Université de Liège
Faculté des Sciences Appliquées

Modélisation multiphysique de l'écoulement des glaciers : analyse et résolution numérique d'un problème couplé non linéaire

BULTHUIS Kevin

Travail de fin d'études réalisé en vue de l'obtention du grade de
Master Ingénieur Civil Physicien

Promoteur :
ARNST Maarten

Année académique : **2014-2015**

Remerciements

Je tiens tout d'abord à remercier sincèrement mon promoteur, Monsieur Maarten Arnst, qui m'a permis de mener à bien l'ensemble de ce travail. Son encadrement, sa disponibilité ainsi que ses nombreuses remarques pertinentes ont donné à ce mémoire un intérêt tout particulier.

Je désire ensuite remercier Messieurs Romain Boman, Christophe Geuzaine et Lionel Favier d'avoir accepté de constituer le jury de ce travail de fin d'études. Je leur souhaite une lecture aussi agréable que possible.

Je tiens également à remercier l'ensemble des personnes qui ont contribué de près ou de loin à la relecture de ce mémoire ainsi que ma famille et mes amis qui m'ont soutenu pour la réalisation de ce travail.

Modélisation multiphysique de l'écoulement des glaciers : analyse et résolution numérique d'un problème couplé non linéaire

BULTHUIS Kevin – Master en ingénieur civil physicien – 2014-2015

Résumé

La modélisation multiphysique s'intéresse au couplage entre différents phénomènes physiques (mécanique, thermique, électromagnétique, hydrodynamique, ...) ou encore à l'interaction entre différentes composantes individuelles. Celle-ci tend à proposer des stratégies de résolution efficaces pour tenir compte des difficultés liées au couplage entre différentes composantes d'un système complexe. La modélisation multiphysique connaît un intérêt sans cesse croissant, car de nombreux problèmes actuels qui intéressent la communauté scientifique sont décrits à partir de différentes composantes monophysiques. Parmi ces problématiques actuelles se trouvent l'étude et la modélisation de l'écoulement des glaciers. L'écoulement de la glace est traditionnellement étudié à partir d'un modèle thermomécanique couplé qui décrit l'écoulement d'un fluide visqueux non linéaire soumis à la gravité et dont la viscosité dépend de manière importante de la température. À partir de cet exemple physique concret, l'objectif de ce travail de fin d'études sera d'établir les fondements physiques, mathématiques et numériques sur lesquels pourraient reposer des études ultérieures dans le domaine de la glaciologie et plus généralement dans le cadre de la simulation multiphysique. Pour atteindre cet objectif, ce travail se divisera en trois parties principales. La première partie abordera la description et la modélisation physiques des glaciers. La deuxième partie est consacrée à une étude théorique des méthodes numériques utilisées pour résoudre les problèmes mécanique et thermique individuels. Les méthodes classiques de résolution par éléments finis seront présentées ainsi que de nouvelles perspectives de recherches. La troisième partie s'intéresse à l'application de méthodes numériques multiphysiques au cas d'un glacier alpin modèle. Les méthodes numériques traditionnelles sur lesquelles sont construites les nouvelles méthodes multiphysiques seront présentées théoriquement et comparées au moyen de simulations numériques. Cette troisième partie mettra en évidence le potentiel des méthodes fortement couplées comme moyen efficace pour la résolution de problèmes multiphysiques.

Mots-clés : modélisation multiphysique, modélisation numérique, problème couplé thermomécanique, non-linéarités, glaciologie.

Multiphysics modelling of glacier flow : analysis and numerical solution of a nonlinear coupled problem

BULTHUIS Kevin – Master in engineering physics – 2014-2015

Multiphysics modelling is concerned with systems that couple multiple physical phenomena (mechanics, heat transfer, electromagnetism, hydrodynamics, ...) as well as with the interaction between multiple individual components. Multiphysics modelling is interested in finding efficient numerical methods to deal with the coupling between the different components of a complex system. Interest in multiphysics simulations is steadily increasing as numerous current scientific problems involve multiple physical components. Among current issues, ice-sheet and glacier numerical models have become essential to predict the evolution of ice sheets as a result of global warming. Ice flow is usually described as a gravity-driven creep flow of a nonlinearly viscous fluid. Detailed glacier models need to account for temperature dependence of ice viscosity, which leads to a thermomechanical coupled model for ice flows. This final-year project addresses the analysis of a stationary thermomechanical glacier model as an example of a multiphysics problem. The goal will be to provide a physical, mathematical and numerical framework for future development in glaciology and in multiphysics simulations more generally. Thus this thesis is divided into three main parts. The first part addresses the physical description and modelling of glaciers. The second part is devoted to the theoretical study of numerical methods for the mechanical and thermal problems. Classical finite element methods will be presented as well as new research perspectives. The third part is concerned with the application of these multiphysics methods to an alpine glacier model. This third part highlights the importance of tightly coupled methods as new opportunities for multiphysics simulations.

Key-words: multiphysics modelling, numerical simulations, thermomechanical coupled problem, nonlinearity, glaciology.

Table des matières

Nomenclature

Grandeurs physiques (lettres grecques)

Γ_b	Interface glace-roche
Γ_s	Interface glace-air
Γ_t	Faces latérales du glacier
$\dot{\varepsilon}$	Tenseur taux de déformation
μ	Viscosité dynamique de la glace
ρ	Masse volumique de la glace
σ	Tenseur des contraintes
σ^D	Tenseur des contraintes déviatorique
σ_e	Contrainte effective

Grandeurs physiques (lettres latines)

a_b	Taux de fonte basale
c	Capacité thermique massique de la glace
C_b	Coefficient de frottement
d_e	Taux de déformation effectif
$\mathbf{g}$	Accélération de pesanteur
k	Conductivité thermique de la glace
L	Chaleur latente de fusion de la glace
$\mathbf{n}$	Normale unitaire extérieure
p	Pression mécanique
q_{geo}	Flux géothermique
$\mathbf{t}_b$	Contrainte basale tangentielle
T	Température dans le glacier
T_m	Température de fusion de la glace
T_s	Température à la surface du glacier
$\mathbf{v}$	Vitesse dans le glacier
$\mathbf{v}_b$	Vitesse basale tangentielle

Notations mathématiques

Γ	Frontière de Ω
Γ_D	Partie de Γ avec conditions aux limites de Dirichlet
Γ_N	Partie de Γ avec conditions aux limites de Neumann
Ω	Ouvert de $\mathbb{R}^n$
$\overline{\Omega}$	Adhérence de Ω
$\mathbb{C}^k(\Omega)$	Espace des fonctions k fois continûment dérivables sur Ω
$\mathbb{H}(\Omega)$	Espace de Hilbert sur Ω
$\mathbb{L}^p(\Omega)$	Espace de Lebesgue sur Ω
$\mathbb{W}^{1,p}(\Omega)$	Espace de Sobolev sur Ω
$\underline{\underline{I}}$	Tenseur identité d'ordre 2

Introduction

Le monde physique qui nous entoure se caractérise par un haut degré de complexité. Le sociologue français Pierre Bourdieu disait souvent que la réalité est complexe et que sa description devait l'être également de manière complexe. Pourtant pendant longtemps, la science et l'ingénierie traditionnelles ont eu tendance à simplifier de manière importante l'étude des phénomènes physiques. Parmi les simplifications opérées, il était coutume d'étudier les processus physiques sous l'angle d'une unique discipline traditionnelle de la physique. Les processus physiques étaient ainsi décrits par exemple dans le cadre de la mécanique des solides, de la mécanique des fluides, de la thermique ou encore de l'électromagnétisme. Cette description des phénomènes physiques à partir de processus monophysiques se révèle avantageuse au niveau des traitements analytique et numérique. Néanmoins, la plupart des phénomènes physiques d'intérêt actuel ne peuvent plus être décrits uniquement à partir d'un seul processus physique. C'est ainsi qu'une approche multiphysique de la réalité qui décrit les phénomènes du point de vue du couplage de processus monophysiques se révèle plus que jamais nécessaire. L'approche multiphysique présente une gamme d'applications très étendue parmi lesquelles nous citons les interactions fluide-structure, les couplages thermomécaniques lors de la mise en forme de pièces mécaniques, le couplage électromécanique dans les microsystèmes électromécaniques ou encore les interactions océan-atmosphère dans les modèles climatiques. Outre ces exemples intuitifs du concept de multiphysique, l'appellation multiphysique s'avère en fait très générale et peut également s'étendre à la description d'une quantité physique par des équations de nature mathématique différente, à une équation qui comprend des termes de nature physique différente ou encore à la résolution de problèmes à différentes échelles.

Une autre simplification des phénomènes généralement réalisée est de considérer que ceux-ci sont décrits par des lois linéaires. Cette hypothèse de linéarité est à l'origine de la loi d'Ohm pour la conduction électrique, de la loi de Fourier pour la conduction thermique ou encore de la loi rhéologique de Newton en mécanique des fluides. Pourtant comme le disait Einstein, les vraies lois de la nature ne peuvent pas être linéaires. Dès lors, malgré le succès indéniable des lois de comportement linéaires, celles-ci doivent laisser place à des lois non linéaires hors de leur domaine de validité comme c'est le cas pour la description des fluides non newtoniens. En réalité, les non-linéarités dans les systèmes peuvent se manifester de nombreuses façons (non-linéarité comportementale, non-linéarité des conditions aux limites, ...) et la prise en compte mathématique et numérique de celles-ci nécessite des outils adéquats.

Ce travail de fin d'études a été motivé par la lecture d'un article de Keyes *et al.* intitulé *« Multiphysics simulations : Challenges and opportunities »* [42] dans lequel de nouvelles stratégies pour la résolution des problèmes multiphysiques sont mises en évidence. L'objectif de ce mémoire sera d'établir les fondements physiques, mathématiques et numériques à partir desquels des études ultérieures dans le domaine de la modélisation multiphysique pourraient être menées. Afin de rendre notre travail plus concret, nous aborderons cette question dans le cadre de l'étude et de la modélisation de l'écoulement des glaciers. L'étude des glaciers constitue depuis quelques décennies un domaine de recherche de plus en plus étudié dans la communauté scientifique en raison notamment des préoccupations actuelles environnementales. L'écoulement de la glace est généralement décrit comme celui d'un fluide fortement

visqueux dont le comportement rhéologique est celui d'un fluide non newtonien. La viscosité de la glace présente une dépendance marquée envers sa température. Une description précise du comportement mécanique de la glace nécessite ainsi la résolution complémentaire d'une équation pour la température. Ceci nous amène naturellement à une description des glaciers à partir d'un problème thermomécanique pour lequel les parties mécanique et thermique s'influencent mutuellement.

La tendance traditionnelle en glaciologie a été pendant longtemps d'étudier l'écoulement des glaciers à partir de modèles simplifiés parmi lesquels les approximations de glace peu profonde (*shallow ice approximation*) et du premier ordre (*first order approximation*) [30]. Toutefois, ces modèles sont insuffisants pour décrire le comportement des glaciers dans des zones critiques telles que la ligne d'ancrage qui sépare une calotte glaciaire de sa plate-forme de glace flottante [17]. La tendance actuelle évolue donc progressivement vers le développement et l'implémentation de modèles plus complets. L'utilisation de ces modèles nécessite de résoudre de nouvelles difficultés tant du point de vue mathématique que numérique.

Dans le cadre de ce travail, nous nous proposons d'étudier l'écoulement des glaciers en régime stationnaire à partir du modèle glaciologique le plus complet à savoir le problème mécanique de Stokes couplé à l'équation de la chaleur. Notre travail analysera ce problème sur la base des aspects physique, mathématique et numérique indissociables à l'étude d'un problème physique concret. Les aspects numériques que nous étudierons s'attacheront essentiellement sur les méthodes de résolution par éléments finis qui constituent indéniablement un des outils les plus performants pour la résolution d'équations aux dérivées partielles dans des domaines aussi variés que la mécanique, la thermique et la mécanique des fluides.

Plus spécifiquement, trois grandes parties pourront être mises en évidence dans ce travail. La première partie mettra l'accent sur une présentation générale de la modélisation physique des écoulements glaciaires. Cela nous permettra d'établir un système d'équations pour le comportement thermomécanique des glaciers. La deuxième partie de ce travail consistera en une étude séparée des problèmes mécanique et thermique. Une attention particulière sera portée à la compréhension des fondements théoriques des méthodes numériques utilisées pour résoudre ces deux problèmes. Pour chacun des deux problèmes, nous mettrons en évidence quelques difficultés numériques et nous présenterons des solutions ou des perspectives de solutions adéquates. Enfin, la troisième partie de ce mémoire sera axée de manière plus concrète sur la présentation de méthodes numériques utilisées pour résoudre des problèmes multiphysiques. Cette dernière partie s'articulera autour de l'étude d'un glacier alpin « jouet » qui nous offrira l'opportunité d'appliquer différentes méthodes numériques à un cas concret.

De manière plus détaillée, ce travail se présentera de la manière suivante :

- Le chapitre 1 constitue une brève introduction à la glaciologie. Nous y présenterons succinctement quelques termes de vocabulaire employés dans ce domaine ainsi que l'importance des glaciers dans le monde et les raisons de s'y intéresser.

- Le chapitre 2 est consacré à la modélisation physique de l'écoulement des glaciers. À partir des équations de la mécanique des milieux continus et des spécificités propres à l'écoulement de la glace, nous obtiendrons les équations thermomécaniques à la base de tous les modèles glaciologiques. Nous présenterons également les lois de comportement pour la glace notamment la loi de Glen [27] communément utilisée en glaciologie.

- Au chapitre 3, nous étudierons le problème de Stokes utilisé pour décrire le comportement mécanique des fluides fortement visqueux dont la glace constitue un exemple typique. En particulier, nous analyserons les conditions sous lesquelles les problèmes de Stokes continu et discret admettent une unique solution pour les champs de vitesse et de pression dans la glace. Ces questions nous mèneront tout naturellement à introduire la condition de Babuska-Brezzi qui détermine le choix des fonctions de base dans la méthode de Galerkin. Nous présenterons ensuite quelques stratégies pour satisfaire cette condition ou au contraire la contourner.

- Le chapitre 4 est consacré à l'étude du problème thermique utilisé pour déterminer le champ de température dans la glace. Ce problème est décrit plus spécifiquement par une équation d'advection-diffusion. Cette équation est soumise à une contrainte thermique qui reflète l'impossibilité pour la glace d'avoir une température supérieure à son point de fusion. En étudiant l'équation d'advection-diffusion, nous verrons que dans le cas d'un transfert thermique dominé par la convection, la méthode classique de Galerkin peut présenter certaines faiblesses, ce qui nous amènera à l'introduction de méthodes de stabilisation. Nous analyserons également comment la contrainte thermique peut être traitée numériquement.

- Le chapitre 5 développe la question des méthodes de résolution pour les systèmes d'équations non linéaires en multiphysique. Bien que l'approche que nous y adopterons soit fort générale, les méthodes introduites sont d'application pour la résolution du problème thermomécanique décrivant le comportement des glaciers. Nous effectuerons une distinction importante entre les méthodes faiblement couplées et fortement couplées. Ces dernières méthodes qui sont mises en évidence par Keyes *et al.* [42] offrent de nouvelles opportunités dans le domaine de la simulation multiphysique.

- Le chapitre 6 exposera ensuite la résolution du problème thermomécanique couplé pour l'étude d'un glacier alpin « jouet ». Nous y aborderons les questions des formulations variationnelle et algébrique de ce problème. Nous verrons également comment appliquer les méthodes numériques introduites au chapitre 5.

- Le chapitre 7 présentera quelques simulations numériques réalisées sur notre glacier alpin « jouet ». Nous y appliquerons concrètement quelques méthodes numériques de résolution présentées tout au long de ce travail et nous en analyserons l'efficacité.

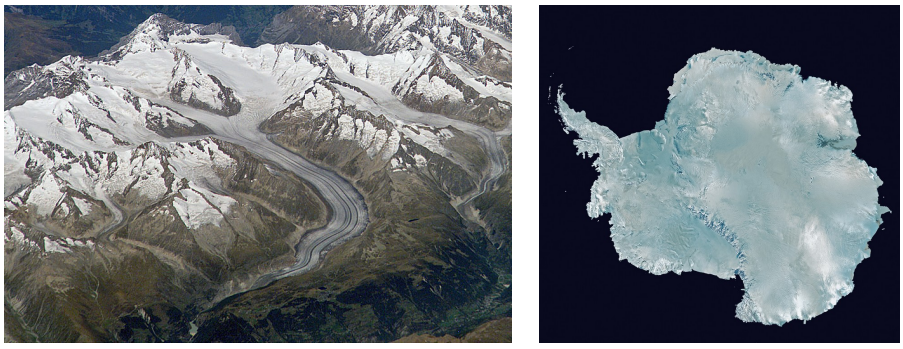
- En guise de fin, le chapitre 8 présentera une liste non exhaustive de possibilités pour approfondir ce mémoire aussi bien dans le domaine de la glaciologie que de la simulation multiphysique en général.

1. Glaciers et calottes polaires

Ce premier chapitre constitue une brève présentation sur les glaciers. Il a pour but de mettre en évidence l'importance des glaciers dans le monde et de souligner l'intérêt d'étudier leur comportement plus en détail. Le chapitre commence à la section 1.1 par une présentation de quelques termes de vocabulaire utilisés dans le domaine de la glaciologie. La section 1.2 poursuit ensuite par quelques arguments qui démontrent l'importance des glaciers dans le système terrestre. Enfin, la section 1.3 rappelle la problématique générale du réchauffement climatique sur la fonte des glaces. Cette problématique est généralement perçue comme une des incitations au développement de modèles pour la dynamique glaciaire.

1.1. Cryosphère terrestre

Une des caractéristiques propres les plus importantes de la Terre est sans nul doute l'abondance d'eau à sa surface. L'hydrosphère terrestre regroupe l'ensemble des zones de la planète où l'eau est présente sous forme liquide, solide ou gazeuse. La cryosphère terrestre qui recouvre actuellement environ 10% de la surface terrestre désigne l'ensemble des zones de la Terre où l'eau est sous forme solide à savoir dans les banquises, les glaciers, les régions enneigées et les lacs gelés.



(a) Glacier d'Aletsch dans les Alpes (image tirée de [66]) .

(b) Image satellite de l'Antarctique (image tirée de [14]).

FIGURE 1.1.: Illustrations de glaciers.

Une part importante de la cryosphère est occupée par les glaciers qui constituent le deuxième réservoir d'eau de la planète loin derrière les océans. Les glaciers contiennent environ 70% des réserves d'eau douce de la Terre. Le terme glacier se réfère de manière générale à une masse de glace qui repose sur le sol au contraire de la banquise qui désigne une étendue d'eau gelée. Sur Terre, les glaciers peuvent avoir des tailles et des formes fort différentes. Il est courant de distinguer les glaciers alpins des glaciers continentaux. Les glaciers alpins tels que le glacier d'Aletsch dans les Alpes (voir figure 1.1(a)) sont des glaciers de petites dimensions dont la morphologie est dictée par le relief [1]. Les glaciers alpins représentent moins de 1% de la quantité

1. Certains auteurs dont Greve et Blatter dans [30] utilisent parfois le terme glacier uniquement pour se référer aux glaciers alpins.

de glace de la cryosphère. Les glaciers continentaux sont quant à eux des glaciers de dimensions importantes dont la morphologie est peu affectée par le relief. Les glaciers continentaux dont la superficie n'excède pas 50 000 [km^2] sont appelés des calottes glaciaires tandis que les glaciers d'une superficie supérieure à 50 000 [km^2] sont nommés calottes polaires ou inlandsis. Les deux seules calottes polaires actuelles sont l'Antarctique et l'inlandsis du Groenland.

1.2. Importance des glaciers

Les glaciers sont des acteurs importants du système terrestre. Nous proposons ci-dessous quelques exemples qui démontrent l'importance des glaciers sur Terre.

- Les glaciers jouent un rôle important dans l'albédo de la Terre. En effet, les glaciers et les régions enneigées ont un pouvoir réfléchissant supérieur à la plupart des surfaces terrestres. Ils contribuent de ce fait à réfléchir une partie non négligeable du flux solaire incident et influencent ainsi la température moyenne de la Terre.

- Les glaciers constituent de grands réservoirs d'eau douce. En retenant l'eau pendant des milliers d'années, les glaciers jouent ainsi un rôle important dans le cycle de l'eau. Les glaciers de montagne sont notamment responsables de l'alimentation des torrents et des rivières de montagne principalement durant l'été.

- Les calottes polaires interagissent avec les océans et influencent la circulation thermohaline qui assure un transport de chaleur à grande échelle au niveau océanique et contribue ainsi à une redistribution de la chaleur sur l'ensemble du globe. Même si l'importance de la circulation thermohaline sur le climat n'est pas encore bien établie, il est admis que celle-ci contribue à un rôle de régulation à l'échelle globale.

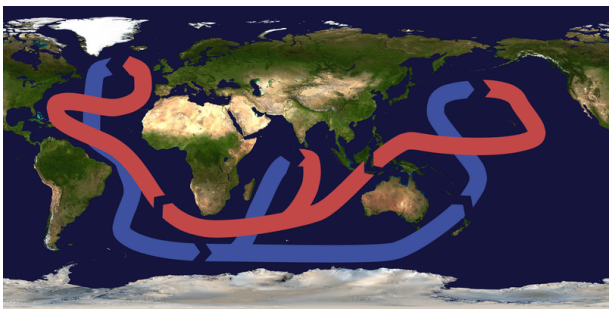


FIGURE 1.2.: La circulation thermohaline est responsable de la circulation des masses d'eau à grande échelle suite à des écarts de température et de salinité (image tirée de [13]).

1.3. Glaciers et réchauffement climatique

Ces dernières décennies, une attention toute particulière a été portée à l'étude des glaciers. En effet, la fonte des glaciers est l'une des premières conséquences du réchauffement climatique global actuel. L'évolution des glaciers a d'ailleurs été choisie par le Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) comme un des indicateurs principaux de

l'évolution du climat au XX^{e} siècle [35]. Afin de mieux percevoir l'importance de développer une meilleure compréhension de la dynamique des glaciers, nous illustrons la suite de cette sous-section par quelques conséquences du réchauffement climatique.

L'une des premières conséquences de la fonte des glaciers est l'augmentation du niveau des océans. Une fonte totale du volume de glace de l'Antarctique correspondrait à une augmentation du niveau marin d'environ 60 [m]. La fonte totale de la glace du Groenland équivaudrait quant à elle à une montée des eaux d'environ 7 [m] tandis que les glaciers alpins et les calottes glaciaires pourraient contribuer à une montée d'environ 0.5 [m] du niveau marin [30]. Bien que les calottes polaires représentent un volume de glace bien plus conséquent que celui des glaciers alpins et des calottes glaciaires, la contribution de celles-ci à la montée du niveau des mers sera moindre à court et à moyen termes. En effet, les glaciers de plus petite taille répondent plus rapidement à des modifications climatiques tandis que les grands glaciers possèdent une plus grande inertie, ce qui les rend moins vulnérables aux modifications environnementales. Dans le cas de l'Antarctique, les basses températures qui y règnent et des précipitations accrues pourraient résulter en un gain de masse positif dans cette région.

À long terme, la fonte des calottes polaires pourrait néanmoins avoir un impact sur l'ensemble des sous-systèmes terrestres. Outre la montée du niveau des mers, la fonte des inlandsis provoquera une diminution de l'albédo moyen de la Terre avec pour répercussion une accélération probable de l'augmentation de la température moyenne de la Terre et de la fonte des glaces. La fonte des calottes polaires entraînera aussi un mélange de l'eau de mer et de l'eau douce, ce qui diminuera la salinité des océans et influencera la circulation thermohaline.

La situation des glaciers alpins est plus préoccupante. Partout dans le monde, les glaciers alpins ont reculé de manière importante ces dernières décennies. Ce recul s'est même accéléré ces dernières années. La fonte des glaciers alpins pourrait provoquer des problèmes d'approvisionnement en eau douce dans certaines régions ainsi qu'affecter les activités agricoles dont l'irrigation dépend de l'eau de fonte des glaciers. Ce recul des glaciers de montagne aura aussi un impact sur les activités économiques de montagne telles que les sports d'hiver. De plus, le réchauffement climatique est susceptible d'accroître les risques de catastrophes glaciaires qui, bien que rares, peuvent avoir des conséquences dramatiques comme lors de la rupture en juillet 1892 d'une poche d'eau nichée dans le glacier de Tête Rousse en Haute-Savoie.

(a) Glacier Reid en 1899. (b) Glacier Reid en 2003.

FIGURE 1.3.: Récession du glacier Reid dans le parc national de Glacier Bay en Alaska. En 104 ans, le glacier a reculé d'environ 3 kilomètres (images tirées de [64]).

2. Modélisation physique de l'écoulement des glaciers

Une des premières étapes de l'étude d'un problème physique concret consiste à développer des modèles de la réalité physique à partir desquels des études mathématiques et numériques pourront être menées. Dans ce deuxième chapitre, nous introduisons les équations utilisées pour décrire le comportement thermomécanique des glaciers. Pour commencer, nous présenterons à la section 2.1 quelques arguments qui justifient de décrire le mouvement des glaciers comme l'écoulement d'un fluide fortement visqueux. Ensuite, le lecteur peu familier avec la mécanique des milieux continus pourra trouver à la section 2.2 quelques rappels généraux sur les équations thermomécaniques de bilan. Dans les sections 2.3 à 2.5, nous présenterons plus spécifiquement les équations thermomécaniques utilisées en glaciologie ainsi que les lois constitutives pour la glace dont la plupart sont de nature non linéaire. Ce chapitre se terminera finalement à la section 2.6 par une présentation du modèle physique que nous étudierons dans le reste de ce travail. Durant la lecture de ce chapitre, le lecteur devra être conscient que la modélisation des glaciers est un sujet relativement complexe pour lequel de nombreuses approches sont présentées dans la littérature scientifique. L'approche que nous exposons est basée plus spécifiquement sur la présentation de Greve et Blatter [30].

2.1. Modèle du glacier

La formation des glaciers résulte de l'accumulation de neige dans une région durant de nombreuses années. Ce processus se produit de préférence dans les régions de hautes latitudes ou de hautes altitudes pour lesquelles le bilan de masse glaciaire[1] est en général positif. L'accumulation de neige d'année en année provoque un tassement des couches des années précédentes et progressivement, la neige se compacte et devient plus dense en expulsant l'air qu'elle contient. Ce processus transforme finalement la neige en glace et donne ainsi naissance à un glacier.

Le mouvement des glaciers a longtemps constitué un mystère, car il paraissait inconcevable qu'un solide comme la glace puisse s'écouler. Ce n'est qu'au début du XX^{e} siècle que l'idée que la glace puisse se comporter comme un fluide visqueux commence à s'imposer. En fait sous l'action de son propre poids, la glace va se déformer par fluage. D'un point de vue rhéologique, la glace se comporte comme un fluide de viscosité élevée (de l'ordre de 10^{13} [Pa s] à 0 [°C]), ce qui explique l'écoulement lent des glaciers sur de longues périodes de temps. Ce comportement visqueux de la glace est illustré à la figure 2.1. Outre la déformation par fluage, les glaciers peuvent aussi glisser sur leur lit rocheux en présence d'eau de fonte.

1. Le bilan de masse glaciaire est la différence entre le gain de masse du glacier suite notamment aux précipitations neigeuses et la perte de masse par fonte de la glace ainsi que par sublimation de la neige et de la glace.

FIGURE 2.1.: Cette image illustre le caractère visqueux des glaciers. Les glaciers se déplacent d'abord dans des vallées étroites au-delà desquelles ils s'étalent à la manière d'un fluide fortement visqueux. Les glaciers de l'illustration se situent près du Surprise Fjord sur l'île Axel Heiberg dans l'Arctique canadien (image tirée de [65]).

Le comportement fluide de la glace ne se manifeste réellement que sur de longues périodes d'observation. Lorsque la glace est soumise à des efforts externes de courte durée, son comportement est bien entendu celui d'un solide. Ce double comportement de la glace est en fait résumé par le nombre de Déborah De défini par

$$De = \frac{t_c}{t_p}, \tag{2.1}$$

où t_c est le temps caractéristique du matériau (temps de relaxation) et t_p est le temps de l'expérience. Ce nombre montre qu'un matériau se comportera comme un solide si le temps de sollicitation est bien plus petit que le temps caractéristique du matériau ($De >> 1$) et comme un fluide si ce matériau est sollicité durant un temps suffisamment long ($De << 1$). C'est ainsi qu'en observant les glaciers sur plusieurs décennies, il est possible de constater leur écoulement.

Les arguments précédents nous permettent de comprendre que la description physique du mouvement des glaciers trouvera sa place dans le cadre de la mécanique des fluides. Avant d'établir les équations mathématiques de l'écoulement des glaciers, nous avons représenté sur la figure 2.2 un schéma simplifié d'un glacier qui nous permet d'introduire les notations que nous utiliserons par la suite. Nous supposerons l'étude d'un glacier représenté par le volume de glace $\Omega(t)$ qui repose sur un sol rocheux. La frontière de ce domaine se décompose en deux parties à savoir l'interface air-glace $\Gamma_s(t)$ et l'interface glace-roche $\Gamma_b(t)$. La surface supérieure du glacier est donnée par l'équation $z = h(x, y, t)$ tandis que l'équation de la base du glacier est donnée $z = b(x, y, t)$. Chacune des quantités introduites dans ce paragraphe est susceptible de varier au cours du temps. Les variations de ces grandeurs pourront en réalité être déterminées à partir des équations que nous allons établir dans le reste de ce chapitre.

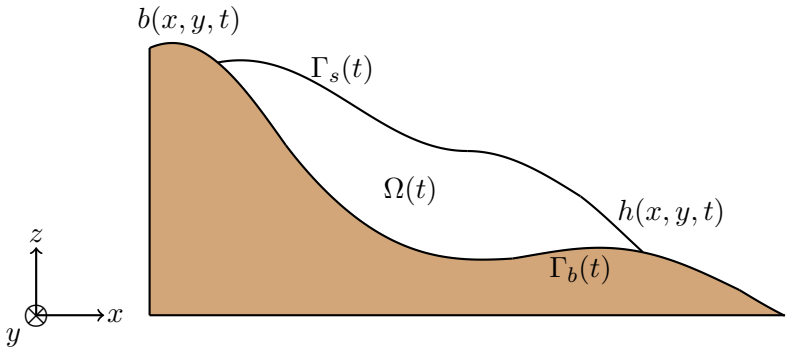


FIGURE 2.2.: Représentation schématique d'un glacier. Le domaine $\Omega(t)$ représente le volume de glace tandis que $\Gamma_s(t)$ est l'interface air-glace et $\Gamma_b(t)$ est l'interface glace-roche.

2.2. Équations générales de bilan

Les équations de bilan en mécanique des milieux continus sont un ensemble d'équations qui expriment localement la conservation de la masse, de la quantité de mouvement et de l'énergie. Pour établir ces équations, nous considérons un volume $\mathcal{V}(t)$ de glace qui contient à tout instant une quantité fixe de matière. Ce volume est caractérisé à chaque instant par sa frontière $\partial\mathcal{V}(t)$ de normale unitaire extérieure $\mathbf{n}$. Il est également plus simple d'envisager ce volume comme un ensemble constitué de particules matérielles [2] qui se déplacent à la vitesse $\mathbf{v}$.

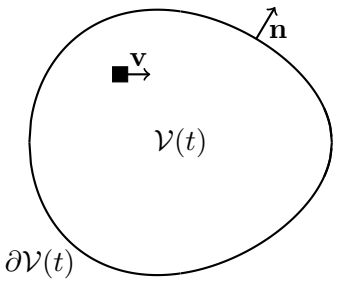


FIGURE 2.3.: Les équations de bilan sont exprimées pour un volume matériel $\mathcal{V}(t)$ constitué de particules matérielles qui se déplacent avec une vitesse $\mathbf{v}$.

Pour la suite de ce chapitre, il est utile d'introduire la dérivée matérielle d'une quantité physique comme l'opérateur

$$\frac{D(\,\cdot\,)}{Dt} = \frac{\partial(\,\cdot\,)}{\partial t} + \mathbf{v}\cdot\nabla(\,\cdot\,). \tag{2.2}$$

Cette dérivée permet de calculer la variation dans le temps d'une grandeur physique associée à une particule matérielle.

2. En mécanique des milieux continus, la notion de particule matérielle se réfère à une portion de matière suffisamment petite pour être considérée comme ponctuelle, mais qui contient suffisamment de particules pour pouvoir lui associer des propriétés macroscopiques.

2.2.1. Équation de conservation de la masse

L'équation de conservation de la masse exprime mathématiquement que la masse d'un système fermé est une quantité conservée. Ainsi, la masse est une grandeur physique qui ne peut être ni créée ni détruite [3].

La conservation de la masse du volume de matière $\mathcal{V}$ s'écrit

$$\frac{d}{dt}\int_{\mathcal{V}} \rho \, dV = 0, \tag{2.3}$$

où ρ représente la densité volumique de masse.

En utilisant le théorème de transport de Reynolds, cette équation peut se réécrire

$$\int_{\mathcal{V}} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) \, dV = 0. \tag{2.4}$$

Comme cette dernière équation est valable pour un volume $\mathcal{V}$ arbitraire, le théorème de localisation nous permet alors d'obtenir la loi de conservation de la masse sous forme locale à savoir

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0. \tag{2.5}$$

2.2.2. Équation de bilan de quantité de mouvement

L'équation de bilan de quantité de mouvement est l'application de la deuxième loi de la mécanique de Newton au volume matériel $\mathcal{V}(t)$. Cette loi exprime que la variation temporelle de la quantité de mouvement du volume matériel $\mathcal{V}(t)$ est égale à la somme des forces extérieures de volume et de surface qui s'exercent sur ce corps. Les forces de volume qui agissent sur le corps sont caractérisées par leur densité volumique $\mathbf{f}$. Les efforts surfaciques de contact entre des volumes matériels voisins sont généralement décrits en fonction de leur densité surfacique $\mathbf{t}$. La caractérisation de cette quantité se base généralement sur le postulat de Cauchy et le théorème des contraintes de Cauchy [33]. Ce dernier théorème introduit un tenseur σ appelé tenseur des contraintes de Cauchy qui relie la densité surfacique de forces qui s'exercent sur une surface à la normale unitaire extérieure $\mathbf{n}$ à cette surface suivant la relation $\mathbf{t} = \sigma \cdot \mathbf{n}$. Il est également possible de montrer qu'en général le tenseur des contraintes est symétrique [4].

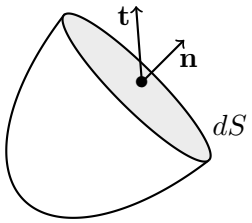


FIGURE 2.4.: Sur un élément de surface dS de normale unitaire extérieure $\mathbf{n}$, les efforts externes sont caractérisés par leur densité surfacique $\mathbf{t}$.

3. Dans le cas de la glace, il serait toutefois légitime de supposer un gain ou une perte de glace par transition de phase. Nous supposerons cependant que ce n'est pas le cas à l'intérieur du domaine Ω et que celui-ci est uniquement composé de glace. Des gains ou des pertes de glace seront juste envisagés sur la frontière du domaine.

4. Cette propriété est une conséquence de l'équation de bilan de moment cinétique. Elle n'est cependant pas valable pour des milieux dans lesquels il existe des couples internes en volume.

En tenant compte des remarques précédentes, l'équation de bilan de quantité de mouvement s'écrit alors

$$\frac{d}{dt}\int_{\mathcal{V}} \rho\,\mathbf{v}\,dV = \int_{\partial\mathcal{V}} \sigma\cdot\mathbf{n}\,dS + \int_{\mathcal{V}} \mathbf{f}\,dV. \tag{2.6}$$

En utilisant le théorème de la divergence de Gauss, cette équation peut encore se réécrire

$$\int_{\mathcal{V}} \rho\,\frac{D\mathbf{v}}{Dt}\,dV = \int_{\mathcal{V}} \operatorname{div}\sigma\,dV + \int_{\mathcal{V}} \mathbf{f}\,dV. \tag{2.7}$$

Sous forme locale, cette expression devient finalement

$$\rho\,\frac{D\mathbf{v}}{Dt} = \rho\,\frac{\partial\mathbf{v}}{\partial t} + \rho\,\mathbf{v}\cdot\nabla\mathbf{v} = \operatorname{div}\sigma + \mathbf{f}. \tag{2.8}$$

2.2.3. Équation de bilan d'énergie

La troisième équation de bilan que nous utiliserons est une équation de nature thermodynamique. Elle consiste à appliquer le principe de conservation de l'énergie au volume $\mathcal{V}(t)$. L'énergie de ce volume est composée d'une énergie cinétique associée au mouvement macroscopique de ses particules matérielles ainsi que d'une énergie interne de densité massique e liée aux énergies potentielle et cinétique des particules microscopiques. La première loi de la thermodynamique appliquée au système fermé $\mathcal{V}(t)$ implique que la variation temporelle de l'énergie contenue dans ce système provient de la puissance développée par les forces externes ainsi que des flux de chaleur au travers de la frontière $\partial\mathcal{V}(t)$ et des sources internes de chaleur telles que des sources radiatives.

Sous forme mathématique, l'équation de bilan d'énergie peut alors s'écrire

$$\frac{d}{dt}\int_{\mathcal{V}} \rho\,(e+\frac{1}{2}\|\mathbf{v}\|^2)\,dV = \int_{\mathcal{V}} \mathbf{f}\cdot\mathbf{v}\,dV + \int_{\partial\mathcal{V}} (\sigma\cdot\mathbf{n})\cdot\mathbf{v}\,dS - \int_{\partial\mathcal{V}} \mathbf{q}\cdot\mathbf{n}\,dS + \int_{\mathcal{V}} r\,dV, \tag{2.9}$$

où $\mathbf{q}$ représente le vecteur densité de flux de chaleur et r les sources internes de chaleur par unité de volume.

En utilisant encore une fois le théorème de la divergence de Gauss, nous obtenons

$$\int_{\mathcal{V}} \rho\,\frac{D}{Dt}(e+\frac{1}{2}\|\mathbf{v}\|^2)\,dV = \int_{\mathcal{V}} \mathbf{f}\cdot\mathbf{v}\,dV + \int_{\mathcal{V}} \operatorname{div}(\sigma\cdot\mathbf{v})\,dV - \int_{\mathcal{V}} \operatorname{div}\mathbf{q}\,dV + \int_{\mathcal{V}} r\,dV. \tag{2.10}$$

Sous forme locale, l'équation de bilan d'énergie devient alors

$$\rho\,\frac{D}{Dt}(e+\frac{1}{2}\|\mathbf{v}\|^2) = \mathbf{f}\cdot\mathbf{v} + \operatorname{div}(\sigma\cdot\mathbf{v}) - \operatorname{div}\mathbf{q} + r. \tag{2.11}$$

Il est courant de soustraire à cette dernière relation l'équation de bilan de quantité de mouvement (2.8) multipliée scalairement par le champ de vitesse $\mathbf{v}$ afin d'obtenir une équation uniquement pour la densité massique d'énergie interne. Ainsi, nous obtenons l'équation de la chaleur

$$\rho\,\frac{De}{Dt} = -\operatorname{div}\mathbf{q} + \sigma : \dot{\varepsilon} + r, \tag{2.12}$$

où nous avons introduit le tenseur taux de déformation $\dot{\varepsilon}(\mathbf{v}) = \frac{1}{2}(\nabla\mathbf{v} + \nabla\mathbf{v}^T)$.

Le comportement thermomécanique de tout milieu continu est ainsi décrit par les trois équations de bilan (2.5), (2.8) et (2.12). L'application de ces équations au cas concret de

l'étude des écoulements glaciaires nécessite toutefois d'introduire des lois comportementales pour la glace. De plus, comme nous le verrons à la section suivante, certains termes de ces équations peuvent être négligés dans le cas spécifique des glaciers.

2.3. Équations thermomécaniques de bilan pour les glaciers

Les équations établies à la section 2.2 sont en fait très générales et relativement complexes. En considérant quelques hypothèses applicables aux écoulements glaciaires, nous pouvons simplifier ces équations. Cette section présente ainsi les hypothèses couramment employées en glaciologie.

La première hypothèse consiste à traiter la glace comme un fluide incompressible. Bien que cette hypothèse paraisse assez naturelle, elle peut sembler assez contradictoire avec la compression de la neige qui est à l'origine des glaciers. Toutefois, ce phénomène de formation de glace se produit principalement à la surface du glacier de sorte que le glacier peut être considéré comme globalement incompressible. Dans ce cas, l'équation (2.5) se réduit à

$$\operatorname{div} \mathbf{v} = 0. \tag{2.13}$$

Bien que l'hypothèse d'incompressibilité du fluide permette de simplifier les équations, nous verrons au chapitre 3 que l'équation (2.13) pose certains problèmes pour la résolution tant mathématique que numérique des équations mécaniques de bilan.

Dans l'équation de bilan de quantité de mouvement, il est commode de décomposer le tenseur des contraintes en sa partie déviatorique σ^D et un tenseur isotrope. Nous écrirons ainsi

$$\sigma = -p\,\underline{\underline{I}} + \sigma^D, \tag{2.14}$$

où nous avons noté par $\underline{\underline{I}}$ le tenseur identité. Nous avons introduit également la pression mécanique [5] $p = -\frac{1}{3}\operatorname{Tr}\sigma$. Le tenseur des contraintes déviatorique peut être relié au tenseur taux de déformation via la relation

$$\sigma^D = 2\mu\dot{\varepsilon}, \tag{2.15}$$

avec μ la viscosité dynamique du fluide qui dépend a priori de l'état de contrainte dans le fluide. La viscosité du fluide est supposée être une grandeur scalaire ce qui est en accord avec l'hypothèse d'isotropie de la glace que nous poserons à la sous-section 2.4.1.

En tenant compte de cette décomposition du tenseur des contraintes, la relation (2.8) peut se réécrire

$$\rho\,\frac{\partial \mathbf{v}}{\partial t} + \rho\,\mathbf{v}\cdot\nabla\mathbf{v} = -\nabla p + \operatorname{div}(2\mu\dot{\varepsilon}) + \mathbf{f}. \tag{2.16}$$

La relation (2.16) peut être simplifiée en considérant les remarques suivantes :

- Les forces de volume qui agissent sur le glacier sont principalement de nature gravifique. Nous écrirons donc $\mathbf{f} = \rho\mathbf{g}$ avec $\mathbf{g}$ l'accélération de pesanteur supposée constante. Pour l'étude des calottes polaires dans un référentiel lié à la Terre, nous pourrions envisager

5. Cette pression n'est pas égale à la pression thermodynamique, car elle ne peut pas être exprimée au moyen d'une équation d'état. Le rôle de la pression pour un fluide incompressible sera discuté au chapitre 3.

d'introduire des termes d'accélération supplémentaires comme par exemple l'accélération de Coriolis à laquelle est associée la force fictive de Coriolis. Toutefois dans le cas des glaciers, les forces fictives d'inertie peuvent être négligées en comparaison de la force gravifique.

- Les glaciers s'écoulant lentement, une deuxième hypothèse naturelle consiste à négliger le terme convectif $\rho\,\mathbf{v}\cdot\nabla\mathbf{v}$ en comparaison du terme visqueux (terme de diffusion de la quantité de mouvement). Cette hypothèse peut être vérifiée à partir d'un argument dimensionnel en introduisant le nombre de Reynold Re qui compare les effets inertiel et visqueux. Pour rappel, nous avons

$$Re = \frac{\rho V L}{\mu}, \tag{2.17}$$

avec V une vitesse caractéristique et L une longueur caractéristique. À titre d'exemple, nous pouvons considérer un glacier de montagne avec une vitesse horizontale de 10 [m/an], une longueur caractéristique de 1 [km], une viscosité de 10^{13} [Pa s] et une masse volumique de 910 [kg/m^3] [6]. Nous obtenons alors $Re \approx 10^{-14}$, ce qui confirme que le terme convectif dans l'équation de bilan de quantité de mouvement est négligeable en comparaison du terme visqueux.

- Le terme de dérivée partielle de la vitesse par rapport au temps est également faible comparé au gradient de pression et au terme visqueux de sorte que ce terme pourra être également négligé. De cette façon, nous adoptons une approximation quasi statique pour l'écoulement.

En tenant compte de ces trois observations, l'équation (2.16) s'écrit finalement

$$-\nabla p + \operatorname{div}(2\mu\dot{\varepsilon}) + \rho\mathbf{g} = \mathbf{0}. \tag{2.18}$$

Les équations (2.13) et (2.18) décrivent le comportement mécanique des glaciers. Ces équations accompagnées de conditions aux limites adéquates constituent le problème de Stokes utilisé pour étudier les écoulements fortement visqueux. L'étude analytique de ce problème sera menée au chapitre 3. En glaciologie, il est de coutume de simplifier le problème de Stokes en négligeant divers termes afin de faciliter les analyses mathématique et numérique. Parmi les approximations, nous citons l'approximation de glace peu profonde [30] et le modèle de Blatter/Pattyn [54]. Ces approximations reposent en général sur la constatation que certaines composantes du tenseur des contraintes sont plus importantes que d'autres ou encore que les dérivées des composantes de la vitesse sont plus importantes dans certaines directions. Malgré l'intérêt indéniable de ces modèles simplifiés pour la résolution numérique des équations du mouvement, nous ne les considérerons pas dans le reste de ce travail où l'accent sera mis sur le problème de Stokes.

La viscosité du fluide dépend notamment de la température. Il en résulte un problème thermomécanique couplé. L'équation qui décrit le champ de température dans le glacier peut être dérivée à partir de l'équation (2.12) pour l'énergie interne. Dans cette équation, nous

6. Les dimensions et vitesses caractéristiques dépendent évidemment du type de glacier. En particulier, la longueur d'un glacier peut aller de quelques centaines de mètres pour un glacier de montagne à plusieurs milliers de kilomètres pour les calottes polaires. Dans tous les cas, le nombre de Reynolds est faible.

négligeons le terme de chauffage radiatif qui n'est réellement important que dans les premiers centimètres de glace exposés directement au Soleil. Pour un fluide incompressible, l'énergie interne est reliée à la température par la capacité thermique massique $c(T) > 0$. L'énergie interne à une température T est ainsi donnée par

$$e = \int_{T_0}^{T} c(\bar{T})\, d\bar{T}. \tag{2.19}$$

De plus, nous supposons que la glace est isotrope d'un point de vue thermique et que le flux diffusif de chaleur $\mathbf{q}$ est décrit par la loi de Fourier

$$\mathbf{q} = -k(T)\, \nabla T, \tag{2.20}$$

où nous avons introduit la conductivité thermique $k(T) > 0$ du matériau.

En tenant compte de l'incompressibilité du fluide et de la décomposition (2.14) du tenseur des contraintes, l'équation pour la température s'écrit finalement

$$\rho c \frac{DT}{Dt} = \operatorname{div}(k \nabla T) + 2\mu \operatorname{Tr} \dot{\varepsilon}^2. \tag{2.21}$$

Pour que cette équation décrive correctement la température dans un volume de glace, il convient d'imposer une restriction sur les valeurs que peut prendre la température T. En effet, la température de la glace ne peut pas excéder son point de fusion. Si nous notons par T_m la température de fusion de la glace, l'équation (2.21) est alors soumise à la contrainte thermique

$$T \leqslant T_m. \tag{2.22}$$

Les équations (2.13), (2.18) et (2.21) constituent le système d'équations qui permet de déterminer les champs de vitesse, de pression et de température pour les écoulements des glaciers. L'utilisation de ces équations nécessite toutefois de disposer d'expressions analytiques pour les propriétés physiques de la glace et d'introduire des conditions aux limites. Ces deux points sont discutés dans les deux sections suivantes.

2.4. Propriétés physiques de la glace

2.4.1. Comportement rhéologique de la glace

La glace de la cryosphère se présente sous la forme d'un solide cristallin de structure hexagonale (glace $\mathrm{I_h}$). Un cristal de glace résulte de la superposition d'un ensemble de plans à structure hexagonale. Il en découle qu'un monocristal de glace est fortement anisotrope. Dans la nature, la glace se manifeste sous la forme d'un polycristal constitué d'un très grand nombre de cristallites (voir figure 2.5). Si la répartition de ces cristallites est aléatoire, il en résulte un polycristal isotrope. Pour la suite de ce mémoire, nous adopterons l'hypothèse d'isotropie de la glace. Toutefois, une description de la glace en tant que solide anisotrope est plus pertinente pour certaines situations telles que la description de la glace en profondeur dans les glaciers épais.

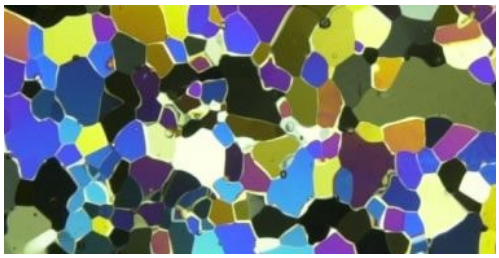


FIGURE 2.5.: Microstructure d'un échantillon de glace polycristalline vu au travers de polariseurs croisés. Les cristallites ont une taille de l'ordre de 1 [cm] et leur couleur apparente dépend de leur orientation (image tirée de [39]).

Comme souligné à la section 2.1, la glace se déforme par fluage. Lors de la déformation par fluage, il est courant de distinguer trois étapes. Celles-ci sont représentées schématiquement sur la courbe de fluage 2.6. La première étape qualifiée de fluage primaire, qui succède directement à une déformation élastique instantanée, est caractérisée par une vitesse de déformation qui diminue au cours du temps avant d'atteindre un minimum. Durant la deuxième étape dite de fluage secondaire, le matériau se déforme à vitesse constante. Enfin dans la dernière étape dite de fluage tertiaire, la vitesse de déformation augmente pour tendre vers une valeur constante. Pour les glaciers, il est courant de supposer que la glace a été sollicitée depuis un temps suffisamment long pour se situer dans les étapes de fluage secondaire et tertiaire.

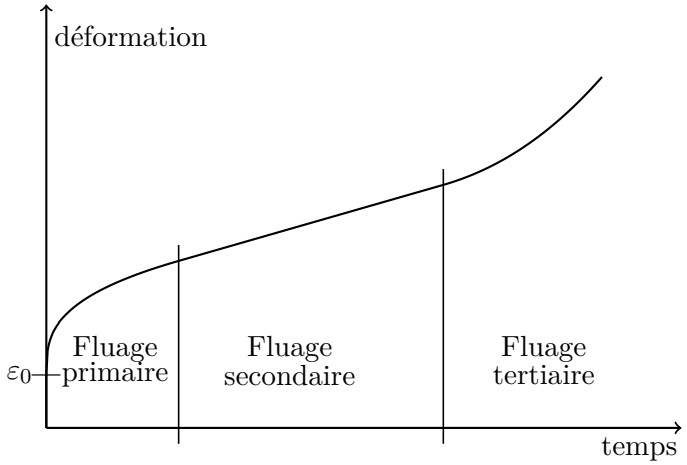


FIGURE 2.6.: Représentation schématique d'une courbe de fluage.

L'établissement d'une loi de fluage pour la glace est en pratique un exercice difficile. En théorie, le comportement de la glace est susceptible de varier d'un glacier à l'autre. Dans ce travail, nous nous limiterons à présenter l'une des lois d'écoulement les plus utilisées en glaciologie à savoir la loi de Glen. Cette loi n'est applicable à strictement parler que dans la phase de fluage secondaire. Le comportement rhéologique de la glace est celui d'un fluide non newtonien de sorte que sa viscosité dépend de son état de contrainte. L'isotropie supposée de la glace signifie que la viscosité est une grandeur scalaire si bien que son expression mathématique doit être indépendante de tout système de coordonnées. Il en résulte que la viscosité dépend du tenseur des contraintes déviatorique σ^D uniquement par l'intermédiaire des invariants de ce tenseur. En particulier, nous introduisons la contrainte effective σ_e qui est égale à la racine carrée du deuxième invariant du tenseur σ^D.

Nous avons dès lors

$$\sigma_e = \sqrt{\frac{1}{2}\,\mathrm{Tr}(\sigma^D)^2}\,. \tag{2.23}$$

La contrainte effective n'est pas une contrainte qui existe réellement dans le matériau, mais elle permet de résumer en une seule quantité scalaire l'état de contrainte décrit par le tenseur des contraintes. Pour un cisaillement simple, la contrainte effective est équivalente à la contrainte de cisaillement. La notion de contrainte effective est souvent utilisée en mécanique du solide pour établir des critères de plasticité [59]. Son expression ne fait pas intervenir la pression du fluide, car celle-ci est supposée ne pas affecter la déformation plastique. L'utilisation d'une contrainte effective apparaît donc comme un choix pertinent pour l'écriture d'une loi comportementale pour la glace.

La loi de Glen suppose que la viscosité de la glace ne dépend que du deuxième invariant du tenseur σ^D [7]. L'expression de la viscosité dans la loi de Glen s'écrit

$$\mu(T', \sigma_e) = \frac{1}{2A(T')\sigma_e^{n-1}}\,, \tag{2.24}$$

avec le facteur d'Arrhénius

$$A(T') = A_0 e^{-\frac{Q}{RT'}}\,, \tag{2.25}$$

où A_0 est une constante pré-exponentielle, Q est l'énergie d'activation pour le fluage, $R = 8.3144\,[\mathrm{J\,mol^{-1}\,K^{-1}}]$ est la constante des gaz parfaits et T' est la température de la glace relativement au point de fusion à savoir $T' = T + \beta\, p$ avec $\beta = 7.42 \times 10^{-8}\,[\mathrm{K\,Pa^{-1}}]$ la constante de Clausius-Clapeyron [8]. Dans l'expression (2.24), la quantité n est appelée l'exposant de Glen. La valeur de cette exposant est encore un sujet de discussion dans le domaine de la glaciologie même si la valeur $n = 3$ est habituellement utilisée [68].

La relation (2.25) traduit la diminution de dureté de la glace quand la température augmente. Au-dessus de $-10\,[^\circ\mathrm{C}]$, la dureté de la glace diminue cependant plus rapidement que ce que ne prédit la loi d'Arrhénius. Ce phénomène est notamment attribué au glissement à la frontière des cristallites ainsi qu'à la présence d'eau liquide au niveau de ces frontières [15]. Afin de rendre compte de cette modification dans le comportement de la glace, il est d'usage de considérer une énergie d'activation différente dans le facteur d'Arrhénius pour $T' \leqslant 263.15\,[\mathrm{K}]$ et $T' > 263.15\,[\mathrm{K}]$. Les paramètres généralement utilisés dans la loi de Glen sont repris dans le tableau 2.1.

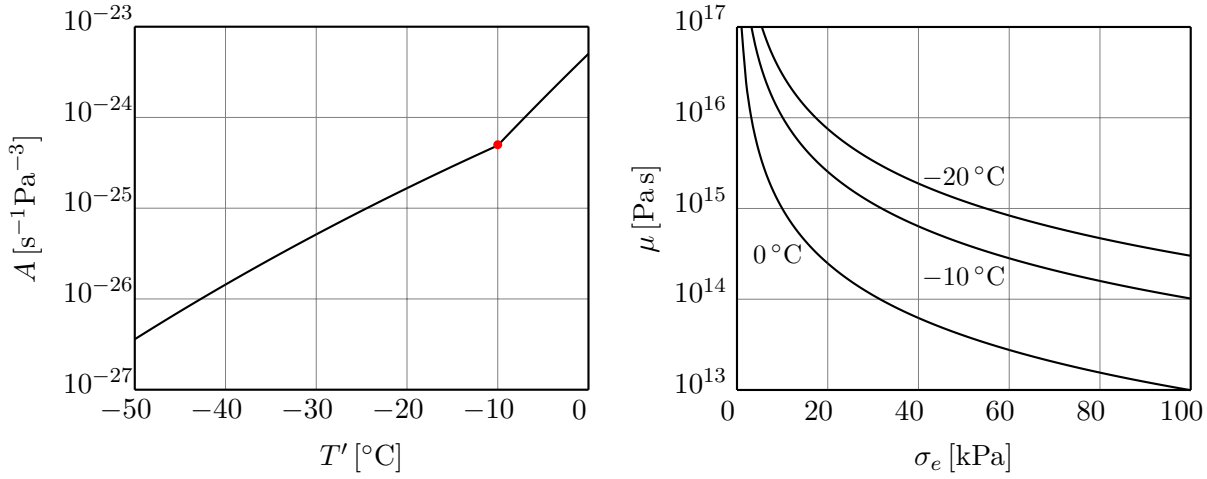
7. Cette hypothèse est encore sujette à discussion. Certaines expériences de fluage réalisées sur de la glace semblent montrer que la réponse du matériau n'est pas totalement indépendante des autres invariants du tenseur (pour plus de détails sur cette hypothèse voir [5, 15])

8. La formule de Clausius-Clayperon en thermodynamique permet de déterminer la pression de fusion d'un corps en fonction de la température. Pour les glaciers, il est possible d'utiliser la relation linéaire approchée $T_m = T_0 - \beta p$ avec $T_0 = 273.15\,[\mathrm{K}]$. Le signe moins rend compte de la propriété particulière de la glace d'avoir une température de fusion qui diminue avec une augmentation de la pression. Notons également que la valeur de $7.42 \times 10^{-8}\,[\mathrm{K\,Pa^{-1}}]$ pour la constante de Clausius-Clapeyron correspond à celle de l'eau pure. Pour une eau saturée en air, cette constante vaut $\beta = 9.8 \times 10^{-8}\,[\mathrm{K\,Pa^{-1}}]$. En général, la glace des glaciers contient des bulles d'air et choisir β supérieur à $7.42 \times 10^{-8}\,[\mathrm{K\,Pa^{-1}}]$ peut être un choix approprié [34].

Paramètres	Valeur	
Exposant de Glen, n	3	
Constante pré-exponentielle, A_0	$3.985 \times 10^{-13}\,[\mathrm{s^{-1}\,Pa^{-3}}]$	$(T' \leqslant 263.15\,[\mathrm{K}])$
	$1.916 \times 10^{3}\,[\mathrm{s^{-1}\,Pa^{-3}}]$	$(T' > 263.15\,[\mathrm{K}])$
Énergie d'activation, Q	$60\,[\mathrm{kJ\,mol^{-1}}]$	$(T' \leqslant 263.15\,[\mathrm{K}])$
	$139\,[\mathrm{kJ\,mol^{-1}}]$	$(T' > 263.15\,[\mathrm{K}])$

TABLE 2.1.: Paramètres physiques pour la loi de Glen [30].

Sur la figure 2.7, nous avons représenté le facteur d'Arrhénius en fonction de la température ainsi que la viscosité de la glace en fonction de la contrainte effective pour différentes températures. Nous constatons à la figure 2.7(a) que le facteur d'Arrhénius présente une sensibilité importante vis-à-vis de la température. Sur la plage de température de $-50\,[°\mathrm{C}]$ à $0\,[°\mathrm{C}]$ typique des glaciers, le facteur d'Arrhénius peut varier d'un facteur 10^3, ce qui signifie que la connaissance de la température dans le glacier est un point essentiel. Bien que le facteur d'Arrhénius soit une fonction continue de la température, la courbe de la figure 2.7(a) montre la présence d'un point anguleux pour une température de $-10\,[°\mathrm{C}]$. Il en résulte donc une discontinuité de la dérivée du facteur d'Arrhénius par rapport à la température. La figure 2.7(b) montre que la viscosité de la glace peut varier de plusieurs ordres de grandeur suivant la valeur de la contrainte effective.



(a) Facteur d'Arrhénius.

(b) Viscosité de la glace.

FIGURE 2.7.: Caractéristiques mécaniques de la glace ($n = 3$).

En pratique, il est plus utile d'exprimer la relation (2.24) à l'aide du tenseur taux de déformation. Dans ce cas, cette expression s'écrit

$$\mu(T', d_e) = \frac{1}{2} A(T')^{-1/n} d_e^{-(1-1/n)}\,, \tag{2.26}$$

où nous avons introduit le taux de déformation effectif

$$d_e = \sqrt{\frac{1}{2}\,\mathrm{Tr}\,\dot{\varepsilon}^2}. \tag{2.27}$$

Dans la relation (2.24), nous constatons que la viscosité devient infinie si σ_e tend vers 0. Cela peut poser un problème notamment au niveau de la résolution numérique. Afin de s'affranchir

de cette complication, il est commode d'introduire une faible contrainte résiduelle $\sigma_0 > 0$ dans la loi de Glen. Dans ce cas, nous obtenons la loi de Glen régularisée

$$\mu(T', \sigma_e) = \frac{1}{2A(T')\left[\sigma_e^{n-1} + \sigma_0^{n-1}\right]}. \tag{2.28}$$

Cette dernière relation peut être exprimée en fonction du taux de déformation effectif en remplaçant σ_e par $2\mu d_e$. Toutefois, il n'est pas possible d'établir de manière générale une formule explicite pour l'expression de la viscosité en fonction du taux de déformation effectif. Il est en fait nécessaire de résoudre analytiquement ou numériquement l'équation suivante [9] :

$$2^n A(T') d_e^{n-1} \mu^n(T', d_e) + 2A(T')\sigma_0^{n-1}\mu(T', d_e) - 1 = 0. \tag{2.29}$$

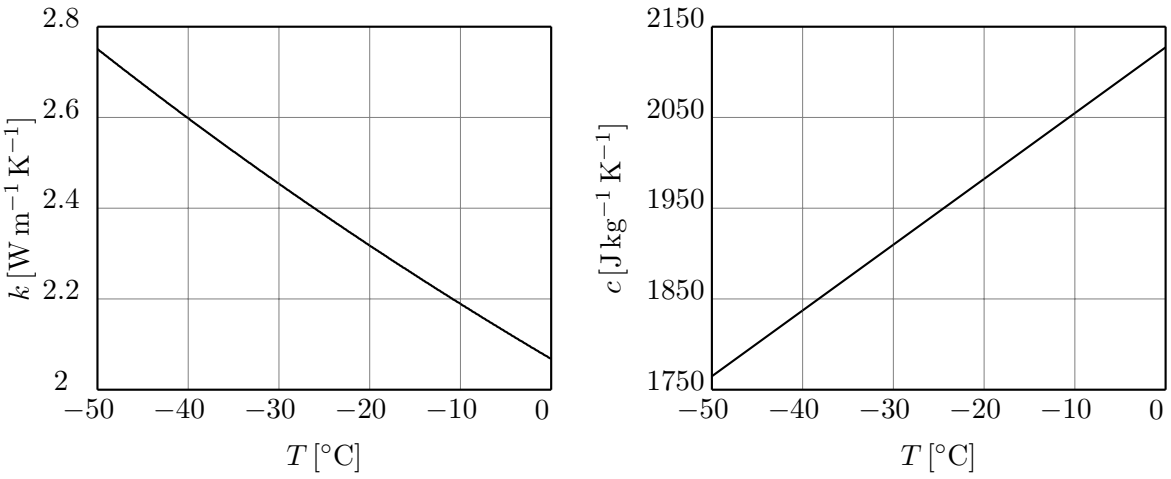
2.4.2. Propriétés thermiques de la glace

La conductivité thermique et la capacité thermique massique de la glace dépendent toutes les deux de la température. La conductivité thermique de la glace est une fonction décroissante de la température au contraire de sa capacité thermique massique. Les dépendances de ces grandeurs par rapport à la température sont données par les relations

$$k(T) = 9.828\, e^{-0.0057T}\ [\mathrm{W\,m^{-1}\,K^{-1}}], \tag{2.30}$$

$$c(T) = (146.3 + 7.253T)\ [\mathrm{J\,kg^{-1}\,K^{-1}}]. \tag{2.31}$$

La variation de ces deux grandeurs avec la température est reprise sur la figure 2.8. Nous observons sur cette figure que la variation de k et de c avec la température reste relativement modérée de sorte qu'il pourrait être raisonnable de travailler avec une valeur moyenne.



(a) Conductivité thermique.

(b) Capacité thermique massique.

FIGURE 2.8.: Propriétés thermiques de la glace de $-50\,[^\circ\mathrm{C}]$ à $0\,[^\circ\mathrm{C}]$.

9. Pour le cas $n = 3$ qui est d'un intérêt particulier, la formule (2.29) est un équation polynomiale du troisième degré du type $x^3 + px + q = 0$ ($d_e \neq 0$). Ces trois solutions peuvent être calculées à partir des formules de Cardan. Il peut être vérifié que l'équation (2.29) admet une unique solution réelle.

2.5. Conditions aux limites

Les équations présentées jusqu'à présent ne sont pas suffisantes pour déterminer le comportement dynamique d'un glacier. À ces équations, il convient d'adjoindre des conditions initiales et des conditions aux limites adéquates. Dans cette section, nous présentons les conditions aux limites les plus fréquemment utilisées.

2.5.1. Équations des interfaces

Le problème de Stokes introduit à la section 2.3 ne comprend aucun terme de dérivée temporelle de telle façon que ces équations semblent décrire un problème mécanique stationnaire. En réalité, le mouvement du glacier induit une évolution des interfaces air-glace et glace-sol au cours du temps. Il convient dès lors de déterminer une équation d'évolution pour $b(x, y, t)$ et $h(x, y, t)$. En tenant compte des interfaces qui évoluent dans le temps, nous obtenons ainsi un problème mécanique quasi statique.

Nous établissons l'équation d'évolution pour le cas de l'interface air-glace. L'équation pour l'interface glace-sol se déduit de manière similaire. L'équation de l'interface air-glace est donnée par $F_s(\mathbf{x}, t) = z - h(x, y, t) = 0$. La normale unitaire à cette surface est notée $\mathbf{n} = \nabla F_s / \|\nabla F_s\|$ et est dirigée vers l'extérieur de la glace. Pour la suite, nous désignons par N_s la norme du gradient de F_s. Nous notons également par $\mathbf{w}$ la vitesse de déplacement de cette interface et par (u_s, v_s, w_s) les composantes de la vitesse de la glace au niveau de cette surface.

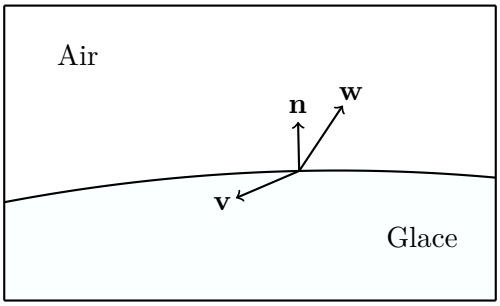


FIGURE 2.9.: Représentation de l'interface entre la glace et l'air. La vitesse $\mathbf{w}$ est la vitesse de l'interface et $\mathbf{v}$ est la vitesse du fluide à l'interface (image tirée de [30]).

Pour une particule fictive qui se déplacerait avec l'interface, celle-ci ne semble pas évoluer au cours du temps, ce qui s'écrit mathématiquement

$$\frac{D_{\mathbf{w}} F_s}{Dt} = \frac{\partial F_s}{\partial t} + \mathbf{w} \cdot \nabla F_s = 0. \tag{2.32}$$

Dans le cas des glaciers, l'interface air-glace peut évoluer suite à des précipitations neigeuses ou à des pertes de glace. Il est donc pertinent d'introduire la quantité $a_s = (\mathbf{w} - \mathbf{v}) \cdot \mathbf{n}$ appelée fonction d'accumulation-ablation qui représente le flux net de matière au travers de l'interface air-glace. Cette quantité est choisie positive pour un gain de masse et négative pour une perte de masse.

Nous pouvons alors réécrire l'équation (2.32) sous la forme

$$\frac{\partial F_s}{\partial t} + \mathbf{v} \cdot \nabla F_s = -N_s a_s, \tag{2.33}$$

ou encore en faisant intervenir $h(x, y, t)$

$$\frac{\partial h}{\partial t} + u_s \frac{\partial h}{\partial x} + v_s \frac{\partial h}{\partial y} - w_s = N_s a_s. \tag{2.34}$$

Cette dernière relation exprime simplement que l'évolution de Γ_s est liée à la fois au mouvement du fluide au niveau de cette interface et au gain ou à la perte de masse au travers de celle-ci. Pour l'interface glace-sol, nous avons simplement

$$\frac{\partial b}{\partial t} + u_b \frac{\partial b}{\partial x} + v_b \frac{\partial b}{\partial y} - w_b = N_b a_b. \tag{2.35}$$

où a_b, appelé le taux de fonte basale, est à présent choisi positif pour une perte de masse (fonte de glace qui s'accompagne d'une infiltration d'eau dans la roche).

2.5.2. Conditions aux limites à l'interface air-glace

L'interface air-glace est traitée comme une surface libre, ce qui revient à supposer que les forces surfaciques qui s'y exercent sont nulles. Nous obtenons alors une condition aux limites de Neumann homogène sur Γ_s à savoir

$$\sigma \cdot \mathbf{n} = \mathbf{0}. \tag{2.36}$$

Pour la résolution de l'équation de la chaleur, nous appliquons une condition aux limites de Dirichlet. Sur Γ_s, nous imposons ainsi

$$T = T_s\,. \tag{2.37}$$

La température de surface T_s est supposée être la température de l'air. En pratique, celle-ci peut être déterminée à partir de mesures de température réalisées à l'emplacement du glacier.

2.5.3. Conditions aux limites à l'interface glace-sol

L'écriture des conditions aux limites à l'interface glace-sol est plus compliquée. En fait, il existe plusieurs façons de les définir. En général, deux cas sont à envisager. Dans le premier cas, la température du sol rocheux est suffisamment basse pour que la glace soit supposée gelée au sol et une condition aux limites de Dirichlet homogène ($\mathbf{v} = \mathbf{0}$) est imposée à la vitesse. Cette situation s'applique principalement aux glaciers dits froids tels que les calottes polaires pour lesquelles la température de la glace est suffisamment basse pour ne pas dépasser son point de fusion. Dans le second cas, la température est proche du point de fusion de la glace de sorte qu'un film d'eau peut se développer entre la glace et la roche. La glace est ainsi susceptible de glisser sur le sol. Cette condition convient surtout pour les glaciers qualifiés de tempérés tels que les glaciers des Alpes dont la température est proche du point de fusion de la glace. Le glissement au niveau du sol est décrit par des lois de glissement empiriques qui lient la contrainte tangentielle $\mathbf{t}_b$ à la base du glacier à sa vitesse tangentielle $\mathbf{v}_b$. À titre d'illustration, nous citons deux types de lois qui sont communément utilisées pour décrire le glissement basal des glaciers :

- La loi de glissement la plus simple suppose une relation linéaire entre la contrainte tangentielle et la vitesse tangentielle. Cette relation s'écrit

  $$\mathbf{t}_b = -C_b\,\mathbf{v}_b, \tag{2.38}$$

  où $C_b > 0$ représente le coefficient de glissement.

- Il est possible de considérer également des lois de glissement non linéaires du type loi de Weertman qui constituent une généralisation du cas linéaire. De telles lois peuvent s'écrire
$$\mathbf{t}_b = -C_b \|\mathbf{v}_b\|^{m-1} \mathbf{v}_b. \tag{2.39}$$
où m est un exposant et C_b le coefficient de glissement qui dépend de m.

Malgré la simplicité apparente de ces lois de glissement, l'étude du glissement basal des glaciers est en fait un sujet fort compliqué. En effet, il est difficile d'observer et d'étudier ce phénomène en réalité. Le glissement est notamment déterminé par la nature et le profil du sol rocheux ainsi que la pente de la base et la pression basale. La détermination des coefficients des lois de glissement présente donc une difficulté particulière.

Pour certains glaciers, la base peut se diviser en des régions où le glacier est gelé au sol et d'autres régions où il glisse sur son lit rocheux. Ce sera le cas typiquement pour les glaciers polythermiques qui combinent les caractéristiques des glaciers froids et tempérés.

Certains auteurs considèrent que le glissement du glacier ne se produit uniquement que pour $T = T_m$. En dessous de cette température, le glacier est gelé à sa base. Cette hypothèse présente plusieurs désavantages. D'une part, elle implique la possibilité d'avoir une discontinuité du champ de vitesse et d'autre part, elle ne représente pas la réalité physique, car le glissement se produit généralement dans une gamme de températures proches du point de fusion pour laquelle des petits films d'eau se développent. Il est toutefois possible de concilier la condition de non-glissement à la base pour de basses températures avec la condition de glissement pour des températures proches du point de fusion. Pour cela, il est intéressant de considérer une loi de glissement qui dépend de la température [22, 55]. Nous considérons dans ce cas une relation du type
$$\mathbf{t}_b = f(\mathbf{v}_b, T). \tag{2.40}$$
Pour l'équation de la chaleur, une condition de Neumann est généralement adoptée au niveau de la base du glacier, ce qui revient à imposer le flux de chaleur au niveau de l'interface glace-sol. Il y a trois contributions principales à ce flux de chaleur. La première est le flux de chaleur géothermique q_{geo}, la deuxième est liée à l'énergie dissipée par frottement sur le sol rocheux et la troisième correspond à une variation d'énergie suite à la fonte de la glace (chaleur latente de fusion L). Cette dernière contribution n'intervient que si la température de la glace atteint son point de fusion. Il est intéressant de remarquer que ce terme de chaleur latente est uniquement pris en compte à la frontière du glacier et non pas dans l'équation en volume pour la température. Nous obtenons ainsi la condition aux limites suivantes
$$k\,\nabla T \cdot \mathbf{n} = q_{\text{geo}} - \mathbf{t}_b \cdot \mathbf{v}_b - \begin{cases} \rho L a_b & \text{si } T = T_m \\ 0 & \text{si } T < T_m \end{cases}. \tag{2.41}$$
Dans cette relation, le flux géothermique q_{geo} doit être considéré comme une donnée du problème. Au contraire, le taux de fonte basale a_b pourra être calculé une fois que les parties de la frontière où la température atteint le point de fusion de la glace auront été déterminées. La prise en compte numérique de la contrainte thermique $T \leqslant T_m$ pour la glace sera étudiée plus en détail dans le chapitre 4.

Nous terminons cette section par une petite remarque. Les conditions aux limites que nous avons écrites font intervenir certains paramètres notamment a_s et q_{geo}. La détermination de ces quantités peut être relativement difficile et introduire des incertitudes importantes dans les modèles. En particulier, le flux de chaleur géothermique q_{geo} est une quantité difficile à déterminer, ce qui peut rendre incertaines les valeurs de température basale calculées.

2.6. Modèle stationnaire des glaciers

Jusqu'à présent, notre présentation s'est voulue relativement générale. Nous introduisons désormais le modèle que nous utiliserons dans les prochains chapitres. Par la suite, nous nous limiterons à l'étude stationnaire des glaciers. Cette hypothèse revient à considérer l'équation de la chaleur en régime stationnaire et à supposer les interfaces avec l'extérieur comme figées dans le temps. Nous nous plaçons ainsi dans le cadre d'une approche dite de diagnostic dont le but est de mettre notamment en évidence certains processus physiques en action dans les glaciers. Pour information, la seconde approche qui est qualifiée de prédictive est utilisée pour prévoir l'évolution future des glaciers ou encore reconstruire le comportement passé de ces derniers.

Sous l'hypothèse de stationnarité, le système d'équations différentielles que nous sommes amené à résoudre sur le domaine Ω est donné par

$$\begin{cases} \operatorname{div} \mathbf{v} = 0, \\ -\operatorname{div}(2\mu\dot{\varepsilon}) + \nabla p = \rho \mathbf{g}, \\ -\operatorname{div}(k\nabla T) + \rho c\, \mathbf{v} \cdot \nabla T = 4\mu d_e^2, \quad \text{avec } T \leqslant T_m \text{ dans } \Omega, \end{cases} \tag{2.42}$$

où μ est donné par la loi de Glen régularisée (2.28) et k et c sont donnés respectivement par les relations (2.30) et (2.31).

Le système (2.42) représente le problème thermomécanique de Stokes en régime stationnaire. Il définit un ensemble d'équations pour les inconnues $\mathbf{v}, p$ et T. Ce problème est caractérisé par un haut degré de non-linéarité à cause de la dépendance des propriétés thermomécaniques de la glace par rapport aux inconnues du problème. De plus, les problèmes mécanique et thermique ne sont pas indépendants, ce qui rend le problème thermomécanique couplé. La dépendance de la partie mécanique par rapport à la température se fait au travers de la viscosité du fluide. La vitesse du fluide apparaît quant à elle dans les termes de convection et de dissipation visqueuse de l'équation de la chaleur.

Sur la frontière du domaine, nous imposons les conditions aux limites suivantes :

- Sur Γ_s, nous adoptons les conditions présentées à la sous-section 2.5.2 à savoir

$$\begin{cases} \sigma \cdot \mathbf{n} = \mathbf{0}, \\ T = T_s. \end{cases} \tag{2.43}$$

- Sur Γ_b, nous avons décidé d'adopter une loi de glissement pour la composante tangentielle de la vitesse avec un coefficient de friction qui dépend de la température. La composante normale de la vitesse peut être déterminée à partir de la relation (2.35) en considérant que l'interface n'évolue pas dans le temps. La condition aux limites pour la température est donnée par la relation (2.41). Ceci nous donne finalement les conditions aux limites suivantes :

$$\begin{cases} \mathbf{v} \cdot \mathbf{n} = a_b, \\ \mathbf{t}_b = -C_b\, \mathbf{v}_b, \\ k\, \nabla T \cdot \mathbf{n} = q_{\text{geo}} - \mathbf{t}_b \cdot \mathbf{v}_b - \begin{cases} \rho L a_b & \text{si } T = T_m \\ 0 & \text{si } T < T_m \end{cases}. \end{cases} \tag{2.44}$$

Le taux de fonte qui apparaît dans la condition $\mathbf{v} \cdot \mathbf{n} = a_b$ dépend évidemment de la température. Si nous nous intéressons uniquement au problème mécanique, cette quantité peut être traitée comme une donnée du problème.

L'expression du coefficient de friction C_b est choisie de manière à rendre compte de la plus grande difficulté de glissement pour des températures éloignées du point de fusion de la glace. Un choix possible est donné par une relation exponentielle du type [55]

$$C_b = C_0 \exp\left(\gamma(T_m - T)\right), \tag{2.45}$$

où C_0 est le coefficient de friction à $T = T_m$ et γ est un paramètre qui est choisi de manière à ce que le glissement soit important uniquement sur une plage de températures proches du point de fusion.

Remarquons que le modèle que nous avons présenté est insuffisant pour assurer l'existence d'une solution unique. En effet, les conditions aux limites fournissent une contrainte uniquement sur la valeur de la composante normale de la vitesse sur Γ_b. Cette condition ne suffit pas à déterminer l'ensemble des composantes du vecteur vitesse. Pour remédier à ce problème, nous devons inclure dans notre modèle une nouvelle région de la frontière où des conditions supplémentaires sur les composantes de la vitesse sont imposées. Nous reviendrons sur ce point à la section 6.1 lorsque nous nous intéresserons au problème couplé.

Dans la suite de ce travail, nous allons étudier comment il est possible de résoudre le système d'équations présentées. Dans les deux prochains chapitres, nous étudierons séparément les parties mécanique et thermique d'un point de vue mathématique afin d'en dégager des éléments importants pour la résolution numérique de ces équations.

3. Étude analytique du problème de Stokes

Ce troisième chapitre a pour objectif d'étudier de manière analytique le problème de Stokes qui décrit le comportement mécanique des glaciers. Nous entamerons le chapitre à la section 3.1 par quelques rappels d'analyse fonctionnelle dont nous ferons usage dans la suite de ce travail. La section 3.2 sera consacrée à l'étude des équations de Stokes linéaires. En considérant la formulation faible de ces équations, nous verrons qu'il est possible de réécrire le problème de Stokes sous la forme d'un problème de minimisation sous contrainte et d'un problème de point de selle. Nous introduirons également la condition de Babuska-Brezzi qui assure l'existence et l'unicité d'une solution au problème de Stokes. Ensuite, le problème continu sera discrétisé par une méthode classique de Galerkin et nous en déduirons certaines restrictions quant au choix des fonctions de base pour la résolution du problème analytique par éléments finis. À la fois dans le cas du problème continu et discret, nous verrons que la difficulté de la résolution sera liée à la condition d'incompressibilité du fluide. La section 3.3 terminera ce chapitre par une extension des résultats pour le problème de Stokes linéaire au problème non linéaire. Cette extension au problème non linéaire sera menée de manière similaire au cas linéaire et nous pourrons constater que les mêmes résultats sont applicables pour les deux cas. Outre son intérêt pour l'application à des fluides fortement visqueux, l'étude du problème de Stokes offre une première approche de la difficulté de la discrétisation par éléments finis dans le cadre général des fluides incompressibles.

3.1. Rappels d'analyse fonctionnelle

L'établissement d'une formulation faible pour les problèmes de Stokes linéaire et non linéaire offre la possibilité d'utiliser les outils de l'analyse fonctionnelle afin de démontrer plusieurs propriétés de la solution notamment son existence et son unicité. Avant de développer le problème de Stokes, nous rappellerons brièvement dans cette section quelques éléments d'analyse fonctionnelle qui seront utilisés par la suite. L'essentiel des définitions que nous présentons dans cette section-ci sont reprises de Brezis [8]. Le lecteur intéressé pourra se référer à cet ouvrage pour une présentation détaillée de l'analyse fonctionnelle.

3.1.1. Espaces vectoriels de fonctions

L'analyse fonctionnelle introduit différents types d'espaces vectoriels. Parmi ceux-ci, les espaces traditionnels qui apparaissent dans la formulation faible d'équations aux dérivées partielles s'inscrivent dans les espaces de Banach qui sont définis comme des espaces vectoriels normés complets. Par la suite, trois types d'espaces vectoriels particuliers nous seront utiles. Pour les définir, nous considérons un ensemble Ω inclus dans $\mathbb{R}^n$.

Le premier type d'espaces vectoriels que nous introduisons sont les espaces de Lebesgue dénotés $\mathbb{L}^p(\Omega)$ avec $1 \leqslant p < \infty$. La définition de ces espaces est reprise ci-dessous [1].

Définition 3.1. On appelle espace de Lebesgue, l'espace vectoriel défini par

$$\mathbb{L}^p(\Omega) = \{f : \Omega \to \mathbb{R};\ \int_\Omega |f|^p \, d\Omega < \infty\}. \tag{3.1}$$

Si l'ensemble Ω est de mesure finie, nous avons la relation d'inclusion $\mathbb{L}^q(\Omega) \subseteq \mathbb{L}^p(\Omega)$ pour $1 \leqslant p \leqslant q < \infty$. Aux éléments de $\mathbb{L}^p(\Omega)$, il est possible d'associer une norme définie par

$$\|f\|_{\mathbb{L}^p(\Omega)} = \left(\int_\Omega |f|^p \, d\Omega\right)^{1/p}. \tag{3.2}$$

Le deuxième type d'espaces vectoriels que nous présentons sont les espaces de Hilbert. Afin de les définir, nous introduisons l'opération de produit scalaire entre deux éléments f et g d'un espace vectoriel. Cette opération sera notée de manière générique sous la forme (f, g) dans la suite de ce chapitre. Les espaces de Hilbert $\mathbb{H}$ sont caractérisés par la définition 3.2.

Définition 3.2. Un espace de Hilbert est un espace vectoriel $\mathbb{H}$ muni d'un produit scalaire (f, g) et qui est complet pour la norme $(f, f)^{1/2}$.

Le dernier type d'espaces vectoriels que nous présentons sont les espaces de Sobolev. Ces espaces vectoriels définissent le cadre adéquat dans lequel seront recherchées les solutions d'une équation aux dérivées partielles exprimée sous une forme faible. Les espaces de Sobolev permettent de définir correctement l'ensemble des expressions qui apparaissent dans la formulation faible. Ces espaces donnent aussi l'occasion de travailler avec des solutions qui sortent de l'ensemble traditionnel $\mathbb{C}^m(\Omega)$ des fonctions m fois continûment dérivables.

La définition formelle de l'espace de Sobolev $\mathbb{W}^{1,p}(\Omega)$ avec $1 \leqslant p < \infty$ est donnée ci-dessous.

Définition 3.3. L'espace de Sobolev $\mathbb{W}^{1,p}(\Omega)$ est défini par

$$\begin{aligned}\mathbb{W}^{1,p}(\Omega) =& \{f \in \mathbb{L}^p(\Omega);\ \exists\, g_1,\, g_2, \ldots,\, g_n \in \mathbb{L}^p(\Omega) \text{ tels que} \\ & \int_\Omega f \frac{\partial \varphi}{\partial x_i} \, d\Omega = -\int_\Omega g_i \, \varphi \, d\Omega \ \forall \varphi \in \mathbb{C}_c^\infty(\Omega)\ \forall i = 1, 2, \ldots, n\}.\end{aligned} \tag{3.3}$$

Dans la définition 3.3, la fonction φ est une fonction test qui appartient à l'ensemble $\mathbb{C}_c^\infty(\Omega)$ des fonctions indéfiniment continûment dérivables sur Ω et qui sont à support compact dans Ω. La quantité $g_i = \frac{\partial f}{\partial x_i}$ peut être vue comme la dérivée partielle au sens des distributions de f. La définition 3.3 signifie alors qu'une fonction $f \in \mathbb{L}^p(\Omega)$ appartient à $\mathbb{W}^{1,p}(\Omega)$ si toutes ses dérivées partielles premières définies au sens des distributions appartiennent à $\mathbb{L}^p(\Omega)$. L'espace de Sobolev $\mathbb{W}^{1,2}(\Omega)$ est un espace de Hilbert qui est noté habituellement $\mathbb{H}^1(\Omega)$.

1. La définition rigoureuse des espaces de Lebesgue demande de travailler avec des fonctions mesurables. Nous supposerons par la suite que cette condition est toujours vérifiée. De plus, deux fonctions de $\mathbb{L}^p(\Omega)$ seront dites équivalentes si elles sont égales presque partout sur Ω. L'ensemble des fonctions équivalentes à f forme une classe d'équivalence notée $[f]$. Les espaces de Lebesgue sont ainsi définis à partir de classes d'équivalence. Il est toutefois courant d'identifier une fonction f à sa classe d'équivalence de sorte que nous adopterons également cette notation. Ainsi, lorsque nous écrirons qu'une fonction f est solution d'une formulation variationnelle, cela signifiera que toutes les fonctions équivalentes à f le seront également. Dès lors, quand nous parlerons de l'unicité de la solution à une formulation variationnelle, cela sous-entendra la possibilité d'avoir deux fonctions non équivalentes qui la satisfont.

L'espace $\mathbb{W}^{1,p}(\Omega)$ est muni de la norme

$$\|f\|_{\mathbb{W}^{1,p}(\Omega)} = \left(\|f\|^p_{\mathbb{L}^p(\Omega)} + \sum_{i=1}^{n} \left\| \frac{\partial f}{\partial x_i} \right\|^p_{\mathbb{L}^p(\Omega)} \right)^{1/p}. \tag{3.4}$$

Notons enfin que les normes (3.2) et (3.4) se généralisent sans problème pour des fonctions à valeurs vectorielles $\mathbf{f} = (f_1, \ldots, f_d)$. Nous écrirons ainsi

$$\|\mathbf{f}\|_{[\mathbb{L}^p(\Omega)]^d} = \left(\sum_{i=1}^{d} \|f_i\|^p_{\mathbb{L}^p(\Omega)} \right)^{1/p}, \tag{3.5}$$

$$\|\mathbf{f}\|_{[\mathbb{W}^{1,p}(\Omega)]^d} = \left(\sum_{i=1}^{d} \|f_i\|^p_{\mathbb{W}^{1,p}(\Omega)} \right)^{1/p}. \tag{3.6}$$

3.1.2. Opérations sur les espaces vectoriels de fonctions

Il est possible d'introduire des opérations sur les éléments d'un espace vectoriel. Dans la suite de ce chapitre, nous serons amené à utiliser des applications définies dans un espace vectoriel et à valeurs dans $\mathbb{R}$. Ces applications particulières seront appelées formes.

Certaines propriétés des formes nous seront utiles par la suite. Une forme $a : V \times V \to \mathbb{R}$ sera qualifiée de [56] :

- *bilinéaire* si elle est linéaire par rapport à ses deux arguments ;
- *continue* s'il existe une constante $M > 0$ telle que
$$|a(u,v)| \leqslant M\|u\|_V\|v\|_V \quad \forall u, v \in V \ ;$$
- *symétrique* si $a(u,v) = a(v,u) \quad \forall u, v \in V$;
- *positive* si $a(v,v) > 0 \quad \forall v \neq 0 \in V$;
- *coercive* s'il existe une constante $\alpha > 0$ telle que
$$a(v,v) \geqslant \alpha\|v\|^2_V \quad \forall v \in V.$$

Nous présentons également le lemme de Lax-Milgram [18] qui est notamment utilisé pour démontrer l'existence et l'unicité de la solution à une formulation variationnelle.

Lemme 3.1 (Lax-Milgram). *Soit $a(\,\cdot\,,\cdot\,)$ une forme bilinéaire sur un espace de Hilbert $\mathbb{H}$ muni d'une norme $\|\cdot\|_{\mathbb{H}}$. Si $a(\,\cdot\,,\cdot\,)$ est continue, c'est-à-dire*

$$\exists\, M > 0 \text{ tel que } |a(w,v)| \leqslant M\|w\|_{\mathbb{H}}\|v\|_{\mathbb{H}} \quad \forall w, v \in \mathbb{H},$$

et si $a(\,\cdot\,,\cdot\,)$ est coercive à savoir

$$\exists\, \alpha > 0 \text{ tel que } a(v,v) \geqslant \alpha\, \|v\|^2_{\mathbb{H}} \quad \forall v \in \mathbb{H},$$

alors pour toute application $l(\,\cdot\,)$ linéaire bornée sur $\mathbb{H}$ (donc continue), c'est-à-dire

$$\exists\, \gamma > 0 \text{ tel que } |l(w)| \leqslant \gamma\|w\|_{\mathbb{H}} \quad \forall w \in \mathbb{H},$$

il existe un unique $u \in \mathbb{H}$ tel que la relation

$$a(w,u) = l(w) \quad \forall w \in \mathbb{H}$$

est satisfaite.

3.2. Problème de Stokes linéaire

3.2.1. Formulation forte du problème de Stokes

Nous considérons à présent l'écoulement d'un volume de glace $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ dont le comportement rhéologique est celui d'un fluide newtonien de viscosité dynamique $\mu > 0$ constante. L'écoulement du glacier est décrit par le problème de Stokes linéaire

$$\begin{cases} -\operatorname{div}\sigma = \mathbf{f} & \text{dans } \Omega, \\ \operatorname{div}\mathbf{v} = 0 & \text{dans } \Omega, \\ \sigma \cdot \mathbf{n} = \mathbf{0} & \text{sur } \Gamma_N, \\ \mathbf{v} = \mathbf{0} & \text{sur } \Gamma_D, \end{cases} \tag{3.7}$$

où les surfaces Γ_N et Γ_D représentent les parties de la frontière Γ de Ω ($\Gamma_N \cup \Gamma_D = \Gamma$, $\Gamma_N \cap \Gamma_D = \varnothing$) sur lesquelles sont appliquées respectivement des conditions aux limites de Neumann et de Dirichlet homogènes. Au système d'équations (3.7), il faut encore adjoindre la loi de comportement $\sigma = -p\underline{\underline{I}} + 2\mu\dot{\varepsilon}$ afin de spécifier complètement le modèle.

Les deux champs inconnus dans le problème de Stokes sont les champs de vitesse $\mathbf{v}$ et de pression p. Cependant, dans le cas d'un fluide incompressible, la pression joue un rôle particulier. En effet, les équations pour un fluide incompressible amènent à résoudre l'équation de quantité de mouvement sous la contrainte que le champ de vitesse soit à divergence nulle. La pression peut ainsi être vue comme un nouveau degré de liberté qui va permettre de satisfaire la contrainte d'incompressibilité. Comme nous le verrons par la suite, la pression peut en fait être considérée comme un multiplicateur de Lagrange.

3.2.2. Formulation faible du problème de Stokes

Pour établir la forme faible du problème de Stokes linéaire, nous supposons que $\mathbf{f} \in \left[\mathbb{L}^2(\Omega)\right]^d$, $p \in Q = \mathbb{L}^2(\Omega)$ [2] et que $\mathbf{v} \in V$ où l'ensemble V est défini par

$$V = \{\mathbf{w} \in \left[\mathbb{H}^1(\Omega)\right]^d : \mathbf{w} = \mathbf{0} \text{ sur } \Gamma_D\}. \tag{3.8}$$

Le choix de ces espaces de fonctions assure que les intégrales de la formulation faible seront correctement définies. Pour obtenir la forme faible du problème de Stokes, nous partons du système d'équations différentielles

$$\begin{cases} -\operatorname{div}\sigma = \mathbf{f}, \\ \operatorname{div}\mathbf{v} = 0. \end{cases} \tag{3.9}$$

Nous considérons à présent un couple de fonctions test $(\mathbf{w}, q) \in V \times Q$ par lequel nous multiplions les équations (3.9) avant de les intégrer sur le domaine Ω. Nous obtenons alors

$$\begin{cases} -\displaystyle\int_\Omega (\operatorname{div}\sigma) \cdot \mathbf{w}\, d\Omega = \int_\Omega \mathbf{f} \cdot \mathbf{w}\, d\Omega, \\ \displaystyle\int_\Omega q \operatorname{div}\mathbf{v}\, d\Omega = 0. \end{cases} \tag{3.10}$$

2. Si les conditions aux limites consistent uniquement en des conditions aux limites de Dirichlet, la pression solution du problème de Stokes sera définie à une constante additive près. Dans ce cas, une condition supplémentaire doit être imposée sur l'espace de fonctions auquel le champ de pression appartient. En pratique, le champ de pression sera choisi pour être en moyenne nul sur le domaine ou sa valeur sera imposée en un point du domaine.

En utilisant le théorème de la divergence, nous pouvons encore écrire

$$\begin{cases} -\displaystyle\int_{\Gamma_D} (\sigma \cdot \mathbf{w}) \cdot \mathbf{n} \, d\Gamma_D - \int_{\Gamma_N} (\sigma \cdot \mathbf{w}) \cdot \mathbf{n} \, d\Gamma_N + \int_\Omega \sigma : \nabla \mathbf{w} \, d\Omega = \int_\Omega \mathbf{f} \cdot \mathbf{w} \, d\Omega, \\ \displaystyle\int_\Omega q \operatorname{div} \mathbf{v} \, d\Omega = 0. \end{cases} \tag{3.11}$$

L'intégrale sur Γ_N disparaît à cause des conditions de Neumann homogènes imposées sur cette frontière. L'intégrale de surface sur Γ_D est également nulle, car nous nous sommes limité à des fonctions test $\mathbf{w}$ qui s'annulent sur Γ_D.

Pour finir, nous exploitons la symétrie du tenseur des contraintes et la loi de comportement pour les contraintes, ce qui nous donne l'expression suivante pour la forme faible :

$$\begin{cases} \displaystyle\int_\Omega 2\mu \, \dot{\varepsilon}(\mathbf{v}) : \dot{\varepsilon}(\mathbf{w}) \, d\Omega - \int_\Omega p \operatorname{div} \mathbf{w} \, d\Omega = \int_\Omega \mathbf{f} \cdot \mathbf{w} \, d\Omega, \\ \displaystyle\int_\Omega q \operatorname{div} \mathbf{v} \, d\Omega = 0. \end{cases} \tag{3.12}$$

Pour le problème linéaire, il est intéressant d'introduire les formes bilinéaires suivantes :

$$\begin{cases} a(\mathbf{v}, \mathbf{w}) = \displaystyle\int_\Omega 2\mu \, \dot{\varepsilon}(\mathbf{v}) : \dot{\varepsilon}(\mathbf{w}) \, d\Omega, \\ b(\mathbf{v}, q) = -\displaystyle\int_\Omega q \operatorname{div} \mathbf{v} \, d\Omega. \end{cases} \tag{3.13}$$

Ces deux formes sont définies respectivement de $V \times V$ dans $\mathbb{R}$ et de $V \times Q$ dans $\mathbb{R}$. De plus, il peut être montré que la forme bilinéaire a est une forme continue, symétrique, positive et coercive. La forme b est quant à elle une forme bilinéaire et continue. Notons que le caractère bilinéaire et symétrique de la forme a n'est valable que si μ ne dépend pas de $\mathbf{v}$. Ces deux propriétés ne seront donc plus applicables quand nous traiterons le problème non linéaire à la section 3.3.

Le produit scalaire $(\mathbf{f}, \mathbf{w})$ entre deux vecteurs peut s'écrire sous la forme

$$(\mathbf{f}, \mathbf{w}) = \int_\Omega \mathbf{f} \cdot \mathbf{w} \, d\Omega. \tag{3.14}$$

La formulation variationnelle du problème de Stokes linéaire peut finalement s'exprimer de la manière suivante :

Formulation variationnelle 3.1. Trouver $\mathbf{v} \in V = \{\mathbf{w} \in \left[\mathbb{H}^1(\Omega)\right]^d : \mathbf{w} = \mathbf{0} \text{ sur } \Gamma_D\}$ et $p \in Q = \mathbb{L}^2(\Omega)$ tels que

$$\begin{cases} a(\mathbf{v}, \mathbf{w}) + b(\mathbf{w}, p) = (\mathbf{f}, \mathbf{w}), \\ b(\mathbf{v}, q) = 0, \end{cases}$$

pour tout couple $(\mathbf{w}, q) \in V \times Q$.

Même si nous avons peu insisté sur ce point jusqu'à présent, la formulation faible du problème de Stokes est une formulation variationnelle mixte dont les inconnues primaires sont les champs de vitesse et de pression dans le glacier. La résolution numérique de ce problème mènera à l'utilisation d'éléments finis mixtes qui devront être choisis de manière appropriée.

3.2.3. Problèmes d'optimisation pour le problème de Stokes

La solution de certaines équations aux dérivées partielles peut s'exprimer comme la solution d'un problème d'optimisation. Dans le cas du problème de Stokes, il est possible d'établir de tels problèmes. En fait, nous allons voir que nous pouvons définir un problème de minimisation et un problème de point de selle pour le problème de Stokes. Pour établir le problème de minimisation, nous allons réécrire la formulation variationnelle 3.1 en fonction uniquement du champ de vitesse. Cela est rendu possible par le fait que l'équation de continuité du fluide correspond à une contrainte imposée au champ de vitesse (champ de vitesse solénoïdal). La vérification de la formulation variationnelle 3.1 par le champ de vitesse assure que la solution $\mathbf{v}$ est bien à divergence nulle. Toutefois, il est possible d'imposer que le champ de vitesse satisfasse d'emblée la contrainte d'incompressibilité. Dans ce cas, nous obtenons une formulation variationnelle uniquement pour le champ de vitesse.

Formulation variationnelle 3.2. Trouver $\mathbf{v} \in V_{\text{div}} = \{\mathbf{w} \in \left[\mathbb{H}^1(\Omega)\right]^d : \mathbf{w} = \mathbf{0} \text{ sur } \Gamma_D \text{ et } \operatorname{div} \mathbf{w} = 0 \text{ dans } \Omega\}$ tel que

$$a(\mathbf{v}, \mathbf{w}) = (\mathbf{f}, \mathbf{w}),$$

pour tout $\mathbf{w} \in V_{\text{div}}$.

La continuité et la coercivité de la forme bilinéaire $a(\,\cdot\,,\,\cdot\,)$ permettent en vertu du lemme de Lax-Milgram 3.1 d'assurer l'existence d'une unique solution $\mathbf{v} \in V_{\text{div}}$ au problème variationnel 3.2 [19]. Une des conséquences du lemme de Lax-Milgram est donnée par le corollaire ci-dessous [56].

Corollaire 3.1. *La solution du problème variationnel 3.2 est bornée de la façon suivante :*

$$\|\mathbf{v}\|_{V_{\text{div}}} \leqslant \frac{1}{\alpha} \|\mathbf{f}\|_{[\mathbb{L}^2(\Omega)]^d},$$

où α représente la constante de coercivité.

Le corollaire 3.1 donne une estimation a priori de la solution et montre que celle-ci est bornée par le terme de force volumique $\mathbf{f}$. De plus, en anticipant déjà sur la question de la résolution numérique du problème de Stokes, il est possible d'appliquer ce corollaire à une approximation $\mathbf{v}_h$ obtenue par la méthode de Galerkin. Ce corollaire montre en particulier que la méthode de Galerkin appliquée à la formulation variationnelle 3.2 est stable dans le sens que la solution approchée dépend continûment des données du problème.

La forme $a(\,\cdot\,,\,\cdot\,)$ est une forme symétrique, ce qui permet de réécrire le problème de Stokes sous la forme d'un problème de minimisation. Pour cela, nous introduisons une fonctionnelle $J(\mathbf{w}) = \frac{1}{2}a(\mathbf{w}, \mathbf{w}) - (\mathbf{f}, \mathbf{w})$ telle que la solution $\mathbf{v} \in V_{\text{div}}$ du problème variationnel 3.2 satisfait

$$J(\mathbf{v}) = \min_{\mathbf{w} \in V_{\text{div}}} J(\mathbf{w}). \tag{3.15}$$

L'équivalence entre la formulation variationnelle 3.2 et le problème de minimisation (3.15) est démontrée en annexe A. Le problème de minimisation (3.15) peut aussi être réécrit sous la forme d'un problème de minimisation sous contrainte équivalent.

La fonction $\mathbf{v} \in V$ satisfait la formulation variationnelle 3.2 si $\mathbf{v}$ est une solution du problème d'optimisation

$$\min_{\mathbf{w}\in V} J(\mathbf{w}) = \min_{\mathbf{w}\in V} \frac{1}{2} a(\mathbf{w},\mathbf{w}) - (\mathbf{f},\mathbf{w})$$

sous la contrainte que

$$\operatorname{div} \mathbf{w} = 0.$$

Ce problème d'optimisation peut en fait être interprété d'un point de vue plus physique. L'expression $a(\mathbf{w},\mathbf{w})$ correspond au taux de dissipation visqueuse dans le fluide (conversion d'énergie cinétique en énergie thermique causée par les tensions visqueuses). La positivité de la forme a est en accord avec le fait que le taux de dissipation visqueuse est également une quantité positive. Le produit scalaire $(\mathbf{f},\mathbf{w})$ représente le taux de variation de l'énergie potentielle des forces de volume. Ainsi, le problème de minimisation consiste en la minimisation d'une fonctionnelle qui est reliée au taux de variation de l'énergie cinétique du fluide. La solution $\mathbf{v}$ du problème de Stokes satisfait la relation $a(\mathbf{v},\mathbf{v}) = (\mathbf{f},\mathbf{v})$, ce qui signifie qu'à l'état stationnaire la perte d'énergie cinétique liée à la dissipation visqueuse est compensée exactement par le gain d'énergie cinétique lié au travail des forces de volume.

Le problème de minimisation sous contrainte (3.15) peut aussi s'exprimer sous la forme d'un problème quasi non contraint si nous introduisons le lagrangien

$$\mathcal{L}(\mathbf{w},q) = \frac{1}{2} a(\mathbf{w},\mathbf{w}) + b(\mathbf{w},q) - (\mathbf{f},\mathbf{w}). \tag{3.16}$$

Dans l'expression du lagrangien, nous constatons que la contrainte d'incompressibilité a été pénalisée. La pression q se présente sous la forme d'un multiplicateur de Lagrange encore appelé variable duale. Dans ce cas, la solution $(\mathbf{v},p)$ du problème de Stokes est un point de selle du lagrangien, c'est-à-dire

$$\forall q \in Q \quad \mathcal{L}(\mathbf{v},q) \leqslant \mathcal{L}(\mathbf{v},p) \leqslant \mathcal{L}(\mathbf{w},p) \quad \forall \mathbf{w} \in V. \tag{3.17}$$

Intuitivement, nous pouvons nous convaincre que si le couple $(\mathbf{v},p)$ est un point de selle du lagrangien, alors $\mathbf{v}$ doit satisfaire le problème d'optimisation sous contrainte (3.15). Pour cela, nous réécrivons l'expression (3.17) sous la forme

$$\forall q \in Q \quad J(\mathbf{v}) + b(\mathbf{v},q) \leqslant J(\mathbf{v}) + b(\mathbf{v},p) \leqslant J(\mathbf{w}) + b(\mathbf{w},p) \quad \forall \mathbf{w} \in V. \tag{3.18}$$

La fonction q étant quelconque dans Q, le seul moyen pour satisfaire la première inégalité est d'avoir $b(\mathbf{v},q) = 0\ \forall q \in Q$, ce qui n'est possible que si $\mathbf{v}$ satisfait la contrainte d'incompressibilité. Si $\mathbf{v}$ est à divergence nulle, la seconde inégalité devient $J(\mathbf{v}) \leqslant J(\mathbf{w}) + b(\mathbf{w},p)$, ce qui implique que $\mathbf{v}$ minimise $J(\mathbf{w})\ \forall \mathbf{w} \in V_{\text{div}}$ et donc que $\mathbf{v}$ est bien une solution du problème de Stokes. Ainsi, la résolution du problème de minimisation sous contrainte est équivalente à la résolution du problème de point de selle.

Il est intéressant de constater la manière dont l'interprétation de la pression change suivant que le fluide soit considéré comme compressible ou bien sous l'hypothèse d'incompressibilité.

Dans le cas d'un fluide compressible, la pression correspond en réalité à une variable d'état thermodynamique du système qui est reliée à d'autres variables d'état telles que la masse volumique et la température par une équation d'état. Pour un fluide incompressible, cette vision thermodynamique de la pression doit être modifiée. En effet, la pression dans le fluide à un instant donné va dépendre du champ de vitesse dans tout l'écoulement au même instant, ce qui est possible comme la vitesse du son est infinie dans un fluide incompressible (vitesse infinie pour la propagation de l'information). Cette dépendance de la pression envers le champ de vitesse dans le fluide peut se comprendre si la pression est interprétée comme un multiplicateur de Lagrange pour la contrainte d'incompressibilité.

3.2.4. Existence et unicité de la solution du problème continu

Précédemment, nous avons vu que la formulation variationnelle 3.2 permettait de montrer l'existence et l'unicité du champ de vitesse $\mathbf{v}$. La question de l'existence et de l'unicité du champ de pression p n'a cependant pas encore été discutée. C'est ce point-ci que nous allons à présent étudier.

L'existence d'un couple unique $(\mathbf{v}, p)$ solution du problème de Stokes est en fait déterminée par la surjectivité de l'opérateur divergence. Afin de mieux comprendre pourquoi cet opérateur peut poser un problème, nous présentons ci-dessous un raisonnement qualitatif qui mettra en évidence le rôle de la surjectivité de l'opérateur divergence dans l'unicité de la solution. Pour ce faire, nous reconsidérons la formulation faible de l'équation de quantité de mouvement du problème de Stokes à savoir

$$\int_\Omega 2\mu\, \dot{\varepsilon}(\mathbf{v}) : \dot{\varepsilon}(\mathbf{w})\, d\Omega - \int_\Omega p \operatorname{div} \mathbf{w}\, d\Omega = \int_\Omega \mathbf{f} \cdot \mathbf{w}\, d\Omega. \tag{3.19}$$

Comme nous l'avons vu à la sous-section 3.2.3, l'existence et l'unicité de $\mathbf{v}$ sont assurées par le lemme de Lax-Milgram appliqué au problème variationnel 3.2. Dès lors, le problème lié à l'existence d'un unique couple $(\mathbf{v}, p)$ provient de la possibilité d'avoir plusieurs fonctions p qui satisfont la formulation faible du problème de Stokes.

Supposons que nous ayons trouvé une solution $(\mathbf{v}, p_1)$ de la formulation variationnelle 3.1 et que nous désirions savoir sous quelles conditions cette solution est unique. Pour cela, nous regardons au terme $\int_\Omega p \operatorname{div} \mathbf{w}\, d\Omega$ qui intervient dans l'équation (3.19). Le champ de pression p peut s'écrire sous la forme $p = p_1 + p_2$. La solution $(\mathbf{v}, p_1)$ ne sera pas unique s'il existe une fonction p_2 différente de 0 sur Ω telle que $\int_\Omega p_2 \operatorname{div} \mathbf{w}\, d\Omega = 0$.

Afin de déterminer si une telle fonction p_2 existe, nous rappelons que les fonctions test $\mathbf{w}$ sont supposées appartenir à l'espace de Hilbert $\mathbb{H}^1(\Omega)$ de sorte que $\operatorname{div} \mathbf{w} \in \mathbb{L}^2(\Omega)$. La fonction p_2 appartient de plus à l'espace de Lebesgue $\mathbb{L}^2(\Omega)$. L'intégrale $\int_\Omega p_2 \operatorname{div} \mathbf{w}\, d\Omega$ correspond donc à un produit scalaire dans $\mathbb{L}^2(\Omega)$. Par simplicité, nous noterons cette intégrale par $(\operatorname{div} \mathbf{w}, p_2)_{\mathbb{L}^2}$.

Nous pouvons rechercher des fonctions $p_2 \in \mathbb{L}^2(\Omega)$ qui satisfont $(\operatorname{div} \mathbf{w}, p_2)_{\mathbb{L}^2} = 0\ \forall \mathbf{w} \in V$. La solution à ce problème est en fait déterminée par l'opérateur divergence. Dans le cas du problème de Stokes, l'opérateur divergence est vu comme une application qui fait correspondre un élément de $\mathbb{H}^1(\Omega)$ à un élément de $\mathbb{L}^2(\Omega)$. A priori, cette application n'est pas surjective de sorte que tout élément de $\mathbb{L}^2(\Omega)$ n'est pas nécessairement l'image d'un élément de $\mathbb{H}^1(\Omega)$ par cet opérateur.

Si l'opérateur divergence est surjectif, alors p_2 est nécessairement nul sur tout Ω et la solution du problème de Stokes est bien unique. En effet, si l'application divergence est surjective, tout élément de $\mathbb{L}^2(\Omega)$ peut être perçu comme l'image d'un élément de $\mathbb{H}^1(\Omega)$. Dans ce cas, rechercher des fonctions $p_2 \in \mathbb{L}^2(\Omega)$ telles que $(\operatorname{div} \mathbf{w}, p_2)_{\mathbb{L}^2} = 0 \ \forall \mathbf{w} \in V$ est équivalent à rechercher des fonctions $p_2 \in \mathbb{L}^2(\Omega)$ telles que $(q, p_2)_{\mathbb{L}^2} = 0 \ \forall q \in \mathbb{L}^2(\Omega)$. Or, le seul vecteur orthogonal à tous les éléments d'un espace vectoriel est le vecteur nul, ce qui signifie que nous avons $p_2 = 0$ sur Ω. Au contraire, si l'opérateur divergence n'est pas surjectif, alors nous pouvons trouver des fonctions $p_2 \neq 0$ sur Ω. Ces fonctions vérifient $(\operatorname{div} \mathbf{w}, p_2)_{\mathbb{L}^2} = 0 \ \forall \mathbf{w} \in V$, mais ne satisfont pas $(q, p_2)_{\mathbb{L}^2} = 0 \ \forall q \in \mathbb{L}^2(\Omega)$.

Le raisonnement qualitatif que nous avons mené a mis en évidence l'importance de la surjectivité de l'opérateur divergence pour l'existence d'une solution unique au problème de Stokes. Le lecteur intéressé par une étude plus détaillée des opérateurs surjectifs pourra consulter des livres de référence appropriés tels que Brezis [8].

Nous pouvons à présent énoncer les conditions sous lesquelles le problème de Stokes admet une unique solution. Pour cela, nous considérons le théorème suivant [19, 20] :

Théorème 3.1 (Babuska-Brezzi). *La formulation variationnelle 3.1 du problème de Stokes admet une unique solution $(\mathbf{v}, p) \in V \times Q$ sous les conditions suivantes :*

- *Les formes bilinéaires a et b sont continues respectivement sur $V \times V$ et $V \times Q$;*
- *La forme bilinéaire a est coercive sur V ;*
- *Il existe une constante $\beta > 0$ telle que la forme bilinéaire b vérifie*

$$\inf_{\substack{q \in Q \\ q \neq 0}} \sup_{\substack{\mathbf{w} \in V \\ \mathbf{w} \neq \mathbf{0}}} \frac{b(\mathbf{w}, q)}{\|q\|_Q \|\mathbf{w}\|_V} \geqslant \beta. \tag{3.20}$$

Dans le cas du problème de Stokes linéaire que nous étudions, les deux premières conditions sont nécessairement satisfaites. La relation (3.20) est connue sous le nom de condition de Ladyzhenskaya-Babuska-Brezzi et également sous les noms de condition de Babuska-Brezzi et de condition inf-sup. C'est la vérification de la condition inf-sup qui assure l'existence et l'unicité d'une solution au problème de Stokes. D'un point de vue mathématique, la condition inf-sup est équivalente à la condition de surjectivité de l'opérateur divergence dont nous avons montré l'importance précédemment.

Nous avons insisté sur la question de la surjectivité de l'application divergence, car elle permet d'avoir une meilleure compréhension intuitive de la condition inf-sup qui apparaît dans l'étude de certaines formulations variationnelles mixtes. Comme nous le verrons par la suite pour le problème discret, la condition inf-sup va restreindre le choix des fonctions de base utilisées pour discrétiser les champs de vitesse et de pression par la méthode des éléments finis. En particulier, l'espace des fonctions de base pour la vitesse devra être suffisamment « riche » comparé à l'espace des fonctions de base pour la pression.

3.2.5. Approximation de Galerkin du problème de Stokes

Jusqu'à présent, nous avons étudié le problème de Stokes continu. Nous abordons désormais ce problème sous l'angle de sa discrétisation par la méthode de Galerkin. Nous allons présenter les conséquences de cette discrétisation sur la résolution numérique du problème de Stokes.

Afin d'établir l'approximation de Galerkin du problème de Stokes, nous considérons les sous-espaces $V_h \subset V$ de dimension N et $Q_h \subset Q$ de dimension M. Nous introduisons respectivement les fonctions de base $\{\boldsymbol{\varphi}_j, j = 1, ..., N\}$ de V_h et $\{\phi_j, j = 1, ..., M\}$ de Q_h sur lesquelles nous pouvons décomposer les éléments de V_h et Q_h suivant :

$$\mathbf{w}_h(\mathbf{x}) = \sum_{j=1}^{N} w_j \, \boldsymbol{\varphi}_j(\mathbf{x}), \quad q_h(\mathbf{x}) = \sum_{k=1}^{M} q_k \, \phi_k(\mathbf{x}). \tag{3.21}$$

En adoptant les mêmes fonctions de base pour la discrétisation des fonctions test et de la solution du problème de Stokes, les coefficients $\{v_j\}$ et $\{p_k\}$ de la solution sont donnés par le système suivant :

$$\begin{cases} \displaystyle\sum_{j=1}^{N} v_j \, a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i) + \sum_{k=1}^{M} p_k \, b(\boldsymbol{\varphi}_i, \phi_k) = (\mathbf{f}, \boldsymbol{\varphi}_i), & i = 1, \ldots, N, \\ \displaystyle\sum_{j=1}^{N} v_j \, b(\boldsymbol{\varphi}_j, \phi_k) = 0, & k = 1, \ldots, M. \end{cases} \tag{3.22}$$

Ces équations peuvent également s'écrire sous forme matricielle à savoir

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix}, \tag{3.23}$$

où $A \in \mathbb{R}^{N \times N}$ $(A_{ij} = a(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i))$, $B \in \mathbb{R}^{M \times N}$ $(B_{ij} = b(\boldsymbol{\varphi}_j, \phi_i))$, $\mathbf{V} \in \mathbb{R}^N$ $(\mathbf{V}_i = v_i)$, $\mathbf{P} \in \mathbb{R}^M$ $(\mathbf{P}_i = p_i)$, $\mathbf{F} \in \mathbb{R}^N$ $(\mathbf{F}_i = (\mathbf{f}, \boldsymbol{\varphi}_i))$. La matrice A est appelée la matrice de viscosité et est l'équivalent en mécanique des fluides de la matrice de raideur en élasticité linéaire.

Le problème de Stokes discret aura une unique solution si la matrice du système (3.23) est non singulière. Cette question constituera l'objet de la prochaine sous-section.

3.2.6. Existence et unicité de la solution au problème discret

Dans le cas du problème continu, nous avons vu que l'existence et l'unicité d'une solution sont déterminées par deux facteurs principaux à savoir la coercivité de la forme bilinéaire a et la surjectivité de l'opérateur divergence. Dans le cas du problème discret, ces deux conditions doivent également être respectées. Bien que le passage du problème continu au problème discret conserve la coercivité de a au travers de la matrice A définie positive, le maintien de la surjectivité de l'application divergence n'est pas garanti. Dès lors, nous pouvons concevoir l'existence de nouvelles conditions d'existence et d'unicité de la solution pour le cas discret.

Nous pouvons montrer que les conditions pour résoudre le système (3.23) permettent de retrouver une condition d'existence et d'unicité de la solution qui est liée à la surjectivité de l'opérateur divergence discrétisé. La coercivité de la forme a implique qu'il existe au moins une solution au problème de Stokes discrétisé [7]. La question qui subsiste consiste à déterminer sous quelles conditions cette solution est unique. La solution du système (3.23) peut s'écrire comme la somme d'une solution homogène et d'une solution particulière. La solution homogène $(\mathbf{V}_h, \mathbf{P}_h)$ est obtenue en résolvant

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_h \\ \mathbf{P}_h \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{3.24}$$

Pour que le système (3.23) possède une solution unique, le système homogène doit avoir la solution nulle comme unique solution. Pour satisfaire $B\mathbf{V}_h = \mathbf{0}$, $\mathbf{V}_h$ doit appartenir au noyau de l'application B que nous noterons $\ker B$. En multipliant la première équation du système (3.24) par $\mathbf{V}_h^T$ et en considérant que $\mathbf{V}_h$ appartient à $\ker B$ nous obtenons

$$\mathbf{V}_h^T A \mathbf{V}_h = \mathbf{0}. \tag{3.25}$$

Comme A est une matrice définie positive, cette dernière relation implique que $\mathbf{V}_h = \mathbf{0}$. La valeur de $\mathbf{P}_h$ est obtenue en résolvant l'équation

$$B^T \mathbf{P}_h = \mathbf{0}. \tag{3.26}$$

Ce système d'équations aura l'unique solution $\mathbf{P}_h = \mathbf{0}$ si $\ker B^T = \{\mathbf{0}\}$. Dans le cas contraire, la solution du problème homogène n'est pas uniquement la solution nulle et le problème de Stokes discrétisé admet plusieurs solutions. Comme pour le cas continu, le problème de l'unicité de la solution est de nouveau lié au terme de pression. En utilisant le théorème du rang, nous en déduisons que la condition $\ker B^T = \{\mathbf{0}\}$ implique d'avoir $N \geqslant M$. Cette condition signifie que le nombre de fonctions de base utilisées pour discrétiser le champ de vitesse doit être supérieur au nombre de fonctions de base utilisées pour le champ de pression.

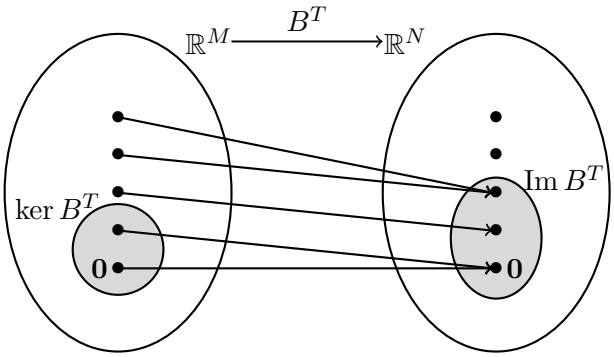


FIGURE 3.1.: Représentation graphique du noyau et de l'image d'une application linéaire.

La condition $\ker B^T = \{\mathbf{0}\}$ peut encore s'écrire sous d'autres formes en considérant les équivalences suivantes [7, 8] :

Théorème 3.2. *Soit B une application de $\mathbb{R}^N$ dans $\mathbb{R}^M$. Les propriétés suivantes sont équivalentes :*

- $\ker B^T = \{\mathbf{0}\}$,
- *B^T est injective,*
- *B est surjective,*
- *B est de rang maximal.*

L'unicité du problème discret nécessite ainsi la surjectivité de l'application B. Or, cette application résulte de la discrétisation de l'opérateur divergence. Ainsi, l'unicité de la solution du problème de Stokes dans le cas continu et discret est dictée par la surjectivité de l'application divergence. La condition de surjectivité de B peut être exprimée sous la forme d'une condition discrète de Babuska-Brezzi. Nous obtenons alors le théorème 3.3 pour l'existence et l'unicité de la solution du problème de Stokes discret.

Théorème 3.3 (Babuska-Brezzi discret). *Le problème de Stokes discret (3.22) admet une unique solution $(\mathbf{v}_h, p_h) \in V_h \times Q_h$ s'il existe une constante $\beta_h > 0$ telle que la forme bilinéaire b vérifie*

$$\inf_{\substack{q_h \in Q_h \\ q_h \neq 0}} \sup_{\substack{\mathbf{w}_h \in V_h \\ \mathbf{w}_h \neq \mathbf{0}}} \frac{b(\mathbf{w}_h, q_h)}{\|q_h\|_{Q_h} \|\mathbf{w}_h\|_{V_h}} \geqslant \beta_h. \tag{3.27}$$

Nous proposons à présent une explication intuitive du lien entre la condition inf-sup et la condition d'injectivité de l'opérateur B^T (une explication plus détaillée pourra être trouvée dans Boffi *et al.* [7]). Nous introduisons la quantité suivante :

$$\sup_{\substack{\mathbf{W} \in \mathbb{R}^N \\ \mathbf{W} \neq \mathbf{0}}} \frac{\mathbf{W}^T B^T \mathbf{Q}}{\|\mathbf{W}\|_{\mathbb{R}^N}}, \tag{3.28}$$

avec $\mathbf{Q} \in \mathbb{R}^M$. L'expression (3.28) est toujours positive ou nulle. Elle sera nulle si $B^T\mathbf{Q} = \mathbf{0}$, c'est-à-dire si $\mathbf{Q} \in \ker B^T$. Or si B^T est un opérateur injectif, $\ker B^T$ ne contient que l'unique vecteur nul. Dans ce cas, il est possible d'introduire une constante $\beta > 0$ telle que

$$\sup_{\substack{\mathbf{W} \in \mathbb{R}^N \\ \mathbf{W} \neq \mathbf{0}}} \frac{\mathbf{W}^T B^T \mathbf{Q}}{\|\mathbf{W}\|_{\mathbb{R}^N}} \geqslant \beta \|\mathbf{Q}\|_{\mathbb{R}^M}, \quad \forall \mathbf{Q} \in \mathbb{R}^M \setminus \{\mathbf{0}\}. \tag{3.29}$$

Cette expression peut encore se réécrire

$$\sup_{\substack{\mathbf{W} \in \mathbb{R}^N \\ \mathbf{W} \neq \mathbf{0}}} \frac{\mathbf{W}^T B^T \mathbf{Q}}{\|\mathbf{Q}\|_{\mathbb{R}^M} \|\mathbf{W}\|_{\mathbb{R}^N}} \geqslant \beta, \quad \forall \mathbf{Q} \in \mathbb{R}^M \setminus \{\mathbf{0}\}. \tag{3.30}$$

Si B^T est injectif, alors la condition (3.30) doit être satisfaite. En particulier, elle doit être vérifiée pour la valeur minimale que peut prendre le membre de gauche de l'expression (3.30). Dans ce cas, nous pouvons reformuler cette expression sous la forme de la condition inf-sup suivante :

$$\inf_{\substack{\mathbf{Q} \in \mathbb{R}^M \\ \mathbf{Q} \neq \mathbf{0}}} \sup_{\substack{\mathbf{W} \in \mathbb{R}^N \\ \mathbf{W} \neq \mathbf{0}}} \frac{\mathbf{W}^T B^T \mathbf{Q}}{\|\mathbf{Q}\|_{\mathbb{R}^M} \|\mathbf{W}\|_{\mathbb{R}^N}} \geqslant \beta. \tag{3.31}$$

3.2.7. Propriétés de la solution numérique

Si les conditions d'existence et d'unicité de la solution sont satisfaites, alors il est possible d'établir certaines propriétés de la solution discrète. En utilisant, la propriété de coercivité de la forme a et l'inégalité de Cauchy-Schwarz, nous pouvons écrire pour $\mathbf{V}$ [3] :

$$\alpha_h \|\mathbf{V}\|^2 \leqslant \mathbf{V}^T A \mathbf{V} = \mathbf{V}^T \mathbf{F} \leqslant \|\mathbf{V}\| \, \|\mathbf{F}\| \Rightarrow \|\mathbf{V}\| \leqslant \frac{1}{\alpha_h} \|\mathbf{F}\|. \tag{3.32}$$

Nous constatons que l'expression (3.32) constitue en fait la forme discrète du corollaire 3.1.

3. Afin d'alléger les notations de cette sous-section-ci, le type de normes que nous employons ne sera pas spécifié. Cette liberté d'écriture se justifie par le fait qu'en dimension finie toutes les normes sont équivalentes. Les relations que nous présentons sont ainsi valides pour tout type de normes. Seule la valeur des constantes qui apparaissent dans les expressions est dépendante de la norme utilisée.

Nous pouvons également introduire une norme matricielle induite par la norme vectorielle. Celle-ci est définie par

$$\|A\| = \sup_{\mathbf{W}\neq\mathbf{0}} \frac{\|A\mathbf{W}\|}{\|\mathbf{W}\|}. \tag{3.33}$$

Nous pouvons supposer l'existence d'une constante M_a telle que $\|A\| \leqslant M_a$. Dans ce cas, nous avons la propriété suivante quel que soit le vecteur $\mathbf{W}$ considéré :

$$\|A\mathbf{W}\| \leqslant M_a\|\mathbf{W}\|. \tag{3.34}$$

En utilisant la relation (3.32), l'inégalité (3.34) peut se réécrire sous la forme

$$\|A\mathbf{V}\| \leqslant \frac{M_a}{\alpha_h}\|\mathbf{F}\|, \tag{3.35}$$

où $\alpha_h \leqslant M_a$, car la norme matricielle induite satisfait $\mathbf{V}^T A\mathbf{V} \leqslant M_a\|\mathbf{V}\|^2$.

Pour obtenir un résultat sur la valeur de $\|\mathbf{P}\|$, nous considérons la formulation suivante de la condition inf-sup [7] :

$$\beta_h\|\mathbf{P}\| \leqslant \|B^T\mathbf{P}\| = \|\mathbf{F} - A\mathbf{V}\|. \tag{3.36}$$

Cette dernière expression peut encore s'écrire

$$\|\mathbf{P}\| \leqslant \frac{1}{\beta_h}\|\mathbf{F} - A\mathbf{V}\| \leqslant \frac{1}{\beta_h}\left(1 + \frac{M_a}{\alpha_h}\right)\|\mathbf{F}\| \leqslant \frac{2M_a}{\alpha_h\beta_h}\|\mathbf{F}\|. \tag{3.37}$$

Les deux relations (3.32) et (3.37) deviennent finalement

$$\|\mathbf{V}\| + \|\mathbf{P}\| \leqslant c_h\|\mathbf{F}\|. \tag{3.38}$$

La relation précédente dépend de la discrétisation du domaine Ω via la constante c_h dont la valeur est liée au paramètre de discrétisation h. Le cas particulier où la constante c_h est indépendante de h est en pratique le plus intéressant. Cela impose notamment que la constante β_h qui intervient dans la condition inf-sup (3.27) soit indépendante de h. Si c_h est indépendant de h, il est possible de construire des matrices A et B de dimension croissante tout en conservant les mêmes constantes de majoration pour la solution numérique, ce qui constitue un avantage pour la méthode numérique utilisée.

3.2.8. Méthodes numériques pour le problème de Stokes

Pour terminer cette section consacrée à l'étude du problème de Stokes linéaire, nous allons regarder comment les résultats tirés des sous-sections précédentes peuvent être utilisés pour élaborer une méthode de résolution par éléments finis qui soit compatible avec les conditions d'existence et d'unicité de la solution discrète.

3.2.8.1. Paire d'éléments finis compatibles

Pour résoudre le problème de Stokes, une première approche est de déterminer simultanément les champs de vitesse et de pression, ce qui mène à la méthode des éléments finis mixtes. Les fonctions de base sont choisies dans les espaces vectoriels V_h et Q_h. Le choix de ces espaces est déterminé de manière à satisfaire la condition inf-sup discrète. Des éléments finis qui satisfont

cette condition seront qualifiés de compatibles [4]. Pour satisfaire la condition inf-sup, l'idée est de travailler avec un espace V_h suffisamment « riche » par rapport à Q_h. Plus l'espace V_h contient de fonctions de base par rapport à Q_h et plus il sera facile de satisfaire la condition inf-sup. Cela revient à imposer que la matrice B^T ait plus de lignes que de colonnes. Intuitivement, nous pouvons comprendre que plus cette matrice aura de lignes comparativement à son nombre de colonnes et plus la possibilité que les colonnes soient des combinaisons linéaires l'une de l'autre sera faible. Dans ce cas, la matrice B^T sera de rang maximal et son noyau sera nul.

Pour enrichir V_h par rapport à Q_h, deux approches sont envisageables. Dans la première approche, les fonctions de base de l'espace V_h sont des polynômes de plus haut degré que les fonctions de base de Q_h. Dans la seconde approche, le maillage est plus raffiné pour le champ de vitesse que pour le champ de pression.

Nous allons citer à titre d'exemples quelques paires d'éléments finis compatibles à deux dimensions. Nous introduisons d'abord l'espace $\mathbb{P}_k$ des polynômes de degré inférieur ou égal à k à savoir

$$\mathbb{P}_k = \{p(x,y) = \sum_{0 \leqslant i+j \leqslant k} a_{ij} x^i y^j,\ a_{ij} \in \mathbb{R}\}. \tag{3.39}$$

Nous considérons également l'espace $\mathbb{Q}_k$ des polynômes de degré inférieur ou égal à k par rapport à chacune des variables. Mathématiquement, nous avons

$$\mathbb{Q}_k = \{q(x,y) = \sum_{0 \leqslant i,j \leqslant k} b_{ij} x^i y^j,\ b_{ij} \in \mathbb{R}\}. \tag{3.40}$$

À titre illustratif, les paires suivantes d'éléments finis sont compatibles :

- Éléments $\mathbb{P}_k/\mathbb{P}_{k-1}$ et $\mathbb{Q}_k/\mathbb{Q}_{k-1}$. Ces éléments sont aussi connus sous le nom d'éléments finis de Taylor-Hood généralisés. Ceux-ci sont compatibles si $k \geqslant 2$. Les éléments de Taylor-Hood présentent de bonnes propriétés de convergence. Les paires $\mathbb{P}_2/\mathbb{P}_1$ et $\mathbb{Q}_2/\mathbb{Q}_1$ qui sont souvent utilisées ont un taux de convergence quadratique.

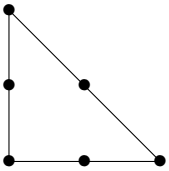
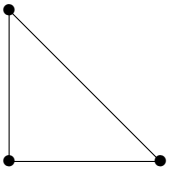
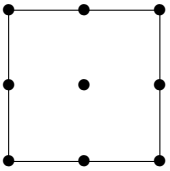
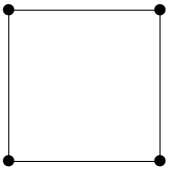
Vitesse	Pression	Vitesse	Pression
$\mathbb{P}_2$	$\mathbb{P}_1$	$\mathbb{Q}_2$	$\mathbb{Q}_1$
			

FIGURE 3.2.: Représentation des éléments finis de Taylor-Hood $\mathbb{P}_2/\mathbb{P}_1$ et $\mathbb{Q}_2/\mathbb{Q}_1$.

4. Même si en théorie le choix des paires d'éléments finis est dicté par la condition inf-sup, l'utilisation pratique d'éléments finis non compatibles peut néanmoins donner des résultats acceptables pour le champ de vitesse (voir Fortin et Brezzi [9] pour une discussion sur l'importance de la condition inf-sup).

- Éléments $\mathbb{P}_1$-iso-$\mathbb{P}_2/\mathbb{P}_1$. Ces éléments constituent une alternative à la paire $\mathbb{P}_2/\mathbb{P}_1$. Dans ce cas, l'approximation $\mathbb{P}_2$ du champ de vitesse est remplacée par une approximation $\mathbb{P}_1$ sur un maillage plus fin. En deux dimensions et pour un maillage $\mathcal{T}_h$ obtenu par triangulation, la vitesse sera discrétisée sur un maillage $\mathcal{T}_{h/2}$ obtenu en divisant chacun des triangles en quatre. Pour un maillage quadrangulaire, les éléments $\mathbb{Q}_1$-iso-$\mathbb{Q}_2/\mathbb{Q}_1$ peuvent être utilisés.

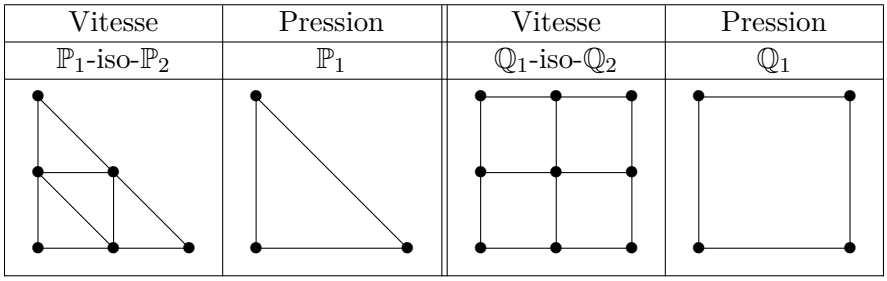
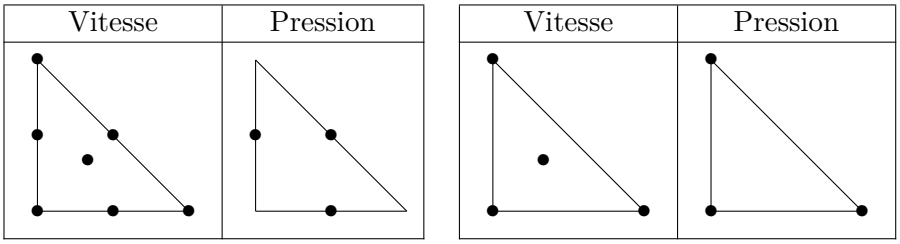


FIGURE 3.3.: Représentation des éléments finis $\mathbb{P}_1$-iso-$\mathbb{P}_2$/ $\mathbb{P}_1$ et $\mathbb{Q}_1$-iso-$\mathbb{Q}_2/\mathbb{Q}_1$.

- Éléments $\mathbb{P}_2/\mathbb{P}_0$. Pour assurer la stabilité de cette paire, une approximation quadratique du champ de vitesse est requise. Cette paire utilise un champ de pression constant par morceaux et donc potentiellement discontinu d'un élément à l'autre. Cette discontinuité ne constitue pas un problème, car le gradient de la pression n'intervient pas dans la formulation variationnelle.

- Utilisation des fonctions bulles. Une autre possibilité pour satisfaire la condition inf-sup consiste à ajouter des fonctions bulles aux éléments finis classiques. Une fonction bulle est une fonction définie à l'intérieur d'un élément et qui s'annule sur sa frontière. L'utilisation de fonctions bulles permet en général d'améliorer la résolution locale de l'approximation par éléments finis et offre de ce fait un moyen de stabilisation pour de nombreux problèmes. Comme exemples d'utilisation des fonctions bulles, nous citons l'élément fini mixte de Crouzeix-Raviart et l'élément Mini. Pour l'élément de Crouzeix-Raviart, le champ de vitesse est discrétisé par des fonctions quadratiques continues d'un élément à l'autre auxquelles est ajoutée une fonction bulle cubique. Le champ de pression est approximé par des fonctions linéaires discontinues d'un élément à l'autre (les degrés de liberté sont par exemple les barycentres des côtés de l'élément). Dans le cas de l'élément Mini, la vitesse est approximée par des fonctions linéaires par morceaux auxquelles une fonction bulle cubique est ajoutée et la pression est approchée par des fonctions linéaires sur chaque élément et continues d'un élément à l'autre.



(a) Élément de Crouzeix-Raviart. (b) Élément Mini.

FIGURE 3.4.: Éléments finis triangulaires avec fonctions bulles.

Pour les éléments quadrangulaires, quelques équivalents de l'élément Mini ont été introduits par Bai [4]. Nous en présentons une paire particulière que nous utiliserons dans nos simulations numériques du chapitre 7. Cette paire d'éléments finis est la paire $\mathbb{Q}_1/\mathbb{Q}_1$ à laquelle sont rajoutés deux degrés de liberté internes pour chacune des composantes de la vitesse. Dans les coordonnées (ξ, η) du quadrilatère de référence, les fonctions de base associées à ces deux nouveaux degrés de liberté sont données par

$$\varphi_1(\xi,\eta) = (1-\xi^2)\,(1-\eta^2), \tag{3.41}$$

$$\varphi_2(\xi,\eta) = (\xi+\eta)\,(1-\xi^2)\,(1-\eta^2). \tag{3.42}$$

Comme nous le constatons, il existe un ensemble assez important de paires d'éléments finis stables. En pratique, le choix d'une paire particulière d'éléments sera dicté par divers arguments pratiques. Le degré des fonctions de base dépendra à la fois de la précision souhaitée et de la complexité des éléments. Des approximations constantes ou linéaires par morceaux sont plus facilement implémentables que des approximations quadratiques ou cubiques. Comme le problème de Stokes présente un couplage entre les champs de vitesse et de pression, l'erreur d'approximation dépendra de la manière de discrétiser ces deux champs. Afin que ces derniers contribuent de manière identique à l'erreur d'approximation, il convient de travailler avec des degrés d'approximation semblables pour la vitesse et la pression. Des paires telles que les paires $\mathbb{P}_k/\mathbb{P}_{k-1}$ et $\mathbb{Q}_k/\mathbb{Q}_{k-1}$ sont relativement efficaces.

3.2.8.2. Choix d'éléments finis à divergence nulle

Nous avons vu à la sous-section 3.2.3 que le problème de Stokes peut s'exprimer uniquement en fonction du champ de vitesse pour autant que la solution à ce problème soit recherchée parmi les fonctions à divergence nulle. Il est dès lors possible de déterminer dans un premier temps la vitesse puis la pression dans le fluide. Pour la détermination du champ de vitesse, l'existence et l'unicité de la solution sont garanties et le problème discret se ramène à une forme matricielle impliquant uniquement une matrice symétrique définie positive. Bien que résoudre le problème de Stokes en traitant séparément les deux champs inconnus puisse sembler une approche intéressante, celle-ci est généralement peu utilisée. En effet, une telle approche nécessite de travailler avec des fonctions de base qui satisfont la contrainte d'incompressibilité a priori, ce qui rend en pratique l'implémentation de la méthode plus difficile.

Deux approches principales peuvent être envisagées pour s'assurer de travailler avec des fonctions à divergence nulle. La première approche est une approche interne, car elle a pour objectif d'utiliser des fonctions qui sont localement indivergentielles. La deuxième approche est une approche externe qui vise à travailler avec des fonctions de base qui sont à divergence nulle en moyenne sur les éléments. Nous allons présenter brièvement ces deux approches en nous focalisant sur le cas bidimensionnel. Une étude plus approfondie de l'approximation des fonctions à divergence nulle peut être trouvée dans Fortin [21].

Dans l'approche interne, nous souhaitons approcher l'espace V_{div} par des sous-espaces $V_{\text{div},h}$. Pour cette méthode, deux problèmes potentiels sont à considérer. Le premier problème est lié à la convergence de l'approximation, c'est-à-dire à la possibilité de se rapprocher arbitrairement près de toute fonction de V_{div} si le facteur de discrétisation h est choisi suffisamment petit. Le second problème provient de la construction effective d'une base de fonctions de $V_{\text{div},h}$. Pour comprendre cette dernière difficulté, nous considérons un domaine discrétisé par triangulation.

Les composantes w_x et w_y du champ de vitesse $\mathbf{w}$ sont approchées sur chaque triangle par des polynômes de degré k à savoir

$$w_x(x,y) = \sum_{0 \leqslant i+j \leqslant k} a_{ij} x^i y^j, \tag{3.43a}$$

$$w_y(x,y) = \sum_{0 \leqslant i+j \leqslant k} b_{ij} x^i y^j. \tag{3.43b}$$

La divergence du champ de vitesse est un polynôme de degré $k-1$ constitué de $k(k+1)/2$ monômes. Comme les monômes sont linéairement indépendants, la contrainte de divergence nulle de la vitesse conduit à $k(k+1)/2$ contraintes sur la valeur des coefficients a_{ij} et b_{ij} et donc sur les valeurs nodales de la solution. Pour mieux comprendre cela, nous envisageons une approximation de la vitesse par des polynômes quadratiques par morceaux ($k = 2$), c'est-à-dire

$$w_x(x,y) = a_{00} + a_{10}\, x + a_{01}\, y + a_{20}\, x^2 + a_{11}\, xy + a_{02}\, y^2, \tag{3.44a}$$

$$w_y(x,y) = b_{00} + b_{10}\, x + b_{01}\, y + b_{20}\, x^2 + b_{11}\, xy + b_{02}\, y^2. \tag{3.44b}$$

La condition de divergence nulle s'écrit

$$a_{10} + b_{10} + (2a_{20} + b_{11})\, x + (a_{11} + 2b_{02})\, y = 0. \tag{3.45}$$

Cette condition peut aussi s'écrire sous la forme du système suivant de $k(k+1)/2 = 3$ relations :

$$\begin{cases} a_{10} + b_{10} = 0, \\ 2a_{20} + b_{11} = 0, \\ a_{11} + 2b_{02} = 0. \end{cases} \tag{3.46}$$

Si la condition de divergence nulle est exprimée pour chacun des éléments du domaine, nous obtenons un système d'équations homogènes pour les valeurs nodales. Il sera possible de définir des fonctions de base localement indivergentielles si ce système possède des solutions non nulles. Cela sera notamment déterminé par le degré des fonctions de base ainsi que le nombre d'éléments dans la triangulation. Si nous employons des polynômes de degré un, la contrainte d'incompressibilité se limite à une seule équation par élément. Si nous considérons un domaine à la frontière duquel sont imposées des conditions de Dirichlet, nous obtiendrons un nombre d'éléments triangulaires plus important que le nombre de nœuds internes. Dans ce cas, le système d'équations pour les valeurs nodales contient plus d'équations que d'inconnues. Dans la plupart des cas, ce système ne possèdera que la solution nulle. Pour avoir une solution non nulle, il faut augmenter le nombre d'inconnues par rapport au nombre d'équations, ce qui revient à augmenter le degré des polynômes d'approximation. Le traitement mathématique peut alors devenir relativement compliqué, ce qui rend la méthode d'une utilisation peu aisée.

Dans l'approche externe, l'idée est de remplacer la condition de divergence nulle par une condition discrète. Les fonctions d'approximation $\mathbf{w}$ sont choisies de manière à satisfaire la condition de divergence nulle sur tout le domaine à savoir

$$\int_\Omega \operatorname{div} \mathbf{w}\, d\Omega = 0. \tag{3.47}$$

Cette relation exprime mathématiquement la conservation de la masse dans le domaine Ω (flux net nul au travers des frontières du domaine). La condition globale (3.47) de divergence

nulle se ramène finalement à une équation algébrique reliant entre elles les valeurs nodales de la solution du problème discret.

Au final, la détermination du champ de vitesse à partir de fonctions test à divergence nulle constitue une approche envisageable pour la résolution du problème de Stokes. Plus généralement, cette approche peut être utilisée pour l'étude des fluides incompressibles ou en électromagnétisme où le champ d'induction magnétique est indivergentiel. Cependant en mécanique des fluides, une résolution par éléments finis mixtes est en général préférée vu la difficulté d'approcher l'espace des fonctions à divergence nulle.

3.2.8.3. Utilisation de méthodes de stabilisation

Nous avons vu précédemment que la condition inf-sup limite le choix des paires d'éléments finis qui peuvent être utilisées pour résoudre numériquement le problème de Stokes. Nous pouvons remarquer que la condition inf-sup ne permet pas d'utiliser des polynômes de même degré pour les champs de vitesse et de pression pour un même nombre d'éléments finis. De plus, l'utilisation de certaines paires telles que la paire $\mathbb{P}_0/\mathbb{P}_1$ de faible degré n'est également pas possible. L'utilisation de certaines paires qui ne respectent pas la condition inf-sup pourrait cependant présenter un intérêt notamment pour diminuer la taille du système d'équations algébriques à résoudre. C'est la raison pour laquelle des méthodes de stabilisation ont été développées afin de contourner la condition inf-sup. Nous nous proposons d'en illustrer un exemple particulier. D'autres exemples de méthodes de stabilisation seront présentés dans le chapitre 4 pour le problème thermique dont la méthode de stabilisation *Galerkin Least-Squares* qui s'applique également au problème de Stokes [18, 36].

La méthode de stabilisation que nous présentons a été proposée par Bochev *et al.* [6] et est applicable à la paire $\mathbb{P}_0/\mathbb{P}_1$. Nous pouvons d'abord montrer sur la base d'un exemple que cette paire ne satisfait pas la condition inf-sup [31]. Pour cela, nous considérons le maillage d'un domaine carré comme représenté sur la figure 3.5. Nous supposons des conditions aux limites de Dirichlet homogènes sur toute la frontière du domaine. Le domaine est caractérisé par $(n+1)^2$ nœuds et $2n^2$ éléments. Si la vitesse est discrétisée en utilisant des éléments $\mathbb{P}_1$, alors il y a $2(n-1)^2$ inconnues associées aux deux composantes du vecteur vitesse. La pression est approximée par des éléments $\mathbb{P}_0$. Le nombre d'inconnues pour la pression vaut le nombre d'éléments moins un pour tenir compte du fait que la valeur moyenne de la pression est par exemple imposée sur le domaine. Il y correspond donc $2n^2 - 1$ inconnues pour la pression. Nous constatons que le nombre d'inconnues pour la vitesse est inférieur au nombre d'inconnues pour la pression, ce qui implique que $\ker B^T \neq \{\mathbf{0}\}$ et donc que la condition inf-sup n'est pas satisfaite.

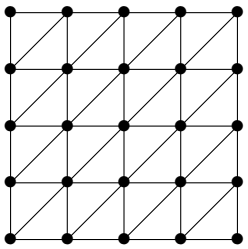


FIGURE 3.5.: Triangulation d'un domaine carré.

Nous présentons à présent la méthode de stabilisation de Bochev *et al.* pour rendre la paire $\mathbb{P}_1/\mathbb{P}_0$ stable. Pour cela, nous considérons un domaine $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ de frontière Γ sur laquelle sont imposées des conditions aux limites de Dirichlet homogènes [5]. Nous introduisons une partition $\mathcal{T}_h$ de Ω en éléments finis Ω_e. La frontière de chacun des éléments finis est constituée de faces en dimension trois ou de bords en dimension deux que nous notons de manière générique γ_f. Nous notons par Γ_h l'ensemble des faces des éléments finis qui sont à l'intérieur de Ω.

Nous considérons à présent les fonctions $\mathbf{w}_h$ et q_h qui appartiennent respectivement aux sous-espaces d'approximation V_h et Q_h. Les espaces V_h et Q_h sont choisis de manière à ce que la vitesse et la pression soient approchées respectivement par des fonctions linéaires et constantes sur chaque élément. Comme nous l'avons vu, la paire $\mathbb{P}_1/\mathbb{P}_0$ ne satisfait pas la condition inf-sup. Il est néanmoins possible de montrer que pour cette paire, il existe deux constantes strictement positives C_1 et C_2 telles que

$$\sup_{\substack{\mathbf{w}_h \in V_h \\ \mathbf{w}_h \neq \mathbf{0}}} \frac{\int_\Omega q_h \operatorname{div} \mathbf{w}_h \, d\Omega}{\|\mathbf{w}_h\|_{\mathbb{H}^1}} \geqslant C_1 \|q_h\|_{\mathbb{L}^2} - C_2 \, h^{1/2} \, \|[q_h]\|_{\Gamma_h}, \quad \forall q_h \in Q_h, \tag{3.48}$$

où

$$\|[q_h]\|_{\Gamma_h} = \left(\sum_{\gamma_f \in \Gamma_h} \int_{\gamma_f} [q_h]^2 \, d\Gamma_h \right). \tag{3.49}$$

La notation $[q_h]$ représente le saut de pression d'un élément à l'autre. La quantité $\|[q_h]\|_{\Gamma_h}$ est donc une mesure de la discontinuité du champ de pression dans tout l'intérieur de Ω. Dans la relation (3.48), le terme $-C_2 \, h^{1/2} \, \|[q_h]\|_{\Gamma_h}$ mesure à quel point la paire d'éléments finis ne satisfait pas la condition inf-sup. En effet, si ce terme est nul, la relation (3.48) se ramène à la condition inf-sup. Ainsi, le terme $C_2 \, h^{1/2} \, \|[q_h]\|_{\Gamma_h}$ joue un rôle déstabilisateur qui est d'autant plus important que la valeur de ce terme est importante. Afin de compenser l'effet déstabilisateur de ce terme, certaines méthodes de stabilisation rajoutent des termes dans la formulation faible du problème de Stokes. Dans Bochev *et al.* [6], la stabilisation de la paire d'éléments est réalisée de la manière suivante. Nous introduisons un opérateur Π défini de $\mathbb{L}^2$ dans $\mathbb{P}_1$. En utilisant cet opérateur, nous pouvons borner le terme déstabilisateur de la façon suivante (avec $C > 0$ une constante indépendante de h) :

$$C h^{1/2} \|[q_h]\|_{\Gamma_h} \leqslant \|q_h - \Pi q_h\|_{\mathbb{L}^2}, \quad \forall q_h \in Q_h. \tag{3.50}$$

Afin de compenser le manque de stabilité de la paire $\mathbb{P}_0/\mathbb{P}_1$, il semble pertinent d'ajouter au lagrangien du problème de Stokes le terme

$$\frac{\alpha}{2} \|(I - \Pi) q_h\|_{\mathbb{L}^2}^2, \tag{3.51}$$

où I représente l'opérateur identité et α correspond juste à un paramètre destiné à rendre ce nouveau terme dimensionnellement correct par rapport aux autres termes du lagrangien. Le nouveau lagrangien s'écrit à présent

$$\tilde{\mathcal{L}}(\mathbf{w}_h, q_h) = \frac{1}{2} a(\mathbf{w}_h, \mathbf{w}_h) + b(\mathbf{w}_h, q_h) - (\mathbf{f}, \mathbf{w}_h) - \frac{\alpha}{2} \|(I - \Pi) q_h\|_{\mathbb{L}^2}^2 . \tag{3.52}$$

5. Le raisonnement est explicité pour le cas de conditions aux limites de Dirichlet homogènes, mais les preuves peuvent aussi être étendues au problème que nous étudions.

Le point de selle $(\mathbf{v}_h, p_h)$ du lagrangien est également solution du problème variationnel suivant :

Trouver $\mathbf{v}_h \in V_h$ et $p_h \in Q_h$ tels que

$$\begin{cases} a(\mathbf{v}_h, \mathbf{w}_h) + b(p_h, \mathbf{w}_h) = (\mathbf{f}, \mathbf{w}_h) & \forall \mathbf{w}_h \in V_h, \\ b(q_h, \mathbf{v}_h) - g(p_h, q_h) = 0 & \forall q_h \in Q_h, \end{cases} \tag{3.53}$$

où

$$g(p_h, q_h) = \int_\Omega \alpha\,(p_h - \Pi p_h)(q_h - \Pi q_h)\,d\Omega. \tag{3.54}$$

L'avantage de cette méthode de stabilisation est notamment qu'elle ne dépend pas d'un paramètre de stabilisation contrairement à d'autres méthodes numériques telles que les méthodes de pénalisation pour lesquelles le choix du paramètre de stabilisation est toujours une question délicate. Un autre avantage réside dans la grande flexibilité au niveau du choix du terme de stabilisation (3.51). En effet, le choix de l'opérateur Π reste relativement libre pour autant qu'il satisfasse quelques propriétés qui assurent la stabilité et la convergence de la méthode. En pratique, le choix sera dicté à la fois par la simplicité de l'opérateur et la possibilité de calculer son action au niveau de chaque élément. Des exemples pratiques de l'opérateur Π pourront être trouvés dans [6].

3.2.8.4. Illustration numérique

Pour terminer cette sous-section consacrée à la stabilisation du problème de Stokes, nous nous proposons de vérifier la stabilité de quelques paires d'éléments finis compatibles ainsi que de comparer leurs performances. Pour ce faire, nous nous inspirons d'un exemple tiré de Donea et Huerta [18]. Nous souhaitons résoudre le problème suivant sur le domaine $\Omega =]0,1[\times]0,1[$:

$$\begin{cases} -\Delta \mathbf{v} + \nabla p = \mathbf{f} & \text{dans } \Omega, \\ \operatorname{div} \mathbf{v} = 0 & \text{dans } \Omega, \\ \mathbf{v} = \mathbf{0} & \text{sur } \Gamma, \\ p(0,0) = 0, \end{cases} \tag{3.55}$$

avec

$$\begin{aligned} f_x =& (12 - 24y)\,x^4 + (-24 + 48y)\,x^3 + (-48y + 72y^2 - 48y^3 + 12)\,x^2 \\ &+ (-2 + 24y - 72y^2 + 48y^3)\,x + 1 - 4y + 12y^2 - 8y^3, \end{aligned} \tag{3.56a}$$

$$\begin{aligned} f_y =& (8 - 48y + 48y^2)\,x^3 + (-12 + 72y - 72y^2)\,x^2 \\ &+ (4 - 24y + 48y^2 - 48y^3 + 24y^4)\,x - 12y^2 + 24y^3 - 12y^4. \end{aligned} \tag{3.56b}$$

La solution de ce problème est donnée par

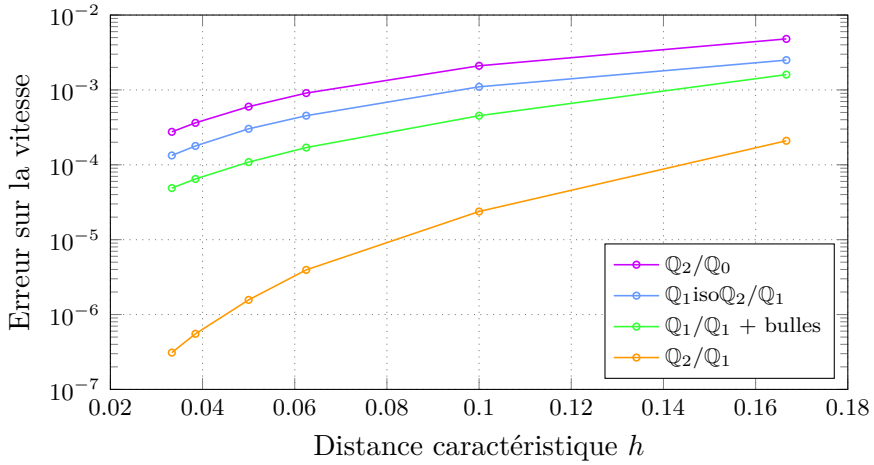
$$v_x = x^2\,(1-x)^2\,(2y - 6y^2 + 4y^3), \tag{3.57a}$$

$$v_y = -\,y^2\,(1-y)^2\,(2x - 6x^2 + 4x^3), \tag{3.57b}$$

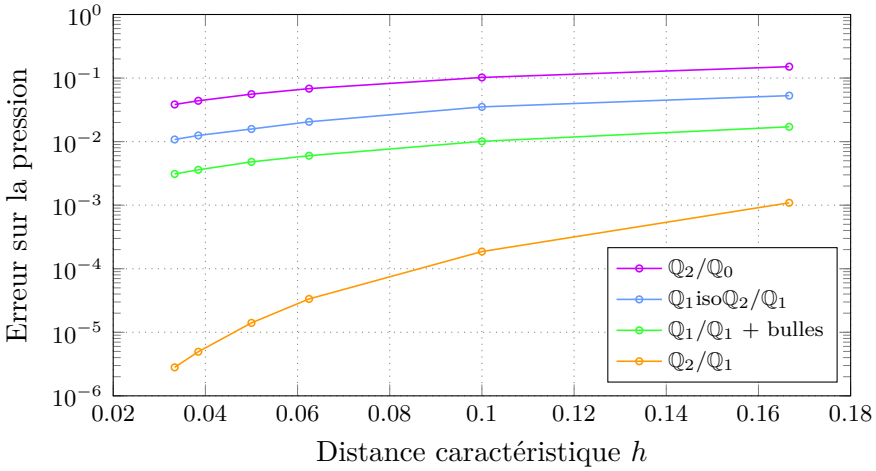
$$p = x\,(1-x). \tag{3.57c}$$

Nous avons résolu le problème (3.55) en utilisant quatre paires d'éléments finis compatibles à savoir les paires $\mathbb{Q}_2/\mathbb{Q}_0$,$\mathbb{Q}_1iso\mathbb{Q}_2/\mathbb{Q}_1$,$\mathbb{Q}_2/\mathbb{Q}_1$ et $\mathbb{Q}_1/\mathbb{Q}_1$ avec les fonctions bulles (3.41)–(3.42).

Pour chacune des paires utilisées, nous avons calculé l'erreur maximale entre les valeurs nodales de la solution numérique et la solution exacte évaluée aux nœuds de discrétisation. Les figures 3.6(a) et 3.6(b) représentent cette erreur respectivement pour la vitesse et la pression en fonction de la distance caractéristique h entre deux nœuds de discrétisation de la vitesse. Nous constatons pour chacune des paires utilisées que l'erreur devient plus petite quand le maillage est raffiné, ce qui tend à montrer que la solution numérique converge bien vers la solution exacte. Nous observons également que la paire $\mathbb{Q}_2/\mathbb{Q}_1$ présente une erreur systématiquement plus petite que les trois autres paires pour lesquelles l'erreur est du même ordre de grandeur. Ceci est en accord avec le fait que la paire $\mathbb{Q}_2/\mathbb{Q}_1$ a une vitesse de convergence en théorie plus élevée que les trois autres. Notons enfin que si nous testons la paire $\mathbb{Q}_1/\mathbb{Q}_1$, nous constatons que celle-ci donne une solution pour la pression qui diverge, ce qui est en accord avec le caractère non compatible de cette paire.



(a) Erreur sur la vitesse.



(b) Erreur sur la pression.

FIGURE 3.6.: Comparaison de quelques paires d'éléments finis compatibles.

3.3. Problème de Stokes non linéaire

L'utilisation concrète des équations de Stokes pour l'étude de l'écoulement des glaciers nécessite l'utilisation d'une loi rhéologique non linéaire. Ainsi, l'étude de la dynamique des glaciers demande d'analyser un problème de Stokes non linéaire. Toutefois, les résultats obtenus pour le problème de Stokes linéaire peuvent être étendus au cas non linéaire. C'est la raison pour laquelle nous nous sommes principalement attardé sur le problème linéaire pour lequel le traitement analytique est plus aisé. Dans cette section-ci, nous nous limiterons donc à la présentation des résultats principaux relatifs au problème non linéaire afin de souligner la correspondance avec le problème linéaire. Le lecteur intéressé pourra trouver plus de détails sur le problème de Stokes non linéaire dans [37, 38].

3.3.1. Formes forte et faible du problème de Stokes non linéaire

Nous considérons un domaine $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ dont la frontière Γ peut être décomposée en deux parties disjointes à savoir Γ_D et Γ_N sur lesquelles sont imposées respectivement des conditions de Dirichlet homogènes (glacier ancré dans la roche) et de Neumann homogènes (surface libre du glacier) [6].

Mathématiquement, la formulation forte du problème de Stokes s'écrit

$$\begin{cases} -\operatorname{div}(2\mu\dot{\varepsilon}(\mathbf{v})) + \nabla p = \mathbf{f} & \text{dans } \Omega, \\ \operatorname{div}\mathbf{v} = 0 & \text{dans } \Omega, \\ \sigma \cdot \mathbf{n} = \mathbf{0} & \text{sur } \Gamma_N, \\ \mathbf{v} = \mathbf{0} & \text{sur } \Gamma_D. \end{cases} \tag{3.58}$$

Dans le système (3.58), nous supposons que la glace obéit à la loi de Glen régularisée qui pour rappel peut s'écrire

$$\frac{1}{2\mu} = A\left(\sigma_0^{n-1} + (2\mu d_e)^{n-1}\right). \tag{3.59}$$

Nous pouvons à présent établir une formulation variationnelle du problème de Stokes. Pour cela, nous introduisons les quantités $r = 1+1/n$ et $r' = 1+n$. La formulation variationnelle sera correctement définie (c'est-à-dire que chacune des intégrales de la formulation faible existe) si la fonction $p \in \mathbb{L}^{r'}(\Omega)$ et si $\mathbf{v}$ appartient à l'ensemble V défini par

$$V = \{\mathbf{w} \in \left[\mathbb{W}^{1,r}(\Omega)\right]^d, \mathbf{w} = \mathbf{0} \text{ sur } \Gamma_D\}. \tag{3.60}$$

La formulation faible s'obtient en multipliant les équations de Stokes par un couple de fonctions test $(\mathbf{w}, q) \in V \times Q$ et en les intégrant sur le domaine Ω. Pour simplifier les notations, nous introduisons les formes suivantes :

$$\begin{cases} c(\mathbf{v}, \mathbf{w}) = \displaystyle\int_\Omega 2\mu(d_e)\,\dot{\varepsilon}(\mathbf{v}) : \dot{\varepsilon}(\mathbf{w})\,d\Omega, \\ b(\mathbf{v}, q) = -\displaystyle\int_\Omega q \operatorname{div}\mathbf{v}\,d\Omega. \end{cases} \tag{3.61}$$

La forme $c(\,\cdot\,,\,\cdot\,)$ pour le problème de Stokes non linéaire a une expression similaire à la forme $a(\,\cdot\,,\,\cdot\,)$ pour le problème linéaire. Toutefois, les propriétés mathématiques de ces deux

6. Les résultats que nous présentons sont également d'application si une loi de glissement est ajoutée sur une partie de la frontière (voir Jouvet [37] à ce sujet).

formes ne sont pas identiques. En particulier, la forme $c(\,\cdot\,,\,\cdot\,)$ n'est pas bilinéaire et symétrique alors que la forme $a(\,\cdot\,,\,\cdot\,)$ l'est bien.

En passant les détails de calcul qui sont similaires à ceux de la section 3.2.2, nous obtenons la formulation variationnelle 3.3 pour le problème de Stokes non linéaire.

Formulation variationnelle 3.3. Trouver $\mathbf{v} \in V = \{\mathbf{w} \in \left[\mathbb{W}^{1,r}(\Omega)\right]^d, \mathbf{w} = \mathbf{0} \text{ sur } \Gamma_D\}$ et $p \in Q = \mathbb{L}^{r'}(\Omega)$ tels que

$$\begin{cases} c(\mathbf{v},\mathbf{w}) + b(\mathbf{w},p) = (\mathbf{f},\mathbf{w}), \\ b(\mathbf{v},q) = 0, \end{cases}$$

pour tout couple $(\mathbf{w},q) \in V \times Q$.

Cette formulation variationnelle est semblable à la formulation variationnelle 3.1 du problème linéaire. De nouveau, il faut rester conscient qu'elle n'est pas parfaitement identique dû notamment au fait que la forme c n'est pas bilinéaire et symétrique.

3.3.2. Existence et unicité de la solution au problème de Stokes non linéaire

Comme dans le cas linéaire, la question de l'existence et de l'unicité d'une solution au problème variationnel constitue une question primordiale. Pour traiter ce problème, il est de nouveau possible d'exprimer le problème de Stokes sous la forme d'un problème d'optimisation. Pour déterminer le champ de vitesse, nous introduisons la quantité suivante :

$$M(x) = \int_0^x s\,\mu(s)\,ds\,. \tag{3.62}$$

Le champ de vitesse solution du problème de Stokes doit satisfaire le problème de minimisation sous contrainte suivant :

La fonction $\mathbf{v} \in V$ satisfait la formulation variationnelle 3.3 si $\mathbf{v}$ est une solution du problème d'optimisation suivant :

$$\min_{\mathbf{w}\in V} J(\mathbf{w}) = \min_{\mathbf{w}\in V} \int_\Omega 2\,M(\sqrt{2}d_e)\,d\Omega - \int_\Omega \mathbf{f}\cdot\mathbf{w}\,d\Omega$$

sous la contrainte que

$$\nabla\cdot\mathbf{w} = 0.$$

Il peut être montré que ce problème de minimisation possède une unique solution [37]. Ce résultat est notamment une conséquence de la stricte convexité de la fonctionnelle $J(\mathbf{w})$. Ainsi, l'existence et l'unicité du champ de vitesse pour le problème de Stokes non linéaire sont assurées sans aucune condition. Comme dans le cas linéaire, l'existence et l'unicité de la solution à la formulation variationnelle du problème de Stokes sont déterminées par le champ de pression et en particulier par le problème de la surjectivité de l'opérateur divergence. Il est donc normal de retrouver une condition inf-sup à savoir

$$\inf_{\substack{q\in Q\\ q\neq 0}} \sup_{\substack{\mathbf{w}\in V\\ \mathbf{w}\neq\mathbf{0}}} \frac{b(\mathbf{w},q)}{\|q\|_Q\|\mathbf{w}\|_V} \geqslant \beta, \quad \beta > 0. \tag{3.63}$$

Tout naturellement, cette condition inf-sup s'applique aussi au cas de l'approximation discrète du problème de Stokes. Ainsi, si les champs de vitesse et de pression sont approximés

respectivement par des fonctions $\mathbf{w}_h \in V_h \subset V$ et $q_h \in Q_h \subset Q$, ces fonctions devront satisfaire une condition inf-sup discrète à savoir

$$\inf_{\substack{q_h \in Q_h \\ q_h \neq 0}} \sup_{\substack{\mathbf{w}_h \in V_h \\ \mathbf{w}_h \neq \mathbf{0}}} \frac{b(\mathbf{w}_h, q_h)}{\|q_h\|_{Q_h} \|\mathbf{w}_h\|_{V_h}} \geqslant \beta_h, \quad \beta_h > 0. \tag{3.64}$$

La plupart des résultats théoriques obtenus dans le cas du problème de Stokes linéaire sont ainsi applicables au problème non linéaire. Toutefois, le problème non linéaire représente un défi plus important notamment pour la détermination des valeurs nodales des inconnues. En effet, la discrétisation du problème non linéaire mène à l'obtention d'un système d'équations non linéaires pour les inconnues. Cela nécessite donc la mise en place de méthodes de résolution adaptées. La problématique de la résolution numérique d'équations algébriques non linéaires sera envisagée dans la suite de ce travail et plus particulièrement au chapitre 5.

4. Étude du problème thermique

Ce quatrième chapitre est consacré à l'étude de l'équation stationnaire de la chaleur qui permet de déterminer le profil de température dans le glacier en régime stationnaire. Dans ce chapitre, l'accent sera placé principalement sur les questions d'ordre numérique même si nous mettrons en exergue quelques résultats mathématiques importants. Dans cette partie, nous envisagerons comme unique inconnue le champ de température dans le glacier et nous supposerons la vitesse du fluide connue. La section 4.1 commencera par un bref rappel sur l'équation de la chaleur. Dans la section 4.2, nous nous intéresserons à quelques propriétés du problème thermique linéaire. À la section 4.3, nous étudierons de manière générale l'équation d'advection-diffusion sous l'angle numérique. Il sera montré que la présence d'un terme de convection dans cette équation peut mener à une instabilité [1] dans la méthode classique de Galerkin. Différentes méthodes de stabilisation (SUPG et GLS) seront alors présentées. À la section 4.4, nous montrerons à partir d'une analyse dimensionnelle la pertinence d'appliquer des méthodes numériques stabilisées pour l'équation de la chaleur dans les glaciers. Enfin, ce chapitre se terminera par la section 4.5 dans laquelle nous analyserons comment la contrainte thermique $T \leqslant T_m$ peut être assurée numériquement.

4.1. Rappel du problème thermique

Au chapitre 2, nous avons vu que l'équation de la chaleur en régime stationnaire pouvait s'écrire de la manière suivante :

$$\underbrace{-\operatorname{div}\left(k(T)\nabla T\right)}_{\text{Diffusion de la chaleur}} + \underbrace{\rho\, c(T)\, \mathbf{v}\cdot\nabla T}_{\text{Transport de chaleur par convection}} = \underbrace{4\,\mu(T)\, d_e^2}_{\text{Source de chaleur}} \quad . \tag{4.1}$$

Cette équation est en fait une équation d'advection-diffusion avec un terme source qui dépend de la température. Les termes qui apparaissent dans l'équation (4.1) sont ainsi de nature différente et se caractérisent donc par des propriétés différentes tant du point de vue mathématique que physique. Cette équation peut ainsi être considérée comme un problème multiphysique dans la mesure où elle couple des termes de nature physique différente. Comme nous le verrons à la section 4.3, le couplage entre ces deux termes peut poser des problèmes numériques quand le terme d'advection est plus important que le terme diffusif. L'importance relative de la convection sur la diffusion au niveau du transfert de chaleur peut se mesurer à partir du nombre sans dimension de Péclet Pe_g. Afin de définir ce nombre, nous introduisons une vitesse caractéristique V, une longueur caractéristique L, une valeur typique k' de la conductivité thermique et une capacité thermique massique typique c'. Le terme de diffusion est de l'ordre de $\frac{k'}{L}\|\nabla T\|$ tandis que le terme de convection est de l'ordre de $\rho c' V\|\nabla T\|$.

1. Une remarque sur le vocabulaire employé aux sections 4.3 et 4.4 s'impose. Classiquement, la notion de stabilité d'une méthode numérique est associée à la continuité de la solution numérique par rapport aux données du problème. De ce point de vue, la méthode classique de Galerkin est stable quant elle est appliquée à l'équation d'advection-diffusion (voir section 4.2). Toutefois dans les deux sections 4.3 et 4.4, nous qualifierons de stable une méthode numérique pour laquelle la solution numérique ne présente pas d'oscillations indésirables. Une méthode de stabilisation aura ainsi pour but de réduire (ou idéalement d'éliminer) ces oscillations indésirables.

Le nombre de Péclet est alors défini comme le rapport du terme de convection sur le terme de diffusion à savoir

$$Pe_g = \frac{\rho c' V}{\frac{k'}{L}} = \frac{VL}{\kappa} \tag{4.2}$$

où $\kappa = k'/(\rho c')$ est une valeur typique pour la diffusivité thermique du fluide. Si le nombre de Péclet est faible, alors la diffusion est bien plus efficace que la convection pour le transfert de chaleur tandis que c'est l'inverse qui se produit si le nombre de Péclet est important. Le nombre de Péclet peut aussi être défini comme le rapport du temps caractéristique L^2/κ de diffusion sur le temps caractéristique L/V de la convection. Le processus physique avec le temps caractéristique le plus faible contribuera plus efficacement à une redistribution de la chaleur en cas de perturbation thermique. Outre son intérêt physique, nous verrons à la section 4.3 que ce nombre joue un rôle important dans la résolution numérique de l'équation d'advection-diffusion. C'est d'ailleurs en prévision de cette section que nous avons associé au nombre de Péclet un indice g qui se réfère au terme global. Cela nous évitera toute ambiguïté au niveau des notations lorsque nous introduirons le nombre de Péclet local ou numérique.

4.2. Équation linéaire d'advection-diffusion

Nous considérons à présent le cas simplifié de l'équation linéaire d'advection-diffusion en régime stationnaire sur un domaine $\Omega \subset \mathbb{R}^d$ avec une diffusivité thermique $\kappa > 0$ constante. Nous supposerons des conditions de Dirichlet homogènes sans perte de généralité. En effet, par un relèvement approprié des conditions aux limites, il est toujours possible de se ramener à des conditions aux limites de Dirichlet homogènes. Le problème que nous étudions est le suivant [2] :

$$\begin{cases} -\kappa\,\Delta T + \mathbf{v}\cdot\nabla T = s(\mathbf{x}) & \text{dans } \Omega, \\ T = 0 & \text{sur } \Gamma_D, \\ \kappa\,\nabla T\cdot\mathbf{n} = q(\mathbf{x}) & \text{sur } \Gamma_N. \end{cases} \tag{4.3}$$

Nous supposerons pour la suite que le champ de vitesse $\mathbf{v}$ vérifie deux hypothèses. La première hypothèse consiste à supposer que le fluide est incompressible à savoir $\operatorname{div}\mathbf{v} = 0$. De plus, nous imposons que $\mathbf{v}\cdot\mathbf{n}$ soit positif ou nul sur Γ_N, ce qui revient à supposer que cette frontière est une sortie pour le fluide ou bien une frontière impénétrable. Si nous associons Γ_N à l'interface glace-roche d'un glacier, cette seconde hypothèse est en réalité plausible. Enfin, nous admettons pour simplifier que $\mathbf{v}$ appartient à $[\mathbb{L}^\infty(\Omega)]^d$ [3].

Pour établir la formulation faible du problème (4.3), nous supposons que la fonction source $s \in \mathbb{L}^2(\Omega)$ et que le flux de chaleur $q \in \mathbb{L}^2(\Gamma_N)$. De plus, nous avons $T \in V$ où l'ensemble V est défini par

$$V = \{w \in \mathbb{H}^1(\Omega) : w = 0 \text{ sur } \Gamma_D\}. \tag{4.4}$$

2. Nous supposons que des conditions de Dirichlet sont imposées au moins sur une partie de la frontière de Ω. Si ce n'était pas le cas, l'ajout d'une fonction constante à la solution du problème thermique donnerait également une solution à ce problème.

3. L'ensemble $\mathbb{L}^\infty(\Omega)$ est défini par $\mathbb{L}^\infty(\Omega) = \{f : \Omega \to \mathbb{R}; f \text{ mesurable et } \exists \text{ une constante } C \text{ telle que } |f| \leqslant C \text{ presque partout sur } \Omega\}$ et $\|f\|_{\mathbb{L}^\infty(\Omega)} = \inf\{C; |f| \leqslant C \text{ presque partout sur } \Omega\}$ [8].

Nous considérons à présent une fonction test $w \in V$ par laquelle nous multiplions l'équation de la chaleur avant de l'intégrer sur Ω. Nous obtenons

$$-\int_\Omega \kappa \, \Delta T \, w \, d\Omega + \int_\Omega \mathbf{v} \cdot \nabla T \, w \, d\Omega = \int_\Omega s \, w \, d\Omega. \tag{4.5}$$

En intégrant ensuite par parties et en tenant compte des conditions aux limites, nous pouvons écrire

$$\int_\Omega \kappa \, \nabla T \cdot \nabla w \, d\Omega - \int_{\Gamma_N} q \, w \, d\Gamma_N + \int_\Omega \mathbf{v} \cdot \nabla T \, w \, d\Omega = \int_\Omega s \, w \, d\Omega. \tag{4.6}$$

Nous pouvons définir la forme bilinéaire

$$g(T, w) = \int_\Omega \kappa \, \nabla T \cdot \nabla w \, d\Omega + \int_\Omega w \, \mathbf{v} \cdot \nabla T \, d\Omega. \tag{4.7}$$

Nous introduisons également la fonctionnelle linéaire $L(w)$ définie par

$$L(w) = \int_\Omega s \, w \, d\Omega + \int_{\Gamma_N} q \, w \, d\Gamma_N. \tag{4.8}$$

Finalement, l'équation d'advection-diffusion s'écrit sous la forme variationnelle 4.1.

Formulation variationnelle 4.1. Trouver $T \in V = \{w \in \mathbb{H}^1(\Omega) : w = 0 \text{ sur } \Gamma_D\}$ tel que

$$g(T, w) = L(w)$$

pour tout $w \in V$.

L'écriture de la formulation variationnelle 4.1 nous permet de démontrer l'existence et l'unicité d'une solution pour le problème thermique linéaire. En effet, la forme bilinéaire $g(T, w)$ est continue et coercive et la fonctionnelle $L(w)$ est linéaire et continue. En utilisant le lemme de Lax-Milgram 3.1, nous pouvons affirmer l'existence et l'unicité de la solution au problème variationnel $g(T, w) = L(w) \, \forall w \in V$. Une vérification plus détaillée des hypothèses d'application du lemme de Lax-Milgram pourra être trouvée en annexe B. Celle-ci repose notamment sur les hypothèses que nous avons formulées pour la vitesse et permet ainsi de comprendre les raisons mathématiques de ces suppositions.

Une conséquence importante du lemme de Lax-Milgram est donnée par le corollaire 4.1 [56] qui donne une estimation a priori de la solution du problème d'advection-diffusion. Il montre en particulier que la solution et son gradient sont bornés par une expression qui dépend de la fonction source s et du flux de chaleur q. Toutefois si la diffusivité thermique κ est faible, la solution pourra présenter de forts gradients. Nous verrons à la section 4.3 que c'est effectivement le cas pour les écoulements dominés par la convection.

Corollaire 4.1. *La solution faible du problème d'advection-diffusion (4.3) est bornée de la façon suivante :*

$$\|T\|_{\mathbb{H}^1(\Omega)} \leqslant \frac{1}{\alpha}(\|s\|_{\mathbb{L}^2(\Omega)} + \|q\|_{\mathbb{L}^2(\Gamma_N)}), \; \|\nabla T\|_{\mathbb{L}^2(\Omega)} \leqslant \frac{C_\Omega}{\kappa}(\|s\|_{\mathbb{L}^2(\Omega)} + \|q\|_{\mathbb{L}^2(\Gamma_N)}),$$

où α représente la constante de coercivité et la constante C_Ω est donnée par l'inégalité de Poincaré à savoir

$$\|T\|_{\mathbb{L}^2(\Omega)} \leqslant C_\Omega \, \|\nabla T\|_{\mathbb{L}^2(\Omega)}.$$

Le corollaire 4.1 s'applique également à la solution T_h issue de la discrétisation par la méthode de Galerkin. De plus, la convergence de la méthode de Galerkin est assurée par le lemme de Céa [58].

Lemme 4.1 (Céa). *La solution de Galerkin T_h du problème discret d'advection-diffusion satisfait la relation suivante :*

$$\|T - T_h\|_V \leqslant \frac{M}{\alpha} \inf_{w_h \in V_h} \|T - w_h\|_V,$$

où α et M sont respectivement les constantes de coercivité et de continuité.

Ainsi, la méthode de Galerkin donne bien une solution qui converge vers la solution exacte du problème continu quand $h \to 0$. De plus, cette méthode est stable dans le sens qu'à une petite variation des données du problème correspond également une petite variation de la solution numérique. Toutefois, la solution numérique peut présenter un écart important par rapport à la solution exacte si le rapport M sur α est important, ce qui sera typiquement le cas si la valeur du rapport $\|\mathbf{v}\|_{\mathbb{L}^\infty(\Omega)}/\kappa$ est élevée, c'est-à-dire quand la convection thermique est bien plus importante que la diffusion de la chaleur.

Il est également intéressant de préciser qu'il est impossible d'exprimer l'équation d'advection-diffusion sous la forme d'un problème d'optimisation comme dans le cas du problème de Stokes. Cette impossibilité est liée au fait que la forme bilinéaire $g(T, w)$ n'est pas une forme symétrique à cause du terme convectif. Toutefois, dans le cas d'un problème de diffusion pure, la forme $g(T, w)$ devient symétrique et la solution du problème diffusif minimise une fonctionnelle liée à l'énergie du système.

Pour finir cette section, nous notons également que la solution du problème (4.3) souscrit au principe du minimum si $s(\mathbf{x}) \geqslant 0$, ce qui est le cas pour les écoulements fluides. Ce principe affirme que la température intérieure d'un glacier ne pourra pas être en dessous de la température minimale sur la frontière de ce glacier. En général, cette température minimale est réalisée sur la surface supérieure du glacier aux endroits soumis à l'environnement climatique le plus froid. Pour une présentation mathématique plus détaillée de ce principe, le lecteur pourra se référer à l'annexe C.

4.3. Résolution numérique de l'équation d'advection-diffusion

4.3.1. Analyse d'un problème modèle unidimensionnel

Afin de percevoir les difficultés liées à la résolution numérique d'une équation d'advection-diffusion, nous étudions tout d'abord le cas d'un problème modèle unidimensionnel défini sur un domaine Ω de dimension unitaire. L'exemple traité est inspiré de Donea et Huerta [18]. L'équation que nous souhaitons résoudre est donnée par

$$\begin{cases} -\nu \dfrac{d^2u}{dx^2} + a \dfrac{du}{dx} = 1 \quad \text{dans }]0,1[, \\ u(0) = u(1) = 0, \end{cases} \tag{4.9}$$

où u représente un champ scalaire quelconque, $\nu > 0$ est une constante qui joue le rôle de diffusivité et la constante $a > 0$ représente la valeur de la vitesse d'un fluide qui transporte la quantité u.

La solution exacte de cette équation différentielle linéaire à coefficients constants est donnée par

$$u(x) = \frac{1}{a}\left[x - \frac{1-\exp(\frac{a}{\nu}x)}{1-\exp(\frac{a}{\nu})}\right]. \tag{4.10}$$

En utilisant le nombre de Péclet Pe_g, cette solution peut également s'écrire

$$u(x) = \frac{1}{a}\left[x - \frac{1-\exp(Pe_g\, x)}{1-\exp(Pe_g)}\right]. \tag{4.11}$$

La figure 4.1 représente la solution (4.11) pour différentes valeurs du nombre de Péclet lorsque la vitesse a est fixée à 1. Nous constatons que lorsque le nombre de Péclet augmente, c'est-à-dire quand le terme convectif devient de plus en plus important, alors il se développe une petite région au voisinage de $x = 1$ dans laquelle la dérivée de la solution devient très importante et devient non bornée pour une valeur du nombre de Péclet infinie. Cette région est appelée couche limite. Nous pouvons déjà appréhender que la présence d'une telle couche limite se révélera problématique pour la résolution numérique.

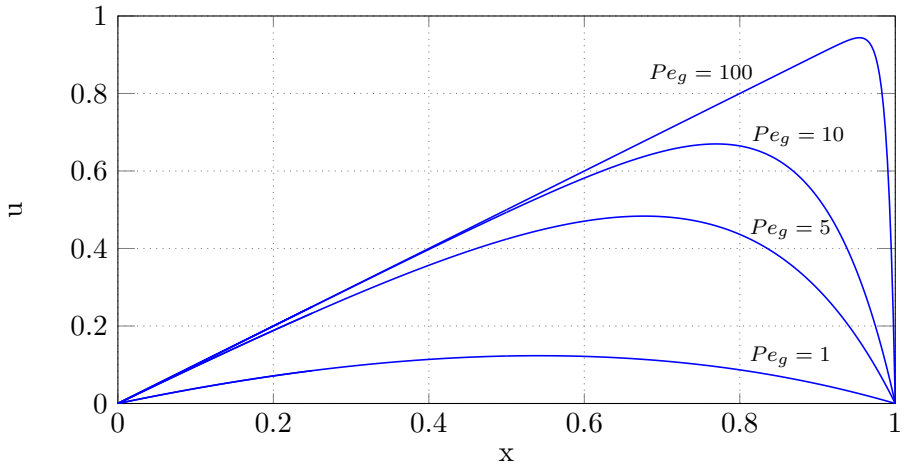


FIGURE 4.1.: Solution exacte du problème (4.9) pour différentes valeurs du nombre de Péclet global.

Afin de résoudre numériquement le problème (4.9), nous partons de la formulation variationnelle suivante [4] :

$$\text{Trouver } u \in \mathbb{H}^1_0(\Omega) \text{ tel que } \int_0^1 \nu \frac{du}{dx}\frac{dv}{dx}\,dx + \int_0^1 a\frac{du}{dx}v\,dx = \int_0^1 v\,dx \quad \forall v \in \mathbb{H}^1_0(\Omega). \tag{4.12}$$

La résolution numérique du problème (4.12) peut être menée au moyen d'une méthode éléments finis basée sur la méthode classique de Galerkin [5]. Un choix naturel de fonctions de base consiste à utiliser des fonctions linéaires par morceaux. Par facilité, nous considérons que le domaine Ω est partitionné en un ensemble de N intervalles de longueur constante h.

4. La notation $\mathbb{H}^1_0(\Omega)$ désigne l'ensemble des fonctions de $\mathbb{H}^1(\Omega)$ qui s'annulent sur la frontière de Ω.
5. Dans la méthode classique de Galerkin aussi connue sous le nom de méthode de Bubnov-Galerkin, les fonctions de base pour la discrétisation des fonctions test et de la solution sont identiques.

Les fonctions u et v sont ainsi approchées par les fonctions u_h et v_h qui peuvent s'écrire sous la forme

$$u_h = \sum_{i=1}^{N+1} u_i \, \varphi_i(x) \quad \text{et} \quad v_h = \sum_{j=1}^{N+1} v_j \, \varphi_j(x), \tag{4.13}$$

avec

$$\varphi_i(x) = \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & \text{sur } [x_{i-1}, x_i], \\ \dfrac{x - x_{i+1}}{x_i - x_{i+1}} & \text{sur } [x_i, x_{i+1}], \\ 0 & \text{sinon.} \end{cases} \tag{4.14}$$

En insérant les décompositions de u_h et v_h dans la formulation variationnelle (4.12), nous trouvons que les coefficients u_i doivent satisfaire ($u_1 = u_{N+1} = 0$)

$$-\nu \left(\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} \right) + a \left(\frac{u_{i+1} - u_{i-1}}{2h} \right) = 1, \; i = 2, \ldots, N. \tag{4.15}$$

La relation (4.15) est équivalente à une discrétisation du problème (4.9) par différences finies en utilisant une différence finie centrée pour la dérivée seconde et pour le terme de convection [6].

Il est utile d'introduire à présent le nombre de Péclet local ou numérique défini par

$$Pe = \frac{ah}{2\nu}. \tag{4.16}$$

Nous constatons que cette expression est similaire à celle définie à la relation (4.2) à la différence près que c'est la taille caractéristique d'un élément fini qui intervient dans l'expression de ce nombre et non plus la taille caractéristique du domaine. Pour distinguer ces deux nombres de Péclet, nous parlerons de nombre de Péclet local et de nombre de Péclet global [7].

La relation (4.15) peut également s'écrire à l'aide du nombre de Péclet local sous la forme

$$-\frac{Pe + 1}{Pe} \, u_{i-1} + \frac{2}{Pe} \, u_i + \frac{Pe - 1}{Pe} \, u_{i+1} = \frac{2h}{a}, \; i = 2, \ldots, N. \tag{4.17}$$

À la figure 4.2, nous avons résolu numériquement le problème (4.9) pour un nombre de Péclet $Pe_g = 100$ et une vitesse $a = 1$ en utilisant la relation (4.17). Nous avons considéré différentes valeurs du nombre de Péclet local en modifiant la valeur du pas de discrétisation. Lorsque le nombre de Péclet local est augmenté, nous constatons l'apparition d'oscillations dans la solution numérique d'autant plus marquées que nous nous trouvons proche de $x = 1$. Ces oscillations n'ont aucune signification physique et sont liées à une méthode de résolution numérique peu adéquate. Pour des nombres de Péclet $Pe < 1$, la solution numérique retrouve un comportement monotone identique à celui de la solution exacte.

6. Les discrétisations éléments finis et différences finies pour un problème unidimensionnel ne donnent pas toujours des résultats équivalents. C'est notamment le cas lorsque l'équation présente un terme source qui dépend de la position spatiale.

7. Pour avoir une définition similaire à celle du nombre de Péclet local, certains auteurs d'analyse numérique définissent le nombre de Péclet global en multipliant la relation (4.2) par un facteur 1/2. Nous avons toutefois décidé de conserver la définition communément utilisée en physique.

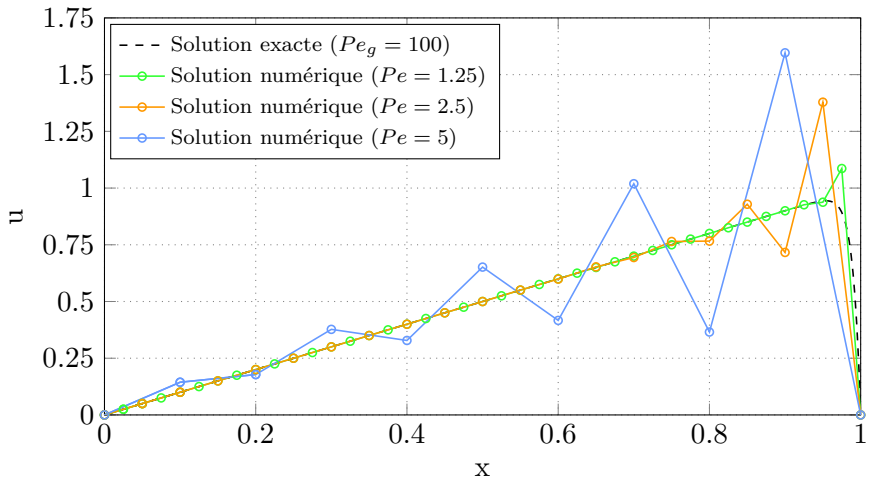


FIGURE 4.2.: Résolution numérique du problème (4.9) pour différentes valeurs du nombre de Péclet local.

Nous pouvons remarquer que le nombre de Péclet local dépend à la fois du nombre de Péclet global et du pas spatial de discrétisation. La solution numérique présentera un comportement oscillant d'autant plus facilement que le problème physique est dominé par la convection et que le pas de discrétisation est grand. Nous pourrions être tenté de choisir un pas de discrétisation suffisamment petit pour rendre la solution numérique stable. Cependant, cette stratégie serait inefficace, car elle mènerait dans certains cas à choisir un pas spatial très petit et donc à un coût de calcul important. C'est pourquoi, il est nécessaire d'utiliser des méthodes de stabilisation adéquates dont nous présenterons quelques exemples à la sous-section 4.3.2.

Pour souligner le fait que le comportement oscillant de la solution numérique est lié à la méthode de discrétisation utilisée, il est intéressant de considérer la solution homogène de l'équation aux différences (4.17). Cette solution s'écrit pour $Pe \neq 1$

$$u_i = C_1 + C_2 \left(\frac{1 + Pe}{1 - Pe} \right)^i , \tag{4.18}$$

avec C_1 et C_2 qui dépendent des conditions aux limites. L'expression (4.18) montre clairement que la solution numérique présente des oscillations pour $Pe > 1$.

Afin d'établir des méthodes de stabilisation, il convient de comprendre pourquoi la méthode classique de Galerkin ne donne pas un solution précise. Pour cela, il est utile de regarder au terme convectif de la relation (4.15) à savoir

$$a \, \frac{u_{i+1} - u_{i-1}}{2h}. \tag{4.19}$$

Il s'agit d'une discrétisation par une différence centrée, ce qui signifie que les valeurs u_{i+1} et u_{i-1} contribuent de manière identique à la valeur de u_i. Cependant, la présence de convection dans un système implique un transfert de l'information suivant une direction privilégiée qui est la direction de l'écoulement. Si un écoulement unidimensionnel se propage vers la droite, nous nous attendons naturellement à ce que la valeur de u_i dépende plus de u_{i-1} que de u_{i+1}.

En fait, il peut être montré (voir [18]) que l'équation aux différences qui fournit la solution exacte aux nœuds de discrétisation est donnée par

$$-\nu\left(\frac{u_{i+1}-2u_i+u_{i-1}}{h^2}\right)+\frac{1-\beta}{2}\left(a\frac{u_{i+1}-u_i}{h}\right)+\frac{1+\beta}{2}\left(a\frac{u_i-u_{i-1}}{h}\right)=1, \tag{4.20}$$

avec

$$\beta=\coth Pe-\frac{1}{Pe}. \tag{4.21}$$

Par rapport à l'expression (4.15), le terme convectif dans la relation (4.20) ne correspond plus à une différence centrée. Les inconnues u_{i-1} et u_{i+1} ont été pondérées différemment de manière à donner plus d'importance au flux convectif à la gauche de u_i. Pour éviter un comportement oscillant de la solution numérique, il convient donc d'introduire de nouvelles méthodes qui permettront de dissymétriser la discrétisation du terme convectif. Cette démarche est en fait naturelle, car elle revient à conserver la dissymétrie du terme convectif dans la formulation faible du problème d'advection-diffusion.

Pour obtenir cet effet de dissymétrie dans la méthode de Galerkin, un choix naturel est d'utiliser des fonctions de base différentes pour la discrétisation de la solution et des fonctions test. La méthode ainsi utilisée est appelée méthode de Petrov-Galerkin et nous verrons que cette méthode est à la base des méthodes de stabilisation présentées à la sous-section 4.3.2. Pour le cas unidimensionnel, un choix possible de fonctions de base pour les fonctions test est illustré à la figure 4.3(b). Cette fonction est obtenue en ajoutant une fonction bulle qui permet de pondérer un nœud en amont de l'écoulement de manière plus importante qu'un nœud en aval.

(a) Fonction de base pour la solution.

(b) Fonction de base pour les fonctions test.

FIGURE 4.3.: Illustration de la méthode de Petrov-Galerkin.

Même si nous avons limité pour l'instant notre exposé à un cas unidimensionnel, il convient de préciser que les problèmes liés à une instabilité de la solution numérique pour des problèmes dominés par la convection sont également présents en dimensions spatiales supérieures. De plus, l'utilisation de fonctions de base d'ordre supérieur ne permet pas de stabiliser la méthode. L'ajout de nœuds internes aux éléments introduit en fait des difficultés supplémentaires, car le comportement des nœuds internes et externes n'est pas identique.

4.3.2. Méthodes de stabilisation

Nous présentons maintenant quelques méthodes de stabilisation qui sont fréquemment utilisées pour les problèmes d'advection-diffusion. Pour cela, nous considérons le problème général suivant :

$$\begin{cases} -\operatorname{div}(\nu\nabla u) + \mathbf{a}\cdot\nabla u = s & \text{dans } \Omega, \\ u = u_D & \text{sur } \Gamma_D, \\ \nu\,\nabla u\cdot\mathbf{n} = h & \text{sur } \Gamma_N, \end{cases} \tag{4.22}$$

où $\mathbf{a}(\mathbf{x})$ correspond au champ de vitesse dans Ω, $s(\mathbf{x})$ est un terme source et ν est le coefficient de diffusion comme à la sous-section 4.3.1.

Pour établir une formulation faible de ce problème, nous introduisons les espaces fonctionnels suivants :

$$S = \{u \in \mathbb{H}^1(\Omega) : u = u_D \text{ sur } \Gamma_D\}, \tag{4.23}$$

$$V = \{w \in \mathbb{H}^1(\Omega) : w = 0 \text{ sur } \Gamma_D\}. \tag{4.24}$$

La formulation faible du problème (4.22) est donnée par

$$\text{Trouver } u \in S \text{ tel que } a(u,w) + c(\mathbf{a};u,w) = (s,w) + (h,w)_{\Gamma_N} \quad \forall w \in V, \tag{4.25}$$

où nous avons introduit les quantités suivantes :

$$\begin{aligned} a(u,w) &= \int_\Omega \nu\,\nabla u\cdot\nabla w\,d\Omega, & (s,w) &= \int_\Omega s\,w\,d\Omega, \\ c(\mathbf{a};u,w) &= \int_\Omega (\mathbf{a}\cdot\nabla u)\,w\,d\Omega, & (h,w)_{\Gamma_N} &= \int_{\Gamma_N} w\,h\,d\Gamma_N. \end{aligned} \tag{4.26}$$

L'application de la méthode classique de Galerkin à la formulation faible (4.25) donne lieu à une solution numérique imprécise pour des nombres de Péclet élevés. L'idée des méthodes de stabilisation est d'ajouter un terme de stabilisation à la formulation variationnelle afin de compenser les déficiences de la méthode classique de Galerkin. Les méthodes de stabilisation couramment utilisées ont été établies de manière à obtenir des méthodes d'approximation fortement consistantes, c'est-à-dire que la solution exacte du problème continu satisfait la méthode d'approximation indépendamment de la discrétisation du domaine.

Les méthodes de stabilisation fortement consistantes peuvent s'écrire sous la forme générale :

Trouver $u \in S$ tel que

$$a(u,w) + c(\mathbf{a};u,w) + \sum_e \int_{\Omega_e} \mathcal{P}(w)\,\tau\,\mathcal{R}(u)\,d\Omega_e = (s,w) + (h,w)_{\Gamma_N} \quad \forall w \in V, \tag{4.27}$$

où $\mathcal{P}(w)$ représente un opérateur qui agit sur les fonctions test w, τ est un paramètre de stabilisation et $\mathcal{R}(u)$ est le résidu fort de l'équation à résoudre à savoir

$$\mathcal{R}(u) = -\operatorname{div}(\nu\nabla u) + \mathbf{a}\cdot\nabla u - s. \tag{4.28}$$

Dans l'expression (4.27), nous constatons que le terme de stabilisation fait intervenir une somme sur chacun des éléments. Cela permet de tenir compte de discontinuités éventuelles

entre les éléments [23] [8]. Il est évident que toute discrétisation de la formulation (4.27) sera satisfaite par la solution exacte du problème, car cette solution exacte annule le résidu $\mathcal{R}(u)$ et satisfait également l'expression (4.25). L'expression (4.27) donnera donc bien lieu à des méthodes fortement consistantes.

La formulation variationnelle (4.27) peut être interprétée comme un cas particulier de la méthode de Petrov-Galerkin. En effet, considérons l'équation différentielle

$$\mathcal{R}(u) = 0. \tag{4.29}$$

Une formulation faible de cette équation peut être établie à partir des fonctions test $w^* = w + \tau\,\mathcal{P}(w)$. En multipliant la relation (4.29) par w^* et en intégrant sur Ω, il vient

$$\int_\Omega w^*\,\mathcal{R}(u)\,d\Omega = 0. \tag{4.30}$$

En pratique, cette dernière expression est décomposée sous la forme

$$\int_\Omega w\,\mathcal{R}(u)\,d\Omega + \sum_e \int_{\Omega_e} \tau\,\mathcal{P}(w)\,\mathcal{R}(u)\,d\Omega_e = 0. \tag{4.31}$$

En intégrant ensuite par parties l'équation (4.31), nous retrouvons la formulation (4.27). Ainsi, la méthode de stabilisation revient à travailler avec des fonctions test $w^* \neq w$. Dès lors, même si w et u sont discrétisés à l'aide des mêmes fonctions de base, l'introduction du terme de stabilisation modifie les fonctions test de manière à travailler en réalité avec une méthode de Petrov-Galerkin.

Différentes méthodes de stabilisation peuvent être définies en fonction de l'expression de l'opérateur $\mathcal{P}$. Nous présentons deux cas particuliers ci-dessous.

- La méthode SUPG (*Streamline Upwind Petrov-Galerkin*) introduite par Brooks et Hughes [10] est définie par

  $$\mathcal{P}(w) = \mathbf{a}\cdot\nabla w. \tag{4.32}$$

  Cet opérateur conduit à la méthode discrétisée suivante :

  Trouver $u_h \in S_h$ tel que

  $$\begin{aligned} a(u_h, w_h) + c(\mathbf{a}; u_h, w_h) + \sum_e \int_{\Omega_e} (\mathbf{a}\cdot\nabla w_h)\,\tau\,\left[\mathbf{a}\cdot\nabla u_h - \operatorname{div}(\nu\nabla u_h)\right] d\Omega_e \\ = (s, w_h) + (h, w_h)_{\Gamma_N} + \sum_e \int_{\Omega_e} (\mathbf{a}\cdot\nabla w_h)\,\tau\,s\,d\Omega_e \quad \forall w_h \in V_h. \end{aligned} \tag{4.33}$$

  Ce choix pour l'opérateur $\mathcal{P}$ est assez naturel dans la mesure où $\mathcal{P}(w)$ représente la dérivée directionnelle des fonctions test dans la direction du vecteur vitesse. De cette façon, les fonctions test sont corrigées de manière à prendre en compte le caractère directionnel du vecteur vitesse.

8. Cette remarque se comprend facilement dans un cas unidimensionnel. Si les fonctions test w sont discrétisées au moyen de fonctions linéaires par éléments, la dérivée seconde de ces fonctions sera nulle au sein des éléments, mais sera égale à une distribution de Dirac à la frontière des éléments. Dès lors $\int_\Omega \frac{d^2w}{dx^2}\,d\Omega \neq \sum_e \int_{\Omega_e} \frac{d^2w}{dx^2}\,d\Omega_e$.

- La méthode GLS (*Galerkin Least-Squares*) introduite par Hughes *et al.* [36] fait usage de l'opérateur

$$\mathcal{P}(w) = \mathbf{a} \cdot \nabla w - \operatorname{div}(\nu \nabla w). \tag{4.34}$$

La formulation variationnelle discrétisée associée à cet opérateur est donnée par :

Trouver $u_h \in S_h$ tel que

$$\begin{aligned} &a(u_h, w_h) + c(\mathbf{a}; u_h, w_h) + \sum_e \int_{\Omega_e} \left[\mathbf{a} \cdot \nabla w_h - \operatorname{div}(\nu \nabla w_h)\right] \tau \left[\mathbf{a} \cdot \nabla u_h - \operatorname{div}(\nu \nabla u_h)\right] d\Omega_e \\ &\quad = (s, w_h) + (h, w_h)_{\Gamma_N} + \sum_e \int_{\Omega_e} \left[\mathbf{a} \cdot \nabla w_h - \operatorname{div}(\nu \nabla w_h)\right] \tau \, s \, d\Omega_e \quad \forall w_h \in V_h. \end{aligned} \tag{4.35}$$

La méthode GLS permet de travailler avec un terme de stabilisation symétrique dans le membre de gauche de la formulation variationnelle. De plus, nous constatons que la méthode GLS se ramène à la méthode SUPG pour des fonctions linéaires par morceaux.

Bien que nous n'ayons pas encore abordé le sujet, le paramètre de stabilisation τ joue un rôle majeur dans la méthode de discrétisation. Il détermine le poids du terme de stabilisation dans la formulation variationnelle. Le choix de ce paramètre est en pratique difficile et constitue toujours un sujet de discussion. Il dépend notamment du type d'éléments utilisés, du problème étudié et doit être évalué élément par élément pour des maillages irréguliers.

Le paramètre τ peut par exemple être défini comme

$$\tau = \frac{\bar{\nu}}{\|\mathbf{a}\|^2}, \tag{4.36}$$

où $\bar{\nu}$ est défini comme un coefficient de diffusion artificiel. Dans le cas d'un élément quadrangulaire bilinéaire, Brooks et Hughes [10] suggèrent d'utiliser l'expression

$$\bar{\nu} = \frac{\bar{\xi} a_\xi h_\xi + \bar{\eta} a_\eta h_\eta}{2} \tag{4.37}$$

avec

$$\bar{\xi} = \coth Pe_\xi - \frac{1}{Pe_\xi}, \qquad \bar{\eta} = \coth Pe_\eta - \frac{1}{Pe_\eta}, \tag{4.38a}$$

$$Pe_\xi = \frac{a_\xi h_\xi}{2\nu}, \qquad Pe_\eta = \frac{a_\eta h_\eta}{2\nu}, \tag{4.38b}$$

$$a_\xi = \mathbf{e}_\xi \cdot \mathbf{a}, \qquad a_\eta = \mathbf{e}_\eta \cdot \mathbf{a}. \tag{4.38c}$$

Les vecteurs unitaires $\mathbf{e}_\xi$ et $\mathbf{e}_\eta$ ainsi que les longueurs h_ξ et h_η sont représentés sur la figure 4.4. Nous constatons que l'expression du coefficient de diffusion artificiel fait intervenir le nombre de Péclet local évalué suivant les directions $\mathbf{e}_\xi$ et $\mathbf{e}_\eta$.

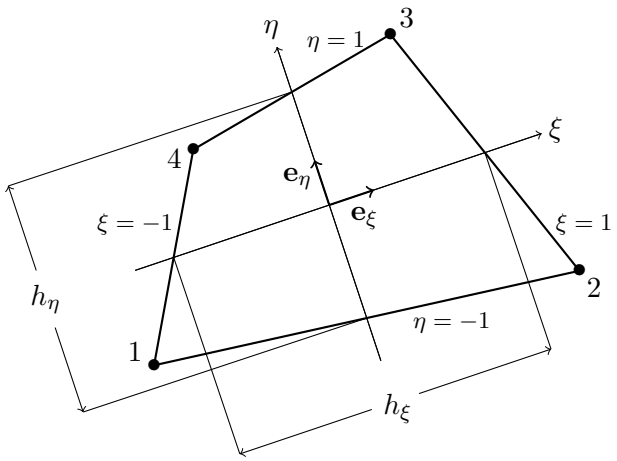


FIGURE 4.4.: Élément quadrangulaire bilinéaire.

Comme le soulignent Brooks et Hughes [10], le choix (4.37) pour le coefficient de diffusion artificiel semble relativement ad hoc. Il constitue en fait une simple généralisation au cas bidimensionnel de la valeur de ce paramètre pour le problème unidimensionnel. Pour ce problème, il est en effet possible de déterminer le paramètre de stabilisation de manière à obtenir une solution exacte aux nœuds de discrétisation. Il semblerait toutefois que la valeur exacte de $\bar{\nu}$ est bien moins importante que la structure des fonctions test.

Enfin, soulignons également que la vitesse $\|\mathbf{a}\|$ qui apparaît dans l'expression (4.36) n'est pas clairement définie. Il est possible de considérer pour cette quantité la vitesse au centre de l'élément, la vitesse moyenne sur l'ensemble des points d'intégration ou encore la vitesse au point d'intégration considéré.

4.3.3. Application numérique

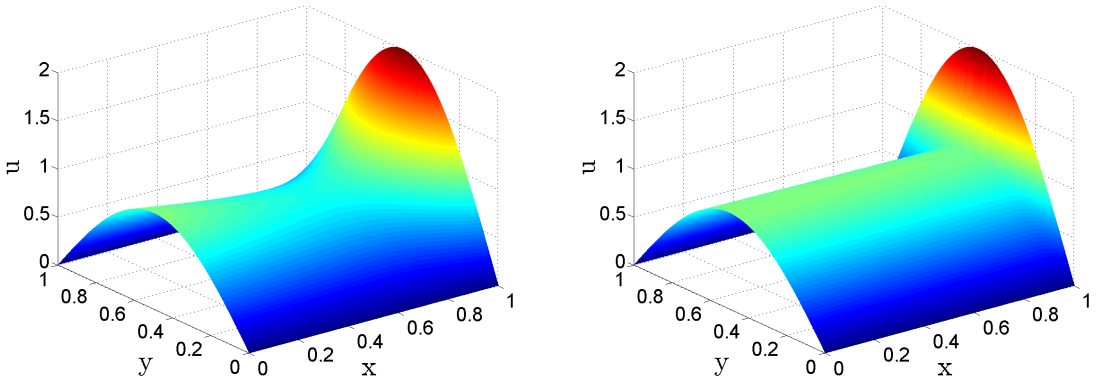
Afin de confirmer l'efficacité des méthodes de stabilisation présentées, nous appliquons celles-ci à une problème proposé par Gupta *et al.* [32, 52]. Le problème est le suivant :

$$\begin{cases} -\nu\Delta u + \mathbf{a}\cdot\nabla u = s & \text{sur } \Omega =]0,1[\times]0,1[\,, \\ u(x,0) = u(x,1) = 0 & \text{pour } x \in [0,1], \\ u(0,y) = \sin(\pi y) & \text{pour } y \in [0,1], \\ u(1,y) = 2\sin(\pi y) & \text{pour } y \in [0,1]. \end{cases} \tag{4.39}$$

Ce problème est étudié en considérant que $\mathbf{a} = (1,0)$. Le terme source $s(\mathbf{x})$ est calculé de manière à ce que la solution exacte soit donnée par

$$u(\mathbf{x}) = \frac{1}{\sinh\sigma}\sin(\pi y)\,\left[2\sinh(\sigma x) + e^{\frac{x}{2\nu}}\,\sinh(\sigma(1-x))\right], \quad \sigma = \sqrt{\pi^2 + \frac{1}{4\nu^2}}\,. \tag{4.40}$$

Sur la figure 4.5, nous avons représenté cette solution pour deux valeurs différentes du coefficient de diffusion ν. Comme nous pouvons le constater, la solution présente peu de variations sur l'ensemble du domaine sauf au voisinage du bord $x = 1$. Sur ce côté, nous remarquons le développement d'une couche limite où la solution présente un gradient d'autant plus important que le coefficient de diffusion est faible.



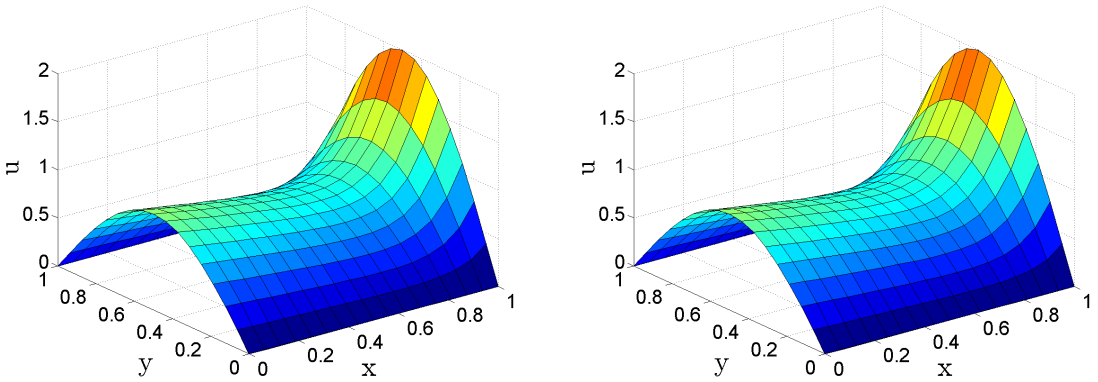
(a) Solution pour $\nu = 0.05$.

(b) Solution pour $\nu = 0.004$.

FIGURE 4.5.: Solution exacte du problème (4.39).

Nous avons résolu numériquement le problème (4.39) par une méthode éléments finis en considérant à la fois la méthode classique de Galerkin et la méthode de stabilisation SUPG. La résolution a été réalisée au moyen d'éléments quadrangulaires bilinéaires pour un pas de discrétisation $h = 1/16$ dans chacune des deux directions spatiales. Dans le cas d'éléments bilinéaires, la méthode GLS est en fait équivalente à la méthode SUPG de sorte que cela ne présente aucun intérêt supplémentaire de considérer cette méthode.

Nous avons d'abord résolu le problème pour $\nu = 0.05$. Pour cette situation, le nombre de Péclet local dans la direction de l'écoulement est donné par $Pe = 0.625$. Comme le nombre de Péclet est faible, nous pouvons nous attendre à obtenir une solution numérique acceptable avec la méthode classique de Galerkin. Cette intuition est confirmée par la figure 4.6 sur laquelle nous constatons que les méthodes stabilisée et non stabilisée donnent des résultats proches de la solution analytique.



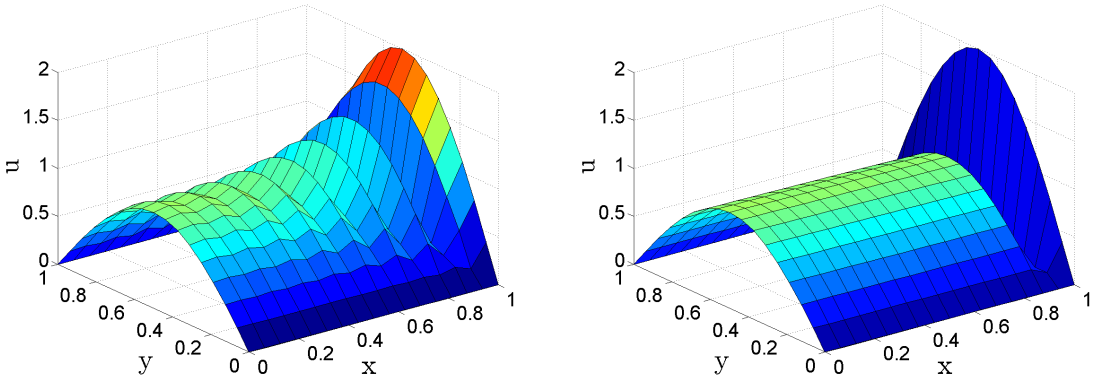
(a) Méthode classique de Galerkin.

(b) Méthode SUPG.

FIGURE 4.6.: Solution numérique du problème (4.39) avec $\nu = 0.05$.

Nous envisageons à présent un second cas pour lequel $\nu = 0.004$. Dans ce cas, le nombre de Péclet local vaut 7.8125 et nous pouvons soupçonner une déficience dans la méthode classique de Galerkin. La figure 4.7(a) montre que cette méthode donne une solution qui présente des oscillations qui n'existent pas dans la solution analytique. Au contraire, la figure 4.7(b) indique

que l'application de la méthode SUPG au problème (4.39) a pour conséquence d'éliminer les modes oscillants introduits par la méthode classique de Galerkin. La solution numérique est à présent bien plus proche de la solution analytique.



(a) Méthode classique de Galerkin.

(b) Méthode SUPG.

FIGURE 4.7.: Solution numérique du problème (4.39) avec $\nu = 0.004$.

4.4. Nombre de Péclet typique pour les glaciers

La section précédente a montré que pour un transfert thermique dominé par la convection, la méthode classique de Galerkin pouvait mener à des oscillations dans la solution numérique, ce qui nécessite l'utilisation d'une méthode de stabilisation. Une question pertinente est de se demander si pour les glaciers le transfert convectif de chaleur est plus important que le transfert diffusif. Dans le cas d'une réponse affirmative, l'utilisation d'une méthode éléments finis stabilisés sera nécessaire pour la résolution numérique du problème thermique.

Pour évaluer l'importance du transfert thermique par convection dans les glaciers, nous allons nous baser sur un argument dimensionnel en calculant un ordre de grandeur pour le nombre de Péclet des glaciers. Les dimensions des glaciers pouvant présenter des différences d'ordre de grandeur fort importantes, nous avons jugé plus convenable de considérer deux cas à savoir le cas d'une calotte polaire et d'un glacier de montagne.

Une calotte polaire est caractérisée par les grandeurs caractéristiques suivantes :

$$\begin{aligned}
\text{Longueur horizontale caractéristique : } L &= 1000\,[\text{km}],\\
\text{Longueur verticale caractéristique : } H &= 1\,[\text{km}],\\
\text{Vitesse horizontale caractéristique : } U &= 100\,[\text{m an}^{-1}],\\
\text{Vitesse verticale caractéristique : } W &= 0.1\,[\text{m an}^{-1}],\\
\text{Masse volumique typique : } \rho &= 910\,[\text{kg m}^{-3}],\\
\text{Conductivité thermique typique : } k &= 2.4\,[\text{W m}^{-1}\,\text{K}^{-1}],\\
\text{Capacité thermique massique typique : } c &= 1900\,[\text{J kg}^{-1}\,\text{K}^{-1}].
\end{aligned}$$

Nous évaluons à présent le nombre de Péclet global pour une calotte polaire en considérant les directions horizontale et verticale. Nous obtenons les valeurs suivantes :

$$\begin{aligned} \text{Nombre de Péclet horizontal } Pe_g &\approx 2 \times 10^6, \\ \text{Nombre de Péclet vertical } Pe_g &\approx 2. \end{aligned}$$

Nous constatons que pour la direction horizontale, le transfert de chaleur par convection est très important comparé au transfert par conduction thermique. Cela peut notamment se justifier par le fait que la dimension horizontale du glacier est si importante que la diffusion thermique devient inefficace pour redistribuer de l'énergie sur de telles distances.

Pour un glacier de montagne, nous avons les ordres de grandeur suivants :

$$\begin{aligned} \text{Longueur horizontale caractéristique : } L &= 1\,[\text{km}], \\ \text{Longueur verticale caractéristique : } H &= 100\,[\text{m}], \\ \text{Vitesse horizontale caractéristique : } U &= 10\,[\text{m an}^{-1}], \\ \text{Vitesse verticale caractéristique : } W &= 0.1\,[\text{m an}^{-1}]. \end{aligned}$$

Nous obtenons alors

$$\begin{aligned} \text{Nombre de Péclet horizontal } Pe_g &\approx 200, \\ \text{Nombre de Péclet vertical } Pe_g &\approx 0.2. \end{aligned}$$

Pour un glacier de montagne, la convection joue également un rôle important sur le profil horizontal de température tandis que suivant la direction verticale, la diffusion thermique est plus efficace que la convection pour le transfert de chaleur. Selon Cuffey et Patterson [15], la diffusion thermique horizontale est en générale plus faible que la diffusion verticale, car les gradients de température sont plus faibles dans cette direction. La convection horizontale dans les glaciers de montagne a pour effet de transporter de la glace plus froide des régions de haute altitude vers les régions de basse altitude et réduit ainsi les différences de température suivant l'altitude.

Les considérations dimensionnelles ci-dessus nous incitent à penser que l'équation de la chaleur pour les glaciers est bien dominée par la convection avec une préférence pour certaines directions. Bien que nous ayons calculé le nombre de Péclet global, il est naturel de penser que le nombre de Péclet local aura lui aussi une valeur élevée si la taille caractéristique des éléments finis n'est pas trop petite. Dès lors, il est pertinent d'appliquer une méthode de stabilisation pour la résolution numérique de l'équation pour la température. L'utilisation de telles méthodes n'est toutefois pas systématique et dépendra du cas concret étudié.

Pour terminer cette sous-section, il est intéressant de constater que dans le cas du problème de Stokes, le terme de viscosité associé à une diffusion de la quantité de mouvement du fluide est bien plus important que le terme convectif considéré comme négligeable, alors que pour le problème thermique, le transfert convectif de la chaleur peut être bien plus important que le transfert diffusif. Cette remarque peut se comprendre à partir du nombre de Prandtl $Pr = \mu\backslash(\rho\kappa)$ qui est le rapport de la diffusivité de quantité de mouvement sur la diffusivité thermique. Comme $\mu/\rho >> \kappa$, nous comprenons alors l'importance plus marquée de la diffusion de quantité de mouvement par rapport à la diffusion thermique.

4.5. Traitement numérique de la contrainte thermique

4.5.1. Inégalité variationnelle pour le problème thermique

Au chapitre 2, nous avions présenté le problème thermique utilisé pour déterminer le champ de température dans un glacier. À ce moment, nous avions souligné que pour avoir un problème mathématique représentatif de la réalité physique, nous devions imposer une contrainte sur la valeur du champ de température dans le glacier afin de prendre en compte le changement de phase de la glace à son point de fusion. Nous allons à présent regarder comment cette contrainte peut être traitée mathématiquement et numériquement. Pour cela, nous supposerons qu'à l'intérieur du glacier la température est strictement inférieure à T_m, mais que cette température peut être atteinte au niveau de l'interface glace-sol. Si la température était égale à la température de fusion de la glace au sein du glacier, il nous faudrait inclure un terme source supplémentaire (terme de chaleur latente) dans l'équation de la chaleur.

Nous souhaitons ainsi examiner plus en détail le problème thermique suivant :

$$\begin{cases} -k\,\Delta T + \rho c\,\mathbf{v}\cdot\nabla T = s & \text{dans } \Omega, \\ T = T_s & \text{sur } \Gamma_s, \\ k\,\nabla T\cdot\mathbf{n} = q - \begin{cases} \rho L a_b & \text{si } T = T_m \\ 0 & \text{si } T < T_m \end{cases} & \text{sur } \Gamma_b. \end{cases} \tag{4.41}$$

où q représente un flux de chaleur générique qui ne dépend pas de la température. Nous constatons que l'équation différentielle à résoudre est une équation linéaire (k, ρ, c et s sont supposés indépendants de la température). Toutefois, le problème (4.41) est non linéaire à cause de la présence de la condition aux limites sur Γ_b. Ce genre de conditions aux limites non linéaires est parfois qualifié de condition de contact par référence aux problèmes mécaniques de contact. Cette non-linéarité est à distinguer des non-linéarités liées par exemple aux lois comportementales. La façon mathématique de traiter ces deux types de non-linéarité est bien entendu différente. Les non-linéarités de type contact présentent une difficulté particulière, car elles s'accompagnent généralement d'une modification brusque du comportement du système.

Il est courant de représenter le comportement du système à l'aide d'un graphe comportemental tel que celui donné à la figure 4.8 qui souligne la forte non-linéarité au niveau des conditions aux limites.

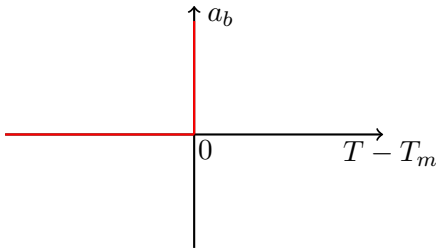


FIGURE 4.8.: Graphe de la loi a_b.

Les conditions aux limites imposent à la solution de vérifier sur Γ_b les trois conditions

$$T \leqslant T_m, \tag{4.42a}$$
$$a_b \geqslant 0, \tag{4.42b}$$
$$(T - T_m)\, a_b = 0. \tag{4.42c}$$

La relation (4.42a) signifie que sur la frontière Γ_b la température ne peut pas excéder T_m, alors que la contrainte (4.42b) impose que le taux de fonte soit positif ou nul sur cette même frontière. La condition (4.42c) est appelée condition de complémentarité. Elle exprime que la fonte de la glace, c'est-à-dire $a_b \neq 0$, ne peut se produire que si $T = T_m$. Les conditions (4.42a) à (4.42c) sont communément appelées conditions de Signorini. Nous présenterons à la sous-section 4.5.3 le raisonnement théorique qui permet d'interpréter les conditions de Signorini comme les conditions d'optimalité de Karush-Kuhn-Tucker (conditions KKT).

En principe, la frontière Γ_b peut se décomposer en deux régions sur lesquelles nous avons respectivement $T = T_m$ et $a_b = 0$. Cependant, ces régions ne sont pas connues a priori et c'est la résolution du problème (4.41) qui permettra de les déterminer. Ce problème peut donc être perçu comme un problème à frontière libre.

Il est courant d'exprimer les problèmes contraints sous la forme générale d'une inégalité variationnelle. Pour ce faire, nous introduisons les espaces de fonctions suivants :

$$H = \{\theta \in \mathbb{H}^1(\Omega) : \theta = 0 \ \text{ sur } \ \Gamma_s\}, \tag{4.43}$$
$$G = \{\theta \in \mathbb{H}^1(\Omega) : \theta = T_s \ \text{ sur } \ \Gamma_s\}, \tag{4.44}$$
$$K = \{\theta \in G : \theta \leqslant T_m \ \text{ sur } \ \Gamma_b\}. \tag{4.45}$$

Il est intéressant de constater que les deux ensembles G et K sont tous les deux convexes. L'ensemble K représente l'ensemble des champs de température admissibles. Pour établir une formulation variationnelle du problème (4.41), nous considérons une fonction arbitraire $\theta \in K$ et nous introduisons $\psi = \theta - T$. La solution $T \in K$ du problème (4.41) satisfait

$$\int_\Omega -k\,\Delta T\,\psi\,d\Omega + \int_\Omega \rho c\,(\mathbf{v}\cdot\nabla T)\,\psi\,d\Omega = \int_\Omega s\,\psi\,d\Omega. \tag{4.46}$$

En intégrant par parties et en éliminant l'intégrale sur Γ_s car $\psi \in H$, il vient

$$\int_\Omega k\,\nabla T\cdot\nabla\psi\,d\Omega - \int_{\Gamma_b} (k\,\nabla T\cdot\mathbf{n})\,\psi\,d\Gamma_b + \int_\Omega \rho c\,(\mathbf{v}\cdot\nabla T)\,\psi\,d\Omega = \int_\Omega s\,\psi\,d\Omega. \tag{4.47}$$

En reprenant les notations (4.26), nous pouvons écrire

$$a(T,\psi) + c(\mathbf{v};T,\psi) = (s,\psi) + (q,\psi)_{\Gamma_b} - \int_{\Gamma_b} \rho L a_b\,\psi\,d\Gamma_b. \tag{4.48}$$

Or, nous avons également

$$-\int_{\Gamma_b} \rho L a_b\,(\theta - T)\,d\Gamma_b = \int_{\Gamma_b} \rho L a_b\,(T - T_m)\,d\Gamma_b + \int_{\Gamma_b} \rho L a_b\,(T_m - \theta)\,d\Gamma_b \tag{4.49a}$$
$$= \int_{\Gamma_b} \rho L a_b\,(T_m - \theta)\,d\Gamma_b \tag{4.49b}$$
$$\geqslant 0. \tag{4.49c}$$

Le passage de (4.49a) à (4.49b) utilise la condition de complémentarité (4.42c) tandis que le passage de (4.49b) à (4.49c) exploite la condition (4.42b) et le fait que θ appartienne à K.

Finalement, le problème (4.41) s'écrit sous la forme de l'inéquation variationnelle suivante :

Formulation variationnelle 4.2. Trouver $T \in K = \{\theta \in G : \theta \leqslant T_m \text{ sur } \Gamma_b\}$ tel que

$$a(T, \theta - T) + c(\mathbf{v}; T, \theta - T) \geqslant (s, \theta - T) + (q, \theta - T)_{\Gamma_b}$$

pour tout $\theta \in K$.

L'étude de l'existence et de l'unicité d'une solution à cette inégalité variationnelle peut être réalisée sur base du théorème de Lions-Stampacchia [49] qui constitue en fait une généralisation du théorème de Lax-Milgram aux inégalités variationnelles. Le raisonnement que nous avons mené permet de montrer que si une solution T satisfait le problème (4.41), alors cette solution vérifie la formulation variationnelle 4.2. Une démarche similaire peut être effectuée en sens inverse pour démontrer qu'une solution suffisamment régulière de l'inégalité variationnelle satisfait nécessairement le problème (4.41) (voir détails en annexe D). Cela permet ainsi de démontrer l'équivalence entre le problème local et sa formulation variationnelle.

La résolution de l'inéquation variationnelle peut être réalisée de différentes façons. Nous en proposons deux ci-dessous. Ces deux méthodes seront présentées dans le cadre d'espaces fonctionnels de dimension infinie. L'application d'une discrétisation de celles-ci par une méthode de Galerkin donne lieu à des méthodes de résolution par éléments finis du problème thermique.

4.5.2. Méthodes de pénalisation

Une première approche pour résoudre une inégalité variationnelle repose sur les méthodes de pénalisation. Ces méthodes sont particulièrement appréciées pour leur simplicité même si elles peuvent poser des difficultés lors de leur application concrète. L'idée de la pénalisation est de transformer le problème contraint initial en un problème approché non contraint via l'introduction d'un terme de pénalisation. Pour ce faire, nous introduisons un opérateur de pénalisation P qui satisfait les propriétés suivantes [28] :

$$\begin{aligned} &\bullet\ P(\theta) = 0 \text{ si } \theta \in K, \\ &\bullet\ P \text{ est monotone, c'est-à-dire} \\ &\quad \int_\Omega \left(P(\theta) - P(\chi)\right)(\theta - \chi)\, d\Omega = (P(\theta) - P(\chi), \theta - \chi) \geqslant 0 \quad \forall \theta, \chi \in G. \end{aligned} \tag{4.50}$$

Dans le cas d'une méthode de pénalisation, le problème variationnel 4.2 est approché par le problème suivant :

Pour $\varepsilon > 0$, trouver $T_\varepsilon \in G$ tel que

$$a(T_\varepsilon, \theta) + c(\mathbf{v}; T_\varepsilon, \theta) + \frac{1}{\varepsilon}(P(T_\varepsilon), \theta)_{\Gamma_b} = (s, \theta) + (q, \theta)_{\Gamma_b} \quad \forall \theta \in H. \tag{4.51}$$

Le problème pénalisé (4.51) fournit une solution approchée de l'inégalité variationnelle. Toutefois, il peut être montré que si $\varepsilon \to 0$, alors $T_\varepsilon \to T$ avec T la solution du problème initial. Nous renvoyons le lecteur à l'annexe E pour la preuve de la convergence de la méthode de pénalisation.

En pratique, l'idée des méthodes de pénalisation consiste à résoudre le problème pénalisé pour une suite décroissante de paramètres de stabilisation ε_k. Quand $\varepsilon_k \to 0$, la contrainte est de plus en plus pénalisée et la solution T_{ε_k} est de plus en plus forcée à se trouver dans l'ensemble K. Dans le cas extrême où le paramètre de pénalisation est nul, le terme de pénalisation devient infini et la seule façon de satisfaire le problème pénalisé est de se trouver dans K.

Il existe différents choix possibles pour l'opérateur P. Nous en présentons deux plus spécifiquement :

- Un premier choix pour le terme de pénalisation est donné par

$$\frac{1}{\varepsilon} P(T_\varepsilon) = \frac{1}{\varepsilon} \mathbf{1}_{\mathbb{R}^+}(T_\varepsilon - T_m) \times (T_\varepsilon - T_m), \tag{4.52}$$

  où nous avons introduit la fonction indicatrice $\mathbf{1}_{\mathbb{R}^+}$ de $\mathbb{R}^+$ à savoir

$$\mathbf{1}_{\mathbb{R}^+}(x) = \begin{cases} 1 & \text{si } x \geqslant 0, \\ 0 & \text{sinon.} \end{cases} \tag{4.53}$$

  Pour ce terme de pénalisation, nous constatons que la fonction n'est pas dérivable en $T = T_m$. La méthode de pénalisation est alors qualifiée de « non lisse ». Il en résulte alors de possibles difficultés au niveau de la résolution numérique [72].

- Une autre possibilité est de considérer une méthode de pénalisation quadratique. Dans ce cas, le terme de pénalisation s'écrit

$$\frac{1}{\varepsilon} P(T_\varepsilon) = \frac{1}{2\varepsilon} \mathbf{1}_{\mathbb{R}^+}(T_\varepsilon - T_m) \times (T_\varepsilon - T_m)^2. \tag{4.54}$$

  Contrairement à la méthode précédente, la fonction de pénalisation quadratique est dérivable à l'origine. Cependant, le problème peut être mal conditionné et poser lui aussi des difficultés numériques [72].

En pratique, les méthodes de pénalisation sont généralement appréciées pour plusieurs raisons. Elles sont tout d'abord assez simples d'implémentation et n'imposent aucune contrainte sur T_ε si ce n'est d'appartenir à l'ensemble G. De plus, ces méthodes ne font intervenir qu'une seule inconnue à savoir T_ε au contraire de l'approche duale que nous introduirons à la sous-section suivante. Le taux de fonte basale a_b qui intervient dans le problème (4.41) peut quant à lui être calculé une fois qu'une approximation de la solution T a été déterminée.

Si le problème (4.41) est dominé par la convection, il convient de tenir compte simultanément de la contrainte thermique et de la nécessité de stabiliser la solution numérique. Cela peut être réalisé en utilisant à la fois une méthode de pénalisation pour la contrainte thermique et une méthode de stabilisation pour réduire les oscillations indésirables dans la solution numérique. Pour combiner ces deux méthodes, nous pouvons simplement ajouter un terme de pénalisation et un terme de stabilisation comme nous l'avons fait aux relations (4.31) et (4.51). Cela peut se justifier par le fait que le terme de pénalisation correspond à une intégrale réalisée sur une partie de la frontière de Ω, alors que le terme de stabilisation introduit uniquement une somme d'intégrales évaluées sur l'intérieur des éléments.

4.5.3. Approche duale

L'idée de l'approche duale est de reformuler l'inéquation variationnelle de départ sous une nouvelle forme en introduisant de nouvelles variables qui jouent le rôle de multiplicateurs de Lagrange. Cela permet de travailler avec un nouveau problème qui est en général plus facile à résoudre. Comme nous l'avions déjà précisé précédemment, le terme de convection rend la formulation faible du problème thermique non symétrique, ce qui signifie que ce problème ne peut pas se réécrire sous la forme d'un problème de minimisation comme c'est le cas en élasticité linéaire. Toutefois, l'approche duale va nous permettre d'écrire le problème thermique contraint sous la forme d'un problème d'optimisation de type point de selle. Les idées de base sur lesquelles repose notre raisonnement sont notamment inspirées de Ekeland et Temam [67]. Notons également que dans cette sous-section, nous considérons uniquement des conditions de Dirichlet homogènes sur Γ_s $(G = H)$ afin de simplifier la présentation.

Pour établir les idées du formalisme dual, nous réécrivons tout d'abord la formulation variationnelle 4.2 sous la forme équivalente suivante :

Trouver $T \in K$ tel que

$$a(T,\theta)+c(\mathbf{v};T,\theta)-(s,\theta)-(q,\theta)_{\Gamma_b} \geqslant a(T,T)+c(\mathbf{v};T,T)-(s,T)-(q,T)_{\Gamma_b} \quad \forall \theta \in K. \tag{4.55}$$

Cette dernière expression indique qu'une inégalité variationnelle peut se présenter comme un problème d'optimisation contraint. En effet, le problème (4.55) est équivalent à

$$T = \arg\min_{\theta \in K} \left[a(T,\theta)+c(\mathbf{v};T,\theta)-(s,\theta)-(q,\theta)_{\Gamma_b}\right]. \tag{4.56}$$

Il est important de souligner que le problème de minimisation (4.56) ne peut pas être utilisé en pratique pour résoudre l'inégalité variationnelle, car la fonction objectif à minimiser dépend explicitement de la solution T. Bien que le problème (4.56) ait l'apparence d'un problème de minimisation, il faut dès lors rester conscient qu'il ne s'agit pas d'un problème d'optimisation classique. Toutefois, si la valeur de T est connue, ce problème d'optimisation a du sens. En fait, l'intérêt principal d'introduire le problème de minimisation (4.56) réside dans le fait qu'il va nous permettre d'établir une formulation variationnelle mixte pour résoudre l'inégalité variationnelle (4.55). Le problème (4.56) est contraint par la condition $T \in K$. Cependant, ce problème peut se réduire à un problème non contraint si nous introduisons la fonction Φ_K telle que

$$\Phi_K(\theta) = \begin{cases} 0 & \text{si } \theta \in K, \\ +\infty & \text{sinon.} \end{cases} \tag{4.57}$$

Dans ce cas, le problème contraint (4.56) s'écrit sous la forme du problème quasi non contraint

$$T = \arg\min_{\theta \in G} \left[a(T,\theta)+c(\mathbf{v};T,\theta)-(s,\theta)-(q,\theta)_{\Gamma_b}+\Phi_K(\theta)\right]. \tag{4.58}$$

Nous introduisons à présent le multiplicateur de Lagrange $\lambda(\mathbf{x})$ encore appelé variable duale. Le multiplicateur de Lagrange $\lambda(\mathbf{x})$ est supposé appartenir au cône convexe A défini par

$$A = \{\lambda \in \mathbb{H}^{1/2}(\Gamma_b) : \lambda \geqslant 0 \ \text{ sur } \Gamma_b\}, \tag{4.59}$$

où $\mathbb{H}^{1/2}(\Gamma_b)$ représente la restriction sur Γ_b des fonctions de $\mathbb{H}^1(\Omega)$. Une présentation plus détaillée de cet espace peut être trouvée dans [8, 48].

En fonction du multiplicateur de Lagrange, Φ_K peut s'écrire

$$\Phi_K(\theta) = \max_{\lambda \in A} \int_{\Gamma_b} (\theta - T_m)\, \lambda \, d\Gamma_b. \tag{4.60}$$

Cette expression de $\Phi_K(\theta)$ est bien en concordance avec la définition (4.57). En effet, le problème d'optimisation (4.60) admet les solutions $\lambda = 0$ si $\theta < T_m$ et $\lambda = +\infty$ si $\theta > T_m$. Si $T = T_m$, la valeur de λ est indéterminée et peut prendre a priori n'importe quelle valeur positive.

Nous introduisons à présent le lagrangien

$$\mathcal{L}(T;\theta,\lambda) = a(T,\theta) + c(\mathbf{v};T,\theta) - (s,\theta) - (q,\theta)_{\Gamma_b} + \int_{\Gamma_b} (\theta - T_m)\, \lambda \, d\Gamma_b. \tag{4.61}$$

La notation que nous avons adoptée pour le lagrangien a été choisie de manière à rappeler que le lagrangien dépend de la variable primale θ et de la variable duale λ tandis que la température T est un paramètre du lagrangien. L'introduction du lagrangien permet d'écrire le problème (4.58) sous la forme du problème de point de selle

$$\min_{\theta \in G} \max_{\lambda \in A} \mathcal{L}(T;\theta,\lambda) = \max_{\lambda \in A} \min_{\theta \in G} \mathcal{L}(T;\theta,\lambda), \tag{4.62}$$

ou encore sous forme équivalente

$$\mathcal{L}(T;T,\lambda) \leqslant \mathcal{L}(T;T,\zeta) \leqslant \mathcal{L}(T;\theta,\zeta) \quad \forall \theta \in G \,\text{et}\, \forall \lambda \in A. \tag{4.63}$$

Nous constatons que $\mathcal{L}(\theta,\lambda)$ est convexe par rapport à θ et concave par rapport à λ. De plus, comme les ensembles G et A sont également convexes, cela explique l'inversion des opérations min et max effectuée à la relation (4.62). Il en résulte donc une équivalence entre les problèmes primal et dual.

La solution (T,ζ) du problème (4.62) doit vérifier les conditions KKT à savoir

$$D_\theta \mathcal{L}(T;T,\zeta)[\delta\theta] = 0 \quad \forall \delta\theta \in G, \tag{4.64a}$$
$$T - T_m \leqslant 0, \tag{4.64b}$$
$$\zeta \geqslant 0, \tag{4.64c}$$
$$\zeta\,(T - T_m) = 0. \tag{4.64d}$$

La dérivée qui apparaît dans l'expression (4.64a) est une dérivée au sens de Gateaux du lagrangien [9]. Vu la convexité du problème que nous étudions, les conditions KKT constituent des conditions nécessaires et suffisantes pour être solution du problème (4.62). Les conditions KKT nous permettent à présent de donner une signification physique aux variables duales. Si nous comparons les contraintes (4.42a)–(4.42c) aux conditions (4.64b)–(4.64d), nous constatons une forte similarité entre celles-ci. Il est dès lors naturel d'interpréter le multiplicateur de Lagrange λ comme un taux de fonte basale. Nous noterons ainsi $\zeta = \rho L a_b$ où a_b est le

9. Lindenstrauss et Preiss [47] donnent la définition suivante de la dérivée de Gateaux : Une fonctionnelle f d'un espace de Banach X vers un autre espace de Banach Y est dite différentiable au sens de Gateaux en x_0 s'il existe un opérateur linéaire N tel que pour tout $u \in X$, on a $\lim_{t\to 0} \dfrac{f(x_0+tu)-f(x_0)}{t} = N[u]$. L'opérateur N est appelé la dérivée de Gateaux de f en x_0.

taux de fonte associé à la solution T du problème thermique contraint (4.41). Le taux de fonte a ainsi une signification particulière, car il correspond à la variable duale de la température et permet d'assurer la contrainte thermique $T \leqslant T_m$ sur Γ_b. La quantité $\rho L a_b$ constitue de fait un nouveau degré de liberté dans le système et joue le rôle d'un flux de chaleur externe variable dont la valeur est ajustée de manière à éviter un échauffement excessif du glacier.

D'un point de vue plus mathématique, le taux de fonte basale peut aussi être interprété comme le taux de variation marginal du lagrangien suite à une modification de la valeur de T_m qui peut résulter physiquement d'une modification de la pression dans le glacier. Pour s'en rendre compte, il suffit d'écrire

$$\frac{\partial \mathcal{L}}{\partial T_m}(T; \theta, \rho L a_b) = -\int_{\Gamma_b} \rho L a_b \, d\Gamma_b \,. \tag{4.65}$$

Cette dernière relation montre ainsi que la quantité $\rho L a_b$ est liée à la variation du lagrangien pour une modification de la contrainte sur la température.

Les conditions KKT (4.64a)–(4.64d) peuvent également se réécrire sous la forme d'une formulation variationnelle dont les inconnues sont la température dans le glacier et le taux de fonte basale. Pour cela, nous développons (4.64a) sous la forme

$$a(T, \theta) + c(\mathbf{v}; T, \theta) = (s, \theta) + (q, \theta)_{\Gamma_b} - \int_{\Gamma_b} \theta \, \rho L a_b \, d\Gamma_b \,. \tag{4.66}$$

Pour $\lambda \in A$, la relation (4.64b) peut aussi s'écrire

$$\int_{\Gamma_b} T \, \lambda \, d\Gamma_b \leqslant \int_{\Gamma_b} T_m \, \lambda \, d\Gamma_b \,. \tag{4.67}$$

En exploitant enfin la condition de complémentarité (4.64d), cette dernière relation devient

$$\int_{\Gamma_b} T \, (\lambda - \rho L a_b) \, d\Gamma_b \leqslant \int_{\Gamma_b} T_m \, (\lambda - \rho L a_b) \, d\Gamma_b \,. \tag{4.68}$$

Le problème thermique (4.41) peut finalement s'écrire sous la forme variationnelle suivante :

Formulation variationnelle 4.3. Trouver $T \in G$ et $a_b \in A$ tels que

$$\begin{cases} a(T, \theta) + c(\mathbf{v}; T, \theta) = (s, \theta) + (q, \theta)_{\Gamma_b} - \displaystyle\int_{\Gamma_b} \theta \, \rho L a_b \, d\Gamma_b, \\ \displaystyle\int_{\Gamma_b} T \, (\lambda - \rho L a_b) \, d\Gamma_b \leqslant \int_{\Gamma_b} T_m \, (\lambda - \rho L a_b) \, d\Gamma_b, \end{cases}$$

pour tout couple $(\theta, \lambda) \in G \times A$.

Cette formulation variationnelle peut servir de base à une discrétisation par éléments finis en considérant des approximations de la solution (T, a_b) et des fonctions test (θ, λ) sur des espaces fonctionnels de dimension finie. Le lecteur intéressé pourra trouver dans Kikuchi et Oden [43] quelques compléments d'informations théoriques sur l'approximation des inégalités variationnelles par la méthode des éléments finis.

Même si les méthodes de pénalisation sont généralement privilégiées pour résoudre le problème thermique en glaciologie, la méthode lagrangienne n'est pas dépourvue d'intérêts. Un des inconvénients majeurs de la méthode de pénalisation que nous avons présentée réside dans le choix du paramètre de pénalisation. Un paramètre trop élevé mène à une solution

qui viole la contrainte thermique de manière importante, alors que le choix d'un paramètre trop petit peut affecter le conditionnement du problème et la convergence de l'algorithme. Trouver le bon paramètre n'est donc pas toujours évident et doit parfois être réalisé sur la base d'essais-erreurs. Dans la méthode lagrangienne, le rôle du multiplicateur de Lagrange est similaire à celui du paramètre de pénalisation dans la mesure où celui-ci est destiné à empêcher une violation de la contrainte. Alors que la valeur du paramètre de pénalisation est fixée par l'utilisateur, la valeur des multiplicateurs de Lagrange est ajustée par la méthode en elle-même. Cela permet ainsi d'obtenir une méthode qui est plus fiable. Au niveau des désavantages, la méthode des multiplicateurs de Lagrange demande de travailler avec un nombre plus important d'inconnues et l'implémentation de la méthode est en général plus compliquée.

Alors que l'étude des problèmes contraints pour le contact mécanique a suscité de nombreuses attentions de la part de la communauté scientifique et que de nombreux algorithmes de résolution ont été développés [71], l'étude des problèmes thermiques contraints, notamment en glaciologie, est relativement peu documentée et les algorithmes de résolution sont généralement limités. Une analyse plus détaillée des méthodes duales pour le problème thermique en glaciologie offrirait sans nul doute de nouvelles opportunités dans le domaine de la simulation numérique.

Pour terminer ce chapitre, il est intéressant de faire un parallèle entre l'approche duale développée au chapitre 3 pour le problème mécanique et celle présentée dans ce chapitre pour le problème thermique. Ces deux problèmes sont caractérisés par une inconnue primaire (variable primale) qui représente respectivement le champ de vitesse et de température dans le glacier. Ces deux inconnues sont toutes les deux soumises à une contrainte qui restreint l'ensemble de fonctions auquel elles doivent appartenir. Ces contraintes sont imposées par la nature physique de la glace à savoir sa propriété d'incompressibilité et sa possibilité de changer d'état à son point de fusion. Nous avons pu montrer également que les problèmes mécanique et thermique pouvaient s'écrire sous la forme d'un problème de minimisation sous contrainte. Afin de s'affranchir de cette contrainte, nous avons introduit dans chacun des deux problèmes un nouveau degré de liberté par l'intermédiaire d'un multiplicateur de Lagrange (variable duale). Outre leur intérêt mathématique, une signification physique a pu être attribuée à ces variables duales à savoir la pression dans le fluide pour le problème mécanique et le taux de fonte basale pour le problème thermique. L'introduction de multiplicateurs de Lagrange permet ainsi de développer une formulation variationnelle mixte qui peut être utilisée pour une résolution numérique par éléments finis. Les similitudes que nous venons de montrer entre les deux problèmes ne sont pas fortuites et sont liées à la ressemblance au niveau de la structure mathématique entre ces deux problèmes. L'approche présentée dans ces deux chapitres est donc assez générale et peut s'étendre à d'autres types de problèmes physiques.

Pour finir cette comparaison entre les chapitres 3 et 4, il convient de se rappeler que pour le problème mécanique, l'unicité de la solution $(\mathbf{v}, p)$ dépendait de la vérification de la condition de Babuska-Brezzi qui revenait à assurer la surjectivité de l'opérateur divergence. Une telle condition se rencontre aussi pour le problème thermique et la contrainte $T \leqslant T_m$ sur Γ_b. Néanmoins pour cette contrainte, l'opérateur qui agit sur la variable T est l'opérateur identité et la surjectivité de celui-ci est nécessairement garantie. Il s'ensuit que la condition de Babuska-Brezzi est toujours satisfaite et donc que la formulation mixte admet toujours une unique paire de solutions.

5. Méthodes numériques pour les problèmes multiphysiques non linéaires

Les chapitres précédents étaient orientés vers la modélisation physique et la compréhension théorique des méthodes numériques pour résoudre les problèmes mécanique et thermique. Ce cinquième chapitre a pour but d'assurer la liaison entre les chapitres antérieurs et les prochains chapitres qui seront dédiés à l'application concrète de méthodes numériques multiphysiques à l'étude d'un modèle de glacier alpin. Notre objectif est de présenter quelques méthodes numériques pour la résolution de systèmes algébriques non linéaires issus par exemple de la discrétisation de problèmes physiques couplés tels que le problème thermomécanique qui décrit l'écoulement des glaciers. Nous débuterons à la section 5.1 par une introduction à la notion de méthodes de résolution faiblement et fortement couplées. La section 5.2 présentera ensuite quelques méthodes numériques pour la résolution d'équations non linéaires couplées dont les méthodes de Gauss-Seidel, de Picard et de Newton. Enfin, la section 5.3 terminera ce chapitre par une discussion qui permettra de savoir quand des méthodes faiblement couplées telles que les méthodes de Jacobi et de Gauss-Seidel sont peu adéquates pour résoudre un problème couplé.

Les méthodes de résolution que nous présenterons dans ce chapitre sont pour la plupart des méthodes traditionnelles pour la résolution de problèmes multiphysiques. Actuellement, le domaine de la simulation multiphysique évolue vers l'utilisation de méthodes entièrement couplées telles que la méthode *Jacobian-free Newton-Krylov* dont nous donnerons un aperçu théorique à la sous-section 5.2.5. Ces nouvelles méthodes numériques vont à l'encontre de l'approche traditionnelle en multiphysique qui repose sur la décomposition d'un problème multiphysique en un ensemble de sous-problèmes qui sont résolus individuellement (*operator splitting strategy*). Cette approche bien qu'autorisant l'utilisation de codes de calcul préexistants peut mener à des problèmes de stabilité numérique particulièrement pour la résolution de problèmes évolutifs [42]. C'est pourquoi de nouvelles approches basées sur des méthodes fortement couplées préconditionnées suscitent un intérêt croissant. Même si ces méthodes ne seront pas abordées directement dans le cadre de ce travail, les méthodes numériques que nous présenterons et implémenterons par la suite permettront de développer une meilleure compréhension des fondements sur lesquels sont construites ces méthodes plus avancées.

5.1. Introduction aux méthodes de résolution faiblement couplées et fortement couplées

Dans ce chapitre, nous nous intéressons à la résolution du système d'équations non linéaires suivant :

$$\mathrm{F}(\mathbf{x}^*) = \begin{bmatrix} \mathrm{F}_1(x_1^*, x_2^*) \\ \mathrm{F}_2(x_1^*, x_2^*) \end{bmatrix} = \mathbf{0}. \tag{5.1}$$

Par simplicité, nous envisageons un système d'équations à deux inconnues. Les méthodes présentées dans ce chapitre se généralisent sans peine à des problèmes de dimension supérieure.

Le système (5.1) peut être interprété comme un problème thermomécanique qui résulte du couplage d'une équation mécanique F_1 pour le déplacement x_1 et d'une équation F_2 pour la température x_2. Plus généralement, ce système d'équations peut résulter de la discrétisation par la méthode de Galerkin d'un ensemble d'équations variationnelles couplées. Les méthodes de résolution du système (5.1) sont des méthodes itératives qui construisent une suite d'approximations $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)})$ de la solution exacte. La quantité $F(\mathbf{x}^{(k)})$ sera appelée le résidu du système. Les fonctions F_1 et F_2 sont alors nommées les composantes du résidu. Idéalement, le résidu doit être suffisamment petit pour que l'itéré $\mathbf{x}^{(k)}$ constitue une estimation précise de la solution exacte.

Comme nous l'avons déjà laissé entendre précédemment, deux stratégies principales sont généralement adoptées pour résoudre un système d'équations non linéaires couplées similaire au problème (5.1). La première approche, qui est celle traditionnellement employée, consiste à voir un problème multiphysique comme l'assemblage de différentes composantes individuelles telles que, par exemple, une composante mécanique et une composante thermique. Dans ce cas, les algorithmes de résolution sont utilisés de manière à préserver l'intégrité des problèmes individuels. Le couplage entre les différents problèmes monophysiques est effectué de manière indirecte durant l'algorithme de résolution. Pour cette première approche, la méthode de résolution présente un couplage qualifié de faible ou de lâche. La seconde approche propose une vision opposée des problèmes multiphysiques. Ceux-ci sont perçus comme intrinsèquement couplés et les problèmes monophysiques correspondent à des cas limites. Les algorithmes de résolution basés sur cette approche tentent de synchroniser, dans la mesure du possible, les variables d'état du système comme le déplacement et la température. Le couplage est alors dit fort ou serré. Keyes *et al.* [42] suggèrent d'adopter prioritairement cette seconde approche tant qu'un examen de la possibilité de découpler le problème n'a pas été mené. Dans la suite de cette section, nous présenterons différents algorithmes de résolution qui peuvent être classés dans un de ces deux types d'approches.

5.2. Résolution numérique de systèmes non linéaires

5.2.1. Méthodes itératives sur les sous-problèmes individuels

Une idée naturelle pour résoudre un système d'équations couplées est de résoudre séparément chacun des problèmes monophysiques à chaque itération. Par exemple, l'équation non linéaire F_1 sera résolue pour la variable x_1 en considérant la variable x_2 comme un paramètre fixe, tandis que l'équation F_2 sera résolue pour la variable x_2 en maintenant la variable x_1 constante.

Il existe différentes façons d'implémenter cette idée en pratique. Une méthode habituellement utilisée est la méthode itérative de Gauss-Seidel (voir algorithme 1). Dans cette méthode, si le problème mécanique est résolu par exemple en premier, alors la nouvelle valeur du champ de déplacement est directement utilisée pour résoudre le problème thermique. La convergence de cette méthode peut être améliorée en cas d'utilisation d'une méthode de relaxation. Pour la méthode itérative de Jacobi, seule la valeur des inconnues à l'itération précédente est utilisée, ce qui autorise une implémentation en parallèle de cette méthode beaucoup plus aisée.

Algorithme 1 Algorithme de Gauss-Seidel

1: Initialiser $(x_1^{(0)}, x_2^{(0)})$
2: **pour** $k = 1, 2, \ldots$ (jusqu'à convergence) **faire**
3: Résoudre $\mathrm{F}_1(x_1, x_2^{(k-1)}) = 0$; Mettre $x_1^{(k)} = x_1$
4: Résoudre $\mathrm{F}_2(x_1^{(k)}, x_2) = 0$; Mettre $x_2^{(k)} = x_2$
5: **fin pour**

Lors d'une itération de la méthode de Gauss-Seidel ou de Jacobi, il est parfois nécessaire de résoudre un système non linéaire pour chacun des systèmes monophysiques. Cela peut être réalisé au moyen de méthodes itératives comme la méthode de Picard ou de Newton que nous introduirons par la suite. Notons que pour les méthodes de Gauss-Seidel et de Jacobi, le nombre d'opérations peut devenir rapidement conséquent si chacun des sous-problèmes est résolu précisément. En pratique, il est courant de se limiter uniquement à quelques itérations lors de la résolution de chacun de ces sous-problèmes.

L'avantage des méthodes présentées dans cette sous-section est sans nul doute le fait qu'elles nécessitent peu d'efforts d'implémentation puisque qu'elles autorisent l'utilisation de solveurs développés séparément pour chacun des problèmes monophysiques. Ces méthodes présentent ainsi une grande modularité. De plus, celles-ci permettent de résoudre des systèmes de moindre dimension vu qu'elles travaillent avec chaque modèle individuellement. Pour utiliser correctement ces méthodes, il suffit que chacun des solveurs individuels rende uniquement la solution du problème multiphysique. D'autres informations telles que la valeur du résidu ne sont pas nécessaires, ce qui peut constituer un attrait supplémentaire de ces méthodes d'autant plus que la valeur des résidus individuels n'est pas toujours disponible.

Malgré les attraits des méthodes présentées, la pertinence de celles-ci est fort dépendante du problème multiphysique traité. Ces méthodes peuvent se révéler inefficaces en cas de couplage fort entre les différentes composantes du système comme nous le verrons à la section 5.3. Dans ce cas, il est plus judicieux de traiter toutes les variables du problème simultanément et donc d'utiliser des méthodes plus fortement couplées.

5.2.2. Méthode du point fixe : méthode de Picard

La méthode de Picard est une des méthodes les plus simples à implémenter et également l'une des plus robustes. Il s'agit d'un cas particulier d'implémentation de la méthode du point fixe. Une autre méthode qui entre dans cette catégorie est la méthode de Newton que nous aborderons à la sous-section suivante. Les résultats théoriques généraux que nous présentons pour la méthode du point fixe sont ainsi applicables pour la méthode de Newton. Notons que la méthode de Picard peut être considérée comme une méthode fortement couplée dans la mesure où les différentes variables du problème sont traitées simultanément à chaque itération.

Dans la méthode de Picard, nous ne cherchons pas à résoudre directement le système (5.1), mais plutôt le système équivalent

$$\mathrm{G}(\mathbf{x}^*) = \mathbf{x}^*. \tag{5.2}$$

La relation (5.2) montre que la solution $\mathbf{x}^*$ est un point fixe de G. L'application G n'a pas une unique expression. Un choix possible est d'écrire G sous la forme

$$\mathrm{G}(\mathbf{x}) = \mathbf{x} - \alpha \mathrm{F}(\mathbf{x}), \quad \alpha > 0. \tag{5.3}$$

La méthode itérative de Picard consiste alors à déterminer $\mathbf{x}^*$ en effectuant l'itération de point fixe

$$\mathbf{x}^{(k+1)} = \mathrm{G}(\mathbf{x}^{(k)}). \tag{5.4}$$

Une étude théorique de la méthode (5.4) peut être réalisée plus spécifiquement si G est une application contractante. Pour définir une telle application, nous adoptons la définition suivante donnée par Quarteroni *et al.* [57] :

Définition 5.1. Une application $\mathrm{G} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ est dite contractante sur l'ensemble $D_0 \subset D$ s'il existe une constante $\eta < 1$ telle que $\|\mathrm{G}(\mathbf{x}) - \mathrm{G}(\mathbf{y})\| \leqslant \eta\|\mathbf{x} - \mathbf{y}\|$ pour tout $\mathbf{x}$, $\mathbf{y}$ dans D_0 où $\|\cdot\|$ représente une norme vectorielle.

Une application contractante est ainsi une application qui rapproche les images vu que la distance entre $\mathrm{G}(\mathbf{x})$ et $\mathrm{G}(\mathbf{y})$ est plus faible que la distance entre $\mathbf{x}$ et $\mathbf{y}$. L'existence et l'unicité d'un point fixe de G ainsi que la convergence de la méthode itérative (5.4) sont assurées par le théorème du point fixe de Banach dont nous reprenons l'énoncé ci-dessous [29, 57].

Théorème 5.1 (Théorème du point fixe de Banach). *Soit* $\mathrm{G} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ *une application contractante sur un ensemble fermé* $D_0 \subset D$ *telle que* $\mathrm{G}(\mathbf{x}) \subset D_0$ *pour tout* $\mathbf{x} \in D_0$. *Alors* G *a un unique point fixe* $\mathbf{x}^*$ *dans* D_0 *et la séquence* $\{\mathbf{x}^{(k)}\}$ *définie par* $\mathbf{x}^{(k+1)} = \mathrm{G}(\mathbf{x}^{(k)})$ *converge vers* $\mathbf{x}^*$.

Un des désavantages de la méthode de Picard est sa faible vitesse de convergence. En effet, sous les hypothèses du théorème 5.1, il peut être montré que la vitesse de convergence est linéaire [74] à savoir

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| \leqslant \eta\|\mathbf{x}^{(k)} - \mathbf{x}^*\|, \quad \eta < 1. \tag{5.5}$$

5.2.3. Méthode du point fixe : méthode de Newton

La méthode de Newton est un algorithme pour la résolution de systèmes non linéaires. L'idée de base est la suivante. L'algorithme est initialisé avec une approximation de la solution exacte qui est si possible pas trop éloignée de cette solution. Au lieu de résoudre le système non linéaire, c'est le système linéarisé autour de l'approximation qui est résolu. Ceci permet d'obtenir une nouvelle approximation de la solution du problème. En réitérant la procédure précédente pour la nouvelle approximation, nous obtenons une méthode itérative susceptible de converger vers la solution du problème.

Pour comprendre comment cette méthode s'applique, nous repartons du système (5.1). Ce problème peut être reformulé différemment. Pour cela, nous considérons une approximation initiale $(x_1^{(0)}, x_2^{(0)})$ de la solution (x_1^*, x_2^*). L'écart $\delta\mathbf{x} = (\delta x_1, \delta x_2)$ à la solution exacte doit satisfaire

$$\begin{bmatrix} \mathrm{F}_1(x_1^{(0)} + \delta x_1, x_2^{(0)} + \delta x_2) \\ \mathrm{F}_2(x_1^{(0)} + \delta x_1, x_2^{(0)} + \delta x_2) \end{bmatrix} = \mathbf{0}. \tag{5.6}$$

Les fonctions F_1 et F_2 peuvent être développées en série de Taylor autour de l'approximation $(x_1^{(0)}, x_2^{(0)})$. En limitant ce développement à l'ordre un, nous obtenons le système

$$\begin{bmatrix} \mathrm{F}_1(x_1^{(0)}, x_2^{(0)}) \\ \mathrm{F}_2(x_1^{(0)}, x_2^{(0)}) \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathrm{F}_1}{\partial x_1}(x_1^{(0)}, x_2^{(0)}) & \frac{\partial \mathrm{F}_1}{\partial x_2}(x_1^{(0)}, x_2^{(0)}) \\ \frac{\partial \mathrm{F}_2}{\partial x_1}(x_1^{(0)}, x_2^{(0)}) & \frac{\partial \mathrm{F}_2}{\partial x_2}(x_1^{(0)}, x_2^{(0)}) \end{bmatrix} \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix} = \mathbf{0}. \tag{5.7}$$

Nous introduisons à présent la matrice jacobienne J du système définie par

$$J(\mathbf{x}) = \begin{bmatrix} \frac{\partial \mathrm{F}_1}{\partial x_1}(\mathbf{x}) & \frac{\partial \mathrm{F}_1}{\partial x_2}(\mathbf{x}) \\ \frac{\partial \mathrm{F}_2}{\partial x_1}(\mathbf{x}) & \frac{\partial \mathrm{F}_2}{\partial x_2}(\mathbf{x}) \end{bmatrix}. \tag{5.8}$$

Le système (5.7) peut alors se réécrire

$$J(\mathbf{x}^{(0)}) \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix} = - \begin{bmatrix} \mathrm{F}_1(x_1^{(0)}, x_2^{(0)}) \\ \mathrm{F}_2(x_1^{(0)}, x_2^{(0)}) \end{bmatrix}. \tag{5.9}$$

La résolution du système (5.9) permet de déterminer les termes de correction δx_1 et δx_2. Ceci nous donne alors une nouvelle approximation ($x_1^{(1)} = x_1^{(0)} + \delta x_1, x_2^{(1)} = x_2^{(0)} + \delta x_2$) de la solution du problème. Les termes de correction ne permettent pas de retrouver la solution exacte, car les équations ont été linéarisées. Toutefois, si la méthode a correctement fonctionné, la nouvelle approximation est plus proche de la solution exacte que l'approximation précédente. Celle-ci peut être utilisée pour déterminer de nouveaux termes de correction. De cette façon, nous obtenons une méthode itérative pour résoudre le problème (5.6).

La procédure générale de la méthode de Newton est finalement reprise à l'algorithme 2.

Algorithme 2 Algorithme de Newton

1: Initialiser $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)})$
2: **pour** $k = 0, 1, 2, \ldots$ (jusqu'à convergence) **faire**
3: Résoudre $J(\mathbf{x}^{(k)})\delta\mathbf{x}^{(k)} = -\mathrm{F}(\mathbf{x}^{(k)})$
4: Mettre $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta\mathbf{x}^{(k)}$
5: **fin pour**

L'algorithme 2 montre en fait que la méthode de Newton peut être interprétée comme une méthode de point fixe pour laquelle nous avons

$$\begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} - J^{-1}(\mathbf{x}^{(k)})\mathrm{F}(\mathbf{x}^{(k)}) \tag{5.10}$$

Le théorème 5.2 [57] fournit un résultat théorique concernant la convergence et la vitesse de convergence de la méthode de Newton.

Théorème 5.2. *Soit* $\mathrm{F} : \mathbb{R}^n \to \mathbb{R}^n$ *une fonction de classe* C^1 *sur un ouvert convexe* D *de* $\mathbb{R}^n$ *qui contient* $\mathbf{x}^*$ *tel que* $\mathrm{F}(\mathbf{x}^*) = \mathbf{0}$. *Supposons que* $J^{-1}(\mathbf{x}^*)$ *existe et qu'il existe des constantes positives* R, C *et* L *telles que* $\|J^{-1}(\mathbf{x}^*)\| \leqslant C$ *et*

$$\|J(\mathbf{x}) - J(\mathbf{y})\| \leqslant L\,\|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in B(\mathbf{x}^*; R),$$

où on a noté par le même symbole $\|\cdot\|$ *une norme vectorielle et une norme matricielle consistante et par* $B(\mathbf{x}^*; R)$ *la boule ouverte de rayon* R *centrée en* $\mathbf{x}^*$ *à savoir* $B(\mathbf{x}^*; R) = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{x}^*\| < R\}$. *Il existe alors* $r > 0$ *tel que, pour tout* $\mathbf{x}^{(0)} \in B(\mathbf{x}^*; r)$, *la suite* (5.10) *est définie de façon unique et converge vers* $\mathbf{x}^*$ *avec*

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| \leqslant CL\,\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2.$$

Au niveau des avantages, la vitesse de convergence de la méthode de Newton est généralement rapide. Le théorème 5.2 met en évidence la possibilité d'obtenir un taux de convergence quadratique sous certaines hypothèses. Parmi les désavantages de cette méthode, celle-ci peut ne pas converger vers le résultat attendu si l'itéré initial n'est pas correctement choisi. En particulier, le théorème 5.2 montre que la convergence n'est assurée que dans une boule ouverte de rayon r centrée en $\mathbf{x}^*$. En général, le rayon de cette boule ouverte sera faible et la méthode convergera pour un itéré initial suffisamment proche de la solution exacte. Pour diminuer la sensibilité de la méthode de Newton par rapport au choix de l'approximation initiale, une méthode plus robuste peut être utilisée pour les premières itérations ou encore une méthode hybride comme expliqué dans le paragraphe ci-dessous. Enfin, la méthode de Newton nécessite de calculer la matrice jacobienne du système, ce qui peut être coûteux en temps de calcul ou bien impossible si le système n'est pas dérivable aux points parcourus durant les itérations. Différentes variantes de la méthode de Newton ont ainsi été proposées pour diminuer les coûts de calcul dont les méthodes de Quasi-Newton sur lesquelles nous reviendrons à la sous-section suivante. Quelques alternatives à la méthode de Newton sont notamment discutées par Ortega et Rheinboldt [53] et Quarteroni *et al.* [57].

Notons que pour augmenter la robustesse de la méthode de Newton, celle-ci peut être combinée avec une autre méthode comme la méthode de Picard. Cela donne lieu à des méthodes de résolution hybrides qui combinent à la fois robustesse et vitesse de convergence élevée. Un moyen simple pour combiner l'itération de Picard (5.4) avec l'itération de Newton (5.10) est d'introduire un coefficient $\gamma \in [0, 1]$ et de définir la méthode hybride suivante :

$$\begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = (1-\gamma)\, G(\mathbf{x}^{(k)}) + \gamma \left(\begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \end{bmatrix} - J^{-1}(\mathbf{x}^{(k)}) \mathrm{F}(\mathbf{x}^{(k)}) \right). \tag{5.11}$$

Pour finir cette sous-section, nous indiquons que la méthode de Newton peut aussi bien être appliquée pour résoudre chacun des problèmes monophysiques individuels lors d'une résolution par une méthode faiblement couplée ou être employée pour résoudre l'ensemble du système couplé. Dans ce dernier cas, la méthode de Newton peut être interprétée comme une méthode fortement couplée. Cette dernière remarque est également valable pour les méthodes de Picard et de Quasi-Newton que nous présentons dans cette section.

5.2.4. Méthodes de Quasi-Newton

Plutôt que de calculer la matrice jacobienne à chaque itération, il peut être préférable de l'approximer par une matrice qui serait mise à jour à chaque itération. Les méthodes dites de Quasi-Newton généralisent l'itération de Newton (5.10) en introduisant la mise à jour

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - [A^{(k)}]^{-1} \mathrm{F}(\mathbf{x}^{(k)}). \tag{5.12}$$

La méthode de Newton peut être vue comme un cas particulier de la méthode de Quasi-Newton pour lequel $A^{(k)} = J^{(k)}$. En général, la matrice $A^{(k)}$ sera choisie de manière à constituer une approximation de la matrice jacobienne du système. Le choix pour une telle matrice n'est bien entendu pas unique. Habituellement, les matrices $A^{(k)}$ sont choisies de manière à satisfaire l'équation sécante [1] à savoir

$$A^{(k)} \mathbf{s}^{(k-1)} = \mathbf{y}^{(k-1)}, \tag{5.13}$$

1. Certains auteurs réservent l'appellation de Quasi-Newton uniquement aux méthodes du type (5.12) pour lesquelles $A^{(k)}$ satisfait l'équation sécante (voir Martínez [50] à ce sujet).

avec

$$\mathbf{s}^{(k-1)} = \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}, \tag{5.14}$$

$$\mathbf{y}^{(k-1)} = \mathrm{F}(\mathbf{x}^{(k)}) - \mathrm{F}(\mathbf{x}^{(k-1)}). \tag{5.15}$$

L'équation sécante généralise dans $\mathbb{R}^n$ la méthode de la sécante utilisée pour trouver les racines d'une fonction dans $\mathbb{R}$.

En dimension supérieure ou égale à 2, l'équation (5.13) admet en général une infinité de solutions pour $A^{(k)}$. Un choix naturel est de choisir la matrice $A^{(k)}$ qui ne soit pas trop éloignée de la valeur de la matrice $A^{(k-1)}$ utilisée à l'itération précédente. Ainsi, $A^{(k)}$ sera choisie en général de manière à minimiser la distance à $A^{(k-1)}$ en terme d'une certaine norme matricielle.

Le théorème 5.3 [16] donne un moyen de calculer $A^{(k)}$ si la distance entre cette matrice et $A^{(k-1)}$ est mesurée à partir de la norme de Frobenius $\|\cdot\|_F$ définie par

$$\|M\|_F = \left(\sum_{i,j=1}^{n} |M_{ij}|^2 \right)^{1/2}. \tag{5.16}$$

Théorème 5.3. *Soit $A^{(k-1)} \in \mathbb{R}^{n\times n}$, $\mathbf{s}^{(k-1)}$, $\mathbf{y}^{(k-1)} \in \mathbb{R}^n$ et Q l'ensemble des matrices carrées d'ordre n vérifiant l'équation sécante, c'est-à-dire $Q = \{M \in \mathbb{R}^{n\times n} : M\mathbf{s}^{(k-1)} = \mathbf{y}^{(k-1)}\}$. Alors, l'unique solution de*

$$\min_{A\in Q} \|A - A^{(k-1)}\|_F \tag{5.17}$$

est donnée par

$$A^{(k)} = A^{(k-1)} + \frac{(\mathbf{y}^{(k-1)} - A^{(k-1)}\mathbf{s}^{(k-1)})(\mathbf{s}^{(k-1)})^T}{(\mathbf{s}^{(k-1)})^T\mathbf{s}^{(k-1)}}. \tag{5.18}$$

La mise à jour (5.18) est à la base de la méthode de Broyden dont le pseudo-code est donné à l'algorithme 3. Cette méthode s'applique tout particulièrement lorsque la matrice jacobienne du système ne possède aucune propriété particulière comme par exemple être symétrique. D'autres exemples de méthodes de Quasi-Newton sont présentés par Dennis et Schnabel [16] ainsi que par Martínez [50].

Algorithme 3 Algorithme de Broyden

1: Initialiser $\mathbf{x}^{(0)}$ et $A^{(0)}$;
   Mettre $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - [A^{(0)}]^{-1}\mathrm{F}(\mathbf{x}^{(0)})$;
   $\mathbf{s}^{(0)} = \mathbf{x}^{(1)} - \mathbf{x}^{(0)}$;
   $\mathbf{y}^{(0)} = \mathrm{F}(\mathbf{x}^{(1)}) - \mathrm{F}(\mathbf{x}^{(0)})$
2: **pour** $k = 1, 2, \ldots$ (jusqu'à convergence) **faire**
3: Mettre

$$A^{(k)} = A^{(k-1)} + \frac{(\mathbf{y}^{(k-1)} - A^{(k-1)}\mathbf{s}^{(k-1)})(\mathbf{s}^{(k-1)})^T}{(\mathbf{s}^{(k-1)})^T\mathbf{s}^{(k-1)}}$$

4: Résoudre $A^{(k)}\mathbf{s}^{(k)} = -\mathrm{F}(\mathbf{x}^{(k)})$
5: Mettre $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{s}^{(k)}$ et $\mathbf{y}^{(k)} = \mathrm{F}(\mathbf{x}^{(k+1)}) - \mathrm{F}(\mathbf{x}^{(k)})$
6: **fin pour**

Au niveau de ses avantages, la méthode de Broyden permet d'éviter le calcul explicite de la matrice jacobienne du système tout en gardant une vitesse de convergence élevée. Sous certaines hypothèses [11, 57], la méthode de Broyden présente localement une convergence superlinéaire à savoir

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\| = o(\|\mathbf{x}^{(k)} - \mathbf{x}^*\|). \tag{5.19}$$

Il est intéressant de souligner que la méthode de Broyden peut présenter une convergence superlinéaire même si la matrice $A^{(k)}$ ne converge pas vers $J(\mathbf{x}^*)$.

Une des difficultés des méthodes de Quasi-Newton peut résider dans l'initialisation de l'algorithme et plus particulièrement dans le choix de la matrice $A^{(0)}$. Une possibilité pourrait être d'initialiser $A^{(0)}$ à $J(\mathbf{x}^{(0)})$ si la matrice jacobienne peut être calculée et puis d'utiliser la méthode de Quasi-Newton pour limiter le nombre d'évaluations de la matrice jacobienne.

Notons enfin que l'algorithme 3 tel qu'il est présenté n'est pas nécessairement optimal d'un point de vue temps de calcul. Nous en présenterons une version alternative à la section 6.5.4 qui s'adapte mieux à la résolution numérique de systèmes creux.

5.2.5. Méthode Jacobian-free Newton-Krylov

Une autre alternative importante à la méthode traditionnelle de Newton est donnée par les méthodes de Newton-Krylov. Ces méthodes combinent deux types de méthodes numériques à savoir la méthode de Newton qui est utilisée pour traiter les non-linéarités du problème et les méthodes de Krylov qui sont employées pour résoudre itérativement le système linéaire obtenu à chaque itération de Newton. De cette façon, le coût de calcul lié à la détermination de $\delta\mathbf{x}^{(k)}$ peut être réduit. Parmi les méthodes de Newton-Krylov, nous avons décidé de présenter la méthode *Jacobian-free Newton-Krylov* (JFNK) [45]. Cette méthode a suscité ces dernières années un intérêt croissant de la part de la communauté scientifique comme moyen d'implémenter des solveurs non linéaires efficaces pour la simulation des écoulements glaciaires [40, 46]. L'idée de cette méthode repose sur le fait que les méthodes de Krylov utilisent uniquement des opérations de type matrice fois vecteur. Dans la méthode JFNK, le produit de la matrice jacobienne avec un vecteur sera approché à l'aide d'une différence finie, ce qui ne nécessitera donc plus le calcul de la matrice jacobienne et son stockage qui peuvent tous les deux se révéler problématiques.

Après cette brève introduction à la méthode JFNK, regardons à présent plus concrètement comment celle-ci fonctionne. Pour cela, nous envisageons de nouveau le problème (5.1) qui peut être résolu avec la méthode de Newton à savoir

$$J^{(k)}\delta\mathbf{x}^{(k)} = -\mathrm{F}(\mathbf{x}^{(k)}), \tag{5.20a}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta\mathbf{x}^{(k)}. \tag{5.20b}$$

Pour une itération k donnée, il est nécessaire de résoudre le système linéaire (5.20a). Cela peut être réalisé au moyen d'une méthode de Krylov dont nous allons présenter brièvement les idées de base afin de mieux en comprendre l'intérêt. Pour plus de détails sur les méthodes de Krylov, nous conseillons entre autres la lecture de Kelley [41] et de van der Vorst [69]. Les méthodes de Krylov sont des méthodes itératives particulièrement adaptées pour les systèmes creux tels que ceux obtenus en général lors d'une discrétisation par une méthode éléments finis.

Pour résoudre le système (5.20a), nous allons construire des sous-espaces vectoriels appelés sous-espaces de Krylov dans lesquels une approximation de la solution $\delta\mathbf{x}^{(k)}$ est recherchée. Nous notons par $\delta\mathbf{x}_0^{(k)}$ une première approximation de cette solution. Dans le cas de la méthode de Newton, il est d'usage de prendre $\delta\mathbf{x}_0^{(k)} = \mathbf{0}$. Ceci s'explique par le fait qu'une correction supposée faible de l'itéré courant est recherchée et que celle-ci doit tendre asymptotiquement vers zéro si la méthode de Newton converge. Le vecteur résidu $\mathbf{r}^{(0)}$ associé à $\delta\mathbf{x}_0^{(k)}$ est donné par

$$\mathbf{r}^{(0)} = J(\mathbf{x}^{(k)})\delta\mathbf{x}_0^{(k)} + \mathrm{F}(\mathbf{x}^{(k)}). \tag{5.21}$$

La définition des sous-espaces de Krylov construits sur $J^{(k)}$ et $\delta\mathbf{x}_0^{(k)}$ est donnée ci-dessous.

Définition 5.2. On appelle espace de Krylov d'ordre p, noté $\mathcal{K}_p$, l'espace vectoriel généré par $\mathbf{r}^{(0)}$ et ses $p-1$ produits itérés par $J^{(k)}$, c'est-à-dire

$$\text{Vect} = \{\mathbf{r}^{(0)}, J^{(k)}\mathbf{r}^{(0)}, [J^{(k)}]^2\mathbf{r}^{(0)}, \ldots, [J^{(k)}]^{p-1}\mathbf{r}^{(0)}\}. \tag{5.22}$$

Les sous-espaces de Krylov forment une famille croissante de sous-espaces vectoriels, c'est-à-dire

$$\mathcal{K}_1 \subseteq \mathcal{K}_2 \subseteq \mathcal{K}_3 \subseteq \ldots \tag{5.23}$$

Les sous-espaces de Krylov sont de dimension finie. La dimension maximale de ces sous-espaces pour $\mathbf{r}^{(0)}$ donné est notée $p_{\max}$. Il peut être montré que la suite des espaces de Krylov est strictement croissante de 1 à $p_{\max}$ puis stagne à partir de $p = p_{\max}$ [60]. Il découle de cette affirmation que pour $p \leqslant p_{\max}$, les vecteurs de l'ensemble $\{\mathbf{r}^{(0)}, J^{(k)}\mathbf{r}^{(0)}, \ldots, [J^{(k)}]^{p-1}\mathbf{r}^{(0)}\}$ sont linéairement indépendants et constituent une base de $\mathcal{K}_p$.

Le théorème suivant [60] donne une information importante sur la solution $\delta\mathbf{x}^{(k)}$.

Théorème 5.4. *La solution du système linéaire $J^{(k)}\delta\mathbf{x}^{(k)} = -\mathrm{F}(\mathbf{x}^{(k)})$ appartient à l'espace affine $\delta\mathbf{x}_0^{(k)} + \mathcal{K}_{p_{max}}$.*

Ce résultat montre qu'il est possible de déterminer la solution du problème linéaire en construisant progressivement les sous-espaces de Krylov. Par exemple, l'itéré $\delta\mathbf{x}_l^{(k)}$ construit à partir de l'espace de Krylov $\mathcal{K}_l$ s'écrira sous la forme

$$\delta\mathbf{x}_l^{(k)} = \delta\mathbf{x}_0^{(k)} + \sum_{i=0}^{l-1} \alpha_i [J^{(k)}]^i \mathbf{r}^{(0)}. \tag{5.24}$$

Le calcul des coefficients α_i peut être réalisé de différentes façons. Deux approches qui sont généralement utilisées sont à mentionner tout particulièrement. La première approche consiste à minimiser la norme du résidu $\mathbf{r}^{(l)}$ et constitue la base des méthodes MINRES (*Minimum Residual*) pour les matrices symétriques et GMRES (*Generalized Minimum Residual*) pour les matrices quelconques. Dans la seconde approche, les coefficients sont déterminés de manière à rendre le résidu $\mathbf{r}^{(l)}$ orthogonal à l'espace de Krylov $\mathcal{K}_l$. Dans cette catégorie se retrouvent les méthodes du gradient conjugué pour les matrices symétriques définies positives et FOM (*Full Orthogonal Method*) pour les matrices quelconques.

Comme le montre la relation (5.24), les méthodes de Krylov ne font intervenir la matrice $J^{(k)}$ uniquement que sous la forme du produit de cette matrice par un vecteur. Ainsi, si les méthodes de Krylov sont utilisées pour résoudre le système (5.20a), l'essentiel n'est pas de

connaître l'expression de la matrice jacobienne mais bien son action sur un vecteur. Dans la méthode JFNK, le produit $J^{(k)}\mathbf{v}$ est ainsi approché par une différence finie. Deux choix possibles pour approcher $J^{(k)}\mathbf{v}$ sont donnés par

$$J^{(k)}\mathbf{v} \approx \frac{\mathrm{F}(\mathbf{x}^{(k)} + \epsilon\mathbf{v}) - \mathrm{F}(\mathbf{x}^{(k)})}{\epsilon}, \tag{5.25}$$

$$J^{(k)}\mathbf{v} \approx \frac{\mathrm{F}(\mathbf{x}^{(k)} + \epsilon\mathbf{v}) - \mathrm{F}(\mathbf{x}^{(k)} - \epsilon\mathbf{v})}{2\epsilon}. \tag{5.26}$$

La relation (5.25) représente une approximation d'ordre un de la matrice jacobienne tandis que la relation (5.26) constitue une approximation d'ordre deux. L'approximation (5.26) est plus précise que l'approximation (5.25), ce qui peut permettre de réduire le nombre d'itérations de la méthode de Krylov. Toutefois, cette deuxième approximation nécessite d'évaluer plus de fois la fonction F.

Le paramètre ϵ joue un rôle important au niveau de l'approximation. Une valeur trop importante de ce paramètre rend les approximations par différence finie peu précises tandis qu'une valeur trop petite fait apparaître des imprécisions numériques importantes. En pratique, il existe plusieurs possibilités pour le choix de ce paramètre dont quelques-unes sont présentées par Knoll et Keyes [45].

L'intérêt de la méthode JNFK est à présent évident. Elle permet tout d'abord de conserver un algorithme de résolution similaire à la méthode de Newton, ce qui permet de garder une vitesse de convergence élevée bien que généralement plus faible que celle obtenue avec la méthode de Newton. Ensuite, elle permet de ne jamais calculer explicitement la matrice jacobienne.

Notons pour finir qu'une méthode de préconditionnement est généralement appliquée au système (5.20a) afin d'obtenir un système équivalent pour lequel le nombre d'itérations des méthodes de Krylov sera réduit. Alors que le préconditionnement de systèmes linéaires est bien connu, la construction de préconditionneurs pour des problèmes multiphysiques reste encore une question difficile. Différentes approches pour le préconditionnement pourront notamment être trouvées dans Knoll et Keyes [45].

5.3. Limites de l'utilisation des méthodes itératives sur les sous-problèmes individuels

À présent que nous avons présenté les méthodes de résolution faiblement et fortement couplées, il est légitime de s'interroger sur la pertinence ou non d'utiliser une méthode de résolution numérique faiblement couplée comme celles présentées à la sous-section 5.2.1 au lieu d'une méthode de résolution numérique plus fortement couplée comme celle de Newton. Bien que ces méthodes faiblement couplées soient attractives d'un point de vue implémentation, elles négligent une partie du couplage entre les différents problèmes monophysiques. Dans certains cas, cela peut devenir problématique. Whiteley *et al.* [70] ont étudié la question pour l'approximation de Gauss-Seidel et ont montré dans quelle mesure les termes hors diagonale de la matrice jacobienne influençaient l'efficacité de cette méthode. Nous proposons ci-dessous une approche différente inspirée des résultats théoriques pour la résolution des systèmes linéaires par des méthodes itératives [61].

Nous envisageons tout d'abord la résolution du système linéaire $A\mathbf{x}^* = \mathbf{b}$. À titre illustratif, nous pouvons considérer le système de dimension deux suivant :

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{5.27}$$

La résolution du système $A\mathbf{x}^* = \mathbf{b}$ peut être coûteuse en temps de calcul. Au lieu de résoudre directement le système (5.27), il est plus intéressant de le formuler sous la forme d'une méthode de point fixe pour laquelle nous serions amené à résoudre des systèmes linéaires plus simples. En général, une décomposition de la matrice A est effectuée en sa partie diagonale D, sa partie triangulaire inférieure stricte L et sa partie triangulaire supérieure stricte U. Nous supposons implicitement que les éléments diagonaux sont tous différents de zéro, ce qui constitue une hypothèse raisonnable si le système d'équations vient du couplage de différents sous-systèmes. Les trois méthodes itératives suivantes sont généralement présentées dans la littérature :

- Méthode de Jacobi : l'itération de point fixe pour la méthode de Jacobi est donnée par

$$\mathbf{x}^{(k+1)} = D^{-1}\left[-(L+U)\,\mathbf{x}^{(k)} + \mathbf{b}\right], \tag{5.28a}$$

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} b_1 - A_{12}x_2^{(k)} \\ b_2 - A_{21}x_1^{(k)} \end{bmatrix}. \tag{5.28b}$$

  Dans la méthode de Jacobi, chacune des équations est résolue pour une des inconnues en fixant les autres inconnues à leur valeur à l'itération précédente.

- Méthode de Gauss-Seidel : l'itération de point fixe pour la méthode de Gauss-Seidel est donnée par

$$\mathbf{x}^{(k+1)} = (D+L)^{-1}\left[-U\,\mathbf{x}^{(k)} + \mathbf{b}\right], \tag{5.29a}$$

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} b_1 - A_{12}x_2^{(k)} \\ b_2 \end{bmatrix}. \tag{5.29b}$$

  Pour la méthode de Gauss-Seidel, la première équation est résolue en fixant x_2 à sa valeur à l'itération précédente et la seconde équation est résolue en utilisant la nouvelle valeur calculée pour l'approximation de x_1.

- Méthode de surrelaxation successive : il s'agit d'une généralisation de la méthode de Gauss-Seidel. Le nouvel itéré $x_i^{(k+1)}$ est obtenu comme $x_i^{(k+1)} = x_i^{(k)} + \omega(x_i - x_i^{(k)})$ où x_i est obtenu par la méthode de Gauss-Seidel et $\omega > 0$ est appelé le paramètre de surrelaxation. Sous forme matricielle, nous avons

$$\mathbf{x}^{(k+1)} = (D+\omega L)^{-1}\left[-(\omega U + (\omega-1)D)\,\mathbf{x}^{(k)} + \omega\mathbf{b}\right], \tag{5.30a}$$

$$\begin{bmatrix} A_{11} & 0 \\ \omega A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} \omega b_1 - \omega A_{12}x_2^{(k)} + (1-\omega)A_{11}x_1^{(k)} \\ \omega b_2 + (1-\omega)A_{22}x_2^{(k)} \end{bmatrix}. \tag{5.30b}$$

Chacune des trois méthodes présentées peut être envisagée comme une méthode du point fixe sur un système préconditionné [2]. De manière générique, nous écrirons cette itération de point fixe sous la forme

$$\mathbf{x}^{(k+1)} = G\,\mathbf{x}^{(k)} + \mathbf{f}, \tag{5.31}$$

où G est appelée la matrice d'itération.

La question qui nous intéresse désormais est de savoir sous quelles conditions les méthodes itératives présentées convergent, ce qui revient à connaître quand de telles méthodes peuvent être utilisées pour résoudre un problème couplé. Pour que la convergence de la méthode (5.31) soit assurée, il suffit en fait que toutes les valeurs propres de G soient inférieures à 1 en module, ce qui signifie que le rayon spectral $\rho(G)$ de la matrice G doit être inférieur à 1. Cette information est reprise plus précisément dans le théorème 5.5 [61].

Théorème 5.5. *Soit G une matrice carrée telle que $\rho(G) < 1$. Alors la matrice $I - G$, où I représente la matrice identité, est non singulière et l'itération (5.31) converge pour tout $\mathbf{f}$ et $\mathbf{x}^{(0)}$. Inversement, si l'itération (5.31) converge pour tout $\mathbf{f}$ et $\mathbf{x}^{(0)}$, alors $\rho(G) < 1$.*

Nous allons regarder à présent plus spécifiquement la méthode de Gauss-Seidel. Toutefois, un raisonnement similaire à celui que nous allons mener peut être effectué pour les méthodes de Jacobi et de surrelaxation successive. Pour la méthode de Gauss-Seidel, la matrice G est donnée par

$$G = \begin{bmatrix} 0 & -A_{11}^{-1}A_{12} \\ 0 & A_{22}^{-1}A_{21}A_{11}^{-1}A_{12} \end{bmatrix}. \tag{5.32}$$

La convergence de la méthode de Gauss-Seidel est assurée si $|A_{22}^{-1}A_{21}A_{11}^{-1}A_{12}| < 1$.

De manière générale, les éléments A_{11}, A_{12}, A_{21} et A_{22} sont des blocs matriciels et les méthodes présentées deviennent les méthodes de Jacobi, de Gauss-Seidel et de surrelaxation successive par blocs. Dans ce cas, la convergence de la méthode de Gauss-Seidel sera garantie si le rayon spectral de la matrice $C = A_{22}^{-1}A_{21}A_{11}^{-1}A_{12}$ est inférieur à 1.

Même si nous avons présenté notre raisonnement pour la résolution d'un système linéaire, celui-ci est généralisable à un système non linéaire. En effet, si nous envisageons la résolution du système (5.1) par une méthode de Newton, nous devons résoudre à l'itération k un système linéaire du type

$$\begin{bmatrix} J_{11}^{(k)} & J_{12}^{(k)} \\ J_{21}^{(k)} & J_{22}^{(k)} \end{bmatrix} \begin{bmatrix} x_1^{(k+1)} \\ x_2^{(k+1)} \end{bmatrix} = \begin{bmatrix} b_1^{(k)} \\ b_2^{(k)} \end{bmatrix}. \tag{5.33}$$

Le système (5.33) est en fait équivalent au système (5.27), si ce n'est que ce système varie d'un itération à l'autre de la méthode de Newton. Le rôle de la matrice A est tenu par la matrice jacobienne. Le système (5.33) peut être résolu par une méthode de Gauss-Seidel. Dans ce cas, l'algorithme obtenu est identique à la version linéarisée de l'algorithme 1 de Gauss-Seidel non linéaire. La condition de convergence de cet algorithme est également donnée par le théorème 5.5. Pour la méthode de Gauss-Seidel, cette condition est équivalente à avoir

$$\rho(C^{(k)}) < 1 \quad \text{avec } C^{(k)} = [J_{22}^{(k)}]^{-1}J_{21}^{(k)}[J_{11}^{(k)}]^{-1}J_{12}^{(k)}. \tag{5.34}$$

2. La méthode de Jacobi résulte par exemple de la multiplication du système $A\mathbf{x}^* = \mathbf{b}$ par la matrice de préconditionnement D^{-1}, tandis que la méthode de Gauss-Seidel considère une matrice de préconditionnement égale à $(D + L)^{-1}$.

De manière similaire, nous pouvons utiliser une norme matricielle subordonnée à une norme vectorielle pour écrire la condition de convergence de la méthode de Gauss-Seidel (voir [70]). Nous obtenons ainsi la condition de convergence

$$\|C^{(k)}\| < 1. \tag{5.35}$$

Or, nous savons que pour une norme matricielle, nous avons

$$\|C^{(k)}\| \leqslant \|[J_{22}^{(k)}]^{-1}\| \, \|J_{21}^{(k)}\| \, \|[J_{11}^{(k)}]^{-1}\| \, \|J_{12}^{(k)}\|. \tag{5.36}$$

Cette dernière relation peut aussi se réécrire en terme des nombres de conditionnement $\kappa(J_{11}^{(k)})$ et $\kappa(J_{22}^{(k)})$ des matrices $J_{11}^{(k)}$ et $J_{22}^{(k)}$. Nous écrivons alors

$$\|C^{(k)}\| \leqslant \kappa(J_{11}^{(k)}) \, \kappa(J_{22}^{(k)}) \, \frac{\|J_{12}^{(k)}\| \, \|J_{21}^{(k)}\|}{\|J_{11}^{(k)}\| \, \|J_{22}^{(k)}\|}. \tag{5.37}$$

L'inégalité (5.37) nous donne une indication sur la pertinence ou non d'utiliser la méthode de Gauss-Seidel pour résoudre un système non linéaire couplé. L'expression (5.37) suggère que la méthode de Gauss-Seidel sera peu appropriée si les problèmes individuels sont mal conditionnés (nombre de conditionnement élevé) ou encore lorsque le rapport des termes hors diagonale de la matrice jacobienne sur les termes diagonals est important. Il est en fait naturel de retrouver ces conclusions. La méthode de Gauss-Seidel se limite en substance à résoudre chacun des problèmes individuels et donc si ceux-ci sont mal conditionnés, la méthode globale le sera également. De plus, il est normal que la matrice jacobienne intervienne dans l'analyse de l'efficacité de la méthode de Gauss-Seidel. En effet, cette matrice exprime mathématiquement la sensibilité de chaque équation d'un problème non linéaire aux variables du problème. Les termes hors diagonale de la matrice jacobienne donnent une mesure de la sensibilité de chacun des problèmes monophysiques aux autres problèmes monophysiques. Plus ces termes hors diagonale auront de l'importance et plus le couplage entre les problèmes individuels sera important. Dès lors, adopter une méthode de résolution qui néglige une partie de ce couplage sera peu efficace.

6. Application des méthodes multiphysiques à l'étude d'un modèle de glacier alpin

Ce sixième chapitre aborde la résolution numérique concrète des équations thermomécaniques de bilan appliquées au cas d'un glacier alpin « jouet ». Ce chapitre établit le cadre algébrique à partir duquel ont été réalisées les simulations numériques présentées au chapitre 7. Le chapitre débute à la section 6.1 par une présentation de la formulation forte des équations de bilan pour le glacier « jouet » étudié. Les sections 6.2 et 6.3 introduiront les formulations faibles de ce problème dont nous déduirons la forme algébrique à la section 6.4. À la section 6.5, nous passerons enfin en revue les différentes méthodes numériques présentées au chapitre 5 et nous expliquerons comment ces algorithmes peuvent s'appliquer concrètement au modèle que nous étudions.

6.1. Formulation forte du problème couplé

Dans les chapitres 6 et 7, nous souhaitons étudier un glacier alpin « jouet » bidimensionnel. Une représentation schématique de ce glacier est donnée à la figure 6.1.

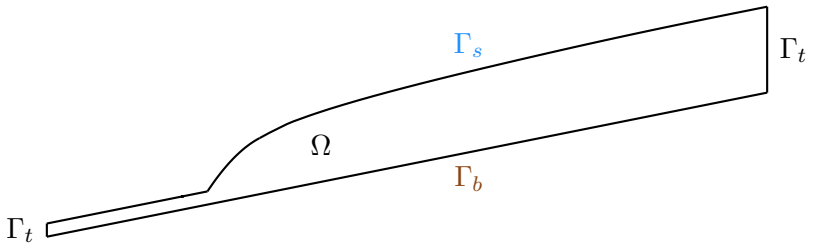


FIGURE 6.1.: Illustration du glacier alpin « jouet ». Les régions Γ_s, Γ_b et Γ_t représentent respectivement les surfaces supérieure, inférieure et latérales du glacier.

La partie du glacier à basse altitude peut représenter la zone d'ablation, tandis que la partie supérieure représente par exemple la zone de transport [1]. Mathématiquement, le glacier est traité comme un domaine $\Omega \in \mathbb{R}^2$ dont la frontière Γ est l'union des régions distinctes Γ_s, Γ_b et Γ_t. Les régions Γ_s et Γ_b constituent comme précédemment les surfaces supérieure (interface air-glace) et inférieure (interface glace-lit rocheux) du glacier. Les extrémités latérales du glacier sont regroupées dans la région Γ_t. Sur cette région, nous avons décidé d'appliquer

1. Dans un glacier, il est possible de distinguer généralement trois zones. La zone d'ablation est la partie du glacier où la fonte de la glace est la plus importante, ce qui entraîne une diminution importante de l'épaisseur du glacier. La zone d'accumulation est la partie du glacier soumise aux précipitations neigeuses les plus importantes. C'est essentiellement dans cette zone que la neige se transforme en glace. Entre les zones d'ablation et d'accumulation se trouve la zone de transport où se produit principalement le transport de la glace des hautes vers les basses altitudes.

des conditions de Dirichlet homogènes pour la vitesse et des conditions de Neumann homogènes pour la température (flux de chaleur nul) [2]. Sur Γ_s et Γ_b, les conditions aux limites sont celles présentées à la section 2.6. Dans cette section, nous avions indiqué que le modèle utilisé présentait une insuffisance au niveau des conditions aux limites pour fixer de manière unique la valeur des composantes de la vitesse. L'introduction de la région Γ_t et les conditions aux limites qui y sont imposées permettent de suppléer à cette insuffisance.

Le modèle mathématique qui décrit le comportement thermodynamique du glacier étudié s'écrit sous la forme forte suivante :

- Modèle mécanique :

$$\begin{cases} -\operatorname{div}(2\mu(\mathbf{v},p,T)\dot{\varepsilon}) + \nabla p = \rho\mathbf{g} & \text{dans } \Omega, \\ \operatorname{div}\mathbf{v} = 0 & \text{dans } \Omega, \\ \sigma\cdot\mathbf{n} = \mathbf{0} & \text{sur } \Gamma_s, \\ \mathbf{v} = \mathbf{0} & \text{sur } \Gamma_t, \\ \mathbf{v}\cdot\mathbf{n} = a_b(T) & \text{sur } \Gamma_b, \\ \mathbf{t}_b = -C_b(T)\,\mathbf{v}_b & \text{sur } \Gamma_b. \end{cases} \tag{6.1}$$

- Modèle thermique :

$$\begin{cases} -\operatorname{div}(k(T)\nabla T) + \rho c(T)\,\mathbf{v}\cdot\nabla T = 4\mu(\mathbf{v},p,T)d_e^2 & \text{dans } \Omega, \\ T = T_s & \text{sur } \Gamma_s, \\ k(T)\nabla T\cdot\mathbf{n} = 0 & \text{sur } \Gamma_t, \\ k(T)\,\nabla T\cdot\mathbf{n} = q_{\text{geo}} - \mathbf{t}_b\cdot\mathbf{v}_b - \begin{cases} \rho L a_b & \text{si } T = T_m(p) \\ 0 & \text{si } T < T_m(p) \end{cases} & \text{sur } \Gamma_b. \end{cases} \tag{6.2}$$

- Équations constitutives :

$$2^n A(p,T)d_e^{n-1}\mu^n(\mathbf{v},p,T) = -2A(p,T)\sigma_0^{n-1}\mu(\mathbf{v},p,T) + 1, \tag{6.3a}$$
$$T_m = 273.15 - \beta p\,[\mathrm{K}], \tag{6.3b}$$
$$k(T) = 9.828\,e^{-0.0057T}\,[\mathrm{W\,m^{-1}\,K^{-1}}], \tag{6.3c}$$
$$c(T) = (146.3 + 7.253T)\,[\mathrm{J\,kg^{-1}\,K^{-1}}], \tag{6.3d}$$
$$C_b(T) = C_0\exp(\gamma(T_m - T))\,. \tag{6.3e}$$

Quelques remarques s'imposent quant à l'écriture de ces équations. Les équations mathématiques ont été séparées en un modèle mécanique et un modèle thermique auxquels nous avons joint les équations constitutives (6.3a)–(6.3e). Cette séparation est artificielle dans la mesure où les deux modèles sont couplés puisque les champs de vitesse et de pression influent sur le champ de température et inversement. Nous avons dès lors un unique système d'équations. Toutefois, la séparation que nous faisons permet de mettre plus en avant le caractère multiphysique du problème qui résulte de la prise en compte simultanée des modèles mécanique et thermique.

2. D'autres conditions aux limites auraient pu être envisagées. Par exemple, la vitesse horizontale et le vecteur traction de surface vertical peuvent être imposés tous les deux à une valeur nulle.

Outre le caractère couplé du modèle, celui-ci est marqué par un haut degré de non-linéarité. Pour souligner cet aspect, nous avons spécifié dans les systèmes (6.1) et (6.2) la dépendance des propriétés physiques du fluide vis-à-vis des inconnues du problème. Cette dépendance est rendue plus explicite dans les lois constitutives (6.3a)–(6.3e). Le terme convectif $\mathbf{v} \cdot \nabla T$ de l'équation de la chaleur est à présent un terme non linéaire, ce qui n'était pas le cas lors de l'étude du problème thermique seul. Enfin, il est intéressant de constater que les conditions aux limites dépendent également de la solution. En particulier, la température de fusion de la glace dépend de la pression tandis que le taux de fonte aura une valeur nulle ou non suivant la valeur de la température.

Pour simplifier le traitement analytique, nous formulons les trois hypothèses suivantes :

- T_m est indépendant de la pression p.

  Notre première hypothèse revient à formuler l'indépendance de la température de fusion de la glace vis-à-vis de la pression. Cette hypothèse est relativement plausible pour un glacier de montagne de faible épaisseur. Pour un glacier de 100 [m] d'épaisseur, la pression hydrostatique au niveau du sol est de l'ordre de 0.9 [MPa] et la variation de la température de fusion est d'environ 0.07 [K] par rapport à la température de fusion à la surface libre. La variation de la température de fusion est suffisamment faible pour pouvoir être négligée en première approximation.

- Nous imposons uniquement $\mathbf{v} \cdot \mathbf{n} = 0$ sur Γ_b.

  Cette deuxième hypothèse consiste à remplacer la condition aux limites $\mathbf{v} \cdot \mathbf{n} = a_b(T)$ sur Γ_b qui dépend de la température par la condition plus simple $\mathbf{v} \cdot \mathbf{n} = 0$. Nous négligeons de cette façon le taux de fonte basale dans le modèle mécanique. Cette hypothèse peut sembler moins pertinente que la première hypothèse que nous avons posée. Toutefois, cette hypothèse simplificatrice est réalisée dans certains travaux notamment [63]. Nous pouvons néanmoins la justifier partiellement à partir d'un raisonnement dimensionnel. Le taux de fonte basale a_b est de l'ordre de $q_{\text{geo}}/(\rho L)$ avec un flux géothermique de l'ordre de $100\,[\text{mW/m}^2]$, $\rho = 910\,[\text{kg/m}^3]$ et $L = 334\,[\text{kJ/kg}]$. Pour ces valeurs, nous avons $a_b \approx 0.01\,[\text{m/an}]$. En première approximation, cette valeur peut être négligée dans le modèle mécanique. Notons également que si le taux de fonte basale n'est pas pris en compte dans le modèle mécanique, sa présence est nécessaire dans le modèle thermique pour assurer que la température de la glace ne dépasse pas son point de fusion.

- L'équation d'advection-diffusion est résolue sans terme de stabilisation.

  Cette troisième hypothèse suppose que l'importance de la convection thermique dans le glacier reste modérée. Dans ce cas, l'équation de la chaleur peut être discrétisée avec la méthode classique de Galerkin. Cette hypothèse sera vérifiée au chapitre 7 par le biais de simulations numériques.

6.2. Formulation variationnelle du problème couplé

Pour établir la formulation variationnelle du problème couplé, nous introduisons les ensembles de fonctions suivants :

$$V = \{\mathbf{w} : \mathbf{w} = \mathbf{0} \text{ sur } \Gamma_t, \mathbf{w}\cdot\mathbf{n} = 0 \text{ sur } \Gamma_b\}, \tag{6.4}$$
$$H = \{\theta : \theta = 0 \text{ sur } \Gamma_s\}, \tag{6.5}$$
$$G = \{\theta : \theta = T_s \text{ sur } \Gamma_s\}, \tag{6.6}$$
$$K = \{\theta \in G : \theta \leqslant T_m \text{ sur } \Gamma_b\}. \tag{6.7}$$

La vitesse $\mathbf{v}$ et la température T, solutions du système (6.1)–(6.2), appartiennent respectivement aux ensembles V et K. Nous supposons également qu'il existe un espace de fonctions Q auquel appartient la pression. L'établissement de la formulation faible pour le problème couplé peut s'effectuer de manière similaire à ce que nous avons présenté aux chapitres 3 et 4 de sorte que nous ne présenterons pas les détails de calcul pour le problème couplé. Nous renvoyons le lecteur à ces chapitres pour plus de détails.

La formulation variationnelle du problème (6.1)–(6.2) est donnée par

Formulation variationnelle 6.1. Trouver $\mathbf{v} \in V$, $p \in Q$ et $T \in K$ tels que

$$\begin{cases} \displaystyle\int_\Omega 2\mu\,\dot\varepsilon(\mathbf{v}) : \dot\varepsilon(\mathbf{w})\,d\Omega + \int_{\Gamma_b} C_b\,\mathbf{v}_b\cdot\mathbf{w}_b\,d\Gamma_b - \int_\Omega p\,\mathrm{div}\,\mathbf{w}\,d\Omega = \int_\Omega \rho\mathbf{g}\cdot\mathbf{w}\,d\Omega, \\ \displaystyle\int_\Omega q\,\mathrm{div}\,\mathbf{v}\,d\Omega = 0, \end{cases}$$

$$\int_\Omega k\,\nabla T\cdot\nabla(\theta - T)\,d\Omega + \int_\Omega \rho c\,(\mathbf{v}\cdot\nabla T)\,(\theta - T)\,d\Omega \geqslant \int_\Omega 4\mu d_e^2\,(\theta - T)\,d\Omega + \int_{\Gamma_b} (q_{\mathrm{geo}} + C_b\,\|\mathbf{v}_b\|^2)\,(\theta - T)\,d\Gamma_b,$$

pour tout triplet $(\mathbf{w}, q, \theta) \in V \times Q \times K$.

6.3. Formulation variationnelle pénalisée du problème couplé

La formulation variationnelle 6.1 n'est pas directement utilisable pour établir une formulation algébrique du problème couplé. Afin de résoudre numériquement le problème thermique contraint, nous décidons de considérer une méthode de résolution par pénalisation (voir sous-section 4.5.2). Dans l'optique d'une utilisation de la méthode de résolution de Newton, nous envisageons un terme de pénalisation de la forme

$$\frac{1}{\varepsilon}P(T_\varepsilon) = \frac{1}{a\varepsilon}\,\mathbf{1}_{\mathbb{R}^+}(T_\varepsilon - T_m) \times (T_\varepsilon - T_m)^a. \tag{6.8}$$

où le paramètre a est strictement supérieur à un pour travailler avec un terme de pénalisation dérivable en $T = T_m$. La valeur $a = 2$ représente le cas d'une pénalisation quadratique, mais d'autres valeurs de a moins utilisées en pratique peuvent aussi être envisagées. Le choix de ce paramètre sera discuté plus précisément à la section 7.3.1. De plus, vu l'hypothèse sur le problème thermique formulée à la section 6.1, nous n'ajoutons pas de terme de stabilisation dans la formulation variationnelle de l'équation de la chaleur. Ces deux remarques nous amènent ainsi à résoudre la formulation variationnelle approchée 6.2.

Formulation variationnelle 6.2. Trouver $\mathbf{v} \in V$, $p \in Q$ et $T_\varepsilon \in G$ tels que

$$\begin{cases} \displaystyle\int_\Omega 2\mu\,\dot\varepsilon(\mathbf{v}) : \dot\varepsilon(\mathbf{w})\,d\Omega + \int_{\Gamma_b} C_b\,\mathbf{v}_b\cdot\mathbf{w}_b\,d\Gamma_b - \int_\Omega p\,\operatorname{div}\mathbf{w}\,d\Omega = \int_\Omega \rho\mathbf{g}\cdot\mathbf{w}\,d\Omega, \\ \displaystyle\int_\Omega q\,\operatorname{div}\mathbf{v}\,d\Omega = 0, \end{cases}$$

$$\int_\Omega k\,\nabla T_\varepsilon\cdot\nabla\theta\,d\Omega + \int_\Omega \rho c\,(\mathbf{v}\cdot\nabla T_\varepsilon)\,\theta\,d\Omega = \int_\Omega 4\mu d_e^2\;\theta\,d\Omega + \int_{\Gamma_b}(q_{\text{geo}} + C_b\,\|\mathbf{v}_b\|^2)\,\theta\,d\Gamma_b - \int_{\Gamma_b}\frac{1}{a\varepsilon}\,\mathbf{1}_{\mathbb{R}^+}(T_\varepsilon - T_m)\times(T_\varepsilon - T_m)^a\,\theta\,d\Gamma_b,$$

pour tout triplet $(\mathbf{w}, q, \theta) \in V \times Q \times H$.

Par la suite afin de ne pas alourdir les notations, nous remplacerons systématiquement la notation T_ε par T. Il faut néanmoins rester conscient de la dépendance de T envers le paramètre de pénalisation.

6.4. Formulation algébrique du problème couplé

Afin d'écrire la formulation variationnelle 6.2 sous forme discrète, nous travaillons avec les espaces de dimension finie $V_h \subset V$, $Q_h \subset Q$, $G_h \subset G$ et $H_h \subset H$. Les inconnues du problème sont approchées par les fonctions $\mathbf{v}_h \in V_h$, $p_h \in Q_h$ et $T_h \in G_h$ tandis que les fonctions test sont représentées par les fonctions $\mathbf{w}_h \in V_h$, $q_h \in Q_h$ et $\theta_h \in H_h$. Ces fonctions sont ensuite décomposées sur une base de fonctions de leur espace respectif. Nous écrirons ainsi

$$\mathbf{v}_h(\mathbf{x}) = \sum_{j=1}^{N} v_j\,\boldsymbol{\varphi}_j(\mathbf{x}), \qquad \mathbf{w}_h(\mathbf{x}) = \sum_{k=1}^{N} w_k\,\boldsymbol{\varphi}_k(\mathbf{x}), \tag{6.9a}$$

$$p_h(\mathbf{x}) = \sum_{j=1}^{M} p_j\,\phi_j(\mathbf{x}), \qquad q_h(\mathbf{x}) = \sum_{k=1}^{M} q_k\,\phi_k(\mathbf{x}), \tag{6.9b}$$

$$T_h(\mathbf{x}) = \sum_{j=1}^{R} T_j\,\chi_j(\mathbf{x}), \qquad \theta_h(\mathbf{x}) = \sum_{k=1}^{R} \theta_k\,\chi_k(\mathbf{x}). \tag{6.9c}$$

Il est utile de préciser que les espaces de discrétisation de la vitesse et de la pression sont choisis de manière à satisfaire la condition inf-sup. De plus, comme nous avons supposé que la convection thermique restait modérée, nous utilisons les mêmes fonctions de base pour T_h et θ_h. En outre dans les expressions (6.9a)–(6.9c), certains degrés de liberté sont fixés par les conditions aux limites. La condition aux limites $\mathbf{w}\cdot\mathbf{n} = 0$ rend également certains degrés de liberté dépendants entre eux. Nous supposons cependant pour établir la formulation algébrique que ces degrés de liberté sont bien indépendants. La prise en compte de ces conditions aux limites pourra être réalisée ultérieurement.

Pour la suite, nous réunissons les coefficients $\{v_j\}$, $\{p_j\}$ et $\{T_j\}$ de la solution dans les vecteurs $\mathbf{V}$, $\mathbf{P}$ et $\mathbf{T}$ tandis que les coefficients $\{w_j\}$, $\{q_j\}$ et $\{\theta_j\}$ des fonctions test sont placés dans les vecteurs $\mathbf{W}$, $\mathbf{Q}$ et $\boldsymbol{\Theta}$.

La formulation algébrique du problème variationnel 6.2 peut s'écrire sous la forme suivante :

$$\begin{bmatrix} A(\mathbf{V},\mathbf{T}) & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix}, \tag{6.10a}$$

$$\Big[C(\mathbf{V},\mathbf{T}) + K(\mathbf{T})\Big]\ \mathbf{T} = \mathbf{G}(\mathbf{V},\mathbf{T}), \tag{6.10b}$$

avec

$$\mathbf{W}^T A\mathbf{V} = \int_\Omega 2\mu(\mathbf{v}_h, T_h)\, \dot\varepsilon(\mathbf{v}_h) : \dot\varepsilon(\mathbf{w}_h)\, d\Omega + \int_{\Gamma_b} C_b(T_h)\, \mathbf{v}_{h,b} \cdot \mathbf{w}_{h,b}\, d\Gamma_b, \tag{6.11a}$$

$$\mathbf{Q}^T B\mathbf{V} = -\int_\Omega q_h \operatorname{div} \mathbf{v}_h\, d\Omega, \tag{6.11b}$$

$$\mathbf{W}^T\mathbf{F} = \int_\Omega \rho\mathbf{g} \cdot \mathbf{w}_h\, d\Omega, \tag{6.11c}$$

$$\mathbf{\Theta}^T C\mathbf{T} = \int_\Omega \rho c(T_h)\, (\mathbf{v}_h \cdot \nabla T_h)\, \theta_h\, d\Omega, \tag{6.11d}$$

$$\mathbf{\Theta}^T K\mathbf{T} = \int_\Omega k(T_h)\, \nabla T_h \cdot \nabla\theta_h\, d\Omega, \tag{6.11e}$$

$$\begin{aligned} \mathbf{\Theta}^T\mathbf{G} = &\int_\Omega 4\mu(\mathbf{v}_h, T_h)\, d_e^2(\mathbf{v}_h)\ \theta_h\, d\Omega + \int_{\Gamma_b} (q_{\text{geo}} + C_b(T_h)\, \|\mathbf{v}_{h,b}\|^2)\, \theta_h\, d\Gamma_b \\ &- \int_{\Gamma_b} \frac{1}{a\varepsilon}\, \mathbf{1}_{\mathbb{R}^+}(T_h - T_m) \times (T_h - T_m)^a\, \theta_h\, d\Gamma_b. \end{aligned} \tag{6.11f}$$

Les matrices A, C et K sont appelées respectivement les matrices de viscosité, de convection et de diffusion (thermique) du système couplé. Bien que le système (6.10a)–(6.10b) ait l'apparence d'un système linéaire, il s'agit en fait d'équations non linéaires, car les matrices A, C et K ainsi que le vecteur $\mathbf{G}$ dépendent des inconnues du problème. Les diverses dépendances sont mises en évidence dans le choix de nos notations pour l'écriture du système (6.10a)–(6.10b).

6.5. Applications des algorithmes de résolution aux écoulements glaciaires

6.5.1. Méthodes de Jacobi et de Gauss-Seidel

Le problème couplé (6.10a)–(6.10b) se prête sans difficultés à une résolution par des méthodes itératives faiblement couplées de type Jacobi et Gauss-Seidel. La procédure pour ces deux méthodes est donnée respectivement par les algorithmes 4 et 5. Ces deux méthodes peuvent être implémentées facilement à partir du moment où nous disposons d'un solveur pour le problème mécanique et d'un autre solveur pour le problème thermique.

Algorithme 4 Algorithme de Jacobi

1: Initialiser $\mathbf{V}^{(0)}$, $\mathbf{P}^{(0)}$ et $\mathbf{T}^{(0)}$
2: **pour** $k = 1, 2, \ldots$ (jusqu'à convergence) **faire**
3: Résoudre $\begin{bmatrix} A(\mathbf{V},\mathbf{T}^{(k-1)}) & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix}$;
Mettre $\mathbf{V}^{(k)} = \mathbf{V}$ et $\mathbf{P}^{(k)} = \mathbf{P}$
4: Résoudre $\Big[C(\mathbf{V}^{(k-1)},\mathbf{T}) + K(\mathbf{T})\Big]\ \mathbf{T} = \mathbf{G}(\mathbf{V}^{(k-1)},\mathbf{T})$;
Mettre $\mathbf{T}^{(k)} = \mathbf{T}$
5: **fin pour**

Algorithme 5 Algorithme de Gauss-Seidel

1: Initialiser $\mathbf{V}^{(0)}$, $\mathbf{P}^{(0)}$ et $\mathbf{T}^{(0)}$
2: **pour** $k = 1, 2, \dots$ (jusqu'à convergence) **faire**
3: Résoudre $\begin{bmatrix} A(\mathbf{V}, \mathbf{T}^{(k-1)}) & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix}$;
Mettre $\mathbf{V}^{(k)} = \mathbf{V}$ et $\mathbf{P}^{(k)} = \mathbf{P}$
4: Résoudre $\left[C(\mathbf{V}^{(k)}, \mathbf{T}) + K(\mathbf{T})\right] \mathbf{T} = \mathbf{G}(\mathbf{V}^{(k)}, \mathbf{T})$;
Mettre $\mathbf{T}^{(k)} = \mathbf{T}$
5: **fin pour**

6.5.2. Méthode de Picard

Pour appliquer l'algorithme de Picard, nous réécrivons le système (6.10a)–(6.10b) sous la forme suivante :

$$\begin{bmatrix} \mathbf{V} \\ \mathbf{P} \\ \mathbf{T} \end{bmatrix} = \begin{bmatrix} A(\mathbf{V}, \mathbf{T}) & B^T & 0 \\ B & 0 & 0 \\ D(\mathbf{V}, \mathbf{T}) & 0 & C(\mathbf{V}, \mathbf{T}) + K(\mathbf{T}) \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \\ \mathbf{H}(\mathbf{T}) \end{bmatrix}, \tag{6.12}$$

où nous avons introduit

$$\boldsymbol{\Theta}^T D\mathbf{V} = -\int_\Omega 4\mu(\mathbf{v}_h, T_h)\, d_e^2(\mathbf{v}_h)\ \theta_h\, d\Omega - \int_{\Gamma_b} C_b(T_h)\, \|\mathbf{v}_{h,b}\|^2\, \theta_h\, d\Gamma_b, \tag{6.13a}$$

$$\boldsymbol{\Theta}^T \mathbf{H} = \int_{\Gamma_b} q_{\text{geo}}\, \theta_h\, d\Gamma_b - \int_{\Gamma_b} \frac{1}{a\varepsilon}\, \mathbf{1}_{\mathbb{R}^+}(T_h - T_m) \times (T_h - T_m)^a\, \theta_h\, d\Gamma_b. \tag{6.13b}$$

L'itération de point fixe (5.4) peut s'appliquer pour résoudre le système d'équations (6.12). Notons que la procédure (6.12) est purement formelle. En pratique, l'inversion de la matrice sera remplacée par des méthodes de résolution de systèmes d'équations linéaires beaucoup plus performantes. La méthode du point fixe permet ainsi de remplacer la résolution d'un système non linéaire par une suite de systèmes linéaires plus faciles à résoudre. L'algorithme général de résolution du problème couplé est résumé à l'algorithme 6.

Algorithme 6 Algorithme de Picard

1: Initialiser $\mathbf{V}^{(0)}$, $\mathbf{P}^{(0)}$ et $\mathbf{T}^{(0)}$
2: **pour** $k = 1, 2, \dots$ (jusqu'à convergence) **faire**
3: Résoudre le système linéaire

$$\begin{bmatrix} A^{(k-1)} & B^T & 0 \\ B & 0 & 0 \\ D^{(k-1)} & 0 & C^{(k-1)} + K^{(k-1)} \end{bmatrix} \begin{bmatrix} \mathbf{V}^{(k)} \\ \mathbf{P}^{(k)} \\ \mathbf{T}^{(k)} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \\ \mathbf{H}^{(k-1)} \end{bmatrix}$$

4: **fin pour**

Notons enfin que la formulation de point fixe n'est pas unique et que nous en avons uniquement présenté un exemple particulier. La formulation que nous avons adoptée fait apparaître le bloc hors diagonale $D(\mathbf{V}, \mathbf{T})$, ce qui permet d'obtenir un couplage plus important entre les différentes inconnues que pour les méthodes de type Jacobi et Gauss-Seidel. Le désavantage est qu'à présent, il est nécessaire de résoudre un système linéaire de plus grande dimension. La méthode de Picard présentée peut être considérée comme fortement couplée, car les variables du problème mécanique et thermique restent à tout instant synchronisées.

6.5.3. Méthode de Newton

Afin d'appliquer la méthode de Newton, nous réécrivons les équations (6.10a)–(6.10b) sous la forme d'un vecteur résidu à savoir

$$\begin{bmatrix} \mathbf{R_V}(\mathbf{V},\mathbf{P},\mathbf{T}) \\ \mathbf{R_P}(\mathbf{V}) \end{bmatrix} = \begin{bmatrix} A(\mathbf{V},\mathbf{T}) & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{P} \end{bmatrix} - \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \tag{6.14a}$$

$$\mathbf{R_T}(\mathbf{V},\mathbf{T}) = \Big[C(\mathbf{V},\mathbf{T}) + K(\mathbf{T})\Big]\,\mathbf{T} - \mathbf{G}(\mathbf{V},\mathbf{T}) = \mathbf{0}. \tag{6.14b}$$

Pour utiliser la méthode de Newton, nous devons évaluer la matrice jacobienne du système (6.14a)–(6.14b). Afin de mieux voir le couplage entre les différents équations et inconnues du problème couplé, il est intéressant d'écrire tout d'abord la matrice jacobienne sous la forme suivante :

$$J = \left[\begin{array}{cc:c} D_{\mathbf{V}}\mathbf{R_V} & D_{\mathbf{P}}\mathbf{R_V} & D_{\mathbf{T}}\mathbf{R_V} \\ D_{\mathbf{V}}\mathbf{R_P} & 0 & 0 \\ \hdashline D_{\mathbf{V}}\mathbf{R_T} & 0 & D_{\mathbf{T}}\mathbf{R_T} \end{array}\right]. \tag{6.15}$$

Dans la matrice (6.15), nous avons mis en exergue les blocs liés à la linéarisation des problèmes mécanique et thermique considérés séparément ainsi que les termes de couplage entre ces deux problèmes monophysiques. Nous constatons que pour le modèle étudié, le couplage entre les deux modèles se fait au moyen des champs de vitesse et de température, c'est-à-dire au travers des termes $D_{\mathbf{T}}\mathbf{R_V}$ et $D_{\mathbf{V}}\mathbf{R_T}$. Vu les hypothèses que nous avons posées au début de ce chapitre, la pression n'affecte plus les propriétés thermiques de la glace et conserve dès lors uniquement son rôle de multiplicateur de Lagrange de la contrainte d'incompressibilité.

Les différentes composantes de la matrice jacobienne (6.15) sont données par [3]

$$\begin{aligned}
\mathbf{W}^T D_{\mathbf{V}}\mathbf{R_V}(\mathbf{V},\mathbf{P},\mathbf{T})\delta\mathbf{V} &= \int_\Omega 2\,\mu(\mathbf{v}_h,T_h)\,\dot\varepsilon(\delta\mathbf{v}_h) : \dot\varepsilon(\mathbf{w}_h)\,d\Omega + \int_{\Gamma_b} C_b(T_h)\,\delta\mathbf{v}_{h,b}\cdot\mathbf{w}_{h,b}\,d\Gamma_b \\
&\quad + \int_\Omega 2\,\nabla_{\mathbf{v}}\mu(\mathbf{v}_h,T_h)\,\delta\mathbf{v}_h\,\dot\varepsilon(\mathbf{v}_h) : \dot\varepsilon(\mathbf{w}_h)\,d\Omega, && (6.16a)\\
\mathbf{W}^T D_{\mathbf{P}}\mathbf{R_V}(\mathbf{V},\mathbf{P},\mathbf{T})\delta\mathbf{P} &= -\int_\Omega \delta p_h\,\mathrm{div}\,\mathbf{w}_h\,d\Omega, && (6.16b)\\
\mathbf{W}^T D_{\mathbf{T}}\mathbf{R_V}(\mathbf{V},\mathbf{P},\mathbf{T})\delta\mathbf{T} &= \int_\Omega 2\,\nabla_T\mu(\mathbf{v}_h,T_h)\,\delta\mathbf{v}_h\,\dot\varepsilon(\mathbf{v}_h) : \dot\varepsilon(\mathbf{w}_h)\,d\Omega && (6.16c)\\
&\quad + \int_{\Gamma_b} \nabla_T C_b(T_h)\,\delta T_h\,\mathbf{v}_{h,b}\cdot\mathbf{w}_{h,b}\,d\Gamma_b, && (6.16d)\\
\mathbf{Q}^T D_{\mathbf{V}}\mathbf{R_P}(\mathbf{V})\delta\mathbf{V} &= -\int_\Omega q_h\,\mathrm{div}\,\delta\mathbf{v}_h\,d\Omega, && (6.16e)\\
\mathbf{\Theta}^T D_{\mathbf{V}}\mathbf{R_T}(\mathbf{V},\mathbf{T})\delta\mathbf{V} &= \int_\Omega \rho c(T_h)\,\delta\mathbf{v}\cdot\nabla T_h\,\theta_h\,d\Omega - \int_{\Gamma_b} 2\,C_b(T_h)\,\delta\mathbf{v}_{h,b}\cdot\mathbf{v}_{h,b}\,\theta_h\,d\Gamma_b \\
&\quad - \int_\Omega 4\,\nabla_{\mathbf{v}}\mu(\mathbf{v}_h,T_h)\,\delta\mathbf{v}_h\,d_e^2(\mathbf{v}_h)\,\theta_h\,d\Omega \\
&\quad - \int_\Omega 4\,\mu(\mathbf{v}_h,T_h)\,\dot\varepsilon(\delta\mathbf{v}_h) : \dot\varepsilon(\mathbf{v}_h)\,\theta_h\,d\Omega, && (6.16f)
\end{aligned}$$

3. Dans les expressions (6.16a)–(6.16g), nous notons par $\nabla_{\mathbf{v}}$ et ∇_T les dérivées respectivement par rapport à la vitesse et à la température.

$$\begin{aligned}\boldsymbol{\Theta}^T D_{\mathbf{T}}\mathbf{R}_{\mathbf{T}}(\mathbf{V},\mathbf{T})\delta\mathbf{T} = &\int_\Omega k(T_h)\,\nabla\delta T_h\cdot\nabla\theta_h\,d\Omega + \int_\Omega \nabla_T k(T_h)\,\delta T_h\,\nabla T_h\cdot\nabla\theta_h\,d\Omega\\ &+\int_\Omega \rho c(T_h)\,\mathbf{v}_h\cdot\nabla\delta T_h\,\theta_h\,d\Omega + \int_\Omega \rho\nabla_T c(T_h)\,\delta T_h\,\mathbf{v}_h\cdot\nabla T_h\,\theta_h\,d\Omega\\ &-\int_\Omega \nabla_T C_b(T_h)\,\delta T_h\,\|\mathbf{v}_{h,b}\|^2\,\theta_h\,d\Omega\\ &+\frac{1}{\varepsilon}\int_{\Gamma_b}\mathbf{1}_{\mathbb{R}^+}(T_h-T_m)\times(T_h-T_m)^{a-1}\,\delta T_h\,\theta_h\,d\Gamma_b\\ &-\int_\Omega 4\,\nabla_T\mu(\mathbf{v}_h,T_h)\,\delta T_h\,d_e^2(\mathbf{v}_h)\,\theta_h\,d\Omega. \end{aligned}\tag{6.16g}$$

Une fois que la matrice jacobienne est connue, nous pouvons appliquer la méthode de Newton comme présentée à la sous-section 5.2.3. La procédure générale pour cette méthode est résumée à l'algorithme 7.

Algorithme 7 Algorithme de Newton

1: Initialiser $\mathbf{V}^{(0)}$, $\mathbf{P}^{(0)}$ et $\mathbf{T}^{(0)}$
2: **pour** $k = 0, 1, 2, \ldots$ (jusqu'à convergence) **faire**
3: Résoudre le système linéaire

$$J(\mathbf{V}^{(k)},\mathbf{P}^{(k)},\mathbf{T}^{(k)})\begin{bmatrix}\delta\mathbf{V}^{(k)}\\ \delta\mathbf{P}^{(k)}\\ \delta\mathbf{T}^{(k)}\end{bmatrix} = -\begin{bmatrix}\mathbf{R}_{\mathbf{V}}(\mathbf{V}^{(k)},\mathbf{P}^{(k)},\mathbf{T}^{(k)})\\ \mathbf{R}_{\mathbf{P}}(\mathbf{V}^{(k)})\\ \mathbf{R}_{\mathbf{T}}(\mathbf{V}^{(k)},\mathbf{T}^{(k)})\end{bmatrix}$$

4: Mettre

$$\begin{bmatrix}\mathbf{V}^{(k+1)}\\ \mathbf{P}^{(k+1)}\\ \mathbf{T}^{(k+1)}\end{bmatrix} = \begin{bmatrix}\mathbf{V}^{(k)}\\ \mathbf{P}^{(k)}\\ \mathbf{T}^{(k)}\end{bmatrix} + \begin{bmatrix}\delta\mathbf{V}^{(k)}\\ \delta\mathbf{P}^{(k)}\\ \delta\mathbf{T}^{(k)}\end{bmatrix}$$

5: **fin pour**

6.5.4. Méthodes de Quasi-Newton

Même si la matrice jacobienne du problème couplé (6.14a)–(6.14b) peut être évaluée analytiquement, il est parfois souhaitable de ne pas devoir la calculer numériquement ou du moins de limiter le nombre de fois où celle-ci est évaluée. Dans ce cas, il est préférable d'adopter des méthodes sans matrice jacobienne (*jacobian free methods*). Une possibilité que nous utiliserons au chapitre 7 est offerte par la méthode de Broyden qui a été introduite à la sous-section 5.2.4. Comme la matrice jacobienne du problème couplé ne présente aucune propriété particulière, l'utilisation de cette méthode peut paraître un choix adéquat. Toutefois, l'application directe de l'algorithme de Broyden 3 n'est pas optimale d'un point de vue numérique. En effet, même si la matrice jacobienne du système est creuse, les différentes approximations $A^{(k)}$ ne le seront pas nécessairement. Dès lors, la résolution du système linéaire à chaque itération de la méthode de Broyden peut rapidement devenir prohibitive. Pour éviter cela, nous avons décidé d'appliquer un algorithme rapporté par Kelley [41] et dont la procédure est reprise à l'algorithme 8. Cet algorithme offre l'avantage de ne devoir résoudre à chaque itération qu'un système linéaire qui implique la matrice $A^{(0)}$. Si cette matrice est creuse, alors le coût de calcul pour la résolution du système linéaire est réduit. La matrice $A^{(0)}$ peut par exemple être initialisée avec la matrice jacobienne du système. Un autre avantage de l'algorithme 8 est qu'il fait uniquement intervenir à chaque itération des opérations matricielles simples telles qu'un produit scalaire.

La complexité de l'algorithme augmente avec le nombre d'itérations. Si celui-ci reste faible, cette reformulation de l'algorithme de Broyden 3 peut se révéler relativement efficace.

Algorithme 8 Algorithme de Broyden (implémentation creuse)

1: Initialiser $\mathbf{X}^{(0)} = [\mathbf{V}^{(0)}, \mathbf{P}^{(0)}, \mathbf{T}^{(0)}]$ et $A^{(0)}$
2: Résoudre le système linéaire

$$A^{(0)}\mathbf{s}^{(0)} = -\begin{bmatrix}\mathbf{R_V}(\mathbf{X}^{(0)})\\ \mathbf{R_P}(\mathbf{X}^{(0)})\\ \mathbf{R_T}(\mathbf{X}^{(0)})\end{bmatrix}$$

3: **pour** $k = 0, 1, 2, \ldots$ (jusqu'à convergence) **faire**
4: Mettre

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} + \mathbf{s}^{(k)}$$

5: Résoudre le système linéaire

$$A^{(0)}\mathbf{z} = -\begin{bmatrix}\mathbf{R_V}(\mathbf{X}^{(k+1)})\\ \mathbf{R_P}(\mathbf{X}^{(k+1)})\\ \mathbf{R_T}(\mathbf{X}^{(k+1)})\end{bmatrix}$$

6: **pour** $j = 0, 1, \ldots, k-1$ **faire**
7: Mettre

$$\mathbf{z} = \mathbf{z} + \mathbf{s}^{(j+1)}\frac{(\mathbf{s}^{(j)})^T\mathbf{z}}{\|\mathbf{s}^{(j)}\|_2^2}$$

8: **fin pour**
9: Mettre

$$\mathbf{s}^{(k+1)} = \frac{\mathbf{z}}{1 - \frac{(\mathbf{s}^{(k)})^T\mathbf{z}}{\|\mathbf{s}^{(k)}\|_2^2}}$$

10: **fin pour**

7. Résultats numériques

Nous terminons ce travail par la présentation de quelques simulations numériques réalisées sur notre modèle « jouet ». Notre objectif sera d'implémenter les méthodes numériques présentées tout au long de ce travail afin de les comparer et de vérifier leur efficacité sur un exemple concret. Nous commencerons à la section 7.1 par présenter les caractéristiques physiques du glacier que nous étudions. Ensuite, les sections 7.2 et 7.3 présenteront quelques résultats numériques pour les problèmes mécanique et thermique individuels. La section 7.4 s'intéressera enfin au cas du problème thermomécanique couplé.

7.1. Présentation du glacier étudié

Nous commençons ce chapitre par la présentation du glacier que nous souhaitons étudier numériquement. Il s'agit d'un glacier de montagne posé sur une pente qui a une inclinaison d'environ 11.3 [°]. Le glacier a une longueur d'approximativement 2550 [m] et une épaisseur maximale d'environ 72.4 [m]. Une illustration de ce glacier est reprise à la figure 7.1. L'exemple que nous traitons est inspiré d'un tutoriel consacré à l'utilisation du logiciel Elmer/Ice qui permet la résolution par éléments finis des équations de Stokes [24] [1]. Cela nous a donné l'opportunité de valider certains aspects de notre propre code ainsi que de disposer de certains paramètres physiques pour notre modèle.

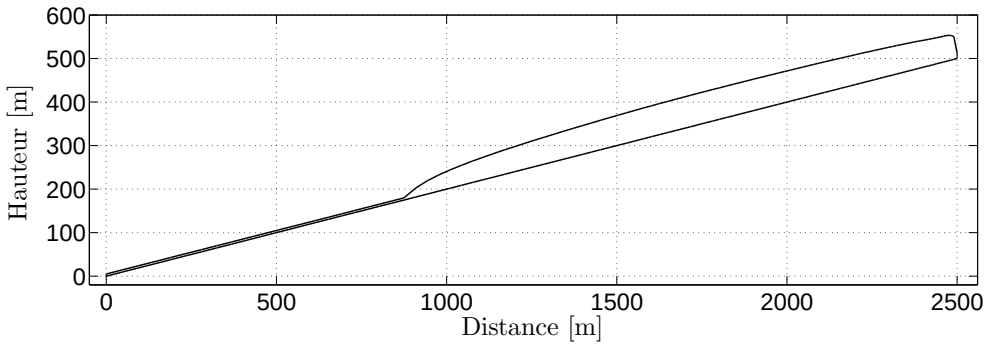


FIGURE 7.1.: Représentation graphique du glacier « jouet » utilisé pour les simulations numériques.

Le maillage du domaine et la visualisation des résultats ont été réalisés à l'aide du logiciel Gmsh [25] [2]. Le code de calcul proprement dit a été réalisé avec le logiciel MATLAB [3]. Pour l'ensemble de nos simulations, nous avons travaillé avec un maillage structuré composé de quadrangles (voir illustration 7.2). Sauf indication contraire, l'ensemble des simulations de ce chapitre a été réalisé sur un maillage composé de 1030 quadrangles. Pour le choix des

1. http://elmerice.elmerfem.org/
2. http://geuz.org/gmsh/
3. http://nl.mathworks.com/

fonctions de base utilisées dans la discrétisation de Galerkin, nous avons opté pour des fonctions bilinéaires par morceaux pour les champs de pression et de température, tandis que le champ de vitesse est discrétisé avec des fonctions bilinéaires par morceaux auxquelles nous avons adjoint les fonctions bulles (3.41) et (3.42) afin de vérifier la condition de stabilité de Babuska-Brezzi. Notons également qu'il est plus pertinent pour les écoulements glaciaires de travailler avec un système d'unités différent du système international. Pour notre part, nous avons choisi de travailler avec les unités suivantes : [an], [m], [MPa] et [K].

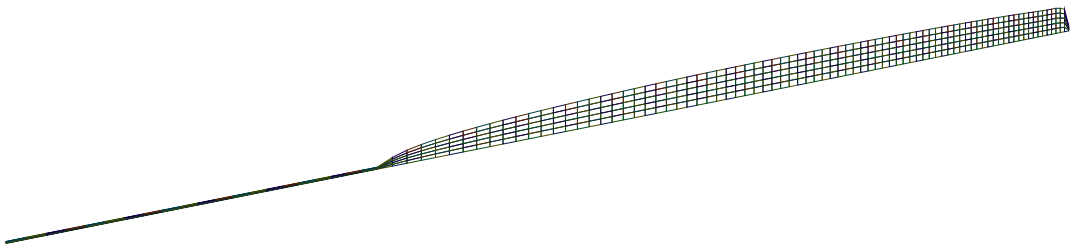


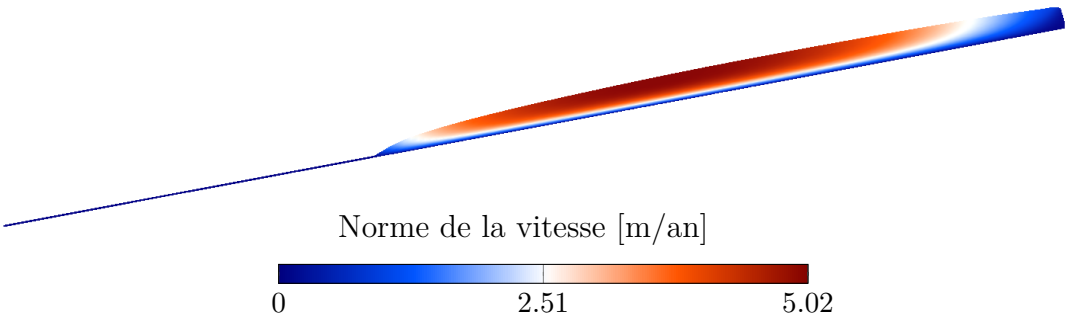
FIGURE 7.2.: Maillage structuré du glacier avec 515 quadrangles.

7.2. Résolution du problème mécanique

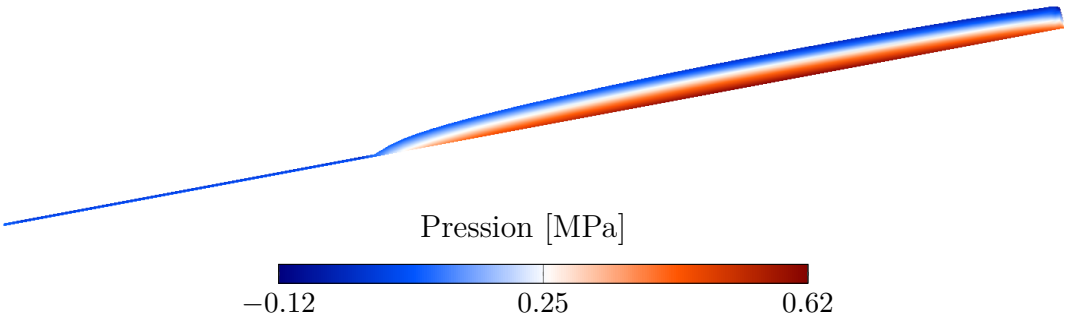
7.2.1. Problème mécanique sans glissement

Nous commençons tout d'abord par résoudre uniquement le problème de Stokes non linéaire en supposant que le glacier est gelé à sa base ($\mathbf{v} = \mathbf{0}$ sur Γ_b). La température de la glace est fixée à une valeur de 270.15 [K]. Le comportement rhéologique de la glace obéit à la loi de Glen régularisée avec un exposant égal à 3. La valeur de la contrainte résiduelle σ_0 a été fixée à 1×10^{-10} [MPa]. Nous présentons pour l'instant uniquement les résultats obtenus. L'analyse des méthodes de résolution de la non-linéarité du problème sera traitée à la sous-section 7.2.3.

Les figures 7.3(a) et 7.3(b) représentent respectivement la norme du vecteur vitesse $\mathbf{v}$ et la pression p dans la glace. Nous constatons que la vitesse de la glace est plus élevée au niveau de la surface libre du glacier. Le profil de pression est quant à lui de nature essentiellement hydrostatique.



(a) Norme de la vitesse.



(b) Pression dans la glace.

FIGURE 7.3.: Étude d'un glacier gelé à sa base.

Nous constatons sur la figure 7.3(b) que la pression dans la glace peut être négative au voisinage de l'interface air-glace. Pour en comprendre la raison, il convient de se remémorer la condition aux limites imposée à cette interface. Pour rappel, nous avions imposé une condition de surface libre, c'est-à-dire

$$\sigma \cdot \mathbf{n} = -p\,\mathbf{n} + 2\mu\,\dot{\varepsilon}(\mathbf{v}) \cdot \mathbf{n} = \mathbf{0}. \tag{7.1}$$

À deux dimensions, il est utile de considérer un système de coordonnées (τ,n) locales à l'interface et dont les vecteurs de base $(\boldsymbol{\tau},\mathbf{n})$ sont respectivement dirigés parallèlement et suivant la normale extérieure à cette surface. La condition 7.1 peut ainsi se réécrire en fonction des composantes normale v_n et tangentielle v_τ de la vitesse à l'interface sous la forme

$$-p + 2\mu\,\mathbf{n} \cdot \nabla v_n = 0, \tag{7.2a}$$
$$\mathbf{n} \cdot \nabla v_\tau + \boldsymbol{\tau} \cdot \nabla v_n = 0. \tag{7.2b}$$

La relation (7.2a) montre que si la dérivée directionnelle de la composante normale à l'interface dans la direction de la normale extérieure est négative, alors la pression doit être négative pour assurer l'absence de contrainte sur la surface du glacier. Nous pouvons vérifier qu'en effet le gradient normal de v_n à l'interface est bien négatif, ce qui permet de justifier la présence d'une pression négative.

7.2.2. Problème mécanique avec glissement

Nous envisageons à présent l'influence du glissement basal sur le déplacement du glacier. Pour rappel, le glissement basal est, avec la déformation par fluage, l'un des deux mécanismes principaux responsables du mouvement des glaciers. Pour notre simulation, nous considérons un

coefficient de glissement C_b indépendant de la température mais qui dépend de la position sur la base. Nous supposons que le coefficient de frottement vaut 0.01 [MPa an/m] pour des altitudes comprises entre 300 [m] et 400 [m] et 1000 [MPa an/m] ailleurs. De cette manière, le glacier peut soit glisser facilement soit difficilement en certains endroits de sa base.

Nous avons représenté sur la figure 7.4 la norme de la vitesse du glacier. En comparant cette figure à l'illustration 7.3(a), nous constatons une augmentation globale marquée de la vitesse du glacier. Cette figure illustre ainsi l'importance du glissement basal sur le mouvement de la glace. Vu l'importance de ce glissement pour certains glaciers, il est indéniable qu'une compréhension plus approfondie de ce phénomène ainsi que le développement de nouveaux modèles peuvent se révéler profitables pour la réalisation de simulations numériques plus précises.

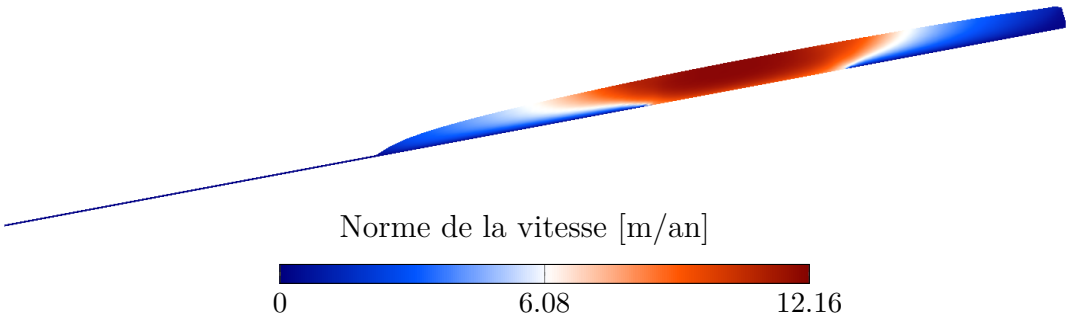


FIGURE 7.4.: Norme de la vitesse pour le glacier avec glissement.

7.2.3. Résolution de la non-linéarité du problème de Stokes

Cette sous-section a pour objectif de comparer les performances de quelques méthodes de résolution pour traiter la non-linéarité du problème de Stokes. Pour cela, nous reconsidérons le cas du glacier gelé à sa base de la sous-section 7.2.1. Ce problème a été résolu à l'aide de cinq méthodes différentes à savoir la méthode de Picard, la méthode de Newton, la méthode hybride avec $\gamma = 0.5$ ainsi que les méthodes de Broyden et de Newton initialisées par quelques itérations de la méthode de Picard. Pour chacune des cinq méthodes, nous avons initialisé les valeurs nodales pour la vitesse à 0.1 [m/an] et celles pour la pression à 0 [MPa]. La méthode de Broyden a quant à elle été initialisée avec la matrice jacobienne évaluée au premier itéré.

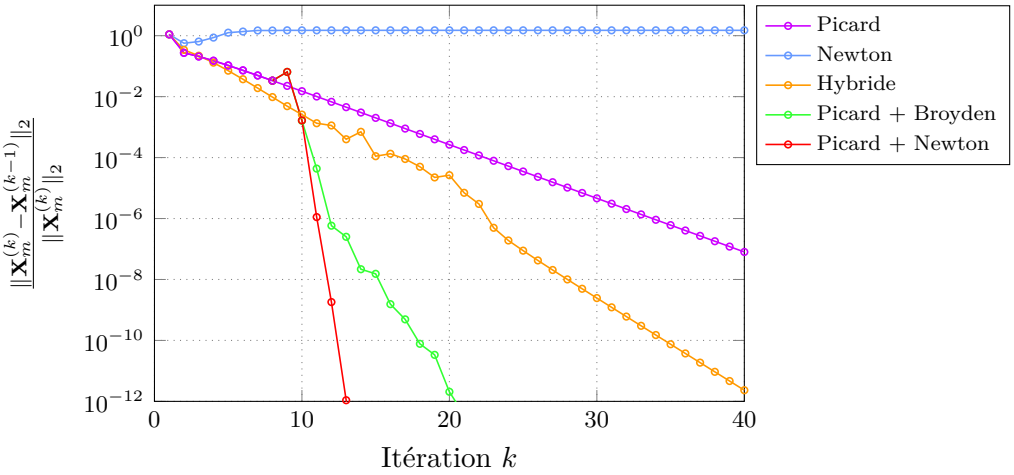
Pour comparer les différentes méthodes de résolution, nous avons décidé de représenter l'erreur relative entre deux itérés successifs à savoir

$$\frac{\|\mathbf{X}_m^{(k)} - \mathbf{X}_m^{(k-1)}\|_2}{\|\mathbf{X}_m^{(k)}\|_2}, \tag{7.3}$$

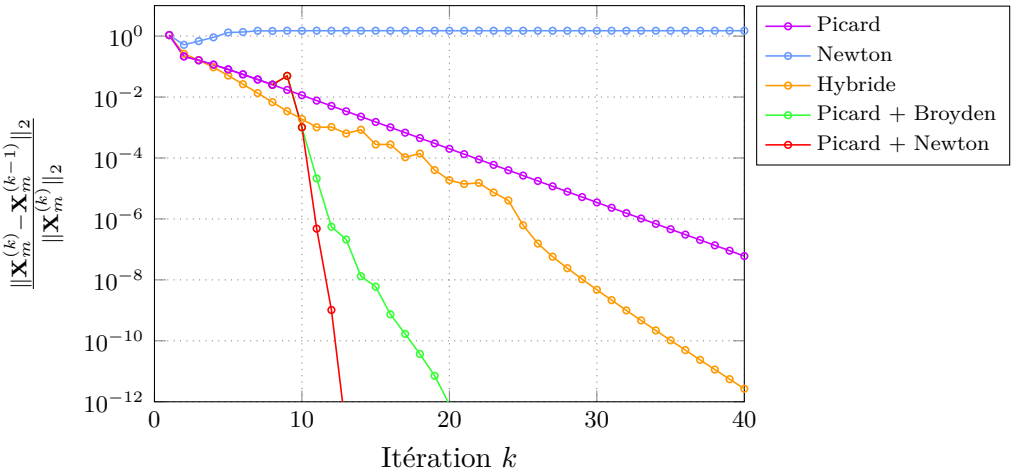
où les notations $\mathbf{X}_m$ et $\|\cdot\|_2$ désignent respectivement le couple mécanique $(\mathbf{V}, \mathbf{P})$ et la norme euclidienne traditionnelle. Nous avons choisi l'erreur relative comme critère de comparaison, car celle-ci constitue un choix possible pour définir un critère d'arrêt et donne de ce fait une indication du nombre d'itérations nécessaires pour obtenir une précision souhaitée sur la solution numérique. Un nombre d'itérations réduit signifie une diminution importante du temps de calcul si le temps pour réaliser une itération est similaire pour chacune des méthodes.

Notons également que d'autres critères de comparaison tels que la valeur du résidu ou l'écart absolu entre deux itérés successifs auraient également pu être considérés.

Les figures 7.5(a) et 7.5(b) représentent l'erreur relative entre deux itérés successifs pour un maillage constitué de respectivement 515 et 1030 quadrangles. Nous constatons pour les deux figures des résultats fort identiques, ce qui tend à montrer que les performances des méthodes de résolution sont indépendantes du maillage du domaine. Nous pouvons remarquer que pour l'itéré initial considéré, la méthode de Newton ne converge pas et la solution diverge. Au contraire, si la méthode de Newton est utilisée après quelques itérations de Picard ou est combinée avec la méthode de Picard, alors cette méthode converge bien. Ceci confirme que la méthode de Newton présente une robustesse assez faible et donc que son utilisation peut se révéler inefficace si elle est utilisée trop loin de la solution exacte. Pour que la méthode de Broyden présente de bonnes propriétés de convergence, il convient aussi de l'utiliser suffisamment proche de la solution exacte.



(a) Maillage avec 515 quadrangles.



(b) Maillage avec 1030 quadrangles.

FIGURE 7.5.: La figure compare l'erreur relative entre deux itérés successifs pour cinq méthodes de résolution numérique (méthode de Picard, méthode de Newton, méthode hybride ($\gamma = 0.5$) et méthodes de Broyden et de Newton initialisées avec quelques itérations de Picard).

La figure 7.5 fournit également une indication sur les vitesses de convergence des différentes méthodes. Nous observons que la méthode de Newton présente la meilleure vitesse de convergence quand elle est initialisée suffisamment proche de la solution exacte. Au contraire, la méthode de Picard possède la vitesse de convergence la plus faible. La méthode hybride qui combine à la fois la méthode de Picard et de Newton présente quant à elle une convergence intermédiaire à celle de ces deux méthodes. Nous constatons également que la méthode de Broyden constitue un choix intéressant, car ses performances se rapprochent de celles de la méthode de Newton tout en évitant un calcul explicite de la matrice jacobienne à toutes les itérations.

Pour juger plus précisément de la vitesse de convergence des différentes méthodes, il est en outre intéressant de représenter l'erreur absolue entre l'itéré courant et la solution exacte à savoir

$$\|\mathbf{X}_m^{(k)} - \mathbf{X}_m^*\|_2 \tag{7.4}$$

Comme la solution exacte du problème est inconnue, nous avons tout d'abord réalisé une première simulation qui nous a permis d'obtenir une estimation précise de la solution exacte. La figure 7.6 représente sur une échelle logarithmique l'erreur absolue en fonction du nombre d'itérations pour les méthodes de Picard, de Broyden et de Newton ainsi que pour la méthode hybride avec $\gamma = 0.5$. Les courbes pour les méthodes de Picard et hybride montrent une vitesse de convergence linéaire. Cette constatation est en accord avec la théorie en ce qui concerne la méthode de Picard. Pour les méthodes de Broyden et de Newton, les graphiques semblent suggérer une vitesse de convergence superlinéaire.

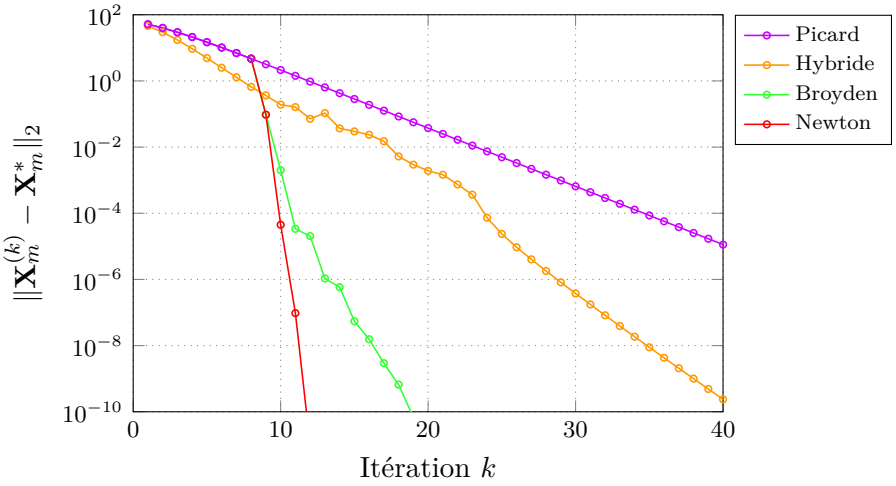


FIGURE 7.6.: La figure montre l'erreur absolue en fonction du nombre d'itérations pour quatre méthodes de résolution numérique (méthode de Picard, méthode hybride ($\gamma = 0.5$), méthodes de Broyden et de Newton précédées de quelques itérations de Picard).

7.3. Résolution du problème thermique

Nous envisageons à présent la résolution du problème thermique uniquement. Pour ce faire, nous supposons que le profil de température à la surface du glacier est donné par

$$T_s = 273.15 - 0.01h\,[\mathrm{K}], \tag{7.5}$$

où h représente l'altitude de la surface libre exprimée en mètres. Cela revient à considérer une diminution de la température de 1 [K] tous les 100 [m], ce qui est une valeur typique pour l'air sec [4]. Nous considérons que la valeur du flux géothermique à l'interface glace-roche est de 200 [mW/m^2]. Sur cette même frontière, nous supposons également que le glacier peut glisser sur la roche avec un coefficient de glissement qui dépend de la température suivant la relation [5]

$$C_b = 0.1 \exp(T_m - T) \, [\text{MPa an/m}]. \tag{7.6}$$

Comme nous nous limitons à résoudre le problème thermique uniquement, il est nécessaire de disposer de la valeur du champ de vitesse dans le glacier. Pour cela, nous avons résolu le problème mécanique en imposant une température intérieure au glacier égale à 270.15 [K] et en fixant la valeur du coefficient de glissement à 0.1[MPa an/m] vu que la température à la base du glacier sera probablement proche du point de fusion. À titre illustratif, nous avons repris sur la figure 7.7 la norme du champ de vitesse que nous utiliserons dans cette section.

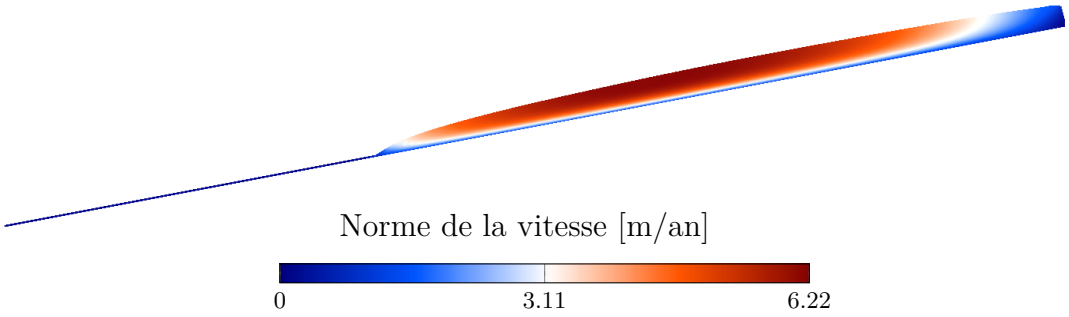


FIGURE 7.7.: Norme du champ de vitesse pour le problème thermique.

Les non-linéarités du problème ont été résolues au moyen de la méthode de Picard dans les sous-sections 7.3.1 et 7.3.2. Comme nous sommes uniquement intéressé par le résultat final de la simulation, une autre méthode numérique aurait pu être utilisée également. Soulignons enfin que pour résoudre le problème thermique, nous avons suivi l'hypothèse formulée au chapitre 6 à savoir que l'importance de la convection thermique demeurait faible pour le modèle « jouet » étudié. Cela signifie que nous n'avons pas rajouté de terme de stabilisation dans la formulation variationnelle de l'équation de la chaleur. La validité de cette hypothèse sera toutefois vérifiée à la sous-section 7.3.2.

7.3.1. Pénalisation du problème thermique

Le problème thermique est tout d'abord résolu sans terme de pénalisation afin de juger de la nécessité ou non de contraindre numériquement le problème. Le profil de température dans le glacier est repris à la figure 7.8. Sur l'échelle de couleur, nous avons indiqué en rouge la température maximale dans le glacier. Comme nous pouvons le constater, la température à la base du glacier excède en certains points la température de fusion de manière importante. Ce dépassement du point de fusion est notamment lié à la valeur élevée du flux géothermique que nous avons choisie. Notons aussi que la température minimale est de 267.62 [K] et que celle-ci est atteinte à l'altitude la plus élevée de la surface libre.

4. Dans l'hypothèse adiabatique, la diminution de la température avec l'altitude est de 10 [K/km] pour l'air sec et de 5.5 [K/km] pour un air saturé en vapeur d'eau. La valeur moyenne observée est de 6.5 [K/km] [62].

5. L'expression que nous employons est inspirée d'une relation pour le coefficient de friction de la calotte polaire antarctique [55].

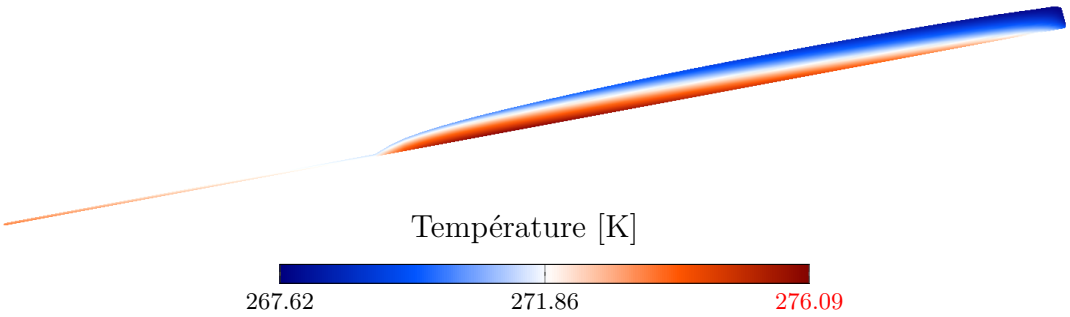


FIGURE 7.8.: Champ de température dans le glacier en l'absence de pénalisation.

Pour empêcher la température dans le glacier de dépasser son point de fusion, nous avons ajouté un terme de pénalisation quadratique à l'équation de la chaleur. La figure 7.9 illustre la solution du problème thermique pour un facteur de pénalisation ε égal à 10^{-7}. Nous avons utilisé la même échelle de couleur que sur la figure 7.8 afin de permettre une comparaison plus aisée de ces deux figures. La température maximale de la glace a de nouveau été indiquée en rouge. Nous voyons que le terme de pénalisation introduit permet effectivement d'obtenir une température inférieure ou proche du point de fusion. Nous avons également représenté sur la figure 7.10 l'écart de température entre le problème contraint (figure 7.9) et le problème non contraint (figure 7.8) afin de rendre plus manifeste l'effet de la pénalisation. Nous constatons que la contrainte thermique a induit un refroidissement quasi général. Le refroidissement est plus marqué au niveau du lit rocheux où la température excédait le point de fusion de manière importante.

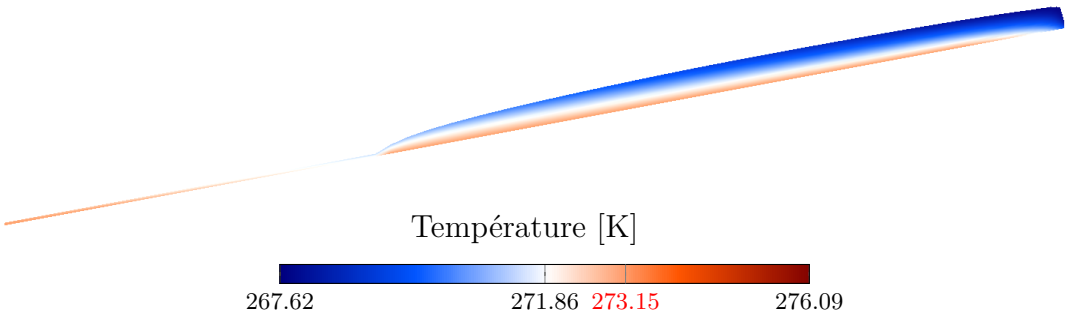


FIGURE 7.9.: Champ de température dans le glacier en présence d'un terme de pénalisation.

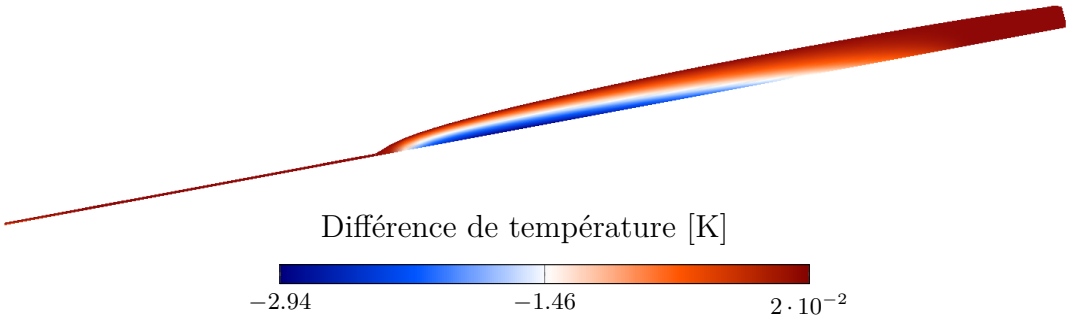


FIGURE 7.10.: La figure montre l'écart de température dans le glacier pour une résolution numérique avec prise en compte de la contrainte thermique (figure 7.9) et sans prise en compte de cette contrainte (figure 7.8).

À la section 6.3, nous avions introduit la fonction de pénalisation (6.8) qui dépendait d'un paramètre a. Nous pouvons juger l'influence de ce paramètre sur la convergence de la méthode numérique employée. Pour cela, nous avons représenté l'erreur relative entre deux solutions approchées successives en fonction du nombre d'itérations pour des valeurs de a égales à 1.6, 1.8 et 2. Le paramètre de pénalisation est choisi égal à 10^{-7}. Nous remarquons que la pénalisation quadratique présente la vitesse de convergence la plus faible, alors que cette vitesse augmente si la valeur de a diminue. Cela suggère que l'utilisation d'un terme de pénalisation quadratique n'est pas la plus adéquate. Nous avons ainsi décidé de choisir une valeur de a égale à 1.6 pour la suite de ce travail. Nous avons pu également constater dans le cas de la résolution du problème couplé que ce choix donnait de meilleurs résultats de convergence qu'un terme de pénalisation quadratique.

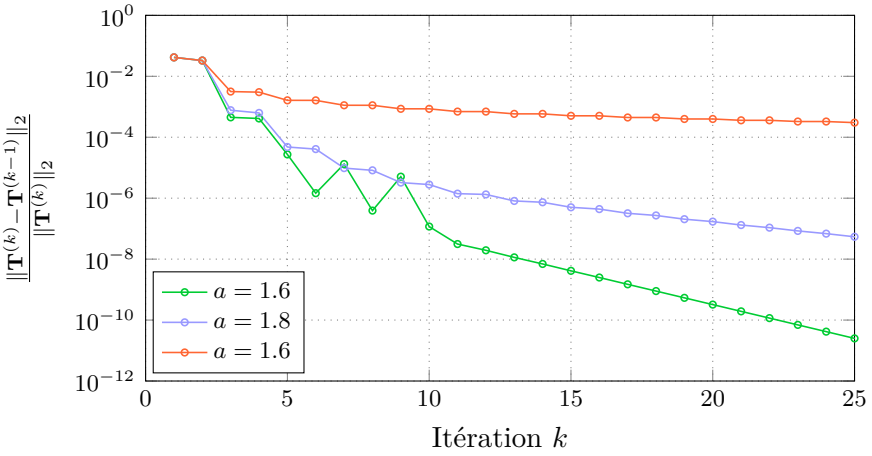


FIGURE 7.11.: Comparaison de trois types de pénalisation.

Nous avons voulu aussi donner un aperçu de l'influence du paramètre de pénalisation sur la solution numérique. Pour cela, nous avons indiqué dans le tableau 7.1 la température maximale de la solution obtenue pour différentes valeurs de ce paramètre. La première colonne de ce tableau indique la température maximale pour le problème non contraint à savoir 276.09 [K]. Nous constatons que l'ajout du terme de pénalisation permet de ramener la température maximale à une valeur proche de la température de fusion de la glace. Comme attendu, la pénalisation est d'autant plus efficace que le paramètre ε est petit. Toutefois, si ce paramètre est trop faible, nous constatons des problèmes d'ordre numérique. Le choix $\varepsilon = 10^{-7}$ que nous adopterons pour la suite semble permettre de satisfaire correctement la contrainte thermique.

ε	\	10^{-2}	10^{-4}	10^{-6}	10^{-7}
$T_{\max}$ [K]	276.09	273.315	273.162	273.151	273.15

TABLE 7.1.: Température maximale dans le glacier en fonction du paramètre de pénalisation.

7.3.2. Importance de la convection dans le modèle thermique

Nous allons à présent vérifier si l'hypothèse qui consiste à ne pas ajouter un terme de stabilisation dans le problème thermique est fondée. Pour cela, nous avons résolu le problème thermique pénalisé une première fois en utilisant la méthode SUPG et une seconde fois sans ajouter de terme de stabilisation. Nous avons pour chacun des deux cas pris la même solution

initiale et réalisé trente itérations de la méthode de Picard. La figure 7.12 représente l'écart de température entre la méthode stabilisée et la méthode non stabilisée. Nous pouvons constater que l'écart entre les deux solutions reste relativement faible et donc que l'utilisation d'une méthode non stabilisée constitue une hypothèse acceptable. Nous pouvons aussi observer que l'écart est le plus important au niveau du rétrécissement brusque de l'épaisseur du glacier là où le gradient de vitesse est le plus important (voir figure 7.7).

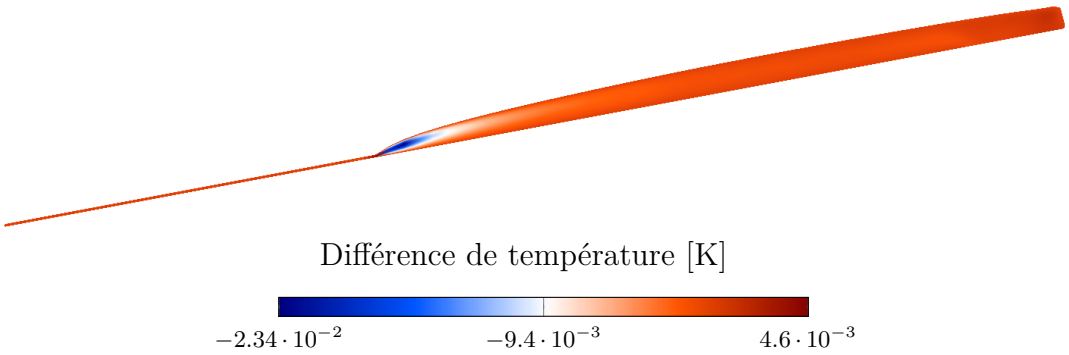


FIGURE 7.12.: La figure montre l'écart de température dans le glacier pour le modèle thermique résolu avec un terme de stabilisation (méthode SUPG) et sans terme de stabilisation.

7.3.3. Résolution de la non-linéarité du problème thermique

Comme pour le problème de Stokes, nous nous proposons à présent de comparer différentes méthodes de résolution numérique pour résoudre le problème thermique non linéaire pénalisé. Pour cela, nous avons résolu ce problème en utilisant les méthodes de Picard et de Newton ainsi que les méthodes hybride ($\gamma = 0.5$) et de Broyden. Dans chacun des cas, nous avons initialisé la solution à 268.15 [K]. La méthode de Broyden est utilisée après quelques itérations de la méthode de Picard afin de garantir sa convergence. La figure 7.13 reprend l'erreur relative entre deux itérés successifs en fonction du nombre d'itérations.

Nous constatons sur la figure 7.13 que chacune des méthodes employées converge pour l'itéré initial considéré. Cependant, il apparaît que la méthode de Picard présente une convergence bien plus faible que pour les trois autres. Toutefois, alors que nous pourrions nous attendre à ce que la méthode de Newton converge le plus rapidement, nous constatons qu'en fait ce sont les méthodes hybride et de Broyden qui présentent la meilleure convergence sur une grande partie des itérations considérées. En fait, la méthode de Newton se révèle peu efficace dans les premières itérations pour lesquelles la solution numérique reste encore fort éloignée de la solution exacte.

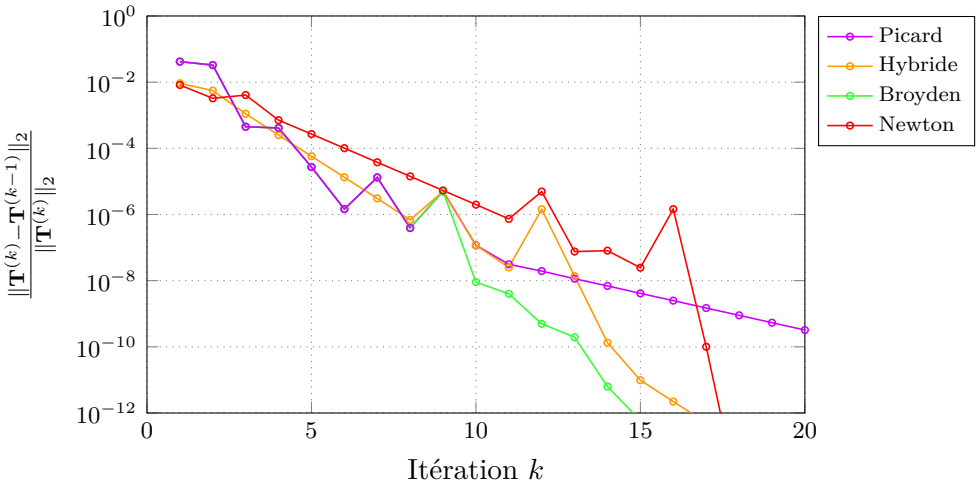


FIGURE 7.13.: Comparaison de l'erreur relative entre deux itérés successifs pour quatre méthodes de résolution numérique (méthode de Picard, méthode hybride ($\gamma = 0.5$), méthode de Broyden et méthode de Newton).

Afin de mieux se faire une idée de la vitesse de convergence des différentes méthodes, nous avons représenté sur la figure 7.14 l'erreur absolue en fonction du nombre d'itérations pour le problème pénalisé. Nous constatons que le comportement des différentes méthodes est fort différent de celui obtenu pour le problème mécanique (voir figure 7.6). Les conclusions que nous pouvons tirer de ce graphique sont assez semblables à celles de la figure 7.13. Nous pouvons néanmoins observer la présence d'un seuil au-delà duquel la vitesse de convergence de la méthode de Newton augmente sensiblement, ce qui confirme une fois encore que cette méthode est efficace suffisamment proche de la solution.

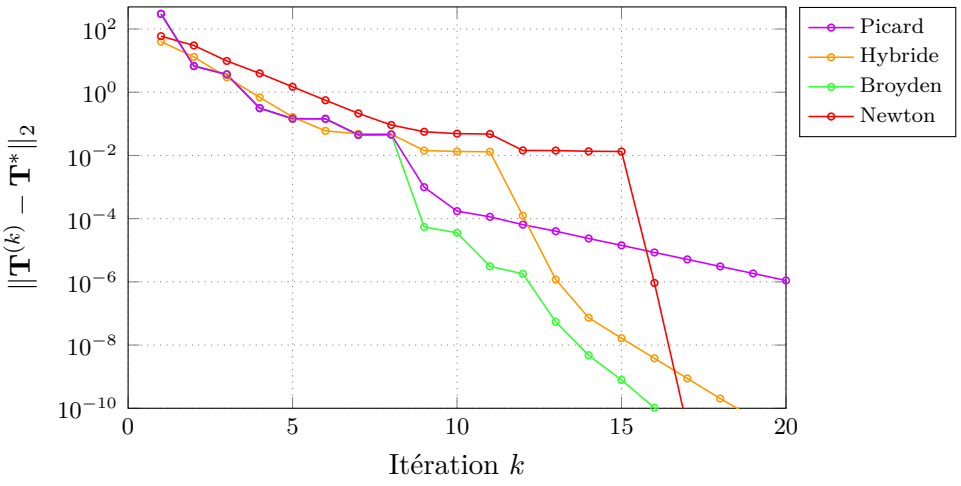


FIGURE 7.14.: Comparaison de l'erreur absolue en fonction du nombre d'itérations pour quatre méthodes de résolution numérique appliquées au problème thermique pénalisé.

Lors de nos simulations, nous avons pu constater que l'ajout d'un terme de pénalisation semblait modifier le comportement des méthodes numériques. Afin de mieux évaluer l'influence de la pénalisation, nous avons représenté sur la figure 7.15 l'erreur absolue pour chacune des quatre méthodes utilisées pour le problème sans terme de pénalisation. Nous observons dans

ce cas une convergence bien plus rapide de l'ensemble des méthodes et qui est plus en accord avec les résultats théoriques. La méthode de Broyden présente toutefois une convergence bien plus lente que les trois autres méthodes, ce qui peut se justifier notamment par un mauvais choix pour l'initialisation de cette méthode.

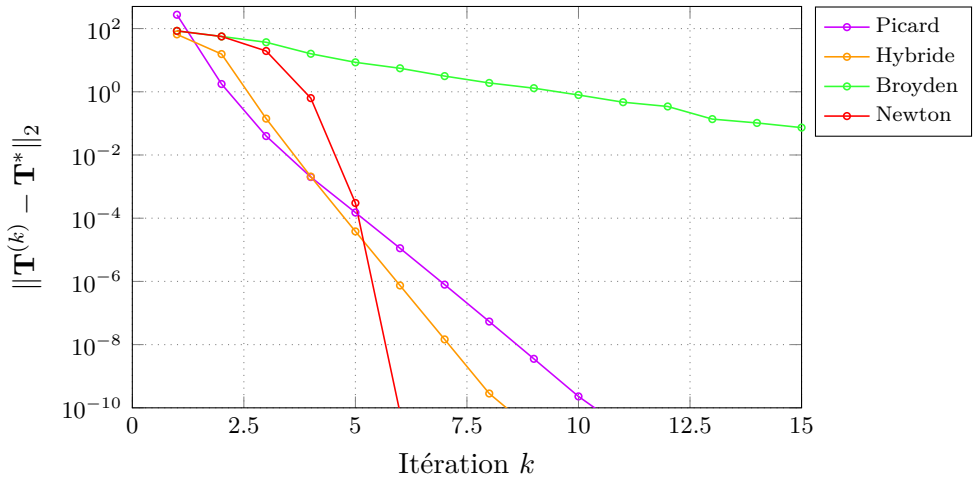


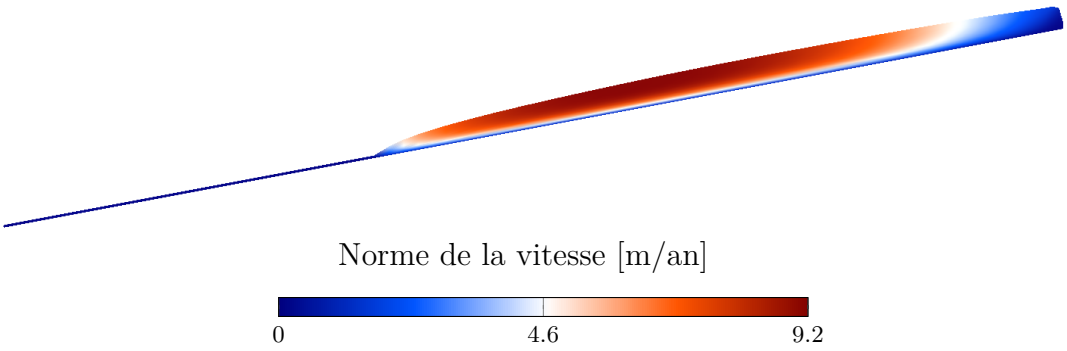
FIGURE 7.15.: Comparaison de l'erreur absolue en fonction du nombre d'itérations pour quatre méthodes numériques appliquées au problème thermique non pénalisé.

7.4. Résolution du problème couplé

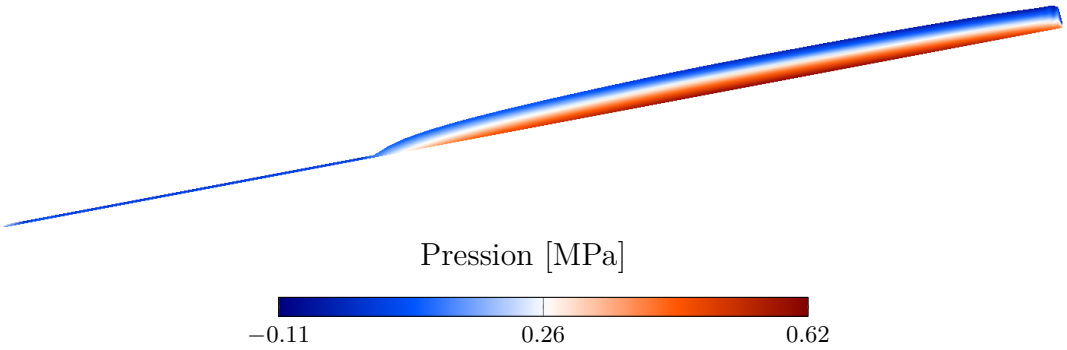
7.4.1. Solution du problème couplé

Nous envisageons à présent la résolution du problème thermomécanique couplé. Nous présentons tout d'abord les résultats numériques avant de s'intéresser plus précisément à différentes méthodes de résolution numérique.

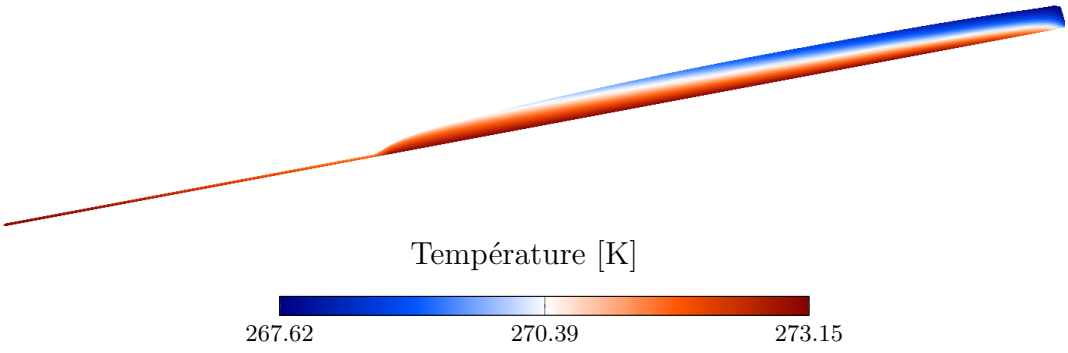
Nous avons représenté sur la figure 7.16 les champs de vitesse, de pression et de température pour le problème couplé. Nous constatons sur les figures 7.16(b) et 7.16(c) que les solutions pour la pression et la température semblent avoir peu changé par rapport aux problèmes mécanique et thermique considérés individuellement. Par contre, nous observons en comparant les figures 7.7 et 7.16(a) une variation marquée de la norme de la vitesse en certains endroits du glacier. Pour rappel, nous avions réalisé la figure 7.7 à une température constante de 270.15 [K]. Il apparaît dès lors que la température de la glace peut avoir une forte influence sur la vitesse d'avancement du glacier. Cela s'explique par la dépendance importante de la viscosité de la glace avec la température. Ainsi, la connaissance précise de la température d'un glacier est un élément important si nous souhaitons réaliser des simulations numériques précises.



(a) Norme de la vitesse.



(b) Pression dans la glace.



(c) Température de la glace.

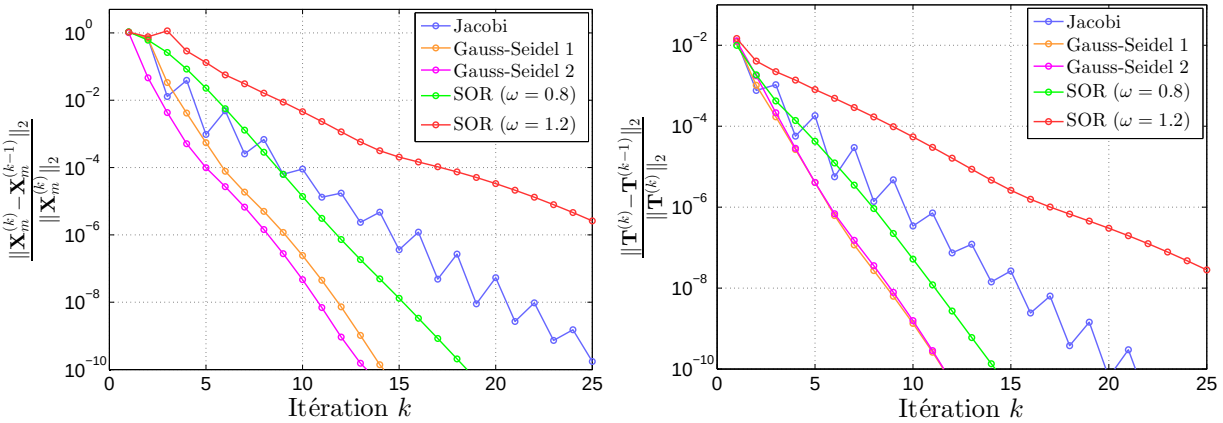
FIGURE 7.16.: Résultats numériques pour le problème couplé.

7.4.2. Comparaison des méthodes numériques faiblement couplées

Nous avons commencé par résoudre le problème couplé en considérant des méthodes numériques faiblement couplées. Les méthodes que nous avons utilisées sont les méthodes de Jacobi, de Gauss-Seidel et de surrelaxation successive avec un paramètre de surrelaxation ω égal à 0.8 et 1.2. Pour les méthodes de Gauss-Seidel et SOR, l'ordre dans lequel sont résolus les problèmes monophysiques a en principe de l'importance sur le comportement de la méthode. Nous avons ainsi décidé pour la méthode de Gauss-Seidel de considérer un premier cas pour lequel le problème mécanique est résolu avant le problème thermique et inversement pour le second cas. Pour les méthodes SOR, le problème mécanique est toujours résolu avant le

problème thermique. Notons aussi que pour les méthodes SOR, le paramètre de surrelaxation a été choisi de manière à assurer la convergence de la méthode.

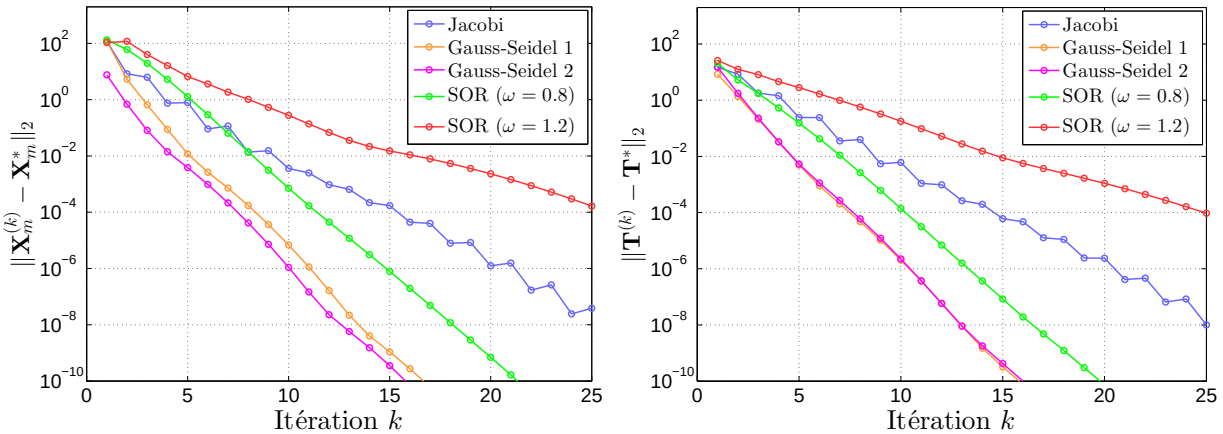
Les figures 7.17 et 7.18 comparent respectivement l'erreur relative entre deux itérés successifs et l'erreur absolue pour chacune des différentes méthodes utilisées. Pour obtenir ces résultats, nous avons résolu chacun des problèmes monophysiques en réalisant dix itérations de la méthode de Picard.



(a) Erreur pour la partie mécanique.

(b) Erreur pour la partie thermique.

FIGURE 7.17.: Comparaison de l'erreur relative entre deux itérés successifs pour quelques méthodes de résolution faiblement couplées.



(a) Erreur pour la partie mécanique.

(b) Erreur pour la partie thermique.

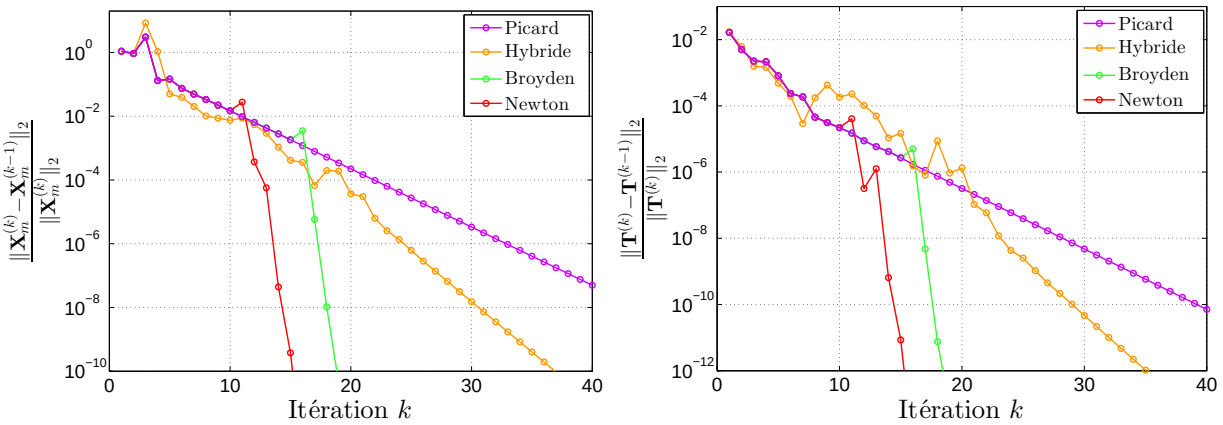
FIGURE 7.18.: Comparaison de l'erreur absolue pour quelques méthodes de résolution faiblement couplées.

Bien que l'ensemble des méthodes semblent converger vers la solution exacte, les vitesses de convergence varient d'une méthode à l'autre. Il apparaît que les méthodes de Gauss-Seidel ont la meilleure vitesse de convergence tandis que la méthode SOR avec $\omega = 1.2$ présente une convergence un peu moins rapide que les autres méthodes. Nous constatons également que l'ordre dans lequel les problèmes monophysiques sont résolus influence peu les résultats pour la méthode de Gauss-Seidel. Les courbes de convergence pour la méthode de Jacobi montrent un

comportement légèrement oscillant qui peut s'expliquer par le fait que cette méthode résout les problèmes monophysiques en utilisant toujours la solution à l'itéré précédent.

7.4.3. Comparaison des méthodes numériques fortement couplées

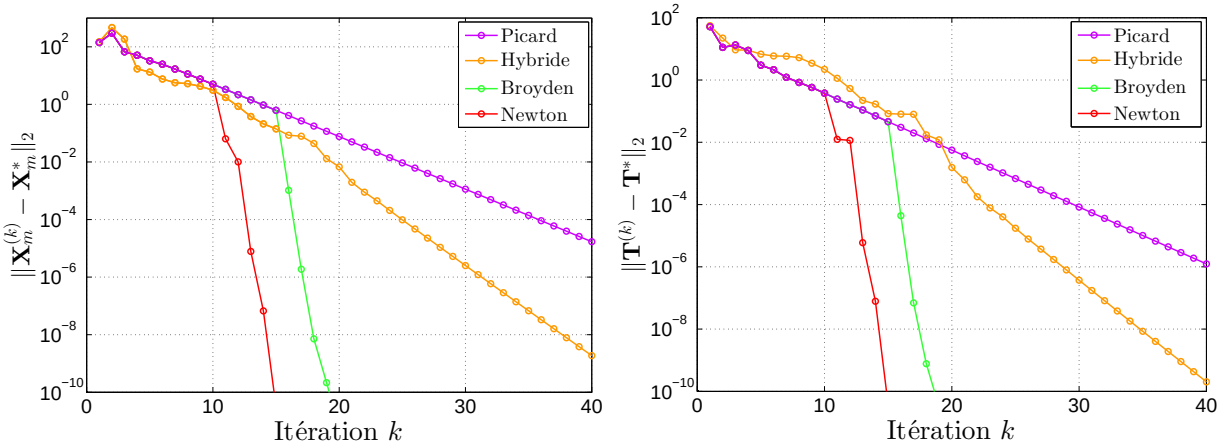
Nous envisageons enfin la résolution numérique du problème couplé par des méthodes fortement couplées. Pour cela, nous avons résolu ce problème en utilisant la méthode de Picard, la méthode hybride avec $\gamma = 0.5$ et les méthodes de Newton et de Broyden qui ont été initialisées par quelques itérations de la méthode de Picard afin d'en assurer la convergence. Les résultats de convergence pour l'erreur relative entre deux itérés successifs et l'erreur absolue sont repris aux figures 7.19 et 7.20.



(a) Erreur pour la partie mécanique.

(b) Erreur pour la partie thermique.

FIGURE 7.19.: Comparaison de l'erreur relative entre deux itérés successifs pour quelques méthodes de résolution fortement couplées.



(a) Erreur pour la partie mécanique.

(b) Erreur pour la partie thermique.

FIGURE 7.20.: Comparaison de l'erreur absolue pour quelques méthodes de résolution fortement couplées.

Si nous comparons tout d'abord les différentes méthodes couplées, nous observons que les résultats de convergence sont en accord avec ce qui est attendu théoriquement. La méthode de Newton présente la vitesse de convergence la plus élevée tandis que la méthode de Picard a

la vitesse de convergence la plus faible. Les méthodes hybride et de Broyden ont un comportement intermédiaire à ces deux méthodes. Les méthodes de Picard et hybride montrent une vitesse de convergence linéaire tandis que les méthodes de Newton et de Broyden semblent présenter une convergence superlinéaire.

Pour terminer, nous pouvons comparer les résultats obtenus pour les méthodes faiblement et fortement couplées. Bien que les deux types de méthodes convergent vers la solution exacte, il apparaît que les méthodes fortement couplées ont une convergence bien plus élevée que les méthodes faiblement couplées en terme du nombre d'itérations. À chaque itération des méthodes faiblement couplées correspond en réalité vingt sous-itérations pour résoudre les deux problèmes monophysiques. Or, nous observons que pour une quarantaine d'itérations des méthodes fortement couplées, nous obtenons déjà une très bonne précision de la solution numérique. Bien qu'à chaque itération d'une méthode fortement couplée il soit nécessaire de résoudre un système linéaire de plus grande dimension, le nombre réduit d'itérations permet en fait de diminuer de manière importante le coût de calcul par rapport aux méthodes faiblement couplées. Notre exemple montre que l'utilisation de méthodes fortement couplées peut se révéler un choix opportun pour résoudre des problèmes couplés malgré que ces méthodes soient plus difficiles à implémenter.

8. Perspectives d'approfondissement

Dans ce travail de fin d'études, nous nous sommes attaché à donner une vue d'ensemble aussi générale que possible des problèmes tant mathématiques que numériques qui pouvaient survenir lors de la résolution d'un problème thermomécanique en glaciologie. Pour ce faire, nous avons étudié tant les problèmes monophysiques que le problème couplé. Il est évident que de nombreux points que nous avons abordés auraient mérité une attention plus soutenue. Dès lors, avant de conclure ce travail, nous souhaitons établir une liste non exhaustive de quelques possibilités d'approfondissement de ce mémoire qui présenteraient à la fois un intérêt pour le domaine de la glaciologie à proprement parler, mais également pour le domaine de la multiphysique dans son ensemble.

- Au chapitre 3, nous avons présenté différentes méthodes pour la résolution du problème de Stokes. Par la suite, nous avons principalement considéré l'utilisation d'une paire particulière d'éléments finis compatibles. Il pourrait être intéressant d'étudier et de comparer numériquement sur un modèle de glacier l'efficacité de différentes paires compatibles ainsi que l'utilisation de méthodes de stabilisation comme celle proposée par Bochev *et al.* [6].

- Au chapitre 4, nous avons proposé deux approches principales pour résoudre la contrainte thermique à savoir les méthodes de pénalisation et duale. Nous avons limité notre étude de l'approche duale à une approche purement théorique. Une implémentation numérique concrète de cette méthode ainsi qu'une comparaison avec les méthodes de pénalisation pourraient constituer un approfondissement intéressant de ce travail. Les méthodes duales constituent une approche encore peu développée dans le domaine de la glaciologie. Elles pourraient néanmoins représenter de nouvelles pistes de recherches.

- Nous avons introduit au chapitre 5 quelques méthodes de résolution pour les problèmes multiphysiques. Nous avons présenté essentiellement les méthodes traditionnellement utilisées dans le domaine de la simulation multiphysique à savoir les méthodes de Jacobi, de Gauss-Seidel, de Picard et de Newton. Une étude et une analyse des nouvelles méthodes numériques multiphysiques telles que la méthode JFNK que nous avons présentée théoriquement peuvent représenter une opportunité pour le développement de codes de calcul plus efficaces. En particulier, l'étude du préconditionnement du système pour la méthode JFNK peut constituer un sujet intéressant d'approfondissement.

- Pour étudier notre glacier modèle, nous avons formulé l'hypothèse que le taux de fonte basale pouvait être négligé dans l'expression des conditions aux limites pour la vitesse. Bien que cette hypothèse puisse se justifier en première approximation, la réalisation de simulations numériques précises nécessite de prendre en compte explicitement cette quantité. Plusieurs approches sont envisageables pour traiter la fonte de la glace. Les modèles actuels s'orientent vers des modèles enthalpiques comme étudiés notamment par Aschwanden *et al.* [2] ainsi que Kleiner *et al.* [44]. Une autre possibilité pourrait être offerte par l'utilisation de méthodes duales pour le problème thermique qui permettrait d'introduire le taux de fonte basale comme inconnue dans le problème.

Conclusion

Ce travail de fin d'études a eu pour objectif de mettre en évidence les défis et les opportunités offerts par la modélisation multiphysique. Ce mémoire a été orienté plus spécifiquement vers une analyse théorique et numérique de l'écoulement des glaciers. La dynamique des glaciers est décrite convenablement à partir du couplage d'un problème mécanique et d'un problème thermique. Notre intention lors de la réalisation de ce mémoire a été de proposer une étude de l'écoulement des glaciers en régime stationnaire sous l'angle de la modélisation multiphysique. Ce travail présente de ce fait une synthèse des bases physiques, mathématiques et numériques à partir desquelles de futurs développements dans le domaine de la multiphysique pourraient être réalisés.

Ce mémoire a été abordé en premier sous l'aspect de la description physique de l'écoulement de la glace dans le cadre général des milieux continus. Cette théorie conduit tout naturellement à l'établissement d'un système d'équations thermomécaniques pour la glace, ce qui souligne la nature multiphysique intrinsèque des écoulements glaciaires. Ce travail s'est poursuivi ensuite par une étude théorique des problèmes mécanique et thermique individuels. Cela nous a permis de mettre en évidence les difficultés numériques liées à la résolution de chacun de ces deux problèmes. Nous avons vu en particulier que chacun des problèmes était caractérisé par une contrainte mécanique ou thermique qui en compliquait la résolution. La fin de ce mémoire a été consacrée à la résolution du problème couplé pour un modèle de glacier alpin. Cette dernière partie a offert une introduction aux méthodes numériques faiblement et fortement couplées pour résoudre des problèmes multiphysiques. Nous avons pu constater à partir de simulations numériques que les méthodes fortement couplées offraient une approche attrayante pour résoudre efficacement les équations thermomécaniques pour les glaciers bien que leur implémentation numérique soit plus compliquée.

Malgré le cadre spécifique dans lequel nous avons abordé ce travail, l'approche méthodologique présentée se révèle en réalité très globale et peut se généraliser à d'autres types de problèmes multiphysiques. En effet, les problèmes multiphysiques s'inscrivent naturellement à l'intersection des approches physique, mathématique et numérique. Ainsi, dans un premier temps, il convient de mettre en évidence l'ensemble des processus physiques qui gouvernent le comportement d'un système complexe. Une étude plus avancée de l'ensemble des phénomènes monophysiques peut ensuite se révéler profitable avant d'aborder le couplage entre ces processus. Enfin, le couplage entre les différentes composantes physiques du système doit être introduit afin de représenter correctement la réalité.

La connaissance de l'évolution des glaciers est appelée dans les prochaines années à connaître un intérêt sans cesse croissant aussi bien de la part de la communauté scientifique que du grand public. La réalisation de simulations numériques de plus en plus précises sur des glaciers de dimensions parfois très importantes nécessitera l'utilisation de ressources informatiques de plus en plus puissantes ainsi que le développement de nouveaux algorithmes de résolution plus performants. Parmi ceux-ci, les méthodes fortement couplées préconditionnées présentent un attrait tout particulier et leur étude théorique ainsi que leur application pourraient constituer une suite naturelle à ce travail de fin d'études.

Annexes

Annexe A.

Problème de minimisation pour le problème de Stokes

Nous montrons dans cette annexe que la résolution de la formulation variationnelle 3.2 est équivalente à la résolution du problème de minimisation (3.15). La démonstration que nous présentons est inspirée de Fortin et Garon [20].

Démonstration.

La preuve procède en deux parties. Dans la première partie, nous montrons que si $\mathbf{v}$ satisfait la formulation variationnelle 3.2, alors $\mathbf{v}$ est une solution du problème de minimisation (3.15). Dans la seconde partie, nous montrons que l'inverse est également valable.

- Pour montrer la première partie de la preuve, nous considérons tout d'abord un élément $\mathbf{w}$ quelconque de V_{div}. Cet élément peut s'écrire sous la forme $\mathbf{w} = \mathbf{v} + (\mathbf{w} - \mathbf{v}) = \mathbf{v} + \mathbf{w}_1$ avec $\mathbf{w}_1$ qui appartient à V_{div}.

  Nous pouvons écrire

  $$J(\mathbf{w}) = J(\mathbf{v} + \mathbf{w}_1) = \frac{1}{2} a(\mathbf{v} + \mathbf{w}_1, \mathbf{v} + \mathbf{w}_1) - (\mathbf{f}, \mathbf{v} + \mathbf{w}_1). \tag{A.1}$$

  En utilisant la bilinéarité de a et la linéarité de l'application $(\mathbf{f}, \cdot)$, nous obtenons

  $$J(\mathbf{w}) = \frac{1}{2} \left[a(\mathbf{v}, \mathbf{v}) + a(\mathbf{v}, \mathbf{w}_1) + a(\mathbf{w}_1, \mathbf{v}) + a(\mathbf{w}_1, \mathbf{w}_1) \right] - (\mathbf{f}, \mathbf{v}) - (\mathbf{f}, \mathbf{w}_1). \tag{A.2}$$

  Comme a est une forme symétrique, nous obtenons alors

  $$\begin{aligned} J(\mathbf{w}) = &\left(\frac{1}{2} a(\mathbf{v}, \mathbf{v}) - (\mathbf{f}, \mathbf{v}) \right) + \left(a(\mathbf{v}, \mathbf{w}_1) - (\mathbf{f}, \mathbf{w}_1) \right) + \frac{1}{2} a(\mathbf{w}_1, \mathbf{w}_1) \\ =&J(\mathbf{v}) + \frac{1}{2} a(\mathbf{w}_1, \mathbf{w}_1) \\ \geqslant&J(\mathbf{v}). \end{aligned} \tag{A.3}$$

  Pour obtenir le résultat final, nous avons utilisé le fait que $\mathbf{v}$ satisfait la formulation variationnelle 3.2 pour annuler l'expression $\left(a(\mathbf{v}, \mathbf{w}_1) - (\mathbf{f}, \mathbf{w}_1) \right)$. Nous avons également exploité la positivité de la forme a à savoir $a(\mathbf{w}_1, \mathbf{w}_1) \geqslant 0$, $\forall \mathbf{w}_1 \in V_{\text{div}}$.

  Comme $\mathbf{w}$ est quelconque, la fonction $\mathbf{v}$ minimise bien la fonctionnelle $J(\mathbf{w})\, \forall \mathbf{w} \in V_{\text{div}}$.

- Pour montrer la seconde partie de cette preuve, nous introduisons une fonction g de la variable réelle ε telle que

$$g(\varepsilon) = J(\mathbf{v} + \varepsilon\mathbf{w}) = \frac{1}{2}a(\mathbf{v} + \varepsilon\mathbf{w}, \mathbf{v} + \varepsilon\mathbf{w}) - (\mathbf{f}, \mathbf{v} + \varepsilon\mathbf{w}), \tag{A.4}$$

où $\mathbf{v}$ est la solution du problème de minimisation (3.15) et $\mathbf{w}$ est une fonction de V_{div}.

En utilisant la bilinéarité de a et la linéarité de l'application $(\mathbf{f}, \cdot)$, nous pouvons écrire

$$g(\varepsilon) = \frac{1}{2}a(\mathbf{v}, \mathbf{v}) - (\mathbf{f}, \mathbf{v}) + \varepsilon\left(a(\mathbf{v}, \mathbf{w}) - (\mathbf{f}, \mathbf{w})\right) + \frac{\varepsilon^2}{2}a(\mathbf{w}, \mathbf{w}). \tag{A.5}$$

Comme $\mathbf{v}$ est solution du problème de minimisation, g doit posséder un minimum en $\varepsilon = 0$, ce qui implique que $g'(0) = 0$ pour tout $\mathbf{w} \in V_{\text{div}}$. Or nous avons $g'(\varepsilon) = \left(a(\mathbf{v}, \mathbf{w}) - (\mathbf{f}, \mathbf{w})\right) + \varepsilon a(\mathbf{w}, \mathbf{w})$ et $g'(0) = a(\mathbf{v}, \mathbf{w}) - (\mathbf{f}, \mathbf{w})$. Comme la dernière expression doit être nulle quel que soit $\mathbf{w}$ dans V_{div}, cela implique que $\mathbf{v}$ satisfait bien la formulation variationnelle 3.2.

□

Annexe B.

Application du lemme de Lax-Milgram à l'équation de la chaleur

Dans cette annexe, nous vérifions que les hypothèses du lemme de Lax-Milgram 3.1 sont bien satisfaites pour la formulation variationnelle 4.1 de l'équation de la chaleur. Le raisonnement que nous tenons est basé sur une démonstration de Quarteroni [56] pour des conditions aux limites de Dirichlet homogènes sur Γ. Nous étendons cette preuve à des conditions aux limites mixtes sur la frontière de Ω.

Afin d'appliquer le théorème de Lax-Milgram, nous devons tout d'abord démontrer la coercivité de la forme g à savoir l'existence d'une constante $\alpha > 0$ telle que

$$g(w, w) \geqslant \alpha \, \|w\|^2_{\mathbb{H}^1(\Omega)} \quad \forall w \in V \subset \mathbb{H}^1(\Omega). \tag{B.1}$$

Pour montrer la coercivité de la forme g, nous décomposons g en ses différentes composantes.

Nous commençons d'abord par regarder au terme diffusif. Par définition de $\|\nabla w\|^2_{\mathbb{L}^2(\Omega)}$, nous avons [1]

$$\int_\Omega \kappa \, \nabla w \cdot \nabla w \, d\Omega = \kappa \, \|\nabla w\|^2_{\mathbb{L}^2(\Omega)}. \tag{B.2}$$

De plus, en vertu de l'inégalité de Poincaré [58], nous savons qu'il existe une constante C_Ω indépendante de w telle que

$$\|w\|_{\mathbb{L}^2(\Omega)} \leqslant C_\Omega \, \|\nabla w\|_{\mathbb{L}^2(\Omega)} \quad \forall w \in V. \tag{B.3}$$

Dès lors

$$\|w\|^2_{\mathbb{H}^1(\Omega)} = \|w\|^2_{\mathbb{L}^2(\Omega)} + \|\nabla w\|^2_{\mathbb{L}^2(\Omega)} \leqslant (1 + C_\Omega^2) \, \|\nabla w\|^2_{\mathbb{L}^2(\Omega)}. \tag{B.4}$$

En tenant compte des relations (B.2) et (B.4), nous avons ainsi

$$\int_\Omega \kappa \, \nabla w \cdot \nabla w \, d\Omega \geqslant \frac{\kappa}{1 + C_\Omega^2} \, \|w\|^2_{\mathbb{H}^1(\Omega)}. \tag{B.5}$$

En ce qui concerne le terme convectif, nous pouvons écrire

$$\begin{aligned} \int_\Omega w \, \mathbf{v} \cdot \nabla w \, d\Omega &= \frac{1}{2} \int_\Omega \mathbf{v} \cdot \nabla w^2 \, d\Omega \\ &= -\frac{1}{2} \int_\Omega w^2 \, \operatorname{div} \mathbf{v} \, d\Omega + \frac{1}{2} \int_\Gamma \mathbf{v} \cdot \mathbf{n} \, w^2 \, d\Gamma \\ &= \frac{1}{2} \int_{\Gamma_N} \mathbf{v} \cdot \mathbf{n} \, w^2 \, d\Gamma_N \\ &\geqslant 0. \end{aligned} \tag{B.6}$$

1. La démonstration est explicitée pour une valeur de κ constante sur Ω. Toutefois, si κ dépendait de la position, nous aurions $\int_\Omega \kappa \, \nabla w \cdot \nabla w \, d\Omega \geqslant \kappa_0 \, \|\nabla w\|^2_{\mathbb{L}^2(\Omega)}$ avec $\kappa \geqslant \kappa_0$ où κ_0 est une constante strictement positive. La preuve de la coercivité de g serait dans ce cas toujours d'application bien que légèrement différente.

Pour parvenir à la dernière inéquality, nous avons exploité le fait que $w = 0$ sur Γ_D ainsi que les hypothèses $\operatorname{div} \mathbf{v} = 0$ dans Ω et $\mathbf{v} \cdot \mathbf{n} \geqslant 0$ sur Γ_N. Dès lors, en posant α égal à $\kappa/(1 + C_\Omega^2)$, nous retrouvons bien la relation (B.1) et donc la coercivité de g.

Pour utiliser le lemme de Lax-Milgram, il faut également que la forme g soit continue, c'est-à-dire qu'il existe une constante $M > 0$ telle que

$$|g(u, w)| \leqslant M \, \|u\|_{\mathbb{H}^1(\Omega)} \, \|w\|_{\mathbb{H}^1(\Omega)} \quad \forall u, w \in V. \tag{B.7}$$

Pour montrer la continuité de g, nous pouvons procéder ainsi. Nous utilisons tout d'abord l'inégalité de Cauchy-Schwarz ce qui nous permet d'écrire

$$\begin{aligned} \left| \int_\Omega \kappa \, \nabla u \cdot \nabla w \, d\Omega \right| &\leqslant \kappa \, \|\nabla u\|_{\mathbb{L}^2(\Omega)} \, \|\nabla w\|_{\mathbb{L}^2(\Omega)} \\ &\leqslant \kappa \, \|u\|_{\mathbb{H}^1(\Omega)} \, \|w\|_{\mathbb{H}^1(\Omega)}. \end{aligned} \tag{B.8}$$

Nous exploitons ensuite l'inégalité de Hölder pour écrire

$$\begin{aligned} \left| \int_\Omega w \, \mathbf{v} \cdot \nabla u \, d\Omega \right| &\leqslant \|\mathbf{v}\|_{\mathbb{L}^\infty(\Omega)} \, \|\nabla u\|_{\mathbb{L}^2(\Omega)} \, \|w\|_{\mathbb{L}^2(\Omega)} \\ &\leqslant \|\mathbf{v}\|_{\mathbb{L}^\infty(\Omega)} \, \|u\|_{\mathbb{H}^1(\Omega)} \, \|w\|_{\mathbb{H}^1(\Omega)}. \end{aligned} \tag{B.9}$$

Pour obtenir la relation (B.7) et ainsi démontrer la continuité de la forme g, il suffit de prendre $M = \kappa + \|\mathbf{v}\|_{\mathbb{L}^\infty(\Omega)}$.

Enfin, nous pouvons montrer que la forme linéaire $L(w)$ est bien bornée en appliquant l'inégalité de Cauchy-Schwarz, ce qui nous donne

$$\left| \int_\Omega s \, w \, d\Omega \right| \leqslant \|s\|_{\mathbb{L}^2(\Omega)} \, \|w\|_{\mathbb{L}^2(\Omega)}, \tag{B.10}$$

$$\left| \int_{\Gamma_N} q \, w \, d\Gamma_N \right| \leqslant \|q\|_{\mathbb{L}^2(\Gamma_N)} \, \|w\|_{\mathbb{L}^2(\Gamma_N)}. \tag{B.11}$$

Les hypothèses du lemme de Lax-Milgram sont ainsi bien vérifiées et la formulation variationnelle 4.1 est donc bien posée.

Annexe C.

Principe du minimum

Les équations elliptiques sont caractérisées par des propriétés qui leur sont propres. Dans le cas d'équations elliptiques du second ordre, il est possible d'introduire un principe du minimum. Ce principe permet d'obtenir une information mathématique et physique sur le comportement de la solution sans avoir à résoudre explicitement l'équation différentielle. D'un point de vue numérique, il offre également un moyen simple pour valider un code de calcul.

Pour introduire ce principe, nous considérons l'équation d'advection-diffusion linéaire suivante :

$$\kappa \, \Delta u - \mathbf{v} \cdot \nabla u = -s(\mathbf{x}). \tag{C.1}$$

Cette équation peut s'écrire de manière simplifiée en introduisant l'opérateur elliptique R tel que

$$Ru = \kappa \, \Delta u - \mathbf{v} \cdot \nabla u. \tag{C.2}$$

Dans le cadre de l'écoulement visqueux d'un fluide, le terme $-s(\mathbf{x})$ de dissipation visqueuse est nécessairement négatif ou nul. Il en découle que la solution de l'équation d'advection-diffusion doit satisfaire $Ru \leqslant 0 \, \forall \, \mathbf{x} \in \Omega$. La fonction u est alors qualifiée de sur-solution du problème $Ru = 0$ dans Ω. Dans ce cas, il est possible d'établir un principe du minimum pour la solution u [1].

Pour établir le principe du minimum, nous considérons l'opérateur elliptique M défini par

$$Mu = \sum_{i,j=1}^{n} a_{ij}(\mathbf{x}) \, \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(\mathbf{x}) \, \frac{\partial u}{\partial x_i}, \quad a_{ij} = a_{ji} \, . \tag{C.3}$$

L'opérateur R défini précédemment n'est en fait qu'un cas particulier de l'opérateur M. Pour que l'opérateur M soit elliptique, la matrice $[a_{ij}(\mathbf{x})]$ doit être définie positive en tout point $\mathbf{x} \in \Omega$. Dans le cas de l'opérateur R cette condition est trivialement vérifiée vu que la matrice $[a_{ij}(\mathbf{x})] = \kappa \, \delta_{ij}$ où δ_{ij} est le symbole de Kronecker et $\kappa > 0$.

Le principe du minimum désigne deux principes à savoir le principe faible et le principe fort du minimum. Le premier principe dit que la borne inférieure de u dans Ω est égale à sa borne inférieure sur la frontière de Ω. Ainsi, la valeur minimale de u dans $\overline{\Omega}$ vaut la valeur minimale de u sur $\partial\Omega$. Le second principe est plus restrictif et affirme quant à lui que si u atteint sa valeur minimale dans Ω, alors u est une fonction constante.

1. Si $Ru \geqslant 0 \, \forall \, \mathbf{x} \in \Omega$, alors la fonction u est qualifiée de sous-solution du problème $Ru = 0$ dans Ω. Dans ce cas, un principe du maximum peut être établi.

Mathématiquement, le principe faible du minimum s'énonce [26] :

Théorème C.1 (**Principle faible du minimum**). *Soit M un opérateur elliptique défini par la relation (C.3) dans un domaine borné Ω. Supposons une fonction u qui vérifie $Mu \leqslant 0$ dans Ω et qui appartient à $C^2(\Omega) \cap C^0(\overline{\Omega})$. Alors le minimum de u dans $\overline{\Omega}$ est réalisé sur $\partial\Omega$, c'est-à-dire*

$$\inf_{\Omega} u = \inf_{\partial\Omega} u.$$

Une preuve de ce principe est donnée à la fin de cette annexe. Bien que le principe faible du minimum soit en général suffisant, le principe fort garantit l'inexistence d'un minimum non trivial à l'intérieur de Ω. Ce principe est repris ci-dessous sans démonstration [26].

Théorème C.2 (**Principe fort du minimum**). *Soit M un opérateur uniformément elliptique défini par la relation (C.3) tel que $Mu \leqslant 0$ dans un domaine Ω (pas nécessairement borné). Si u réalise son minimum à l'intérieur de Ω, alors u est constant dans Ω.*

Démonstration.

Cette démonstration du principe faible du minimum est inspirée d'Auphan [3] ainsi que de Gilbarg et Trudinger [26]. La démonstration procède en deux étapes :

1. Nous supposons tout d'abord que $Mu(\mathbf{x}) < 0 \, \forall \mathbf{x} \in \Omega$. Dans ce cas, nous pouvons montrer que le minimum de u ne peut pas être atteint à l'intérieur de $\overline{\Omega}$. En effet, supposons qu'il existe un point $\mathbf{x}_0 \in \Omega$ qui soit un minimum de u à l'intérieur de $\overline{\Omega}$. La condition de minimum local de u en $\mathbf{x}_0$ implique les deux conséquences suivantes :

   - $\nabla u(\mathbf{x}_0) = \mathbf{0}$ (condition de stationnarité du point $\mathbf{x}_0$) ;
   - la matrice hessienne $H(\mathbf{x})$ évaluée en $\mathbf{x}_0$ doit être semi-définie positive (condition nécessaire pour avoir un minimum).

   En tenant compte de ces deux conséquences, nous pouvons écrire

   $$\begin{aligned} Mu(\mathbf{x}_0) &= \sum_{i,j=1}^{n} a_{ij}(\mathbf{x}_0) \frac{\partial^2 u}{\partial x_i \partial x_j}(\mathbf{x}_0) + \sum_{i=1}^{n} b_i(\mathbf{x}_0) \frac{\partial u}{\partial x_i}(\mathbf{x}_0) \\ &= \sum_{i,j=1}^{n} a_{ij}(\mathbf{x}_0) \frac{\partial^2 u}{\partial x_i \partial x_j}(\mathbf{x}_0) \\ &= \mathrm{Tr}([a_{ij}(\mathbf{x}_0)] H(\mathbf{x}_0)] \geqslant 0. \end{aligned}$$

   L'inégalité résulte du fait que la matrice $[a_{ij}(\mathbf{x}_0)]$ est définie positive et que $H(\mathbf{x}_0)$ est semi-définie positive. Cette inégalité est en contradiction avec l'hypothèse selon laquelle $Mu < 0 \, \forall \mathbf{x} \in \Omega$. Dès lors, si $Mu < 0 \, \forall \mathbf{x} \in \Omega$ alors u ne peut pas atteindre son minimum à l'intérieur de $\overline{\Omega}$.

2. Pour la seconde partie de la preuve, nous supposons qu'en tout point de Ω nous avons $|b_i(\mathbf{x})|/\lambda(\mathbf{x}) \leqslant b_0 \, (i = 1, 2, \ldots, n)$ où $\lambda(\mathbf{x})$ représente la plus petite valeur propre de la

matrice $[a_{ij}(\mathbf{x})]$. Cette condition est en fait une condition assez naturelle qui revient à borner les termes de plus petit ordre par rapport aux termes de plus grand ordre. Pour l'équation d'advection-diffusion, cela revient à borner les termes de convection par rapport à ceux de diffusion. Comme l'opérateur M est elliptique, nous savons que $a_{11}(\mathbf{x}) \geqslant \lambda(\mathbf{x}) > 0$. Nous considérons à présent la fonction $v(\mathbf{x}) = e^{\gamma x_1}$. Si γ est suffisamment grand, nous pouvons écrire

$$Mv(\mathbf{x}) = Me^{\gamma x_1} = (\gamma^2 a_{11} + \gamma b_1)\, e^{\gamma x_1} \geqslant \lambda\,(\gamma^2 - \gamma b_0) e^{\gamma x_1} > 0.$$

Dès lors pour n'importe quel $\varepsilon > 0$, $M(u - \varepsilon e^{\gamma x_1}) < 0$ sur Ω (car $Mu \leqslant 0$ par hypothèse et $Me^{\gamma x_1} > 0$) et en utilisant la première partie de la preuve, nous obtenons alors

$$\inf_{\Omega}(u - \varepsilon e^{\gamma x_1}) = \inf_{\partial\Omega}(u - \varepsilon e^{\gamma x_1}).$$

Si $\varepsilon \to 0$, nous obtenons bien le résultat attendu à savoir $\inf\limits_{\Omega} u = \inf\limits_{\partial\Omega} u$.

□

Annexe D.

Équivalence entre la forme faible et la forme forte du problème thermique contraint

Nous montrons dans cette annexe que si une solution classique $T \in \mathbb{C}^2(\overline{\Omega})$ satisfait la formulation variationnelle 4.2, alors cette solution satisfait aussi la formulation forte (4.41). Une solution classique est ainsi une solution pour laquelle l'ensemble des dérivées qui apparaissent dans la forme forte existent au sens habituel et sont continues. Nous supposerons également que les fonctions $\mathbf{v}$ et s sont continues sur $\overline{\Omega}$ [1]. Les quantités physiques ρ, c et k seront de plus traitées comme des constantes. Pour développer la démonstration ci-dessous, nous nous sommes basé sur un développement établi pour un problème de contact mécanique [73] et nous l'avons appliqué au cas d'un problème thermique.

Démonstration.

Pour commencer cette preuve, nous notons que $T \in K$ et dès lors $T = T_s$ sur Γ_s et $T \leqslant T_m$ sur Γ_b. Ensuite, nous considérons une fonction arbitraire $\phi \in \mathbb{C}^\infty(\Omega)$ qui s'annule sur la frontière de Ω. La fonction $\theta = T \pm \phi$ appartient à l'ensemble K. Si T est solution de l'inéquation variationnelle, alors T satisfait

$$a(T, \pm\phi) + c(\mathbf{v}; T, \pm\phi) \geqslant (s, \pm\phi) + (q, \pm\phi)_{\Gamma_b}. \tag{D.1}$$

Comme ϕ s'annule sur la frontière de Ω, l'inégalité se réécrit aussi

$$a(T, \pm\phi) + c(\mathbf{v}; T, \pm\phi) \geqslant (s, \pm\phi). \tag{D.2}$$

Cette dernière inégalité implique que nous ayons

$$a(T, \phi) + c(\mathbf{v}; T, \phi) = (s, \phi), \tag{D.3}$$

$$\int_\Omega k\, \nabla T \cdot \nabla \phi \, d\Omega + \int_\Omega \rho c\, (\mathbf{v} \cdot \nabla T)\, \phi \, d\Omega = \int_\Omega s\, \phi \, d\Omega. \tag{D.4}$$

En utilisant le théorème de la divergence, nous avons alors

$$\int_\Omega -k\, \Delta T\, \phi \, d\Omega + \int_\Omega \rho c\, (\mathbf{v} \cdot \nabla T)\, \phi \, d\Omega = \int_\Omega s\, \phi \, d\Omega. \tag{D.5}$$

Comme les fonctions ϕ sont arbitraires, il en résulte que T satisfait l'équation de la chaleur

$$-k\, \Delta T + \rho c\, \mathbf{v} \cdot \nabla T = s. \tag{D.6}$$

Pour continuer, il est pratique de décomposer Γ_b en une région $\Gamma_{b,1}$ sur laquelle $T < T_m$ et une région $\Gamma_{b,2}$ sur laquelle $T = T_m$. Nous introduisons aussi les fonctions $\theta = T \pm \phi \in K$

1. Le lecteur intéressé trouvera dans Brezis [8] une présentation plus détaillée du cadre fonctionnel dans lequel s'inscrit la démonstration de l'équivalence entre formes faible et forte dans le cas d'égalités variationnelles.

où les fonctions ϕ sont des fonctions arbitraires qui s'annulent sur $\Gamma_s \cup \Gamma_{b,2}$. Les fonctions θ sont correctement définies. En effet, nous avons $T < T_m$ sur $\Gamma_{b,1}$ et dès lors si ϕ est choisi suffisamment petit sur $\Gamma_{b,1}$, alors $\theta = T \pm \phi$ appartient bien à K. Cela nous permet d'écrire la relation suivante :

$$a(T, \pm\phi) + c(\mathbf{v}; T, \pm\phi) \geqslant (s, \pm\phi) + (q, \pm\phi)_{\Gamma_{b,1}}. \tag{D.7}$$

De cette inégalité, il résulte

$$a(T, \phi) + c(\mathbf{v}; T, \phi) = (s, \phi) + (q, \phi)_{\Gamma_{b,1}}, \tag{D.8}$$

$$\int_\Omega k\, \nabla T \cdot \nabla \phi \, d\Omega + \int_\Omega \rho c\, (\mathbf{v} \cdot \nabla T)\, \phi \, d\Omega = \int_\Omega s\, \phi \, d\Omega + \int_{\Gamma_{b,1}} q\, \phi \, d\Gamma_{b,1}. \tag{D.9}$$

En utilisant encore une fois le théorème de la divergence, nous avons

$$\int_\Omega -k\, \Delta T\, \phi \, d\Omega + \int_\Omega \rho c\, (\mathbf{v} \cdot \nabla T)\, \phi \, d\Omega + \int_{\Gamma_{b,1}} (k \nabla T \cdot \mathbf{n})\, \phi \, d\Gamma_{b,1} = \int_\Omega s\, \phi \, d\Omega + \int_{\Gamma_{b,1}} q\, \phi \, d\Gamma_{b,1}. \tag{D.10}$$

Les intégrales de volume disparaissent puisque nous avons montré précédemment que T satisfait l'équation (D.6). Il nous reste alors

$$\int_{\Gamma_{b,1}} (k\, \nabla T \cdot \mathbf{n})\, \phi \, d\Gamma_{b,1} = \int_{\Gamma_{b,1}} q\, \phi \, d\Gamma_{b,1}. \tag{D.11}$$

Comme les fonctions ϕ sont arbitraires, nous obtenons donc

$$k\, \nabla T \cdot \mathbf{n} = q \ \text{ sur } \Gamma_{b,1}. \tag{D.12}$$

Nous considérons maintenant une fonction quelconque $\theta \in K$. Nous pouvons écrire

$$\begin{aligned} a(T, \theta - T) + c(\mathbf{v}; T, \theta - T) = & \int_\Omega -k\, \Delta T\, (\theta - T)\, d\Omega + \int_\Omega \rho c\, (\mathbf{v} \cdot \nabla T)\, (\theta - T)\, d\Omega \\ & + \int_{\Gamma_{b,1}} (k\, \nabla T \cdot \mathbf{n})\, (\theta - T)\, d\Gamma_{b,1} + \int_{\Gamma_{b,2}} (k\, \nabla T \cdot \mathbf{n})\, (\theta - T)\, d\Gamma_{b,2}. \end{aligned} \tag{D.13}$$

Vu les relations (D.6) et (D.12), nous obtenons

$$\begin{aligned} a(T, \theta - T) + c(\mathbf{v}; T, \theta - T) = & \int_\Omega s\, (\theta - T)\, d\Omega + \int_{\Gamma_{b,1}} q\, (\theta - T)\, d\Gamma_{b,1} \\ & + \int_{\Gamma_{b,2}} (k\, \nabla T \cdot \mathbf{n})\, (\theta - T)\, d\Gamma_{b,2}. \end{aligned} \tag{D.14}$$

Sur $\Gamma_{b,2}$, nous pouvons réaliser la décomposition suivante

$$k\, \nabla T \cdot \mathbf{n} = q - \rho L a_b, \tag{D.15}$$

où a_b est une quantité a priori quelconque (pas de contrainte sur le signe). En tenant compte de cette décomposition, nous réécrivons (D.14) comme

$$a(T, \theta - T) + c(\mathbf{v}; T, \theta - T) = (s, \theta - T) + (q, \theta - T)_{\Gamma_b} - \int_{\Gamma_{b,2}} \rho L a_b\, (\theta - T)\, d\Gamma_{b,2}. \tag{D.16}$$

Or, nous savons que $\forall \theta \in K$, T satisfait l'inégalité variationnelle

$$a(T, \theta - T) + c(\mathbf{v}; T, \theta - T) \geqslant (s, \theta - T) + (q, \theta - T)_{\Gamma_b}. \tag{D.17}$$

Il en résulte donc

$$\int_{\Gamma_{b,2}} \rho L a_b \left(T - \theta\right) d\Gamma_{b,2} \geqslant 0. \tag{D.18}$$

Par définition de $\Gamma_{b,2}$, $T = T_m$ sur $\Gamma_{b,2}$ et l'inégalité (D.18) correspond à

$$\int_{\Gamma_{b,2}} \rho L a_b \left(T_m - \theta\right) d\Gamma_{b,2} \geqslant 0. \tag{D.19}$$

Si $\theta \in K$, nous avons nécessairement $T_m - \theta \geqslant 0$. De plus, comme θ est arbitraire dans K, il en résulte

$$a_b \geqslant 0 \ \text{ sur } \Gamma_{b,2}. \tag{D.20}$$

Enfin, la relation de complémentarité $a_b \left(T - T_m\right) = 0$ est satisfaite de manière évidente par la solution T si nous considérons que nous pouvons écrire de manière générale $k \nabla T \cdot \mathbf{n} = q - \rho L a_b$ sur Γ_b. Cette remarque conclut notre démonstration de l'équivalence entre formes faible et forte pour l'inégalité variationnelle de la thermique.

□

Annexe E.

Convergence de la méthode de pénalisation

Démonstration.

Cette preuve est basée sur une démonstration tirée de Glowinski *et al.* [28]. Nous commençons tout d'abord par réécrire (4.51) sous la forme

$$(P(T_\varepsilon),\theta)_{\Gamma_b} = \varepsilon\left[(s,\theta)+(q,\theta)_{\Gamma_b} - a(T_\varepsilon,\theta) - c(\mathbf{v};T_\varepsilon,\theta)\right].$$

Nous avons alors

$$\|P(T_\varepsilon)\|_{\Gamma_b} = O(\varepsilon),$$

et donc $\|P(T_\varepsilon)\|_{\Gamma_b} \to 0$ si $\varepsilon \to 0$. Dans ce cas, nous pouvons construire une suite de fonctions $\{T_\varepsilon\}$ de G qui converge vers une fonction T qui appartient à l'ensemble K car $P(T) = 0$. Il nous reste à présent à montrer que T est solution de l'inéquation variationnelle.

Pour cela, nous considérons à nouveau l'expression (4.51) dans laquelle nous remplaçons θ par $\chi - T_\varepsilon$ avec $\chi \in K$. Cela donne

$$\begin{aligned} a(T_\varepsilon,\chi-T_\varepsilon)+c(\mathbf{v};T_\varepsilon,\chi-T_\varepsilon)-(s,\chi-T_\varepsilon)-(q,\chi-T_\varepsilon)_{\Gamma_b} &= \frac{1}{\varepsilon}(P(T_\varepsilon),T_\varepsilon-\chi)_{\Gamma_b} \\ &= \frac{1}{\varepsilon}(P(T_\varepsilon)-P(\chi),T_\varepsilon-\chi)_{\Gamma_b} \\ &\geqslant 0. \end{aligned}$$

Dans ces relations, nous avons exploité le fait que $P(\chi) = 0$ car $\chi \in K$ ainsi que le caractère monotone de l'opérateur P.

En prenant la limite pour $\varepsilon \to 0$, nous trouvons alors pour $T \in K$

$$a(T,\chi-T)+c(\mathbf{v};T,\chi-T)-(s,\chi-T)-(q,\chi-T)_{\Gamma_b} \geqslant 0 \quad \forall \chi \in K.$$

Nous constatons alors que la solution du problème pénalisé est bien solution de la formulation variationnelle 4.2 quand $\varepsilon \to 0$.

□

Bibliographie

[1] G. Allaire et M. Schoenauer. *Conception optimale de structures*, volume 58 de *Mathématiques & Applications*. Springer, 2007.

[2] A. Aschwanden, E. Bueler, C. Khroulev et H. Blatter. An enthalpy formulation for glaciers and ice sheets. *Journal of Glaciology*, 58(209):441–457, 2012.

[3] T. Auphan. Equations aux dérivées Partielles. Notes de cours, Février 2012. http://www.latp.univ-mrs.fr/~tauphan/lib/exe/fetch.php?media=edp.pdf.

[4] W. Bai. The quadrilateral 'Mini' finite element for the Stokes problem. *Computer Methods in Applied Mechanics and Engineering*, 143(1):41–47, 1997.

[5] R. W. Baker. Is the creep of ice really independent of the third deviatoric stress invariant ? *The Physical Basis of Ice Sheet Modelling*, 1983.

[6] P. B. Bochev, C. R. Dohrmann et M. D. Gunzburger. Stabilization of low-order mixed finite elements for the Stokes equations. *SIAM Journal on Numerical Analysis*, 44(1):82–101, 2006.

[7] D. Boffi, F. Brezzi et M. Fortin. *Mixed Finite Element Methods and Applications*, volume 44 de *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, 2013.

[8] H. Brezis. *Analyse fonctionnelle : théorie et applications*. Collection mathématiques appliquées pour la maîtrise. Masson Paris, 1987.

[9] F. Brezzi et M. Fortin. *Mixed and Hybrid Finite Element Methods*, volume 15 de *Springer Series in Computational Mathematics*. Springer-Verlag, New-York, 1991.

[10] A. N. Brooks et T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 32(1):199–259, 1982.

[11] C. G. Broyden, J. E. Dennis et J. J Moré. On the Local and Superlinear Convergence of Quasi-Newton Methods. *IMA Journal of Applied Mathematics*, 12:223–245, 1973.

[12] L. Champaney. Contact unilatéral entre solides élastiques. Notes de cours "Eléments finis" du DESS Dynamique des Structures Modernes dans leur Environnement, 2005. http://laurent.champaney.free.fr/Cours/Contact.pdf.

[13] CNES. «ArgOcéan : océan et variabilité climatique» [En ligne], http://www.cnes.fr/web/CNES-fr/7164-argocean.php (consulté le 12.04.2015).

[14] S. Coutterand. «Calottes glaciaires et inlandsis» [En ligne], http://www.glaciers-climat.com/calottes-glaciaires-et-inlandsis.html (publié le 26.03.2012).

[15] K. M. Cuffey et W. S. B. Paterson. *The Physics of Glaciers*. Butterworth-Heinemann/Elsevier, 4^{e} édition, 2010.

[16] J. E. Dennis et R. B. Schnabel. Least Change Secant Updates for Quasi-Newton Methods. *SIAM Review*, 21(4):443–459, 1979.

[17] D. Docquier. *Representing grounding-line dynamics in Antarctic ice-sheet models*. Thèse de doctorat, Université Libre de Bruxelles, 2013.

[18] J. Donea et A. Huerta. *Finite Element Methods for Flow Problems*. John Wiley & Sons, Chichester, 2003.

[19] A. Ern et J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159 de *Applied Mathematical Sciences*. Springer Science & Business Media, New York, 2004.

[20] A. Fortin et A. Garon. Les éléments finis : de la théorie à la pratique. Notes de cours, Université Laval, 2011. http://www.giref.ulaval.ca/~afortin/cours_elements_finis/documents/notes_elements_finis.pdf.

[21] M. Fortin. Utilisation de la méthode des éléments finis en mécanique des fluides, I. *Calcolo*, 12(4):405–441, 1975.

[22] A. C. Fowler. Sub-temperate basal sliding. *Journal of Glaciology*, 32(110):3–5, 1986.

[23] T.-P. Fries et H. G. Matthies. A review of Petrov-Galerkin Stabilization Approaches and an Extension to Meshfree Methods. 2004.

[24] O. Gagliardini, T. Zwinger, F. Gillet-Chaulet, G. Durand, L. Favier, B. de Fleurian, R. Greve, M. Malinen, C Martín, P. Råback, J. Ruokolainen, M. Sacchettini, M. Schäfer, H. Seddik et J. Thies. Capabilities and performance of Elmer/Ice, a new-generation ice sheet model. *Geoscientific Model Development*, 6(4):1299–1318, 2013.

[25] C. Geuzaine et J.-F. Remacle. Gmsh : a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.

[26] D. Gilbarg et N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 de *Classics in mathematics*. Springer-Verlag, Berlin Heidelberg, 2001.

[27] J. W. Glen. The creep of polycrystalline ice. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 228(1175):519–538, 1955.

[28] R. Glowinski, J.-L. Lions et R. Trémolières. *Numerical Analysis of Variational Inequalities*, volume 8 de *Studies in Mathematics and Its Applications*. Elsevier, North-Holland, 1981.

[29] A. Granas et J. Dugundji. *Fixed Point Theory*. Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[30] R. Greve et H. Blatter. *Dynamics of Ice Sheets and Glaciers*. Advances in geophysical and environmental mechanics and mathematics. Springer-Verlag, Berlin, 2009.

[31] M. D. Gunzburger. *Finite Element Methods for Viscous Incompressible Flows : A Guide to Theory, Practice, and Algorithms*. Computer Science and Scientific Computing. Academic Press, 1989.

[32] M. M. Gupta, R. P. Manohar et J. W. Stephenson. High-Order Difference Schemes for Two-Dimensional Elliptic Equations. *Numerical Methods for Partial Differential Equations*, 1(1):71–80, 1985.

[33] M. E. Gurtin, E. Fried et L. Anand. *The Mechanics and Thermodynamics of Continua*. Cambridge University Press, 2010.

[34] R. L. Hooke. *Principles of Glacier Mechanics*. Cambridge University Press, 2^{e} édition, 2005.

[35] J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell et C. A. Johnson. Climate Change 2001 : The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. *Cambridge University Press*, 2001.

[36] T. J. R. Hughes, L. P. Franca et G. M. Hulbert. A new finite element formulation for computational fluid dynamics : VIII. The Galerkin/least-squares method for advective-diffusive equations. *Computer Methods in Applied Mechanics and Engineering*, 73(2):173–189, 1989.

[37] G. Jouvet. *Modélisation, analyse mathématique et simulation numérique de la dynamique des glaciers*. Thèse de doctorat, École polytechnique fédérale de Lausanne, 2010.

[38] G. Jouvet et J. Rappaz. Analysis and Finite Element Approximation of a Nonlinear Stationary Stokes Problem Arising in Glaciology. *Advances in Numerical Analysis*, Volume 2011, 2011.

[39] F. Kalfoun et L. Augustin. «Archives climatiques des 740.000 dernières années» [En ligne], http://planet-terre.ens-lyon.fr/article/archives-climatiques.xml (publié le 19.10.2004).

[40] J. A. Kallen-Brown. *Computational methods for ice flow simulation*. Thèse de doctorat, Eidgenössische Technische Hochschule ETH Zürich, 2011.

[41] C. T. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, Philadelphia, 1995.

[42] D. E. Keyes, L. C. McInnes, C. Woodward, W. Gropp, E. Myra, M. Pernice, J. Bell, J. Brown, A. Clo, J. Connors *et al.* Multiphysics simulations : Challenges and opportunities. *International Journal of High Performance Computing Applications*, 27(1):4–83, 2013.

[43] N. Kikuchi et J. T. Oden. *Contact Problems in Elasticity. A Study of Variational Inequalities and Finite Element Methods*, volume 8 de *Studies in Applied and Numerical Mathematics*. SIAM, 1988.

[44] T. Kleiner, M. Rückamp, J. H. Bondzio et A. Humbert. Enthalpy benchmark experiments for numerical ice sheet models. *The Cryosphere*, 9(1):217–228, 2015.

[45] D. A. Knoll et D. E. Keyes. Jacobian-free Newton–Krylov methods : a survey of approaches and applications. *Journal of Computational Physics*, 193(2):357–397, 2004.

[46] J.-F. Lemieux, S. F. Price, K. J. Evans, D. Knoll, A. G. Salinger, D. M. Holland et A. J. Payne. Implementation of the Jacobian-free Newton–Krylov method for solving the first-order ice sheet momentum balance. *Journal of Computational Physics*, 230(17):6531–6545, 2011.

[47] J. Lindenstrauss et D. Preiss. On Fréchet differentiability of Lipschitz maps between Banach spaces. *Annals of mathematics*, 157(1):257–288, 2003.

[48] J.-L. Lions et E. Magenes. *Problèmes aux limites non homogènes et applications*, volume 17 de *Travaux et recherches mathématiques*. Dunod, Paris, 1968.

[49] J.-L. Lions et G. Stampacchia. Variational Inequalities. *Communications on Pure and Applied Mathematics*, 20(3):493–519, 1967.

[50] J. M. Martínez. Practical quasi-Newton methods for solving nonlinear systems. *Journal of Computational and Applied Mathematics*, 124(1):97–121, 2000.

[51] G. Meurant. *Méthodes de Krylov pour la résolution des systèmes linéaires*. Ed. Techniques Ingénieur, 2007.

[52] S. A. Mohamed, N. A. Mohamed, A. F. A. Gawad et M. S. Matbuly. A modified diffusion coefficient technique for the convection diffusion equation. *Applied Mathematics and Computation*, 219(17):9317–9330, 2013.

[53] J. M. Ortega et W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables*, volume 30 de *Classics in applied mathematics*. SIAM, 1970.

[54] F. Pattyn. A new three-dimensional higher-order thermomechanical ice sheet model : Basic sensitivity, ice stream development, and ice flow across subglacial lakes. *Journal of Geophysical Research*, 108(B8), 2003.

[55] F. Pattyn, B. de Smedt et R. Souchez. Influence of subglacial Vostok lake on the regional ice dynamics of the Antarctic ice sheet : a model study. *Journal of Glaciology*, 50(171):583–589, 2004.

[56] A. Quarteroni. *Numerical Models for Differential Problems*, volume 2 de *MS&A - Modeling, Simulation and Applications*. Springer, Milan, 2009.

[57] A. Quarteroni, R. Sacco et F. Saleri. *Numerical Mathematics*, volume 37 de *Texts in Applied Mathematics*. Springer-Verlag, New York, 2000.

[58] A. Quarteroni et A. Valli. *Numerical Approximation of Partial Differential Equations*, volume 23 de *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin Heidelberg, 2008.

[59] D. W. A. Rees. *Basic Engineering Plasticity : An Introduction with Engineering and Manufacturing Applications*. Butterworth-Heinemann, 2006.

[60] F.-X. Roux. Chapitre 6 : Méthodes de Krylov. Notes de cours, Université Pierre et Marie Curie. https://www.ljll.math.upmc.fr/MathInter/tutorial-1.php?M=enseignement (consulté le 25.04.2015).

[61] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, 2^{e} édition, 2003.

[62] M. L. Salby. *Physics of the Atmosphere and Climate*. Cambridge University Press, 2012.

[63] H. Seroussi. *Modeling ice flow dynamics with advanced multi-model formulations*. Thèse de doctorat, Ecole Centrale Paris, 2011.

[64] U.S. Geological Survey. «Glacier and Landscape Change in Response to Changing Climate : Repeat Photography of Alaskan Glaciers» [En ligne], http://www.usgs.gov/climate_landuse/glaciers/repeat_photography.asp (dernière modification le 30.05.2012).

[65] SwissEduc. «Glaciers online» [En ligne], http://swisseduc.ch/glaciers/earth_icy_planet/glaciers02-en.html?id=9 (consulté le 18.03.2015).

[66] SwissEduc. «Glaciers online» [En ligne], http://www.swisseduc.ch/glaciers/alps/grosser_aletschgletscher/aletschgletscher-en.html?id=0 (consulté le 12.04.2015).

[67] R. Temam et I. Ekeland. *Analyse convexe et problèmes variationnels*. Dunod-Gauthier-Villars, Paris, 1974.

[68] C. J. van der Veen. *Fundamentals of Glacier Dynamics*. CRC Press, 2e édition, 2013.

[69] H. A. van der Vorst. *Iterative Krylov Methods for Large Linear Systems*, volume 13 de *Cambridge monographs on applied and computational mathematics*. Cambridge University Press, 2003.

[70] J. P. Whiteley, K. Gillow, S. J. Tavener et A. C. Walter. Error bounds on block Gauss–Seidel solutions of coupled multiphysics problems. *International Journal for Numerical Methods in Engineering*, 88(12):1219–1237, 2011.

[71] P. Wriggers. *Computational Contact Mechanics*. Springer, 2e édition, 2006.

[72] S. J. Wright et J. Nocedal. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006.

[73] F. M. Youbissi. *Résolution par éléments finis du problème de contact unilatéral par des méthodes d'optimisation convexe*. Thèse de doctorat, Université Laval, 2006.

[74] E. Zeidler. *Nonlinear Functional Analysis and its Applications I : Fixed-Point Theorems*. Springer-Verlag, New York, 1986.