Delta t}{4} \|E_p^{A,N}\|_{L^2}^2 \\ &\quad + 4d\alpha^2 \left( \frac{1}{\epsilon_r} \sum_{n=1}^{N-1} \int_{t^{n-1}}^{t^n} \|\partial_t E_{\mathbf{u}}^I\|_{H^1}^2 + \int_{t^{N-1}}^{t^N} \|\partial_t E_{\mathbf{u}}^I\|_{H^1}^2 \right). \end{aligned}$$

Afin de retrouver la norme en temps définie dans la section 2.2 par  $\|g\|_{L^2(T)}^2 = \int_{t^0}^{t^N} g(s)^2 ds$ , les termes manquants dans l'intégrale sont ajoutés au membre de droite, ceux-ci étant positifs :

$$\begin{aligned} \left| 2 \sum_{n=1}^N \alpha \Delta t \int_{\Omega} E_p^{A,n} \operatorname{div} (D^- E_{\mathbf{u}}^{I,n}) \right| &\leq \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_p^{A,n}\|_{L^2}^2 + \frac{\Delta t}{4} \|E_p^{A,N}\|_{L^2}^2 + 4d\alpha^2 \left( \frac{1}{\epsilon_r} \|\partial_t E_{\mathbf{u}}^I\|_{L^2(0,T;H^1)}^2 + \|\partial_t E_{\mathbf{u}}^I\|_{L^2(0,T;H^1)}^2 \right) \\ &\leq \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_p^{A,n}\|_{L^2}^2 + \frac{\Delta t}{4} \|E_p^{A,N}\|_{L^2}^2 + 4d\alpha^2 \left( \frac{1}{\tilde{\epsilon}_r} \|\partial_t E_{\mathbf{u}}^I\|_{L^2(0,T;H^1)}^2 \right). \end{aligned}$$

Ceci achève la majoration de la première intégrale. Concernant la seconde, une sommation par parties est d'abord effectuée

$$2 \sum_{n=1}^N \Delta t \int_{\Omega} \alpha \operatorname{div} (D^- E_{\mathbf{u}}^{A,n}) E_p^{I,n} = 2 \int_{\Omega} \alpha \operatorname{div} (E_{\mathbf{u}}^{A,N}) E_p^{I,N} - 2 \int_{\Omega} \alpha \operatorname{div} (E_{\mathbf{u}}^{A,0}) E_p^{I,1} - 2 \sum_{n=1}^{N-1} \Delta t \int_{\Omega} \alpha \operatorname{div} (E_{\mathbf{u}}^{A,n}) D^- E_p^{I,n+1}$$

puis, en suivant la même procédure que précédemment on aboutit à

$$\begin{aligned}
 2 \left| \int_{\Omega} \alpha \operatorname{div} \left( E_{\mathbf{u}}^{A,N} \right) E_p^{I,N} \right| &\leq \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,N}\|_{H^1}^2 + \frac{4d\alpha^2}{\epsilon_l} \|E_p^{I,N}\|_{L^2}^2, \\
 2 \left| \int_{\Omega} \alpha \operatorname{div} \left( E_{\mathbf{u}}^{A,0} \right) E_p^{I,1} \right| &\leq \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2 + \frac{4d\alpha^2}{\epsilon_l} \|E_p^{I,1}\|_{L^2}^2, \\
 2 \left| \sum_{n=1}^{N-1} \alpha \Delta t \int_{\Omega} \operatorname{div} \left( E_{\mathbf{u}}^{A,n} \right) D^- E_p^{I,n+1} \right| &\leq \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_{\mathbf{u}}^{A,n}\|_{H^1}^2 + \frac{4d\alpha^2}{\epsilon_r} \|\partial_t E_p^I\|_{L^2(0,T;L^2)}^2.
 \end{aligned}$$

Il ne reste qu'à combiner les deux termes pour obtenir l'estimation sur  $B_1$  :

$$\begin{aligned}
 \left| 2\Delta t \sum_{n=1}^N B_1 \right| &\leq \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_p^{A,n}\|_{L^2}^2 + \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_{\mathbf{u}}^{A,n}\|_{H^1}^2 + \frac{\Delta t}{4} \|E_p^{A,N}\|_{L^2}^2 + \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,N}\|_{H^1}^2 + \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2 \\
 &\quad + 4d\alpha^2 \left( \frac{\|\partial_t E_{\mathbf{u}}^I\|_{L^2(0,T;H^1)}^2}{\tilde{\epsilon}_r} + \frac{\|E_p^{I,N}\|_{L^2}^2}{\epsilon_l} + \frac{\|E_p^{I,1}\|_{L^2}^2}{\epsilon_l} + \frac{\|\partial_t E_p^I\|_{L^2(0,T;L^2)}^2}{\epsilon_r} \right). \quad (4.27)
 \end{aligned}$$

- Majoration de  $\Delta t B_2^n = \Delta t \sum_{K \in \tau_h} a^K(-E_{\mathbf{u}}^{\pi,n}, D^- E_{\mathbf{u}}^{A,n}) + \Delta t \sum_{K \in \tau_h} a_h(E_{\mathbf{u}}^{\pi,n} - E_{\mathbf{u}}^{I,n}, D^- E_{\mathbf{u}}^{A,n})$ .

On commence par traiter les formes bilinéaire  $a^K$ , dont on note  $M_K$  les constantes de continuité. Pour une fonction  $\mathbf{u}$  localement dans  $(H^1(K))^d$  et une fonction  $\mathbf{v}$  globalement dans  $(H^1(\Omega))^d$ , étant donné  $\epsilon > 0$ ,

$$\begin{aligned}
 \left| \sum_{K \in \tau_h} a^K(\mathbf{u}, \mathbf{v}) \right| &\leq \sum_{K \in \tau_h} (M_K \|\mathbf{u}\|_{H^1(K)} \|\mathbf{v}\|_{H^1(K)}) \leq \sum_{K \in \tau_h} \left( \frac{M_K^2}{2\epsilon} \|\mathbf{u}\|_{H^1(K)}^2 + \frac{\epsilon}{2} \|\mathbf{v}\|_{H^1(K)}^2 \right) \\
 &\leq \frac{M^2}{2\epsilon} \sum_{K \in \tau_h} \|\mathbf{u}\|_{H^1(K)}^2 + \frac{\epsilon}{2} \sum_{K \in \tau_h} \|\mathbf{v}\|_{H^1(K)}^2 = \frac{M^2}{2\epsilon} \|\mathbf{u}\|_{H^1(\tau)}^2 + \frac{\epsilon}{2} \|\mathbf{v}\|_{H^1(\Omega)}^2
 \end{aligned}$$

avec  $M = \max_K M_K$ . Puisque  $a^K$  est symétrique, une sommation par partie peut-être effectuée

$$\begin{aligned}
 2 \sum_{n=1}^N \Delta t \sum_{K \in \tau_h} a^K(-E_{\mathbf{u}}^{\pi,n}, D^- E_{\mathbf{u}}^{A,n}) &= 2 \sum_{K \in \tau_h} a^K(-E_{\mathbf{u}}^{\pi,N}, E_{\mathbf{u}}^{A,N}) - 2 \sum_{K \in \tau_h} a^K(-E_{\mathbf{u}}^{\pi,1}, E_{\mathbf{u}}^{A,0}) \\
 &\quad - 2 \sum_{n=1}^{N-1} \sum_{K \in \tau_h} a^K(-(E_{\mathbf{u}}^{\pi,n+1} - E_{\mathbf{u}}^{\pi,n}), E_{\mathbf{u}}^{A,n})
 \end{aligned}$$

puis combinée avec le résultat précédent de sorte que

$$\begin{aligned}
 \left| 2 \sum_{K \in \tau_h} a^K(-E_{\mathbf{u}}^{\pi,N}, E_{\mathbf{u}}^{A,N}) \right| &\leq \frac{8M^2}{\epsilon_l} \|E_{\mathbf{u}}^{\pi,N}\|_{H^1(\tau)}^2 + \frac{\epsilon_l}{8} \|E_{\mathbf{u}}^{A,N}\|_{H^1}^2, \\
 \left| 2 \sum_{K \in \tau_h} a^K(-E_{\mathbf{u}}^{\pi,1}, E_{\mathbf{u}}^{A,0}) \right| &\leq \frac{8M^2}{\epsilon_l} \|E_{\mathbf{u}}^{\pi,1}\|_{H^1(\tau)}^2 + \frac{\epsilon_l}{8} \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2.
 \end{aligned}$$

Pour le troisième terme, en suivant l'idée de l'outil 5,

$$\begin{aligned}
 2 \sum_{K \in \tau_h} |a^K(E_{\mathbf{u}}^{\pi, n+1} - E_{\mathbf{u}}^{\pi, n}, E_{\mathbf{u}}^{A, n})| &= 2 \sum_{K \in \tau_h} \left| a^K \left( \int_{t^n}^{t^{n+1}} \partial_t E_{\mathbf{u}}^{\pi} \right), E_{\mathbf{u}}^{A, n} \right| \leq 2 \sum_{K \in \tau_h} M_K \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{u}}^{\pi}\|_{H^1(K)} \|E_{\mathbf{u}}^{A, n}\|_{H^1(K)} \\
 &\leq 2 \sum_{K \in \tau_h} M_K \left( \Delta t \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{u}}^{\pi}\|_{H^1(K)}^2 \right)^{\frac{1}{2}} \|E_{\mathbf{u}}^{A, n}\|_{H^1(K)} \\
 &\leq \sum_{K \in \tau_h} \frac{8M^2}{\epsilon_r} \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{u}}^{\pi}\|_{H^1(K)}^2 + \Delta t \frac{\epsilon_r}{8} \|E_{\mathbf{u}}^{A, n}\|_{H^1(K)}^2 \\
 &= \frac{8M^2}{\epsilon_r} \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{u}}^{\pi}\|_{H^1(\tau)}^2 + \Delta t \frac{\epsilon_r}{8} \|E_{\mathbf{u}}^{A, n}\|_{H^1}^2.
 \end{aligned}$$

Les mêmes méthodes sont appliquées aux formes bilinéaires  $a_h^K$ , avec les notations  $\tilde{M}_K$  pour les constantes de continuité et  $\tilde{M} = \max_K M_K$ ,

$$\begin{aligned}
 2 \sum_{n=1}^N \Delta t \sum_{K \in \tau_h} a_h^K(E_{\mathbf{u}}^{\pi, n} - E_{\mathbf{u}}^{I, n}, D^- E_{\mathbf{u}}^{A, n}) &= 2 \sum_{K \in \tau_h} a_h^K(E_{\mathbf{u}}^{\pi, N} - E_{\mathbf{u}}^{I, N}, E_{\mathbf{u}}^{A, N}) - 2 \sum_{K \in \tau_h} a_h^K(E_{\mathbf{u}}^{\pi, 1} - E_{\mathbf{u}}^{I, 1}, E_{\mathbf{u}}^{A, 0}) \\
 &\quad - 2 \sum_{n=1}^{N-1} \sum_{K \in \tau_h} a_h^K((E_{\mathbf{u}}^{\pi, n+1} - E_{\mathbf{u}}^{\pi, n}) - (E_{\mathbf{u}}^{I, n+1} - E_{\mathbf{u}}^{I, n}), E_{\mathbf{u}}^{A, n}),
 \end{aligned}$$

les erreurs  $E^I$  et  $E^{\pi}$  étant traitées séparément :

$$\begin{aligned}
 2 \sum_{K \in \tau_h} |a_h^K(E_{\mathbf{u}}^{\pi, N} - E_{\mathbf{u}}^{I, N}, E_{\mathbf{u}}^{A, N})| &\leq \frac{8\tilde{M}^2}{\epsilon_l} \|E_{\mathbf{u}}^{\pi, N}\|_{H^1(\tau)}^2 + \frac{4\tilde{M}^2}{\epsilon_l} \|E_{\mathbf{u}}^{I, N}\|_{H^1}^2 + \left(\frac{\epsilon_l}{8} + \frac{\epsilon_l}{4}\right) \|E_{\mathbf{u}}^{A, N}\|_{H^1}^2, \\
 2 \sum_{K \in \tau_h} |a_h^K(E_{\mathbf{u}}^{\pi, 1} - E_{\mathbf{u}}^{I, 1}, E_{\mathbf{u}}^{A, 0})| &\leq \frac{8\tilde{M}^2}{\epsilon_l} \|E_{\mathbf{u}}^{\pi, 1}\|_{H^1(\tau)}^2 + \frac{4\tilde{M}^2}{\epsilon_l} \|E_{\mathbf{u}}^{I, 1}\|_{H^1}^2 + \left(\frac{\epsilon_l}{8} + \frac{\epsilon_l}{4}\right) \|E_{\mathbf{u}}^{A, 0}\|_{H^1}^2, \\
 2 \sum_{K \in \tau_h} a_h^K |(E_{\mathbf{u}}^{\pi, n+1} - E_{\mathbf{u}}^{\pi, n}, E_{\mathbf{u}}^{A, n})| &\leq \frac{8\tilde{M}^2}{\epsilon_r} \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{u}}^{\pi}\|_{H^1(\tau)}^2 + \frac{\epsilon_r}{8} \Delta t \|E_{\mathbf{u}}^{A, n}\|_{H^1}^2, \\
 2 \sum_{K \in \tau_h} a_h^K |(E_{\mathbf{u}}^{I, n+1} - E_{\mathbf{u}}^{I, n}, E_{\mathbf{u}}^{A, n})| &\leq \frac{4\tilde{M}^2}{\epsilon_r} \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{u}}^I\|_{H^1}^2 + \frac{\epsilon_r}{4} \Delta t \|E_{\mathbf{u}}^{A, n}\|_{H^1}^2.
 \end{aligned}$$

Finalement, on obtient pour  $B_2$  l'estimation

$$\begin{aligned}
 \left| 2\Delta t \sum_{n=1}^N B_2 \right| &\leq \frac{\epsilon_r}{2} \sum_{n=1}^{N-1} \Delta t \|E_{\mathbf{u}}^{A, n}\|_{H^1}^2 + \frac{\epsilon_l}{2} \|E_{\mathbf{u}}^{A, N}\|_{H^1}^2 + \frac{\epsilon_l}{2} \|E_{\mathbf{u}}^{A, 0}\|_{H^1}^2 + 4\tilde{M}^2 \left( \frac{\|E_{\mathbf{u}}^{I, 1}\|_{H^1}^2}{\epsilon_l} + \frac{\|E_{\mathbf{u}}^{I, N}\|_{H^1}^2}{\epsilon_l} + \frac{\|\partial_t E_{\mathbf{u}}^I\|_{L^2(0, T; H^1)}^2}{\epsilon_r} \right) \\
 &\quad + 8(M^2 + \tilde{M}^2) \left( \frac{\|E_{\mathbf{u}}^{\pi, 1}\|_{H^1(\tau)}^2}{\epsilon_l} + \frac{\|E_{\mathbf{u}}^{\pi, N}\|_{H^1(\tau)}^2}{\epsilon_l} + \frac{\|\partial_t E_{\mathbf{u}}^{\pi}\|_{L^2(0, T; H^1(\tau))}^2}{\epsilon_r} \right). \quad (4.28)
 \end{aligned}$$

- Majoration de  $\Delta t B_3^n = \Delta t \int_{\Omega} c_0 (D^- p^n - \partial_t p^n) E_p^{A, n} + \Delta t \int_{\Omega} \alpha \operatorname{div} (D^- \mathbf{u}^n - \partial_t \mathbf{u}^n) E_p^{A, n}$ .

Les différences entre les dérivées continues et discrètes sont remplacées par les développements de Taylor

(suivant l'outil 3)

$$\Delta t B_3^n = \int_{\Omega} c_0 \int_{t^{n-1}}^{t^n} \partial_{tt} p(s) (t^{n-1} - s)^2 ds E_p^{A,n} + \int_{\Omega} \alpha \operatorname{div} \left( \int_{t^{n-1}}^{t^n} \partial_{tt} \mathbf{u}(s) (t^{n-1} - s)^2 ds \right) E_p^{A,n}$$

puis la double identité de Cauchy-Schwarz (outil 5) est appliquée en notant cette fois que  $\int_{t^{n-1}}^{t^n} (t^{n-1} - s)^2 ds = \frac{\Delta t^3}{3}$ . Enfin, l'opérateur de divergence est majoré grâce à l'outil 2 ce qui donne

$$|\Delta t B_3^n| \leq c_0 \|E_p^{A,n}\|_{L^2} \left( \frac{\Delta t^3}{3} \int_{t^{n-1}}^{t^n} \|\partial_{tt} p(s)\|_{L^2}^2 ds \right)^{\frac{1}{2}} + \alpha \|E_p^{A,n}\|_{L^2} \left( \frac{d\Delta t^3}{3} \int_{t^{n-1}}^{t^n} \|\partial_{tt} \mathbf{u}(s)\|_{H^1}^2 ds \right)^{\frac{1}{2}}.$$

Procédant comme lors de l'estimation de  $B_1$ , l'inégalité de Young est appliquée avec  $\epsilon = \frac{1}{4}$  pour le dernier terme de la somme en temps et  $\epsilon = \frac{\epsilon_r}{4}$  pour les autres, d'où l'estimation de  $B_3$

$$\left| 2\Delta t \sum_{n=1}^N B_3^n \right| \leq \frac{2\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_p^{A,n}\|_{L^2}^2 + \frac{2\Delta t}{4} \|E_p^{A,N}\|_{L^2}^2 + \frac{4}{3} \Delta t^2 \left( c_0^2 \frac{\|\partial_{tt} p\|_{L^2(0,T;L^2)}^2}{\tilde{\epsilon}_r} + d\alpha^2 \frac{\|\partial_{tt} \mathbf{u}\|_{L^2(0,T;H^1)}^2}{\tilde{\epsilon}_r} \right). \quad (4.29)$$

- Majoration de  $\Delta t B_4^n = \Delta t \int_{\Omega} (\mathbf{f}^n - \mathbf{f}_h^n) \cdot D^- E_{\mathbf{u}}^{A,n}$ .

La différence  $(\mathbf{f}^n - \mathbf{f}_h^n)$  est assimilée à une erreur d'interpolation et notée  $E_{\mathbf{f}}^{I,n}$ . En sommant par parties,

$$2\Delta t \sum_{n=1}^N \int_{\Omega} E_{\mathbf{f}}^{I,n} \cdot D^- E_{\mathbf{u}}^{A,n} = 2 \int_{\Omega} E_{\mathbf{f}}^{I,N} E_{\mathbf{u}}^{A,N} - \int_{\Omega} 2E_{\mathbf{f}}^{I,1} E_{\mathbf{u}}^{A,0} - 2 \sum_{n=1}^{N-1} \int_{\Omega} (E_{\mathbf{f}}^{I,n+1} - E_{\mathbf{f}}^{I,n}) \cdot E_{\mathbf{u}}^{A,n}.$$

Chaque terme est majoré grâce aux inégalités de Cauchy-Schwarz et de Young

$$\begin{aligned} 2 \left| \int_{\Omega} E_{\mathbf{f}}^{I,N} E_{\mathbf{u}}^{A,N} \right| &\leq 2 \|E_{\mathbf{f}}^{I,N}\|_{L^2} \|E_{\mathbf{u}}^{A,N}\|_{L^2} \leq \frac{4}{\epsilon_l} \|E_{\mathbf{f}}^{I,N}\|_{L^2}^2 + \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,N}\|_{H^1}^2, \\ 2 \left| \int_{\Omega} E_{\mathbf{f}}^{I,1} E_{\mathbf{u}}^{A,0} \right| &\leq 2 \|E_{\mathbf{f}}^{I,1}\|_{L^2} \|E_{\mathbf{u}}^{A,0}\|_{L^2} \leq \frac{4}{\epsilon_l} \|E_{\mathbf{f}}^{I,1}\|_{L^2}^2 + \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2, \\ 2 \left| \int_{\Omega} (E_{\mathbf{f}}^{I,n+1} - E_{\mathbf{f}}^{I,n}) \cdot E_{\mathbf{u}}^{A,n} \right| &\leq \|E_{\mathbf{u}}^{A,n}\|_{L^2} \left( \Delta t \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{f}}^I\|_{L^2}^2 \right)^{\frac{1}{2}} \leq \frac{4}{\epsilon_r} \int_{t^n}^{t^{n+1}} \|\partial_t E_{\mathbf{f}}^I\|_{L^2}^2 + \frac{\epsilon_r}{4} \Delta t \|E_{\mathbf{u}}^{A,n}\|_{H^1}^2, \end{aligned}$$

puis il reste à remarquer que  $\|E_{\mathbf{u}}^A\|_{L^2} \leq \|E_{\mathbf{u}}^A\|_{H^1}$  pour obtenir la majoration sur  $B_4$  :

$$\left| 2\Delta t \sum_{n=1}^N B_4^n \right| \leq \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_{\mathbf{u}}^{A,n}\|_{H^1}^2 + \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,N}\|_{H^1}^2 + \frac{4}{\epsilon_r} \|\partial_t E_{\mathbf{f}}^I\|_{L^2(0,T;L^2)}^2 + \frac{4}{\epsilon_l} (\|E_{\mathbf{f}}^{I,N}\|_{L^2}^2 + \|E_{\mathbf{f}}^{I,1}\|_{L^2}^2) + \frac{\epsilon_l}{4} \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2. \quad (4.30)$$

- Majoration de  $\Delta t B_5^n = -\Delta t \sum_K \sum_{f \in \partial K} (\overline{F_{Kf}^n} - F_{Kf}^{*,n}) E_p^{A,n}$ .

Pour ce terme, on utilise la consistance des flux  $\overline{F_{Kf}^n} = -F_{Lf}^n$  et  $F_{Kf}^{*,n} = -F_{Lf}^{*,n}$ . Ceci permet de réorganiser la

double somme de la manière suivante, où  $K$  désigne l'une des deux mailles voisines de la face  $f$  :

$$\begin{aligned} \sum_K \sum_{f \subset \partial K} (\overline{F_{Kf}^n} - F_{Kf}^{*,n}) E_p^{A,n} &= \sum_{f \subset \overline{\Omega} \setminus \Gamma_{N_p}} (\overline{F_{Kf}^n} - F_{Kf}^{*,n}) D_f(E_p^{A,n}) \\ &\leq \left( \sum_{f \subset \overline{\Omega} \setminus \Gamma_{N_p}} T_f^{-1} |\overline{F_{Kf}^n} - F_{Kf}^{*,n}|^2 \right)^{\frac{1}{2}} \left( \sum_{f \subset \overline{\Omega} \setminus \Gamma_{N_p}} T_f D_f(E_p^{A,n})^2 \right)^{\frac{1}{2}} \quad (\text{Cauchy-Schwarz}) \\ &\leq \frac{|E_p^{A,n}|_{1,\tau}^2}{2} + \sum_{f \subset \overline{\Omega} \setminus \Gamma_{N_p}} \frac{|\overline{F_{Kf}^n} - F_{Kf}^{*,n}|^2}{2T_f} \quad (\text{Young}). \end{aligned}$$

Le terme  $\overline{F_{Kf}^n} - F_{Kf}^{*,n}$  ne dépend que de la solution exacte et fait donc partie des erreurs d'interpolation. Il demeure au membre de droite, et la majoration sur  $B_5$  est directement donnée par

$$\left| 2\Delta t \sum_{n=1}^N B_5^n \right| \leq \sum_{n=1}^N \Delta t |E_p^{A,n}|_{1,\tau}^2 + \sum_{n=1}^N \Delta t \sum_{f \subset \overline{\Omega} \setminus \Gamma_{N_p}} \frac{|\overline{F_{Kf}^n} - F_{Kf}^{*,n}|^2}{T_f}. \quad (4.31)$$

- Majoration de  $\Delta t B_6^n = -\Delta t c_0 \int_{\Omega} E_p^{A,n} D^- E_p^{I,n}$ .

Ce terme se traite simplement en utilisant la double identité de Cauchy-Schwarz puis l'inégalité de Young

$$|\Delta t B_6^n| \leq c_0 \|E_p^{A,n}\|_{L^2} \left( \Delta t \int_{t^{n-1}}^{t^n} \|\partial_t E_p^I\|_{L^2}^2 \right)^{\frac{1}{2}} \leq \frac{\epsilon \Delta t}{2} \|E_p^{A,n}\|_{L^2}^2 + \frac{c_0^2}{2\epsilon} \int_{t^{n-1}}^{t^n} \|\partial_t E_p^I\|_{L^2}^2,$$

puis en fixant  $\epsilon = \frac{1}{4}$  pour  $n = N$  et  $\epsilon = \frac{\epsilon_r}{4}$  pour les autres  $n$ . On obtient donc l'estimation sur  $B_6$

$$\left| 2\Delta t \sum_{n=1}^N B_6^n \right| \leq \frac{\epsilon_r}{4} \sum_{n=1}^{N-1} \Delta t \|E_p^{A,n}\|_{L^2}^2 + \frac{\Delta t}{4} \|E_p^{A,N}\|_{L^2}^2 + \frac{4c_0^2}{\epsilon_r} \|\partial_t E_p^I\|_{L^2(0,T;L^2)}^2. \quad (4.32)$$

- En combinant les inégalités (4.27) à (4.32),

$$(\tilde{\mu} - \epsilon_l) \|E_{\mathbf{u}}^{A,N}\|_{H^1}^2 + (c_0 - \Delta t) \|E_p^{A,N}\|_{L^2}^2 + \sum_{n=0}^N \Delta t |E_p^{A,n}|_{1,\tau}^2 \leq \epsilon_r \sum_{n=1}^{N-1} \Delta t \left( \|E_{\mathbf{u}}^{A,n}\|_{H^1}^2 + \|E_p^{A,n}\|_{L^2}^2 \right) + \tilde{B} \Delta t^2 + \tilde{B}' \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2 + \tilde{E}^I$$

où  $\tilde{B}$  et  $\tilde{B}'$  dépendent de  $c_0, \alpha, d, \tilde{\mu}, \partial_{tt} p, \partial_{tt} \mathbf{u}, \epsilon_l$  et  $\epsilon_r$ . Le terme  $\tilde{E}^I$  regroupe les erreurs d'interpolation ou assimilées ainsi que l'erreur liée aux volumes finis (on note que même en dimension 3, le terme  $T_f^{-1}$  d'ordre  $\mathcal{O}(h^{-1})$  s'élimine avec l'erreur due au flux et ne diverge donc pas) :

$$\begin{aligned} \tilde{E}^I &= \frac{2\tilde{M}}{\epsilon_l} \|E_{\mathbf{u}}^I\|_{L^\infty(0,T;H^1)}^2 + \left( \frac{4\tilde{M}^2}{\epsilon_r} + \frac{4d\alpha^2}{\tilde{\epsilon}_r} \right) \|\partial_t E_{\mathbf{u}}^I\|_{L^2(0,T;H^1)}^2 \\ &+ \frac{4(M^2 + \tilde{M}^2)}{\epsilon_l} \|E_{\mathbf{u}}^{\pi}\|_{L^\infty(0,T;H^1(\tau))}^2 + \frac{8(M^2 + \tilde{M}^2)}{\epsilon_r} \|\partial_t E_{\mathbf{u}}^{\pi}\|_{L^2(0,T;H^1(\tau))}^2 + \frac{2}{\epsilon_l} \|E_{\mathbf{f}}^I\|_{L^\infty(0,T;L^2)}^2 + \frac{4}{\epsilon_r} \|\partial_t E_{\mathbf{f}}^I\|_{L^2(0,T;L^2)}^2 \\ &+ \frac{2d\alpha^2}{\epsilon_l} \|E_p^I\|_{L^\infty(0,T;L^2)}^2 + \left( \frac{4d\alpha^2}{\epsilon_r} + \frac{4c_0^2}{\tilde{\epsilon}_r} \right) \|\partial_t E_p^I\|_{L^2(0,T;L^2)}^2 + \sum_{n=1}^N \Delta t \sum_{f \subset \overline{\Omega} \setminus \Gamma_{N_p}} \frac{|\overline{F_{Kf}^n} - F_{Kf}^{*,n}|^2}{T_f}. \end{aligned}$$

### Troisième partie – conclusion

On utilise à présent l'hypothèse  $\Delta t \leq \frac{c_0}{2}$  qui implique que  $\frac{c_0}{2} \leq c_0 - \Delta t$  et permet d'absorber le  $\Delta t$  dans le membre de gauche. On fixe ensuite  $\epsilon_l$  de sorte que  $\epsilon_l < \tilde{\mu}$  puis  $\epsilon_r = \min\{\tilde{\mu} - \epsilon_l, \frac{c_0}{2}\}$ . En divisant l'inégalité précédente par  $\epsilon_r$ , il vient

$$\|E_{\mathbf{u}}^{A,N}\|_{H^1}^2 + \|E_p^{A,N}\|_{L^2}^2 + \sum_{n=0}^N \Delta t |E_p^{A,n}|_{1,\tau}^2 \leq \sum_{n=0}^{N-1} \Delta t \left( \|E_{\mathbf{u}}^{A,n}\|_{H^1}^2 + \|E_p^{A,n}\|_{L^2}^2 \right) + B\Delta t^2 + B' \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2 + E^I$$

avec  $B = \frac{\tilde{B}}{\epsilon_r}$  et  $B' = \frac{\tilde{B}'}{\epsilon_r}$  dépendant de  $c_0, \alpha, d, \tilde{\mu}, \partial_{tt}p$ , et  $\partial_{tt}\mathbf{u}$  et

$$E^I = \frac{1}{\epsilon_r} \tilde{E}^I. \quad (4.33)$$

La quantité positive  $\Delta t (\|E_{\mathbf{u}}^{A,0}\|_{H^1}^2 + \|E_p^{A,0}\|_{L^2}^2)$  a été ajoutée au membre de droite de manière à conclure grâce à une version discrète du lemme de Grönwall :

**Lemme 4.3** Soit  $(a_n)_{n \geq 0}, (b_n)_{n \geq 0}$  deux suites de réels positifs et  $C \in \mathbb{R}^+$ . Supposons que

$$a_0 + b_0 \leq C$$

et qu'il existe  $\lambda > 0$  tel que

$$\forall N \geq 1, \quad a_N + b_N \leq \lambda \sum_{m=0}^{N-1} a_m + C.$$

Alors,

$$\forall N \geq 0, \quad a_N + b_N \leq C \exp(\lambda N).$$

En appliquant le lemme avec  $a_n = \|E_{\mathbf{u}}^{A,n}\|_{H^1}^2 + \|E_p^{A,n}\|_{L^2}^2$ ,  $b_n = \sum_{m=0}^n \Delta t |E_p^{A,m}|_{1,\tau}^2$ ,  $\lambda = \Delta t$  et  $C = B\Delta t^2 + B' \|E_{\mathbf{u}}^{A,0}\|_{H^1}^2 + E^I$ , on aboutit au résultat (4.19).

#### 4.2.3 Majoration des erreurs d'interpolation

Pour achever l'analyse d'erreur, il reste à borner les quantités regroupées dans le terme  $E^I$ . Dans cette section,  $C$  et  $\tilde{C}$  désignent des constantes positives indépendantes du pas du maillage variant selon le contexte.

#### Opérateurs d'interpolation

Dans un premier temps, on remarque que les opérateurs d'interpolation et de dérivation en temps commutent : en effet, puisque les fonctions de base  $\varphi$  ne dépendent pas du temps, on a sur chaque élément  $K$ ,

$$\frac{\partial(\tilde{P}\mathbf{u})}{\partial t} = \sum_{i=1}^{M^K} \frac{\partial}{\partial t} \left( u_i^x \varphi_i^x(\mathbf{x}) + u_i^y \varphi_i^y(\mathbf{x}) + u_i^z \varphi_i^z(\mathbf{x}) \right) = \tilde{P}(\partial_t \mathbf{u}).$$

De la même manière, en développant l'expression de  $\pi^K$  sur chaque maille on montre que  $\partial_t \pi(\mathbf{u}) = \pi(\partial_t \mathbf{u})$ , l'argument principal étant que le gradient en espace et la dérivée en temps commutent pour des fonctions suffisamment régulières. La démonstration est également directe pour l'interpolation de la pression :  $\partial_t P(p) = P(\partial_t p)$ .



L'erreur d'interpolation  $E_{\mathbf{u}}^I = \mathbf{u} - \tilde{P}\mathbf{u}$  est majorée en utilisant les résultats classiques d'interpolation, voir par exemple [All12, prop. 6.3.16] ou [DS80] : si  $\mathbf{u} \in H^2(\Omega)$ , il existe une constante  $C$  indépendante de  $h$  et de  $\mathbf{u}$  telle que

$$\|\mathbf{u} - \tilde{P}\mathbf{u}\|_{H^1(\Omega)} = \|E_{\mathbf{u}}^I\|_{H^1(\Omega)} \leq Ch\|\mathbf{u}\|_{H^2(\Omega)}. \quad (4.34)$$

Puisque l'erreur sur le second membre fait intervenir le même opérateur d'interpolation, on a cette fois en norme  $L^2$

$$\|\mathbf{f} - \tilde{P}\mathbf{f}\|_{L^2(\Omega)} = \|E_{\mathbf{f}}^I\|_{L^2(\Omega)} \leq Ch^2\|\mathbf{f}\|_{H^2(\Omega)}.$$

Les deux derniers opérateurs  $P$  et  $\pi^K$  conservant respectivement les constantes et les fonctions affines, d'après la théorie de Scott-Dupont [DS80],

$$\|p - Pp\|_{L_2(\Omega)} = \|E_p^I\|_{L_2(\Omega)} \leq Ch\|p\|_{H^1(\Omega)}$$

et, sur chaque maille  $K$ ,

$$\|\mathbf{u}|_K - \pi^K \mathbf{u}|_K\|_{H^1(K)} \leq Ch_K\|\mathbf{u}\|_{H^2(K)} \quad \text{d'où} \quad \|E_{\mathbf{u}}^\pi\|_{H^1(\tau)} \leq Ch\|\mathbf{u}\|_{H^2(\Omega)}.$$

D'après ce qui précède, on a des estimations similaires pour les dérivées temporelles  $\|\partial_t(E_{\mathbf{u}}^I)\|_{H^1(\Omega)}$ ,  $\|\partial_t(E_p^I)\|_{L^2(\Omega)}$ ,  $\|\partial_t(E_p^I)\|_{L^2(\Omega)}$  et  $\|\partial_t(E_{\mathbf{u}}^\pi)\|_{h,H^1(\Omega)}$  avec cette fois la norme des dérivées  $\partial_t \mathbf{u}$ ,  $\partial_t \mathbf{f}$  et  $\partial_t p$  dans le second membre.

Ces bornes sont valables à chaque pas de temps, ce qui permet de majorer les normes  $L^\infty$  et  $L^2$  dans  $[0, T]$  indépendamment pas de temps : formellement, pour  $p = 1$  ou  $p = 2$ ,

$$\|E_v\|_{L^\infty(0,T;H^p(\Omega))} = \sup_{n \in [1,N]} \|E_v^n\|_{H^p(\Omega)} \leq \sup_{n \in [1,N]} C^n h \|v^n\|_{H^{p+1}(\Omega)} \leq \tilde{C}h \text{ avec } \tilde{C} = \sup_{n \in [1,N]} C^n \|v^n\|_{H^{p+1}(\Omega)}$$

et

$$\|E_v\|_{L^2(0,T;H^p(\Omega))} = \left( \sum_{n=1}^N \Delta t \|E_v^n\|_{H^p}^2 \right)^{\frac{1}{2}} \leq \left( \sum_{n=1}^N \Delta t (C^n h \|v^n\|_{H^{p+1}})^2 \right)^{\frac{1}{2}} \leq \tilde{C}h \text{ avec } \tilde{C} = t^N \sup_{n \in [1,N]} C^n \|v^n\|_{H^{p+1}(\Omega)}.$$

### Erreur d'interpolation issue des volumes finis

Le terme issu de la discrétisation volumes finis s'écrit

$$\sum_{n=1}^N \Delta t \sum_{f \in \overline{\Omega} \setminus \Gamma_{N_p}} \frac{|F_{Kf}^n - F_{Kf}^{*,n}|^2}{T_f} = \sum_{n=1}^N \Delta t \sum_{f \in \overline{\Omega} \setminus \Gamma_{N_p}} \frac{|f| d_{KL}}{\kappa^{eq}} \frac{|F_{Kf}^n - F_{Kf}^{*,n}|^2}{|f|^2}.$$

L'analyse du schéma volumes finis à deux points [EGH00, section 3.1.5] montre l'existence d'une constante  $C \in \mathbb{R}^+$  dépendant seulement de la pression exacte  $p(\mathbf{x}, t)$  telle que  $\frac{|F_{Kf}^n - F_{Kf}^{*,n}|}{|f|} \leq Ch$ . Cette constante inclut la norme des dérivées spatiales jusqu'au second ordre de  $p$ . De plus,  $\sum |f| d_{KL} \leq d|\Omega|$ , d'où en notant  $\tilde{C} = \frac{d|\Omega|C^2}{\min(\kappa^{eq})}$ ,

$$\sum_{n=1}^N \Delta t \sum_{f \in \overline{\Omega} \setminus \Gamma_{N_p}} \frac{|F_{Kf}^n - F_{Kf}^{*,n}|^2}{T_f} \leq \sum_{n=1}^N \Delta t \frac{C^2 h^2}{\min(\kappa^{eq})} \sum_{f \in \overline{\Omega} \setminus \Gamma_{N_p}} |f| d_{KL} \leq \sum_{n=1}^N \Delta t \frac{C^2 h^2 d |\Omega|}{\min(\kappa^{eq})} \leq \tilde{C} t^N h^2.$$

L'erreur d'interpolation (au carré) liée au schéma volumes finis est donc également de l'ordre de  $h^2$ .

Enfin, la solution initiale du déplacement  $\mathbf{u}^0$  étant calculée en résolvant par éléments virtuels le problème élastique seul, le lemme 3.1 permet de majorer  $\|E_{\mathbf{u}}^{A,0}\|_{H^1}^2$  par un terme lui aussi d'ordre  $h^2$ . Il en résulte le théorème 4.4 :

**Théorème 4.4** *Sous les hypothèses de la proposition 4.2, l'erreur d'approximation du schéma satisfait*

$$\left( \|E_{\mathbf{u}}^A\|_{L^\infty(0,T;H^1(\Omega))}^2 + \|E_p^A\|_{L^\infty(0,T;L^2(\Omega))}^2 + |E_p^A|_{L^2(0,T;H^1(\tau))}^2 \right)^{\frac{1}{2}} \leq C(h + \Delta t) \quad (4.35)$$

où  $C$  dépend des constantes d'ellipticité de  $a$  et  $a_h$ , des paramètres physiques  $c_0, \alpha, d, |\Omega|, t^N$  et des normes de la solution exacte  $(\mathbf{u}, p)$  et de ses dérivées.

### 4.3 Implémentation et validation numérique

Pour ce premier schéma couplé, la partie la plus compliquée à implémenter reste celle des éléments virtuels. Celle-ci a justement été détaillée dans le chapitre précédent. Dans cette section, on ajoute quelques détails concernant la forme du système linéaire couplé, les développements relatifs aux volumes finis deux points et le traitement des termes de couplage. Quelques vérifications et premiers tests sont effectués, avant d'illustrer numériquement l'ordre de convergence obtenu dans le théorème 4.4.

#### 4.3.1 Implémentation du schéma

De la même manière que pour la discrétisation du système élastique, le vecteur déplacement  $\mathbf{u}_h$  s'écrit dans la base des éléments virtuels  $\mathbf{u}_h = \sum_{i \in \mathcal{M}} u_i^x \boldsymbol{\varphi}_i^x + u_i^y \boldsymbol{\varphi}_i^y + u_i^z \boldsymbol{\varphi}_i^z$  où  $\mathcal{M}$  représente l'ensemble des indices de sommets. De même, en introduisant pour la pression l'ensemble  $\mathfrak{M}$  des index des  $\# \tau_h$  mailles et la base  $\{\phi_k\}_{k \in \mathfrak{M}}$  dont les fonctions sont définies par

$$\phi_k(\mathbf{x}) = \begin{cases} 1 & \text{si } \mathbf{x} \text{ appartient à la maille d'indice } k \\ 0 & \text{sinon,} \end{cases}$$

on a la décomposition  $p_h = \sum_{k \in \mathfrak{M}} p_k \phi_k$ .

En regroupant les inconnues du déplacement et celles de la pression sous la forme d'un vecteur  $[\mathcal{U}, \mathcal{P}]$ , le schéma (4.7) s'écrit sous forme matricielle

$$\begin{bmatrix} \mathcal{A} & -\mathcal{B} \\ \mathcal{B}^T & \mathcal{F} \end{bmatrix} \begin{bmatrix} \mathcal{U} \\ \mathcal{P} \end{bmatrix} = \begin{bmatrix} \mathcal{L}^u \\ \mathcal{L}^p + \mathcal{L}^c \end{bmatrix} \quad (4.36)$$

où les blocs correspondent aux quantités suivantes :

- La matrice  $\mathcal{A}$  carrée de taille  $d \cdot \text{card}(\mathcal{M}_{\overline{D}_d})$  est exactement la matrice de rigidité des éléments virtuels qui aurait été obtenue pour le problème élastique, telle que définie dans la section 3.1.2 et calculée par la méthode décrite dans la section 3.2.
- La matrice  $\mathcal{F}$  carrée de taille  $\text{card}(\mathfrak{M})$  est la matrice associée au schéma volumes finis deux points et serait celle obtenue pour la discrétisation de l'équation de conservation seule. Ses coefficients s'écrivent, en notant  $\nu(K)$  l'ensemble des mailles voisines de  $K$  et  $T_f$  la transmissivité entre les mailles  $K$  et  $L$  définie par (4.11),

$$\mathcal{F}_{KK} = c_0 |K| + \Delta t \sum_{L \in \nu(K)} T_f \quad \text{et} \quad \mathcal{F}_{KL} = -\Delta t T_f. \quad (4.37)$$

- Le bloc  $\mathcal{B}$  est une matrice de couplage rectangulaire de taille  $d \cdot \text{card}(\mathcal{M}_{\overline{D_d}}) \times \text{card}(\mathfrak{M})$  définie par

$$\begin{aligned} \mathcal{B}(\boldsymbol{\varphi}_j^l, \phi_k) &= \alpha \int_K \phi_k \text{div}(\boldsymbol{\varphi}_j^l) = \alpha \int_K \text{div}(\boldsymbol{\varphi}_j^l) \text{ par définition de } \phi_k \\ &= \alpha |K| \langle \nabla \varphi_j \rangle^K \cdot \mathbf{e}_l. \end{aligned} \quad (4.38)$$

où le calcul de la quantité  $\langle \nabla \varphi_i \rangle^K$  sur la maille  $K$  d'indice  $k$  a déjà été détaillé section 3.1.3. Le calcul de ce terme est de toute façon requis pour l'assemblage de la matrice des éléments virtuels et n'entraîne donc pas de coût de calcul supplémentaire.

- Le vecteur  $\mathcal{L}^u$  de taille  $d \cdot \text{card}(\mathcal{M}_{\overline{D_d}})$  est le second membre de l'équation élastique. Il est de la même forme que pour le système élastique, la subtilité étant que le terme imposé de Neumann représente ici la contrainte totale et non la contrainte effective. Dans le cas  $\alpha = 0$ , on retrouve le même second membre que celui du problème élastique.
- Le vecteur  $\mathcal{L}^p + \mathcal{L}^c$  de taille  $\text{card}(\mathfrak{M})$  constitue le second membre de l'équation de conservation. La première partie,  $\mathcal{L}^p$ , correspond au second membre de l'équation de conservation considérée seule, tandis que la seconde partie  $\mathcal{L}^c$  permet de tenir compte du terme de couplage  $\alpha \int_K \text{div}(\mathbf{u}_h^{n-1})$  issu de la discrétisation en temps. Ces deux termes s'expriment respectivement selon

$$\mathcal{L}_K^p = |K|(c_0 p_K^{n-1} + \Delta t q_K^n) + |f| \Delta t \sum_{f \subset \partial K \cap \Gamma_{N_p}} \phi_N^n(\mathbf{x}^f) \quad \text{et} \quad \mathcal{L}_K^c = |K| \alpha \sum_{i \in \mathcal{M}_K} \langle \nabla \varphi_i \rangle^K \cdot \begin{pmatrix} u_i^x \\ u_i^y \\ u_i^z \end{pmatrix}^{n-1}.$$

Dans le terme  $\mathcal{L}^p$ ,  $f$  désigne une face de bord en trois dimensions ou une arête de bord en deux dimensions. Les conditions aux limites de Neumann sont donc intégrées en évaluant la fonction  $\phi_N^n$  au barycentre de l'entité de bord. Notons que dans le cas non couplé,  $\mathcal{L}^c$  est nul et le second membre est réduit à  $\mathcal{L}^p$ .

L'assemblage du système couplé réutilise l'implémentation des éléments virtuels (voir section 3.2) et est décrit dans l'algorithme 2. Par rapport au chapitre précédent, il reste à détailler l'assemblage de la matrice associée aux volumes finis.

Dans un premier temps, une boucle sur les faces du maillage est réalisée. En interrogeant la connectivité de l'objet face, on peut obtenir les deux objets cellules  $K$  et  $L$  séparées par cette face (ou l'unique objet  $K$  dans le cas où il s'agit d'une face de bord). Ceci permet, à l'aide des caractéristiques des cellules, de calculer la transmissivité  $T_f$  d'après la formule (4.11). Cette transmissivité est ensuite ajoutée dans la matrice  $\mathcal{F}$  selon

$$\begin{aligned} \mathcal{F}_{KK} &+= \Delta t T_f & \mathcal{F}_{KL} &-= \Delta t T_f \\ \mathcal{F}_{LL} &+= \Delta t T_f & \mathcal{F}_{LK} &-= \Delta t T_f \end{aligned}$$

dans le cas où la face  $f$  est interne et selon

$$\mathcal{F}_{KK} += \Delta t T_f \quad \text{et} \quad \mathcal{L}_K^p += \Delta t T_f p_D(\mathbf{x}^f)$$

dans le cas où  $f$  est une face de Dirichlet. Dans un second temps, une boucle sur les cellules est effectuée pour ajouter le terme  $\mathcal{F}_{KK} += c_0 |K|$  à la matrice.

Notons qu'en pratique, les termes  $\langle \nabla \varphi_i \rangle^K$  et les transmissivités  $T_f$  sont calculés une seule fois en début de simulation et ne sont recalculés que si la géométrie est modifiée.

**Algorithm 2** Implémentation du schéma couplé VEM-TFPA

---

```

1 Read mesh and compute each cell  $\diamond$  geometry
2 Initialisation :  $p^0$  being given, run algorithm 1 with load term  $\mathbf{f} - \alpha \mathbf{div}(p\bar{l}_d)$  to find  $\mathbf{u}^0$ 
3 while  $t < T$  do
4   for  $\diamond$  in mesh do
5     Compute local Vem matrix  $\mathcal{A}^\diamond$ . Store  $\langle \nabla \bullet \rangle^\diamond$  values! ▷ see section 3.2
6     Compute local coupling vector  $B^\diamond$ .
7     Compute local rhs  $\mathcal{L}^u_\diamond$ 
8     Compute  $\mathcal{L}^p(\diamond)$  and  $\mathcal{L}^c(\diamond)$  using  $\langle \nabla \bullet \rangle^\diamond$ ,  $\mathcal{P}^{n-1}$  and  $\mathcal{U}^{n-1}$ 
9     for non Dirichlet dof  $\bullet$  in  $\diamond$  do
10       $\mathcal{L}^u(\bullet) += \mathcal{L}^u_\diamond(\bullet)$ 
11      for non Dirichlet  $\bullet$  in  $\diamond$  do
12         $\mathcal{A}(\bullet, \bullet) += \mathcal{A}^\diamond(\bullet, \bullet)$ 
13        Set  $\mathcal{B}(\bullet, \diamond) = B^\diamond(\bullet)$ 
14      for Dirichlet  $\bullet$  in  $\diamond$  do
15         $\mathcal{L}^u(\bullet) -= \mathcal{A}^\diamond(\bullet, \bullet)g_D(\bullet)$ 
16         $\mathcal{L}^c(\diamond) -= B^\diamond(\bullet)g_D(\bullet)$ 
17   for each face  $l$  in mesh do
18     Compute transmissivity and assemble  $\mathcal{F}$  ▷ see above
19   Solve  $(d \cdot \text{card}(\mathcal{M}_{D_d}^-) + \text{card}(\mathfrak{M}))$  linear system  $\begin{bmatrix} \mathcal{A} & -B \\ B^T & \mathcal{F} \end{bmatrix} \begin{bmatrix} \mathcal{U} \\ \mathcal{P} \end{bmatrix} = \begin{bmatrix} \mathcal{L}^u \\ \mathcal{L}^p + \mathcal{L}^c \end{bmatrix}$ 
20   Complete  $\mathcal{U}$  with Dirichlet  $\bullet$  values
21   for  $\diamond$  in mesh do
22     Compute stress vector and flux vector
23   Compute errors
24    $t \leftarrow t + \Delta t$ 

```

---

Concernant le post traitement, le calcul de la contrainte mécanique a déjà été détaillé dans le paragraphe 3.2. Pour la variable en pression, la vitesse  $v_f$  est reconstruite sur chaque maille  $K$  à l'aide de la formule

$$v_f^K = \frac{1}{|K|} \sum_{f \subset \partial K} \mathcal{F}_{Kf}^n(\mathbf{x}^f - \mathbf{x}^K) \quad (4.39)$$

où  $\mathcal{F}_{Kf}^n$  est le flux numérique de la méthode défini par l'expression (4.11) et évalué avec la pression calculée, et où  $\mathbf{x}^f$  et  $\mathbf{x}^K$  désignent respectivement le barycentre de la face  $f$  et de la maille  $K$ . En pratique, de la même manière que lors de l'assemblage du système, une boucle sur les faces du maillage est effectuée pour calculer les flux numériques et ajouter cette contribution aux mailles attenantes à la face.

En plus des écarts en déplacement et en contrainte définis pour l'élasticité (3.26), on définit les écarts absolus et relatifs en pression  $E_{\text{abs}}^p$  et  $E_{\text{rel}}^p$  – comparables à des écarts dans l'espace  $L^2$  – ainsi que les écarts absolus et relatifs en flux  $E_{\text{abs}}^{\nabla p}$  et  $E_{\text{rel}}^{\nabla p}$  par les expressions

$$E_{\text{abs}}^p = \|p_h - p\|_{L^2}, \quad E_{\text{rel}}^p = \frac{\|p_h - p\|_{L^2}}{\|p\|_{L^2}}, \quad E_{\text{abs}}^{\nabla p} = \|\bar{\kappa} \nabla p_h - \bar{\kappa} \nabla p\|_{L^2} \quad \text{et} \quad E_{\text{rel}}^{\nabla p} = \frac{\|\bar{\kappa} \nabla p_h - \bar{\kappa} \nabla p\|_{L^2}}{\|\bar{\kappa} \nabla p\|_{L^2}}. \quad (4.40)$$

Ces écarts sont calculés en approchant l'intégrale sur chaque maille par la valeur des pressions et des gradients au centre de ces mailles : par exemple,

$$\|p_h - p\|_{L^2} = \left( \sum_{K \in \tau_h} \int_K (p_h - p)^2 \right)^{\frac{1}{2}} \approx \left( \sum_{K \in \tau_h} |K| (p_{h_K} - p(\mathbf{x}^K))^2 \right)^{\frac{1}{2}}.$$

La même technique est utilisée pour la norme  $L^2$  de la différence des gradients, en utilisant la reconstruction par cellule du gradient définie par l'équation (4.39). Néanmoins, la grandeur  $E_{\text{abs}}^{\nabla p}$  ne fait pas forcément office de norme  $H^1$  et cet indicateur est donc complété par un dernier écart calculé à partir de la semi-norme  $H^1$  discrète spécifique au schéma définie par (4.13) :

$$E_{\text{abs}}^{\tau_v} = |p_h - p|_{1,\tau} = \left( \sum_{f \subset \Omega} T_f \left( (p_{h_K} - p(\mathbf{x}^K)) - (p_{h_L} - p(\mathbf{x}^L)) \right)^2 + \sum_{f \subset \Gamma_{Dp}} T_f \left( p_{h_K} - p(\mathbf{x}^K) \right)^2 \right)^{\frac{1}{2}} \quad (4.41)$$

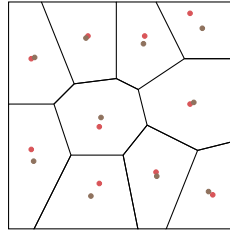
et

$$E_{\text{rel}}^{\tau_v} = \frac{E_{\text{abs}}^{\tau_v}}{|p|_{1,\tau}}.$$

### 4.3.2 Patch test et limites de l'approximation à deux points

**Validation : système découplé** Dans le cas où  $\alpha = 0$ , on a souligné que les seconds membres sont ceux des deux problèmes considérés seuls. Il est clair que, dans ce cas, la matrice  $\mathcal{B}$  est également nulle. On a alors un système diagonal par bloc, dont la résolution équivaut à résoudre séparément un problème de mécanique et un problème d'écoulement indépendants. Cette propriété permet de vérifier l'implémentation en comparant le code couplé aux codes éléments virtuels pour l'élasticité et volumes finis pour l'écoulement.

**Patch test pour le système couplé** Dans la section 3.3, la capacité des éléments virtuels à satisfaire le *patch test* était évoquée, c'est à dire à retrouver la solution exacte du problème lorsque celle-ci est polynomiale de degré un. On peut se demander si la discrétisation employée pour le système poroélastique satisfait également cette propriété. Cependant, avant même de considérer le système couplé, il faut avoir à l'esprit une limite des maillages de Voronoï. Un maillage de Voronoï est construit à partir d'un ensemble de points, appelés centres de Voronoï, de telle manière que chaque arête du maillage soit la médiatrice du segment formé par les deux centres de Voronoï des mailles adjacentes. Ces points sont donc particulièrement adaptés pour la discrétisation volumes finis deux points puisque par construction, ils vérifient la condition d'orthogonalité. En revanche, ces centres de Voronoï sont en général distincts des centres de masse des mailles comme illustré sur la figure 4.1. Or, lorsqu'une intégration numérique est nécessaire (par exemple pour les



• Centre de masse • Centre de Voronoï

FIG. 4.1 : **Maillage Voronoï** simplifié avec représentation des centres de Voronoï et des centres de masses. Ces points sont distincts dans le cas général.

calculs d'erreurs), la formule d'approximation utilisée se base sur le centre de masse :  $\int_K f \approx |K|f(\mathbf{x}^K)$ . Par conséquent, des erreurs numériques supplémentaires apparaissent lors de l'intégration.

Pour illustrer ce phénomène, on considère tout d'abord sur le carré unité l'équation de conservation seule. On choisit la solution exacte linéaire  $p = -x + 2y$ , le coefficient  $\kappa$  est fixé à 1 et des conditions de Dirichlet sont imposées sur tous les bords du domaine. Avec un maillage carré, les écarts  $E_{\text{rel}}^p$  et  $E_{\text{rel}}^{\nabla p}$  sont bien nuls. Ce résultat, de référence, traduit la capacité du schéma volumes finis à approcher exactement des fonctions linéaires. Avec un maillage de Voronoï, plusieurs cas sont à distinguer. Si les centres de masse sont utilisés comme points de discrétisation pour la pression, les écarts  $E_{\text{rel}}^p$  et  $E_{\text{rel}}^{\nabla p}$  sont significativement non nuls (de l'ordre de  $10^{-2}$ ). En effet, la condition d'orthogonalité n'étant pas respectée dans cette situation, il en résulte une mauvaise approximation des flux numériques. Si les centres de Voronoï sont utilisés pour discrétiser la pression, la condition d'orthogonalité est bien respectée. Cependant, on observe dans ce cas un écart en pression  $E_{\text{rel}}^p$  non nul et un écart en flux  $E_{\text{rel}}^{\nabla p}$  nul. En réalité, comparer la solution calculée à la solution exacte sur les points de discrétisation permet de se rendre compte que les deux valeurs sont identiques. L'écart sur  $E_{\text{rel}}^p$  provient de la formule d'intégration utilisée pour le calcul de l'erreur, qui calcule la solution de référence aux centres de masse. D'ailleurs, modifier la formule d'intégration pour forcer l'évaluation de la solution exacte aux points de Voronoï permet d'obtenir un écart nul. Malheureusement, dans le cas général, modifier cette formule d'intégration ne donne plus une approximation consistante de l'intégrale. On note également que l'écart nul pour le gradient de pression s'explique par le fait que  $\nabla p$  est constant sur ce cas et que la différence de position entre les deux types de centres n'a donc pas d'importance.

Ce problème se répercute également sur la résolution de cas couplés. On complète l'équation précédente par l'équation mécanique avec une solution exacte linéaire donnée par  $\mathbf{u} = (3x - 2y, x + y)$ , et des coefficients de couplage  $\alpha = 1, c_0 = 0.5, E = 2.5, \nu = 0.25, \kappa = 1$ . Avec un maillage constitué de carrés, tous

Méthode	$E_{\text{rel}}^p$	$E_{\text{rel}}^{\nabla p}$	Commentaire
Carré	= 0	= 0	Référence
Voronoi, calcul de $p_h$ aux centres de masse	$\neq 0$	$\neq 0$	Orthogonalité non respectée
Voronoi, calcul de $p_h$ aux centres de Voronoi	$\neq 0$	= 0	Erreur due à l'intégration
Voronoi, calcul de $p_h$ et $p$ aux centres de Voronoi	= 0	= 0	Erreurs nulles, mais perte de consistance pour l'intégration numérique

TAB. 4.1 : Limites liées à la non coïncidence des centres de masse et de Voronoi lors de l'utilisation de l'approximation à deux points.

les écarts  $E_{\text{rel}}^{\cdot}$  sont bien nuls. Avec un maillage de Voronoi, on retrouve pour les écarts en pression les résultats observés précédemment (cf tableau 4.1). En revanche, les écarts en déplacement  $E_{\text{rel}}^{\mathbf{u}}$  et  $E_{\text{rel}}^{\sigma}$  sont non nuls sur ce cas. En effet, à la différence du cas non couplé, la matrice  $\mathcal{B}$  est à présent non nulle. Il a été vu que ses coefficients sont de la forme  $\int_K p \operatorname{div}(\mathbf{u})$  ce qui ajoute à nouveau une erreur d'intégration en raison du décalage des centres. Cette intégrale intervenant dans l'équation de la mécanique, on observe un report de l'erreur d'intégration sur le champ de déplacement. En conséquence, l'usage du schéma VEM-TPFA ne satisfait pas le *patch test* sur les maillages de Voronoi.

Néanmoins, la solution recherchée n'étant pas linéaire dans le cas général, cette apparente incompatibilité entre l'intégration numérique et la condition d'orthogonalité devrait être invisible, dans la mesure où aussi bien l'approximation du flux que celle de l'intégrale sont d'ordre 1 en espace.

### 4.3.3 Validation numérique de l'estimation d'erreur

Afin justement d'illustrer cet ordre d'approximation et de retrouver le résultat de l'estimation d'erreur du théorème 4.4, on considère un cas académique bidimensionnel issu de [Lem13]. Soit le carré unité constitué d'un matériau poreux homogène, de paramètres  $E = 2.5$ ,  $\nu = 0.25$  pour le solide  $\bar{\kappa} = \kappa \bar{l}_2$  avec  $\kappa = 1$  pour le fluide, couplés par les coefficients poroélastiques  $\alpha = 1$  et  $c_0 = 0.5$ . L'intervalle de temps pour les simulations démarre à 0 pour s'achever à  $T = 1$ . On considère la solution analytique

$$\mathbf{u}(\mathbf{x}, t) = 10^{-2} \exp(-t) \begin{pmatrix} x^2 y \\ -x y^2 \end{pmatrix}, \quad p(\mathbf{x}, t) = \exp(-t) \sin\left(\frac{x}{\sqrt{2}}\right) \sin\left(\frac{y}{\sqrt{2}}\right),$$

que l'on utilise pour calculer les chargements

$$\mathbf{f}(\mathbf{x}, t) = \begin{pmatrix} \frac{\alpha \exp(-t)}{\sqrt{2}} \cos(x/\sqrt{2}) \sin(y/\sqrt{2}) - 2 \cdot 10^{-2} \exp(-t) y \\ \frac{\alpha \exp(-t)}{\sqrt{2}} \sin(x/\sqrt{2}) \cos(y/\sqrt{2}) + 2 \cdot 10^{-2} \exp(-t) x \end{pmatrix} \text{ et } q(\mathbf{x}, t) = (\kappa - c_0) \exp(-t) \sin\left(\frac{x}{\sqrt{2}}\right) \sin\left(\frac{y}{\sqrt{2}}\right).$$

Des conditions aux limites de Dirichlet sont imposées sur les frontières pour les deux équations. Le domaine  $\Omega$  est discrétisé par des mailles de Voronoi qui satisfont par définition la condition d'orthogonalité. On génère plusieurs maillages de diamètres maximums respectifs  $h = 1.2 \cdot 10^{-1}$ ,  $5.5 \cdot 10^{-2}$ ,  $1.4 \cdot 10^{-2}$ ,  $4.5 \cdot 10^{-3}$  et  $1.6 \cdot 10^{-3}$  et, simultanément, on considère plusieurs valeurs pour le pas de temps :  $\Delta t = 0.2, 0.1, 0.02, 0.01$ . Pour chaque couple  $(h, \Delta t)$ , l'erreur d'approximation  $E^A$  définie par l'équation (4.35) est calculée. Son évolution est représentée sur la figure 4.2 en fonction du pas du maillage  $h$  pour un pas de temps donné (à gauche) ou en fonction de pas de temps (à droite) pour un maillage donné. Deux comportements asymptotiques peuvent être distingués. D'un côté, si le pas de temps est trop grand, réduire le pas du maillage ne permet

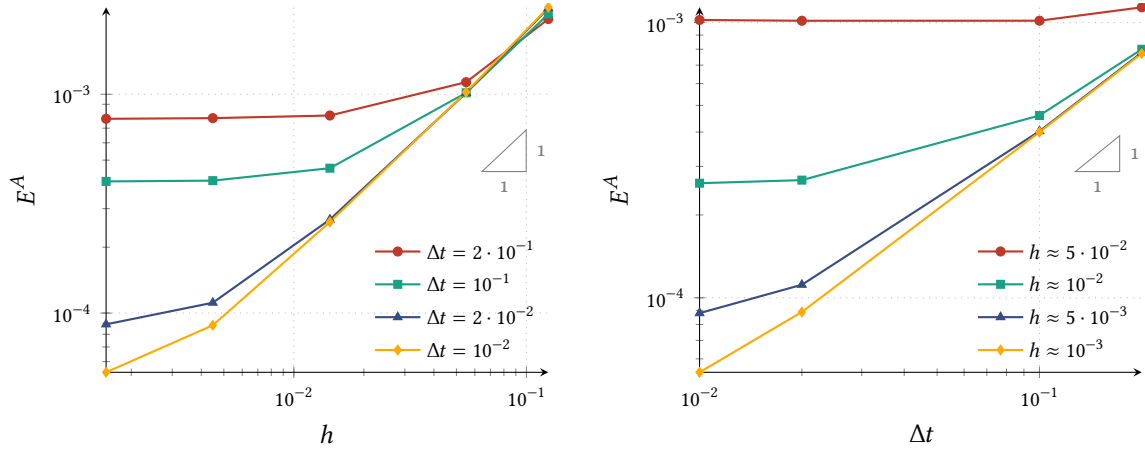


FIG. 4.2 : Évolution de l'erreur d'approximation en fonction du pas d'espace (gauche) ou du pas de temps (droite).

plus d'améliorer l'erreur. Ceci s'observe grâce à la courbe rouge du graphe de gauche ou grâce à la partie droite du graphe de droite. Dans cette situation, le terme en  $h$  est négligeable par rapport au terme en  $\Delta t$  dans l'estimation d'erreur. De l'autre côté, si le pas du maillage est trop grossier, réduire le pas de temps n'entraîne pas une réduction de l'erreur. Ceci s'observe dans la partie droite du graphe de gauche ou à l'aide de la courbe rouge du graphe de droite. Dans cette situation, le terme en  $\Delta t$  est négligeable devant le terme en  $h$  dans l'estimation d'erreur. Sur les deux graphes, on retrouve l'ordre de convergence attendu en regardant les courbes jaunes, où  $h$  (resp.  $\Delta t$ ) est suffisamment petit pour ne pas perturber l'évolution de l'erreur en fonction de  $\Delta t$  (resp.  $h$ ). La pente des droites jaunes est égale à un, comme le prévoit la formule d'estimation d'erreur. Entre les courbes limites rouges et jaunes, l'erreur d'approximation peut être diminuée en réduisant au choix la taille du maillage ou le pas de temps. Cette double dépendance de l'erreur s'observe également sur la figure 4.3 représentant les lignes de niveaux de l'erreur d'approximation dans le plan  $(h, \Delta t)$ . Encore une fois, on peut observer les régions où l'erreur est due en majeure partie à  $\Delta t$  (lignes horizontales) et celles où elle est due en majeure partie à  $h$  (lignes verticales).

## 4.4 Calculs réalistes de problèmes physiques

Afin d'illustrer la méthode dans un cadre plus réaliste, on propose dans cette section la résolution numérique des deux problèmes physiques détaillés dans la section 2.2.3.

### 4.4.1 Cas de compaction

On s'intéresse dans cette partie au cas de compaction schématisé sur la figure 2.1 dont la solution exacte est connue. Les paramètres mécaniques sont ceux exposés sur le schéma, et on adopte un pas de temps de 10000 ans pour une durée totale de simulation de 5 millions d'années (soit 500 pas de temps). Le taux de sédimentation  $\sigma_a$  est fixé à  $6.4 \cdot 10^{-8} \text{ Pa}\cdot\text{s}^{-1}$  ce qui équivaut à un dépôt d'environ 100 mètres de sédiments par million d'années. Le domaine est maillé dans un premier temps par des carrés, ce qui produit un maillage de 6000 cellules, puis par des éléments de Voronoï, ce qui produit un maillage de 5000 cellules représenté sur la figure 4.4 et utilisé pour les illustrations dans la suite.



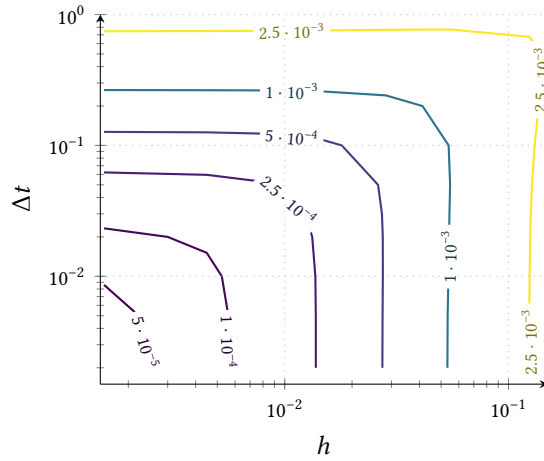


FIG. 4.3 : **Erreur d'approximation dans le plan** ( $h, \Delta t$ ). Les deux paramètres de discrétisation peuvent être diminués pour améliorer l'erreur, jusqu'au point où l'un devient négligeable par rapport à l'autre.

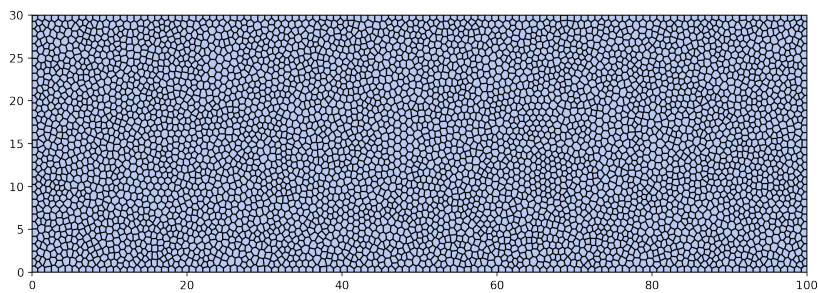


FIG. 4.4 : **Maillage de Voronoï** utilisé pour illustrer le schéma VEM-TPFA sur le cas de compaction. Ce maillage satisfait la condition d'orthogonalité.

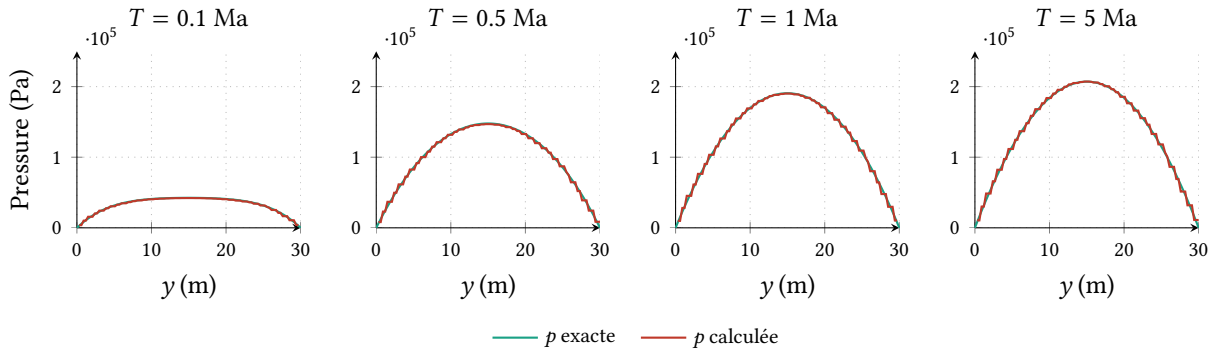


FIG. 4.5 : **Pression calculée et exacte** à différents instants sur le cas de compaction. En dehors d'un effet d'escalier dû à la représentation constante par maille de la solution, les deux pressions sont proches.

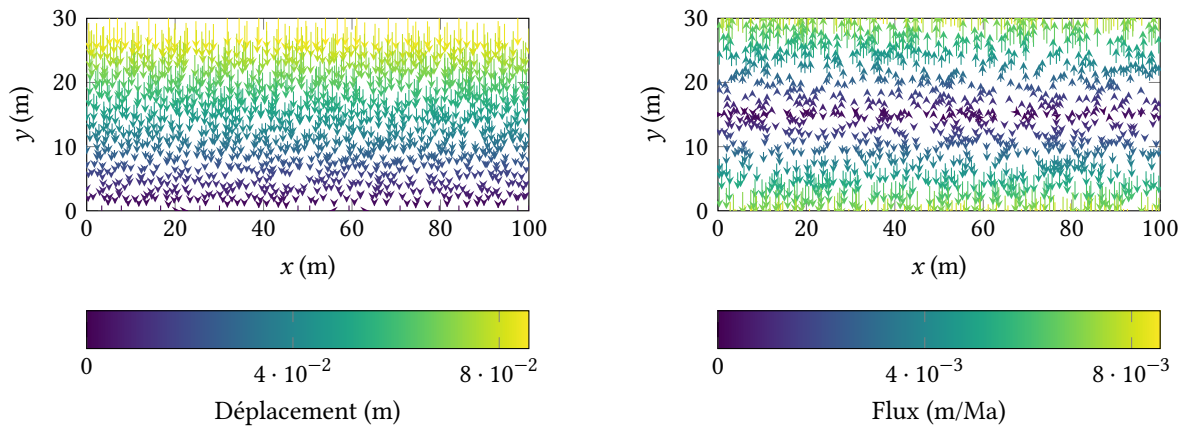


FIG. 4.6 : **Champ de déplacement (gauche) et flux (droite)** au temps final sur le cas de compaction. Les vecteurs sont colorés selon leur norme, et leur longueur est proportionnelle à leur norme.

Puisque la pression analytique donnée par l'équation (2.18) ne dépend pas de la composante spatiale  $x$ , elle peut être représentée le long de l'axe vertical médian  $x = 50$  m. Pour différents instants, la figure 4.5 représente sur cet axe et en fonction de l'altitude  $y$  la pression analytique et la pression calculée. On observe que pour un temps suffisamment grand, la pression a une forme parabolique. Cette forme correspond à l'expression analytique dans laquelle le terme temporel est devenu négligeable en raison de l'exponentielle décroissante. Pour chaque pas de temps représenté, la pression calculée est fidèle à la pression analytique. Seul apparaît un effet d'escalier dû à la faible résolution du maillage, cet effet s'atténuant pour un maillage plus fin. D'autre part, la figure 4.6 représente, au temps final et sur l'ensemble du domaine, le champ de déplacement  $\mathbf{u}_h$  (à gauche) et le flux du fluide (à droite). Pour une meilleure visibilité, un nombre restreint de vecteurs sont affichés et leur taille est multipliée de manière homogène par un certain coefficient. Le champ de déplacement montre que comme attendu, le milieu ne se déplace que dans la direction  $\mathbf{e}_y$ , c'est à dire dans la même direction que la contrainte qu'il subit. Le flux montre que sous l'effet de ce déplacement, le fluide s'échappe du domaine par les deux seuls plans non imperméables  $y = 0$  et  $y = 30$  m. Le flux présente de plus les mêmes symétries que le champ de pression. Pour approfondir la comparaison entre les grandeurs calculées et les grandeurs exactes, des maillages plus raffinés sont créés. Cependant, pour

diminuer le temps de calcul, la longueur  $L$  est ramenée à 30 m au lieu de 100 m, ce qui ne modifie pas les résultats puisque la pression ne dépend que de  $y$ . Pour chaque famille de maillages et pour chaque niveau de raffinement, le tableau 4.2 indique le nombre de degrés de liberté du système couplé.

$h_{max}$	$8.4 E^{-1}$	$4.2 E^{-1}$	$2.1 E^{-1}$	$1.1 E^{-1}$	$7.1 E^{-2}$	$h_{max}$	1.0	$5.0 E^{-1}$	$2.6 E^{-1}$	$1.3 E^{-1}$	$8.9 E^{-2}$
#dofs	7.7K	30K	121K	482K	1.1M	#dofs	13K	49K	196K	798K	1.8M
(a) Maillage carré						(b) Maillage Voronoï					

TAB. 4.2 : Diamètre maximal et nombre de degrés de libertés pour les deux familles de maillages

Pour ces deux familles de maillages, la figure 4.7 représente l'évolution des erreurs  $E_{rel}^u$ ,  $E_{rel}^p$ ,  $E_{rel}^\sigma$  et  $E_{rel}^{\tau_v}$  en fonction du diamètre du maillage au temps final de la simulation. Ceci permet d'illustrer l'ordre de convergence du schéma sur un cas plus complexe que le problème académique du paragraphe 4.3.3. En norme  $L^2$ , on observe un ordre de convergence égal à deux pour le déplacement et pour la pression sur les deux familles de maillages. Même si l'analyse d'erreur effectuée plus tôt dans ce chapitre pour le problème couplé ne donne un résultat que sur une norme équivalente à la norme  $H^1$ , cette observation rejoint les ordres habituellement relevés pour les problèmes découplés : la méthode des éléments virtuels convergeait à l'ordre 2 dans le chapitre 3, tandis que les volumes finis deux points sont connus pour converger à l'ordre 2 en norme  $L^2$ , même si ceci n'a pas été démontré de manière théorique [DN17]. Pour obtenir ces courbes, la pression analytique a été calculée au centres de Voronoï. Si la pression analytique est calculée aux centres de masses, l'erreur  $E_{rel}^p$  est plus élevée, ce qui rejoint les observations du paragraphe 4.3.2, et l'ordre de convergence n'est plus que de 1 pour l'erreur  $L^2$  en pression (voir la courbe bleue sur la figure). En norme  $H^1$ , l'erreur en contraintes converge à l'ordre 1 sur les deux familles de maillage. Pour la variable de pression, on observe un phénomène de superconvergence avec les maillages carrés puisque la pente de la droite vaut 1,5. Cette pente est bien égale à un sur les maillages de Voronoï.

#### 4.4.2 Cas Mandel

Le cas de consolidation de Mandel, décrit dans la section 2.2.3, permet également d'illustrer le comportement du code poroélastique. En particulier, cet exemple permet d'observer l'influence du choix des conditions initiales du problème. Les caractéristiques géométriques et mécaniques utilisées pour ce problème sont décrites sur la figure 2.2. L'intensité de la force  $F$  est choisie de manière à avoir une pression de l'ordre du bar :  $F = 100$  MPa. Les simulations sont cette fois effectuées sur un maillage constitué de  $100 \times 30$  carrés.

Dans un premier temps, le problème est initialisé selon la procédure suivie dans les cas précédents : la pression  $p^0(\mathbf{x})$  est calculée à partir de la solution exacte (2.20) à l'instant  $t = 0^+$  (celle-ci est donc non nulle), puis les déplacements initiaux  $\mathbf{u}^0(\mathbf{x})$  sont calculés en résolvant le problème d'équilibre mécanique. La simulation court jusqu'au temps final  $T = 10^{14}$  s  $\approx 3$  Ma avec un pas de temps  $\Delta t = 10^{12}$  s  $\approx 30$  000 ans. L'évolution de la pression et du déplacement  $u_x$  (qui ne dépendent que de  $x$ ) le long de l'axe médian  $x$  ainsi que l'évolution du déplacement  $u_y$  (qui ne dépend que de  $y$ ) le long de l'axe  $y$  médian sont représentés sur la figure 4.8 à différents instants. Avec cette approche, la solution approchée représentée par les symboles est très proche de la solution exacte représentée par les traits pleins. De plus, l'effet Mandel-Creyer est bien capturé par la pression approchée : ceci peut être observé sur la figure 4.9 qui représente l'évolution de la pression au point  $(x = 0, y = 0)$  en fonction du temps. À l'extrémité gauche du domaine de simulation (et donc sur la verticale médiane du domaine), la pression augmente légèrement en début de simulation avant

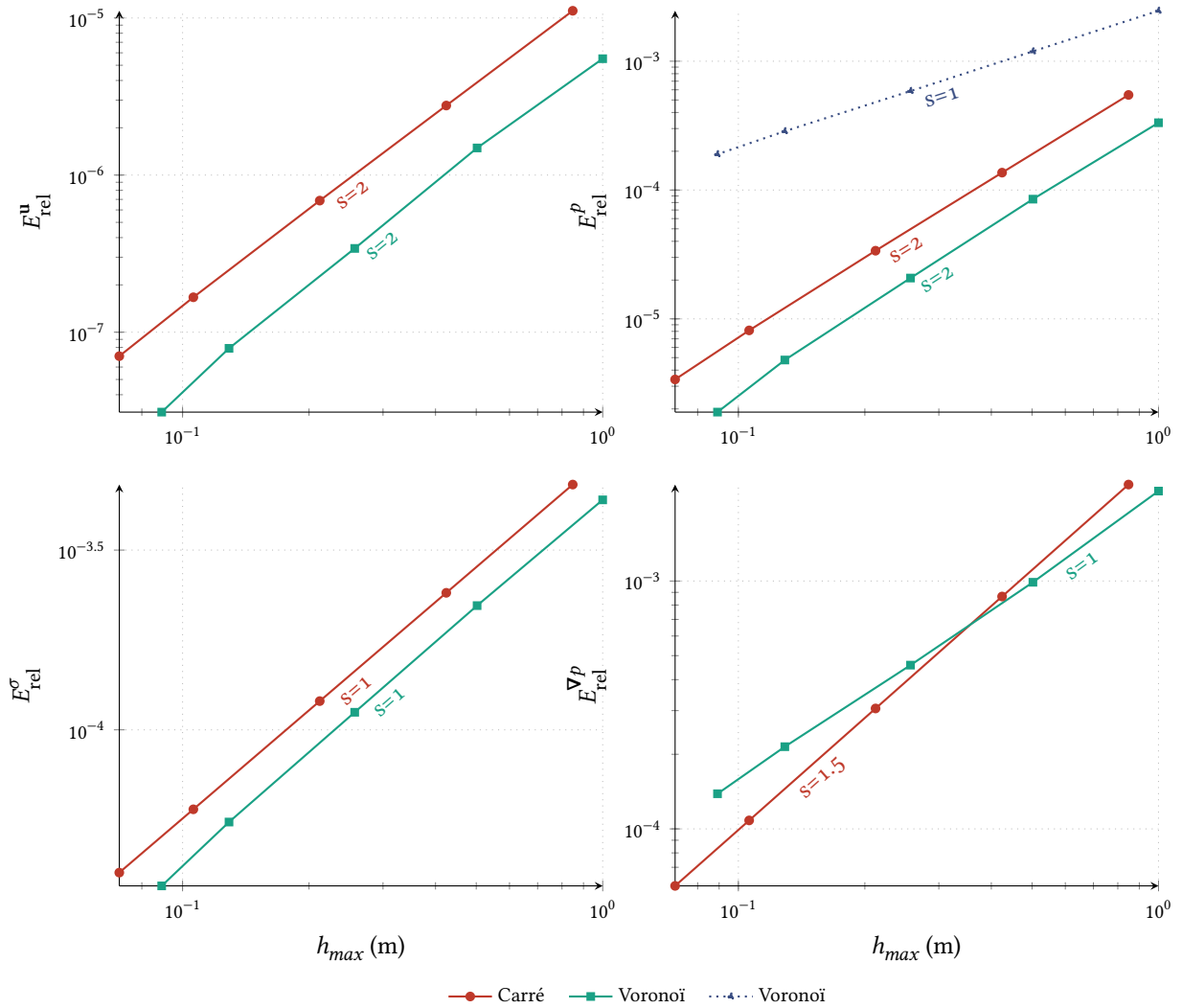


FIG. 4.7 : **Convergence en normes  $L^2$  et  $H^1$  des erreurs en déplacement et en pression.** Les ordres de convergence correspondent à ceux que l'on obtient classiquement sur les problèmes découplés, sauf celui de l'erreur  $E_{rel}^{\nabla p}$  sur les maillages carrés qui est plus élevé qu'attendu. De plus, on observe (droite bleue) que l'ordre de convergence se dégrade pour  $E_{rel}^p$  sur le maillage de Voronoï si la pression analytique n'est pas évaluée en ces mêmes points.

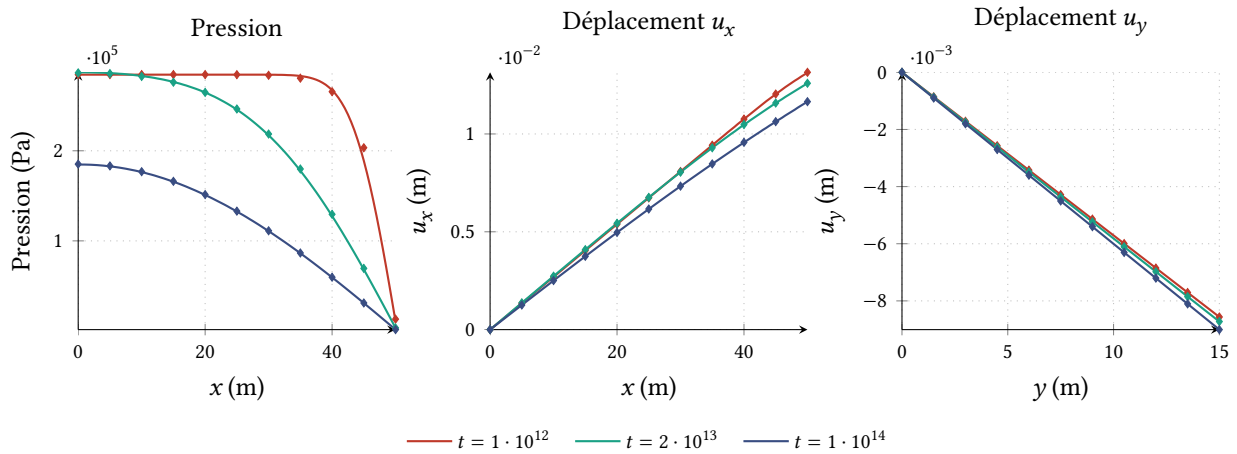


FIG. 4.8 : Comparaison de la solution calculée et analytique du cas Mandel pour la pression  $p$  et les déplacement  $u_x$  et  $u_y$  à différents instants. La solution exacte, en traits pleins, et la solution calculée, représentée par les symboles, sont cohérentes à chaque instant de la simulation.

de décroître vers 0.

Dans un second temps, en plus de la méthode utilisée ci-dessus pour initialiser les solutions, deux autres approches sont testées de manière à se rapprocher de la description physique du problème, qui indique que la pression et les déplacements sont nuls à l'instant initial. La deuxième approche consiste ainsi à fixer la pression initiale à 0 et à laisser la méthode calculer les déplacements au temps  $t = 0$ . Il s'agit en quelque sorte d'un compromis entre le fonctionnement de la méthode numérique et les conditions du problème. Enfin, dans une troisième approche, la pression initiale et les déplacements initiaux sont pris nuls. Notons que cette troisième approche ne correspond pas au comportement "naturel" de l'algorithme et que celui-ci est donc modifié pour pouvoir imposer un déplacement initial. Pour ces trois approches, la figure 4.10 représente à nouveau les valeurs de  $p(x)$ ,  $u_x(x)$  et  $u_y(y)$  au temps  $t = 5 \cdot 10^{12}$  s, c'est à dire en début de simulation. La courbe noire, qui fait office de référence, est calculée à partir des expressions analytiques. Comme observé précédemment, la première approche donne des résultats numériques proches des solutions analytiques, mais cette approche ne correspond pas aux paramètres du problème. Avec la deuxième approche, les quantités calculées pour  $p$  et  $u_x$  diffèrent des solutions exactes. Ce comportement est probablement dû à la condition imposée au déplacement sur le bord  $y = h$ . En effet, cette condition est non nulle quand  $t = 0$ , et perturbe donc le problème de mécanique qui calcule une solution non nulle sous cette condition. En d'autres termes, la condition au bord à l'instant initial n'est valable qu'à partir de l'instant  $t = 0^+$  et est donc elle aussi incompatible avec la description physique du problème. Cette incompatibilité est levée par la troisième approche puisque celle-ci n'utilise plus la condition à l'instant initial pour calculer les déplacements. Avec cette méthode, on retrouve un résultat coïncidant avec la pression et les déplacements exacts. Cette configuration permet de valider le code de calcul, qui à partir de la description physique du problème ( $p^0 = 0$ ,  $\mathbf{u}^0 = \mathbf{0}$ ) parvient à retrouver la solution analytique qui est pourtant asymptotiquement différente de 0.

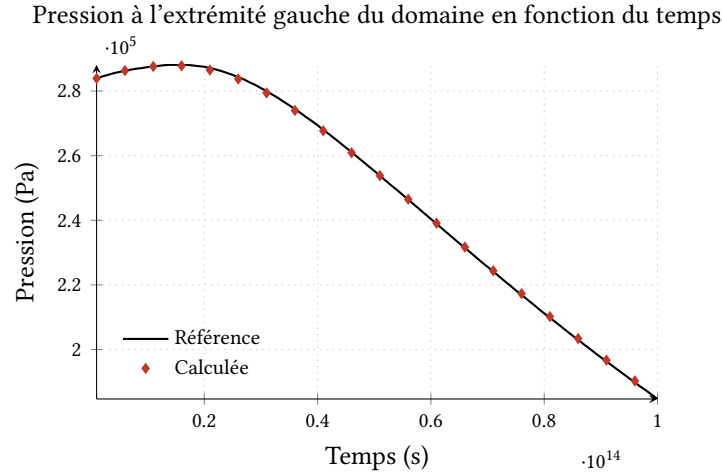


FIG. 4.9 : **Effet Mandel-Creyer** : en début de simulation, la pression à l'extrémité gauche du domaine augmente avant de décroître. L'effet est bien capturé par la pression calculée (symboles), qui reste très proche de la pression exacte (trait plein).

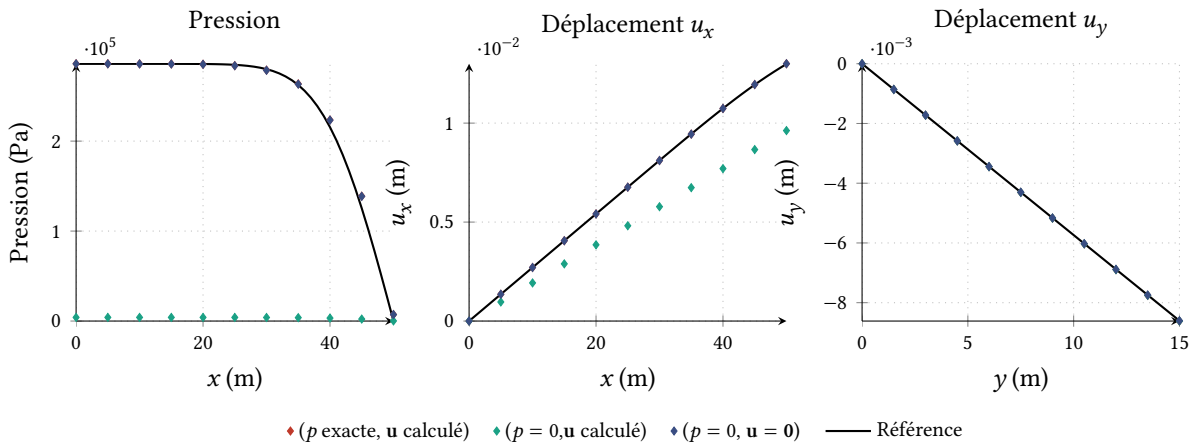


FIG. 4.10 : **Comparaison de différentes méthodes d'initialisation pour le cas Mandel**. Les trois approches (symboles) sont comparées à la pression et aux déplacements exacts (trait plein) au premier pas de temps  $t = 5 \cdot 10^{12}$  s. La troisième approche, qui correspond à la description physique du problème, donne des résultats proches de la solution exacte malgré la discontinuité entre cette condition initiale et la valeur asymptotique des solutions.



# 5

## Vers des maillages plus généraux avec une discrétisation VEM-MPFA

Dans le chapitre précédent, un premier schéma couplé a été construit pour approcher le système de poroélasticité. L'utilisation de l'approximation à deux points pour le flux numérique constituait un bon choix pour construire, implémenter et étudier mathématiquement le schéma. Celui-ci a été illustré sur des cas analytiques et réalistes et a ouvert de nouvelles possibilités, par exemple celle de gérer des maillages de type Voronoï. L'objectif suivant, qui fait l'objet de ce chapitre, est de traiter des maillages plus généraux, par exemple constitués de polytopes quelconques ou bien issus de domaines souterrains réels. La première section de ce chapitre revient justement sur les limites du schéma précédent et décrit les caractéristiques spécifiques des maillages utilisés en géosciences. Dans la seconde section, un schéma volumes finis multipoints est introduit pour la discrétisation de l'équation de conservation de la masse. Ce nouveau schéma couplé est adapté à l'utilisation de maillages polygonaux ou polyédriques. Cependant, pour pouvoir traiter les maillages utilisés en géosciences, la méthode des éléments virtuels doit être adaptée : les modifications qui y sont apportées sont décrites dans la troisième section. Finalement, une illustration réaliste inspirée d'un problème de géomécanique est présentée dans la quatrième section du chapitre.

### Sommaire

---

<b>5.1</b>	<b>Motivations</b>	<b>88</b>
5.1.1	$\bar{\kappa}$ -orthogonalité	88
5.1.2	Maillages géosciences	90
<b>5.2</b>	<b>Approximation multipoints pour le schéma volumes finis</b>	<b>91</b>
5.2.1	Approximation MPFA pour les flux numériques	91
5.2.2	Implémentation et couplage avec les éléments virtuels	93
5.2.3	Illustrations numériques	94
<b>5.3</b>	<b>Modification des VEMs pour les faces non planes</b>	<b>96</b>
5.3.1	Transformation en maillage polyédrique	96
5.3.2	Un espace d'approximation modifié pour les éléments virtuels	99
5.3.3	Un mot sur les volumes finis	101



<b>5.4</b>	<b>Illustration sur une application géosciences . . . . .</b>	<b>102</b>
5.4.1	Paramètres du cas . . . . .	102
5.4.2	Résultats numériques . . . . .	103

---

## 5.1 Motivations

Deux limites sur la compatibilité des maillages peuvent être opposées au schéma VEM-TPFA. La première, imputable aux volumes finis, est la condition d’orthogonalité qui doit être respectée par les maillages. La seconde, plutôt spécifique au contexte applicatif, provient des caractéristiques spécifiques des maillages utilisés en géosciences. Cette section présente ces deux incompatibilités qui poussent à modifier le schéma numérique.

### 5.1.1 $\bar{\kappa}$ -orthogonalité

L’approximation à deux points du flux numérique dans la méthode des volumes finis, décrite dans la section 4.1.2, n’est consistante que si la condition dite de  $\bar{\kappa}$ -orthogonalité est respectée : pour toute face  $f$  séparant deux mailles  $K$  et  $L$  dans lesquelles la pression est approchée aux points  $\mathbf{x}_K$  et  $\mathbf{x}_L$ , la propriété

$$\bar{\kappa}^{-1}(\mathbf{x}_K \mathbf{x}_L) \perp f \quad (5.1)$$

doit être satisfaite pour que l’approximation soit consistante.

Même dans le cas le plus simple où  $\bar{\kappa}$  est proportionnel à l’identité, seuls certains types de maillages vérifient cette condition. Sur les quatre familles de maillages discrétisant le carré unité présentées précédemment sur la figure 3.4, seuls les maillages carrés la satisfont. L’orthogonalité est proche d’être vérifiée sur les maillages de Lloyd, qui sont des maillages de Voronoï régularisés, ces derniers étant orthogonaux par construction. Pour illustrer les conséquences de l’absence de cette propriété, le cas test défini dans la section 4.3.3 est à nouveau considéré. Pour des coefficients physiques prenant respectivement les valeurs  $E = 2.5$ ,  $\nu = 0.25$ ,  $\bar{\kappa} = \bar{I}_2$ ,  $\alpha = 1$  et  $c_0 = 0.5$ , le problème dont la solution analytique est donné par

$$\mathbf{u}(\mathbf{x}, t) = 10^{-2} \exp(-t) \begin{pmatrix} x^2 y \\ -x y^2 \end{pmatrix}, \quad p(\mathbf{x}, t) = \exp(-t) \sin\left(\frac{x}{\sqrt{2}}\right) \sin\left(\frac{y}{\sqrt{2}}\right),$$

est résolu pour  $t \in [0, 1]$  en utilisant ces différentes familles de maillages. L’évolution de l’erreur d’approximation  $E^A$  définie par l’équation (4.35) est représentée sur la figure 5.1. À gauche, le pas de temps  $\Delta t$  est suffisamment faible pour que l’erreur soit due au pas d’espace  $h$ . À droite, le pas d’espace est suffisamment faible pour que l’erreur soit due au pas de temps. Dans les deux cas, on observe la non convergence du schéma VEM-TPFA et une erreur d’approximation élevée pour les maillages triangles et twisted, qui ne respectent pas la condition d’orthogonalité.

Dans le cas plus complexe où  $\bar{\kappa}$  est anisotrope, construire des maillages satisfaisant la condition de  $\bar{\kappa}$ -orthogonalité (5.1) est un sujet de recherche en soi [Rou16]. Inversement, lorsque le maillage est créé en amont et donné tel quel, il peut être très délicat de choisir des points de calcul de manière à ce que la  $\bar{\kappa}$ -orthogonalité soit respectée. Ceci est illustré sur la figure 5.2 où différents tenseurs  $\bar{\kappa}$  sont considérés. Dans les trois cas, on cherche à positionner les centres de calcul des trois cellules de manière à ce que la  $\bar{\kappa}$ -orthogonalité soit satisfaite pour les trois arêtes  $f$ ,  $f'$  et  $f''$ . Ce cas est simplifié puisque sur un maillage de ce type, la condition devrait normalement être respectée par les six arêtes de chaque hexagone. Malgré cela, lorsque la matrice de mobilité n’est pas multiple de l’identité, le placement des points n’est pas trivial

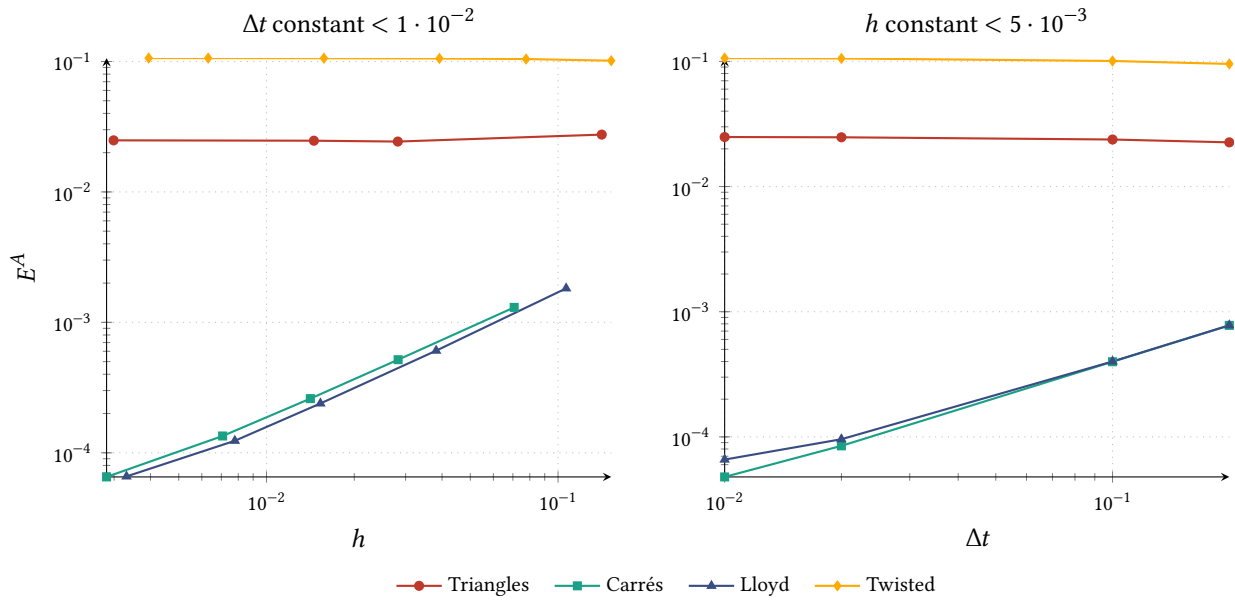


FIG. 5.1 : Évolution de l'erreur d'approximation pour 4 maillages en fonction du pas d'espace (gauche) ou du pas de temps (droite). Seuls les maillages satisfaisant la condition d'orthogonalité convergent.

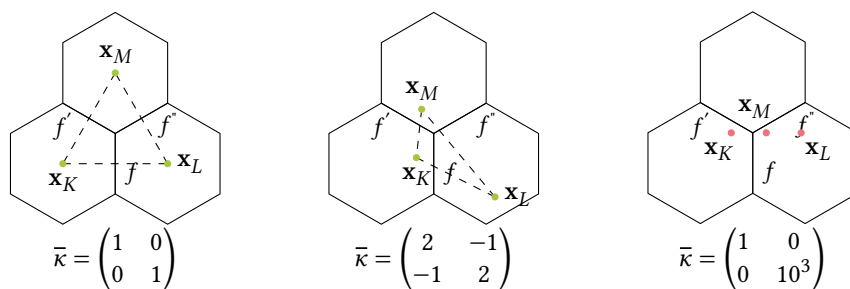


FIG. 5.2 : Positions possibles pour les centres de calculs des cellules  $K$ ,  $L$  et  $M$  de manière à ce que la condition d'orthogonalité soit respectée pour différents tenseurs  $\bar{\kappa}$ . Même sur ce cas simplifié où seulement trois conditions sont imposées (pour six coordonnées inconnues), il est parfois difficile de maintenir les centres de calcul à l'intérieur des cellules.

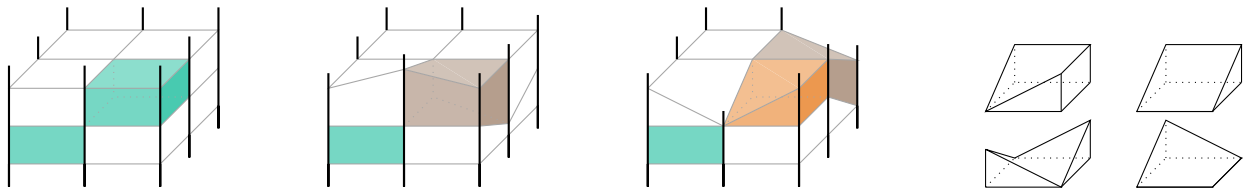


FIG. 5.3 : **Méthode des piliers utilisée pour mailler les domaines en géosciences.** En fonction de l'altitude des points, le maillage peut contenir des cellules **hexaédriques** classiques, des mailles à faces **non planes** ou bien des hexaèdres **dégénérés**. Certaines de ces mailles dégénérées sont représentées à droite.

(figure du milieu). Dans certains cas, il est même impossible de satisfaire la condition en maintenant les centres de calcul à l'intérieur de leurs cellules. C'est par exemple le cas sur la figure de droite. Il va sans dire que ce problème est encore plus complexe en trois dimensions ou lorsque les coefficients de  $\bar{\kappa}$  dépendent de l'espace. Cette contrainte d'orthogonalité constitue donc une première restriction à l'usage du schéma VEM-TPFA.

### 5.1.2 Maillages géosciences

Une seconde restriction apparaît lorsque l'on souhaite utiliser les types de maillage standards dans le domaine des géosciences. En modélisation de bassin, mailler un domaine géologique est un processus complexe : à partir de la configuration actuelle du domaine, le temps est remonté pour reconstruire la configuration initiale du domaine. Cette restauration prend en compte les phénomènes physiques qui modifient le milieu tels que l'érosion, le dépôt de sédiments, la subsidence ou le compactage vertical. Pour décrire ces phénomènes, majoritairement verticaux, la grille tridimensionnelle est bâtie à partir de droites verticales appelées piliers. L'ensemble de ces piliers forme une structure bidimensionnelle cartésienne, chacun d'entre eux étant repéré par ses coordonnées  $(i, j)$  dans le plan horizontal. Sur chaque pilier se trouve un même nombre de points repérés par leur profondeur qui a été calculée lors de la reconstruction géométrique. Finalement, chaque région délimitée par les 4 piliers  $(i, j)$ ,  $(i + 1, j)$ ,  $(i + 1, j + 1)$  et  $(i, j + 1)$  et par les  $k$ -ième et  $(k + 1)$ -ième points définit une maille du domaine, qui s'apparente donc à un hexaèdre à 8 sommets [Pon89]. D'ailleurs, si les profondeurs  $z_{ij}^k$  étaient équi-réparties le long du pilier  $(i, j)$  et identiques d'un pilier à l'autre, le maillage serait en réalité constitué de parallélépipèdes. La difficulté provient du fait que ces profondeurs varient d'un pilier à l'autre, puisque ce sont celles-ci qui permettent une représentation fine des strates géologiques, des érosions ou des failles. En particulier, rien n'assure que les 4 points  $z_{i',j'}^k$  (avec  $i' \in \{i, i + 1\}$  et  $j' \in \{j, j + 1\}$ ) se trouvent dans le même plan horizontal. Dans ce cas, la maille n'étant définie que par ses huit sommets, elle décrit un "hexaèdre" à face(s) non plane(s). Pire encore, sur un pilier donné les points  $z_{ij}^k$  et  $z_{ij}^{k+1}$  peuvent être confondus, ce qui conduit à des hexaèdres dit dégénérés. Ces situations sont schématisées sur la figure 5.3. À droite de celle-ci, on représente les différentes mailles dégénérées pouvant être présentes dans le maillage selon si une, deux ou trois paires de sommets sont confondues. Pour compléter cette illustration, un maillage réel construit par cette méthode est représenté sur la figure 5.4. Sur 72500 cellules, plus de 3000 sont des hexaèdres dégénérés, extraits du maillage sur la figure 5.4b et colorées en fonction de leur catégorie. Le maillage comporte également un grand nombre de faces non planes : plus de 90% des faces "horizontales", soit quasiment un tiers des faces du maillage. Bien que la méthode des éléments virtuels soit définie pour des polyèdres quelconques, les maillages issus des géosciences sortent de son cadre d'application puisque les mailles ne sont même plus des polyèdres. Ce cas d'application spécifique constitue donc une seconde restriction à l'usage du schéma défini dans le chapitre

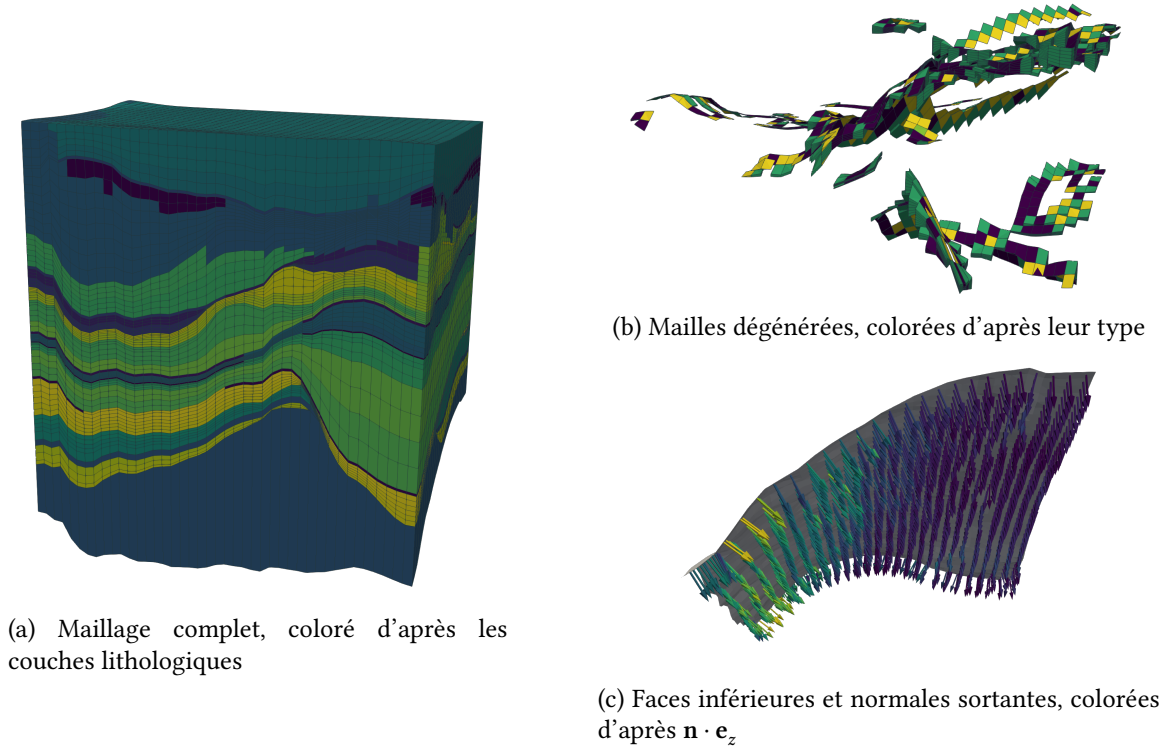


FIG. 5.4 : **Maillage issu des géosciences.** Le respect des propriétés géométriques (a) engendre des mailles dégénérées, extraites du domaine sur la figure (b). De plus, les bords du domaine ne sont pas tous plans (c).

précédent.

## 5.2 Approximation multipoints pour le schéma volumes finis

Dans cette section, on cherche à résoudre la première limite soulevée dans le paragraphe 5.1.1 en utilisant un schéma volumes finis insensible à la condition de  $\bar{\kappa}$ -orthogonalité. En restant dans le formalisme générique de la discrétisation VEM-FV fourni par l'équation (4.7), on étudie une autre manière d'assembler les flux numériques  $\mathcal{F}_{Kf}^n$ . Même s'il serait à terme intéressant d'utiliser des schémas récents et prometteurs (schéma hybride, TPFA non linéaire), la classe étudiée ici est la plus standard après le schéma TPFA dans les codes d'écoulement en milieux poreux. En particulier, les changements par rapport à la méthode précédente restent limités et une implémentation de ces méthodes est déjà disponible dans la plateforme Arcane.

### 5.2.1 Approximation MPFA pour les flux numériques

Plus précisément, l'approximation considérée ici appartient à la famille des approximations multipoints, dénotées MPFA (multipoint flux approximation en anglais), qui sont également des méthodes *cell centered* utilisant des inconnues localisées dans les mailles. Ainsi, par rapport au schéma VEM-TPFA présenté dans le chapitre précédent, seule la définition du flux numérique est à revoir. De manière générale pour les méthodes d'approximation multipoints, ce flux est scindé en deux demi-flux qui sont calculés en

utilisant une interpolation linéaire de l'inconnue autour du sommet associé. En pratique, des inconnues supplémentaires sont introduites sur les faces du maillage pour approcher le gradient puis éliminées par continuité de la solution et conservativité des flux. Parmi les différentes variantes des méthodes multipoints, qui se distinguent par le choix de ces équations de continuité et de conservativité [Dro14], celle présentée ici est la version MPFA-O détaillée par exemple dans [Aavo2]. Pour simplifier les développements, on se limite à un maillage bidimensionnel en notant que les mêmes idées s'appliquent en trois dimensions. Certaines des quantités nécessaires et leurs notations sont schématisées sur la figure 5.5.

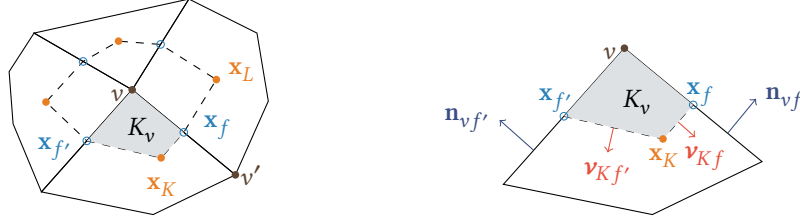


FIG. 5.5 : **Grandeurs et notations** nécessaires à la construction d'une approximation MPFA-O. Représentée ici en deux dimensions, la construction est similaire en trois dimensions.

Entre les mailles  $K$  et  $L$  séparées par la face  $f$ , le flux est divisé en deux demi-flux selon

$$\mathcal{F}_{Kf} = \mathcal{F}_{Kfv} + \mathcal{F}_{Kfv'} \quad (5.2)$$

où  $\mathbf{x}^f$  est le milieu de l'arête  $f$  qui relie deux sommets  $v$  et  $v'$ . Pour calculer le demi flux  $\mathcal{F}_{Kfv}$ , une *zone d'interaction* est construite autour du sommet  $v$  en reliant les centres des mailles avec les milieux des arêtes connectés à  $v$  (en pointillés sur la figure). C'est à ce stade qu'une hypothèse de continuité est effectuée pour supposer que l'approximation de la pression, linéaire dans chaque sous-cellule, est continue au milieu des arêtes et donc définir les inconnues intermédiaires  $p_{h_f}$ . Grâce à ces inconnues supplémentaires, le gradient constant dans chaque sous-cellule se calcule selon

$$\nabla_{K_v} p_h = -\frac{1}{2A_{K_v}} \left( \mathbf{v}_{Kf'}(p_{h_f} - p_{h_K}) + \mathbf{v}_{Kf}(p_{h_{f'}} - p_{h_K}) \right) \quad (5.3)$$

où  $A_{K_v}$  désigne l'aire du triangle  $\mathbf{x}_K \mathbf{x}_f \mathbf{x}_{f'}$ , et  $\mathbf{v}_{Kf}$  (resp.  $\mathbf{v}_{Kf'}$ ) le vecteur normal à l'arête  $(\mathbf{x}_K \mathbf{x}_f)$  (resp.  $(\mathbf{x}_K \mathbf{x}_{f'})$ ) orienté vers l'extérieur de ce triangle. Ces normales ne sont pas unitaires mais de norme  $d(\mathbf{x}_K, \mathbf{x}_f)$  et  $d(\mathbf{x}_K, \mathbf{x}_{f'})$ . Pour cette sous-cellule, les demi-flux sortant des arêtes connectées au sommet  $v$  s'expriment donc selon

$$\begin{bmatrix} \mathcal{F}_{Kfv} \\ \mathcal{F}_{Kf'v} \end{bmatrix} = \bar{B}_{K_v} \begin{bmatrix} p_{h_f} - p_{h_K} \\ p_{h_{f'}} - p_{h_K} \end{bmatrix} \quad \text{avec} \quad \bar{B}_{K_v} = -\frac{1}{2A_{K_v}} \begin{bmatrix} |\mathbf{x}_v \mathbf{x}_f| \mathbf{n}_{vf}^T \bar{\kappa}_K \mathbf{v}_{Kf} & |\mathbf{x}_v \mathbf{x}_f| \mathbf{n}_{vf}^T \bar{\kappa}_K \mathbf{v}_{Kf} \\ |\mathbf{x}_v \mathbf{x}_{f'}| \mathbf{n}_{vf'}^T \bar{\kappa}_K \mathbf{v}_{Kf'} & |\mathbf{x}_v \mathbf{x}_{f'}| \mathbf{n}_{vf'}^T \bar{\kappa}_K \mathbf{v}_{Kf'} \end{bmatrix}, \quad (5.4)$$

les notations  $\bar{\kappa}_K$  et  $|\mathbf{x}_v \mathbf{x}_f|$  désignant respectivement la matrice  $\bar{\kappa}$  évaluée au point  $\mathbf{x}_K$  et la longueur des demi-arêtes  $(\mathbf{x}_v \mathbf{x}_f)$ . La même méthode est appliquée à chaque sous-cellule de la zone d'interaction autour de  $v$ , puis la conservativité des flux  $\mathcal{F}_{Kfv} + \mathcal{F}_{Lfv} = 0$  est imposée pour chaque demi-arête interne  $f$  séparant les mailles  $K$  et  $L$ . On aboutit ainsi à un système linéaire carré de taille  $\text{card}(\tau_h(v))$  où  $\tau_h(v)$  est l'ensemble des mailles partageant le sommet  $v$ . La résolution de ce système permet d'exprimer les inconnues de face en fonction des inconnues de centre et donc de réécrire les demi-flux dans (5.4) en ne faisant intervenir que les inconnues de centre. L'ensemble de la procédure étant répété autour du sommet  $v'$ , on peut au final

sommer suivant (5.2) les deux demi-flux  $\mathcal{F}_{Kfv}$  et  $\mathcal{F}_{Kfv'}$  pour exprimer le flux numérique  $\mathcal{F}_{Kf}$  à travers la face  $f$ . Pour les mailles adjacentes à un bord du domaine, une telle élimination des inconnues de face est plus difficile : en particulier, lorsque des conditions aux limites de Neumann sont imposées, l'inconnue de face dépend de cette condition et n'est donc pas connue au moment de la construction des flux. Plutôt que de réévaluer les flux à chaque pas de temps (pour tenir compte de cette condition), une stratégie courante consiste à réemployer l'approximation à deux points pour calculer le flux à travers les faces de bord.

Avec cette méthode, l'inconnue  $p_{h_K}$  de la maille  $K$  est liée aux inconnues de toutes les mailles qui partagent un sommet avec une des arêtes de la maille  $K$ . Le stencil de la matrice  $\mathcal{F}$  résultante est donc plus large que celui obtenu avec l'approximation à deux points, et la résolution est en conséquences plus coûteuse. Lorsque les mailles sont très déformées, ce schéma rencontre des problèmes de stabilité : les matrices intermédiaires deviennent non inversibles, le principe du maximum n'est plus respecté ou la coercivité n'est plus assurée. En dehors de ces deux inconvénients, ce schéma a l'avantage de ne pas dépendre de la condition d'orthogonalité et d'être localement conservatif. Cette construction s'étend au cas tridimensionnel sur des maillages polyédriques quelconques, voir par exemple [AGM10].

### 5.2.2 Implémentation et couplage avec les éléments virtuels

Comme indiqué précédemment, le fait que les inconnues soient toujours (après élimination des inconnues temporaires de faces) *cell centered* rend l'implémentation de ce schéma VEM-MPFA très proche de celle du schéma VEM-TPFA détaillée dans la section 4.3. Pour réutiliser l'algorithme 2, seuls certains éléments internes doivent être adaptés.

**Assemblage matrice et second membre volumes finis** Comme dans le cas précédent, cet assemblage est réalisé par une boucle sur les faces du maillage durant laquelle sont calculées l'équivalent des transmissivités du schéma à deux points. En effet, en appliquant la méthodologie décrite dans le paragraphe précédent, le flux numérique s'écrit comme une combinaison linéaire des inconnues des mailles partageant l'un des sommets de la face : pour une face  $f$  interne,

$$\mathcal{F}_f^n = \sum_l T_f^l p_l^n \text{ avec } l \in \{\cup_v \tau_h(v) \text{ pour } v \text{ sommet de } f\}.$$

Pour une face  $f$  de bord, le flux s'exprime de la même manière que celui du schéma TPFA, voir l'équation (4.11). Ces coefficients  $T_f^l$  sont ensuite ajoutés dans la matrice dans chaque colonne correspondant à la maille d'indice  $l$ , et ce pour les deux lignes d'indices  $k_1$  et  $k_2$  si la face sépare les mailles  $K_1$  et  $K_2$  :

$$\mathcal{F}_{K_1L} += \Delta t T_f^l \quad \mathcal{F}_{K_2L} += \Delta t T_f^l$$

En particulier, les inconnues associées aux deux mailles  $K_1$  et  $K_2$  interviennent dans cette somme, ce qui ajoute une contribution aux termes diagonaux, mais ce ne sont pas les seules (contrairement au schéma deux points). L'autre contribution aux termes diagonaux provient, comme avec la méthode TPFA, du terme issu de la discrétisation en temps qui est ajouté au moyen d'une boucle sur les cellules :  $\mathcal{F}_{KK} += c_0 |K|$ .

Le calcul de la partie  $\mathcal{L}^p$  du second membre est identique à la méthode utilisée avec l'approximation à deux points puisque les flux de bord utilisent cette approximation.

**Termes de couplage aux éléments virtuels** La matrice de couplage  $\mathcal{B}$  ainsi que la seconde partie  $\mathcal{L}^c$  du second membre de l'équation de conservation sont assemblées exactement de la même manière que précédemment.

**Calcul du gradient et de la norme  $H^1$**  Pour calculer le gradient, ou plus précisément le flux  $v_f^K$  sur chaque maille  $K$ , on utilise à nouveau la formule de reconstruction (4.39).

$$v_f^K = \frac{1}{|K|} \sum_{f \in \partial K} \mathcal{F}_{Kf}^n (\mathbf{x}^f - \mathbf{x}^K).$$

Encore une fois, on procède au moyen d'une boucle sur les faces du maillage pour évaluer le flux numérique  $\mathcal{F}_{Kf}^n$ . La pression approchée ayant été calculée, il suffit pour chaque face de réappliquer la relation  $\mathcal{F}_{Kf}^n = \sum_l T_f^l p_l^n$  en utilisant les informations déjà calculées sur la connectivité et les transmissivités. Une fois les flux connus sur chaque face, une boucle sur les cellules est effectuée pour appliquer la formule d'interpolation.

Si le schéma multipoints permet ainsi de reconstruire un gradient discret constant par maille, il n'admet en revanche pas de définition naturelle pour la semi-norme  $H^1$  discrète qui avait été introduite avec l'approximation TPFA. Afin de pouvoir néanmoins comparer ce schéma au précédent, on construit une quantité similaire empruntant à nouveau la quantité définie par (4.13). Autrement dit, après résolution avec les flux numériques du schéma multipoints, les transmissivités définies par la méthode à deux points sont calculées de manière indépendante et utilisées pour calculer la semi-norme  $H^1$  discrète sur la solution selon la relation (4.41).

### 5.2.3 Illustrations numériques

Le comportement de cette nouvelle combinaison pour le schéma couplé est illustré sur deux problèmes. En premier lieu, on revient sur le *patch test* qui n'était pas satisfait avec le schéma VEM-TPFA. En second lieu, on observe numériquement la vitesse de convergence de ce schéma.

#### Patch test pour le schéma couplé sur maillages polyédriques

Dans la section 4.3.2 du chapitre précédent, un maillage test de type Voronoï avait été utilisé pour souligner l'impact de la non coïncidence des centres de Voronoï, qui vérifient la condition d'orthogonalité et sont donc utilisés comme centres de calcul, et des barycentres qui peuvent être utilisés pour intégrer exactement les fonctions linéaires. Cette distinction posait un problème au schéma couplé VEM-TPFA, qui a besoin de la condition d'orthogonalité pour rester consistant mais qui requiert aussi le calcul d'intégrales pour assembler les termes de couplage.

Le schéma multipoints n'étant pas soumis à la condition d'orthogonalité, cette distinction n'a plus lieu d'être : les centres de masses peuvent être utilisés comme centre de calcul. On teste à nouveau la capacité du schéma à calculer exactement ou non lorsque la solution exacte est linéaire, en l'occurrence  $\mathbf{u} = (3x - 2y, x + y)$  et  $p = -x + 2y$ . Les coefficients de couplage sont fixés de la manière suivante :  $\alpha = 1, c_0 = 0.5, E = 2.5, \nu = 0.25$  et  $\kappa = 1$ . Le domaine de calcul est le carré unité sur lequel des conditions aux limites de Dirichlet sont imposées sur tous les bords pour le déplacement et la pression. Dans un premier temps, on travaille avec le maillage test de Voronoï illustré sur la figure 4.1. Cependant, les points de calcul utilisés pour la pression sont cette fois les centres de masse des cellules. Dans cette configuration, on a déjà vu que le schéma VEM-TPFA retourne des écarts non nuls en normes  $L^2$  et  $H^1$  pour  $\mathbf{u}$  et  $p$ . Utiliser le schéma VEM-MPFA ne donne pas non plus un résultat entièrement satisfaisant : les écarts  $E_{\text{rel}}^p$  et  $E_{\text{rel}}^{\nabla p}$  obtenus sont nuls, ce qui démontre l'intérêt de cette nouvelle approximation pour le flux numérique, mais ce n'est pas le cas des écarts  $E_{\text{rel}}^{\mathbf{u}}$  et  $E_{\text{rel}}^{\sigma}$ . En réalité, la faute incombe au traitement du second membre de l'équation d'élasticité. Pour ce problème analytique, on a

$$\mathbf{f} = -\text{div}(\overline{\overline{\mathbf{C}}\overline{\overline{\mathbf{e}}}}) + \alpha \text{div}(p\overline{\overline{\mathbf{I}}}) = \alpha \nabla p.$$

Maillage	Méthode de quadrature	Méthode du potentiel
Carré	Satisfait	Satisfait
Triangles	Satisfait	Satisfait
Voronoi	En échec	Satisfait
Twisted	En échec	Satisfait

TAB. 5.1 : Patch test pour le problème couplé sur différents maillages, en fonction de la méthode utilisée pour approcher le second membre de l'équation d'élasticité.

On rappelle qu'au moment d'intégrer ce second membre contre les fonctions de bases, la méthode des éléments virtuels effectue une approximation : la contribution de la fonction de base  $\varphi_i^l$  sur la maille  $K$  se traduit par l'ajout du terme

$$\mathcal{L}_i^u += \sum_{j \in \mathcal{M}_K} \omega_j^K \mathbf{f}(V_j) \cdot \varphi_i^l(V_j) = \omega_i^K \mathbf{f}(V_i) \cdot \mathbf{e}_l \approx \int_K \mathbf{f} \cdot \varphi_i^l,$$

voir les équations (3.12) et (3.25). Autrement dit, dans le cas où  $\mathbf{f}$  est constante, on a réalisé l'approximation  $\int_K \varphi_i^l \approx \omega_i^K$ . Dans cette situation, mieux vaut employer une autre méthode suggérée par [ANR17] pour l'approximation du second membre. Celle-ci peut s'appliquer à chaque fois que la force  $\mathbf{f}$  dérive d'un potentiel, c'est-à-dire lorsqu'il existe  $F$  tel que  $\mathbf{f} = \nabla F$ , et se base sur le fait que

$$\forall \mathbf{v}_h \in V_{h,0}, \quad \int_{\Omega} \mathbf{f} \cdot \mathbf{v}_h = \int_{\Omega} \nabla F \cdot \mathbf{v}_h = - \int_{\Omega} F \operatorname{div}(\mathbf{v}_h)$$

car  $\mathbf{v}_h = \mathbf{0}$  sur  $\Gamma_{D_d} \equiv \partial\Omega$  pour ce test. Pour une fonction de base  $\varphi_i^l$  donnée, cette méthode propose l'approximation

$$\int_K \mathbf{f} \cdot \varphi_i^l = - \int_K F \operatorname{div}(\varphi_i^l) \approx -F(\mathbf{x}^K) \int_K \operatorname{div}(\varphi_i^l) = -F(\mathbf{x}^K) |K| \langle \nabla \varphi_i^l \rangle^K \cdot \mathbf{e}_l. \quad (5.5)$$

On a en effet déjà vu au paragraphe 4.3, lors de l'assemblage de la matrice de couplage  $\mathcal{B}$ , que l'intégrale de la divergence des fonctions de bases est une quantité connue dans la méthode des éléments virtuels. Cette technique tire partie de la définition "naturelle" de l'opérateur de divergence discret. Avec cette approche, l'opérateur de discrétisation du second membre est compatible avec l'opérateur de discrétisation du gradient, utilisé pour traiter le terme  $\alpha \nabla p$  lors de l'assemblage de la matrice de couplage. La conservation de cette notion de potentiel au niveau discret permet de ne pas introduire d'erreur lors de l'approximation du second membre.

En appliquant cette méthode pour discrétiser le second membre avec  $F = \alpha p = -x + 2y$ , on obtient en effet des écarts nuls en déplacement et en pression pour le problème linéaire. On note que sur certains maillages, le patch test est également satisfait sans la méthode potentielle pour l'intégration du second membre  $\mathbf{f}$ , comme indiqué dans le tableau 5.1. Quitte à adopter cette méthode pour approcher l'intégrale, le nouveau schéma couplé VEM-MPFA satisfait donc le patch test pour le système couplé sur les différents maillages dont on dispose. C'est également le cas en trois dimensions sur des maillages cubiques, tétraédriques et Voronoï.



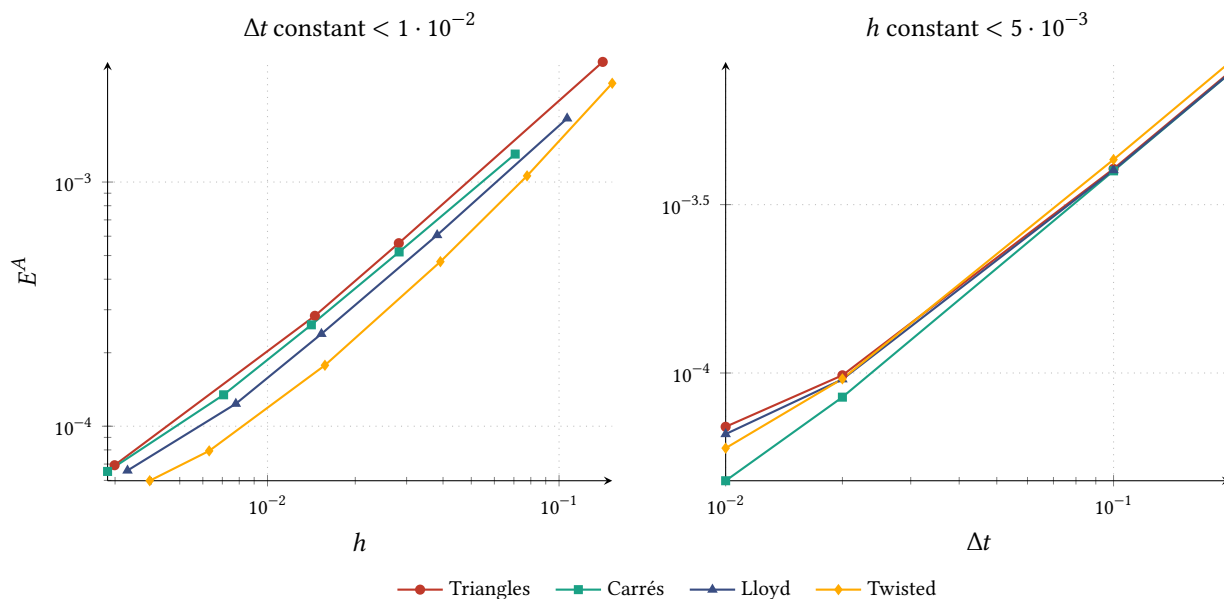


FIG. 5.6 : Évolution de l'erreur d'approximation pour 4 maillages en fonction du pas d'espace (gauche) ou du pas de temps (droite). Avec le schéma VEM-MPFA, l'erreur converge à l'ordre 1 sur les quatre catégories de maillages.

### Ordre de convergence du schéma VEM-MPFA

Même si l'étude mathématique du schéma VEM-MPFA n'a pas été effectuée (les schémas multipoints ne disposant pas à ce jour des estimations d'erreur nécessaires à la démonstration), on illustre tout de même la vitesse de convergence de celui-ci sur le cas présenté en début de chapitre. On conserve comme grandeur de travail l'erreur  $E^A$  définie par l'équation (4.35), la méthode utilisée pour calculer ses différentes composantes avec le schéma VEM-MPFA étant décrite dans le paragraphe 5.2.2. Son évolution est à nouveau représentée en fonction du pas d'espace pour un pas de temps suffisamment faible (à gauche) ou en fonction du pas de temps pour un pas spatial suffisamment faible (à droite) sur la figure 5.6. Avec le schéma multipoints pour l'approximation du flux des volumes finis, un résultat satisfaisant est obtenu sur tous les maillages, que la condition d'orthogonalité soit satisfaite ou non. De plus, on observe numériquement que l'ordre de convergence de ce schéma est de un en espace et un en temps, ce qui est identique au schéma utilisant l'approximation à deux points. Au prix de ses inconvénients précédemment mentionnés, ce schéma permet donc de conserver la même précision tout en admettant des maillages plus généraux.

## 5.3 Modification des VEMs pour les faces non planes

L'objectif de cette section est de résoudre la seconde limite soulevée dans le paragraphe 5.1.1 en adaptant la méthode des éléments virtuels aux mailles présentes dans les maillages utilisés en géosciences.

### 5.3.1 Transformation en maillage polyédrique

Pour retomber dans le cadre d'application de la méthode, la première étape consiste à se ramener à un maillage réellement polyédrique, et donc à éliminer les faces non planes. De manière assez classique, on

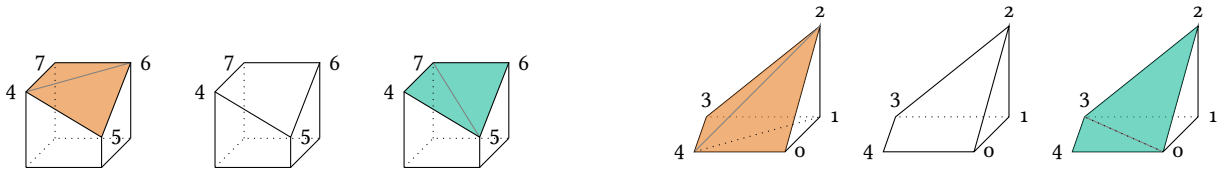


FIG. 5.7 : **Obtention de mailles polyédriques à partir d'éléments à faces non planes.** Pour chaque face non plane à 4 sommets, deux couples de sous-faces peuvent être obtenus. L'illustration est simplifiée en supposant que la face inférieure des cellules est déjà plane.

peut procéder à cette élimination en divisant les faces non planes en sous-faces planes. Se posent alors deux questions sur la manière de procéder, la première pratique, la seconde stratégique. Sur le plan pratique, un tel découpage, en admettant qu'il existe, peut s'inscrire à différentes étapes de la chaîne de simulation. Une première solution serait, à partir d'un maillage d'entrée à faces non planes, de créer en amont et sauvegarder un maillage de sortie dans lequel les mailles sont polyédriques. Cette solution nécessite un format de maillage capable de décrire des polyèdres ayant un nombre de faces arbitraire ainsi qu'une manipulation des fichiers, mais elle permet ensuite d'appliquer la méthode numérique en l'état sur le nouveau maillage. Une seconde solution consisterait à travailler sur le maillage existant au cours de la simulation, en découpant à la volée les faces non planes et en adaptant la méthode numérique à ce découpage. Cette approche élimine le prétraitement et permet de s'affranchir d'éventuelles contraintes liées au format de maillage, mais déporte les difficultés techniques dans l'adaptation de la méthode.

Sur le plan stratégique, il existe différentes divisions possibles et il faut donc choisir la technique de découpage en sous-faces planes. Une première possibilité exploite le fait que les types de mailles sont connus *a priori* et que leur diversité est limitée. En effet, dans les maillages utilisés en géosciences, toutes les faces non planes sont définies par quatre sommets : il n'y a pas de faces à plus de 4 sommets dans le maillage et celles qui n'en ont que 3 (car appartenant à une maille dégénérée) sont forcément planes. On peut donc les découper en deux sous-faces triangulaires selon l'une ou l'autre des deux diagonales, comme illustré sur la figure 5.7. De plus, le fait que seules les faces supérieures et inférieures des hexaèdres sont potentiellement non planes couplé à la structure du maillage qui adopte la même numérotation locale des sommets d'une maille à l'autre facilite ce découpage. En particulier, on peut assez facilement s'assurer qu'une face donnée est divisée de la même manière du point de vue des deux cellules qu'elle sépare, condition qui doit bien sûr être respectée pour que la nouvelle géométrie soit correctement définie. Ce choix stratégique pour la division en sous-faces restreint le choix pratique décrit dans le paragraphe précédent. En effet, l'approche consistant à réécrire en amont le maillage n'est pas compatible avec ce choix. La raison provient du support limité des mailles polyédriques par le simulateur : les hexaèdres peuvent bien être redéfinis comme des polyèdres à 8 faces, en revanche le découpage vert de l'hexaèdre dégénéré représenté à droite de la figure 5.7 n'est pas autorisé par la structure de données. On ne peut en effet pas, au sein de la même maille, déclarer deux faces différentes par la même liste de sommets (ici formées par les sommets 0, 3 et 4 vus "de dessous" ou "de dessus"). Or, dans certaines situations, il faut découper suivant cet axe pour respecter la conformité du maillage. La seule possibilité est donc de modifier la méthode numérique pour tenir compte de ce découpage. En pratique, la partie à modifier dans la méthode des éléments virtuels est le calcul de la projection des fonctions de bases, et plus spécifiquement celui du gradient moyen de ces fonctions de base dont l'expression est donnée par (3.18). Considérons par exemple le découpage vert de la maille hexaédrique de la figure 5.7. On note  $T_a$  le triangle formé par les nœuds 4, 5 et 7 et  $T_b$  celui formé par les nœuds 5, 6 et 7. La projection des fonctions de bases associées aux nœuds de la face inférieure est inchangée puisque

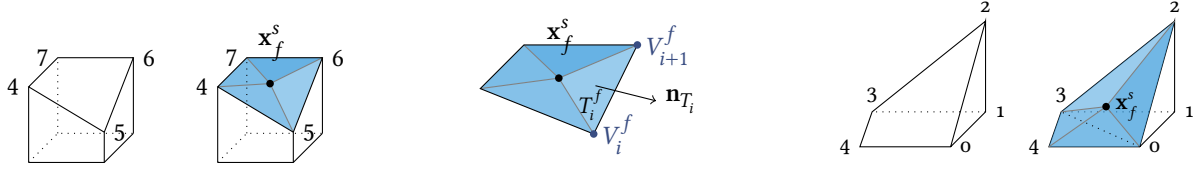


FIG. 5.8 : **Obtention de mailles polyédriques à partir d'éléments à faces non planes.** Chaque face non plane est triangulée par rapport à la moyenne arithmétique des sommets  $\mathbf{x}_s^f$ . L'illustration est simplifiée en supposant que seule la face supérieure des cellules n'est pas plane. Au centre, le schéma indique les notations utilisées dans la suite.

toutes les faces auxquelles appartiennent ces nœuds sont planes. Pour le nœud 4, qui appartient à une des deux sous-faces, et pour le nœud 5, qui appartient aux deux sous-faces, on a respectivement

$$\frac{1}{|K|} \int_K \nabla \varphi_4 = \frac{1}{|K|} \sum_{f \in F_{4,K}^p} \left( \frac{|f|}{M_f} + \frac{1}{2} \left( |e_{f,4}^-| \mathbf{n}_{f,4}^- + |e_{f,4}^+| \mathbf{n}_{f,4}^+ \right) \cdot (\mathbf{x}^f - \bar{\mathbf{x}}^f) \right) \mathbf{n}_f + \frac{1}{|K|} \frac{|T_a|}{3} \mathbf{n}_{T_a}$$

$$\frac{1}{|K|} \int_K \nabla \varphi_5 = \frac{1}{|K|} \sum_{f \in F_{5,K}^p} \left( \frac{|f|}{M_f} + \frac{1}{2} \left( |e_{f,5}^-| \mathbf{n}_{f,5}^- + |e_{f,5}^+| \mathbf{n}_{f,5}^+ \right) \cdot (\mathbf{x}^f - \bar{\mathbf{x}}^f) \right) \mathbf{n}_f + \frac{1}{|K|} \frac{|T_a|}{3} \mathbf{n}_{T_a} + \frac{1}{|K|} \frac{|T_b|}{3} \mathbf{n}_{T_b}$$

où l'on a noté  $F_{i,K}^p$  le sous-ensemble des faces de  $F_{i,K}$  qui étaient déjà planes (sur cet exemple, il s'agit des faces latérales), et où on a utilisé le fait que  $\mathbf{x}^f = \bar{\mathbf{x}}^f$  sur une face triangulaire pour simplifier l'expression. Il suffit donc de modifier le calcul des termes  $\mathbf{q}_i$  (cf équation (3.22)) dans l'assemblage des matrices locales. Le calcul des volumes  $|K|$  est également adapté pour tenir compte de la nouvelle géométrie des mailles. Cette stratégie de division est donc utilisable lorsqu'elle est implémentée directement dans le code, en ne nécessitant de plus que peu d'ajustements. Cependant, si elle permet de s'affranchir de la limite technique posée par la structure de donnée, elle manipule implicitement des objets qui ne sont pas des polyèdres, ce qui peut conceptuellement être gênant. De plus, elle n'est pas directement généralisable à des faces non planes présentant plus de 4 sommets ou à des mailles dont les faces latérales seraient également non planes.

On introduit donc une deuxième stratégie de division en sous-mailles plus générale, qui triangule les faces non planes par rapport à un point  $\mathbf{x}_s^f$ . La définition de ce point étant relativement libre, on le calcule à partir de la moyenne arithmétique des sommets de la face non plane :  $\mathbf{x}_s^f = \frac{1}{M_f} \sum_{j \in \mathcal{M}_f} \mathbf{x}(V_j)$ . Chaque face non plane est donc divisée en  $M_f$  triangles  $T_i^f$  de normales sortantes  $\mathbf{n}_{T_i}$ . Cette triangulation est schématisée sur la figure 5.8 pour les mêmes éléments que précédemment, mais elle s'étend évidemment aux faces comportant un nombre quelconque de sommets. De plus, contrairement à l'approche précédente, celle-ci permettrait d'adopter la stratégie pratique consistant à réécrire le maillage en tant qu'éléments polyédrique (elle est du moins compatible avec les formats de maillages disponibles), ou la stratégie de modification du schéma numérique à la volée (en s'inspirant du paragraphe précédent). Dans les deux cas, la difficulté technique provient de l'introduction de sommets supplémentaires qui ajoutent des inconnues au système linéaire. Modifier le maillage demande l'introduction de ces sommets, et modifier la méthode demande l'indexation de ces sommets. Même si cette difficulté est surmontable, il n'empêche que rajouter des sommets a un coût très important en terme de nombre d'inconnues puisque chaque sommet génère trois inconnues en déplacement. Appliquer cette méthode sur le maillage d'exemple (figure 5.4) reviendrait à peu près à doubler le nombre d'inconnues en déplacement. Pour éviter cela, la section suivante présente une modifi-

Méthode de division	Par modification du maillage	Par modification de la méthode	Commentaire
Diagonale	Impossible*	Réalisé	Généralisation à $M_f > 4$ délicate
Triangulation	Possible	Possible	Inconnues supplémentaires à éviter (cf 5.3.2)

(\*) Avec les formats de maillage disponibles

TAB. 5.2 : Résumé de la faisabilité des combinaisons permettant la transformation en maillage polyédrique.

cation de la méthode des éléments virtuels qui utilise cette triangulation mais n'ajoute pas d'inconnues au système.

### 5.3.2 Un espace d'approximation modifié pour les éléments virtuels

La méthode proposée pour éviter d'ajouter des inconnues aux points de triangulation provient de [CBP17], où des problèmes de déformations finies sont considérés. Pour ce genre de problèmes, même si les faces sont planes dans la configuration initiale, il est possible qu'elles ne le soient plus après application de la déformation ce qui justifie ce besoin de triangulation. L'idée est de modifier l'espace d'approximation de manière à pouvoir exprimer les déplacements aux centres de triangulations en fonction des déplacements des autres sommets de la face. Ainsi, l'espace d'approximation local  $V_{h,scal}^K$  des éléments virtuels n'est plus défini par (3.3) mais est remplacé par

$$V_{h,scal}^K = \left\{ v_h \in H^1(K) \text{ tel que } v_h|_{\partial K} \in \mathcal{C}^0(\partial K), \quad v_h|_{T_j^f} \in \mathbb{P}_1 \quad \forall T_j^f \subset \partial K, \right. \\ \left. v_h(\mathbf{x}_s^f) = \frac{1}{M_f} \sum_{j \in \mathcal{M}_f} v_h(V_j^f) \quad \forall f \in \partial K \quad \text{et} \quad \Delta v_h = 0 \text{ dans } K \right\}. \quad (5.6)$$

Cet espace satisfait toujours les propriétés énoncées au début de la section 3.1.1 :

- Si  $q$  appartient à  $\mathbb{P}_1(K)$ , il est évident que  $q|_{T_j^f} \in \mathbb{P}_1(K)$  et  $\Delta q = 0$ . De plus, par définition de  $\mathbf{x}_s^f$  et par linéarité,  $q(\mathbf{x}_s^f) = q\left(\frac{1}{M_f} \sum_{j \in \mathcal{M}_f} V_j^f\right) = \frac{1}{M_f} \sum_{j \in \mathcal{M}_f} q(V_j^f)$  donc  $q \in V_{h,scal}^K$ .
- Si deux fonctions  $v^a \in V_{h,scal}^K$  et  $v^b \in V_{h,scal}^K$  partageant les mêmes valeurs sur les  $M_K$  sommets de  $K$ , alors  $v^a(\mathbf{x}_s^f) = v^b(\mathbf{x}_s^f)$ . Sur chaque face triangulaire,  $v^- := v^a - v^b = 0$  car  $v^- \in \mathbb{P}_1$  et s'annule en trois points distincts. Puisque  $\Delta v^- = 0$  dans  $K$ , il s'ensuit que  $v^- = 0$  et donc que  $v^a = v^b$ .
- Pour une face  $f$  donnée, la trace de  $v_h$  sur  $f$  ne dépend que des valeurs de  $v_h$  sur les sommets de  $f$  et en  $\mathbf{x}_s^f$ . Puisque cette dernière valeur est une combinaison linéaire des premières, la trace de  $v_h$  ne dépend finalement que des valeurs de  $v_h$  sur les sommets de  $f$ .

Une fois cette définition adoptée pour  $V_{h,scal}^K$  l'espace global scalaire ainsi que les espaces vectoriels  $V_h^K$  et  $V_h$  sont construits de la même manière que dans le chapitre 3. En particulier, les degrés de liberté sont toujours les valeurs aux sommets originaux des faces et les fonctions de bases (vectorielles) sont toujours définies par  $\varphi_i^k = \varphi_i \mathbf{e}_k$  ( $k = x, y, z$ ) avec  $\varphi_i(V_j) = \delta_{i,j} \quad \forall j \in \mathcal{M}$ .

La construction de la forme bilinéaire approchée, celle de la formulation variationnelle approchée ainsi que l'assemblage du système linéaire ne sont pas modifiés par rapport aux méthodes décrites dans la section

3.1.2. En revanche, comme avec la stratégie de division précédente, le calcul de la projection des fonctions de bases doit être adapté : en effet, le nouvel espace d'approximation défini par (5.6) ne contient plus explicitement la propriété  $\int_f \pi^{f,0} v_h = \int_f v_h$ , qui avait été ajoutée dans le but spécifique de pouvoir calculer les projections à partir des degrés de liberté. Plus précisément, le calcul des termes  $\mathbf{q}_i = \frac{1}{|K|} \int_K \nabla \varphi_i$  doit être adapté en utilisant les informations contenues dans l'espace d'approximation. En l'occurrence, la formule de Stokes est cette fois appliquée sur les sous-faces triangulaires

$$\mathbf{q}_i = \frac{1}{|K|} \int_K \nabla \varphi_i = \frac{1}{|K|} \sum_f \sum_{j=1}^{M_f} \int_{T_j^f} \varphi_i \mathbf{n}_{T_j} = \frac{1}{|K|} \sum_f \sum_{j=1}^{M_f} \frac{|T_j^f|}{3} \left( \varphi_i(V_j) + \varphi_i(V_{j+1}) + \varphi_i(\mathbf{x}_s^f) \right) \mathbf{n}_{T_j}$$

sur lesquelles  $\varphi_i$  peut être intégrée exactement en tant que quantité linéaire. En utilisant l'autre propriété de l'espace d'approximation  $\varphi_i(\mathbf{x}_s^f) = \frac{1}{M_f} \sum_{j \in \mathcal{M}_f} \varphi_i(V_j^f) = \frac{1}{M_f}$  car  $\varphi_i(V_j^f) = \delta_{ij}$ , on obtient finalement

$$\mathbf{q}_i = \frac{1}{3|K|} \sum_{f \in \mathcal{F}_{i,K}} \left( |T_i^f| \mathbf{n}_{T_i^f} + |T_{i+1}^f| \mathbf{n}_{T_{i+1}^f} + \frac{1}{M_f} \sum_{j=1}^{M_f} |T_j^f| \mathbf{n}_{T_j} \right). \quad (5.7)$$

*Remarque.* Dans le cas où les faces originales sont déjà planes, les espaces définis par (3.3) et (5.6) sont équivalents au sens où ils partagent la même dimension et les mêmes degrés de liberté : un ensemble de valeurs aux sommets donné identifie de manière unique une fonction de l'espace original ou de l'espace modifié. La contribution de la face  $f$  au vecteur  $\mathbf{q}_i$  peut donc être calculée ou bien avec l'expression originale (3.18), ou bien avec la nouvelle expression (5.7). Dans l'implémentation effectuée, chaque face originale possède un attribut booléen permettant d'indiquer si elle est plane ou non. Seules les faces non planes sont triangulées et contribuent à  $\mathbf{q}_i$  via la technique développée dans ce paragraphe.

Les mailles à faces non planes amènent également une difficulté supplémentaire dans la prise en compte des coefficients aux limites de Neumann. D'une part parce que les faces de bord n'étant plus planes, le vecteur traction  $\mathbf{t}_N = \overline{\sigma^e} \cdot \mathbf{n}$  ne peut plus être facilement calculé en amont puis imposé en tant que condition aux limites : d'autre part parce que la formule d'approximation du second membre (3.25) avait été établie pour une face de bord plane. On propose donc une autre approximation des conditions de Neumann qui calcule le vecteur  $\mathbf{t}_N$  sur chaque sous-face et qui exploite l'appartenance de  $\varphi_i$  à  $\mathcal{P}_1$  pour raisonner comme lors du calcul de  $\mathbf{q}_i$  : en utilisant la symétrie de  $\overline{\sigma^e}$ ,

$$\begin{aligned} \int_f \overline{\sigma^e} \mathbf{n} \cdot \boldsymbol{\varphi}_i^k &= \sum_{j=1}^{M_f} \int_{T_j^f} \overline{\sigma^e} \mathbf{n} \cdot \varphi_i \mathbf{e}_k \approx \sum_{j=1}^{M_f} \overline{\sigma^e}(V_j) \mathbf{e}_k \cdot \int_{T_j^f} \varphi_i \mathbf{n}_{T_j} \\ &= \overline{\sigma^e}(V_i) \mathbf{e}_k \cdot \left( |T_i^f| \mathbf{n}_{T_i^f} + |T_{i+1}^f| \mathbf{n}_{T_{i+1}^f} + \frac{1}{M_f} \sum_{j=1}^{M_f} |T_j^f| \mathbf{n}_{T_j} \right). \end{aligned} \quad (5.8)$$

Pour mesurer l'amélioration apportée par ce nouvel espace et par les méthodes de calcul des projections et des conditions aux limites de Neumann qui y sont associées, on considère une fois encore un patch-test pour le problème de mécanique. La solution exacte à atteindre est  $\mathbf{u} = 10^{-3} z \mathbf{e}_z$ , avec des coefficients élastiques fixés à  $E = 1$ ,  $\nu = 0.25$ . Le maillage considéré est toujours le maillage géosciences de test représenté sur la figure 5.4. Des conditions aux limites de Dirichlet sont imposées sur les faces latérales  $X_{\max}$  et  $Y_{\max}$  ainsi que sur la face supérieure, tandis que des conditions aux limites de Neumann, calculées

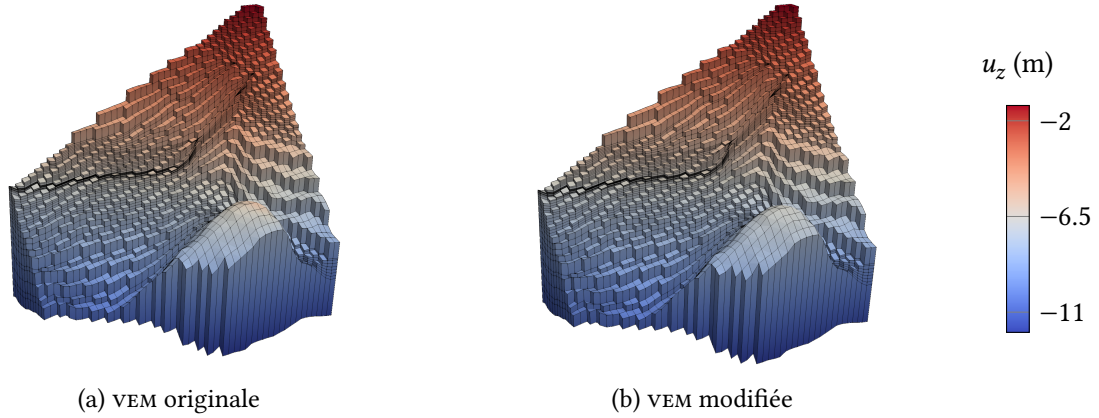


FIG. 5.9 : **Comparaison de la version modifiée des éléments virtuels à la version originale** sur un maillage typique des géosciences (vue en coupe). Ce problème patch test dont la solution exacte vaut  $\mathbf{u} = 10^{-3}z\mathbf{e}_z$  est satisfait avec la version modifiée ( $E_{\text{rel}}^{\mathbf{u}} = 10^{-12}$ ), mais pas avec la version originale ( $E_{\text{rel}}^{\mathbf{u}} = 10^{-2}$ ), même si les deux solutions sont visuellement proches.

à partir de la solution exacte, sont imposées sur les autres faces. Dans la version dite originale, on prescrit le vecteur traction sur chacune des 3 autres faces du domaine :

$$\mathbf{t}_N = \begin{cases} -4 \cdot 10^{-4} \mathbf{e}_x & \text{sur } X = 0, \\ -4 \cdot 10^{-4} \mathbf{e}_y & \text{sur } Y = 0, \\ -1.2 \cdot 10^{-3} \mathbf{e}_z & \text{sur la face inférieure.} \end{cases}$$

On calcule ensuite les projections avec la méthode développée pour les faces planes dans le chapitre 3. Toutefois, les grandeurs géométriques telles que aires, volumes et normales ont été calculées en tenant compte des faces non planes et en les triangulant, ce qui explique que le résultat ne soit pas si éloigné de la solution attendue. Dans la version dite modifiée, la face inférieure n'étant pas plane (voir figure 5.4c), on adopte la technique développée dans l'équation (5.8) en fournissant pour cette face la matrice des contraintes  $\hat{\sigma}^e = (4 \cdot 10^{-4}, 4 \cdot 10^{-4}, 1.2 \cdot 10^{-3}, 0, 0, 0)^T$ . De même, le calcul des matrices locales est effectué en utilisant l'espace et la méthode détaillés dans ce paragraphe. Pour ces deux méthodes, le champ de déplacement  $u^z$  est représenté sur la figure 5.9, vue en coupe avec reconstruction des mailles traversées par le plan de coupe. La différence entre les deux approches est distinguable à l'œil nu même si les deux résultats semblent proches. C'est en calculant les écarts avec la solution exacte  $E_{\text{rel}}^{\mathbf{u}}$  et  $E_{\text{rel}}^{\sigma}$  que l'on s'aperçoit que la version modifiée est bien plus précise : on obtient des écarts de l'ordre de  $10^{-12}$  pour les déplacements et  $10^{-12}$  pour les contraintes avec la méthode modifiée, qui satisfait donc le patch test, alors que la méthode originale produit des écarts de l'ordre de  $10^{-2}$  pour déplacements et  $10^{-1}$  pour les contraintes.

### 5.3.3 Un mot sur les volumes finis

Tout comme des modifications ont été apportées à la méthode des éléments virtuels dans le but de traiter les faces non planes, la méthode des volumes finis doit être adaptée à cette même fin pour conserver un schéma couplé performant. Sans proposer un traitement général comme pour la méthode des éléments virtuels, ni rentrer dans les détails, on relève dans [AEK06] une modification assez simple à apporter au schéma MPFA-O pour traiter les faces quadrilatères non planes : il suffit d'adopter une certaine expression

pour le calcul des vecteurs normaux aux faces. Ceci revient à transformer les quadrilatères non plans en des parallélogrammes plans. D'après cette même référence, avec une telle modification, le patch test est satisfait lorsque le tenseur  $\bar{\kappa}$  est homogène ; dans le cas d'un tenseur hétérogène, la pression et le flux calculés convergent respectivement à l'ordre 2 et 1 en espace. En réutilisant l'implémentation des schémas volumes finis multipoints disponible dans la plateforme Arcane, on constate sur le maillage pris pour exemple, avec une solution linéaire  $p = 10^{-3}z$  et un tenseur homogène  $\bar{\kappa} = \bar{I}_3$ , que le schéma multipoints ne retrouve pas la solution exacte. Néanmoins, on observe un gain de deux ordres de grandeur sur l'écart  $E_{\text{rel}}^p$  par rapport au schéma à deux points.

## 5.4 Illustration sur une application géosciences

Les différents ingrédients introduits dans ce chapitre ont permis de proposer un second schéma couplé conçu pour lever les limitations dont souffrait le schéma VEM-TPFA. Pour compléter les exemples synthétiques des sections précédentes, on propose une illustration sur une application réaliste de géosciences, à défaut de disposer d'une situation réelle. Il s'agit de simuler l'extraction, via un puits vertical, d'un fluide aqueux initialement piégé dans un milieu poreux souterrain stratifié.

### 5.4.1 Paramètres du cas

On réutilise pour ce cas le maillage précédent, mais en redéfinissant les matériaux rocheux présents dans le domaine : la répartition des strates ainsi que les caractéristiques des trois matériaux utilisés sont détaillées sur la figure 5.10. La porosité  $\phi$  est utilisée pour calculer le coefficient de stockage  $c_0$  selon la relation  $c_0 = \phi c_f$ , où  $c_f = 4.5 \cdot 10^{-10} \text{ Pa}^{-1}$  est le module de compressibilité de l'eau. La perméabilité varie fortement d'un matériau à l'autre : le réservoir et les grès, relativement perméables, sont séparés par des couches d'argile très peu perméables. Ceci implique plusieurs ordres de grandeur de différence entre les matériaux pour la matrice de mobilité, définie comme le rapport entre la perméabilité et la viscosité dynamique de l'eau  $\mu_f = 10^{-3} \text{ Pa}\cdot\text{s}$ . En raison du poids exercé par les couches supérieures, le réservoir se compacte durant la production ce qui génère un phénomène de subsidence. Celui-ci devrait mécaniquement se modéliser par des déformations plastiques. Néanmoins, le modèle physique utilisé ici permettant uniquement la description de déformations élastiques linéaires, on procède autrement pour prendre en compte cet effet : le module de Young est abaissé à  $E = 1 \text{ GPa}$  dans le réservoir contre  $E = 5 \text{ GPa}$  dans le reste du domaine. Dans le but de simplifier cette étude, les coefficients de Poisson  $\nu$  et de Biot  $\alpha$  sont maintenus constants.

Pour ce genre de problème où la gravité intervient, il est d'usage de raisonner en terme de surpression  $\tilde{p} = p - (P_0 - \rho_f g|z|)$  où  $\rho_f = 1000 \text{ kg}\cdot\text{m}^{-3}$  est la masse volumique de l'eau et  $P_0 = 1 \text{ bar}$  la pression atmosphérique. Ceci permet d'isoler les effets liés aux sollicitations mécaniques de la pression hydrostatique due à la profondeur. De plus, avec ce point de vue, le schéma volumes finis pour la partie écoulement n'a pas à tenir compte du terme de gravité lors du calcul des flux. Au passage, pour une configuration physique de ce genre où le second membre de l'équation de mécanique est le terme de gravité  $\mathbf{f} = \rho \mathbf{g} = (\phi \rho_f + (1 - \phi) \rho_s) \mathbf{g}$ ,  $\rho$  désignant la masse volumique de l'ensemble et  $\rho_s$  celle du solide, le système (2.17) se réécrit en variables  $(\mathbf{u}, \tilde{p})$

$$\begin{aligned} -\mathbf{div} \left( \bar{\bar{C}} \bar{\bar{\epsilon}}_u - \alpha \bar{\bar{p}} \bar{\bar{I}}_d \right) &= ((1 - \phi) \rho_s - (\alpha - \phi) \rho_f) \mathbf{g} \\ \partial_t (c_0 \tilde{p}) + \alpha \mathbf{div}(\mathbf{u}) + \mathbf{div}(-\bar{\kappa}(\nabla \tilde{p})) &= q. \end{aligned}$$

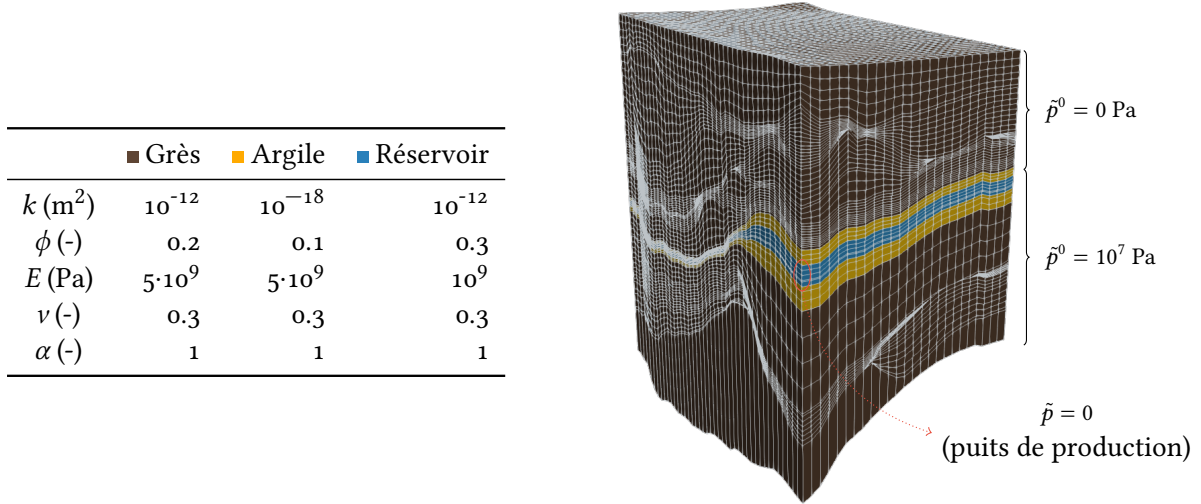


FIG. 5.10 : **Maillage et caractéristiques des strates géologiques** utilisées pour le cas de subsidence. Les conditions initiales en surpression et la position du puits de production sont détaillées à droite de la figure.

Pour ce problème, on prend  $\rho_s = 2650 \text{ kg.m}^{-3}$  dans la couche réservoir et on choisit  $\rho_s$  dans les deux autres matériaux de façon à ce que le produit  $(1 - \phi)\rho_s$  soit constant dans le domaine (soit  $\rho_s \approx 2320 \text{ kg.m}^{-3}$  pour le grès et  $\rho_s \approx 2060 \text{ kg.m}^{-3}$  pour l'argile). On signale pour clore cet aparté qu'une telle force  $\mathbf{f}$  dérive d'un potentiel : puisque  $\mathbf{g} = -|g|\nabla z$ , on a  $\mathbf{f} = \nabla(-|g|((1 - \phi)\rho_s + (\alpha - \phi)\rho_f)z)$  qui peut donc être approchée selon l'équation (5.5). Dans ce formalisme, la pression initiale imposée est une surpression de 100 bars dans le réservoir et dans les couches inférieures d'argile et de grès. Dans les couches recouvrant le réservoir, la surpression initiale est nulle (en d'autres termes, la pression est égale à la pression hydrostatique). Cette surpression permet de calculer un champ de déplacement initial en résolvant le problème de mécanique. Une fois cet état initial déterminé, en notant  $\mathbf{n}$  la normale sortante unitaire aux faces, les conditions aux limites sont les suivantes :

- sur les bords latéraux, le déplacement dans la direction orthogonale à la face ainsi que le flux sont fixés à 0 :  $\mathbf{u} \cdot \mathbf{n} = 0$  et  $\nabla \tilde{p} \cdot \mathbf{n} = 0$ . Pour les autres composantes du déplacement, une condition de contrainte libre est adoptée ;
- sur la face inférieure, le déplacement vertical et le flux sont fixés à 0 :  $u_z = 0$  et  $\nabla \tilde{p} \cdot \mathbf{n} = 0$  ;
- sur la face supérieure, les tractions mécaniques et la surpression sont nulles :  $\overline{\sigma^e} \mathbf{n} = \mathbf{0}$  et  $\tilde{p} = 0$  ;
- dans un des coins du domaine et au niveau du réservoir, la surpression est fixée à zéro de manière à modéliser une connectivité hydraulique infinie avec la surface et donc la partie productive du puits. En pratique, la condition ( $\tilde{p} = 0$ ) est imposée sur un petit groupe de 6 faces latérales appartenant au réservoir et ayant un sommet sur la ligne verticale  $(x_{\max}, y_{\max})$ , voir figure 5.10.

La simulation couvre une période d'un an divisée en 20 pas de temps.

#### 5.4.2 Résultats numériques

Dès que la condition de surpression nulle est imposée dans le coin du réservoir, le fluide piégé à l'intérieur de celui-ci commence à se déplacer vers le puits de production et est progressivement drainé au



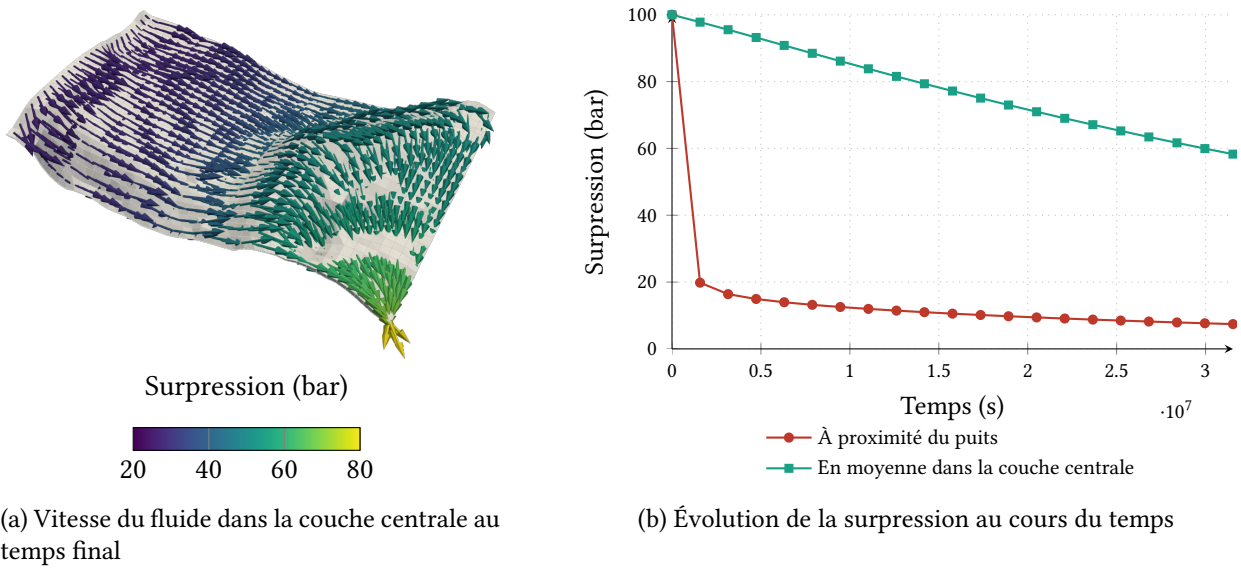


FIG. 5.11 : **Surpression et flux calculés dans la couche centrale du réservoir**, en fonction de l'espace au temps final (gauche), ou moyennée en espace et tracée en fonction du temps (droite).

fur et mesure que le temps s'écoule. Ce phénomène est représenté sur la figure 5.11. Au temps final de la simulation, les vecteurs flux  $\mathbf{v}_f$  colorés d'après la valeur de la surpression sont représentés dans la couche centrale du réservoir, en fonction de l'espace, sur la figure 5.11a. La surpression tend vers 0 à proximité du puits, alors qu'elle est toujours proche de sa valeur initiale de 100 bars à l'extrémité opposée du réservoir. Inversement, la surpression en une maille adjacente au puits ainsi que sa valeur moyenne dans toute la couche centrale du réservoir sont représentées en fonction du temps sur la figure 5.11b. On observe qu'à proximité du puits, la surpression se stabilise très rapidement (dès le premier pas de temps) autour de la valeur imposée. En revanche, la valeur moyenne dans la couche diminue de manière régulière durant l'extraction du fluide. Avec les paramètres choisis, cette valeur est à peine divisée par deux en fin de simulation. On note que dans les autres matériaux, en raison des barrières de perméabilité que constituent les couches d'argile, la surpression reste quasiment constante.

La conséquence de cette chute de pression est une diminution de la contrainte effective  $\bar{\sigma}^e = \bar{\sigma} + \alpha p \bar{l}_3$ . Cette chute est bien visible à proximité du puits, mais moins remarquable à distance de celui-ci puisque la surpression varie moins fortement au milieu du domaine. Ceci peut être observé sur la figure 5.12 où les contraintes totales (au signe près) sont tracées en fonction de la profondeur le long d'une ligne horizontale, située soit dans le coin productif, soit au centre du domaine. Sur ce graphe, la pression représentée est la pression absolue qui tient compte du terme hydrostatique et non la surpression. Les contraintes effectives peuvent donc être graphiquement retrouvées dans la différence entre les courbes de contraintes totales et la courbe de pression. À distance du réservoir, les contraintes horizontales  $\sigma_{xx}$  et  $\sigma_{yy}$  sont presque confondues dans la partie supérieure du domaine : la distinction apparaît dans la partie profonde du domaine, probablement en raison de la condition limite imposée sur la face inférieure : puisque seul le déplacement dans la direction  $\mathbf{e}_z$  est fixé, il peut y avoir un glissement dans les autres directions qui ne sont pas géométriquement équivalentes, la face inférieure n'étant pas un plan horizontal.

La diminution des contraintes effectives au niveau du puits implique les déformations et le compactage du réservoir. Pour observer les effets de l'extraction du fluide, on définit les déplacements relatifs  $\Delta \mathbf{u}$  par

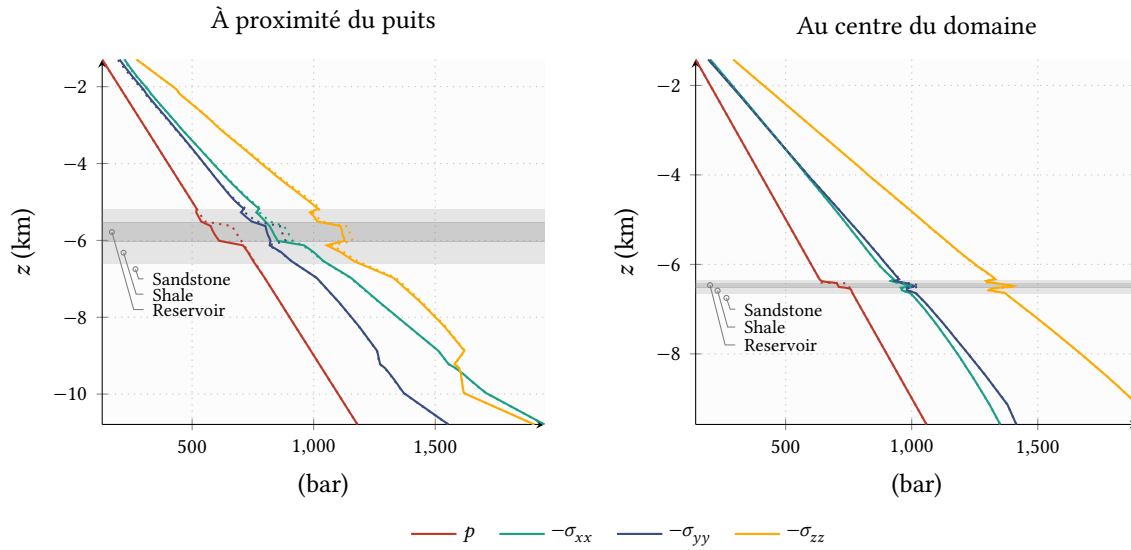


FIG. 5.12 : Évolution de la pression et des composantes diagonales du tenseur des contraintes le long d'une ligne verticale située à proximité du puits (gauche) ou au centre du domaine (droite). En pointillés, valeurs au temps initial ; en traits pleins, valeurs au temps final.

rapport à l'état initial en soustrayant le déplacement initial  $\mathbf{u}^0$  (du à la gravité) au champ de déplacements final. Au temps final, les déplacements relatifs dans les trois directions sont représentés sur la figure 5.13. L'effet prédominant s'observe sur le déplacement vertical  $\Delta u_z$  : en raison de l'extraction du fluide, la partie située au dessus du réservoir s'affaisse d'environ 1,5 mètres. C'est le phénomène de subsidence. En parallèle, la partie située en dessous du réservoir s'élève d'environ un demi-mètre. Cette légère décompaction est due au fait que la masse du réservoir est plus faible après extraction d'une partie du fluide. En plus de ce déplacement vertical, on observe que la subsidence entraîne la flexion de la partie supérieure du réservoir. Ce phénomène, dont l'intensité est tout de même moindre, est particulièrement visible sur les déplacements  $\Delta u_x$ .

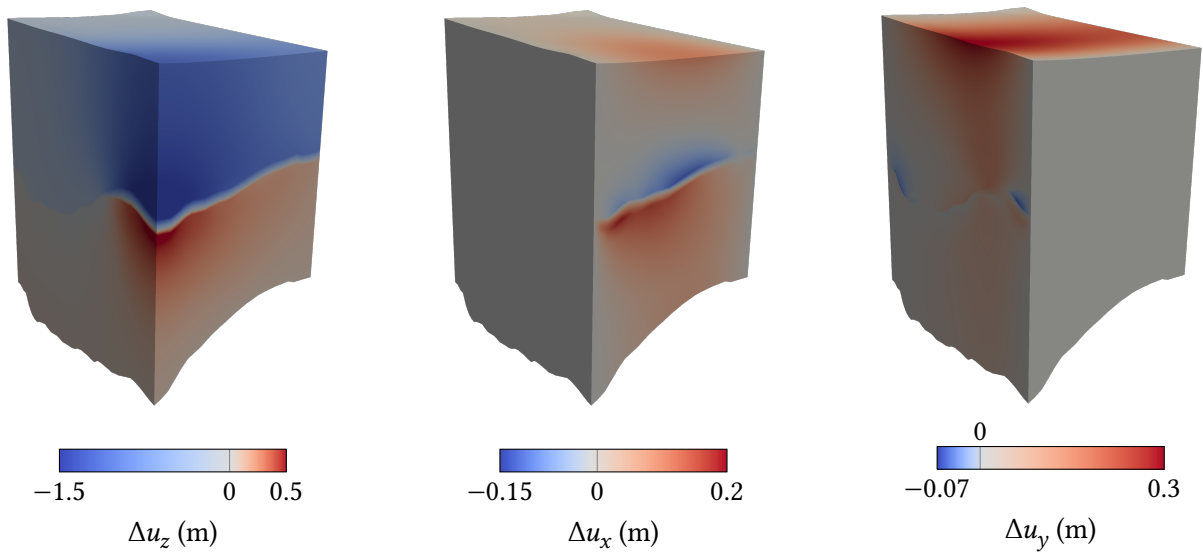


FIG. 5.13 : **Déplacements relatifs au temps final de la simulation.** Le phénomène dominant est la subsidence, affaissement de la partie supérieure du domaine.

# 6

## Au-delà de la discrétisation : stratégies de résolution du système poroélastique

Dans les chapitres précédents, une méthode de discrétisation pour le système de poroélasticité a été construite, étudiée et améliorée pour répondre aux problématiques spécifiques aux applications. Ce chapitre aborde le système poroélastique à travers un autre angle d'approche : celui de la résolution du système linéaire issu de la discrétisation. Ainsi, jusqu'à la fin de ce chapitre, la discrétisation est simplifiée en utilisant le schéma volumes finis deux points plutôt que le schéma multipoints pour la partie écoulement. Ces travaux sont motivés par l'incapacité des méthodes classiques à résoudre efficacement les systèmes obtenus dans des cas réels tels que celui illustrant le chapitre 5. Ce point est détaillé dans la première section du présent chapitre, qui donne également un aperçu des stratégies envisageables. Parmi celle-ci, la méthode séquentielle dite *fixed-stress* est détaillée dans la section 6.2. En réinterprétant celle-ci comme un problème d'annulation de résidu, la section 6.3 en propose une variante accélérée. Le principal intérêt derrière cette méthode étant de pouvoir choisir un sous-solveur pour chaque sous-problème, la section 6.4 est dédiée à la recherche de deux solveurs, chacun performant sur l'un des deux sous-problèmes physiques. On qualifie ainsi une méthode multigrille algébrique pour le sous-problème en pression, et une méthode de décomposition de domaine multi-niveaux pour le sous-problème en déplacements. Ces deux méthodes sont finalement mises en place et utilisées en tant que sous-solveurs dans le cadre de l'algorithme *fixed-stress* dans la section 6.5.

### Sommaire

---

<b>6.1</b>	<b>Panorama des différentes méthodes . . . . .</b>	<b>108</b>
6.1.1	Méthode directe ou méthode itérative . . . . .	108
6.1.2	Résolution monolithique ou résolution séquentielle . . . . .	109
6.1.3	Choix de la stratégie . . . . .	110
<b>6.2</b>	<b>Stratégie séquentielle <i>fixed-stress</i> . . . . .</b>	<b>112</b>
6.2.1	Formulation de l'algorithme <i>fixed-stress</i> . . . . .	112
6.2.2	Interprétation de la stratégie <i>fixed-stress</i> comme une méthode de point fixe . . . . .	114
6.2.3	Comparaison des différentes variantes . . . . .	116

<b>6.3</b>	<b>Accélération de la stratégie fixed-stress</b>	<b>117</b>
6.3.1	Interprétation de la stratégie fixed-stress comme une méthode de Krylov.	118
6.3.2	Comparaison numérique des algorithmes	120
<b>6.4</b>	<b>Recherche de sous-solveurs efficaces pour les sous-problèmes.</b>	<b>120</b>
6.4.1	Préconditionneurs ILU et AMG	120
6.4.2	Méthode de décomposition de domaine pour l'élasticité	123
<b>6.5</b>	<b>Utilisation des sous-solveurs spécialisés parallèles dans le cadre de la stratégie fixed-stress.</b>	<b>126</b>
6.5.1	Utilisation de la librairie hpddm pour la décomposition de domaine.	126
6.5.2	Comportement sur le cas synthétique hétérogène.	129
6.5.3	Application au problème tridimensionnel réaliste.	131

---

## 6.1 Panorama des différentes méthodes

Différentes techniques vont être abordées dans ce chapitre, parfois de manière indépendante, parfois en tant qu'élément d'une approche englobante. Pour faciliter la compréhension des concepts manipulés et de leurs interactions, cette section propose d'entrée de jeu un aperçu des stratégies envisageables. Le choix de la stratégie à suivre est justifié en fin de section à partir du cas d'application tridimensionnel proposé dans le chapitre précédent.

### 6.1.1 Méthode directe ou méthode itérative

Du point de vue le plus proche de la résolution, et quelle que soit la méthode utilisée en amont, le cœur de l'algèbre linéaire consiste à résoudre un système linéaire réel de la forme  $Ax = b$  avec  $A \in \mathbb{R}^{m \times n}$ ,  $x \in \mathbb{R}^n$  et  $b \in \mathbb{R}^m$ . On classe en général les algorithmes de résolution en deux grandes catégories : les méthodes *directes* et les méthodes *itératives* [AKo8].

Avec une méthode directe, la solution "exacte" du système peut être déterminée en un nombre fini d'opérations. Les guillemets sont utilisés pour souligner le fait que même si l'algorithme donne la vraie solution en arithmétique exacte, l'utilisation de l'outil numérique introduit inévitablement des erreurs d'arrondis lors de la manipulation de coefficients réels (de l'ordre de  $10^{-16}$  par opération sur une machine standard). Un algorithme célèbre est la factorisation LU, qui consiste à réécrire la matrice  $A$  comme le produit d'une matrice triangulaire inférieure avec une matrice triangulaire supérieure en suivant des étapes inspirées de la méthode de Gauss, pour ensuite obtenir la solution en résolvant deux systèmes triangulaires (ce qui est une opération relativement peu coûteuse par rapport à la factorisation). Les méthodes directes sont les plus précises et sont également robustes : leur efficacité dépend relativement peu des coefficients de la matrice. En revanche, elles sont coûteuses en nombre d'opérations – de l'ordre de  $\mathcal{O}(n^3)$  pour une matrice dense – et exigent une grande quantité de mémoire. De plus, la parallélisation de ces méthodes est délicate.

En pratique, les méthodes directes deviennent rapidement inutilisables lorsque la taille du système augmente. On leur préfère pour cette raison les méthodes dites itératives. Les méthodes itératives consistent à appliquer une procédure un certain nombre de fois sur une solution approchée jusqu'à ce que celle-ci soit suffisamment proche de la solution exacte. Ces méthodes ne donnent pas la solution exacte, mais une approximation de la solution, l'écart entre ces deux grandeurs pouvant être contrôlé à l'aide du critère d'arrêt de la méthode. Il existe une grande variété de méthodes itératives [Saa03], qui sont en général bien plus

économiques en consommation de mémoire (du moins pour les systèmes creux tels que ceux issus de la discrétisation de problèmes physiques), moins coûteuses en nombre d'opérations et facilement parallélisables. L'inconvénient principal de ces méthodes est qu'elles sont bien moins robustes que les méthodes directes : le nombre d'itérations nécessaires à la convergence de ces méthodes dépend fortement des coefficients de la matrice. En réalité, une méthode itérative est toujours utilisée avec un préconditionnement : étant donné une matrice  $M$  inversible, on résout le système  $M^{-1}Ax = M^{-1}b$  dont la solution est la même que celle du système  $Ax = b$ . Tout l'intérêt est de choisir une technique de construction de  $M$  de sorte que la résolution de ce système soit plus rapide ou plus robuste que celle du système original, ce qui est un sujet de recherche en soi. Ce point sera développé dans les sections suivantes. En pratique, il est possible d'appliquer le préconditionnement à la droite de la matrice plutôt qu'à sa gauche : le système effectivement résolu est alors  $AM^{-1}y = b$ , et la solution est déduite selon  $x = M^{-1}y$ . Cette approche est à préférer pour comparer deux préconditionneurs car avec celle-ci, le critère d'arrêt ne dépend pas du préconditionnement.

### 6.1.2 Résolution monolithique ou résolution séquentielle

En prenant du recul par rapport au système linéaire et en se plaçant du point de vue plus éloigné de la méthode numérique étudiée, la discrétisation du problème poroélastique s'achève avec l'obtention du système

$$\begin{bmatrix} \mathcal{A} & -\mathcal{B} \\ \mathcal{B}^T & \mathcal{F} \end{bmatrix} \begin{bmatrix} \mathcal{U} \\ \mathcal{P} \end{bmatrix} = \begin{bmatrix} \mathcal{L}^u \\ \mathcal{L}^p + \mathcal{L}^c \end{bmatrix} \quad (6.1)$$

où l'on cherche à déterminer les déplacements et la pression inconnus au temps  $t^n$  en fonction du second membre, qui fait intervenir ces quantités au temps  $t^{n-1}$  (connues) ainsi que les données du problème. On parlera de résolution *externe* pour désigner le processus permettant de déterminer ces inconnues et donc de passer au pas de temps suivant de la simulation. Différentes stratégies peuvent être adoptées pour la résolution externe, catégorisées en deux familles principales : les stratégies de résolution *monolithique* et les stratégies de résolution *séquentielle*. Il est évidemment souhaitable que toutes ces stratégies produisent au final une même solution à l'issue du temps  $t^n$ , *a minima* à une certaine tolérance près.

#### Résolution monolithique

La stratégie monolithique, aussi appelée *fully coupled*, est la plus naturelle. Elle consiste à voir (6.1) comme un grand système linéaire  $Ax = b$  dont l'inconnue est un vecteur stockant les déplacements et la pression. Cette stratégie ne nécessite qu'une seule itération externe pour valider un pas de temps : la résolution de ce système. Même s'il n'y a dans ce cas pas de confusion possible, on parle de sous-solveur pour désigner la méthode de résolution. Celle-ci peut être l'une ou l'autre des méthodes développées dans la section 6.1.1 : directe, auquel cas il y a une seule itération *interne*, ou itérative, auquel cas le sous-solveur effectue plusieurs itérations internes.

#### Résolution séquentielle

La seconde approche, la stratégie séquentielle, consiste à reformuler le système (6.1) de manière à découpler les deux équations. L'idée est de retomber sur deux sous-problèmes dont les inconnues sont respectivement les déplacements et la pression. Ces deux sous-problèmes sont résolus de manière alternée jusqu'à ce que la solution converge vers celle du système couplé. À chaque pas de temps, le processus de résolution va donc effectuer un certain nombre d'itérations externes, chacune d'entre elles nécessitant la résolution des sous-problèmes. À nouveau, on parle de sous-solveurs pour désigner les méthodes utilisées

	Sous-solveurs directs		Sous-solveurs itératifs		Combinaisons		
Stratégie monolithique	$l = 1$	$s = 1$	$l = 1$	$s > 1$	$\emptyset$		
Stratégie séquentielle	$l > 1$	$s^u = 1$	$s^p = 1$	$l > 1$	$s^u > 1$	$s^p > 1$	$l > 1$ $s^u = 1$ $s^p > 1$ ou $l > 1$ $s^u > 1$ $s^p = 1$

TAB. 6.1 : Nombre d’itérations externes  $l$  et internes  $s$  en fonction du couple stratégie – sous-solveur choisi. Dans le cas de la stratégie séquentielle, le sous-solveur utilisé pour la pression peut être différent du sous-solveur utilisé pour les déplacements.

pour résoudre ceux-ci. Comme dans le cas précédent, cette résolution peut être directe ou itérative, sachant qu’il y a cette fois deux sous-solveurs à choisir (un pour chaque problème). On parlera donc d’itérations internes en déplacements et d’itérations internes en pression. Au passage, il est possible de mélanger les types des sous-solveurs en optant par exemple pour une méthode directe en pression et pour une méthode itérative en déplacement, même si ce choix n’est pas fréquent. Il peut en revanche être intéressant d’utiliser deux méthodes itératives mais d’adapter le solveur ou le préconditionneur au sous-problème considéré. La stratégie séquentielle permet en particulier de réutiliser des codes de calcul existant pour chaque sous-problème et donc de réduire l’effort de développement. Il existe différentes approches pour découpler les deux équations, qui ne présentent pas toutes les mêmes propriétés de stabilité ou de vitesse de convergence (voir par exemple [MW13]). Parmi ces approches, on s’intéressera particulièrement à la méthode *fixed-stress split* qui a l’avantage d’être inconditionnellement stable et qui a récemment fait l’objet de nombreux travaux.

Les différentes combinaisons entre la stratégie de résolution et la méthode employée pour le ou les sous-solveurs sont résumées dans le tableau 6.1.

### 6.1.3 Choix de la stratégie

La plateforme Arcane est interfacée avec la bibliothèque *open source* PETSc<sup>1</sup> qui rassemble différents solveurs linéaires. Ceux-ci sont accessibles via une interface commune, de sorte qu’il suffit d’indiquer dans le jeu de données le solveur à utiliser et ses éventuels paramètres (critère d’arrêt, nombre maximum d’itérations, préconditionneur, etc.).

Dans les cas tests présentés jusqu’à maintenant, la résolution du système linéaire a été réalisée en choisissant l’approche conceptuellement la plus simple : la stratégie monolithique avec résolution directe du système. Cependant, les limites de cette approche apparaissent rapidement, par exemple lors des calculs sur le cas physique tridimensionnel réaliste de la section 5.4. Pour ce système couplé totalisant environ 300 milliers d’inconnues, la résolution par méthode directe est laborieuse avec les deux solveurs testés. Avec le premier solveur direct, SuperLU [Lio5], la résolution échoue ; avec le second solveur direct, MUMPS [ADLKo1], la résolution est un succès mais au prix d’une consommation importante de mémoire vive (plus de 20 Go) et d’un temps de résolution extrêmement long (environ 1h30 sur un cœur cadencé à 3,6 GHz).

Pour accélérer la résolution, le plus simple serait de conserver la stratégie de résolution monolithique tout en utilisant une méthode itérative pour résoudre le système couplé. La matrice n’étant pas symétrique, on utilise la méthode du BiCGStab avec un critère d’arrêt fixé à  $10^{-8}$  et un nombre d’itérations maximal de 3000. Plusieurs préconditionnements sont testés pour ce problème : les méthodes ILU( $k$ ) avec  $k = 0$  et

1. <https://www.mcs.anl.gov/petsc/>

	Block-ILU(o)			Block-ILU(4)			AMG		
	$N_{it}^{int}$	$T_{it}^{ext}$	$T_{it}^{int}$	$N_{it}^{int}$	$T_{it}^{ext}$	$T_{it}^{int}$	$N_{it}^{int}$	$T_{it}^{ext}$	$T_{it}^{int}$
1p	1027	161	0,16	21	564	27			
2p	1716	179	0,10	684	795	1,16	En échec		
4p	1947	137	0,07	1225	615	0,50			
8p	2518	112	0,04	1667	386	0,23			

TAB. 6.2 : Sur le cas géosciences, en fonction du préconditionneur et du nombre de cœurs utilisés : nombre d'itérations  $N_{it}^{int}$  et temps  $T_{it}^{ext}$  (en secondes) nécessaires à la convergence de l'algorithme BiCGStab, et temps passé dans chaque itération interne  $T_{it}^{int} = T_{it}^{ext}/N_{it}^{int}$ . Les valeurs sont des moyennes obtenues sur 10 pas de temps.

$k = 4$  (lorsque  $k$  augmente, la qualité du préconditionnement est meilleure mais le coût de construction du préconditionneur est plus élevé), et la méthode multigrille AMG via son implémentation dans la bibliothèque Hypr<sup>2</sup>. Pour une introduction à ces méthodes de préconditionnement, voir par exemple [Saa03]. Les méthodes sont utilisées avec leurs paramètres par défaut, et le préconditionneur est appliqué à droite de la matrice. Sur cet exemple, la simulation est stoppée à mi-parcours (après 10 pas de temps). Pour chaque pas de temps, le nombre d'itérations externe est de 1 puisque la méthode est monolithique. On relève la durée de résolution du système linéaire, c'est à dire la durée nécessaire à une itération externe  $T_{it}^{ext}$ , le nombre d'itérations internes du BiCGStab  $N_{it}^{int}$  et le temps nécessaire à une itération interne du solveur  $T_{it}^{int} = \frac{T_{it}^{ext}}{N_{it}^{int}}$ . Toutes ces valeurs sont ensuite moyennées sur les 10 pas de temps. Cette procédure est répétée en faisant varier le nombre de cœurs utilisés pour le calcul, et les résultats sont présentés dans le tableau 6.2. On peut observer que les préconditionneurs ILU( $k$ ) ne sont pas stables vis à vis du parallélisme : lorsque le nombre de cœurs augmente, le nombre d'itérations augmente également. En particulier, on observe un écart important entre l'utilisation séquentielle et parallèle, par exemple avec le préconditionneur ILU(4) qui converge en 21 itérations en séquentiel contre presque 700 sur deux cœurs. Même si le temps passé dans une itération interne diminue grâce au parallélisme, la durée d'une itération externe (équivalente ici à la durée nécessaire à la validation d'un pas de temps) peut être plus élevée en parallèle qu'en séquentiel. Pour le préconditionneur AMG, le comportement est encore pire puisqu'en séquentiel comme en parallèle, le solveur atteint la limite des 3000 itérations sans satisfaire le critère d'arrêt, et ce dès le premier pas de temps. Ce cas illustre donc un frein à l'utilisation d'un solveur itératif avec la stratégie monolithique : les préconditionneurs classiques ne sont pas efficaces pour ce problème couplé. Cette limite est en réalité connue, et on compte de nombreux travaux visant à en proposer des nouveaux adaptées à ce système, tels que [BFG07 ; WB11 ; HOL12].

À moins de construire un préconditionneur efficace pour le système couplé, la stratégie monolithique est donc insuffisante. On choisit donc de se tourner vers une stratégie séquentielle, en faisant le pari suivant : certes, pour chaque pas de temps, plusieurs itérations externes seront nécessaires, mais on espère

1. que le nombre d'itérations externes reste raisonnable,
2. que pour chaque sous-problème, le sous-solveur employé soit robuste et rapide,

2. <https://computation.llnl.gov/projects/hypr-scalable-linear-solvers-multigrid-methods>



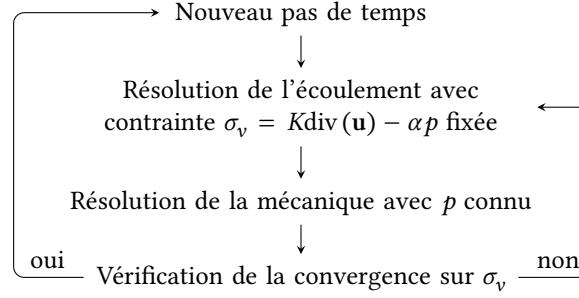


FIG. 6.1 : **Idée générale de la méthode fixed-stress.** Le coefficient  $K$  intervenant dans l'expression de la contrainte  $\sigma_v$  peut être ajusté de différentes manières.

de manière à être compétitif par rapport à la méthode monolithique, en particulier lors de simulations sur plusieurs cœurs.

## 6.2 Stratégie séquentielle fixed-stress

Pour mettre en place une stratégie séquentielle, on considère la méthode appelée fixed-stress. Introduite dans sa formulation historique puis réinterprétée de manière à faire le lien avec l'algorithme de point fixe sous-jacent, différentes variantes de la stratégie fixed-stress sont décrites et comparées dans cette section.

### 6.2.1 Formulation de l'algorithme fixed-stress

Historiquement, cette manière de découpler les deux équations est motivée par la physique du problème [SM94]. L'idée générale est décrite par l'organigramme présenté sur la figure 6.1, en utilisant ici les variables et opérateurs continus. À chaque pas de temps, on alterne entre la résolution d'un problème d'écoulement en supposant qu'une certaine quantité est fixée (cette quantité est assimilable à une contrainte, d'où le nom de l'algorithme), et la résolution d'un problème de mécanique en prenant pour la pression la valeur issue de l'étape précédente. Le paramètre  $K$  intervenant dans la définition de la contrainte  $\sigma_v$  peut être ajusté : d'abord choisi de manière à rester au plus proche de la physique, d'autres valeurs ont ensuite été proposées dans le but d'accélérer la convergence de la méthode. Pour une comparaison des différentes valeurs admissibles, voir par exemple [BK19]. Dans ce chapitre, on se contentera du choix guidé par la physique en prenant  $K = \lambda$ , ce paramètre ayant été défini dans la section 2.1.

Avec la méthode fixed-stress, les matrices et seconds membres sont tout de même modifiés par rapport à ceux qui seraient issus de problèmes initialement non couplés. Habituellement (voir par exemple [GWAD19]), on fait apparaître dans l'équation d'écoulement la contrainte moyenne  $\sigma_v = \lambda \operatorname{div}(\mathbf{u}) - \alpha p$  : formellement, en repartant de l'équation semi-discrète,

$$\left(c_0 + \frac{\alpha^2}{\lambda}\right) p^n + \frac{\alpha}{\lambda} \underbrace{(\lambda \operatorname{div}(\mathbf{u}^n) - \alpha p^n)}_{\sigma_v^n} + \Delta t \operatorname{div}(-\bar{\kappa} \nabla p^n) = \Delta t q^n + c_0 p^{n-1} + \alpha \operatorname{div}(\mathbf{u}^{n-1}).$$

Le système est ensuite résolu par une méthode de point fixe en plaçant  $\sigma_v^n$  au membre de droite : en notant

l les itérations de point fixe, l'équation d'écoulement devient

$$\left(c_0 + \frac{\alpha^2}{\lambda}\right) p^{n,l} + \Delta t \operatorname{div}(-\bar{\kappa} \nabla p^{n,l}) = \Delta t q^n + c_0 p^{n-1} + \alpha \operatorname{div}(\mathbf{u}^{n-1}) - \frac{\alpha}{\lambda} \underbrace{(\lambda \operatorname{div}(\mathbf{u}^{n,l-1}) - \alpha p^{n,l-1})}_{\sigma_v^{n,l-1}}. \quad (6.2)$$

Ainsi, seule la grandeur  $p^{n,l}$  est une inconnue dans l'équation d'écoulement. Pour l'équation de mécanique, il est clair le terme  $p^{n,l}$  venant d'être déterminé, seuls les déplacements sont à calculer :

$$-\mathbf{div}(\bar{C} \bar{\epsilon}_{u^{n,l}}) = \mathbf{f} - \alpha \nabla p^{n,l}. \quad (6.3)$$

On propose ici une manière différente d'aboutir aux sous-systèmes, qui se base sur la formulation matricielle du système original couplé (6.1) et qui sera utile pour formuler différentes variantes de la méthode fixed-stress. Cette approche consiste à reformuler le système linéaire en introduisant l'inconnue supplémentaire discrète  $\mathfrak{S}_v$  (représentant la contrainte  $\sigma_v$ ) et l'équation qui lui est associée : puisque  $\mathcal{B}^T(\phi_k, \boldsymbol{\varphi}_j^l) = \alpha_K \int_K \operatorname{div}(\boldsymbol{\varphi}_j^l)$  d'après (4.38), à chaque pas de temps,

$$\mathfrak{S}_v = \frac{\lambda_K}{\alpha_K |K|} \mathcal{B}^T \mathcal{U} - \alpha_K \mathcal{P}. \quad (6.4)$$

L'abus de notation  $\cdot_K$  dans cette équation fait référence au fait que la ligne  $k$  des matrices, qui correspond à la cellule  $K$ , est multipliée par les coefficients réels  $\cdot_K$  dans le cas où ceux-ci ne sont pas constants sur l'ensemble du domaine. La notation  $\mathcal{I}$  désigne la matrice identité de taille  $\operatorname{card}(\mathfrak{M})$ . Ajouter cette équation au système couplé permet de reformuler la seconde ligne selon

$$\mathcal{L}^p + \mathcal{L}^c = \mathcal{B}^T \mathcal{U} + \mathcal{F} \mathcal{P} = \frac{\alpha_K |K|}{\lambda_K} \mathfrak{S}_v + \left( \mathcal{F} + \frac{\alpha_K^2 |K|}{\lambda_K} \mathcal{I} \right) \mathcal{P}$$

et donc d'obtenir le système issu des trois équations

$$\begin{bmatrix} \mathcal{I} & -\frac{\lambda_K}{\alpha_K |K|} \mathcal{B}^T & \alpha_K \mathcal{I} \\ 0 & \mathcal{A} & -\mathcal{B} \\ \frac{\alpha_K |K|}{\lambda_K} \mathcal{I} & 0 & \tilde{\mathcal{F}} \end{bmatrix} \begin{bmatrix} \mathfrak{S}_v \\ \mathcal{U} \\ \mathcal{P} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{L}^u \\ \mathcal{L}^p + \mathcal{L}^c \end{bmatrix} \quad \text{où} \quad \tilde{\mathcal{F}} = \mathcal{F} + \frac{\alpha_K^2 |K|}{\lambda_K} \mathcal{I}. \quad (6.5)$$

La technique du complément de Schur permet alors de définir deux formulations différentes de l'algorithme fixed-stress. En effet, (6.5) est de la forme

$$\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ G_2 \end{bmatrix}$$

avec

$$Q_{11} = \mathcal{I}, \quad Q_{12} = \left[ -\frac{\lambda_K}{\alpha_K |K|} \mathcal{B}^T \quad \alpha_K \mathcal{I} \right], \quad Q_{21} = \begin{bmatrix} 0 \\ \frac{\alpha_K |K|}{\lambda_K} \mathcal{I} \end{bmatrix}, \quad Q_{22} = \begin{bmatrix} \mathcal{A} & -\mathcal{B} \\ 0 & \tilde{\mathcal{F}} \end{bmatrix}, \quad X_1 = \mathfrak{S}_v \quad \text{et} \quad X_2 = \begin{bmatrix} \mathcal{U} \\ \mathcal{P} \end{bmatrix},$$

où  $Q_{11}$  et  $Q_{22}$  sont inversibles, et peut donc être réécrit sous la forme du système

$$\begin{aligned} X_1 &= -Q_{12} X_2 \\ X_2 &= Q_{22}^{-1} (G_2 - Q_{21} X_1). \end{aligned} \quad (6.6)$$

Éliminer l'inconnue  $X_1$  dans (6.6) revient à travailler sur un problème dont le vecteur d'inconnues est  $X_2 = [\mathcal{U}, \mathcal{P}]^T$  :

$$X_2 = Q_{22}^{-1} (G_2 + Q_{21}Q_{12}X_2) := C(X_2). \quad (6.7)$$

Inversement, éliminer l'inconnue  $X_2$  dans (6.6) revient à travailler sur un problème dont le vecteur d'inconnues est  $X_1 = [\mathcal{S}_v]$  :

$$X_1 = -Q_{12}Q_{22}^{-1} (G_2 - Q_{21}X_1) := \tilde{C}(X_1). \quad (6.8)$$

Il est donc possible de travailler sur l'une ou l'autre des deux variables pour les itérations externes de la méthode fixed-stress. Néanmoins, quelque soit l'alternative choisie, l'étape au centre de l'algorithme consiste à résoudre des systèmes linéaires dont la matrice associée est  $Q_{22}$ .

### 6.2.2 Interprétation de la stratégie fixed-stress comme une méthode de point fixe

La méthode classique pour progresser dans la résolution externe consiste à appliquer une méthode de point fixe : selon la variable de travail choisie, pour une valeur initiale  $X_1^0$  – resp.  $X_2^0$  donnée, l'opération  $X_1^l = \tilde{C}(X_1^{l-1})$  – resp.  $X_2^l = C(X_2^{l-1})$  est appliquée jusqu'à convergence. Cette méthode est détaillée dans l'algorithme 3. Dans le cas où la variable de travail est  $X_2 = [\mathcal{U} \ \mathcal{P}]^T$ , cette algorithme est équivalent à la

---

#### Algorithm 3 Stratégie fixed-stress vue comme une méthode de point fixe

---

<ol style="list-style-type: none"> <li>1 Choose initial unknown <math>X_1^0 = \mathcal{S}_v^0</math></li> <li>2 Compute initial residual <math>r_0</math></li> <li>3 <b>repeat</b></li> <li>4     <math>X_1^l = \tilde{C}(X_1^{l-1})</math> :</li> <li>5     Compute <math>r</math></li> <li>6     <math>X_1^{l-1} = X_1^l</math></li> <li>7 <b>until</b> <math>r \leq r_0 \epsilon^{\text{ext}}</math> or <math>l &gt; N_{\text{max}}</math></li> <li>8 <b>function</b> <math>\tilde{C}(X)</math></li> <li>9     <math>X_1 = X</math></li> <li>10    <math>X_2 = Q_{22}^{-1}(G_2 - Q_{21}X_1)</math></li> <li>11    <math>X_1 = -Q_{12}X_2</math></li> </ol>	<ol style="list-style-type: none"> <li>1 Choose initial unknown <math>X_2^0 = [\mathcal{U}^0 \ \mathcal{P}^0]^T</math></li> <li>2 Compute initial residual <math>r_0</math></li> <li>3 <b>repeat</b></li> <li>4     <math>X_2^l = C(X_2^{l-1})</math> :</li> <li>5     Compute <math>r</math></li> <li>6     <math>X_2^{l-1} = X_2^l</math></li> <li>7 <b>until</b> <math>r \leq r_0 \epsilon^{\text{ext}}</math> or <math>l &gt; N_{\text{max}}</math></li> <li>8 <b>function</b> <math>C(X)</math></li> <li>9     <math>X_2 = X</math></li> <li>10    <math>X_1 = -Q_{12}X_2</math></li> <li>11    <math>X_2 = Q_{22}^{-1}(G_2 - Q_{21}X_1)</math></li> </ol>
---	---

---

formulation historique de la méthode fixed-stress : en effet, (6.7) se réécrit

$$\begin{bmatrix} \mathcal{U}^l \\ \mathcal{P}^l \end{bmatrix} = \begin{bmatrix} \mathcal{A} & -\mathcal{B} \\ 0 & \tilde{\mathcal{F}} \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{L}^u \\ \mathcal{L}^p + \mathcal{L}^c + \mathcal{L}^{\sigma, l-1} \end{bmatrix} \quad \text{avec} \quad \mathcal{L}^{\sigma, l-1} = -\mathcal{B}^T \mathcal{U}^{l-1} + \frac{\alpha_K^2 |K|}{\lambda_K} \mathcal{P}^{l-1},$$

et n'est autre que la formulation matricielle issue de la discrétisation des équations (6.2)–(6.3). Finalement, par rapport à un problème fluide indépendant, l'équation en pression est modifiée en ajoutant au second membre la quantité  $\mathcal{L}^c + \mathcal{L}^\sigma$  (calculée à partir de l'itération fixed-stress précédente pour  $\mathcal{L}^\sigma$  et du pas de temps précédent pour  $\mathcal{L}^c$ ), et à la matrice  $\mathcal{F}$  la quantité  $\frac{\alpha_K^2 |K|}{\lambda_K} \mathcal{I}$ . L'équation de mécanique est résolue dans un second temps en utilisant la pression qui vient d'être calculée :  $\mathcal{A} \mathcal{U}^l = \mathcal{L}^u + \mathcal{B} \mathcal{P}^l$ .

Dans le cas où la variable de travail est  $X_1 = \mathcal{S}_v$ , le lien avec la méthode fixed-stress historique est soumis à condition : (6.7) se réécrit

$$\mathcal{S}_v^l = \begin{bmatrix} \frac{\lambda_K}{\alpha_K |K|} \mathcal{B}^T & -\alpha_K \mathcal{I} \end{bmatrix} \begin{bmatrix} \mathcal{A} & -\mathcal{B} \\ 0 & \tilde{\mathcal{F}} \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{L}^u \\ \mathcal{L}^p + \mathcal{L}^c + \mathcal{L}^{\sigma, l-1} \end{bmatrix} \quad \text{avec} \quad \mathcal{L}^{\sigma, l-1} = -\frac{\alpha_K |K|}{\lambda_K} \mathcal{S}_v^{l-1}.$$

La résolution du système linéaire donne donc deux variables temporaires  $\mathcal{U}^l$  et  $\mathcal{P}^l$ , correspondant à  $X_2$  dans la partie gauche de l'algorithme 3 (l 10), et utilisées pour mettre à jour  $\mathfrak{S}_v^l$ . *A priori*, les deux versions de la méthode fixed-stress produisent des grandeurs itérées différentes à chaque itération ; cependant, dans le cas particulier où

$$\mathfrak{S}_v^0 = -Q_{12} [\mathcal{U}^0 \quad \mathcal{P}^0]^T, \quad (6.9)$$

et donc où les variables sont compatibles à l'instant initial, les deux versions de l'algorithme produisent en réalité à chaque itération les mêmes grandeurs  $X_1^l$  et  $X_2^l$ . Pour montrer cette propriété, on utilise la définition des fonctions  $\tilde{C}$  et  $C$  dans l'algorithme 3 (l9-l13) en notant  $\tilde{X}_1$  et  $\tilde{X}_2$  les grandeurs de la version itérant sur  $X_1$  pour les différencier.

La démonstration s'effectue par récurrence : supposons qu'à l'instant  $l-1$ ,  $\tilde{X}_1^{l-1} = -Q_{12}X_2^{l-1}$ . Alors

$$\tilde{X}_2^l = Q_{22}^{-1}(G_2 - Q_{21}\tilde{X}_1^{l-1}) = Q_{22}^{-1}(G_2 + Q_{21}Q_{12}X_2^{l-1}) = Q_{22}^{-1}(G_2 - Q_{21}X_1^l) = X_2^l$$

puis

$$\tilde{X}_1^l = -Q_{12}\tilde{X}_2^l = -Q_{12}X_2^l = X_1^l.$$

En tenant compte de l'hypothèse initiale (6.9)  $\tilde{X}_1^0 = -Q_{12}X_2^0$ , la seconde égalité établit la récurrence et montre que à chaque itération  $l$ ,  $\tilde{X}_1^l = X_1^l$  et  $\tilde{X}_2^l = X_2^l$ . Pour conserver l'équivalence de ces deux variantes, on prendra donc comme valeur initiale la grandeur issue du pas de temps précédent :

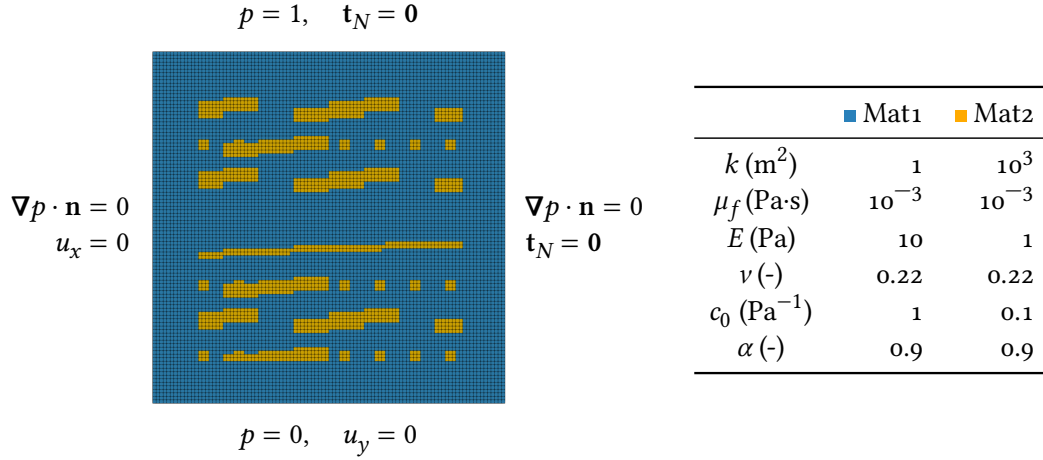
$$\begin{aligned} X_2^{n,0} &= [\mathcal{U}^{n-1} \quad \mathcal{P}^{n-1}]^T && \text{pour la formulation en } X_2 \\ X_1^{n,0} &= \mathfrak{S}_v^{n-1} = -Q_{12}[\mathcal{U}^{n-1} \quad \mathcal{P}^{n-1}]^T && \text{pour la formulation en } X_1. \end{aligned}$$

Pour achever la description de la méthode, il reste à choisir la méthode à utiliser pour calculer les résidus  $r_0$  et  $r$ . Le choix standard pour un algorithme de point fixe consiste à calculer la différence entre deux itérations successives :

$$\begin{aligned} r(X_2) &= \left\| C \left( \begin{bmatrix} \mathcal{U}^{l-1} \\ \mathcal{P}^{l-1} \end{bmatrix} \right) - \begin{bmatrix} \mathcal{U}^{l-1} \\ \mathcal{P}^{l-1} \end{bmatrix} \right\|_{L^2} && \text{et } r_0 = \|C(0)\|_{L^2} \text{ pour la formulation en } X_2 \\ r(X_1) &= \|\tilde{C}(\mathfrak{S}_v^{l-1}) - \mathfrak{S}_v^{l-1}\|_{L^2} && \text{et } r_0 = \|\tilde{C}(0)\|_{L^2} \text{ pour la formulation en } X_1. \end{aligned}$$

Dans notre cas, lorsque la variable de travail utilisée est  $X_2$ , d'autres possibilités sont offertes. Il est par exemple possible avant et après l'itération de calculer les contraintes  $\mathfrak{S}_v^{l-1}$  et  $\mathfrak{S}_v^l$  selon la relation (6.4). On peut alors calculer le résidu sur la différence entre la variable de contrainte plutôt que sur celle entre les variables primaires selon  $r = \|\mathfrak{S}_v^l - \mathfrak{S}_v^{l-1}\|_{L^2}$ . À nouveau, on normalise en appliquant la méthode au vecteur nul :  $r_0 = \|\tilde{C}(0)\|_{L^2}$ . Calculer le résidu sur la contrainte alors que les itérations sont effectuées sur les variables primaires est une approche parfois adoptée dans la littérature, et qui trouve ici sa justification dans la formulation en  $X_1 = \mathfrak{S}_v$  de la méthode. Une troisième possibilité, pouvant être utile à des fins de tests, consiste à assembler la matrice monolithique  $M^{\text{mono}}$  et le second membre monolithique  $\text{rhs}^{\text{mono}}$  même en cas de résolution séquentielle, et à utiliser ces quantités pour calculer un résidu selon  $r = \|M^{\text{mono}} [\mathcal{U}^l \quad \mathcal{P}^l]^T - \text{rhs}^{\text{mono}}\|_{L^2}$ . Dans ce cas, la quantité utilisée pour normaliser est calculée avec la même formule appliquée au vecteur nul de sorte que  $r_0 = \|\text{rhs}^{\text{mono}}\|_{L^2}$ .

Lorsque la variable de travail utilisée est  $X_1$ , il n'y a pas vraiment d'autre moyen de définir un résidu en dehors de celui propre à la méthode de point fixe défini plus haut. En effet, on ne sait pas remonter aux vecteurs  $\mathcal{U}$  et  $\mathcal{P}$  à partir d'un vecteur  $\mathfrak{S}_v$  donné. Il est notamment impossible de réutiliser le système monolithique comme dans le cas précédent.


 FIG. 6.2 : **Domaine hétérogène bidimensionnel** et caractéristiques physiques des deux matériaux.

### 6.2.3 Comparaison des différentes variantes

Pour une première illustration de l'algorithme fixed-stress, de l'influence du choix de la variable inconnue et de celui du critère d'arrêt, on considère un nouveau cas test bidimensionnel hétérogène inspiré de [BV16] et [CKHT19]. Le carré unité, domaine de calcul, est constitué de deux matériaux et est représenté sur la figure 6.2. Celui-ci est discrétisé par  $400 \times 400$  mailles carrées (pour plus de lisibilité, la figure ne représente qu'un maillage de taille  $100 \times 100$ ). Les constantes mécaniques des matériaux sont également indiquées sur la figure. À l'instant initial  $t = 0$ , la pression est nulle. L'évolution de ce problème est due à ses conditions aux limites, représentées sur la figure. Le système est résolu au bout d'un unique pas de temps  $\Delta t = T = 5$  secondes.

Le nombre d'inconnues est suffisamment faible pour résoudre les sous-problèmes de manière directe. La tolérance de l'algorithme fixed-stress est pour sa part fixée à  $10^{-6}$ . Dans un premier temps, on compare les trois manières de calculer le résidu (et donc le critère d'arrêt) en utilisant la formulation en variables  $(\mathbf{u}, p)$ . L'évolution du résidu en fonction du nombre d'itérations externes est représentée à gauche de la figure 6.3. Les deux méthodes qui calculent la norme  $L^2$  de la différence entre deux itérations fixed-stress successives donnent un résultat proche : l'algorithme converge en 13 itérations lorsque la différence est directement calculée sur  $(\mathbf{u}, p)$  et en 14 itérations lorsqu'on déduit  $\sigma_v$  à partir des variables primaires pour calculer la différence. La méthode faisant appel au système monolithique donne des résultats différents, et semble converger plus rapidement. Cela dit, la comparaison est difficile puisque dans ce dernier cas la valeur du résidu est affectée par la matrice  $M^{\text{mono}}$ , et sa normalisation est différente : on observe sur le graphe un écart de deux ordres de grandeurs. Pour ce cas, on dispose d'une solution de référence calculée avec la stratégie monolithique-résolution directe. Il est donc possible de vérifier en fin de calcul l'écart sur les variables  $\mathbf{u}$  et  $p$  entre la solution de référence et l'approche séquentielle. Pour les deux premières manières de calculer le résidu, l'écart est de l'ordre de  $10^{-7}$  alors qu'il est de l'ordre de  $10^{-5}$  avec la troisième. Dans ce dernier cas, la méthode s'est donc stoppée plus tôt (d'où le nombre d'itérations plus faible) mais a atteint une solution moins précise. Quoi qu'il en soit, cette troisième méthode n'est pas réellement utilisable dans une application réelle puisqu'elle demande l'assemblage du système monolithique en plus des sous-systèmes et est donc plus coûteuse. Dans un second temps, on compare le choix de la variable utilisée en tant que quantité à faire converger par la méthode de point fixe. Le résidu est calculé selon la première méthode, et son évolution en fonction du nombre d'itérations externes est représentée à droite de la figure

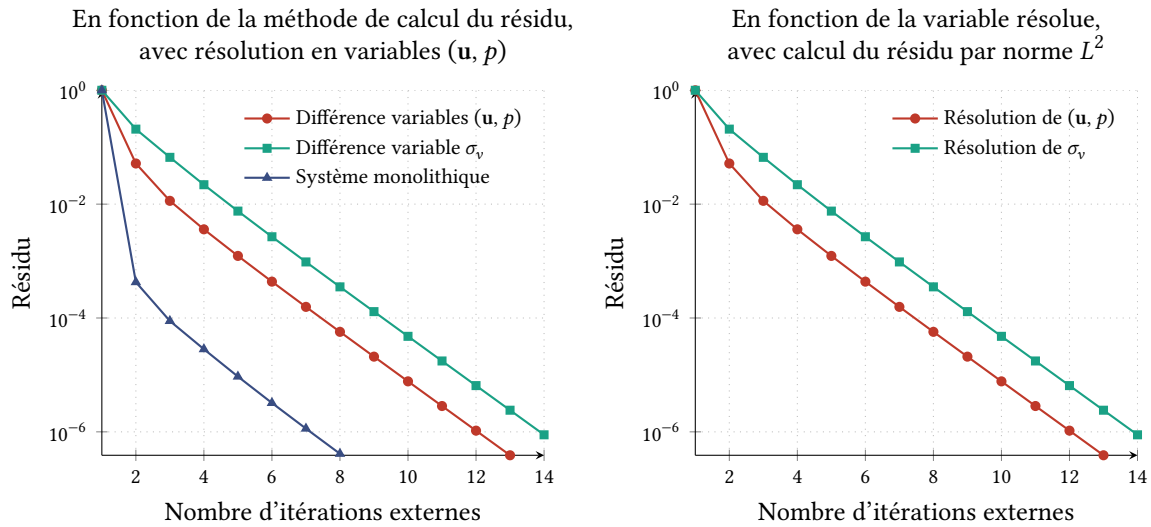


FIG. 6.3 : **Comportement de différentes variantes de l'algorithme fixed-stress** implémenté en tant que méthode de point fixe sur le cas hétérogène 2D.

6.3. Le nombre d'itérations des deux formulations est presque identique. On retrouve également le fait que l'utilisation de la variable  $\sigma_v$  avec son calcul de résidu standard (courbe verte à droite) donne les mêmes résultats que l'utilisation des variables  $(\mathbf{u}, p)$  avec calcul de résidu sur  $\sigma_v$  (courbe verte à gauche), ceci étant dû à l'équivalence des deux formulations. Compte tenu de la proximité des deux approches, il peut être préférable de privilégier la formulation en variable  $\sigma_v$ , qui exécute l'algorithme de point fixe sur un vecteur de taille plus petite (une inconnue par cellule contre une inconnue par cellule plus  $d$  inconnues par nœud pour la formulation  $(\mathbf{u}, p)$ ). Il faut toutefois être conscient que la charge de calcul est majoritairement portée par la résolution des sous-problèmes, et donc ne pas attendre de ce choix une amélioration significative des performances.

Pour achever cette section, on relève la durée de simulation de la stratégie séquentielle : pour ses 14 itérations, la méthode de point fixe a consommé 260 secondes, soit moins de 20 secondes pour chaque itération de point fixe. À titre de comparaison, l'approche monolithique nécessite sur ce cas 80 secondes pour résoudre le système couplé. Résoudre deux "petits" systèmes découplés est donc plus rapide que résoudre un "grand" système couplé. La méthode fixed-stress est au final plus lente en raison de son nombre important d'itérations, qui dépend d'ailleurs de la précision exigée. Même si la robustesse reste l'objectif prioritairement recherché lors de la mise en place de l'approche séquentielle (voir le paragraphe 6.1.3), il est toujours souhaitable d'améliorer la vitesse de la méthode.

### 6.3 Accélération de la stratégie fixed-stress

Une des pistes envisageables pour réduire le temps de calcul de la stratégie fixed-stress consiste à faire diminuer le nombre d'itérations de celle-ci. Il s'agit de l'objet de cette section, où l'algorithme est réécrit dans le formalisme des méthodes de Krylov.

### 6.3.1 Interprétation de la stratégie fixed-stress comme une méthode de Krylov

Dans la section précédente, la stratégie fixed-stress a été présentée comme un algorithme visant à résoudre le problème  $C(X) = X$ , avec  $X$  désignant la variable de travail choisie, cf paragraphe 6.2.1. La méthode alors adoptée pour résoudre ce problème, en accord avec la présentation classique de l'algorithme fixed-stress, était une méthode de point fixe (voir l'algorithme 3). Or, on peut reformuler le même problème comme la résolution du problème linéaire  $\mathcal{R}(X) = 0$  avec  $\mathcal{R}(X) = C(X) - X$ , et utiliser une méthode de Krylov pour en calculer la solution.

Dans les ouvrages, les méthodes de Krylov sont souvent formulées dans le but de résoudre un système linéaire  $Ay = b$  et font donc intervenir la matrice  $A$  et le second membre  $b$ . Une première étape consiste donc à réécrire ces algorithmes dans leur version "matrix-vector product free", c'est à dire en ne faisant intervenir que la fonction résidu  $\mathcal{R}$  (alors définie par  $\mathcal{R}(y) = b - Ay$  dans le cas de ce système). Pour la méthode du BiCGStab, ce travail est effectué dans l'algorithme 4. Dans celui ci, la notation  $(u, v)$  désigne le produit scalaire discret de  $L^2$ . Dans cet algorithme, il existe un lien entre le résidu  $r^k$  utilisé dans le test

---

**Algorithm 4** BiCGStab dans sa version originale (gauche) et matrix-vector product free (droite)

---

<ol style="list-style-type: none"> <li>1 <math>X^0</math> given</li> <li>2 <math>r^0 = b - AX^0</math></li> <li>3 <math>\hat{r}^0 = p^0 = r^0</math></li> <li>4 <b>repeat</b></li> <li>5     <math>\alpha = (r^k, \hat{r}^0) / (Ap^k, \hat{r}^0)</math></li> <li>6     <math>s^k = r^k - \alpha Ap^k</math></li> <li>7     <math>\omega = \frac{(As^k, s^k)}{\ As^k\ ^2}</math></li> <li>8     <math>X^k \leftarrow X^k + \alpha p^k + \omega s^k</math></li> <li>9     <math>r^{k+1} = s^k - \omega As^k</math></li> <li>10    <math>\beta = \alpha(r^{k+1}, \hat{r}^0) / \omega(r^k, \hat{r}^0)</math></li> <li>11    <math>p^k \leftarrow r^{k+1} + \beta(p^k - \omega Ap^k)</math></li> <li>12 <b>until</b> <math>\ r^{k+1}\  \leq \epsilon^{\text{ext}} \ r_0\ </math> or <math>k &gt; k_{\text{max}}</math></li> </ol>	<ol style="list-style-type: none"> <li>1 <math>X^0</math> given</li> <li>2 <math>r^0 = \mathcal{R}(X^0); b = \mathcal{R}(0)</math></li> <li>3 <math>\hat{r}^0 = p^0 = r^0</math></li> <li>4 <b>repeat</b></li> <li>5     <math>\alpha = (r^k, \hat{r}^0) / (b - \mathcal{R}(p^k), \hat{r}^0)</math></li> <li>6     <math>s^k = r^k - \alpha(b - \mathcal{R}(p^k))</math></li> <li>7     <math>\omega = \frac{(b - \mathcal{R}(s^k), s^k)}{\ b - \mathcal{R}(s^k)\ ^2}</math></li> <li>8     <math>X^k \leftarrow X^k + \alpha p^k + \omega s^k</math></li> <li>9     <math>r^{k+1} = s^k - \omega(b - \mathcal{R}(s^k))</math></li> <li>10    <math>\beta = \alpha(r^{k+1}, \hat{r}^0) / \omega(r^k, \hat{r}^0)</math></li> <li>11    <math>p^k \leftarrow r^{k+1} + \beta(p^k - \omega(b - \mathcal{R}(p^k)))</math></li> <li>12 <b>until</b> <math>\ r^{k+1}\  \leq \epsilon^{\text{ext}} \ r_0\ </math> or <math>k &gt; k_{\text{max}}</math></li> </ol>
--	---

---

d'arrêt et l'opérateur résidu  $\mathcal{R}$  utilisé dans la boucle retranscrit par la proposition suivante :

**Proposition 6.1** Si l'opérateur  $\mathcal{R}$  est affine ie  $\forall X, \mathcal{R}(X) = \mathcal{R}(0) + \mathcal{R}^{\text{lin}}(X)$  avec  $\mathcal{R}^{\text{lin}}$  linéaire, alors

$$r^k = \mathcal{R}(X^k)$$

dans l'algorithme du BiCGStab.

La démonstration s'effectue par récurrence : par définition,  $r^0 = \mathcal{R}(X^0)$ . En supposant que la proposition est vrai au rang  $k$ ,

$$\begin{aligned}
 r^{k+1} &= r^k - \alpha(\mathcal{R}(0) - \mathcal{R}(p^k)) - \omega(\mathcal{R}(0) - \mathcal{R}(s^k)) \\
 &= \mathcal{R}(X^k) - \alpha(\mathcal{R}(0) - \mathcal{R}(p^k)) - \omega(\mathcal{R}(0) - \mathcal{R}(s^k)) \\
 &= \mathcal{R}(0) + \mathcal{R}^{\text{lin}}(X^k) + \alpha\mathcal{R}^{\text{lin}}(p^k) + \omega\mathcal{R}^{\text{lin}}(s^k) \\
 &= \mathcal{R}(0) + \mathcal{R}^{\text{lin}}(X^{k+1}) = \mathcal{R}(X^{k+1}).
 \end{aligned}$$

Ce lien permet d'assurer que lorsque le critère d'arrêt de la méthode de Krylov est satisfait, le problème  $\mathcal{R}(X) = 0$  est bien résolu (à la tolérance près).

De la même manière, la méthode Gmres [Saa03, section 6.5], autre méthode de Krylov, est implémentée dans sa version sans produit matrice vecteur afin de l'utiliser pour résoudre le problème de minimisation du résidu fixed-stress. Cette procédure est décrite dans l'algorithme 5. Par rapport au BiCGStab, cette méthode

---

**Algorithm 5** Algorithme Gmres dans sa version matrix-vector product free
 

---

```

1  $X^0$  given
2  $r = \mathcal{R}(X^0)$ ;  $b = r^0 = \mathcal{R}(0)$ ;
3  $\beta, c$  and  $s$  lists of size  $k_{\max} + 1$  ▷ Initialise empty arrays
4  $V$  array of  $k_{\max}$  vectors sized as  $X$ 
5  $\beta_0 = \|r\|_{L^2}$ ;  $V^0 = r/\|r\|_{L^2}$ 
6 repeat
7    $\omega = b - \mathcal{R}(V^k)$ 
8   for  $i = 0, \dots, k$  do ▷ Build an orthonormalized basis
9      $H_{ik} = (\omega, V^i)$ 
10     $\omega -= H_{ik}V^i$ 
11    $H_{k+1,k} = \|\omega\|_{L^2}$ 
12   if  $\|\omega\|_{L^2} \neq 0$  then  $V^{k+1} = \omega/\|\omega\|_{L^2}$ 
13   for  $i = 0, \dots, k - 1$  do ▷ Apply Given's rotations to triangulate  $H$ 
14      $H_{ik} = c_i H_{ik} + s_i H_{i+1,k}$ 
15      $H_{i+1,k} = -s_i H_{ik} + c_i H_{i+1,k}$ 
16      $c_k = \frac{H_{kk}}{\sqrt{H_{kk}^2 + H_{k+1,k}^2}}$  and  $s_k = \frac{H_{k+1,k}}{\sqrt{H_{kk}^2 + H_{k+1,k}^2}}$ 
17      $H_{kk} = c_k H_{kk} + s_k H_{k+1,k}$ 
18      $H_{k+1,k} = 0$ 
19      $\beta_{k+1} = -s_k \beta_k$ 
20      $\beta_k = c_k \beta_k$ 
21      $\|r\| = |\beta_{k+1}|$  ▷ Get the residual norm for free !
22   until  $\|r\| \leq \epsilon^{\text{ext}} \|r^0\|$  or  $k > k_{\max}$ 
23   Solve upper triangular system  $Hy = \beta$  ▷ Gmres is converged, now let's recover  $X$ 
24    $X = X^0 + \sum_{i=0}^k y_i V^i$ 

```

---

consomme plus mémoire puisqu'à l'itération  $k$ , elle requiert tous les vecteurs  $V^i, i < k$  pour construire la base orthonormalisée (18) et doit donc conserver ces vecteurs de taille égale à celle de  $X$ . En revanche, elle a l'avantage de n'appeler qu'une seule fois la procédure de calcul du résidu  $\mathcal{R}$  (17). Cette avantage est dû à l'utilisation des rotations de Given (113) pour résoudre le problème des moindres carrés  $Hy = \beta$ , qui permet de connaître la norme du résidu  $\|r\|_{L^2}$  gratuitement *avant* d'avoir calculé la solution  $X$  (118). Une version moins sophistiquée consisterait à résoudre à chaque itération le problème des moindres carrés, déterminer  $X^k$  puis appeler  $\mathcal{R}(X^k) = r$  pour vérifier si l'algorithme a convergé ou non, mais cette solution nécessite deux appels à  $\mathcal{R}$  par itération.

Dans le cas de la stratégie séquentielle fixed-stress, il suffit donc d'appliquer la version sans produit matrice vecteur de l'algorithme BiCGStab ou de l'algorithme Gmres en utilisant la procédure  $C$  définie par (6.7) ou (6.8) selon le choix de la variable  $X$  lors des appels à la fonction résidu  $\mathcal{R}(X) = C(X) - X$ . Comme dans la méthode du point fixe, le vecteur initial  $X^0$  est calculé à partir des inconnues au pas de temps précédent et, lors du test d'arrêt, la valeur  $r_0$  utilisée pour normaliser correspond à  $\mathcal{R}(0)$ . Compte



tenu des besoins en mémoire plus importants de Gmres, mais directement proportionnels à la taille du vecteur inconnu, il est préférable d'utiliser la formulation en variable  $\sigma_v$  qui est de taille bien inférieure au vecteur d'inconnues  $(\mathbf{u}, p)$  puisque les deux formulations sont équivalentes d'après la section 6.2. Enfin, on note que lorsque la méthode séquentielle s'arrête, le dernier appel à la procédure  $C$  a été effectué avec une variable temporaire :  $p^k$  dans le cas du BiCGStab et  $V^k$  dans le cas de Gmres. Dans l'implémentation effectuée, il faut rappeler  $C$  sur le dernier vecteur  $X^k$  pour mettre à jour correctement les déplacements et la pression.

### 6.3.2 Comparaison numérique des algorithmes

Les stratégies fixed-stress en tant que méthode de Krylov ou en tant qu'algorithme de point fixe sont comparées dans les conditions du cas test de la section 6.2.3. Pour les trois approches, la variable de calcul choisie est  $X = \zeta_v$  et le résidu est calculé par l'approche naturelle à chaque méthode. À gauche de la figure 6.4, l'évolution du résidu est représenté au fil des itérations de la méthode fixed-stress jusqu'à satisfaire le critère de convergence avec une tolérance  $\epsilon^{\text{ext}}$  fixée à  $10^{-6}$ . Contrairement à la méthode de point fixe, dont la convergence est linéaire, les méthodes de Krylov atteignent plus rapidement le critère de convergence : 5 itérations pour le BiCGStab et 8 pour Gmres. Le temps moyen passé dans chaque itération est de 19 secondes pour la méthode de point fixe, 38 secondes pour l'algorithme du BiCGStab et 21 secondes pour la méthode Gmres. Comme attendu, la méthode du BiCGStab est pénalisée par le fait qu'elle appelle deux fois la résolution des sous-systèmes, partie la plus coûteuse, à chaque itération ; si bien qu'au final, la durée totale de la méthode fixed-stress est moins longue avec Gmres qu'avec le BiCGStab.

Les écarts entre la méthode de point fixe et les méthodes de Krylov se creusent lorsque la tolérance  $\epsilon^{\text{ext}}$  est abaissée, comme illustré par le graphique de droite sur la figure 6.4 qui représente le nombre d'itérations des trois méthodes en fonction de la tolérance. Si l'on tient compte du fait que la méthode du BiCGStab appelle deux fois la fonction résidu pour chaque itération (courbe en pointillés verts), la méthode Gmres semble être la plus économe. L'écart entre les deux méthodes de Krylov reste toutefois faible par rapport à l'écart entre une méthode de Krylov et l'algorithme de point fixe.

Cette accélération permet donc de réduire le nombre d'itérations de la méthode fixed-stress. Il s'agit également d'un moyen d'améliorer la fiabilité de l'algorithme fixed-stress : si la méthode de point fixe peut parfois ne pas converger, les méthodes de Krylov sont connues pour être plus robustes. Pour compléter cette optimisation, il reste à choisir des sous-solveurs itératifs robustes pour faire face aux difficultés relevées dans la section 6.1.3 et remplacer les sous-solveurs directs qui ne sont pas utilisables dans les cas réels.

## 6.4 Recherche de sous-solveurs efficaces pour les sous-problèmes

On cherche dans cette section à qualifier les sous-solveurs à utiliser pour la résolution des sous-problèmes de la méthode fixed-stress. Pour effectuer cette sélection, on sort du cadre poroélastique en considérant des problèmes découplés.

### 6.4.1 Préconditionneurs ILU et AMG

#### Problème en pression

Le problème découplé d'écoulement est créé à partir du cas hétérogène décrit dans la section 6.2.3. Le maillage et la répartition en deux matériaux est conservée, avec les mêmes conditions aux limites et initiales pour la pression. La simulation court toujours sur un unique pas de temps ( $\Delta t = 5$  s), et la grandeur

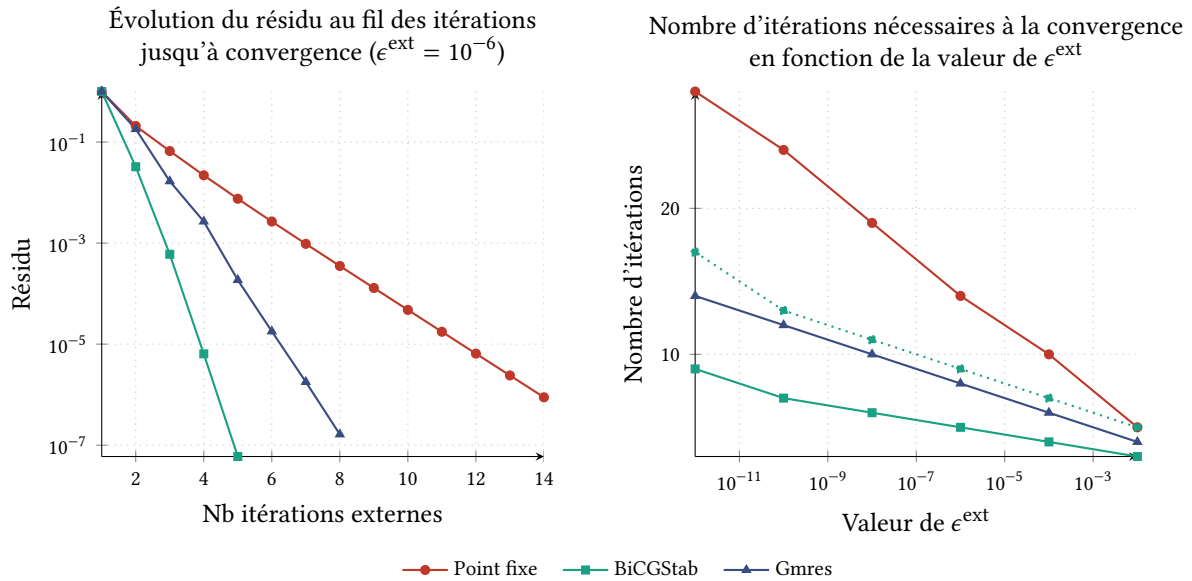


FIG. 6.4 : **Comparaison de la méthode de point fixe et des méthodes de Krylov** pour faire converger l'algorithme fixed-stress. À droite en pointillés, les itérations du BiCGStab sont multipliées par 2 pour tenir compte du double appel à la fonction  $\mathcal{R}$  spécifique à cette méthode. Sur le cas synthétique hétérogène 2D, les méthodes de Krylov nécessitent moins d'itérations.

physique variant selon le domaine est la mobilité  $\bar{\kappa}$  (supposée isotrope,  $\bar{\kappa} = \kappa \bar{l}_2$ ) : dans le cas dit homogène,  $\kappa = 10^{-3} \text{ m}^2(\text{Pa}\cdot\text{s})^{-1}$  dans tout le domaine, tandis que dans le cas dit hétérogène,  $\kappa = 10^{-3} \text{ m}^2(\text{Pa}\cdot\text{s})^{-1}$  dans la zone bleue et  $\kappa = 1 \text{ m}^2(\text{Pa}\cdot\text{s})^{-1}$  dans la zone jaune. Dans les deux cas,  $c_0 = 1 \text{ Pa}^{-1}$  dans l'ensemble du domaine. Le problème est résolu par le solveur itératif BiCGStab avec une tolérance  $\epsilon^{\text{int}}$  de  $10^{-8}$  et un préconditionnement à droite. À nouveau, les préconditionneurs ILU( $k$ ),  $k = 0$  ou  $k = 4$  ainsi que le préconditionneur AMG sont testés pour un calcul séquentiel ou parallèle jusqu'à 8 cœurs. Le nombre d'itérations du solveur selon la configuration est rassemblé dans le tableau 6.3. Dans le cas homogène, en dehors de ILU(4) qui souffre du passage en parallèle, tous les préconditionneurs testés sont relativement robustes lorsque le nombre de cœurs augmente : leur nombre d'itérations reste plus ou moins stable. En revanche, dans le cas hétérogène, la résolution séquentielle est déjà très difficile pour les préconditionneurs ILU, qui effectuent quasiment 10 fois plus d'itérations que dans le cas homogène, et devient impossible pour ceux-ci à partir d'un certain nombre de cœurs. Le préconditionnement par la méthode AMG s'avère au contraire extrêmement robuste pour ce problème. Le nombre d'itérations reste constant quelque soit la configuration et est de plus extrêmement faible par rapport aux autres méthodes. Ce préconditionneur semble donc utilisable pour la résolution du problème d'écoulement de fluide dans le cadre de la méthode séquentielle.

### Problème en déplacements

Pour créer un problème découplé de mécanique à partir du cas hétérogène couplé, il faut modifier les conditions aux limites ou le chargement qui étaient tous les deux nuls, de manière à créer une déformation non nulle. On impose donc sur la face supérieure une traction  $\mathbf{t}_N = -0.1\mathbf{e}_y$ , les autres conditions aux limites étant inchangées. À l'instar du problème de fluide, les constantes des matériaux sont ajustables de manière

	homogène				hétérogène			
	1p	2p	4p	8p	1p	2p	4p	8p
Aucun	389	431	391	390	x	x	x	x
Block-ILU(o)	110	127	121	131	1424	>	x	x
Block-ILU(4)	36	115	121	120	360	1899	1360	x
AMG	3	3	3	3	4	3	3	3

TAB. 6.3 : Comparaison du nombre d'itérations du BiCGStab sur le problème d'écoulement (160K inconnues) en fonction du préconditionnement. Le symbole (x) indique un échec de la résolution, tandis que le symbole (>) signifie que le nombre d'itérations maximum (2000) est dépassé. Le préconditionneur AMG est robuste pour ce problème.

	homogène				hétérogène			
	1p	2p	4p	8p	1p	2p	4p	8p
Aucun	>	>	1989	>	x	>	x	x
Block-ILU(o)	352	455	434	456	1519	x	>	>
Block-ILU(4)	132	422	442	464	368	>	x	1964
AMG	124	126	127	120	365	355	372	341

TAB. 6.4 : Comparaison du nombre d'itérations du BiCGStab sur le problème mécanique (321K inconnues) en fonction du préconditionnement. Le symbole (x) indique un échec de la résolution, tandis que le symbole (>) signifie que le nombre d'itérations maximum (2000) est dépassé. Aucun des préconditionneurs n'est robuste pour ce problème.

à obtenir un problème homogène ou hétérogène. C'est ici le module de Young qui permet d'aboutir à ces deux situations, en le fixant à 1 Pa dans l'ensemble du domaine dans le cas homogène ou bien en le fixant à 1 Pa dans le matériau bleu et 100 Pa dans le matériau jaune dans le cas hétérogène. Le coefficient de Poisson  $\nu = 0.25$  reste homogène dans les deux cas. On note qu'il n'y a en mécanique plus de notion de pas de temps puisque le problème est statique. Comme dans le cas précédent, les préconditionneurs ILU(o), ILU(4) et AMG sont évalués avec le solveur BiCGStab sur un ou plusieurs cœurs. Le critère d'arrêt est cette fois fixé à  $10^{-6}$  pour accélérer les tests. Le nombre d'itérations du solveur selon la configuration est retranscrit dans le tableau 6.4. Par rapport au problème d'écoulement, on note que le problème mécanique est bien plus difficile à résoudre. Même dans le cas homogène, où tous les préconditionneurs convergent, le nombre d'itérations est élevé dans l'absolu. En particulier, le préconditionneur AMG, qui était très efficace pour le problème de fluide, nécessite ici autour de 120 itérations. C'est certes moins que les méthodes ILU(k), mais cela reste assez coûteux. Cette quantité a toutefois le mérite de rester quasiment constante lorsque le nombre de cœurs augmente. En revanche, même le préconditionneur AMG supporte mal le passage au cas hétérogène, qui se traduit par un triplement du nombre de ses itérations. À la différence du problème d'écoulement, ce préconditionneur n'est pas stable vis à vis des coefficients physiques. On note toutefois que celui-ci a été utilisé avec ses options par défaut, et que ses résultats auraient pu être améliorés avec un paramétrage plus fin, par exemple en indiquant les vecteurs associés aux mouvements de corps rigide (cf section 2.1.1) qui sont dans le noyau de la matrice. Les préconditionneurs ILU(k) sont quand à eux presque systématiquement en échec sur les cas parallèles et donc insuffisamment fiables pour

être appliqué à la résolution du problème de mécanique. La conclusion de ce test diffère donc de celle du précédent : les préconditionneurs classiquement utilisés ne semblent pas assez robustes pour envisager leur usage comme sous-solveur de la partie mécanique dans la stratégie séquentielle.

### 6.4.2 Méthode de décomposition de domaine pour l'élasticité

Au tout début de ce chapitre, deux catégories de méthodes de résolution de système linéaire ont été mentionnées : les méthodes directes et les méthodes itératives. Pour trouver un préconditionneur robuste pour le système d'élasticité, les méthodes à présent investiguées appartiennent à une troisième catégorie visant à combiner les avantages des deux précédentes : les méthodes hybrides (catégorie à laquelle appartient également les méthodes multigrilles algébriques, soit dit en passant). Il s'agit ici de méthodes de décomposition de domaine, décrites en détails dans [DJN15], qui combinent des solveurs directs (robustes) utilisés sur des sous-domaines pour construire le préconditionneur (naturellement parallèle) d'une méthode itérative. L'étude de la robustesse de ces méthodes a fait l'objet de plusieurs travaux, parmi lesquels la thèse de N. Spillane [Spi14], traitant entre autres le système d'élasticité discrétisé par éléments finis. Il est ainsi connu que les méthodes de décomposition de domaine à un niveau peuvent souffrir de problèmes de robustesse, mais qu'il existe un cadre théorique permettant de corriger ce comportement. Ce problème de robustesse est dû à l'échange limité d'informations entre les sous-domaines, celles-ci n'étant transmises que de voisin à voisin. Le remède prend donc la forme de la résolution d'un problème dit grossier, commun à tous les sous-domaines, et soigneusement choisi afin de représenter une approximation du problème initial. On parle de méthode de décomposition de domaine à deux niveaux pour désigner l'utilisation de ce problème grossier. Dans cette section, on décrit brièvement les méthodes utilisées pour la création d'un préconditionneur utilisant un ou deux niveaux de décomposition de domaine. Une illustration similaire à celle de la section précédente est ensuite proposée.

#### Méthode à un niveau

Le préconditionneur à un niveau est issu de la méthode de Schwarz additive restrictive (voir [Gano8] pour une description et un historique des méthodes de Schwarz), sous-catégorie populaire des méthodes de décomposition de domaine, et sa description fait appel à la terminologie suivante.

Le maillage  $\tau_h$  est décomposé en  $N_p$  sous-maillages  $\tau_{h,i}$  pouvant éventuellement se recouvrir de sorte que  $\tau_h = \cup_{i=1}^{N_p} \tau_{h,i}$ . Cette décomposition permet de définir naturellement sur chaque sous-domaine un espace d'approximation local  $V_{h,i}$ . On obtient donc une décomposition des degrés de liberté globaux, notés ici  $\mathcal{N}$ , en degrés de libertés locaux  $\mathcal{N}_i$  définis sur l'espace d'approximation associé au sous-domaine  $\tau_{h,i}$  :  $\mathcal{N} = \cup_{i=1}^{N_p} \mathcal{N}_i$ . L'opérateur permettant de passer d'un vecteur global  $\mathcal{U} \in \mathbb{R}^{\#\mathcal{N}}$  à un vecteur local  $\mathcal{U}_i \in \mathbb{R}^{\#\mathcal{N}_i}$  est appelé opérateur de restriction. Il s'agit d'une matrice rectangulaire de taille  $\text{card}(\mathcal{N}_i) \times \text{card}(\mathcal{N})$  notée  $R_i$  dont les coefficients sont booléens. Inversement, la transposée  $R_i^T$  de cette matrice est l'opérateur permettant de passer d'un vecteur local à un vecteur global par ajout de zéros et est appelé opérateur de prolongement. Enfin, on définit sur chaque sous-domaine une matrice diagonale  $D_i$  de taille  $\text{card}(\mathcal{N}_i)$  de manière à ce que la famille  $(D_i)_{1 \leq i \leq N_p}$  forme une partition de l'unité :

$$I = \sum_{i=1}^{N_p} R_i^T D_i R_i \quad \text{avec } I \text{ matrice identité de taille } \text{card}(\mathcal{N}).$$

Il existe plusieurs moyens de définir les matrices  $D_i$ , qui diffèrent par le traitement appliqué aux degrés de liberté dupliqués appartenant à  $k$  sous-domaines ( $k > 1$ ). Pour ces degrés de liberté, on peut par exemple

choisir la valeur  $\frac{1}{k}$  pour le coefficient associé dans chacune des matrices  $D_i$  partageant ce degré de liberté, ou bien imposer arbitrairement 1 dans l'une de ces matrices et 0 dans toutes les autres.

Ces différents éléments permettent de construire sur chaque sous-domaine une matrice appelée matrice de Dirichlet locale  $A_i^D$  à partir de la matrice globale  $A$  selon

$$A_i^D = R_i A R_i^T. \quad (6.10)$$

Le préconditionneur à un niveau de la méthode de Schwarz additive restrictive (RAS, [CS99]) s'obtient grâce à l'équation (6.11) qui fait intervenir l'inverse des matrices de Dirichlet locales :

$$M_1^{-1} = \sum_{i=1}^{N_p} R_i^T D_i (A_i^D)^{-1} R_i. \quad (6.11)$$

En pratique, les inverses des matrices locales ne sont pas effectivement calculés ; ce sont des systèmes  $A_i^D y = b$  qui sont résolus. Si un nombre suffisant de sous-domaines est utilisé, ces systèmes sont de taille raisonnable et peuvent être résolus de manière directe.

### Méthode à deux niveaux

Une méthode à deux niveaux se construit à partir d'une méthode à un niveau et fait intervenir deux ingrédients supplémentaires principaux : un problème grossier, construit de manière à englober tous les sous-domaines pour améliorer la robustesse tout en étant peu coûteux à résoudre, et une formule de correction combinant ce problème grossier au préconditionneur à un niveau. Concernant le choix du problème grossier, on s'intéresse à une méthode algébrique permettant de construire l'espace grossier indépendamment de la décomposition en sous-domaines et de la distribution des coefficients physiques du problème, particulièrement adaptée aux problèmes hétérogènes : la méthode GenEO (**G**eneralized **E**igenvalue in the **O**verlap, [Spi+14]), qui construit son espace grossier à partir de la résolution de problèmes aux valeurs propres. En plus des matrices introduites précédemment, la méthode fait intervenir sur chaque sous-domaine une seconde matrice locale appelée matrice de Neumann et notée  $A_i^N$ . Il s'agit de la matrice qui aurait été obtenue si le problème physique était discrétisé uniquement sur le sous-domaine en question, en utilisant des conditions aux limites de traction nulles sur les bords du sous-domaine (en dehors des bords correspondant à une vraie frontière physique, qui conservent eux leurs conditions aux limites originales), d'où le nom de cette matrice. La méthode nécessite la résolution, dans chaque sous-domaine, du problème aux valeurs propres généralisé

$$A_i^N v = \lambda (D_i A_i^D D_i) v. \quad (6.12)$$

Si les valeurs propres de ce problème sont ordonnées dans l'ordre croissant, l'espace grossier local au sous-domaine  $i$  est l'espace  $Z_i$  engendré par les vecteurs propres associés aux  $n_i$  premières valeurs propres, cet entier pouvant être ajusté. L'espace grossier global est engendré par les colonnes de la matrice rectangulaire  $Z$ , de taille  $\text{card}(\mathcal{N}) \times N_c$  avec  $N_c = \sum_i n_i$ , qui s'obtient à partir des espaces  $Z_i$  selon

$$Z = \left[ R_1^T D_1 Z_1, R_2^T D_2 Z_2, \dots, R_{N_p}^T D_{N_p} Z_{N_p} \right]$$

et permet de définir la matrice grossière  $A_c = Z^T A Z$  de taille  $N_c \times N_c$ . Ainsi, lorsque les  $n_i$  augmentent, la correction apportée par l'espace grossier est plus importante, mais la taille du problème associé est plus grande. Une fois cet espace grossier construit et la matrice  $A_c$  obtenue, il reste à appliquer une formule de correction pour obtenir le préconditionneur de la méthode à deux niveaux. La correction employée dans la suite est la méthode *Balancing Neumann-Neumann* [Man93], qui calcule le préconditionneur selon

$$M_2^{-1} = (I - QA) M_1^{-1} (I - QA) + Q \quad \text{avec } Q = Z A_c^{-1} Z^T. \quad (6.13)$$

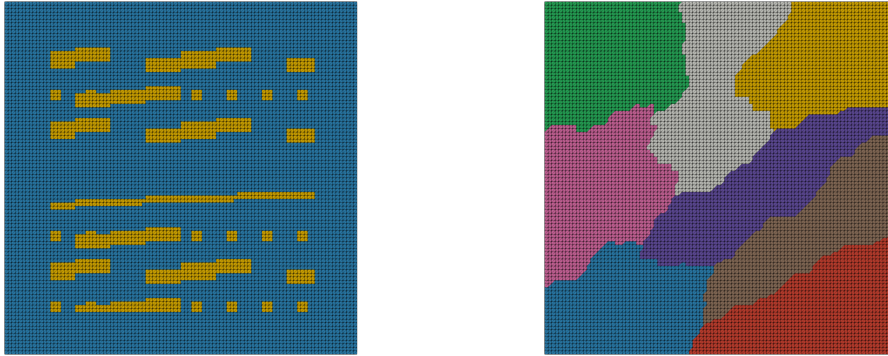


FIG. 6.5 : **Discrétisation en triangles du domaine hétérogène bidimensionnel** et exemple de partitionnement en sous-domaines, généré par Metis.

	homogène					hétérogène				
	2p	4p	8p	12p	16p	2p	4p	8p	12p	16p
1 niveau	84	102	132	148	161	139	241	345	430	482
2 niveaux	35	38	43	39	38	52	79	65	68	68

TAB. 6.5 : Comparaison du nombre d'itérations du solveur Gmres sur le problème mécanique en fonction du nombre de sous-domaines. Les méthodes de décomposition de domaine à deux niveaux semblent bien plus robustes que les préconditionneurs précédemment évalués sur ce problème.

### Application au cas de mécanique hétérogène

Pour évaluer ces méthodes de décomposition de domaine, le logiciel FreeFem++ est à nouveau utilisé. En effet, celui-ci embarque un module de décomposition de domaine<sup>3</sup> qui permet d'utiliser simplement ces méthodes, la construction des matrices et du préconditionneur étant gérée par l'outil. Le cas précédemment utilisé pour l'évaluation des solveurs appliqués au problème de mécanique demande quelques adaptations pour être utilisé avec FreeFem++ : les  $400 \times 400$  mailles carrées de la grille hétérogène sont chacune divisées en deux triangles de manière à obtenir un maillage triangulaire structuré tel que représenté sur la figure 6.5 (pour plus de lisibilité, la figure ne représente qu'un maillage de  $100 \times 100 \times 2$  triangles). La discrétisation du problème se fait par éléments finis  $P_1$ , qui sont ici équivalents à la méthode des éléments virtuels puisque la grille est triangulaire (*cf* section 3.3). Les valeurs des coefficients physiques sont les mêmes que lors des simulations précédentes. En revanche, la méthode itérative appliquée pour résoudre le système est l'algorithme Gmres au lieu de l'algorithme du BiCGStab. Le critère de convergence de  $10^{-6}$  est néanmoins conservé, tout comme le fait de préconditionner le système à droite pour comparer le nombre d'itérations. Le nombre d'itérations nécessaire à la convergence de la méthode Gmres, avec application du préconditionneur à un (6.11) ou à deux (6.13) niveaux est rassemblé dans le tableau 6.5. Dans le second cas, le nombre de vecteurs propres utilisés pour la construction de l'espace grossier est de 10 par sous-domaine.

Ce cas illustre bien le manque de robustesse de la méthode à un niveau : dans le cas hétérogène, le nombre d'itérations augmente fortement avec le nombre de sous-domaines. Cette augmentation se re-

3. <https://doc.freefem.org/documentation/ffdm/>

trouve d'ailleurs dans le cas homogène, même si elle est pour celui-ci d'intensité plus tenue. Dans les deux cas, l'amélioration apportée par le problème grossier est significative : le nombre d'itérations du solveur est extrêmement stable dans le cas homogène, et relativement stable dans le cas hétérogène (seules les simulations sur 2 ou 4 sous-domaines donnent un résultat légèrement plus éloigné de la moyenne). Cet apport ne coûte pas plus de temps de calcul : la résolution est plus rapide avec la méthode à deux niveaux qu'avec la méthode à un niveau. Dans le cas hétérogène à 8 sous-domaines, la durée totale de résolution est de 67 secondes pour la méthode à un niveau contre 13 secondes pour la méthode à deux niveaux. Pour celle-ci, la résolution du problème grossier est négligeable (moins de 0.02 s), et le temps passé à résoudre les problèmes aux valeurs propres (5 s) est largement compensé par celui économisé dans le solveur Gmres (5 secondes, contre 64 pour la méthode à un niveau). Enfin, dans l'absolu, le nombre d'itérations de la méthode à deux niveaux est non seulement inférieur à celui de la méthode à un niveau, mais il est également largement inférieur à celui qui avait été obtenu avec le préconditionnement AMG (voir tableau 6.4), même si la différence entre les solveurs et les maillages d'un cas sur l'autre rend cette comparaison délicate.

## 6.5 Utilisation des sous-solveurs spécialisés parallèles dans le cadre de la stratégie fixed-stress

La section précédente, en considérant des problèmes découplés, a permis de qualifier deux sous-solveurs prometteurs pour la résolution des sous-problèmes. Il reste à mettre effectivement en place ces solveurs dans le cadre de la stratégie séquentielle, et à vérifier leur comportement sur les sous-problèmes issus de cas couplés.

### 6.5.1 Utilisation de la librairie hpddm pour la décomposition de domaine

Cette partie, dont le sujet est à cheval entre les sections 6.4 et 6.5, explique comment utiliser effectivement les sous-solveurs dans la plateforme Arcane dans le cadre de la stratégie séquentielle. Puisque la méthode qualifiée pour le sous-problème d'écoulement, le préconditionnement AMG, est déjà accessible via PETSc, on se concentre en réalité sur la mise en place des méthodes de décomposition de domaine pour l'élasticité.

Pour mettre en place ces méthodes, on passe par la librairie hpddm<sup>4</sup> [JHNP13] qui en propose une implémentation parallèle performante. Si à terme l'objectif est de connecter cette librairie au simulateur, on ne dispose aujourd'hui que d'une installation indépendante, accessible via une interface python. Néanmoins, l'usage de l'interface est très proche de celui qui sera fait de l'API de la librairie lorsqu'elle sera connectée : dans les deux cas, la majorité des développements est à effectuer en amont pour préparer les objets attendus par la librairie (matrices de Dirichlet, de Neumann, partition de l'unité, etc.). Le processus actuel est donc le suivant : dans un premier temps, une simulation sur  $N_p$  cœurs est lancée dans le code de calcul, qui se charge du partitionnement et qui, durant l'assemblage du système linéaire, exporte un certain nombre d'objets dans des fichiers. Dans un second temps, un script python est lancé en parallèle sur  $N_p$  cœurs. Celui-ci lit les informations préalablement stockées, et appelle les fonctions contenues dans hpddm pour résoudre le système.

---

4. <https://github.com/hpddm/hpddm>

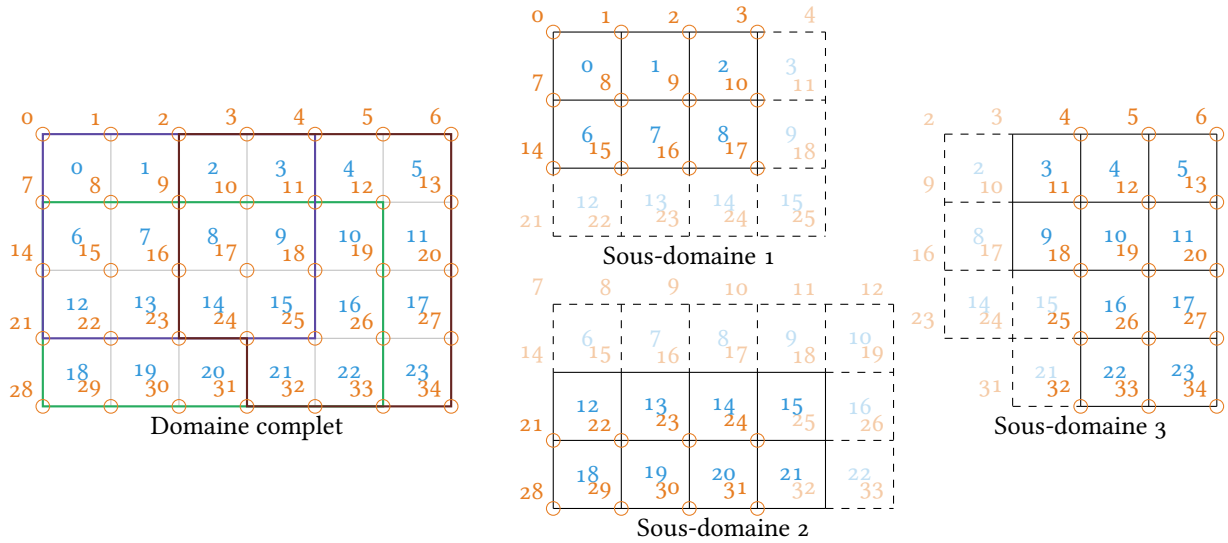


FIG. 6.6 : Schématisation du partitionnement en sous-domaines. Les cellules et les nœuds sont numérotés de manière globale. Dans les sous-domaines, les éléments own sont représentés par des cercles pour les nœuds et par des traits pleins pour les cellules. Les autres éléments affichés sont connus mais ghost.

### Décomposition en sous-domaines

La décomposition en sous-domaines est déléguée à Metis et aux routines d'Arcane. En raison du recouvrement, certains nœuds ou cellules se retrouvent dans plusieurs sous-domaines. Pour différencier ces éléments dupliqués, il existe une notion d'appartenance : chaque sous-domaine dispose de sa liste d'éléments propres (own) et d'éléments connus mais propres à un sous-domaine voisin (ghost). Cette situation est schématisée sur la figure 6.6. Dans cet exemple, le sous-domaine 3 possède 15 nœuds propres, et "voit" 6 nœuds propres au sous-domaine 1 (d'indices globaux 2, 3, 9, 10, 16 et 17) ainsi que 3 nœuds propres au sous-domaine 2 (d'indices 23, 24 et 31). En plus de cette numérotation globale, il existe une numérotation locale à chaque sous-domaine. L'objet de type dictionnaire permettant de passer de la numérotation locale à la numérotation globale est exploité pour construire un ingrédient de base requis par hpddm : la liste des nœuds connus chez chaque voisin, dans la numérotation locale du voisin. Pour ce faire, le processus  $P_i$  réceptionne le dictionnaire de chaque autre processus  $P_j$  envoyé à l'aide de `mpi_bcast`. Si l'intersection entre les clés de son dictionnaire et celles du dictionnaire reçu de  $P_j$  est non nulle, alors les deux sous-domaines sont voisins. Dans ce cas, l'ensemble des valeurs du dictionnaire reçu pour les clés communes est la liste de connectivité demandée par hpddm. La notion de own et ghost permet également de construire les matrices de partition de l'unité  $D_i$ . Dans chaque sous-domaine, l'entrée de  $D_i$  correspondant à un nœud donné vaut 1 si le nœud en question est own, et 0 sinon.

*Remarque.* En pratique, on ne travaille pas sur des nœuds mais sur des degrés de liberté : dans les dictionnaires, une clé (nœud numéroté globalement) est associée à un tableau de degrés de liberté, qui compte au plus 3 éléments, mais qui peut en avoir moins lorsqu'une ou plusieurs composantes du déplacement sont imposées par des conditions aux limites de Dirichlet sur ce nœud. De même, la matrice  $D_i$  comporte jusqu'à trois entrées par nœud.



### Matrices locales aux sous-domaines

Les sous-matrices les plus simples à construire sont les matrices de Neumann  $A_i^N$ . Il s'agit en effet des matrices qui auraient été obtenues si le problème n'était défini que sur le sous-domaine, avec des conditions aux limites de Neumann sur les frontières créées par la partition. Puisque ces conditions aux limites ne modifient pas la matrice, il suffit en réalité d'utiliser la procédure d'assemblage standard (voir algorithme 1) en utilisant cette fois la numérotation locale pour construire ces matrices.

La construction des sous-matrices de Dirichlet  $A_i^D$  est moins directe. Celles-ci sont définies comme la restriction de la matrice globale  $A$  au sous-domaine : de manière grossière, il s'agit de supprimer toutes les lignes et colonnes de la matrice globale qui ne sont pas associées à un degré de liberté connu du sous-domaine. Une solution consiste d'ailleurs à partir de la matrice globale, que l'on sait construire en parallèle, et d'en extraire les matrices  $A_i^D$  au prix de quelques manipulations des différentes indexations. Si cette solution est applicable via un script externe après écriture de la matrice dans un fichier, elle est en revanche délicate à mettre en place au sein du simulateur : on ne peut en effet pas accéder simplement aux coefficients de la matrice globale une fois celle-ci assemblée. L'approche décrite ici permet de construire à la volée les matrices  $A_i^D$  en même temps que les matrices  $A_i^N$ . Dans un premier temps, les matrices locales sont utilisées pour assembler les sous-matrices de Dirichlet, à ceci près que l'assemblage n'est réalisé que pour les nœuds own. Les entrées des matrices associées aux nœuds ghost vont en effet récupérer les valeurs issues des autres sous-domaines. Pour ce faire, dans un second temps, chaque processus  $P_i$  reçoit des autres processus  $P_j$  une liste de nœuds own dans le domaine  $i$  et ghost dans le domaine  $j$ . Après quoi, le processus  $P_i$  parcourt sa matrice, vérifie quelles lignes correspondent à un nœud présent dans la liste demandée, et envoie en retour l'intégralité de ces lignes au processus  $P_j$ , dans une numérotation globale. Il reste au processus  $P_j$  à vérifier, pour chaque ligne reçue, quelles colonnes correspondent également à un nœud présent dans son domaine (own ou ghost), puis à ajouter dans sa propre matrice  $A_i^D$  les coefficients retenus après avoir converti leur numérotation dans son indexation locale. Ces opérations sont effectuées en utilisant des conteneurs de données optimisés pour la recherche d'éléments, et les échanges se font par communications non bloquantes (`mpi_sendRecv`). Sur l'exemple de la figure 6.6, le processus 3 recevrait la liste [4, 11, 18, 25] du processus 1. Il renverrait en retour au processus 1 (parmi d'autres lignes) les lignes de coefficients associées au nœud 4, soit, en supposant pour simplifier qu'il n'y ait qu'un seul degré de liberté par nœud, les coefficients [ $A_3^D(4, 3)$ ,  $A_3^D(4, 4)$ ,  $A_3^D(4, 5)$ ,  $A_3^D(4, 10)$ ,  $A_3^D(4, 11)$ ,  $A_3^D(4, 12)$ ], le reste de cette ligne étant nul. Pour finir, le processus 1 aurait à filtrer cette liste pour ignorer les coefficients  $A_3^D(4, 5)$  et  $A_3^D(4, 12)$ , puisque les nœuds 5 et 11 lui sont totalement inconnus, et à ajouter les autres à sa matrice  $A_1^D$ .

Enfin, il reste à construire sur chaque sous-domaine un second membre local  $b_i$  et une solution initiale locale  $X_i^0$ . Pour ces quantités, l'extraction depuis les vecteurs globaux  $b$  et  $x_0$  pendant la simulation étant possible, c'est cette approche qui est utilisée. Il suffit donc pour chaque processus d'aller chercher dans les vecteurs globaux les coefficients correspondant aux nœuds présents dans son sous-domaine.

### Passage dans hpddm

Une fois ces objets construits, l'utilisation de l'interface de hpddm est relativement simple et est résumée dans la figure 6.7. L'option `schwarz_coarse_correction` déclenche la prise en compte du second niveau dans la méthode de décomposition de domaine, le paramètre `nu` servant à régler le nombre de valeurs propres à conserver dans chaque sous-domaine. Après résolution, chaque processus connaît la solution locale à son sous-domaine stockée dans le vecteur `sol`. Pour reconstruire la solution globale, la contribution de chaque sous-domaine aux nœuds partagés est pondérée à l'aide de la partition de l'unité. Ainsi, avec les choix effectués précédemment, c'est le sous-domaine propriétaire d'un nœud qui fixe la solution

```

1 import hpddm
2 A = hpddm.schwarzCreate(AiD, neighbors, connectivity)
3 hpddm.schwarzInitialize(A, Di)
4 if hpddm.optionSet(opt, b'schwarz_coarse_correction'):
5     hpddm.schwarzSolveGEVP(A, AiN)
6     hpddm.initializeCoarseOperator(hpddm.schwarzPreconditioner(A), nu)
7     hpddm.schwarzBuildCoarseOperator(A, mpi_communicator)
8 hpddm.schwarzCallNumfact(A)
9
10 hpddm.solve(A, rhs, sol, mpi_communicator)

```

FIG. 6.7 : **Utilisation de l'interface de la librairie hpddm** pour la résolution parallèle d'un problème par décomposition de domaine. Pour plus de lisibilité, les procédures ou arguments non essentiels ne figurent pas sur cet extrait.

sur celui-ci. Le rapatriement des solutions locales vers un vecteur solution globale s'effectue à l'aide de la fonction `mpi_gather`.

### 6.5.2 Comportement sur le cas synthétique hétérogène

Les sous-solveurs qualifiés pour la résolution des sous-problèmes étant maintenant utilisables, il reste à régler leurs paramètres, typiquement leur tolérance  $\epsilon^{\text{int}}$ , et à vérifier leur comportement dans le cadre de la stratégie séquentielle fixed-stress. On considère donc à nouveau le problème hétérogène synthétique de la section 6.2.3 pour lequel, jusqu'à maintenant, les sous-problèmes étaient résolus par une méthode directe.

Pour régler les paramètres des sous-solveurs, l'algorithme utilisé est la version point fixe avec itération sur la variable  $\sigma_v$ . Pour les deux sous-problèmes, le sous-solveur est le BiCGStab préconditionné par la méthode ILU(o). Même si ces choix sont loin d'être optimaux du point de vue des performances et de la robustesse (voir sections précédentes), on n'accorde pas d'importance à cet aspect dans cette étude préliminaire. Pour la même raison, les expériences sont menées en séquentiel. La tolérance de la méthode fixed-stress  $\epsilon^{\text{ext}}$  étant fixée à  $10^{-6}$ , on étudie l'impact de la tolérance des sous-solveurs  $\epsilon^{\text{int}}$  sur le nombre d'itérations externes et sur la qualité de la solution. Pour mesurer cette dernière, les déplacements et la pression après convergence sont comparés aux valeurs calculées par la résolution monolithique. Enfin, on mesure le nombre moyen d'itérations des sous-solveurs pour une itération externe. Les résultats sont rassemblés dans le tableau 6.6. Si la tolérance interne est trop faible par rapport à la tolérance externe, alors la méthode fixed-stress converge très rapidement vers une mauvaise solution. Lorsque la tolérance interne diminue, la convergence de la méthode fixed-stress nécessite plus d'itérations mais la précision des solutions s'améliore. En toute logique, ces deux quantités tendent vers un palier correspondant aux résultats obtenus avec un sous-solveur direct pour les sous-problèmes. Diminuer la tolérance interne améliore la solution, mais augmente le coût global de la résolution : non seulement le nombre d'itérations externes est plus élevé, mais le nombre d'itérations internes augmente également. Au vu de ces résultats, la tolérance interne doit être au moins de deux ordres de grandeur inférieure à la tolérance externe. En revanche, à partir de 4 ordres de grandeurs d'écart, l'amélioration de la solution semble insuffisante vis-à-vis de l'augmentation du nombre d'itérations des sous-solveurs. Dans la suite, on maintiendra donc trois ordres de

$\epsilon^{\text{int}}$	Itérations externes	Écart sur $p$	Écart sur $\mathbf{u}$	Itérations internes P	Itérations internes U
$10^{-4}$	4	$1.4 \cdot 10^{-1}$	$1.8 \cdot 10^{-1}$	15	106
$10^{-6}$	6	$4.8 \cdot 10^{-3}$	$7.8 \cdot 10^{-3}$	110	283
$10^{-8}$	9	$5.5 \cdot 10^{-5}$	$8.2 \cdot 10^{-5}$	197	324
$10^{-10}$	13	$9.8 \cdot 10^{-7}$	$1.8 \cdot 10^{-6}$	373	458
$10^{-12}$	14	$2.4 \cdot 10^{-7}$	$5.4 \cdot 10^{-7}$	714	545
Référence	14	$2.4 \cdot 10^{-7}$	$5.3 \cdot 10^{-7}$		

TAB. 6.6 : Nombre d'itérations de la méthode fixed-stress avec  $\epsilon^{\text{ext}} = 10^{-6}$  selon la tolérance des sous-solveurs  $\epsilon^{\text{int}}$ . Le tableau indique également les écarts avec les solutions du schéma monolithique, et le nombre d'itérations internes des sous-solveurs pour chaque itération externe (valeur moyenne).

grandeur d'écart entre la tolérance externe et les tolérances internes.

Par ailleurs, à chaque itération externe, les sous-solveurs itératifs internes nécessitent un vecteur pour initier leur algorithme. Fournir comme vecteur initial la solution issue de l'itération externe précédente  $\mathbf{u}^{n,l-1}$  ou  $p^{n,l-1}$  permet aux sous-solveurs internes de converger de plus en plus rapidement : la différence entre deux itérations externes successives diminue de plus en plus lorsque l'algorithme fixed-stress est sur le point de converger. Par exemple, dans le cas précédent avec  $\epsilon^{\text{int}} = 10^{-8}$ , le sous-solveur de l'écoulement nécessite 485 itérations internes pour converger pendant la première itération externe, mais converge en 2 itérations pendant la dernière itération externe. Pour le sous-solveur des déplacements, on passe de 503 à 2 itérations internes.

L'étude préliminaire étant à présent terminée, on sait comment paramétrer les sous-solveurs : ceux-ci doivent démarrer leurs itérations sur la solution de l'itération externe précédente et avoir un critère d'arrêt de trois ordres de grandeur inférieur à la tolérance externe. Il subsiste une limitation due à l'usage externalisé de la librairie hpddm : il n'est actuellement pas possible de faire remonter la solution de hpddm vers le simulateur et donc de l'utiliser pour valider l'itération externe. Pour évaluer tout de même les sous-solveurs spécialisés, la stratégie retenue est la suivante : en même temps que ceux-ci résolvent les sous-problèmes, deux sous-solveurs directs sont utilisés pour effectuer la même résolution. Ce sont les solutions issues des sous-solveurs directs qui permettent l'avancée des itérations de la méthode séquentielle, les sous-solveurs spécialisés n'étant utilisés que pour étudier leur comportement. Ainsi, on utilise pour la pression la méthode du BiCGStab préconditionnée par AMG (via une interface interne à Arcane), et pour les déplacements la méthode Gmres préconditionnée par la méthode de décomposition de domaine à deux niveaux (via l'interface externalisée). Dans les deux cas, la tolérance est fixée à  $\epsilon^{\text{int}} = 10^{-9}$ . Selon le nombre de cœurs utilisés pour la résolution, le nombre d'itérations des sous-solveurs pour chacune des 14 itérations nécessaires à la convergence de l'algorithme fixed-stress est représenté sur la figure 6.8. Même dans le cadre du problème couplé, le sous-solveur pour la pression reste extrêmement robuste : ce n'est que pour deux sous-systèmes que le nombre d'itérations varie selon le nombre de sous-domaines, l'écart n'étant alors que de une itération. Pour le sous-problème en déplacements, la méthode de décomposition de domaine à deux niveaux reste relativement robuste puisque le solveur converge dans tous les cas, avec une variation de son nombre d'itérations qui reste faible par rapport à celles qui avaient pu être relevées dans les cas qualifiés de problématique (tableaux 6.2 et 6.4).

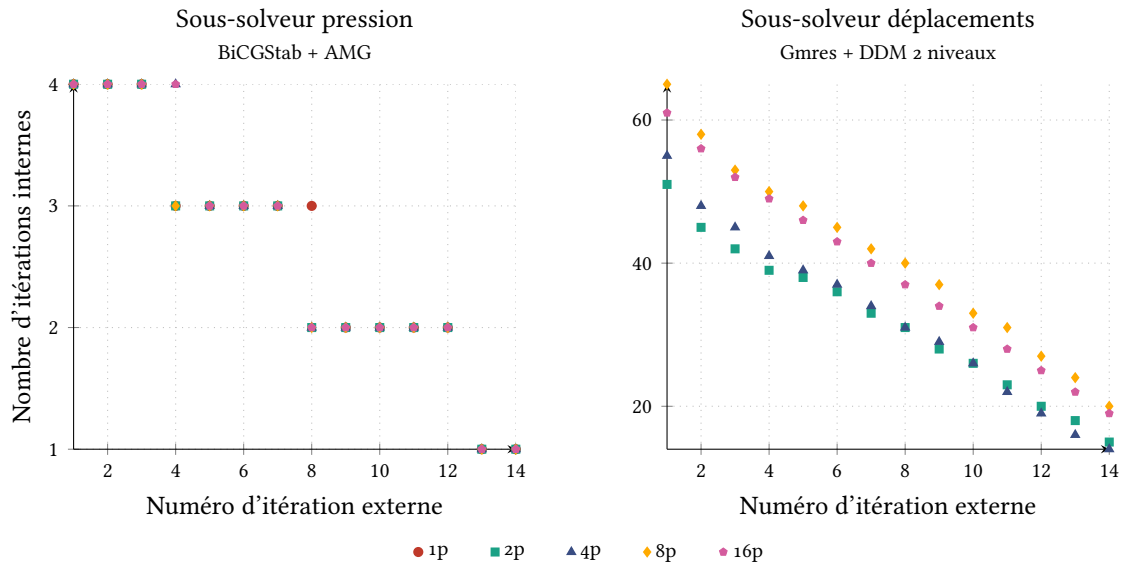


FIG. 6.8 : Nombre d'itérations des sous-solveurs pour chaque itération externe sur le cas hétérogène 2D, selon le nombre de cœurs employés. Même dans le cadre du couplage, les sous-solveurs restent robustes.

### 6.5.3 Application au problème tridimensionnel réaliste

Les différentes méthodes détaillées dans ce chapitre ainsi que les observations effectuées sont à présent mises à contribution pour tenter de résoudre le problème réel tridimensionnel, qui avait été utilisé dans la section 6.1.3 pour choisir la stratégie à adopter.

Le système est donc résolu par stratégie séquentielle, accélérée par la méthode de Krylov Gmres, en itérant sur la variable  $\sigma_v$ . La tolérance externe est fixée à  $10^{-6}$ . Comme dans le cas synthétique hétérogène de la section précédente, les sous-solveurs utilisés sont l'algorithme du BiCGStab préconditionné par AMG pour la pression, et la méthode Gmres préconditionnée par la décomposition de domaine à deux niveaux pour les déplacements. Pour ces deux méthodes, la tolérance interne est fixée à  $10^{-9}$  et le vecteur initial est pris nul. À nouveau, pour s'affranchir de l'intégration partielle de la librairie hpddm, des sous-solveurs directs permettent d'avancer la résolution. Pour la même raison ainsi que pour diminuer le temps de calcul, seul le premier pas de temps du système est résolu. Sur ce premier pas de temps, l'accélération de la méthode fixed-stress permet à celle-ci de converger en 11 itérations. À titre indicatif, 24 itérations sont nécessaires avec l'algorithme de point fixe. Comme dans le cas précédent, le nombre d'itérations des sous-solveurs pour chacune de ces 11 itérations est représenté sur la figure 6.9, chaque courbe correspondant à un certain nombre de sous-domaines. Comme pour le cas synthétique présenté dans la section précédente, le sous-solveur en pression (préconditionné par AMG) est extrêmement stable : pour une itération externe (et donc pour un système) donnée, son nombre d'itérations est le même quelque soit le nombre de sous-domaines utilisés. Ce nombre varie également très peu au fil des itérations externes puisqu'il reste fixé à 4 (excepté pour la première résolution), ce qui est faible : la convergence de la méthode est très rapide. Pour la résolution du sous-problème élastique par décomposition de domaine, on observe cette même stabilité au fil des itérations externes pour un nombre de sous-domaines donné. L'écart selon le nombre de cœurs utilisés est sensiblement identique à celui qui avait été obtenu sur le cas bidimensionnel précédent, de même

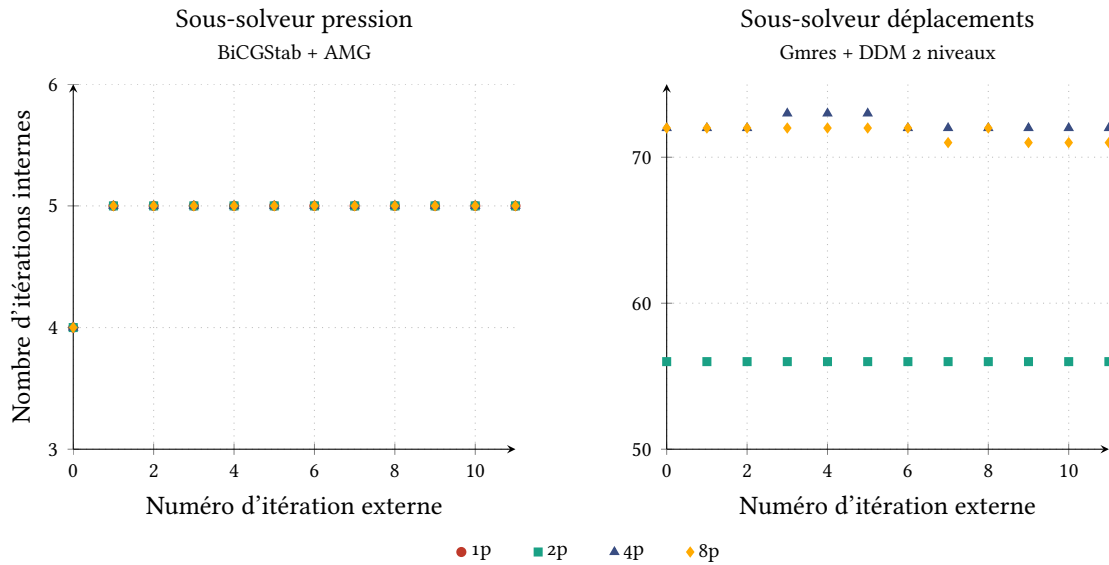


FIG. 6.9 : **Nombre d'itérations des sous-solveurs pour chaque itération externe sur le cas réaliste 3D, selon le nombre de cœurs employés.** Les conclusions sont semblables à celles du cas synthétique 2D : le sous-solveur en pression est extrêmement robuste, le sous-solveur en déplacement relativement robuste.

que la valeur du nombre d'itérations : le passage au cas réel n'entraîne pas de changements significatifs quant à la robustesse de la méthode.

Pour revenir au pari effectué en début de chapitre, les résultats obtenus sur ce cas sont plutôt encourageants. Du point de vue du nombre de systèmes à résoudre, l'accélération par méthode de Krylov de la stratégie fixed-stress a permis de diminuer le nombre d'itérations externes. Du point de vue de la robustesse, le sous-problème des déplacements, sur lequel a été concentrée la majorité des efforts, semble pouvoir être résolu de manière robuste par les méthodes de décomposition de domaine à deux niveaux. Pour le sous-problème d'écoulement, le solveur préconditionné par AMG demeure très robuste et devrait donc être suffisant. Dans tous les cas, il reste à poursuivre les efforts d'intégration de la librairie hpddm au simulateur pour répondre à l'aspect performances de la question : soit de manière directe, soit en attendant son implémentation dans une bibliothèque telle que PETSc.

# 7

## Conclusion et perspectives

Ce chapitre résume le contenu présenté dans ce manuscrit, et propose quelques pistes pouvant permettre l'amélioration des travaux effectués.

### Sommaire

---

<b>7.1 Travaux effectués . . . . .</b>	<b>133</b>
<b>7.2 Perspectives . . . . .</b>	<b>134</b>
<b>7.3 Réalisations et publications. . . . .</b>	<b>135</b>

---

### 7.1 Travaux effectués

L'objectif de cette thèse était de concevoir et d'étudier de nouveaux schémas de discrétisation et de résolution du système poroélastique, compatibles avec des maillages généraux de manière à répondre aux problématiques spécifiques du domaine des géosciences. Pour cela, des travaux ont été effectués dans trois directions : la méthode des éléments virtuels a été étudiée pour la discrétisation de l'équation d'élasticité sur des maillages généraux ; deux schémas couplés ont été créés pour la discrétisation du problème poroélastique ; une stratégie de résolution a été proposée pour calculer de manière efficace la solution des équations discrétisées.

L'étude de la méthode des éléments virtuels de degré un pour le problème élastique, exposée dans le chapitre 3, a permis de s'assurer de la capacité de cette méthode à remplacer celle des éléments finis pour la discrétisation de la partie mécanique. La construction de la méthode des éléments virtuels d'ordre un, en deux ou trois dimensions, a été détaillée, tout comme les techniques permettant son implémentation (par exemple calcul des coefficients des matrices locales, procédure d'assemblage ou traitement des conditions aux limites). Les études numériques effectuées ont illustré la capacité de la méthode à traiter des maillages bien plus généraux que les maillages simpliciaux, tels que les maillages polyédriques ou non conformes, mais ont également permis d'observer certains avantages de la méthode des éléments virtuels par rapport à celle des éléments finis, par exemple dans le traitement de mailles rectangulaires de ratio d'aspect défavorable ou dans le cas de matériaux tendant vers le domaine incompressible.

Le développement de schémas couplés pour la discrétisation du problème poroélastique a été réalisé

en discrétisant l'équation de conservation de la masse du fluide dans le formalisme des volumes finis et en réutilisant la méthode des éléments virtuels introduite précédemment pour l'équation d'équilibre mécanique. À partir de ce schéma couplé générique, deux variantes ont été obtenues selon le choix effectué pour l'approximation volumes finis. Dans le chapitre 4, le premier schéma couplé VEM-TPFA a été construit en utilisant l'approximation à deux points pour les volumes finis. La simplicité relative de la construction de ce schéma a permis d'en effectuer une analyse mathématique rigoureuse, en démontrant l'existence et l'unicité de la solution discrète et en établissant une estimation *a priori* de l'erreur d'approximation. Ce schéma a été illustré aussi bien sur un cas académique, illustrant ainsi numériquement l'ordre de convergence du schéma, que sur des problèmes classiques de poromécanique. Dans le but de traiter des maillages plus généraux et représentatifs de ceux utilisés dans les applications métiers, le second schéma couplé VEM-MPFA a été développé dans le chapitre 5. Utilisant une approximation multipoints pour les volumes finis, et une méthode de calcul légèrement modifiée pour les éléments virtuels, ce nouveau schéma a été appliqué avec succès sur un maillage réellement issu d'un domaine sous-terrain, représentatif de ceux utilisés en simulation de bassin. Le choix de paramètres réalistes pour ce cas numérique a permis de constater la qualité de la solution obtenue avec cette méthode de discrétisation.

Pour résoudre efficacement le système linéaire obtenu à l'issue de l'étape de discrétisation, une stratégie de résolution a été proposée dans le chapitre 6. L'intérêt pour cette thématique a été justifiée par les limitations des méthodes aujourd'hui disponibles, qui souffrent de problèmes de robustesse lorsqu'elles sont utilisées dans le cadre du calcul parallèle. Dans un premier temps, une méthode de résolution séquentielle basée sur l'algorithme fixed-stress split, qui découple les deux équations, a été mise en place. La présentation de cette méthode comme un problème d'annulation de résidu a permis d'en améliorer la vitesse de convergence et la stabilité, en utilisant une méthode de Krylov pour faire progresser les itérations de couplage plutôt que les algorithmes de point fixe appliqués dans l'approche classique. Dans un second temps, des sous-solveurs spécialisés pour chacun des deux sous-problèmes ont été recherchés et utilisés dans le cadre de la stratégie fixed-stress : une méthode de type multigrille algébrique pour l'équation d'écoulement et une méthode de décomposition de domaine à deux niveaux pour le problème de mécanique. Cette seconde méthode a nécessité la construction de sous-matrices et d'informations de connectivité dans chaque sous-domaine, requises par la librairie parallèle externe à laquelle est confiée la résolution. Les premiers tests numériques laissent à penser que ces choix permettent d'obtenir une stratégie de résolution robuste vis-à-vis du nombre de sous-domaines et des hétérogénéités.

## 7.2 Perspectives

Divers aspects de ces travaux pourraient être améliorés ou enrichis. Sur le plan de la modélisation physique, la loi d'élasticité linéaire exclut un certain nombre de matériaux qui ont un comportement plus complexe. Considérer des lois de comportement plus complexes, comme l'élastoplasticité ou la viscoplasticité, permettrait donc un enrichissement du modèle. L'application de la méthode des éléments virtuels à des lois de comportement plus générales fait justement l'objet d'études actives, voir par exemple [HAW19] et ses références pour un aperçu de celles-ci. Concernant l'équation de conservation de la masse de fluide, un enrichissement de la modélisation prendrait par exemple la forme du traitement d'écoulements polyphasiques voire compositionnels. Les schémas volumes finis utilisés actuellement pourraient suffire à traiter ce type de situations.

Cela dit, sur le plan de la discrétisation, il existe aujourd'hui des alternatives au schéma multipoints, plus robustes que celui-ci. On pense par exemple aux schémas hybrides [EGH09] ou aux schémas volumes finis non linéaires [YS08]. L'une de ces méthodes pourrait remplacer l'approximation multipoints

du flux pour améliorer la stabilité du schéma couplé. On note également la possibilité d'augmenter la vitesse de convergence du schéma couplé, en utilisant une méthode des éléments virtuels d'ordre plus élevé [ABLS17a], à condition d'être capable de gérer le couplage avec l'équation d'écoulement, éventuellement discrétisée elle aussi par un schéma d'ordre élevé.

Enfin, sur le plan de la résolution du système linéaire, deux actions sont envisageables. La première consisterait à réduire à nouveau le coût de la stratégie fixed-stress : par exemple en utilisant un pas de temps différent pour chaque sous-problème (la partie mécanique, plus coûteuse à résoudre, évoluant moins rapidement que la partie écoulement), en suivant la méthodologie développée dans [Alm+16] ; ou bien en travaillant sur une version parallèle en temps de la méthode, en se basant sur les travaux récents [Bor+18]. La seconde serait de rester sur une résolution monolithique du système, en tentant d'y appliquer une méthode de décomposition de domaine à deux niveaux, cette approche n'ayant à notre connaissance jamais été évaluée.

### 7.3 Réalisations et publications

Concrètement, la valorisation de ces travaux de thèse se retrouve sous deux formes : la mise à disposition aux équipes de recherche d'IFPEN d'un module de calcul dans la plateforme de simulation Arcane, et plusieurs communications scientifiques.

Le module de calcul développé permet la résolution du problème élastique ou poroélastique sur des maillages polyédriques bidimensionnels ou tridimensionnels, en utilisant la méthode des éléments virtuels pour l'équation d'élasticité et une méthode de volumes finis à deux points ou multipoints pour l'(éventuelle) équation d'écoulement. Pour le problème de poroélasticité, le choix d'une résolution monolithique ou séquentielle par la méthode fixed-stress split est laissé à l'utilisateur, qui peut dans le second cas choisir entre une méthode de type point fixe, l'algorithme BiCGStab ou l'algorithme Gmres pour les itérations de couplage. La résolution peut-être effectuée par l'un des solveurs de PETSc interfacé avec la plateforme ou, pour le (sous-)problème d'élasticité, en laissant le module générer les objets nécessaires à la décomposition de domaine. De plus, le module tire parti de plusieurs briques de base de la plateforme pour proposer des fonctionnalités avancées, parmi lesquelles : gestion (partielle) de maillages évolutifs variant au cours du temps, coefficients physiques hétérogènes, prise en compte de plusieurs types de conditions aux limites possiblement différentes dans chaque direction (pour la mécanique) et sur chaque groupe de faces de bord, analyse syntaxique de fonctions analytiques pour les données d'entrée du problème, parallélisation Mpi.

Une partie des résultats obtenus ont été publiés dans les actes d'une conférence puis dans une revue, et ont fait l'objet de communications orales dans des congrès ainsi que d'un poster, présenté en annexe de ce document.

#### Publications

- J. COULET, I. FAILLE, V. GIRAULT, N. GUY et F. NATAF. « A fully coupled scheme using Virtual Element Method and Finite Volume for poroelasticity ». In : *Computational Geosciences* (2019), p. 1-23. DOI : [10.1007/s10596-019-09831-w](https://doi.org/10.1007/s10596-019-09831-w)
- J. COULET, I. FAILLE, V. GIRAULT, N. GUY et F. NATAF. « Fully coupled schemes using Virtual Element and Finite Volume Discretisations for Biot equation modelling ». In : *Proceedings, 16th European Conference on the Mathematics of Oil Recovery*. 2018. DOI : [10.3997/2214-4609.201802271](https://doi.org/10.3997/2214-4609.201802271)



### **Communications**

- *Coupling Virtual Elements and Finite Volume Methods for Geomechanics*; Conférence POEMS, France (2019) [Poster]
- *Coupling Virtual Element Methods and Finite Volume Schemes for Computational Geomechanics on General Meshes*; Congrès SIAM-GS, États-Unis (2019)
- *Fully Coupled Schemes Using Virtual Element And Finite Volume Discretisations For Biot Equations Modelling*; Conférence ECMOR XVI, Espagne (2018)
- *Méthode des éléments virtuels pour les calculs géomécaniques*; 8<sup>e</sup> congrès de la SMAI, France (2017)

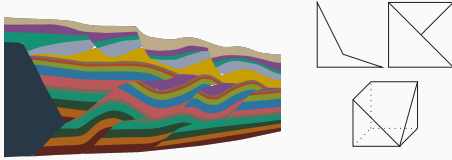
## Geomechanics' challenges

Involves two physics

→ Requires two spatial discretization methods working on a **single** mesh

Highly deformed media

→ Methods must work on **poor quality** meshes



## Governing equations

Under the assumptions of quasistatic strains and slightly compressible single-phase flow, the mechanical equilibrium and the fluid mass conservation are coupled through Biot's equations.

$$-\text{div}(\bar{\bar{C}}\bar{\bar{\epsilon}}_u - \alpha p \bar{I}_d) = \mathbf{f} \quad \text{with } \bar{\bar{\epsilon}}_u = \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T)$$

$$\partial_t(c_0 p + \alpha \text{div}(\mathbf{u})) + \text{div}(-\bar{\kappa}(\nabla p - \rho_f \mathbf{g})) = q$$

where  $\mathbf{u}$  is the solid displacement and  $p$  the fluid pressure.

In this work, we assume that the solid has a linear elastic behaviour described by the stiffness tensor  $\bar{\bar{C}}$ .

$\alpha$	Biot parameter
$c_0$	Constrained specific storage coefficient
$\rho_f$	fluid density
$\bar{\kappa}$	Mobility matrix

## Spatial discretization

Working in a **variational** framework for the mechanics and in the **finite volume** framework for the fluid flow,

$$\alpha(\mathbf{u}_h^n, \mathbf{v}_h) - \sum_K \alpha \int_K \text{div}(\mathbf{v}_h) p_h^n = \int_\Omega \mathbf{f}^n \cdot \mathbf{v}_h$$

$$\int_K (c_0 p_h^n + \alpha \text{div}(\mathbf{u}_h^n)) + \Delta t \sum_{f \subset \partial K} \mathcal{F}_{Kf}^n = \int_K (\Delta t q^n + c_0 p_h^{n-1} + \alpha \text{div}(\mathbf{u}_h^{n-1}))$$

Given two bases for the discrete spaces, the equivalent matrix form is given by

$$\begin{bmatrix} \mathcal{A} & -\mathcal{B} \\ \mathcal{B}^T & \mathcal{F} \end{bmatrix} \begin{bmatrix} \mathcal{U} \\ \mathcal{P} \end{bmatrix} = \begin{bmatrix} \mathcal{L}^u \\ \mathcal{L}^p + \mathcal{L}^c \end{bmatrix}$$

where  $\mathcal{F}$  depends only on flow problem,  $\mathcal{A}$  depends only on the elasticity problem,  $\mathcal{B}$  stores the coupling terms  $\int p_h^n \text{div}(\mathbf{v}_h)$ .

We can now choose a discretization method for  $\mathcal{A}$  and another for  $\mathcal{F}$ . Indeed, the scheme can be customized picking the following elements:

### A) A Virtual Element Method

Key idea of VEM[1]: on each element, substitute  $a$  with an approximate discrete form  $a_h$

- consistent (ensures accuracy)
- stable (ensures coercivity)
- computable from the *dofs*.

To have the consistency and the stability, set

$$a_h^K(\mathbf{u}, \mathbf{v}) := a^K(\pi^K \mathbf{u}, \pi^K \mathbf{v}) + h_K^{d-2} \max_{|S|} |\bar{\bar{C}}|_S^K (\mathbf{u} - \pi^K \mathbf{u}, \mathbf{v} - \pi^K \mathbf{v})$$

where  $\pi^K \mathbf{u}$  is a projection to the polynomial part of  $\mathbf{u}$  and where  $s^K$  stabilizes the non polynomial part. For the lowest order of accuracy, we can take

$$\pi^K \mathbf{v}(\mathbf{x}) = \left( \frac{1}{|K|} \int_K \nabla \mathbf{v} \right) (\mathbf{x} - \bar{\mathbf{x}}_K) + \frac{1}{M_K} \sum_{i \in \mathcal{M}_K} \mathbf{v}(V_i)$$

which is computable from the *dofs*. With this approximate form, basis functions are *virtually* defined over any general element.

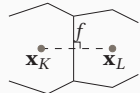
### B) A Finite Volume scheme

Choosing any FV scheme with cell-centered unknowns allows an easy treatment of the coupling terms. From [2] we highlight:

▷ *Two-Point Flux Approximation*

Use only two points to compute the fluxes

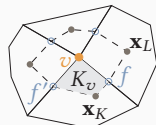
$$\mathcal{F}_{Kf} = |f| \kappa_{KL} d_{KL}^{-1} (p_K - p_L)$$



- ♥ Low computational cost
- ♥ High stability
- ♥ Requires orthogonality condition  $(KL) \perp f$

▷ *Multi-Point Flux Approximation*

Compute half fluxes  $\mathcal{F}_{Kfv}$ , then eliminate face unknowns using consistency  $\mathcal{F}_{Kfv} + \mathcal{F}_{Lfv} = 0$  and finally sum half fluxes to get  $\mathcal{F}_{Kf} = \mathcal{F}_{Kfv} + \mathcal{F}_{Lfv}$ .



- ♥ Allows more general meshes
- ♥ Larger stencil (more costly)
- ♥ Sometimes unstable

### C) A Solution Strategy

Use either a **fully** coupled (monolithic) or an **iteratively** coupled strategy.

▷ *Fully coupled resolution*

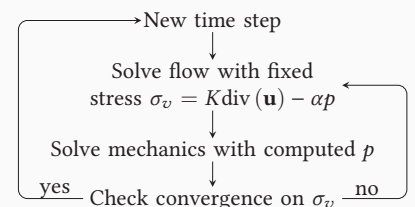
Both equations are simultaneously solved.

♥ Unconditionally stable

♥ Huge matrix, no efficient solver available

▷ *Iteratively coupled resolution*

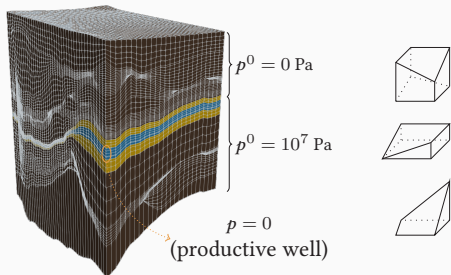
Use a splitting [3] such as the fixed-stress split:



- ♥ More iterations, needs a convergence criteria
- ♥ Pick specialized solvers for each subproblem
- ♥ Krylov-like methods can speed up convergence

## Numerical illustration

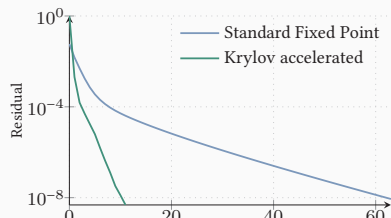
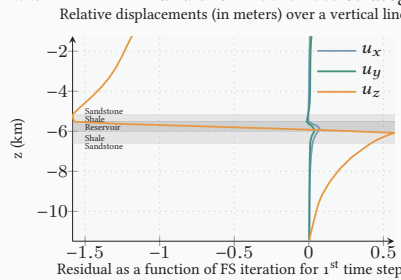
Layered 3D mesh including degenerated cells



	$\kappa$	$\phi$	$E$	$\nu$	$\alpha$
Sandstone	$10^{-9}$	0.2	$5 \cdot 10^9$	0.3	1
Shale	$10^{-15}$	0.1	$5 \cdot 10^9$	0.3	1
Reservoir	$10^{-9}$	0.3	$10^9$	0.3	1

Bottom and lateral faces:  $\nabla p \cdot \mathbf{n} = 0, \mathbf{u} \cdot \mathbf{n} = 0$   
Top face:  $p = 0$  and free stress condition

→ Solve ~300K dofs over 20 time steps (T=1 year) with VEM-MPFA and the fixed-stress strategy:



## References

- [1] L. Beirão Da Veiga, C. Lovadina, and D. Mora. "A Virtual Element Method for elastic and inelastic problems on polytope meshes". In: *Computer Methods in Applied Mechanics and Engineering* 295 (2015), pp. 327–346.
- [2] J. Droniou. "Finite volume schemes for diffusion equations: Introduction to and review of modern methods". In: *Math. Models & Methods in App. Sciences* 24.8, SI (2014), pp. 1575–1619.
- [3] A. Mikelić, B. Wang, and M. F. Wheeler. "Numerical convergence study of iterative coupling for coupled flow and geomechanics". In: *Comp. Geosciences* 18.3 (2014), pp. 325–341.

## Future works

Use domain decomposition methods as an efficient preconditioner to solve  $\mathbf{u}$   
More general mechanical laws



# Bibliographie

- [Aav02] I. AAVATSMARK. « An Introduction to Multipoint Flux Approximations for Quadrilateral Grids ». In : *Computational Geosciences* 6.3 (2002), p. 405-432. DOI : [10.1023/A:1021291114475](https://doi.org/10.1023/A:1021291114475).
- [ABLS17a] E. ARTIOLI, L. BEIRÃO DA VEIGA, C. LOVADINA et E. SACCO. « Arbitrary order 2D virtual elements for polygonal meshes : part I, elastic problem ». In : *Computational Mechanics* 60.3 (2017), p. 355-377. DOI : [10.1007/s00466-017-1404-5](https://doi.org/10.1007/s00466-017-1404-5).
- [ABLS17b] E. ARTIOLI, L. BEIRÃO DA VEIGA, C. LOVADINA et E. SACCO. « Arbitrary order 2D virtual elements for polygonal meshes : part II, inelastic problem ». In : *Computational Mechanics* 60.4 (2017), p. 643-657. DOI : [10.1007/s00466-017-1429-9](https://doi.org/10.1007/s00466-017-1429-9).
- [Abo+96] Y. ABOUSLEIMAN, A. H.-D. CHENG, L. CUI, E. DETOURNAY et J.-C. ROEGIERS. « Mandel's problem revisited ». In : *Géotechnique* 46.2 (1996), p. 187-195. DOI : [10.1680/geot.1996.46.2.187](https://doi.org/10.1680/geot.1996.46.2.187).
- [ADLK01] P. AMESTOY, I. DUFF, J.-Y. L'EXCELLENT et J. KOSTER. « A Fully Asynchronous Multifrontal Solver Using Distributed Dynamic Scheduling ». In : *SIAM Journal on Matrix Analysis and Applications* 23.1 (2001), p. 15-41. DOI : [10.1137/S0895479899358194](https://doi.org/10.1137/S0895479899358194).
- [AEK06] I. AAVATSMARK, G. T. EIGESTAD et R. A. KLAUSEN. « Numerical Convergence of the MPFA O-Method for General Quadrilateral Grids in Two and Three Dimensions ». In : *Compatible Spatial Discretizations*, sous la dir. de D. N. ARNOLD et al. Springer New York, 2006, p. 1-21. ISBN : 978-0-387-38034-6.
- [AGM10] L. AGÉLAS, C. GUICHARD et R. MASSON. « Convergence of Finite Volume MPFA O type Schemes for Heterogeneous Anisotropic Diffusion Problems of General Meshes ». In : *International Journal of Finite Volumes* 7.2 (2010), p. 1-34. HAL : [hal-00340159v2](https://hal.archives-ouvertes.fr/hal-00340159v2).
- [Ahm+13] B. AHMAD, A. ALSAEDI, F. BREZZI, L. MARINI et A. RUSSO. « Equivalent projectors for virtual element methods ». In : *Computers and Mathematics with Applications* 66.3 (2013), p. 376-391. DOI : [10.1016/j.camwa.2013.05.015](https://doi.org/10.1016/j.camwa.2013.05.015).
- [AK08] G. ALLAIRE et S. M. KABER. *Numerical Linear Algebra*. T. 55. Texts in Applied Mathematics. Springer, 2008. ISBN : 978-0-387-34159-0.
- [All12] G. ALLAIRE. *Analyse numérique et optimisation*. 2<sup>e</sup> éd. Les éditions de l'école Polytechnique, 2012. ISBN : 978-2-7302-1255-7.
- [Alm+16] T. ALMANI, K. KUMAR, A. DOGRU, G. SINGH et M. F. WHEELER. « Convergence analysis of multirate fixed-stress split iterative schemes for coupling flow with geomechanics ». In : *Computer Methods in Applied Mechanics and Engineering* 311 (2016), p. 180-207. DOI : [10.1016/j.cma.2016.07.036](https://doi.org/10.1016/j.cma.2016.07.036).

- [ALM16] B. AYUSO DE DIOS, K. LIPNIKOV et G. MANZINI. « The nonconforming virtual element method ». In : *Mathematical Modelling and Numerical Analysis* 50.3 (2016), p. 879-904. DOI : [10.1051/m2an/2015090](https://doi.org/10.1051/m2an/2015090).
- [AML19] E. ARTIOLI, S. de MIRANDA, C. LOVADINA et L. PATRUNO. « An equilibrium-based stress recovery procedure for the VEM ». In : *International Journal for Numerical Methods in Engineering* 117.8 (2019), p. 885-900. DOI : [10.1002/nme.5983](https://doi.org/10.1002/nme.5983).
- [AMS18] E. ARTIOLI, S. MARFIA et E. SACCO. « High-order virtual element method for the homogenization of long fiber nonlinear composites ». In : *Computer Methods in Applied Mechanics and Engineering* 341 (2018), p. 571-585. DOI : [10.1016/j.cma.2018.07.012](https://doi.org/10.1016/j.cma.2018.07.012).
- [AMV18] P. F. ANTONIETTI, L. MASCOTTO et M. VERANI. « A multigrid algorithm for the p-version of the virtual element method ». In : *Mathematical Modelling and Numerical Analysis* 52.1 (2018), p. 337-364. DOI : [10.1051/m2an/2018007](https://doi.org/10.1051/m2an/2018007).
- [ANR17] O. ANDERSEN, H. M. NILSEN et X. RAYNAUD. « Virtual element method for geomechanical simulations of reservoir models ». In : *Computational Geosciences* 21.5 (2017), p. 877-893. DOI : [10.1007/s10596-017-9636-1](https://doi.org/10.1007/s10596-017-9636-1).
- [Art18] E. ARTIOLI. « Asymptotic homogenization of fibre-reinforced composites : a virtual element method approach ». In : *Meccanica* 53.6 (2018), p. 1187-1201. DOI : [10.1007/s11012-018-0818-2](https://doi.org/10.1007/s11012-018-0818-2).
- [BB17a] S. BERRONE et A. BORIO. « Orthogonal polynomials in badly shaped polygonal elements for the Virtual Element Method ». In : *Finite Elements in Analysis and Design* 129 (2017), p. 14-31. DOI : [10.1016/j.finel.2017.01.006](https://doi.org/10.1016/j.finel.2017.01.006).
- [BB17b] S. BERRONE et A. BORIO. « A residual a posteriori error estimate for the Virtual Element Method ». In : *Mathematical Models and Methods in Applied Sciences* 27.08 (2017), p. 1423-1458. DOI : [10.1142/S0218202517500233](https://doi.org/10.1142/S0218202517500233).
- [BBKT17] L. BERGER, R. BORDAS, D. KAY et S. TAVENER. « A stabilized finite element method for finite-strain three-field poroelasticity ». In : *Computational Mechanics* 60.1 (2017), p. 51-68. DOI : [10.1007/s00466-017-1381-8](https://doi.org/10.1007/s00466-017-1381-8).
- [BBM13] L. BEIRÃO DA VEIGA, F. BREZZI et L. MARINI. « Virtual Elements for Linear Elasticity Problems ». In : *SIAM Journal on Numerical Analysis* 51.2 (2013), p. 794-812. DOI : [10.1137/120874746](https://doi.org/10.1137/120874746).
- [BBM14] N. BELLOMO, F. BREZZI et G. MANZINI. « Recent techniques for PDE discretizations on polyhedral meshes ». In : *Mathematical Models and Methods in Applied Sciences* 24.08 (2014), p. 1453-1455. DOI : [10.1142/S0218202514030018](https://doi.org/10.1142/S0218202514030018).
- [BBMR14] L. BEIRÃO DA VEIGA, F. BREZZI, L. MARINI et A. RUSSO. « The hitchhiker's guide to the Virtual Element Method ». In : *Mathematical Models & Methods in applied sciences* 24.8 (2014), p. 1541-1573. DOI : [10.1142/S021820251440003X](https://doi.org/10.1142/S021820251440003X).
- [BBMR16a] L. BEIRÃO DA VEIGA, F. BREZZI, L. MARINI et A. RUSSO. «  $H(\text{div})$  and  $H(\text{curl})$ -conforming virtual element methods ». In : *Numerische Mathematik* 133.2 (2016), p. 303-332. DOI : [10.1007/s00211-015-0746-1](https://doi.org/10.1007/s00211-015-0746-1).
- [BBMR16b] L. BEIRÃO DA VEIGA, F. BREZZI, L. MARINI et A. RUSSO. « Serendipity Nodal VEM spaces ». In : *Computers & Fluids* 141 (2016), p. 2-12. DOI : [10.1016/j.compfluid.2016.02.015](https://doi.org/10.1016/j.compfluid.2016.02.015).

- [BDR17] L. BEIRÃO DA VEIGA, F. DASSI et A. RUSSO. « High-order Virtual Element Method on polyhedral meshes ». In : *Computers & Mathematics with Applications* 74.5 (2017), p. 1110-1122. DOI : [10.1016/j.camwa.2017.03.021](https://doi.org/10.1016/j.camwa.2017.03.021).
- [BE16] L. BEIRÃO DA VEIGA et A. ERN. « Preface ». In : *Mathematical Modelling and Numerical Analysis* 50.3 (2016), p. 633-634. DOI : [10.1051/m2an/2016034](https://doi.org/10.1051/m2an/2016034).
- [Bei+13] L. BEIRÃO DA VEIGA, F. BREZZI, A. CANGIANI, G. MANZINI, L. MARINI et A. RUSSO. « Basic principles of virtual element methods ». In : *Mathematical Models & Methods in applied sciences* 23.1 (2013), p. 199-214. DOI : [10.1142/S0218202512500492](https://doi.org/10.1142/S0218202512500492).
- [Ben02] M. BENZI. « Preconditioning Techniques for Large Linear Systems : A Survey ». In : *Journal of Computational Physics* 182.2 (2002), p. 418-477. DOI : [10.1006/jcph.2002.7176](https://doi.org/10.1006/jcph.2002.7176).
- [BFG07] L. BERGAMASCHI, M. FERRONATO et G. GAMBOLATI. « Novel preconditioners for the iterative solution to FE-discretized coupled consolidation equations ». In : *Computer Methods in Applied Mechanics and Engineering* 196.25 (2007), p. 2647-2656. DOI : [10.1016/j.cma.2007.01.013](https://doi.org/10.1016/j.cma.2007.01.013).
- [Bio41] M. A. BIOT. « General Theory of Three-Dimensional Consolidation ». In : *Journal of Applied Physics* 12.2 (1941), p. 155-164. DOI : [10.1063/1.1712886](https://doi.org/10.1063/1.1712886).
- [Bis14] J. BISHOP. « A displacement-based finite element formulation for general polyhedra using harmonic shape functions ». In : *International Journal for Numerical Methods in Engineering* 97.1 (2014), p. 1-31. DOI : [10.1002/nme.4562](https://doi.org/10.1002/nme.4562).
- [BK19] J. W. BOTH et U. KÖCHER. « Numerical Investigation on the Fixed-Stress Splitting Scheme for Biot's Equations : Optimality of the Tuning Parameter ». In : *Numerical Mathematics and Advanced Applications ENUMATH 2017*, sous la dir. de F. A. RADU et al. Cham : Springer International Publishing, 2019, p. 789-797. ISBN : 978-3-319-96415-7.
- [BLM14] L. BEIRÃO DA VEIGA, K. LIPNIKOV et G. MANZINI. *The Mimetic Finite Difference Method for Elliptic Problems*. T. 11. Modeling, Simulation & Applications. Springer, 2014. ISBN : 978-3-319-02662-6.
- [BLM15] L. BEIRÃO DA VEIGA, C. LOVADINA et D. MORA. « A Virtual Element Method for elastic and inelastic problems on polytope meshes ». In : *Computer Methods in Applied Mechanics and Engineering* 295 (2015), p. 327-346. DOI : [10.1016/j.cma.2015.07.013](https://doi.org/10.1016/j.cma.2015.07.013).
- [BLR17] L. BEIRÃO DA VEIGA, C. LOVADINA et A. RUSSO. « Stability analysis for the virtual element method ». In : *Mathematical Models and Methods in Applied Sciences* 27.13 (2017), p. 2557-2594. DOI : [10.1142/S021820251750052X](https://doi.org/10.1142/S021820251750052X).
- [BM12] L. BERGAMASCHI et Á. MARTÍNEZ. « RMCP : Relaxed Mixed Constraint Preconditioners for saddle point linear systems arising in geomechanics ». In : *Computer Methods in Applied Mechanics and Engineering* 221-222 (2012), p. 54-62. DOI : [10.1016/j.cma.2012.02.004](https://doi.org/10.1016/j.cma.2012.02.004).
- [BM13] F. BREZZI et L. MARINI. « Virtual Element Methods for plate bending problems ». In : *Computer Methods in Applied Mechanics and Engineering* 253 (2013), p. 455-462. DOI : [10.1016/j.cma.2012.09.012](https://doi.org/10.1016/j.cma.2012.09.012).
- [Bor+18] M. BORREGALES, K. KUMAR, F. A. RADU, C. RODRIGO et F. J. GASPAR. « A parallel-in-time fixed-stress splitting method for Biot's consolidation model ». In : *ArXiv e-prints* (2018), p. 1-19. arXiv : [1802.00949](https://arxiv.org/abs/1802.00949).

- [BPP17] S. BERTOLUZZA, M. PENNACCHIO et D. PRADA. « BDDC and FETI-DP for the virtual element method ». In : *Calcolo* 54.4 (2017), p. 1565-1593. DOI : [10.1007/s10092-017-0242-3](https://doi.org/10.1007/s10092-017-0242-3).
- [BQL17] S. C. BRENNER, G. QINGGUANG et S. LI-YENG. « Some Estimates for Virtual Element Methods ». In : *Computational Methods in Applied Mathematics* 17.4 (2017), p. 553-574. DOI : [10.1515/cmam-2017-0008](https://doi.org/10.1515/cmam-2017-0008).
- [BRKN18] M. BORREGALES, F. A. RADU, K. KUMAR et J. M. NORDBOTTEN. « Robust iterative schemes for non-linear poromechanics ». In : *Computational Geosciences* 22.4 (2018), p. 1021-1038. DOI : [10.1007/s10596-018-9736-6](https://doi.org/10.1007/s10596-018-9736-6).
- [BRV19] L. BEIRÃO DA VEIGA, A. RUSSO et G. VACCA. « The Virtual Element Method with curved edges ». In : *Mathematical Modelling and Numerical Analysis* 53.2 (2019), p. 375-404. DOI : [10.1051/m2an/2018052](https://doi.org/10.1051/m2an/2018052).
- [BS08] S. C. BRENNER et L. R. SCOTT. *The Mathematical Theory of Finite Element Methods*. 3<sup>e</sup> éd. T. 15. Texts in Applied Mathematics. Springer, 2008. ISBN : 978-0-387-75933-3.
- [BS18] S. C. BRENNER et L.-Y. SUNG. « Virtual element methods on meshes with small edges or faces ». In : *Mathematical Models and Methods in Applied Sciences* 28.07 (2018), p. 1291-1336. DOI : [10.1142/S0218202518500355](https://doi.org/10.1142/S0218202518500355).
- [BV16] D. L. BROWN et M. VASILYEVA. « A Generalized Multiscale Finite Element Method for poroelasticity problems I : Linear problems ». In : *Journal of Computational and Applied Mathematics* 294 (2016), p. 372-388. DOI : [10.1016/j.cam.2015.08.007](https://doi.org/10.1016/j.cam.2015.08.007).
- [BWHZ18] M. L. de BELLIS, P. WRIGGERS, B. HUDOBIVNIK et G. ZAVARISE. « Virtual element formulation for isotropic damage ». In : *Finite Elements in Analysis and Design* 144 (2018), p. 38-48. DOI : [10.1016/j.finel.2018.01.002](https://doi.org/10.1016/j.finel.2018.01.002).
- [Cal19] J. G. CALVO. « An overlapping Schwarz method for virtual element discretizations in two dimensions ». In : *Computers & Mathematics with Applications* 77.4 (2019), p. 1163-1177. DOI : [10.1016/j.camwa.2018.10.043](https://doi.org/10.1016/j.camwa.2018.10.043).
- [CBP17] H. CHI, L. BEIRÃO DA VEIGA et G. PAULINO. « Some basic formulations of the virtual element method (VEM) for finite deformations ». In : *Computer Methods in Applied Mechanics and Engineering* 318 (2017), p. 148-192. DOI : [10.1016/j.cma.2016.12.020](https://doi.org/10.1016/j.cma.2016.12.020).
- [CBP19] H. CHI, L. BEIRÃO DA VEIGA et G. H. PAULINO. « A simple and effective gradient recovery scheme and a posteriori error estimator for the Virtual Element Method (VEM) ». In : *Computer Methods in Applied Mechanics and Engineering* 347 (2019), p. 21-58. DOI : [10.1016/j.cma.2018.08.014](https://doi.org/10.1016/j.cma.2018.08.014).
- [CGL09] B. COCKBURN, J. GOPALAKRISHNAN et R. LAZAROV. « Unified Hybridization of Discontinuous Galerkin, Mixed, and Continuous Galerkin Methods for Second Order Elliptic Problems ». In : *SIAM Journal on Numerical Analysis* 47.2 (2009), p. 1319-1365. DOI : [10.1137/070706616](https://doi.org/10.1137/070706616).
- [CGPS17] A. CANGIANI, E. H. GEORGIOULIS, T. PRYER et O. J. SUTTON. « A posteriori error estimates for the virtual element method ». In : *Numerische Mathematik* 137.4 (2017), p. 857-893. DOI : [10.1007/s00211-017-0891-9](https://doi.org/10.1007/s00211-017-0891-9).
- [Chi18] C. CHINOSI. « Virtual Elements for the Reissner-Mindlin plate problem ». In : *Numerical Methods for Partial Differential Equations* 34.4 (2018), p. 1117-1144. DOI : [10.1002/num.22248](https://doi.org/10.1002/num.22248).

- [CJ59] H. CARSLAW et J. JAEGER. *Conduction of heat in solids*. 2<sup>e</sup> éd. Oxford : Oxford Univers. Press, 1959. ISBN : 978-0-19-853368-9.
- [CKHT19] N. CASTELLETTO, S. KLEVTSOV, H. HAJIBEYGI et H. A. TCHELEPI. « Multiscale two-stage solver for Biot's poroelasticity equations in subsurface media ». In : *Computational Geosciences* 23.2 (2019), p. 207-224. DOI : [10.1007/s10596-018-9791-z](https://doi.org/10.1007/s10596-018-9791-z).
- [CL18] J. CHOO et S. LEE. « Enriched Galerkin finite elements for coupled poromechanics with local mass conservation ». In : *Computer Methods in Applied Mechanics and Engineering* 341 (2018), p. 311-332. DOI : [10.1016/j.cma.2018.06.022](https://doi.org/10.1016/j.cma.2018.06.022).
- [CMRS15] A. CANGIANI, G. MANZINI, A. RUSSO et N. SUKUMAR. « Hourglass stabilization and the virtual element method ». In : *International Journal for Numerical Methods in Engineering* 102.3-4 (2015), p. 404-436. DOI : [10.1002/nme.4854](https://doi.org/10.1002/nme.4854).
- [Cou+18] J. COULET, I. FAILLE, V. GIRAULT, N. GUY et F. NATAF. « Fully coupled schemes using Virtual Element and Finite Volume Discretisations for Biot equation modelling ». In : *Proceedings, 16th European Conference on the Mathematics of Oil Recovery*. 2018. DOI : [10.3997/2214-4609.201802271](https://doi.org/10.3997/2214-4609.201802271).
- [Cou+19] J. COULET, I. FAILLE, V. GIRAULT, N. GUY et F. NATAF. « A fully coupled scheme using Virtual Element Method and Finite Volume for poroelasticity ». In : *Computational Geosciences* (2019), p. 1-23. DOI : [10.1007/s10596-019-09831-w](https://doi.org/10.1007/s10596-019-09831-w).
- [CPL01] S. H. CHAN, K. K. PHOON et F. H. LEE. « A modified Jacobi preconditioner for solving ill-conditioned Biot's consolidation equations using symmetric quasi-minimal residual method ». In : *International Journal for Numerical and Analytical Methods in Geomechanics* 25.10 (2001), p. 1001-1025. DOI : [10.1002/nag.164](https://doi.org/10.1002/nag.164).
- [CS99] X. CAI et M. SARKIS. « A Restricted Additive Schwarz Preconditioner for General Sparse Linear Systems ». In : *SIAM Journal on Scientific Computing* 21.2 (1999), p. 792-797. DOI : [10.1137/S106482759732678X](https://doi.org/10.1137/S106482759732678X).
- [CTP06] X. CHEN, K. C. TOH et K. K. PHOON. « A modified SSOR preconditioner for sparse symmetric indefinite linear systems of equations ». In : *International Journal for Numerical Methods in Engineering* 65.6 (2006), p. 785-807. DOI : [10.1002/nme.1461](https://doi.org/10.1002/nme.1461).
- [DE12] D. A. DI PIETRO et A. ERN. *Mathematical Aspects of Discontinuous Galerkin Methods*. T. 69. Mathématiques et Applications. Springer, 2012. ISBN : 978-3-642-22979-4.
- [DEL16] D. A. DI PIETRO, A. ERN et S. LEMAIRE. « A Review of Hybrid High-Order Methods : Formulations, Computational Aspects, Comparison with Other Methods ». In : *Building Bridges : Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, sous la dir. de G. R. BARRENECHEA et al. Springer, 2016, p. 205-236. ISBN : 978-3-319-41640-3.
- [DGSM06] R. H. DEAN, X. GIA, C. M. STONE et S. E. MINKOFF. « A Comparison of Techniques for Coupling Porous Flow and Geomechanics ». In : *SPE Journal* 11.1 (2006), p. 132-140. DOI : [10.2118/79709-PA](https://doi.org/10.2118/79709-PA).
- [DGW18] S. DANA, B. GANIS et M. F. WHEELER. « A multiscale fixed stress split iterative scheme for coupled flow and poromechanics in deep subsurface reservoirs ». In : *Journal of Computational Physics* 352 (2018), p. 1-22. DOI : [10.1016/j.jcp.2017.09.049](https://doi.org/10.1016/j.jcp.2017.09.049).



- [DJN15] V. DOLEAN, P. JOLIVET et F. NATAF. *An introduction to domain decomposition methods*. Algorithms, theory, and parallel implementation. Society for Industrial et Applied Mathematics (SIAM), Philadelphia, PA, 2015. ISBN : 978-1-611974-05-8.
- [DN17] J. DRONIOU et N. NATARAJ. « Improved  $L^2$  estimate for gradient schemes and super-convergence of the TPGA finite volume scheme ». In : *IMA Journal of Numerical Analysis* 38.3 (2017), p. 1254-1293. DOI : [10.1093/imanum/drx028](https://doi.org/10.1093/imanum/drx028).
- [DN19] V. DHANUSH et S. NATARAJAN. « Implementation of the virtual element method for coupled thermo-elasticity in Abaqus ». In : *Numerical Algorithms* 80.3 (2019), p. 1037-1058. DOI : [10.1007/s11075-018-0516-0](https://doi.org/10.1007/s11075-018-0516-0).
- [Dro14] J. DRONIOU. « Finite volume schemes for diffusion equations : Introduction to and review of modern methods ». In : *Mathematical Models & Methods in Applied Sciences* 24.8 (2014), p. 1575-1619. DOI : [10.1142/S0218202514400041](https://doi.org/10.1142/S0218202514400041).
- [DS80] T. DUPONT et R. SCOTT. « Polynomial Approximation of Functions in Sobolev Spaces ». In : *Mathematics of Computation* 34.150 (1980), p. 441-463. DOI : [10.1090/S0025-5718-1980-0559195-7](https://doi.org/10.1090/S0025-5718-1980-0559195-7).
- [DVS98] J. DELLINGER, D. VASICEK et C. SONDERGELD. « Kelvin Notation for Stabilizing Elastic-Constant Inversion ». In : *OGST – Revue d'IFP Energies Nouvelles* 53.5 (1998), p. 709-719. DOI : [10.2516/ogst:1998063](https://doi.org/10.2516/ogst:1998063).
- [EGH00] R. EYMARD, T. GALLOUËT et R. HERBIN. « Finite volume methods ». In : *Solution of Equation in  $\mathbb{R}^n$  (Part 3), Techniques of Scientific Computing (Part 3)*. T. 7. Handbook of Numerical Analysis Supplement C. Elsevier, 2000, p. 713-1018. ISBN : 978-0-444-50350-3.
- [EGH09] R. EYMARD, T. GALLOUËT et R. HERBIN. « Discretization of heterogeneous and anisotropic diffusion problems on general nonconforming meshes SUSHI : a scheme using stabilization and hybrid interfaces ». In : *IMA Journal of Numerical Analysis* 30.4 (2009), p. 1009-1043. DOI : [10.1093/imanum/drn084](https://doi.org/10.1093/imanum/drn084).
- [FGT01] M. FERRONATO, G. GAMBOLATI et P. TEATINI. « Ill-conditioning of finite element poroelasticity equations ». In : *International Journal of Solids and Structures* 38.34 (2001), p. 5995-6014. DOI : [10.1016/S0020-7683\(00\)00352-8](https://doi.org/10.1016/S0020-7683(00)00352-8).
- [FV05] J. M. FIGUEIREDO et J. VIANO. *Finite elements  $Q_1$ -Lagrange for the linear elasticity problem*. Rapp. tech. 2005. eprint : <http://hdl.handle.net/1822/5198>.
- [Gano8] M. J. GANDER. « Schwarz methods over the course of time ». In : *Electronic Transactions on Numerical Analysis* 31 (2008), p. 228-255. eprint : <https://zbmath.org/?q=1171.65020>.
- [GPDM15] A. L. GAIN, G. H. PAULINO, L. S. DUARTE et I. F. MENEZES. « Topology optimization using polytopes ». In : *Computer Methods in Applied Mechanics and Engineering* 293 (2015), p. 411-430. DOI : [10.1016/j.cma.2015.05.007](https://doi.org/10.1016/j.cma.2015.05.007).
- [GPF03] G. GAMBOLATI, G. PINI et M. FERRONATO. « Scaling improves stability of preconditioned CG-like solvers for FE consolidation equations ». In : *International Journal for Numerical and Analytical Methods in Geomechanics* 27.12 (2003), p. 1043-1056. DOI : [10.1002/nag.311](https://doi.org/10.1002/nag.311).
- [GTP14] A. L. GAIN, C. TALISCHI et G. H. PAULINO. « On the Virtual Element Method for three-dimensional linear elasticity problems on arbitrary polyhedral meshes ». In : *Computer Methods in Applied Mechanics and Engineering* 282 (2014), p. 132-160. DOI : [10.1016/j.cma.2014.05.005](https://doi.org/10.1016/j.cma.2014.05.005).

- [GWAD19] V. GIRAULT, M. F. WHEELER, T. ALMANI et S. DANA. « A priori error estimates for a discretized poro-elastic-elastic system solved by a fixed-stress algorithm ». In : *OGST – Revue d'IFP Energies Nouvelles* 74 (2019), p. 24. DOI : [10.2516/ogst/2018071](https://doi.org/10.2516/ogst/2018071).
- [HAW19] B. HUDOBIVNIK, F. ALDAKHEEL et P. WRIGGERS. « A low order 3D virtual element formulation for finite elasto-plastic deformations ». In : *Computational Mechanics* 63.2 (2019), p. 253-269. DOI : [10.1007/s00466-018-1593-6](https://doi.org/10.1007/s00466-018-1593-6).
- [Hec12] F. HECHT. « New development in FreeFem++ ». In : *Journal of Numerical Mathematics* 20.3-4 (2012), p. 251-265. DOI : [10.1515/jnum-2012-0013](https://doi.org/10.1515/jnum-2012-0013).
- [HOL12] J. B. HAGA, H. OSNES et H. P. LANGTANGEN. « A parallel block preconditioner for large-scale poroelasticity with highly heterogeneous material parameters ». In : *Computational Geosciences* 16.3 (2012), p. 723-734. DOI : [10.1007/s10596-012-9284-4](https://doi.org/10.1007/s10596-012-9284-4).
- [HRGZ17] X. HU, C. RODRIGO, F. J. GASPAR et L. T. ZIKATANOV. « A nonconforming finite element method for the Biot's consolidation model in poroelasticity ». In : *Journal of Computational and Applied Mathematics* 310 (2017), p. 143-154. DOI : [10.1016/j.cam.2016.06.003](https://doi.org/10.1016/j.cam.2016.06.003).
- [Hus+19] A. HUSSEIN, F. ALDAKHEEL, B. HUDOBIVNIK, P. WRIGGERS, P.-A. GUIDAULT et O. ALLIX. « A computational framework for brittle crack-propagation based on efficient virtual element method ». In : *Finite Elements in Analysis and Design* 159 (2019), p. 15-32. DOI : [10.1016/j.finel.2019.03.001](https://doi.org/10.1016/j.finel.2019.03.001).
- [JHNP13] P. JOLIVET, F. HECHT, F. NATAF et C. PRUD'HOMME. « Scalable Domain Decomposition Preconditioners for Heterogeneous Elliptic Problems ». In : *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*. Denver, Colorado : ACM, 2013, p. 1-11. ISBN : 978-1-4503-2378-9.
- [KTJ11a] J. KIM, H. TCHELEPI et R. JUANES. « Stability and convergence of sequential methods for coupled flow and geomechanics : Drained and undrained splits ». In : *Computer Methods in Applied Mechanics and Engineering* 200.23 (2011), p. 2094-2116. DOI : [10.1016/j.cma.2011.02.011](https://doi.org/10.1016/j.cma.2011.02.011).
- [KTJ11b] J. KIM, H. TCHELEPI et R. JUANES. « Stability and convergence of sequential methods for coupled flow and geomechanics : Fixed-stress and fixed-strain splits ». In : *Computer Methods in Applied Mechanics and Engineering* 200.13 (2011), p. 1591-1606. DOI : [10.1016/j.cma.2010.12.022](https://doi.org/10.1016/j.cma.2010.12.022).
- [LC90] J. LEMAITRE et J.-L. CHABOCHE. *Mechanics of Solid Materials*. Cambridge University Press, 1990. ISBN : 978-1-13916797-0.
- [LDNo7] G. LIU, K. DAI et T. NGUYEN. « A Smoothed Finite Element Method for Mechanics Problems ». In : *Computational Mechanics* 39.6 (2007), p. 859-877. DOI : [10.1007/s00466-006-0075-4](https://doi.org/10.1007/s00466-006-0075-4).
- [Lee16] J. J. LEE. « Robust Error Analysis of Coupled Mixed Methods for Biot's Consolidation Model ». In : *Journal of Scientific Computing* 69.2 (2016), p. 610-632. DOI : [10.1007/s10915-016-0210-0](https://doi.org/10.1007/s10915-016-0210-0).
- [Lem13] S. LEMAIRE. « Nonconforming discretizations of a poromechanical model on general meshes ». Thèse de doct. Université Paris-Est, 2013. HAL : [tel-00957292](https://hal.archives-ouvertes.fr/tel-00957292).
- [Lio5] X. S. LI. « An Overview of SuperLU : Algorithms, Implementation, and User Interface ». In : *ACM Transactions on Mathematical Software* 31.3 (2005), p. 302-325. DOI : [10.1145/1089014.1089017](https://doi.org/10.1145/1089014.1089017).

- [LJ19] Q. LI et L. JIANG. « A multiscale virtual element method for elliptic problems in heterogeneous porous media ». In : *Journal of Computational Physics* 388 (2019), p. 394-415. DOI : [10.1016/j.jcp.2019.03.031](https://doi.org/10.1016/j.jcp.2019.03.031).
- [LRGO15] P. LUO, C. RODRIGO, F. J. GASPAR et C. W. OOSTERLEE. « Multigrid method for nonlinear poroelasticity equations ». In : *Computing and Visualization in Science* 17.5 (2015), p. 255-265. DOI : [10.1007/s00791-016-0260-8](https://doi.org/10.1007/s00791-016-0260-8).
- [Man93] J. MANDEL. « Balancing domain decomposition ». In : *Communications in Numerical Methods in Engineering* 9.3 (1993), p. 233-241. DOI : [10.1002/cnm.1640090307](https://doi.org/10.1002/cnm.1640090307).
- [Mar12] B. MARTIN. « Computational 2D/3D finite volume solvers applied to linear elasticity ». Thèse de doct. École normale supérieure de Cachan, 2012. HAL : [tel-00798769](https://hal.archives-ouvertes.fr/tel-00798769).
- [Mas18] L. MASCOTTO. « Ill-conditioning in the virtual element method : Stabilizations and bases ». In : *Numerical Methods for Partial Differential Equations* 34.4 (2018), p. 1258-1281. DOI : [10.1002/num.22257](https://doi.org/10.1002/num.22257).
- [Mer16] A. MERXHANI. « An introduction to linear poroelasticity ». In : *ArXiv e-prints* (2016), p. 1-38. arXiv : [1607.04274](https://arxiv.org/abs/1607.04274).
- [ML94] M. A. MURAD et A. F. D. LOULA. « On stability and convergence of finite element approximations of Biot's consolidation problem ». In : *International Journal for Numerical Methods in Engineering* 37.4 (1994), p. 645-667. DOI : [10.1002/nme.1620370407](https://doi.org/10.1002/nme.1620370407).
- [MW13] A. MIKELIĆ et M. F. WHEELER. « Convergence of iterative coupling for coupled flow and geomechanics ». In : *Computational Geosciences* 17.3 (2013), p. 455-461. DOI : [10.1007/s10596-012-9318-y](https://doi.org/10.1007/s10596-012-9318-y).
- [NNR18] H. M. NILSEN, J. NORDBOTTEN et X. RAYNAUD. « Comparison between cell-centered and nodal-based discretization schemes for linear elasticity ». In : *Computational Geosciences* 22.1 (2018), p. 233-260. DOI : [10.1007/s10596-017-9687-3](https://doi.org/10.1007/s10596-017-9687-3).
- [Nor14] J. M. NORDBOTTEN. « Finite volume hydromechanical simulation in porous media ». In : *Water Resources Research* 50.5 (2014), p. 4379-4394. DOI : [10.1002/2013WR015179](https://doi.org/10.1002/2013WR015179).
- [OGP17] Z. OURAGA, N. GUY et A. POUYA. « Modeling of natural fracture initiation and propagation in basin sedimentation context ». In : *Journal of Geophysical Research : Solid Earth* 122.1 (2017), p. 247-261. DOI : [10.1002/2016JB013511](https://doi.org/10.1002/2016JB013511).
- [PG15] G. H. PAULINO et A. L. GAIN. « Bridging art and engineering using Escher-based virtual elements ». In : *Structural and Multidisciplinary Optimization* 51.4 (2015), p. 867-883. DOI : [10.1007/s00158-014-1179-7](https://doi.org/10.1007/s00158-014-1179-7).
- [Pon89] D. K. PONTING. « Corner Point Grid Geometry in Reservoir Simulation ». In : *Proceedings, First European Conference on the Mathematics of Oil Recovery*. 1989. DOI : [10.3997/2214-4609.201411305](https://doi.org/10.3997/2214-4609.201411305).
- [PW07] P. J. PHILLIPS et M. F. WHEELER. « A coupling of mixed and continuous Galerkin finite element methods for poroelasticity I : the continuous in time case ». In : *Computational Geosciences* 11.2 (2007), p. 131-144. DOI : [10.1007/s10596-007-9045-y](https://doi.org/10.1007/s10596-007-9045-y).
- [PW09] P. J. PHILLIPS et M. F. WHEELER. « Overcoming the problem of locking in linear elasticity and poroelasticity : an heuristic approach ». In : *Computational Geosciences* 13.1 (2009), p. 5-12. DOI : [10.1007/s10596-008-9114-x](https://doi.org/10.1007/s10596-008-9114-x).

- [QDB11] R. QUEY, P. DAWSON et F. BARBE. « Large-scale 3D random polycrystals for the finite element method : Generation, meshing and remeshing ». In : *Computer Methods in Applied Mechanics and Engineering* 200 (2011), p. 1729-1745. DOI : [10.1016/j.cma.2011.01.002](https://doi.org/10.1016/j.cma.2011.01.002).
- [RBLE19] F. RIVAROLA, M. BENEDETTO, N. LABANDA et G. ETSE. « A multiscale approach with the Virtual Element Method : Towards a VE2 setting ». In : *Finite Elements in Analysis and Design* 158 (2019), p. 1-16. DOI : [10.1016/j.finel.2019.01.011](https://doi.org/10.1016/j.finel.2019.01.011).
- [Rou16] M. ROUXEL-LABBE. « Anisotropic mesh generation ». Thèse de doct. Université Côte d'Azur, 2016. HAL : [tel-01419457](https://hal.archives-ouvertes.fr/tel-01419457).
- [RTT17] B. RIVIÈRE, J. TAN et T. THOMPSON. « Error analysis of primal discontinuous Galerkin methods for a mixed formulation of the Biot equations ». In : *Computers & Mathematics with Applications* 73.4 (2017), p. 666-683. DOI : [10.1016/j.camwa.2016.12.030](https://doi.org/10.1016/j.camwa.2016.12.030).
- [Saa03] Y. SAAD. *Iterative Methods for Sparse Linear Systems*. 2<sup>e</sup> éd. Society for Industrial et Applied Mathematics, 2003. ISBN : 978-0-89871-534-7.
- [Sh000] R. SHOWALTER. « Diffusion in Poro-Elastic Media ». In : *Journal of Mathematical Analysis and Applications* 251.1 (2000), p. 310-340. DOI : [10.1006/jmaa.2000.7048](https://doi.org/10.1006/jmaa.2000.7048).
- [SMo6] N. SUKUMAR et E. MALSCH. « Recent advances in the construction of polygonal finite element interpolants ». In : *Archives of Computational Methods in Engineering* 13.1 (2006), p. 129. DOI : [10.1007/BF02905933](https://doi.org/10.1007/BF02905933).
- [SM94] A. SETTARI et F. MOURITS. « Coupling of geomechanics and reservoir simulation models ». In : *Proceedings, 8th International Conference on Computer Methods and Advances in Geomechanics*. T. 3. Balkema, 1994, p. 2151-2158. ISBN : 978-90-5410-380-6.
- [Spi+14] N. SPILLANE, V. DOLEAN, P. HAURET, F. NATAF, C. PECHSTEIN et R. SCHEICHL. « Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps ». In : *Numerische Mathematik* 126.4 (2014), p. 741-770. DOI : [10.1007/s00211-013-0576-y](https://doi.org/10.1007/s00211-013-0576-y).
- [Spi14] N. SPILLANE. « Robust domain decomposition methods for symmetric positive definite problems ». Thèse de doct. Université Pierre et Marie Curie - Paris VI, 2014. HAL : [tel-00958252](https://hal.archives-ouvertes.fr/tel-00958252).
- [Sut17] O. J. SUTTON. « The virtual element method in 50 lines of MATLAB ». In : *Numerical Algorithms* 75.4 (2017), p. 1141-1159. DOI : [10.1007/s11075-016-0235-3](https://doi.org/10.1007/s11075-016-0235-3).
- [Ter43] K. TERZAGHI. *Theoretical Soil Mechanics*. John Wiley & Sons Inc, 1943. ISBN : 978-0-471-85305-3.
- [TG51] S. TIMOSHENKO et J. GOODIER. *Theory of Elasticity*. Second. McGraw-Hill Book Company, 1951. ISBN : 978-0-07-064270-6.
- [Wac10] E. WACHSPRESS. « Rational bases for convex polyhedra ». In : *Computers & Mathematics with Applications* 59.6 (2010), p. 1953-1956. DOI : [10.1016/j.camwa.2009.11.013](https://doi.org/10.1016/j.camwa.2009.11.013).
- [WB11] J. A. WHITE et R. I. BORJA. « Block-preconditioned Newton–Krylov solvers for fully coupled flow and geomechanics ». In : *Computational Geosciences* 15.4 (2011), p. 647. DOI : [10.1007/s10596-011-9233-7](https://doi.org/10.1007/s10596-011-9233-7).
- [WCT16] J. A. WHITE, N. CASTELLETTO et H. A. TCHELEPI. « Block-partitioned solvers for coupled poromechanics : A unified framework ». In : *Computer Methods in Applied Mechanics and Engineering* 303 (2016), p. 55-74. DOI : [10.1016/j.cma.2016.01.008](https://doi.org/10.1016/j.cma.2016.01.008).

- [WRR16] P. WRIGGERS, W. RUST et B. REDDY. « A virtual element method for contact ». In : *Computational Mechanics* 58.6 (2016), p. 1039-1050. DOI : [10.1007/s00466-016-1331-x](https://doi.org/10.1007/s00466-016-1331-x).
- [WY14] J. WANG et X. YE. « A weak Galerkin mixed finite element method for second order elliptic problems ». In : *Mathematics of Computation* 83.289 (2014), p. 2101-2126. DOI : [10.1090/S0025-5718-2014-02852-4](https://doi.org/10.1090/S0025-5718-2014-02852-4).
- [YS08] G. YUAN et Z. SHENG. « Monotone finite volume schemes for diffusion equations on polygonal meshes ». In : *Journal of Computational Physics* 227.12 (2008), p. 6288-6312. DOI : [10.1016/j.jcp.2008.03.007](https://doi.org/10.1016/j.jcp.2008.03.007).
- [ZF18] B. ZHANG et M. FENG. « Virtual element method for two-dimensional linear elasticity problem in mixed weakly symmetric formulation ». In : *Applied Mathematics and Computation* 328 (2018), p. 1-25. DOI : [10.1016/j.amc.2018.01.023](https://doi.org/10.1016/j.amc.2018.01.023).
- [Zob07] M. D. ZOBACK. *Reservoir Geomechanics*. Cambridge University Press, 2007. ISBN : 978-0-521-14619-7.

---

**Titre :** Méthode des Éléments Virtuels pour le calcul de la déformation mécanique couplée aux écoulements en milieux poreux

**Résumé :** Cette thèse porte sur le développement et l'étude de méthodes numériques pour le problème de poroélasticité, modélisé par les équations de Biot. La discrétisation historique de ce système consiste à utiliser une méthode de type éléments finis pour traiter l'équation d'équilibre mécanique, et une méthode de type volumes finis pour traiter l'équation de conservation de la masse du fluide. Or, en géosciences, les procédures de maillage utilisées pour reproduire les propriétés géométriques des milieux telles que les hétérogénéités, discontinuités ou failles produisent des cellules incompatibles avec la discrétisation par éléments finis. En conséquence, on s'intéresse à la récente méthode des éléments virtuels qui paraît capable de traiter des maillages plus généraux, et qui permettrait donc l'utilisation d'un unique maillage pour les deux modélisations. Après avoir appliqué la méthode des éléments virtuels au problème d'élasticité, on effectue le couplage avec l'écoulement discrétisé par volumes finis. Selon l'approximation choisie pour cette méthode, deux schémas sont créés : le schéma VEM-TPFA, qui utilise l'approximation volumes finis la plus simple et permet d'en proposer une analyse mathématique, et le schéma VEM-MPFA qui se base sur une approximation plus complète de manière à traiter des maillages plus généraux. En plus de la discrétisation, on s'intéresse à la stratégie de résolution du problème couplé dans le but d'en améliorer l'efficacité et la robustesse, en particulier en parallèle. On considère ainsi la stratégie fixed-stress split qui permet de se ramener à la résolution de deux sous-problèmes, pour lesquels on exhibe des solveurs performants.

**Mots clés :** maillages polyédriques, poroélasticité, discrétisation, schémas couplés, systèmes linéaires, géomécanique

---

**Title:** Coupling Virtual Element Methods and finite volume schemes for computational geomechanics

**Abstract:** This manuscript focuses on the design and the study of numerical schemes for the poroelasticity problem modeled through Biot's equations. The classical way to numerically solve this system is to use a finite element method for the mechanical equilibrium equation and a finite volume method for the fluid mass conservation equation. However, to capture specific properties of the underground medium such as heterogeneities, discontinuities, and faults, meshing procedures commonly lead to badly shaped cells for finite element-based modeling. Consequently, we investigate the use of the recent virtual element method which appears as a potential discretization method for the mechanical part and could therefore allow the use of a unique mesh for both the mechanical and fluid flow modeling. Starting from a first insight into virtual element method applied to the elastic problem in the context of geomechanical simulations, we apply in addition a finite volume method to take care of the fluid conservation equation. Depending on the selected finite volume method, we create two different coupled schemes: the VEM-TPFA scheme, using a two-point flux approximation for the fluid flow and for which we provide a mathematical analysis, and the VEM-MPFA scheme, which uses a multipoint flux approximation and can thus handle more general meshes. In addition to the discretization, we also investigate the way to solve the coupled problem in order to improve efficiency and robustness in the context of parallel computing. For this purpose, we consider the fixed-stress split strategy in order to fall back to the resolution of two subproblems for which we can use efficient solvers.

**Keywords:** polyhedral grids, poroelasticity, discretization, coupled schemes, linear systems, geomechanics

---