UNIVERSITÉ AIX-MARSEILLE II - MEDITERRANEE
U.F.R. de Mathématiques, Informatique et Mécanique

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE LA MEDITERRANEE

Discipline : Mathématiques
École Doctorale Mathématiques et Informatique de Marseille
E.D. numéro 184

présentée et soutenue publiquement
par

Matthieu JOBELIN

le 16 octobre 2006

Méthodes de projection pour le calcul d'écoulements incompressibles ou dilatables

JURY

Mme	C.	BERNARDI	Laboratoire J.-L. LIONS, PARIS VI	Rapporteur
MM.	J.-L.	GUERMOND	Dep. of Mathematics, Texas A&M University	Rapporteur
	Ph.	ANGOT	LATP, Aix-Marseille I	Directeur de thèse
	J.-C.	LATCHÉ	IRSN, Cadarache	Encadrant
	J.-P.	CALTAGIRONE	ENSCP Bordeaux	Examinateur
	C.	GALUSINSKY	Université de Toulon	Examinateur
Mme	R.	HERBIN	LATP, Aix-Marseille I	Présidente

Table des matières

Le manuscript est organisé de la manière suivante : le premier chapitre est un résumé de l'ensemble des travaux effectués, puis les quatre chapitres suivants détaillent respectivement la construction de la méthode de projection-pénalité pour les écoulements incompressibles, son analyse, l'extrapolation de cette méthode aux écoulements dilatables et, enfin, la mise en œuvre de la technique d'éléments finis joints.

Chapitre I

Synthèse générale

I.1 Introduction

Nous nous intéressons dans cette thèse à une classe de problèmes obtenue en adjoignant aux équations de bilan de masse et de quantité de mouvement une équation de transport diffusion d'une variable additionnelle, z, dont la masse volumique se déduit :

$$\left|\begin{array}{ll} \dfrac{\partial \varrho z}{\partial t} + \nabla \cdot (\varrho z u) = \nabla \cdot \mathcal{D} \nabla z & \text{dans } [0,T] \times \Omega \\ \varrho = \mathcal{G}(z) & \\ \dfrac{\partial \varrho u}{\partial t} + \nabla \cdot (\varrho u \otimes u) = \nabla \cdot \tau(u) - \nabla p + f & \text{dans } [0,T] \times \Omega \\ \dfrac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho u) = 0 & \text{dans } [0,T] \times \Omega \end{array}\right. \tag{I.1.1}$$

La variable ϱ désigne ici la masse volumique, t le temps, u la vitesse, p la pression, f une force volumique répartie, $\mathcal{D}$ est un coefficient de diffusion et τ est le tenseur des contraintes visqueuses. Pour fixer les idées, on peut voir la fonction $\mathcal{G}(\cdot)$ comme une loi d'état. Ce cadre formel inclut les écoulements incompressibles où la masse volumique est constante.

Des représentations mathématiques de ce type sont rencontrées lors de la modélisation d'une grande variété de phénomènes physiques. Les problèmes de convection naturelle à faible nombre de Mach, par exemple, rentrent dans ce cadre, lorsqu'ils sont traités en utilisant un modèle asymptotique, *i.e.* d'un système d'équations vérifié par les champs de vitesse et pression dans l'écoulement lorsque le nombre de Mach tend vers zéro, tel que décrit par Majda et Sethian [25]. La variable z représente alors la température et la loi $\mathcal{G}(\cdot)$ est déduite de la loi d'état, moyennant le calcul préalable de la pression thermodynamique lorsque le système physique considéré est clos. Si l'on identifie z à une concentration et $\varrho = \mathcal{G}(z)$ à une loi de mélange, on obtient les équations de la convection solutale. De la même manière, certains problèmes très simplifiés de combustion s'écrivent sous la forme du système [I.1.1] ; la variable z prend alors la signification d'une variable d'avancement [1].

Ces problématiques physiques font partie des phénomènes d'intérêt dans le domaine de la sûreté nucléaire et, à ce titre, sont traitées à la Direction de la Prévention des Accidents Majeurs (DPAM) de l'Institut de Radioprotection et de Sûreté Nucléaire (IRSN). Pour ne citer que quelques exemples, la modélisation des incendies fait l'objet d'un programme de recherche ambitieux, combinant expériences à grande

échelle et développement de logiciels de simulation, dont le code ISIS [1] ; cette même approche expérimentale et analytique est également mise en œuvre pour l'étude des phénomènes de dissolution des pastilles de combustible par la gaine lors d'un accident de fusion de cœur. Le présent travail de thèse s'inscrit dans ce cadre général : il a pour objet de développer des schémas numériques en soutien à ces développements de logiciels de simulation.

Du fait que la masse volumique du fluide est supposée indépendante de la pression, cette dernière joue d'un point de vue mathématique un rôle similaire à celui qu'elle tient dans les équations de Navier-Stokes incompressible. Pour s'en convaincre, il suffit de réécrire la première équation en fonction de la variable $q = \varrho u$ et les deux dernières équations du système [I.1.1] retrouvent la structure classique d'un problème mixte. En conséquence, il est naturel de mettre en œuvre pour la résolution numérique de ce problème des schémas initialement développés dans le contexte des écoulements incompressibles. Parmi ceux-ci, les méthodes de projection ont, depuis les travaux originels de Chorin et Témam ([8], [34]), acquis une popularité croissante. Ce succès tient dans le fait qu'elles découplent à chaque pas de temps les équations de bilan de quantité de mouvement et de bilan de masse, substituant ainsi à un problème mixte, de résolution difficile, une succession de problèmes elliptiques plus aisés à résoudre.

Le principe de ces méthodes est le suivant. Dans une première étape, on obtient une prédiction de la vitesse par la résolution de l'équation de bilan de quantité de mouvement, dans laquelle la pression est ignorée (méthode originelle) ou approchée par une formule explicite (méthode dite incrémentale). La seconde étape consiste à projeter la vitesse prédite dans l'espace des fonctions solénoïdales ; cette étape s'apparente à un problème de Darcy, qui est classiquement réécrit comme un problème elliptique pour la pression (méthode originelle) ou l'incrément de pression (méthode incrémentale). Sur cette idée de base sont venues, au fil des années, se greffer de multiples variantes. La méthode de projection incrémentale semble due à Goda [12], le premier schéma formellement de second ordre en temps à Van Kan [37]. Dans l'étape de projection, les conditions aux limites appliquées à l'incrément de pression sont artificielles, *i.e.* ne sont pas vérifiées par la solution du problème, ce qui induit des pertes de précision, particulièrement graves pour les écoulements visqueux obéissant à des conditions aux limites ouvertes sur une partie de la frontière [19]. Ce phénomène est corrigé dans une variante proposée par Timmermans *et al.* [35] puis analysée par Guermond et Shen [21], qui a reçu le nom de méthode rotationnelle. On trouvera une revue de ces différents schémas et de leurs propriétés de convergence respectives dans [18].

Si le découplage des équations de bilan de quantité de mouvement et de bilan de masse simplifie la résolution, il introduit également une erreur numérique, dite erreur de fractionnement, par rapport aux solutions de schéma (semi-)implicite comme la méthode de Lagrangien augmenté [10], erreur qui devient, pour des discrétisations en temps d'ordre deux, importante voire dominante à fort pas de temps, comme nous le montrerons par la suite. Cette erreur de fractionnement disparaitrait si, par un choix judicieux de la pression approchée, la vitesse prédite vérifiait la contrainte de divergence : la vitesse prédite serait alors la même que celle obtenue par un schéma couplé, et l'étape de projection serait sans objet. Bien sûr, ce comportement n'est pas accessible dans la pratique. Il peut toutefois être approché en rajoutant dans la première étape un terme de pénalisation associé à la contrainte de divergence, ana-

logue à celui utilisé dans les techniques de Lagrangien augmenté : c'est le principe des méthodes de projection-pénalité. Le premier schéma de ce type semble avoir été suggéré par Shen [29, chapitre 6], puis l'idée reprise indépendamment par Caltagirone et Breil [7]. Cette dernière méthode s'appuie sur une discrétisation spatiale en volumes finis et une étape de projection particulière, dite "projection vectorielle" ; elle a été extensivement employée dans le code AQUILON.

La concrétisation de ces idées dans le contexte des éléments finis, pour des écoulements incompressibles et dilatables (*i.e.* avec une masse volumique variable mais indépendante de la pression, donc rentrant dans le cadre défini ci-dessus), constitue le principal apport de ce travail de thèse. Nous décrivons ainsi dans la prochaine section la construction de la méthode de projection-pénalité pour les équations de Navier-Stokes incompressibles, ainsi qu'une batterie de tests permettant de vérifier ses propriétés de convergence, avec des conditions aux limites de Dirichlet ou ouvertes ; l'analyse de cette méthode, sur le problème simplifié que constituent les équations de Stokes instationnaire, fait ensuite l'objet de la section suivante. Enfin, nous étendons les techniques développées dans le cadre incompressible aux écoulements dilatables dans une troisième partie.

Le second thème abordé dans cette thèse est la mise en œuvre, pour la résolution de problèmes multi-physiques, d'une technique de décomposition de domaines particulièrement générale : la méthode des éléments finis joints. Après une rapide présentation de cette méthode et quelques tests sur des problèmes modèles, dont l'un avec un second membre irrégulier (mesure de Dirac), nous traitons ainsi, dans une dernière partie, un problème couplant convection naturelle dans un domaine fermé avec la conduction de la chaleur dans les parois qui l'entourent.

I.2 La méthode de projection-pénalité pour les écoulements incompressibles

L'algorithme semi-discrétisé en temps

Les équations de Navier-Stokes pour des écoulements incompressibles et isothermes de fluide Newtonien s'écrivent :

$$\varrho\Big[\frac{\partial u}{\partial t} + (u\cdot\nabla)u\Big] = -\nabla p + \nabla\cdot\tau(u) + f \qquad \text{dans} \quad [0,T]\times\Omega \tag{I.2.2a}$$

$$\nabla\cdot u = 0 \qquad \text{dans} \quad [0,T]\times\Omega \tag{I.2.2b}$$

$$u = u_{\mathrm{D}} \qquad \text{sur} \quad [0,T]\times\partial\Omega_{\mathrm{D}} \tag{I.2.2c}$$

$$-pn + \tau(u)\cdot n = f_{\mathrm{N}} \qquad \text{sur} \quad [0,T]\times\partial\Omega_{\mathrm{N}} \tag{I.2.2d}$$

$$u = u_0 \qquad \text{dans} \quad \{0\}\times\Omega \tag{I.2.2e}$$

où, en plus des notations déjà introduites, nous représentons par u_0 la vitesse initiale du fluide et par f la somme des forces extérieures. La masse volumique ϱ est supposée strictement positive et constante. Le domaine de calcul Ω est un ouvert borné régulier sous-ensemble de $\mathbb{R}^d$ avec $d = 2$ ou $d = 3$. Nous supposons que la frontière $\partial\Omega$ de Ω est partitionnée en deux sous-ensembles $\partial\Omega_{\mathrm{D}}$ et $\partial\Omega_{\mathrm{N}}$, de vecteur normal sortant n.

Sur $\partial\Omega_{\mathrm{D}}$, la vitesse est fixée à la valeur u_{D} ; la force par unité de surface exprimée à chaque point de la frontière $\partial\Omega_{\mathrm{N}}$ est donnée et égale à f_{N}. Le tenseur τ représente la partie visqueuse du tenseur de contrainte, dont la divergence est donnée par l'une des expressions suivantes :

$$\nabla \cdot \tau(u) = \mu \Delta u \tag{I.2.3a}$$

$$\text{où} \qquad \nabla \cdot \tau(u) = \nabla \cdot \mu(\nabla u + \nabla u^{\mathrm{T}}) \tag{I.2.3b}$$

où μ représente la viscosité dynamique du fluide. L'équation (I.2.3a) n'a de sens physique que si μ est constante.

Soit $0 = t^0 < t^1 < \cdots < t^N = T$ une partition du temps de calcul $[0, T]$, que nous supposerons uniforme dans un souci de simplification. Nous noterons Δt le pas de temps, *i.e.* l'intervalle constant entre deux pas de temps consécutifs t^n et t^{n+1}, $0 \leq n \leq N-1$. Une semi-discrétisation en temps semi-implicite du système d'équations (I.2.2) est donnée par :

$$\varrho\Big[\frac{Du^{n+1}}{\Delta t} + (u^{\star,n+1} \cdot \nabla)u^{n+1}\Big] = -\nabla p^{n+1} + \nabla \cdot \tau(u^{n+1}) + f^{n+1} \quad \text{dans } \Omega \tag{I.2.4a}$$

$$\nabla \cdot u^{n+1} = 0 \quad \text{dans } \Omega \tag{I.2.4b}$$

$$u^{n+1} = u_{\mathrm{D}}^{n+1} \quad \text{sur } \partial\Omega_{\mathrm{D}} \tag{I.2.4c}$$

$$-p^{n+1} n + \tau(u^{n+1}) \cdot n = f_{\mathrm{N}}^{n+1} \quad \text{sur } \partial\Omega_{\mathrm{N}} \tag{I.2.4d}$$

où $f^{n+1} = f(t^{n+1})$, $u_{\mathrm{D}}^{n+1} = u_{\mathrm{D}}(t^{n+1})$ et $f_{\mathrm{N}}^{n+1} = f_{\mathrm{N}}(t^{n+1})$, u^n et p^n représentent une approximation respectivement de la vitesse u et de la pression p à $t = t^n$, $u^{\star,n+1}$ est une extrapolation de la vitesse à t^{n+1} et $\dfrac{Du^{n+1}}{\Delta t}$ fournit une approximation de la dérivé en temps de la vitesse à t^{n+1} par une "formule de différentiation rétrograde" (en abrégé, BDF pour *Backward Differentiation Formula*), qui prend la forme générale suivante :

$$Du^{n+1} \stackrel{\text{def}}{=} \beta_q u^{n+1} - \sum_{j=0}^{q-1} \beta_j u^{n-j}$$

Dans la pratique, les choix suivants sont communément effectués :

$$Du^{n+1} = u^{n+1} - u^n \qquad u^{\star,n+1} = u^n \tag{I.2.5}$$

$$Du^{n+1} = \frac{3}{2}u^{n+1} - 2u^n + \frac{1}{2}u^{n-1} \qquad u^{\star,n+1} = 2u^n - u^{n-1} \tag{I.2.6}$$

Le premier choix (I.2.5) correspond à une approximation d'ordre un, le second (I.2.6) à une approximation d'ordre deux.

Du fait de la nature de point-selle du problème (I.2.4), sa résolution est très consommatrice en temps CPU, ce qui rend attractives les stratégies reposant sur un découplage en plusieurs étapes. La méthode projection-pénalité appartient à cette famille de schéma. Comme il est usuel dans les méthodes de correction de pression, la première étape consiste à résoudre l'équation de quantité de mouvement en utilisant la valeur de la pression au temps précédent. L'originalité du schéma que nous proposons tient dans l'ajout durant cette étape d'un terme de pénalisation construit

à partir de la contrainte de divergence nulle, ce qui conduit au problème elliptique suivant pour la vitesse prédite $\tilde{u}^{n+1}$ au temps t^{n+1} :

$$\begin{aligned} &\varrho\Big[\frac{\beta_q\tilde{u}^{n+1}-\sum_{j=0}^{q-1}\beta_j u^{n-j}}{\Delta t}+(u^{\star,n+1}\cdot\nabla)\tilde{u}^{n+1}\Big] \\ &\qquad -r\nabla(\nabla\cdot\tilde{u}^{n+1}) = -\nabla p^n+\nabla\cdot\tau(\tilde{u}^{n+1})+f^{n+1} \end{aligned} \qquad \text{dans } \Omega \qquad \text{(I.2.7a)}$$

$$\tilde{u}^{n+1} = u_{\mathrm{D}}^{n+1} \qquad \text{sur } \partial\Omega_{\mathrm{D}} \qquad \text{(I.2.7b)}$$

$$-p^n n+\tau(\tilde{u}^{n+1})\cdot n = f_{\mathrm{N}}^{n+1} \qquad \text{sur } \partial\Omega_{\mathrm{N}} \qquad \text{(I.2.7c)}$$

où r est un paramètre de pénalisation à définir.

Soit H la variété affine de champs de vecteur à divergence nulle :

$$\mathrm{H} = \{v\in[\mathrm{L}^2(\Omega)]^d,\ \nabla\cdot v=0,\ v\cdot n = u_{\mathrm{D}}^{n+1}\cdot n \text{ sur } \partial\Omega_{\mathrm{D}}\}$$

La seconde étape de la méthode de projection est choisie de manière à réaliser la projection orthogonale par rapport au produit scalaire de L^2 de la vitesse prédite $\tilde{u}^{n+1}$ sur H, qui prend la forme générale suivante :

$$\beta_q\varrho\,\frac{u^{n+1}-\tilde{u}^{n+1}}{\Delta t}+\nabla\phi = 0 \qquad \text{dans } \Omega \qquad \text{(I.2.8a)}$$

$$\nabla\cdot u^{n+1} = 0 \qquad \text{dans } \Omega \qquad \text{(I.2.8b)}$$

Pour des raisons d'efficacité, ce système de Darcy est reformulé en prenant la divergence de la première relation pour obtenir un problème de Poisson pour ϕ, auquel il convient d'adjoindre des conditions aux limites sur $\partial\Omega_{\mathrm{D}}$ et $\partial\Omega_{\mathrm{N}}$. La première vient de la définition de H :

$$u^{n+1}\cdot n = \tilde{u}^{n+1}\cdot n = u_{\mathrm{D}}^{n+1}\cdot n \qquad \text{sur } \partial\Omega_{\mathrm{D}}$$

et par conséquent :

$$\nabla\phi\cdot n = 0 \qquad \text{sur } \partial\Omega_{\mathrm{D}}$$

La condition aux limites sur $\partial\Omega_{\mathrm{N}}$ est établie en exploitant la condition de L^2-orthogonalité pour la projection sur H, qui s'écrit :

$$\int_\Omega (u^{n+1}-\tilde{u}^{n+1})\cdot(u^{n+1}-v) = 0 \qquad \forall v\in\mathrm{H}$$

Gràce à la relation (I.2.8a), en intégrant par partie et en utilisant la définition de H, on obtient :

$$\begin{aligned} 0 &= \int_\Omega \nabla\phi\cdot(u^{n+1}-v) \\ &= \int_{\partial\Omega}\phi\ (u^{n+1}-v)\cdot n - \int_\Omega \phi\ \nabla\cdot(u^{n+1}-v) \\ &= \int_{\partial\Omega_{\mathrm{N}}}\phi\ (u^{n+1}-v)\cdot n \qquad\qquad \forall v\in\mathrm{H} \end{aligned}$$

ce qui est satisfait si la condition de Dirichlet sur la frontière pour ϕ est :

$$\phi = 0 \qquad \text{sur } \partial\Omega_{\mathrm{N}}$$

L'inconnue ϕ est donc la solution du problème elliptique suivant :

$$\Delta\phi = \frac{\beta_q\ \varrho}{\Delta t}\ \nabla\cdot\tilde{u}^{n+1} \qquad \text{dans } \Omega \qquad \text{(I.2.9a)}$$

$$\nabla\phi\cdot n = 0 \qquad \text{sur } \partial\Omega_{\mathrm{D}} \qquad \text{(I.2.9b)}$$

$$\phi = 0 \qquad \text{sur } \partial\Omega_{\mathrm{N}} \qquad \text{(I.2.9c)}$$

La vitesse de fin de pas est calculée *a posteriori* par la relation (I.2.8a) :

$$u^{n+1} = \tilde{u}^{n+1} - \frac{\Delta t}{\beta_q\ \varrho}\ \nabla\phi \qquad \text{(I.2.10)}$$

Pour obtenir une expression pour une approximation de la pression au temps t^{n+1}, nous reconstruisons une discrétisation de l'équation de quantité de mouvement au temps t^{n+1} en ajoutant les relations (I.2.7a) et (I.2.8a) :

$$\varrho\Big[\frac{\beta_q u^{n+1} - \sum_{j=0}^{q-1}\beta_j u^{n-j}}{\Delta t} + (u^{\star,n+1}\cdot\nabla)\tilde{u}^{n+1}\Big] = -\nabla(p^n - r\nabla\cdot\tilde{u}^{n+1} + \phi) + \nabla\cdot\tau(\tilde{u}^{n+1}) + f^{n+1} \qquad \text{dans } \Omega \quad \text{(I.2.11)}$$

Ce qui suggère l'expression suivante pour p^{n+1} :

$$p^{n+1} = p^n - r\nabla\cdot\tilde{u}^{n+1} + \phi \qquad \text{(I.2.12)}$$

En résumé, l'algorithme correspondant à un pas de temps pour la méthode de projection-pénalité consiste à résoudre successivement les équations (I.2.7) et (I.2.9) puis à réactualiser la vitesse et la pression de fin de pas avec (I.2.10) et (I.2.12), respectivement.

Discrétisation par la méthode des éléments finis

En fait, la formulation semi-discrète en temps présentée précédemment souffre du défaut suivant. L'analogue variationnel du terme de pénalisation s'écrit :

$$r\int_\Omega(\nabla\cdot u)(\nabla\cdot v)$$

Or la contrainte $\nabla\cdot u = 0$ n'est imposée qu'au sens faible suivant :

$$\int_\Omega(\nabla\cdot u)\,t = 0 \qquad \forall t\in M_h$$

ce qui fait que, à moins que l'espace constitué des divergences des fonctions de l'espace d'approximation de vitesse coïncide avec l'espace d'approximation de la pression M_h, ce qui est faux en général, le terme de pénalisation ne s'annule pas pour la solution. Cette dernière va donc dépendre de la valeur de r, ce qui est dommageable dans la mesure où ce paramètre a vocation à prendre des valeurs importantes. L'expérience montre que, lorsque c'est le cas, la solution est gravement perturbée.

Pour contourner cette difficulté, il convient de construire le terme de pénalisation à partir de la contrainte discrète. Cette dernière s'écrit sous la forme :

$$\mathbf{B}\mathbf{U}_{\mathrm{F}} = \mathbf{G}$$

où $\mathbf{U}_{\mathrm{F}}$ désigne le vecteur rassemblant les degrés de liberté non fixés (*i.e.* d'où sont exclus les noeuds situés sur une fontière du domaine où la vitesse est prescrite) et l'opérateur de divergence discret $\mathbf{B}$ ainsi que le second membre $\mathbf{G}$ sont construits par la technique habituelle en éléments finis. Le terme de pénalisation s'obtient alors en prémultipliant cette équation par $\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}$:

$$\text{terme de pénalisation} = r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}\mathbf{U}_{\mathrm{F}} - \mathbf{G})$$

où $\mathbf{M}_{\mathbf{p}}$ est un opérateur de "rescaling", dont la présence a pour objet de rendre les propriétés du système algébrique aussi indépendantes du pas de maillage que possible ; typiquement, on obtient $\mathbf{M}_{\mathbf{p}}$ à partir de la matrice de masse de pression par condensation sur la diagonale.

La première étape devient alors :

$$\frac{\beta_q}{\Delta t}\mathbf{M}\widetilde{\mathbf{U}}_{\mathrm{F}} + \mathbf{A}\widetilde{\mathbf{U}}_{\mathrm{F}} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathrm{exp}} = \mathbf{F} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{G} \tag{I.2.13}$$

où la matrice de masse $\mathbf{M}$, l'opérateur de convection-diffusion discret $\mathbf{A}$ et le second membre $\mathbf{F}$ résultent, encore une fois, de l'utilisation de la machinerie éléments finis usuelle.

La partie analogue à l'étape de projection (I.2.8) s'écrit sous la forme algébrique :

$$\begin{cases} \dfrac{\beta_q}{\Delta t}\mathbf{M}(\mathbf{U}_{\mathrm{F}} - \widetilde{\mathbf{U}}_{\mathrm{F}}) + \mathbf{B}^{\mathrm{T}}\mathbf{\Phi} = 0 \\ \mathbf{B}\mathbf{U}_{\mathrm{F}} = \mathbf{G} \end{cases} \tag{I.2.14}$$

Lorsque l'on calcule $\mathbf{U}_{\mathrm{F}}$ par la première équation, on multiplie le résultat par $\mathbf{B}$ et on utilise la relation $\mathbf{B}\mathbf{U}_{\mathrm{F}} = \mathbf{G}$, on obtient l'équation suivante, où seule la variable de projection $\mathbf{\Phi}$ apparaît :

$$\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{\Phi} = \frac{\beta_q}{\Delta t}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G})$$

La matrice de masse de vitesse $\mathbf{M}$ n'étant pas, dans le cas général, diagonale, la résolution d'un tel système est coûteuse en temps de calcul ; l'écriture du schéma sous forme semi-discrète (relation (I.2.9)) suggère de lui substituer le problème suivant :

$$\mathbf{L}\mathbf{\Phi} = \frac{\beta_q\,\varrho}{\Delta t}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \tag{I.2.15}$$

où $\mathbf{L}$ est l'opérateur discret associé au problème de Poisson avec des conditions aux limites de Dirichlet homogène sur $\partial\Omega_{\mathrm{N}}$ et de Neumann homogène sur $\partial\Omega_{\mathrm{D}}$. Une fois $\mathbf{\Phi}$ calculé, il ne reste plus qu'à réactualiser la pression grâce à la relation :

$$\mathbf{P} = \mathbf{P}_{\mathrm{exp}} + \mathbf{\Phi} + r\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \tag{I.2.16}$$

Au final, la formulation algébrique complète d'un pas de temps de la méthode de projection-pénalité s'écrit :

$$\begin{cases} \big[\frac{\beta_q}{\Delta t}\mathbf{M} + \mathbf{A} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{B}\big]\widetilde{\mathbf{U}}_{\mathrm{F}} + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathrm{exp}} = \mathbf{F} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{G} \\ \mathbf{L}\boldsymbol{\Phi} = \frac{\beta_q\,\varrho}{\Delta t}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \\ \mathbf{U}_{\mathrm{F}} = \widetilde{\mathbf{U}}_{\mathrm{F}} - \frac{\Delta t}{\beta_q}\mathbf{M}^{-1}\mathbf{B}^{\mathrm{T}}\boldsymbol{\Phi} \\ \mathbf{P} = \mathbf{P}_{\mathrm{exp}} + \boldsymbol{\Phi} + r\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \end{cases} \tag{I.2.17}$$

Expérimentations numériques

Nous avons effectué différents tests numériques, dans le but de tester les propriétés de convergence de la méthode de projection-pénalité ; à cette occasion, nous comparons son comportement avec celui de méthodes classiques.

Dans tous les tests présentés ici, vitesse et pression sont discrétisées par des éléments finis $\mathbf{P}_2$-$\mathbf{P}_1$ (appelés éléments finis de Taylor-Hood, *c.f.* par exemple [11], [9, chap. 5], [27, chap. 9]) et la discrétisation en temps est d'ordre deux.

Un problème de Navier-Stokes avec des conditions aux limites de type Dirichlet. Le premier cas test effectué est celui dit des "tourbillons de Green-Taylor" ([33], [23]), problème classique possèdant une solution analytique. Les figures I.1 et I.2 donnent, en fonction du pas de temps, la norme de la différence entre la solution analytique et les résultats numériques à l'instant final du calcul :

- avec la méthode de projection incrémentale usuelle, introduite pour la première fois, semble-t-il par Goda [12], analysée dans le cas semi-discret en temps par Shen [28, 30, 32] et dans le cas discret par Guermond et Quartapelle [16, 15],
- avec la méthode de projection dite "rotationnelle" introduite par Timmermans *et al.* [35] et analysée par Guermond et Shen [21],
- avec la méthode de pénalité projection, pour différentes valeurs du paramètre de pénalisation,
- et avec la méthode semi-implicite, *i.e.* où l'on réalise la résolution couplée du bilan de quantité de mouvement, avec une vitesse d'advection explicite, et de la contrainte divergentielle (système (I.2.4)).

Ces courbes montrent tout d'abord une décroissance de l'erreur avec le pas de temps, puis survient un plateau correspondant à l'erreur résiduelle de la discrétisation spatiale. Nous observons que, à fort pas de temps, les méthodes de projection non pénalisées (*i.e.* la méthode incrémentale et rotationnelle) sont beaucoup moins précises que la méthode semi-implicite (erreur jusqu'à 100 fois plus grande) ; la pénalisation a pour effet de réduire cet écart, jusqu'à ce que les résultats de la méthode de projection-pénalité et ceux de la méthode semi-implicite coïncident, pour des valeurs de r typiquement de l'ordre de 10^3. Le schéma semi-implicite converge avec un ordre deux en temps, tant pour la vitesse que pour la pression. Pour la méthode de projection incrémentale, conformément aux résultats prouvés dans [17], la convergence de la vitesse est d'ordre deux, tandis que celle de la pression est plus lente. Les résultats de le méthode de projection-pénalité sont toujours compris entre les courbes correspondant à ces deux schémas, c'est à dire que la méthode de projection-pénalité est toujours plus précise que la méthode de projection incrémentale et moins précise

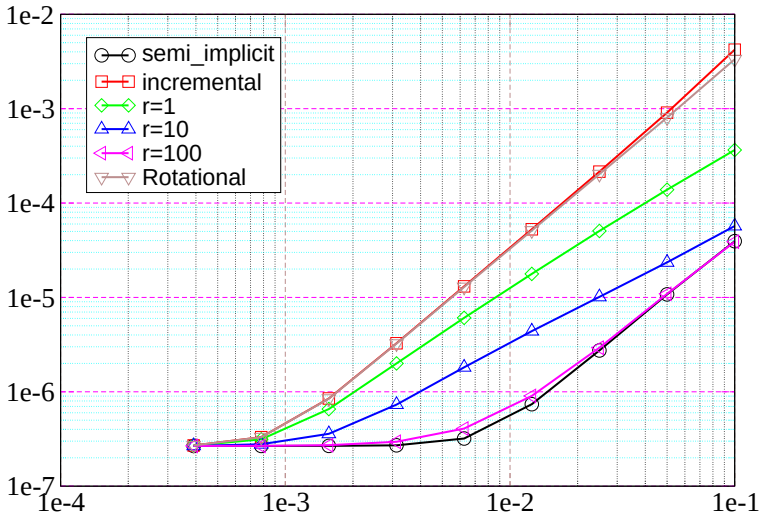


FIG. I.1 – Tourbillon de Green-Taylor - Erreur en norme L^2 pour la vitesse à un instant donné en fonction du pas de temps, pour les méthodes de projection incrémentale, rotationnelle, la méthode de projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et le schéma semi-implicite.

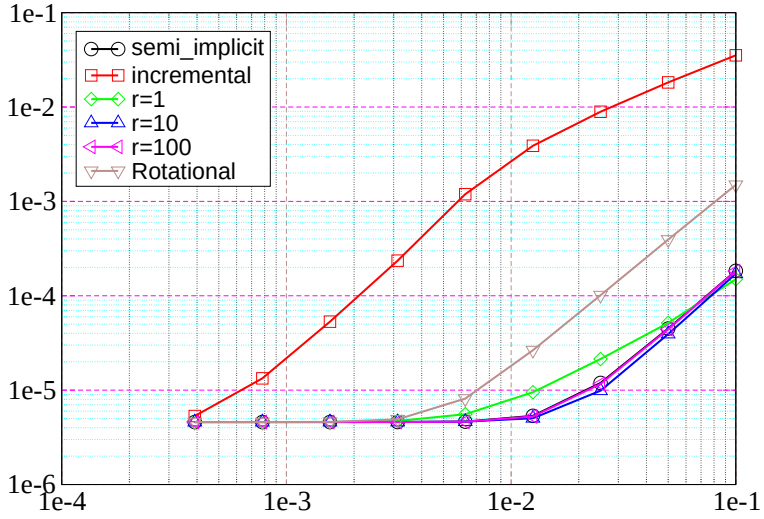


FIG. I.2 – Tourbillon de Green-Taylor - Erreur en norme L^2 pour la pression à un instant donné en fonction du pas de temps, pour les méthodes de projection incrémentale, rotationnelle, la méthode de projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et le schéma semi-implicite.

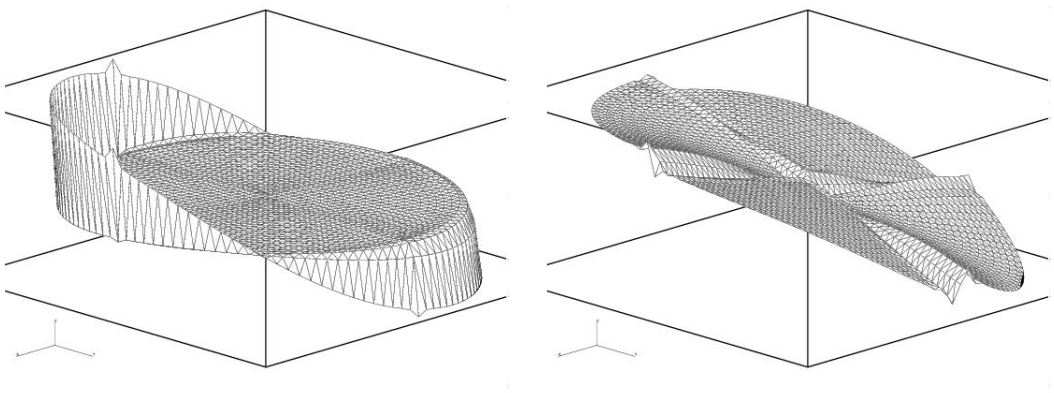
que le schéma semi-implicite. Pour des valeurs significatives du paramètre de pénalité r, les gains en précision sont très importants ; il est à noter, toutefois, que le fait d'augmenter r a pour conséquence de dégrader le conditionnement du système linéaire, d'où une augmentation sensible du temps de calcul lorsque l'on utilise, pour la résolution des systèmes linéaires, des méthodes de Krylov.

Couche limite de pression pour un problème de Stokes avec des conditions aux limites de type Dirichlet. Nous avons vu (équation (I.2.9)) que la conservation par l'étape de projection de la vitesse de normale sur les frontières de

Dirichlet conduit à imposer des conditions aux limites de type Neumann homogène sur la variable de projection ϕ. Dans la mesure où, pour la projection incrémentale, ϕ n'est autre que l'incrément de pression, ces conditions aux limites se transmettent, par récurrence d'un pas de temps sur l'autre, à la pression elle-même, d'où l'apparition d'une couche limite sur l'erreur en pression. Ce phénomème affecte les méthodes utilisant, comme ici, un laplacien de pression [13, 14, 27, 36], [20, section 6] comme celles, dites algébriques, effectuant une résolution directe de l'étape de projection comme un problème de Darcy [36, pp. 53-54], [27]).

Pour la méthode de projection-pénalité, la formulation de l'incrément de pression ne reporte pas les conditions aux limites de ϕ sur la pression, ce qui laisse espérer que ce défaut soit corrigé.

La figure I.3 représente la distribution spatiale de l'erreur de pression obtenue pour un problème de Stokes instationnaire, à un instant et pour un pas de temps donné, pour le méthode de projection incrémentale et de projection-pénalité, avec un paramètre de pénalisation r égal à 10^2 ; cette figure confirme bien la disparition des couches limites d'erreur de pression.



projection incrémentale
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 1.53\,10^{-2}$

projection-pénalité ($r = 10^2$)
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 2.84\,10^{-4}$

FIG. I.3 – Distribution de l'erreur en pression à un instant donné pour un problème de Stokes avec des conditions aux limites de Dirichlet

Un problème de Stokes avec des conditions aux limites ouvertes. Sur les frontières où l'on impose des conditions aux limites de Neumann (dites également "conditions aux limites ouvertes"), la variable de projection ϕ est imposée à zéro. Pour la même raison que précédemment, cette condition aux limites se transmet à la pression pour la méthode de projection incrémentale ; Guermond et Shen [19] démontrent que, dans ce cas, la convergence spatiale est drastiquement réduite (ordre 1/2 en norme L^2 pour la pression). Par contre, la pression n'est pas directement affectée par cette condition aux limites artificielle dans la méthode de projection-pénalité.

Les figures I.4 et I.5 représentent les normes de la différence entre la solution analytique d'un problème de Stokes instationnaire et les solutions obtenues avec les différentes méthodes considérées ici, en fonction du pas de temps et à un instant donné. On observe la perte de convergence prédite par la théorie pour la méthode incrémentale ; la méthode de projection-pénalité ne souffre pas du même défaut.

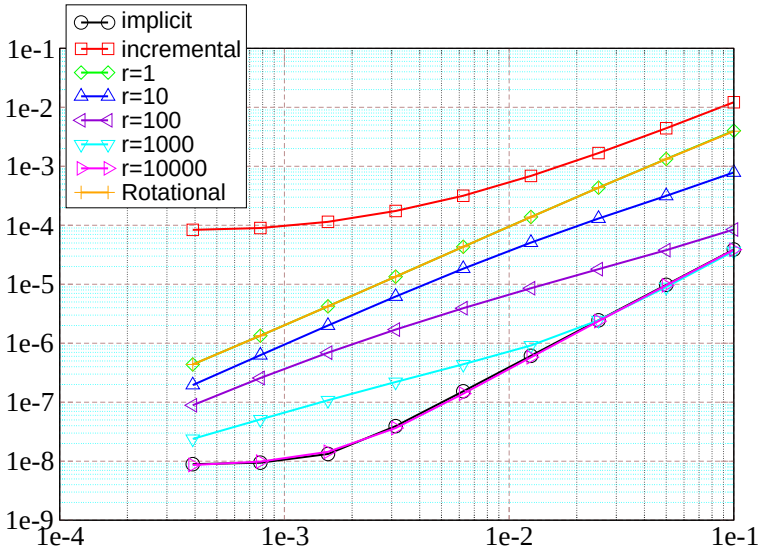


FIG. I.4 – Problème de Stokes avec des conditions aux limites ouvertes - Erreur en norme L^2 pour la vitesse à un instant donné en fonction du pas de temps, pour les méthodes de projection incrémentale, rotationnelle, la méthode de projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et le schéma semi-implicite.

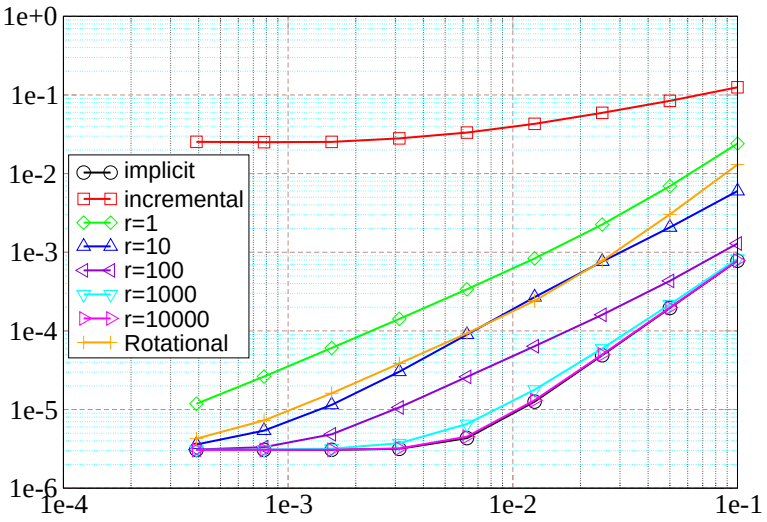


FIG. I.5 – Problème de Stokes avec des conditions aux limites ouvertes - Erreur en norme L^2 pour la pression à un instant donné en fonction du pas de temps, pour les méthodes de projection incrémentale, rotationnelle, la méthode de projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et le schéma semi-implicite.

Ecoulement autour d'un cylindre. Une série de tests a été effectuée sur un problème plus concret, à savoir l'écoulement d'un fluide Newtonien autour d'un cylindre. Le maillage utilisé est tracé sur la figure I.6. Nous rapportons sur la figure I.7 la norme L^2 de l'erreur de fractionnement (*i.e.* la différence entre la solution obtenue par la méthode de projection-pénalité et celle obtenue avec le schéma semi-implicite) à un instant donné, pour plusieurs pas de temps et en fonction du paramètre de pénalité r. Nous observons que l'erreur de fractionnement diminue avec le paramètre

de pénalisation quelque soit le pas de temps choisi, avec un comportement asymptotique en $1/r$.

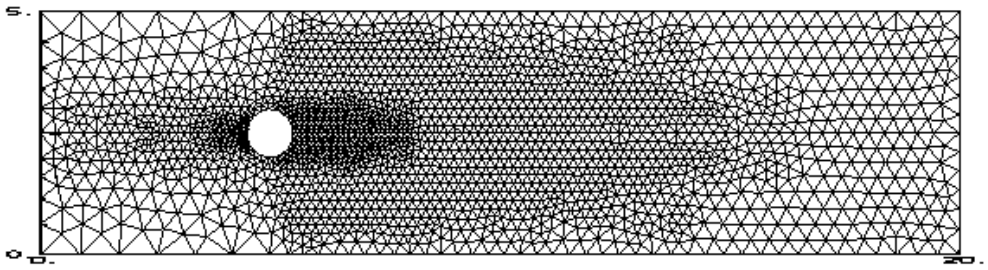


FIG. I.6 – Ecoulement derrière un cylindre - maillage utilisé.

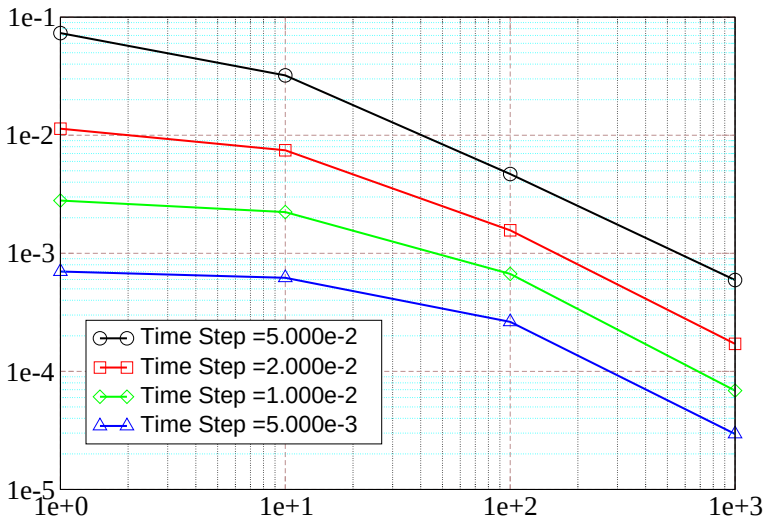


FIG. I.7 – Ecoulement derrière un cylindre - Erreur de fractionnement en norme L^2 pour la vitesse en fonction du paramètre de pénalisation r, pour la méthode de projection-pénalité.

I.3 Analyse de la projection-pénalité pour le problème de Stokes instationnaire

Cette partie est consacrée à l'analyse de la méthode de projection-pénalité, sur le cas modèle du problème de Stokes instationnaire. La technique adoptée est la suivante : nous supposons que la régularité du problème discret associé à une méthode implicite est telle que les dérivées discrètes première et seconde de la solution restent bornées, et nous majorons l'erreur de fractionnement, c'est à dire la différence entre les solutions obtenues par le schéma couplé et la méthode de projection-pénalité. Outre le fait qu'elle simplifie la présentation, cette approche a le mérite essentiel de fournir potentiellement des résultats valables pour des discrétisations en temps d'ordre un ou deux : en effet, on trouvera dans [17] la preuve que l'erreur de frac-

tionnement est d'ordre deux même pour le simple schéma d'Euler implicite. C'est ainsi ce dernier que nous étudions.

Soit $(\bar{u}^k)_{k=0,N}$ et $(\bar{p}^k)_{k=0,N}$ la solution obtenue par le schéma implicite et u^k, $\tilde{u}^k$ et p^k celle obtenue par la méthode de projection-pénalité. On définit alors :

$$e^k = u^k - \bar{u}^k, \qquad \tilde{e}^k = \tilde{u}^k - \bar{u}^k, \qquad \epsilon^k = p^k - \bar{p}^k$$

En utilisant des techniques sensiblement inspirées de l'analyse de la méthode de projection rotationnelle [21] pour les faibles valeurs du paramètre de pénalisation et de l'analyse des méthodes de pénalité [31] pour les fortes valeurs de r, nous obtenons les deux résultats suivants.

Theorem I.3.1 (Faibles valeurs du paramètre de pénalisation r)**.** *Les majorations suivantes sont obtenues pour* $1 \leq n \leq N$ *:*

$$\left[\sum_{k=0}^{n} \Delta t \, \|e^{k+1}\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \Delta t \, \|\tilde{e}^{k+1}\|_0^2\right]^{1/2} \leq c \ \min(\Delta t^2, \frac{\Delta t^{3/2}}{r^{1/2}})$$
$$\left[\sum_{k=0}^{n} \Delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \Delta t \, \|\epsilon^{k+1}\|_0^2\right]^{1/2} \leq c \ \max(1, \frac{1}{r^{1/2}}) \ \Delta t^{3/2}$$

Theorem I.3.2 (Fortes valeurs du paramètre de pénalisation r)**.** *Pour toutes valeurs strictement positive du paramètre de pénalisation* r*, nous obtenons les majorations suivantes pour* $1 \leq n \leq N$ *:*

$$\|e^n\|_0 \ + \ \|\tilde{e}^n\|_0 \ + \left[\sum_{k=1}^{n} \Delta t \, \|\nabla \tilde{e}^k\|_0^2\right]^{1/2} \leq c \ \frac{\Delta t^{1/2}}{r}$$
$$\left[\sum_{k=1}^{n} \Delta t \, \|e^k\|_0^2\right]^{1/2} + \left[\sum_{k=1}^{n} \Delta t \, \|\tilde{e}^k\|_0^2\right]^{1/2} \leq c \ \frac{\Delta t}{r}$$

Ces résultats confirment que la méthode de projection-pénalité se comporte comme on pouvait l'espérer, à savoir comme la méthode de pénalité lorsque r prend des valeurs importantes et comme la méthode de projection rotationnelle lorsque r reste faible. Sur ce dernier point, il ressort également de l'analyse que la méthode est stable quelque soit la valeur que prend le paramètre r dans la relation donnant l'incrément de pression, alors que la méthode de projection rotationnelle, formellement identique à l'omission près du terme de pénalité dans l'étape de prédiction de vitesse, ne l'est que pour $r \leq \mu$; l'ajout de ce terme de pénalisation peut ainsi être vu comme une stabilisation de la méthode rotationnelle, potentiellement utile lorsque la viscosité est variable en temps et/ou en espace.

Les résultats des tests numériques sont illustrés sur les figures I.8 et I.9 ; ils sont conformes à l'analyse.

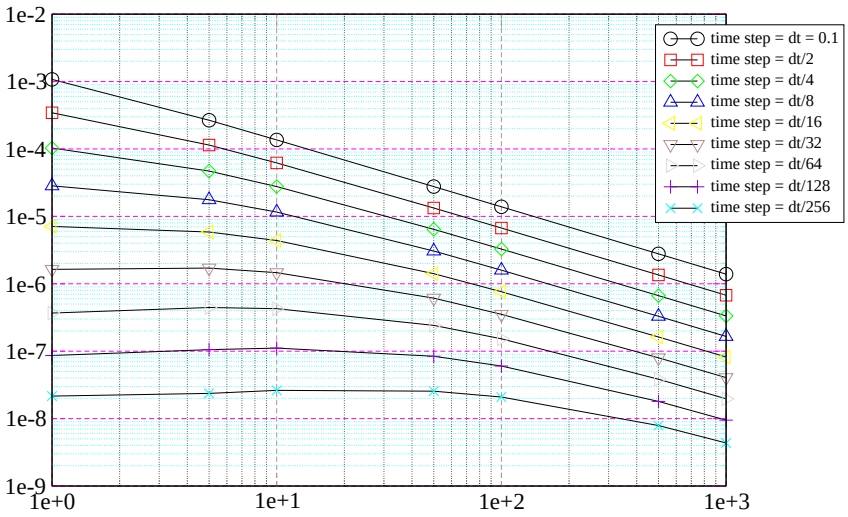


FIG. I.8 – Norme L^2 de $\tilde{e}$ à un instant donné, en fonction du paramètre de pénalisation.

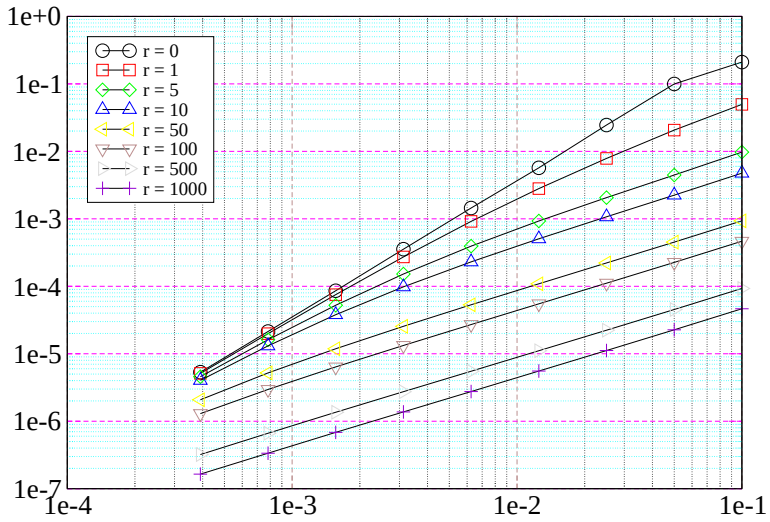


FIG. I.9 – Norme L^2 de ϵ à un instant donné, en fonction du pas de temps.

I.4 Une méthode de projection-pénalité pour un écoulement dilatable

Description de la méthode

L'objet de cette section est le développement d'une méthode de projection-pénalité pour la résolution du problème suivant :

$$\left|\begin{array}{ll} \dfrac{\partial \varrho u}{\partial t} + \nabla \cdot (\varrho u \otimes u) = \nabla \cdot \tau(u) - \nabla p + f & \text{dans } [0,T] \times \Omega \\ \nabla \cdot (\varrho u) + \dfrac{\partial \varrho}{\partial t} = 0 & \text{dans } [0,T] \times \Omega \\ u = u_{\mathrm{D}} & \text{sur } [0,T] \times \partial\Omega_{\mathrm{D}} \\ -pn + \tau(u) \cdot n = f_{\mathrm{N}} & \text{sur } [0,T] \times \partial\Omega_{\mathrm{N}} \end{array}\right.$$

où $\varrho = \varrho(x,t)$ est une fonction donnée.

Par principe des méthodes de projection, l'équation bilan de quantité de mouvement est résolue dans une première étape en utilisant la pression au temps précédent pour obtenir une prédiction de vitesse, avant de corriger cette denière pour vérifier le bilan de masse dans une seconde étape. L'étape de prédiction mise en œuvre ici s'écrit :

$$\left|\begin{array}{ll} \dfrac{\varrho^{n+1}\tilde{u}^{n+1} - \sum_{j=0}^{q-1} \beta_j q_h^{n-j}}{\Delta t} + \nabla \cdot (q^{\star,n+1} \otimes \tilde{u}^{n+1}) & \\ \quad -r\nabla(\nabla \cdot \varrho^{n+1}\tilde{u}^{n+1} + \dfrac{D\varrho^{n+1}}{\Delta t}) = \nabla \cdot \tau(\tilde{u}^{n+1}) - \nabla p^n + f^{n+1} & \text{dans } \Omega \\ \tilde{u}^{n+1} = u_{\mathrm{D}}^{n+1} & \text{sur } \partial\Omega_{\mathrm{D}} \\ -p^n n + \tau(\tilde{u}^{n+1}) \cdot n = f_{\mathrm{N}}^{n+1} & \text{sur } \partial\Omega_{\mathrm{N}} \end{array}\right. \tag{I.4.18}$$

où $q = (\varrho u)$ est une nouvelle variable représentant le débit.

La spécificité de cette méthode réside dans l'ajout du terme de pénalisation :

$$-r\nabla(\nabla \cdot \varrho^{n+1}\tilde{u}^{n+1} + \frac{D\varrho^{n+1}}{\Delta t})$$

Ce terme est obtenu en appliquant l'opérateur gradient à l'équation de bilan de masse ; cette opération est bien sûr formelle pour le problème différentiel, mais ses analogues naturels variationnel et surtout, pour le système discret, algébrique peuvent, quant à eux, être précisément définis.

Du fait notamment de la présence des termes de viscosité, l'inconnue naturelle de cette première étape est la vitesse $\tilde{u}^{n+1}$. Il est à noter alors que, contrairement à ce qui se passe lors d'une augmentation classique, le terme rajouté a pour contrepartie variationnelle une forme bilinéaire s'appliquant à la vitesse qui n'est ni symétrique, ni positive. Cette caractéristique pourrait être corrigée en choisissant, à la place du terme proposé, l'expression suivante :

$$-r\varrho^{n+1}\nabla(\nabla \cdot \varrho^{n+1}\tilde{u}^{n+1} + \frac{D\varrho^{n+1}}{\Delta t})$$

Ce dernier choix présente l'inconvénient majeur que, ne pouvant plus s'écrire comme le gradient d'une quantité, il ne s'intègre pas naturellement dans l'incrément de pression. En outre, les techniques classiques d'étude des problèmes elliptiques permettent de prouver que, pour un pas de temps plus petit qu'une valeur seuil Δt_0 indépendante du paramètre de pénalisation r, l'équation aux dérivées partielles (I.4.18) admet une solution et une seule.

Pour compléter l'algorithme, il convient maintenant de construire l'étape de projection. Soit H l'espace affine défini comme suit :

$$H = \{q \in [\mathrm{L}(\Omega)]^d, \quad \nabla \cdot q = -\frac{D\varrho^{n+1}}{\Delta t}, \quad q \cdot n = \varrho^{n+1}\, u_{\mathrm{D}} \cdot n \text{ sur } \partial\Omega_{\mathrm{D}}\}$$

La seconde étape revient alors à effectuer une projection orthogonale pour le produit scalaire de L^2 du débit prédit (*i.e.* $\varrho^{n+1}\tilde{u}^{n+1}$) sur H ; on obtient alors le système suivant :

$$\left| \begin{array}{ll} \beta_q \dfrac{q_h^{n+1} - \varrho^{n+1}\tilde{u}^{n+1}}{\Delta t} + \nabla\phi = 0 & \text{dans } \Omega \\ \nabla \cdot q_h^{n+1} = -\dfrac{D\varrho^{n+1}}{\Delta t} & \text{dans } \Omega \end{array} \right. \tag{I.4.19}$$

auquel il convient d'adjoindre des conditions aux limites pour ϕ. Sur les frontières où la vitesse est fixée, on a, d'après la définition de H :

$$q_h^{n+1} \cdot n = \varrho^{n+1}\tilde{u}^{n+1} \cdot n = \varrho^{n+1} u_{\mathrm{D}}^{n+1} \cdot n$$

Par conséquent :

$$\nabla\phi \cdot n = 0 \qquad \text{sur } \partial\Omega_{\mathrm{D}}$$

Sur les frontières de Neumann, la condition aux limites provient de la condition de L^2-orthogonalité de la projection sur H. Cette dernière s'écrit :

$$\int_\Omega (q_h^{n+1} - \varrho^{n+1}\tilde{u}^{n+1}) \cdot (q_h^{n+1} - v) = 0 \quad \forall v \in H$$

Grâce à la première relation de (I.4.19) que l'on intègre par partie, puis par définition de H, on a :

$$\begin{aligned} 0 &= \int_\Omega \nabla\phi \cdot (q_h^{n+1} - v) \\ &= -\int_\Omega \phi\ \nabla \cdot (q_h^{n+1} - v) + \int_{\partial\Omega_{\mathrm{N}}} \phi\ (q_h^{n+1} - v) \cdot n \\ &= \int_{\partial\Omega_{\mathrm{N}}} \phi\ (q_h^{n+1} - v) \cdot n \end{aligned}$$

Cette relation nous donne la condition aux limites à appliquer à ϕ sur la frontière $\partial\Omega_{\mathrm{N}}$:

$$\phi = 0 \qquad \text{sur } \partial\Omega_{\mathrm{N}}$$

En rassemblant les relations obtenues, l'inconnue ϕ est solution du problème suivant :

$$\left| \begin{array}{ll} \Delta\phi = \dfrac{\beta_q}{\Delta t}\left(\nabla \cdot \varrho^{n+1}\tilde{u}^{n+1} + \dfrac{D\varrho^{n+1}}{\Delta t}\right) & \text{dans } \Omega \\ \nabla\phi \cdot n = 0 & \text{sur } \partial\Omega_{\mathrm{D}} \\ \phi = 0 & \text{sur } \partial\Omega_{\mathrm{N}} \end{array} \right. \tag{I.4.20}$$

Une fois ϕ calculée, la première relation de (I.4.19) permet de réactualiser le débit :

$$q_h^{n+1} = \varrho^{n+1}\tilde{u}^{n+1} + \frac{\Delta t}{\beta_q}\nabla\phi \quad \text{dans } \Omega \tag{I.4.21}$$

Enfin, si l'on somme cette même relation avec l'équation de prédiction (I.4.18), on obtient :

$$\begin{aligned}&\frac{\varrho^{n+1}\tilde{u}^{n+1} - \sum_{j=0}^{q-1}\beta_j q_h^{n-j}}{\Delta t} + \nabla\cdot(q^{\star,n+1}\otimes\tilde{u}^{n+1})\\ &\quad = \nabla\cdot\tau(\tilde{u}^{n+1}) - \nabla(p^n - r\left[\nabla\cdot\varrho^{n+1}\tilde{u}^{n+1} + \frac{D\varrho^{n+1}}{\Delta t}\right] + \phi) + f^{n+1}\end{aligned}$$

Cette relation n'est rien d'autre que la reconstitution de l'équation de quantité de mouvement, ce qui suggère l'expression suivante pour la pression en fin de pas :

$$p^{n+1} = p^n - r\left[\nabla\cdot\varrho^{n+1}\tilde{u}^{n+1} + \frac{D\varrho^{n+1}}{\Delta t}\right] + \phi \quad \text{dans } \Omega \tag{I.4.22}$$

En conclusion, effectuer un pas de temps consiste donc à résoudre en séquence les problèmes elliptiques (I.4.18) et (I.4.20) puis à réactualiser le débit et la pression par respectivement (I.4.21) et (I.4.22). Toutefois, comme précédemment exposé dans le cas d'un écoulement incompressible, du fait que le bilan de masse n'est vérifié qu'au sens faible, il convient de construire le terme de pénalisation à partir du système discret. Cette construction, pour laquelle on procède strictement de la même manière qu'en incompressible, n'est pas détaillée ici.

Les résultats d'une expérimentation numérique sont reportés sur les figures I.10 et I.11. Ils correspondent à la résolution des équations de Navier-Stokes avec des conditions aux limites de Dirichlet. On retrouve des résultats similaires à ceux rencontrés pour des écoulements incompressibles. Ce constat reste encore valable si l'on impose des conditions aux limites de Neumann sur une partie de la frontière.

Application à la résolution des équations gouvernant un écoulement à faible nombre de Mach

Le système d'équations aux dérivées partielles que nous étudions dans cette section est un modèle asymptotique établi à partir des équations de la dynamique des fluides compressibles en faisant l'hypothèse que la vitesse dans l'écoulement reste faible devant la vitesse des ondes de pression (vitesse du son) [26] ; le nombre de Mach est défini comme le rapport entre la vitesse matérielle et la vitesse du son et un tel écoulement est dit "à faible nombre de Mach". Ce système d'équations s'écrit sous forme adimensionnée de la manière suivante :

$$\left|\begin{array}{l}\dfrac{\partial\varrho}{\partial t} + \nabla\cdot(\varrho u) = 0\\ \dfrac{\partial\varrho u}{\partial t} + \nabla\cdot(\varrho u\otimes u) + \nabla p = \dfrac{1}{\mathrm{Re}}\nabla\cdot(\mu(\nabla u + (\nabla u)^{\mathrm{T}}))\\ \qquad\qquad -\dfrac{2}{3}\,\dfrac{1}{\mathrm{Re}}\,\nabla(\mu\nabla\cdot u) + \dfrac{1}{\mathrm{Fr}^2}\varrho(-\vec{z})\\ \varrho\left(\dfrac{\partial T}{\partial t} + u\cdot\nabla T\right) = \dfrac{1}{\mathrm{RePr}}\nabla\cdot(\lambda\nabla T) + \dfrac{\gamma-1}{\gamma}\dfrac{dP_{th}}{dt}\\ P_{th} = \varrho T\end{array}\right. \tag{I.4.23}$$

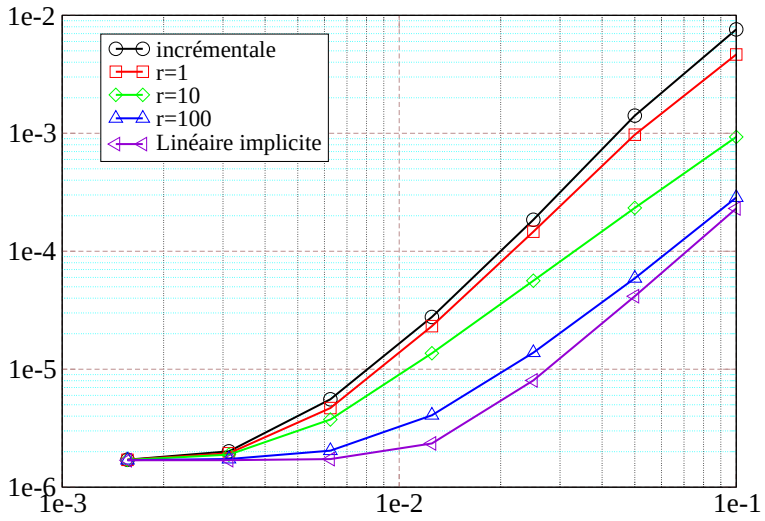


FIG. I.10 – Cas à masse volumique variable – Norme L^2 de l'erreur pour le débit à un instant donné et en fonction du pas de temps pour la méthode de projection incrementale, de projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et le schéma semi-implicite.

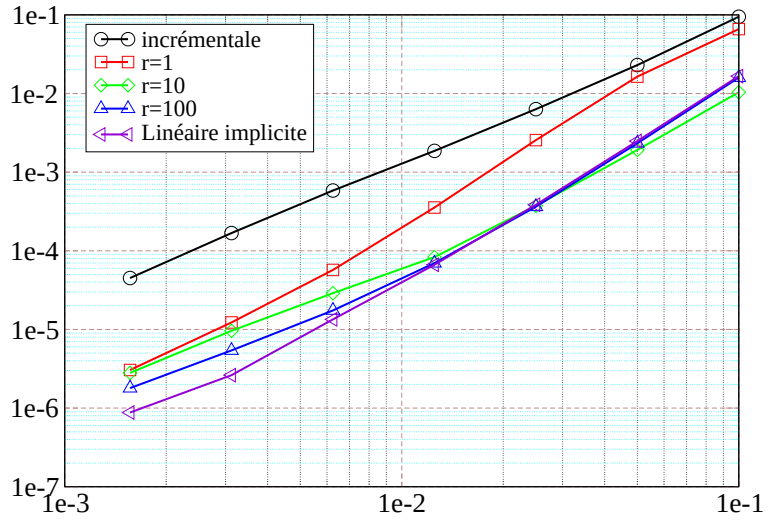


FIG. I.11 – Cas à masse volumique variable – Norme L^2 de l'erreur pour la pression à un instant donné et en fonction du pas de temps pour la méthode de projection incrementale, de projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et le schéma semi-implicite.

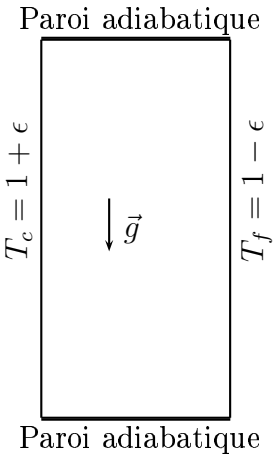
où T désigne la température, P_{th} est une quantité dépendant seulement du temps (*i.e.* constante en espace) dite pression thermodynamique, dont le calcul nécessite, dans un domaine fermé, une équation supplémentaire qui peut être simplement la conservation de la masse totale. Les nombres sans dimension intervenant dans ce

système sont les nombres de Froude, de Reynolds, de Prandlt :

$$\mathrm{Fr} = \frac{u_0}{L_0 g} \qquad \mathrm{Re} = \frac{\varrho_0 u_0 L_0}{\mu_0} \qquad \mathrm{Pr} = \frac{\mu_0 C_p}{\lambda_0}$$

où u_0, L_0, ϱ_0 et μ_0 sont des grandeurs caractéristiques pour respectivement la vitesse, la taille du domaine, la masse volumique et la viscosité, et C_p désigne la capacité calorifique du fluide, supposée constante. Du fait que la masse volumique ne dépend

dans ce modèle que de la pression thermodynamique et plus de la pression dynamique (*i.e.* de la pression qui apparaît dans le bilan de quantité de mouvement), ce système d'équations aux dérivées partielles entre dans le cadre défini en introduction (système (I.1.1)).



Le cas que nous traitons ici est un écoulement de convection naturelle dans une cavité différentiellement chauffée, de rapport d'aspect H/L égal à 4 (soit $\Omega = (0,1) \times (0,4)$).
Les nombres adimensionnels caractéristiques de l'écoulement prennent les valeurs suivantes :

$$\mathrm{Re} = 10^3,\ \mathrm{Fr} = 0.923,\ \mathrm{Pr} = 0.71,\ \gamma = 1.4$$

Des conditions aux limites de Dirichlet homogènes sont imposées à la vitesse tout le long de la frontière. En ce qui concerne le bilan d'énergie, les frontières inférieure et supérieure sont supposées adiabatiques (*i.e.* condition de Neumann homogène), tandis que la température est imposée sur les frontières verticales aux valeurs suivantes : $T_f = 1 - \epsilon$ sur la droite $x = 0$, $T_c = 1 + \epsilon$ sur la droite $x = 1$, ϵ étant un paramètre fixé à 0.6.
Le comportement thermique de la viscosité dynamique adimensionnée ainsi que la conductivité thermique adimensionnée est régi par la loi de Sutherland :

$$\mu(T) = \lambda(T) = T^{\frac{3}{2}} \left(\frac{1+S}{T+S} \right)$$

avec $S = 1.1/6$.

Ce problème reprend exactement les données du benchmark organisé par P. Le Queré et H. Paillère [24], à ceci près que la géométrie est modifiée ($\Omega = (0,1) \times (0,1)$ dans [24]). Par contre, alors que dans [24], l'objectif est le calcul de l'état stationnaire, nous nous intéressons ici essentiellement au transitoire d'établissement depuis l'état initial défini par $u = 0$ et $T = 1$.

Le schéma numérique utilisé pour la résolution de ce problème résulte du simple rajout d'une étape de résolution du bilan d'énergie aux schémas introduits et testés dans les sections précédentes. Il s'écrit de la manière suivante :

1 – Résolution de l'équation bilan d'énergie

$$\frac{\varrho^n DT^{n+1}}{\Delta t} + q^{\star,n+1}\cdot\nabla T^{n+1} - \frac{\lambda}{\mathrm{RePr}}\Delta T^{n+1} = \frac{\gamma-1}{\gamma}\frac{DP_{th}^n}{\Delta t}$$

2 – Réactualisation de la pression thermodynamique puis de la masse volumique

$$\left[\int_\Omega \frac{1}{T^{n+1}}\right] P_{th}^{n+1} = \int_\Omega \frac{P_{th}^n}{T^n}, \qquad \varrho^{n+1} = \frac{P_{th}^{n+1}}{T^{n+1}}$$

3 – Résolution de l'équation bilan de quantitité de mouvement (prédiction)

$$\frac{\varrho^{n+1}\tilde{u}^{n+1} - \sum_{j=0}^{q-1}\beta_j q_h^{n-j}}{\Delta t} + \nabla\cdot(q^{\star,n+1}\otimes\tilde{u}^{n+1}) - \nabla\cdot\tau(\tilde{u}^{n+1}) = -\nabla p^n + \frac{1}{\mathrm{Fr}^2}\varrho^{n+1}(-\vec{z})$$

4 – Etape de projection

$$\left|\begin{array}{l} -\Delta\phi = \nabla\cdot\varrho^{n+1}\tilde{u}^{n+1} + \dfrac{D\varrho^{n+1}}{\Delta t} \\ p^{n+1} = p^n + \phi \\ q^{n+1} = \varrho^{n+1}\tilde{u}^{n+1} + \dfrac{\Delta t}{\beta_q}\nabla\phi \end{array}\right.$$

La discrétisation en temps utilisée est d'ordre deux ; compte-tenu des limitations du pas de temps dues à des problèmes de précision et, semble-t-il, de stabilité, l'ajout d'un terme de pénalisation dans l'étape de prédiction n'est pas nécessaire. Vitesse et pression sont discrétisées en espace par l'élément fini de Taylor-Hood (P2-P_1) ; l'approximation de la température est du même degré que celle des composantes de la vitesse (P2).

L'écoulement tend vers un état permanent, après un transitoire d'établissement assez long (plus de 200 unités de temps), marqué par des instabilités au voisinage de la couche limite chaude. Ces instabilités sont illustrées par les figures I.12, où l'on trace les courbes isothermes à quelques instants consécutifs ; de manière qualitative, ces instabilités correspondent à des détachements de poches froides de la zone inférieure de la cavité. Les tests de sensibilité à la discrétisation spatiale et temporelle confirment que le calcul présenté, qui compte plus de 150 000 mailles pour près d'un million de degrés de liberté (en comptabilisant vitesse, pression et température), a atteint la convergence.

I.5 Résolution d'un problème de convection naturelle par une méthode d'éléments finis joints

Nous présenterons tout d'abord les principes de la méthode d'éléments finis joints (MEFJ) introduite et étudiée dans [5, 2, 3, 4, 39, 38, 6]. Puis nous exposerons les résultats obtenus par cette méthode sur un problème couplant un écoulement de

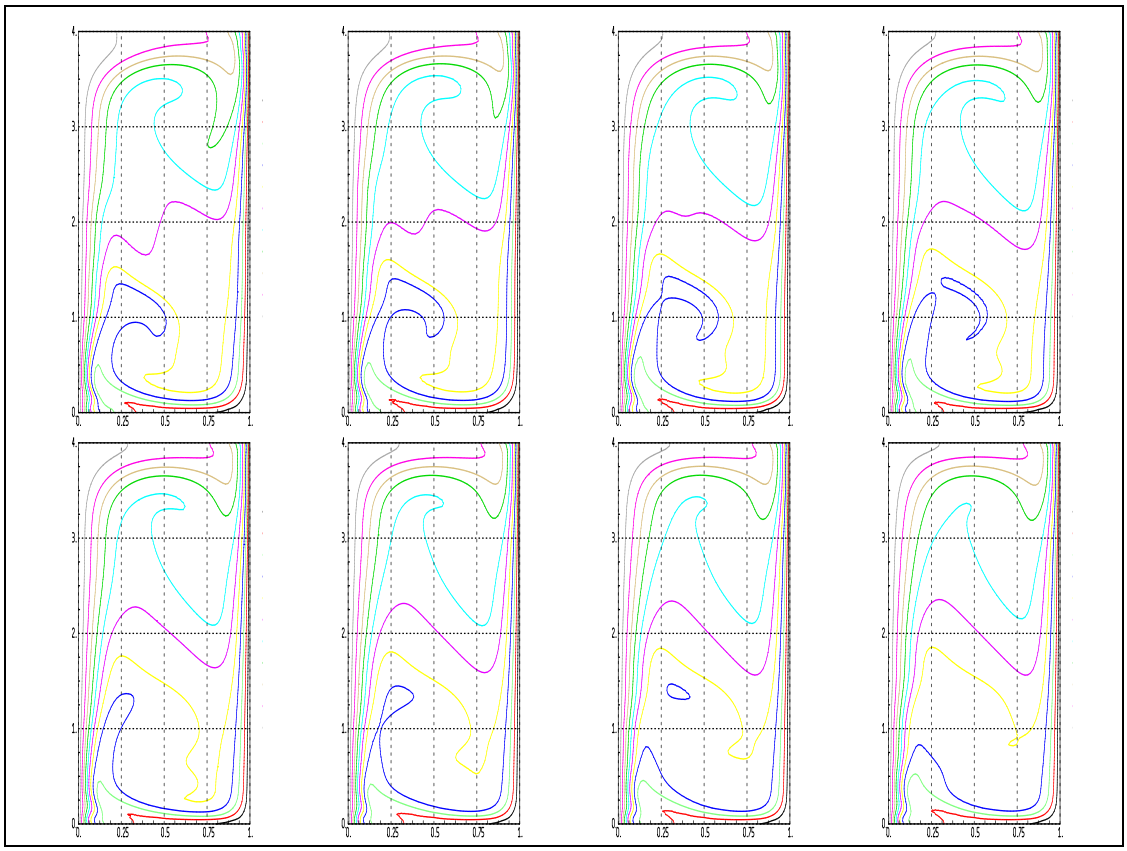


FIG. I.12 – Courbes isothermes de t=13 à t=20 par pas de 1.

convection naturelle avec la conduction de chaleur dans les parois limitant l'écoulement.

Description de la méthode

Nous choisissons, pour présenter la méthode, de traiter un problème modèle, à savoir l'équation de convection-diffusion, associée à des conditions aux limites de Dirichlet homogène. Nous ne précisons dans la suite que les hypothèses du problèmes nécessaires à la compréhension de la méthode multidomaine considérée. Soit donc à résoudre le problème variationnel suivant :

Trouver $u \in H_0^1(\Omega)$ tel que, $\forall v \in H_0^1(\Omega)$ $$\int_\Omega \epsilon \nabla u.\nabla v + \int_\Omega (\vec{b}.\nabla u)v = \int_\Omega f\ v$$

avec f second membre, $\vec{b}$ vitesse d'advection et ϵ coefficient de diffusion donnés.

On se donne deux sous-domaines disjoints Ω_1 et Ω_2 formant une partition du domaine de calcul, de frontière commune notée Γ. Soit V l'espace des fonctions définies sur Ω, dont la restriction à Ω_1 et Ω_2 appartient, respectivement, à $H^1(\Omega_1)$ et $H^1(\Omega_2)$, et de trace nulle sur la frontière du domaine de calcul Ω. Les fonctions de V peuvent être discontinues le long de l'interface Γ.

On démontre alors que le problème elliptique initial est équivalent au problème mixte

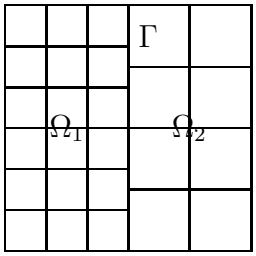
suivant :

Trouver $u \in V$ et $\lambda \in M$ tel que :

$$\left| \begin{array}{ll} \displaystyle\int_\Omega \epsilon \nabla u.\nabla v + \int_\Omega (\vec{b}.\nabla u)\ v + \int_\Gamma [v]\ \lambda = \int_\Omega f\ v & \forall v \in V \\ \displaystyle\int_\Gamma [u]\ \mu = 0 & \forall \mu \in M \end{array} \right.$$

où $[v]$ désigne le saut de la fonction v à travers l'interface et M est un espace de fonctions définies sur Γ assez peu régulières (plus grand que les traces sur Γ des fonctions de V). La continuité de la solution n'est plus imposée *a priori*, par le choix de l'espace dans lequel on la cherche, mais *a posteriori* (seconde équation), par une technique de multiplicateur de Lagrange. On retrouve le problème elliptique initial en restreignant dans la première équation l'espace des fonctions tests aux fonctions de $H_0^1(\Omega)$, pour lesquelles le terme de saut s'annule. L'inconnue λ s'identifie au flux normal à l'interface Γ.

La méthode des éléments finis joints consiste en une discrétisation de ce problème mixte. Du fait que la définition de l'espace V découple les restrictions des fonctions sur chacun des sous-domaines, les maillages sur ces derniers sont indépendants. Le fait que la continuité de la solution n'est demandée qu'au sens faible conduit, au niveau discret, à obtenir une solution généralement discontinue. Toute approximation de M permettant la vérification de la condition de stabilité usuelle du problème mixte est en théorie admissible ; en pratique, on choisit ici l'espace des traces des fonctions d'un des espaces discrets, soit celui associé à Ω_1, soit celui associé à Ω_2.



Le problème mixte précédent est résolu avec l'algorithme d'Uzawa utilisé avec une technique de Lagrangien augmenté.

Un cas d'interaction thermique fluide-structure

Nous appliquons cette méthode sur un problème couplant un problème de conduction thermique dans un solide à un problème de convection naturelle dans la cavité qu'il délimite. Nous souhaitons, par la mise en oeuvre de la technique des éléments finis joints, profiter de ce découpage naturel du domaine de calcul.

La géométrie de ce problème est présentée sur figure I.5. Nous adimensionnons le problème et nous obtenons les problèmes suivants sur les différents domaines.

Le champ de température dans le solide est gouverné par l'équation de la chaleur :

$$\frac{\lambda_r}{a_r}\frac{\partial T}{\partial t} - \nabla \cdot (\lambda_r \nabla T) = 0$$

où λ_r représente le rapport des conductivités thermiques entre la partie fluide et la partie solide et a_r le rapport des diffusivités thermiques.

Dans la cavité nous avons un problème de convection naturelle pour lequel nous supposons valide l'approximation de Boussinesq, si bien que l'écoulement obéit au

système d'équations de bilan suivant :

$$\left|\begin{array}{l} \dfrac{\partial T}{\partial t} + (u\cdot\nabla)T - \nabla\cdot(\nabla T) = 0 \\ \dfrac{\partial u}{\partial t} + (u\cdot\nabla)u - Pr\nabla\cdot(\nabla u + (\nabla)^t u) = PrRa(T-0.5)\vec{y} \\ \nabla\cdot(u) = 0 \end{array}\right.$$

A l'interface fluide-solide, la vitesse dans le fluide s'annule tandis que les conditions de transmission pour l'équation de bilan d'énergie s'écrivent :

$$T_{\text{mur}} = T_{\text{fluide}}\ , \qquad \lambda_r \frac{\partial T_{\text{mur}}}{\partial n} = \frac{\partial T_{\text{fluide}}}{\partial n}$$

La semi-discrétisation en temps utilisée pour résoudre ce problème est la même que précédemment : bilan d'énergie et équations du mouvement sont résolus en deux étapes successives ; pour la seconde, on utilise une méthode de projection. Le couplage entre fluide et solide n'apparaît que dans la première étape. Il est traité en résolvant les problèmes de conduction dans le solide et d'advection-diffusion dans le fluide par des solveurs et des maillages indépendants et en couplant ces deux résolutions par la MEFJ.

Nous pouvons voir sur la figure I.13, les résultats obtenus pour ce cas test dont l'allure générale est proche des résultats obtenus dans [22].

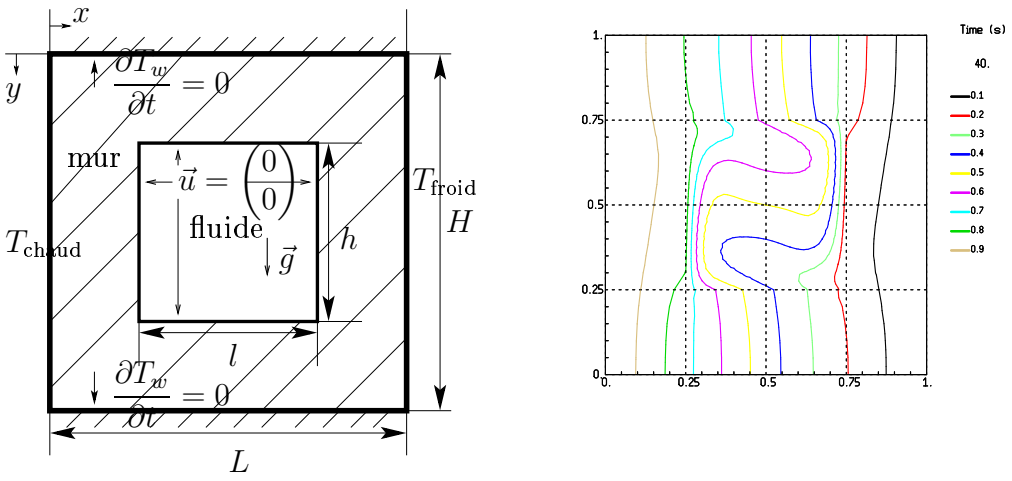


FIG. I.13 – Géométrie du problème et isothermes en régime permanent

I.6 Conclusion

Nous nous intéressons dans cette thèse à une famille particulière de méthodes de projection, les schémas dits de projection-pénalité. Leur principe de base consiste à réintroduire la contrainte de divergence dans l'étape de prédiction de vitesse, via un terme analogue au terme d'augmentation classiquement utilisé dans les méthodes de Lagrangien augmenté. Nous avons développé un schéma de ce type, présentant des similitudes avec des méthodes déjà décrites dans la littérature [29] et [7], mais,

à notre connaissance, original. Ce schéma a été analysé sur le problème modèle de Stokes instationnaire ; nous obtenons des estimations en norme énergie de l'erreur de fractionnement (différence entre la solution et celle obtenue par une méthode couplée) optimales dans le sens suivant : pour les faibles valeurs du paramètre de pénalité r, nous retrouvons l'estimation d'ordre deux par rapport au pas de temps Δt de la méthode de projection classique, tandis que, à fort r, le comportement de l'erreur en $\Delta t/r$ des méthodes de pénalité est observé.

Cette étude théorique est prolongée par une expérimentation numérique, portant sur la résolution des équations de Navier-Stokes avec des conditions aux limites de Dirichlet ou "ouvertes". Cette étude confirme que l'erreur de fractionnement vérifie les estimations obtenues. En outre, nous observons que cette dernière est, pour des discrétisations d'ordre deux en temps, l'erreur dominante à fort pas de temps ; la pénalisation introduite permet alors de gagner jusqu'à deux ordres de grandeur en précision. Enfin, comme pour les variantes dites "rotationnelles", la pression obtenue n'a plus à satisfaire les conditions aux limites artificielles caractéristiques des méthodes de projection standard ; les couches limites parasites de pression pour les problèmes de Dirichlet disparaissent alors, et l'on retrouve une convergence optimale en espace pour les conditions aux limites ouvertes. En revanche, l'ajout d'un terme de pénalisation dans l'étape de prédiction dégrade le conditionnement du système algébrique, ce qui suggère l'emploi de méthodes multi-grilles ; c'est l'un des prolongements possibles de la présente étude.

Les méthodes de projection développées sont ensuite étendues aux cas des écoulements dilatables, pour lesquels on constate une convergence similaire. Nous démontrons ainsi le bon comportement de ce type de méthode en traitant des problèmes de convection naturelle dans le cadre du modèle asymptotique pour les écoulements à faible nombre de Mach. En particulier, nous étudions le régime transitoire depuis l'état au repos pour une cavité verticale différentiellement chauffée de rapport de forme égal à 4, à grand nombre de Rayleigh ; la convergence au pas de temps et au maillage semble avoir été obtenue dans ces calculs, au prix de la mise en œuvre d'éléments finis quadratiques s'appuyant sur des maillages importants (jusqu'à 10^6 degrés de liberté au total). Cette étude dans le cas dilatable pourrait être prolongée sur le plan théorique par un examen plus approfondi des propriétés de stabilité et de convergence. Sur un autre plan, les applications industrielles de ce type de méthodes à pas fractionnaires pour les écoulements dilatables sont multiples ; les schémas développés dans cette thèse ont ainsi été adaptés à l'IRSN à la problématique de la simulation des incendies [1], ce qui, compte-tenu de la complexité des modèles utilisés, ouvre un vaste champ de développements algorithmiques.

Enfin nous nous sommes intéressés à une méthode de décomposition de domaine appelée méthode d'éléments finis joints, qui a pour principe de n'imposer qu'au sens faible la continuité des variables aux interfaces des sous-domaines. Nous tirons parti de cette caractéristique pour traiter un problème "multi-physiques", à savoir un écoulement en convection naturelle dans une cavité aux parois d'épaisseur finie et conductrices de chaleur, avec des maillages non concordants pour la cavité et ses parois. Les résultats obtenus sont comparés avec succès aux données disponibles pour ce problème dans la littérature. Les travaux menés ici sont toutefois très préliminaires sur le plan de l'algorithmique de résolution du problème multi-domaines, et un effort important sur ce plan serait à mener pour permettre la mise en œuvre des idées présentées dans ce manuscript à des cas industriels.

Bibliographie

[1] F. Babik, T. Gallouët, J.-C. Latché, S. Suard, and D. Vola. On some fractional step schemes for combustion problems. In *Finite Volumes for Complex Applications IV (FVCA IV)*. Éditions Hermès, Paris, 2005.

[2] F. Ben Belgacem and Y. Maday. The mortar finite element method for three dimensional finite elements. *Mathematical Modelling and Numerical Analysis*, 31(2) :289–302, 1997.

[3] Faker Ben Belgacem. The mortar finite element method with Lagrange multipliers. *Numerische Mathematik*, 84 :173–197, 1999.

[4] Faker Ben Belgacem. The mixed mortar finite element method for the incompressible Stokes problem : Convergence analysis. *SIAM Journal on Numerical Analysis*, 37(4) :1085–1100, 2000.

[5] C. Bernardi, Y. Maday, and A.T. Patera. Domain decomposition by the mortar element method. In H. G. Kaper and M. Garbey, editors, *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters*, pages 269–286. Kluwer Academic Publishers, NATO ASI Series, 1993.

[6] David Braess, Wolfgang Dahmen, and Christian Wieners. A multigrid algorithm for the mortar finite element method. *SIAM Journal on Numerical Analysis*, 37(1) :48–69, 1999.

[7] Jean-Paul Caltagirone and Jérôme Breil. Sur une méthode de projection vectorielle pour la résolution des équations de Navier-Stokes. *Comptes-Rendus de l'académie des Sciences, Paris – Série II*, 327 :1179–1184, 1999.

[8] Alexandre Joel Chorin. Numerical solution of the Navier-Stokes equations. *Mathematics of Computation*, 22 :745–762, 1968.

[9] Alexandre Ern and Jean-Luc Guermond. *Eléments finis : théorie, applications, mise en œuvre*, volume 36 of *Mathématiques & Applications*. Springer, 2002.

[10] M. Fortin and R. Glowinski. *Méthodes de Lagrangien Augmenté*. Dunod, Paris, 1982.

[11] Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms.*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, 1986.

[12] Katuhiko Goda. A multistep technique with implicit difference schemes for calculating two- or three-dimensional cavity flows. *Journal of Computational Physics*, 30 :76–95, 1979.

[13] Philip M. Gresho. On the theory of semi-implicit projection methods for viscous incompressible flow and its implementation via a finite element method that also introduces a nearly consistent mass matrix. part 1 : Theory. *International Journal for Numerical Methods in Fluids*, 11 :587–620, 1990.

[14] Philip M. Gresho. On the theory of semi-implicit projection methods for viscous incompressible flow and its implementation via a finite element method that also introduces a nearly consistent mass matrix. part 2 : Implementation. *International Journal for Numerical Methods in Fluids*, 11 :621–659, 1990.

[15] J.-L. Guermond and L. Quartapelle. On the approximation of the unsteady Navier-Stokes equations by finite element projection methods. *Numerische Mathematik*, 80 :207–238, 1998.

[16] Jean-Luc Guermond. Some implementations of projection methods for Navier-Stokes equations. *Mathematical Modelling and Numerical Analysis*, 30(5) :637–667, 1996.

[17] Jean-Luc Guermond. Un résultat de convergence d'ordre deux en temps pour l'approximation des équations de Navier-Stokes par une technique de projection incrémentale. *Mathematical Modelling and Numerical Analysis*, 33(1) :169–189, 1999.

[18] J.L. Guermond, P. Minev, and J. Shen. An overview of projection methods for incompressible flows, 2005.

[19] J.L. Guermond, P. Minev, and Jie Shen. Error analysis of pressure-correction schemes for the Navier-Stokes equations with open boundary conditions. *submitted to SIAM Journal on Numerical Analysis*, 2004.

[20] J.L. Guermond, P. Minev, and Jie Shen. An overview of projection methods for incompressible flows, 2004. submitted to International Journal on Numerical Methods in Engineering.

[21] J.L. Guermond and Jie Shen. On the error estimates for the rotational pressure-correction projection methods. *Mathematics of Computation*, 73(248) :1719–1737, 2003.

[22] D. M. Kim and R. Viskanta. Effect of wall heat conduction on natural convection heat transfer in a square enclosure. *Journal of Heat Transfer*, 107 :139–146, feb 1985.

[23] J. Kim and P. Moin. Application of a fractional-step method to incompressible Navier-Stokes equations. *Journal of Computational Physics*, 59 :308–323, 1985.

[24] P. Le Quéré, C. Weisman, H. Paillère, J. Vierendeels, E. Dick, R. Becker, M. Braack, and J. Locke. Modelling of natural convection flows with large temperature differences : A benchmark problem for low mach number solvers. part 1. reference solutions. *Mathematical Modelling and Numerical Analysis*, 39(3) :609–616, 2005.

[25] A. Majda and J. Sethian. The derivation and numerical solution of the equations for zero Mach number solution. *Combustion Science and Techniques*, 42 :185–205, 1985.

[26] B. Müller. Low mach number asymptotics of the Navier-Stokes equations and numerical implications. In *30th Computational Fluid Dynamics, March 8-12 1999*, number 1999-03 in Lecture Series. von Karman Institute for Fluid Dynamics, 1999.

[27] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, 1994.

[28] Jie Shen. On error estimates of projection methods for Navier-Stokes equations : First-order schemes. *SIAM Journal on Numerical Analysis*, 29(1) :57–77, 1992.

[29] Jie Shen. On error estimates of some higher order projection and penalty-projection methods for Navier-Stokes equations. *Numerische Mathematik*, 62 :49–73, 1992.

[30] Jie Shen. Remarks on the pressure estimates for the projection methods. *Numerische Mathematik*, 67 :513–520, 1994.

[31] Jie Shen. On error estimates of the penalty method for unsteady Navier-Stokes equations. *SIAM Journal on Numerical Analysis*, 32(2) :386–403, 1995.

[32] Jie Shen. On error estimates of projection methods for Navier-Stokes equations : Second-order schemes. *Mathematics of Computation*, 65(215) :1039–1065, 1996.

[33] G.I. Taylor and B.A. Green. Mechanism of the production of small eddies from large ones. *Proceedings of the Royal Society of London A*, 158 :499–521, 1935.

[34] R. Temam. Sur l'approximation de la solution des Équations de Navier-Stokes par la méthode des pas fractionnaires (II). *Archive for Rational Mechanics and Analysis*, 33 :377–385, 1969.

[35] L.J.P. Timmermans, P.D. Minev, and F.N. Van de Vosse. An approximate projection scheme for incompressible flow using spectral elements. *International Journal for Numerical Methods in Fluids*, 22 :673–688, 1996.

[36] Stefan Turek. *Efficient Solvers for Incompressible Flow Problems : An Algorithmic Approach in View of Computational Aspects*. Springer, 1999.

[37] J. Van Kan. A second-order accurate pressure-correction scheme for viscous incompressible flow. *SIAM Journal on Scientific and Statistical Computing*, 7(3) :870–891, 1986.

[38] Barbara I. Wohlmuth. Hierarchical a posteriori error estimators for mortar finite element methods with Lagrange multipliers. *SIAM Journal on Numerical Analysis*, 36(5) :1636–1658, 1999.

[39] Barbara I. Wohlmuth. A residual based error estimator for mortar finite element discretizations. *Numerische Mathematik*, 84 :143–171, 1999.

Chapitre II

Analyse de deux variantes de la méthode de Projection sur le problème de Stokes

II.1 Introduction and presentation of the numerical schemes

Since the pioneering work of Chorin [6] and Temam [27] in the late sixties, projection methods have received a lot of attention and fractional step schemes falling in this category are probably nowadays the most popular ones for the solution of the unstationary two and three-dimensional Navier-Stokes equations. Indeed, schemes of this type have proved to be extremely efficient, essentially because, at each time step, they reduce without loss of stability the solution of a saddle-point type problem to a sequence of "uncoupled" elliptic equations for the velocity and pressure. This feature makes them particularly attractive for industrial applications, as for instance nuclear safety studies which are the context of this work.

The aim of the present paper is to analyse some variants of the projection method of which the main advantage is, concisely speaking, to offer the possibility to reduce the splitting error, *i.e.* the difference between the solution of the fractional step scheme and the solution of the coupled one, up to make it negligible. The basic idea behind the development of these schemes originates from a paper of Shen in 1992 [22] and consists in adding to the velocity prediction step a term similar to the augmentation term used in the so-called Augmented Lagrangian method (e.g. [10]), which constrains the tentative velocity to remain almost divergence free. The same idea has been exploited independently later, in 1999, by Caltagirone and Breil [4].

To be more specific, let us consider as model problem the unstationary Stokes problem with homogeneous Dirichlet boundary conditions :

$$\left|\begin{array}{ll} \dfrac{\partial u}{\partial t} - \Delta u + \nabla p = f & \text{in } \Omega\times]0,T[\\ \nabla \cdot u = 0 & \text{in } \Omega\times]0,T[\\ u = 0 & \text{on } \partial\Omega\times]0,T[\\ u(x,0) = u_0(x) & \text{in } \Omega \end{array}\right. \tag{II.1.1}$$

where u stands for the velocity vector, p the pressure, f a forcing term, Ω is a smooth domain of $\mathbb{R}^d$, $d = 2$ or $d = 3$, of boundary $\partial\Omega$, u_0 is the (divergence free) initial velocity field and the problem is posed over a finite time interval $]0, T[$.

With standard notations for Sobolev spaces and their norms (see for instance [1]), we define the functional space W by :

$$W = \{v \text{ such that } v \in \mathrm{L}^2(0,T;\mathrm{H}^1_0(\Omega)^d),\ \frac{\partial v}{\partial t} \in \mathrm{L}^2(0,T;\mathrm{H}^{-1}(\Omega)^d)\}$$

Throughout this paper, we will use $(\cdot,\cdot)$ to note the standard scalar product in $\mathrm{L}^2(\Omega)$ or $\mathrm{L}^2(\Omega)^d$, as well as in $\mathbb{R}^n$. With these notations, a variational form of the Stokes problem reads [8, pp. 265–277] :

Find $u \in W$ and $p \in \mathrm{L}^2(0,T;\mathrm{L}^2_0(\Omega))$ such that :

$$\left|\begin{array}{ll} < \frac{\partial u}{\partial t}, v > + (\nabla u, \nabla v) - (p, \nabla \cdot v) = (f, v) & \forall v \in \mathrm{H}^1_0(\Omega)^d, \quad \text{a.e. } t \in]0,T[\\ (-\nabla \cdot u, q) = 0 & \forall q \in \mathrm{L}^2_0(\Omega), \quad \text{a.e. } t \in]0,T[\\ u|_{t=0} = u_0 & \end{array}\right.$$

where $\mathrm{L}^2_0(\Omega)$ is the space of $\mathrm{L}^2(\Omega)$ functions which mean value is zero, $< \cdot, \cdot >$ stands for the duality product between $\mathrm{H}^{-1}(\Omega)^d$ and $\mathrm{H}^1_0(\Omega)^d$, the forcing term f is supposed to lie in $\mathrm{C}^0([0,T];\mathrm{L}^2(\Omega)^d)$ and u_0 in $\mathrm{H}^1_0(\Omega)^d$. Note that this poblem, which well-posedness is proven in [8, proposition 7.2.3, p. 267], requires more regularity on the data than usually and, in turn, yields a more regular pressure.

Let the problem be discretized in time by a first-order backward Euler method with a constant time step and let the discretization in space be given, by a finite volume or finite element method. We then obtain at each time step an algebraic system of the form :

$$\left|\begin{array}{l} \frac{1}{\delta t}(\mathbf{M}_v \mathbf{U}^{k+1} - \mathbf{M}_v \mathbf{U}^k) + \mathbf{A}\mathbf{U}^{k+1} + \mathbf{B}^t \mathbf{P}^{k+1} = \mathrm{F}^{k+1} \\ \mathbf{B}\mathbf{U}^{k+1} = 0 \end{array}\right. \tag{II.1.2}$$

where $\mathbf{M}_v$ stands for the (regular) velocity mass matrix, $\mathbf{A}$ and $\mathbf{B}^t$ for the algebraic operators associated respectively to the Laplace (*i.e.* $-\Delta$) and gradient operators, $\mathbf{U}^k$ and $\mathbf{P}^k$ are the velocity and pressure degrees of freedom vectors at time $t_k = k\,\delta t$ and F^k stands for the vector associated to the forcing term at time t_k. System (II.1.2) takes the form of a generic saddle-point problem, which has to be solved by specific algorithms.

The method proposed by Caltagirone and Breil [4] is a fractional step algorithm of the form :

$$\begin{array}{ll} \text{(step 1)} & \frac{1}{\delta t}(\mathbf{M}_v \tilde{\mathbf{U}}^{k+1} - \mathbf{M}_v \mathbf{U}^k) + (\mathbf{A} + r\mathbf{B}^t \mathbf{M}_p^{-1} \mathbf{B})\tilde{\mathbf{U}}^{k+1} + \mathbf{B}^t \mathbf{P}^k = \mathrm{F}^{k+1} \\ \text{(step 2)} & \text{compute } \mathbf{U}^{k+1} \text{ from } \tilde{\mathbf{U}}^{k+1} \text{ such that} \quad \mathbf{B}\mathbf{U}^{k+1} = 0 \\ \text{(step 3)} & \mathbf{M}_p \mathbf{P}^{k+1} = \mathbf{M}_p \mathbf{P}^k + r\mathbf{B}\tilde{\mathbf{U}}^{k+1} \end{array} \tag{II.1.3}$$

where r is a positive parameter (the so-called penalty parameter), the matrix $\mathbf{M}_p$ is supposed to be symmetric, and positive definite. In the initial paper, $\mathbf{M}_p$ was set to the identity and the second step was performed by solving an original but singular algebraic system called the vector projection. The spatial discretization was performed by the finite volume MAC scheme (e.g. [9]). This numerical scheme is used in the AQUILON code (http ://www.trefle.u-bordeaux1.fr/aquilon), for a wide range of industrial applications.

Replacing this projection step by a more conventional one, we obtain the following scheme :

$$\begin{aligned}
&\text{(step 1)} \qquad \frac{1}{\delta t}(\mathbf{M}_v\tilde{\mathbf{U}}^{k+1} - \mathbf{M}_v\mathbf{U}^k) + (\mathbf{A} + r\mathbf{B}^t\mathbf{M}_p^{-1}\mathbf{B})\tilde{\mathbf{U}}^{k+1} + \mathbf{B}^t\mathbf{P}^k = \mathrm{F}^{\mathrm{k+1}} \\
&\text{(step 2)} \qquad \mathbf{L}\Phi = \frac{1}{\delta t}\mathbf{B}\tilde{\mathbf{U}}^{k+1}\ , \qquad \mathbf{M}_v\mathbf{U}^{k+1} = \mathbf{M}_v\tilde{\mathbf{U}}^{k+1} - \delta t\mathbf{B}^t\Phi \\
&\text{(step 3)} \qquad \mathbf{M}_p\mathbf{P}^{k+1} = \mathbf{M}_p\mathbf{P}^k + r\mathbf{B}\tilde{\mathbf{U}}^{k+1}
\end{aligned} \tag{II.1.4}$$

where $\mathbf{L}$ is the algebraic operator associated to the pressure laplacian and, in order to make the properties of the algorithm as less as possible mesh-dependent, we use the standard scaling process which consists in taking for $\mathbf{M}_p$ an approximate pressure mass matrix, which, for computational efficiency reasons, can be chosen diagonal. This scheme is the first one that we will study in the present paper.

By analogy with standard incremental projection methods, the "projector" Φ appears to be a rather natural candidate to contribute to the pressure increment, which suggests the following second algorithm, differing from the preceding one only in the last step :

$$\begin{aligned}
&\text{(step 1)} \qquad \frac{1}{\delta t}(\mathbf{M}_v\tilde{\mathbf{U}}^{k+1} - \mathbf{M}_v\mathbf{U}^k) + (\mathbf{A} + r\mathbf{B}^t\mathbf{M}_p^{-1}\mathbf{B})\tilde{\mathbf{U}}^{k+1} + \mathbf{B}^t\tilde{\mathrm{P}}^k = \mathrm{F}^{\mathrm{k+1}} \\
&\text{(step 2)} \qquad \mathbf{L}\Phi = \frac{1}{\delta t}\mathbf{B}\tilde{\mathbf{U}}^{k+1}\ , \qquad \mathbf{M}_v\mathbf{U}^{k+1} = \mathbf{M}_v\tilde{\mathbf{U}}^{k+1} - \delta t\mathbf{B}^t\Phi \\
&\text{(step 3)} \qquad \mathbf{M}_p\tilde{\mathrm{P}}^{k+1} = \mathbf{M}_p(\tilde{\mathrm{P}}^k + \Phi)
\end{aligned}$$

This modification yields a discrete version of the so-called "penalty-projection scheme", introduced by Shen in [22, section 6]. In this algorithm, the quantity $\tilde{\mathrm{P}}^{k+1}$ is only an auxiliary quantity, from which the degrees of freedom of the actual approximation of the pressure can be deduced by the equation :

$$\mathbf{P}^{k+1} = \tilde{\mathrm{P}}^{k+1} + r\mathbf{B}\tilde{\mathbf{U}}^{k+1}$$

As usually in the penalty methods framework, the penalty parameter r varies in Shen's penalty-projection method with the time step; optimal error bounds are proven for $r = 1/\delta t^2$ in [22].

In fact, in view of the last equation, it seems rather natural to use the approximation

of the pressure in the velocity prediction step, to obtain the following algorithm :

$$\begin{aligned}
&\text{(step 1)} \qquad \frac{1}{\delta t}(\mathbf{M}_v\tilde{\mathbf{U}}^{k+1} - \mathbf{M}_v\mathbf{U}^k) + (\mathbf{A} + r\mathbf{B}^t\mathbf{M}_p^{-1}\mathbf{B})\tilde{\mathbf{U}}^{k+1} + \mathbf{B}^t\mathbf{P}^k = \mathrm{F}^{k+1} \\
&\text{(step 2)} \qquad \mathbf{L}\Phi = \frac{1}{\delta t}\mathbf{B}\tilde{\mathbf{U}}^{k+1} \ , \qquad \mathbf{M}_v\mathbf{U}^{k+1} = \mathbf{M}_v\tilde{\mathbf{U}}^{k+1} - \delta t\mathbf{B}^t\Phi \\
&\text{(step 3)} \qquad \mathbf{M}_p\mathbf{P}^{k+1} = \mathbf{M}_p(\mathbf{P}^k + \Phi) + r\mathbf{B}\tilde{\mathbf{U}}^{k+1}
\end{aligned} \tag{II.1.5}$$

This is the second scheme addressed in the present work. Taking $r = 0$ in this method yields the well-known incremental projection scheme, we thus can hope that this method will be convergent whatever the penalty parameter may be. In addition, the operator $\mathbf{A} + r\mathbf{B}^t\mathbf{M}_p^{-1}\mathbf{B}$ is known to be poorly conditionned at high values of r, and, consequently, keeping this parameter within reasonable bounds may be an efficient strategy from a computational point of view. Thus we choose here to completely disconnect the parameter r from the time step δt.

Besides the two above quoted references, the schemes (II.1.4) and (II.1.5) present also some analogy with other already presented numerical algorithms, from both the literature of penalty and pressure correction methods.

First, the algorithm (II.1.4) can be obtained by adding a projection step to a penalty (or quasi-incompressibility) method ; as far as this latter class of methods is concerned, we refer to [26] for the seminal work and to [24] for an analysis.

As far as the algorithm (II.1.5) is concerned, setting $r = 0$ in the velocity prediction step and $r = 1$ for the present model problem (r equal to the viscosity for actual Navier-Stokes equations), yields a scheme proposed by Timmermans *et al.* [28] as an alternative to the usual incremental projection method (concerning this latter scheme, see Goda [12] for the original setting, Shen [21, 23, 25] for an analysis in the time semi-discrete case and Guermond and Quartapelle [14, 13, 15] for an analysis of the fully discrete case). The properties of the Timmermans *et al.* scheme were further investigated by Brown *et al.* through a normal mode analysis for a particular problem in [3]. Finally, energy norm estimates for the time-discrete case were obtained by Guermond and Shen [16], which gave to the scheme the name of "rotational pressure-correction projection method". Note also that an equation to update the pressure similar to the third step in (II.1.4), with the penalty parameter r once again replaced by the viscosity, was used by Prohl [19, chapter 8] in an algorithm which received the name of "Chorin-Uzawa scheme", since this equation is reminiscent of the pressure update step in the so-called Uzawa method (e.g. [10]).

In view of the above mentioned literature, we will refer to the schemes (II.1.5) and (II.1.4) as respectively *the standard penalty-projection method* and *the Uzawa variant.*

Our goal here is to perform an analysis in energy norm of these two algorithms. To this purpose, we will estimate the so-called splitting error, *i.e.* the difference between the results (in velocity and pressure) obtained by the projection methods under consideration and by the coupled scheme (II.1.2). Indeed, this quantity has been shown by Guermond [15] to be rather insensitive to the order of the time discretization of the unstationary term in the momentum balance equation, so we

can hope that the present analysis for a first-order time discretization will also apply to second-order schemes.

This paper is organized as follows. We begin by setting the considered numerical schemes within a variational framework suitable for an error analysis in energy norms (section II.2). Then, after some preliminaries (section II.3), the Uzawa variant and the standard penalty-projection method are addressed in respectively section II.4 and section II.5. Finally, some numerical tests are presented in section II.6.

II.2 A variational framework

The aim of this section is to provide a variational framework for the three schemes under consideration, namely the implicit Euler method (II.1.2), the Uzawa variant of the penalty-projection (II.1.4) scheme and the standard penalty-projection method (II.1.5).

In the first case, the variational setting is standard and reads, for each time step :

$$\text{Find } (\bar{u}^{k+1}, \bar{p}^{k+1}) \in V_h \times M_h \quad \text{such that}$$

$$\left| \begin{array}{lll} (i) & \dfrac{1}{\delta t}\,(\bar{u}^{k+1} - \bar{u}^{k}, v) + (\nabla \bar{u}^{k+1}, \nabla v) - (\nabla \cdot v, \bar{p}^{k+1}) = (f^{k+1}, v) & \forall v \in V_h \\ (ii) & -(\nabla \cdot \bar{u}^{k+1}, q) = 0 & \forall q \in M_h \end{array} \right. \tag{II.2.6}$$

where V_h and M_h are internal approximations of $\mathrm{H}^1_0(\Omega)^d$ and $\mathrm{L}^2_0(\Omega)$ respectively.

To associate a discrete variational setting to the algebraic formulation of the Uzawa variant of the penalty-projection scheme (II.1.4), we face three difficulties, namely to introduce the pressure Poisson problem, to deal with the pressure mass matrix lumping and, finally, to derive a variational analogue of the penalty term added in the velocity prediction step. The first difficulty has been solved by Guermond [14], and its solution consists in searching for the end-of-step velocity in a non $\mathrm{H}^1_0(\Omega)^d$-conforming space X_h which is spanned by the functions of V_h and the gradient of the functions of M_h (which is usually expressed by the notation $X_h = V_h + \nabla M_h$). The second step of (II.1.4) then reads, with obvious notations for the discrete functions :

$$\text{Find } (u^{k+1}, \varphi) \in X_h \times M_h \quad \text{such that}$$

$$\left| \begin{array}{ll} \dfrac{1}{\delta t}\,(u^{k+1} - \tilde{u}^{k+1}, v) + (\nabla \varphi, v) = 0 & \forall v \in X_h \\ (u^{k+1}, \nabla q) = 0 & \forall q \in M_h \end{array} \right.$$

The divergence of the function u^{k+1} does not lie in $\mathrm{L}^2(\Omega)$, and we can no more write the divergence constraint under its standard form $(\nabla \cdot u^{k+1}, q) = 0,\ \forall q \in M_h$; in counterpart, the space M_h is required to be included in $\mathrm{H}^1(\Omega)$, which gives sense to the substitute $(u^{k+1}, \nabla q) = 0,\ \forall q \in M_h$. The projection step then decomposes in two decoupled sub-problems : choosing $v = \nabla q,\ q \in M_h$ in the first equation and using the second one to eliminate the term $(u^{k+1}, \nabla q)$ yields the usual Poisson problem

for the pressure update (first equation of step 2 in (II.1.4) and (II.1.5)) ; then taking $v \in V_h$ in the first equation gives the variational equation which allows to compute the restriction to V_h (defined as the L^2-projection onto V_h) of the end-of-step velocity (second equation of step 2 in (II.1.4) and (II.1.5)).

Then we associate to the discrete operator $\mathbf{M}_p$ an approximate $L^2(\Omega)$ scalar product, noted $(\cdot,\cdot)_h$, *i.e.* , with obvious notations :

$$(p,q)_h = (\mathbf{M}_p \mathrm{P}, \mathrm{Q}) \qquad \forall p,q \in M_h$$

This allows to write the pressure update as :

$$(p^{k+1},q)_h = (p^k,q)_h + r(\tilde{u}^{k+1},\nabla q)$$

Let the operator B_h, acting from V_h to M_h, be defined by :

$$u \mapsto B_h u \quad \text{such that} \quad (B_h u, q)_h = (u, \nabla q) \quad \forall q \in M_h$$

We can see that, for any function u in V_h, the vector of degrees of freedom associated to $B_h u$, noted $\mathbf{B}_h \mathrm{U}$, reads :

$$\mathbf{B}_h \mathrm{U} = \mathbf{M}_p^{-1} \mathbf{B} \mathrm{U}$$

Consequently, observing that the penalty term in the first step of (II.1.4) satisfies the following property :

$$(\mathbf{B}^t \mathbf{M}_p^{-1} \mathbf{B} \mathrm{U}, \mathrm{V}) = (\mathbf{M}_p^{-1} \mathbf{B} \mathrm{U}, \mathbf{B} \mathrm{V}) = (\mathbf{M}_p^{-1} \mathbf{B} \mathrm{U}, \mathbf{M}_p \left[\mathbf{M}_p^{-1} \mathbf{B} \mathrm{V}\right]) = (\mathbf{M}_p \mathbf{B}_h \mathrm{U}, \mathbf{B}_h \mathrm{V})$$

we obtain that this term stems from the following variational counterpart :

$$c_h(u,v) = (B_h u, B_h v)_h \qquad \forall u,v \in V_h \tag{II.2.7}$$

As the matrix $\mathbf{M}_p$ is supposed to be symmetric and positive definite, the bilinear form $c_h(\cdot,\cdot)$ is symmetric and positive.

We then obtain the following variational algorithm for the Uzawa variant of the penalty-projection scheme :

Find $(\tilde{u}^{k+1}, u^{k+1}, \varphi, p^{k+1}) \in V_h \times X_h \times M_h \times M_h$ such that

$$\left|\begin{array}{lll}
(i) & \dfrac{1}{\delta t}\,(\tilde{u}^{k+1} - u^k, v) + (\nabla \tilde{u}^{k+1}, \nabla v) + r\; c_h(\tilde{u}^{k+1}, v) + (\nabla p^k, v) = (f^{k+1}, v) & \forall v \in V_h \\
(ii) & \dfrac{1}{\delta t}\,(u^{k+1} - \tilde{u}^{k+1}, v) + (\nabla \varphi, v) = 0 & \forall v \in X_h \\
(iii) & (u^{k+1}, \nabla q) = 0 & \forall q \in M_h \\
(iv) & (p^{k+1}, q)_h = (p^k, q)_h + r(\tilde{u}^{k+1}, \nabla q) & \forall q \in M_h
\end{array}\right. \tag{II.2.8}$$

Using the definition of $c_h(\cdot,\cdot)$, the definition of B_h and equation (II.2.8-(iv)), we see that, $\forall v \in V_h$:

$$r\, c_h(\tilde{u}^{k+1}, v) = r\, (B_h \tilde{u}^{k+1}, B_h v)_h = r\, (\tilde{u}^{k+1}, \nabla B_h v) = (p^{k+1} - p^k, B_h v)_h = (\nabla (p^{k+1} - p^k), v)$$

and the equation (II.2.8-(i)) equivalently reads :

$$\frac{1}{\delta t}\,(\tilde{u}^{k+1}-u^k,v)+(\nabla\tilde{u}^{k+1},\nabla v)+(\nabla p^{k+1},v)=(f^{k+1},v)\qquad\forall v\in V_h$$

Using similar arguments, the variational formulation associated to the standard penalty-projection method reads :

Find $(\tilde{u}^{k+1},u^{k+1},p^{k+1})\in V_h\times X_h\times M_h$ such that

$$\left|\begin{array}{lll}(i) & \dfrac{1}{\delta t}\,(\tilde{u}^{k+1}-u^k,v)+(\nabla\tilde{u}^{k+1},\nabla v)+r\;c_h(\tilde{u}^{k+1},v)+(\nabla p^k,v)=(f^{k+1},v) & \forall v\in V_h\\ (ii) & \dfrac{1}{\delta t}\,(u^{k+1}-\tilde{u}^{k+1},v)+(\nabla(p^{k+1}-p^k-rB_h\tilde{u}^{k+1}),v)=0 & \forall v\in X_h\\ (iii) & (u^{k+1},\nabla q)=0 & \forall q\in M_h\end{array}\right. \tag{II.2.9}$$

or, equivalently :

Find $(\tilde{u}^{k+1},\tilde{p}^{k+1},u^{k+1},p^{k+1})\in V_h\times M_h\times X_h\times M_h$ s.t.

$$\left|\begin{array}{lll}(i) & \dfrac{1}{\delta t}\,(\tilde{u}^{k+1}-u^k,v)+(\nabla\tilde{u}^{k+1},\nabla v)+(\nabla\tilde{p}^{k+1},v)=(f^{k+1},v) & \forall v\in V_h\\ (ii) & (\tilde{p}^{k+1},q)_h=(p^k,q)_h+r(\tilde{u}^{k+1},\nabla q) & \forall q\in M_h\\ (iii) & \dfrac{1}{\delta t}\,(u^{k+1}-\tilde{u}^{k+1},v)+(\nabla(p^{k+1}-\tilde{p}^{k+1}),v)=0 & \forall v\in X_h\\ (iv) & (u^{k+1},\nabla q)=0 & \forall q\in M_h\end{array}\right. \tag{II.2.10}$$

II.3 Preliminaries

We begin by collecting the assumptions relative to the discretization spaces. We suppose that V_h and M_h are conforming approximations in respectively $\mathrm{H}_0^1(\Omega)^d$ and $\mathrm{H}^1(\Omega)$, satisfying the so-called Babuska-Brezzi or inf-sup condition (e.g. [11]), and that the following approximation property holds for the space M_h :

$$\forall\bar{\varphi}\in\mathrm{H}^1(\Omega),\quad\inf_{\varphi\in M_h}\|\bar{\varphi}-\varphi\|_0\;\le ch\,|\bar{\varphi}|_1$$

where, here and throughout the remaining of the paper, unless explicitely stated, c stands for a positive real number independent of time or space variables or mesh steps.

We assume in addition that the following inverse inequality holds for any function φ in M_h :

$$\|\nabla\varphi\|_0\;\le\frac{c}{h}\,\|\varphi\|_0$$

Both preceding inequalities are valid, for instance, for the usual Lagrange piecewise linear elements [7] and families of quasi-uniform meshes.

As far as the continous problem is concerned, we suppose that the regularity of the computational domain is such that the Stokes problem is regularizing, in the sense that, if the right hand side lies in $\mathrm{L}^2(\Omega)^d$, the solution lies respectively in $\mathrm{H}^2(\Omega)^d$ for the velocity and in $\mathrm{H}^1(\Omega)$ for the pressure [5, 2].

To simplify the presentation, we will assume in addition that the forcing term f is regular, and that the initial condition u_0 satisfies the compatibility conditions which ensure a full regularity of the problem since the initial time (for an in-depth discussion of the general case, see [19] and references herein, in particular [17]). Consequently, the Euler implicit scheme is first-order accurate in time, and the increments of pressure are such that, $\forall k \geq 0$:

$$\begin{aligned} &\|\delta\bar{p}^{k+1}\|_0 \equiv \|\bar{p}^{k+1} - \bar{p}^k\|_0 \leq c\ \delta t \\ &\|\nabla\delta\bar{p}^{k+1}\|_0 \leq c\ \delta t \\ &\|\delta\delta\bar{p}^{k+1}\|_0 \equiv \|\delta\bar{p}^{k+1} - \delta\bar{p}^k\|_0 \leq c\ \delta t^2 \end{aligned} \tag{II.3.11}$$

where c neither depends on the time nor on the time step.

In the course of this paper, we will make use of the inverse of the discrete Stokes operator, noted S_h, defined by :

$$\begin{aligned} S_h : \quad & V_h \to V_h \\ & u \mapsto S_h u \quad \text{such that :} \end{aligned} \tag{II.3.12}$$

$$\left| \begin{array}{ll} (\nabla S_h u, \nabla v) + (\nabla\varphi, v) = (u, v) & \forall v \in V_h \\ (S_h u, \nabla q) = 0 & \forall q \in M_h \end{array} \right.$$

The following properties are proven in [15, section 4.1] :

Lemma II.3.1. *Prop. 1 : The bilinear form defined over $V_h \times V_h$ by $(u, v) \mapsto (S_h u, v)$ is symmetric positive and defines a semi-norm in V_h which will be noted by $(S_h u, u) = \|u\|_s^2$.*

Prop. 2 : Let us define the space H_h by :

$$H_h = \{v \in X_h \text{ such that } (v, \nabla q) = 0,\ \forall q \in M_h\}$$

Then, for any strictly positive real number α and any function u in V_h, the following inequality holds :

$$(\nabla S_h u, \nabla u) \geq (1 - \alpha)\, \|u\|_0^2 - c(\alpha) \inf_{w \in H_h} \|u - w\|_0^2$$

Proof. Let u be an element of V_h and $\bar{s} \in \mathrm{H}_0^1(\Omega)^d$ and $\bar{\varphi} \in \mathrm{L}_0^2(\Omega)$ be the solution of the following continuous Stokes problem :

$$\left| \begin{array}{ll} (\nabla\bar{s}, \nabla w) - (\bar{\varphi}, \nabla \cdot w) = (u, w) & \forall w \in \mathrm{H}_0^1(\Omega)^d \\ (q, \nabla \cdot \bar{s}) = 0 & \forall q \in \mathrm{L}_0^2(\Omega) \end{array} \right.$$

From the regularity of the Stokes problem, we know that $(\bar{s}, \bar{\varphi})$ belongs to $\mathrm{H}_0^2(\Omega)^d \times \mathrm{H}^1(\Omega)$ and :

$$|\bar{s}|_2 + |\bar{\varphi}|_1 \leq c\, \|u\|_0$$

In addition, if $(S_h u, \varphi)$ are given by (II.3.12) :

$$\|\bar{\varphi} - \varphi\|_0 \leq ch\,|\bar{\varphi}|_1 \leq ch\,\|u\|_0$$

Let $\Pi_h\bar{\varphi}$ be a projection of $\bar{\varphi}$ onto M_h such that :

$$\|\bar{\varphi} - \Pi_h\bar{\varphi}\|_0 + h\,|\bar{\varphi} - \Pi_h\bar{\varphi}|_1 \leq ch\,|\bar{\varphi}|_1 \leq ch\,\|u\|_0$$

Such a projection exists by the assumptions on the space M_h (for instance, the result of an orthogonal projection with respect to the $\mathrm{H}^1(\Omega)$ norm). We then have, using an inverse inequality and the three previous estimates :

$$\begin{aligned}
\|\nabla\varphi\|_0 &\leq \|\nabla(\varphi - \Pi_h\bar{\varphi})\|_0 + \|\nabla\Pi_h\bar{\varphi}\|_0 \\
&\leq \frac{c}{h}\,\|\varphi - \Pi_h\bar{\varphi}\|_0 + \|\nabla\Pi_h\bar{\varphi}\|_0 \\
&\leq \frac{c}{h}\,[\,\|\varphi - \bar{\varphi}\|_0 + \|\bar{\varphi} - \Pi_h\bar{\varphi}\|_0\,] + \|\nabla(\Pi_h\bar{\varphi} - \bar{\varphi})\|_0 + \|\nabla\bar{\varphi}\|_0 \\
&\leq c\,\|u\|_0
\end{aligned} \tag{II.3.13}$$

Using the Cauchy-Schwarz inequality, we then get from the definition of $S_h u$ that for any $w \in H_h$:

$$(\nabla S_h u, \nabla u) = -(u, \nabla\varphi) + \|u\|_0^2 = -(u - w, \nabla\varphi) + \|u\|_0^2 \geq \|u\|_0^2 - \|\nabla\varphi\|_0\,\|u - w\|_0$$

From (II.3.13), we thus have from Young inequality :

$$(\nabla S_h u, \nabla u) \geq \|u\|_0^2 - c\,\|u\|_0\,\|u - w\|_0 \geq \|u\|_0^2 - \alpha\,\|u\|_0^2 - \frac{c^2}{4\alpha}\,\|u - w\|_0^2$$

which is the bound we are searching for. □

We will need to assume in the following that the norm associated to the scalar product $(\cdot,\cdot)_h$ is equivalent on V_h to the standard $\mathrm{L}^2(\Omega)$ one :

$$\exists\ \gamma_h \geq 1 \text{ such that } \quad \frac{1}{\gamma_h}\,(v, v)_h \leq (v, v) \leq \gamma_h\,(v, v)_h \qquad \forall v \in V_h \tag{II.3.14}$$

This inequality holds in particular for the lumped mass matrix associated to the usual P_1 discretization.

Consequently, we have the following result, which is a consequence of the inf-sup stability of the discretization :

Lemma II.3.2. *There exists a positive constant c such that, for all $\psi \in M_h$, one can find $v_\psi \in V_h$ such that :*

$$(\nabla \cdot v_\psi, q) = (\psi, q)_h \quad \forall q \in M_h, \qquad \|v_\psi\|_1 \leq c\,\|\psi\|_0$$

Remark II.3.3 (An additional assumption on $(\cdot,\cdot)$ in case of Dirichlet boundary conditions)**.** The imbedding $M_h \subset \mathrm{L}^2_0(\Omega)$ never holds in practice for finite element approximation and the restriction of the pressure space to zero mean value functions is obtained through the properties of the algorithms used to solve the discrete problems. In other words, one must check that the algorithm employed at the algebraic level let at each step the pressure be an element of $\mathrm{L}^2_0(\Omega)$; this property is known to hold, for instance, for the standard Uzawa algorithm, together with some of its variants which can be seen as Krylov methods applied to the pressure Schur complement problem.

In the present case, this property must also hold for the scalar product $(\cdot,\cdot)$ for the lemma II.3.2 to be valid, in the following sense : the Riez representation in M_h of the linear form $(\psi,\cdot)$ must be an element of $\mathrm{L}^2_0(\Omega)$ whenever ψ lies in $\mathrm{L}^2_0(\Omega)$. This condition simply reads :

$$\forall \psi \in \mathrm{L}^2_0(\Omega) \cap M_h, \quad (\psi, 1_h)_h = 0$$

where 1_h stands, in the preceding equation, for the constant function of M_h equal to 1 everywhere. One can easily check that this condition holds when $(\cdot,\cdot)$ is associated to the lumped mass matrix, and should be checked for any other choice.

Finally, throughout this paper, we will repeatedly make use of the discrete Gronwall lemma, a version of which reads [20, p. 14] :

Lemma II.3.4. *Let $(h_k)_{k=0,\dots,n}$ and $(f_k)_{k=0,\dots,n}$ be two families of non-negative real numbers, g_0 a positive real number and $(\theta_k)_{k=1,\dots,n}$ a family of real numbers. We suppose that :*

$$\left| \begin{array}{ll} \theta_0 \leq g_0 & \\ \theta_k \leq g_0 + \displaystyle\sum_{i=0}^{k-1} f_i + \sum_{i=0}^{k-1} h_i \theta_i & \forall k = 1, \dots, n \end{array} \right.$$

Then the following bound holds :

$$\theta_k \leq (g_0 + \sum_{i=0}^{k-1} f_i)\ \exp(\sum_{i=0}^{k-1} h_i) \qquad \forall k = 1, \dots, n$$

II.4 Analysis of the Uzawa variant

Throughout this section, u^k, $\tilde{u}^k$ and p^k will stand for the solution obtained at step k (*i.e.* at time $t = k\ \delta t$) by the Uzawa variant of the penalty-projection method (II.2.8). We note e^k, $\tilde{e}^k$, ϵ^k the splitting errors, that are the differences $e^k = u^k - \bar{u}^k$, $\tilde{e}^k = \tilde{u}^k - \bar{u}^k$, $\epsilon^k = p^k - \bar{p}^k$ where $\bar{u}^k$ and $\bar{p}^k$ are solution of the coupled algorithm (II.2.6). We suppose that both algorithms are initialized by the same approximations of the initial data, that is $e^0 = 0$, and that $\epsilon^0 = 0$. The number N stands for the total number of time steps ($N = T/\delta t$).

This section is devoted to the analysis of the Uzawa variant, which results are gathered in the following theorem.

Theorem II.4.1. *For any strictly positive value of the penalty parameter r, the following bounds hold for $1 \leq n \leq N$:*

$$\|e^n\|_0 + \|\tilde{e}^n\|_0 + \left[\sum_{k=1}^{n} \delta t \, \|\nabla \tilde{e}^k\|_0^2\right]^{1/2} \leq c \, \frac{\delta t^{1/2}}{r}$$

$$\left[\sum_{k=1}^{n} \delta t \, \|e^k\|_0^2\right]^{1/2} + \left[\sum_{k=1}^{n} \delta t \, \|\tilde{e}^k\|_0^2\right]^{1/2} \leq c \, \frac{\delta t}{r}$$

$$\|\epsilon^n\|_h \leq c \, \frac{1}{r^{1/2}}$$

$$\left[\sum_{k=1}^{n} \delta t \, \|\epsilon^k\|_0^2\right]^{1/2} \leq c \, \frac{1}{r}$$

In addition, we have the following uniform estimates with respect to r :

$$\|e^n\|_0 + \|\tilde{e}^n\|_0 + (\sum_{k=1}^{n} \delta t \, \|\nabla \tilde{e}^k\|_0^2)^{1/2} \leq c \, l(h) \, \delta t^{1/2}$$

$$\left[\sum_{k=1}^{n} \delta t \, \|e^k\|_0^2\right]^{1/2} + \left[\sum_{k=1}^{n} \delta t \, \|\tilde{e}^k\|_0^2\right]^{1/2} \leq c \, l(h) \, \delta t$$

However, these latter estimates are not uniform with respect to h, i.e. *the function $l(h)$ may go to infinity when h tends to zero (see remark II.4.3 below).*

The proof of this theorem is obtained by several steps. We begin by establishing the system controlling the splitting errors, obtained by taking differences between the equations of system (II.2.8) and (II.2.6). For $k = 0, \ldots, N-1$, we have :

$$\left|\begin{array}{lll}
(i) & \dfrac{1}{\delta t}\,(\tilde{e}^{k+1} - e^k, v) + (\nabla \tilde{e}^{k+1}, \nabla v) + (\nabla \epsilon^{k+1}, v) = 0 & \forall v \in V_h \\
(ii) & \dfrac{1}{\delta t}\,(e^{k+1} - \tilde{e}^{k+1}, v) + (\nabla \varphi, v) = 0 & \forall v \in X_h \\
(iii) & (e^{k+1}, \nabla q) = 0 & \forall q \in M_h \\
(iv) & -(\tilde{e}^{k+1}, \nabla q) + \dfrac{1}{r}(\epsilon^{k+1} - \epsilon^k, q)_h = -\dfrac{1}{r}(\delta \bar{p}^{k+1}, q)_h & \forall q \in M_h
\end{array}\right. \tag{II.4.15}$$

Then we prove a first set of estimates for the velocity and pressure errors, valid for any value of the penalty parameter.

Lemma II.4.2. *For any strictly positive value of the penalty parameter r, the following bounds hold for $1 \leq n \leq N$:*

$$\|e^n\|_0 + \|\tilde{e}^n\|_0 + \left[\sum_{k=1}^{n} \delta t \, \|\nabla \tilde{e}^k\|_0^2\right]^{1/2} \leq c \, \frac{\delta t^{1/2}}{r}$$

$$\left[\sum_{k=1}^{n} \delta t \, \|e^k - \tilde{e}^k\|_0^2\right]^{1/2} + \left[\sum_{k=1}^{n} \delta t \, \|\tilde{e}^k - e^{k-1}\|_0^2\right]^{1/2} \leq c \, \frac{\delta t}{r}$$

$$\|\epsilon^n\|_h \leq c \, \frac{1}{r^{1/2}}$$

In addition, the following estimates address low values of r, including $r = 0$:

$$\|e^n\|_0 \,+\, \|\tilde{e}^n\|_0 \,+ (\sum_{k=1}^{n} \delta t \,\|\nabla \tilde{e}^k\|_0^2\,)^{1/2} \leq c\; l(h)\, \delta t^{1/2}$$

$$\left[\sum_{k=1}^{n} \delta t \,\|e^k - \tilde{e}^k\|_0^2\right]^{1/2} \leq c\; l(h)\, \delta t$$

However, these estimates are not uniform with respect to h, i.e. *the function $l(h)$ may go to infinity when h tends to zero (see remark II.4.3 below).*

Proof. Choosing $v = 2\,\delta t\,\tilde{e}^{k+1}$ in (II.4.15-(i)) and using the identity $2(a-b)a = a^2 + (a-b)^2 - b^2$, we get for $k = 1,\ \ldots,\ N-1$:

$$\|\tilde{e}^{k+1}\|_0^2 \,+\, \|\tilde{e}^{k+1} - e^k\|_0^2 \,-\, \|e^k\|_0^2 \,+ 2\delta t\,\|\nabla \tilde{e}^{k+1}\|_0^2 \,+ 2\delta t(\tilde{e}^{k+1}, \nabla \epsilon^{k+1}) = 0 \qquad \text{(II.4.16)}$$

Taking $q = 2\,\delta t\,\epsilon^{k+1}$ in (II.4.15-(iv)) yields :

$$-2\delta t(\tilde{e}^{k+1}, \nabla \epsilon^{k+1}) + \frac{\delta t}{r}\left[\,\|\epsilon^{k+1}\|_h^2 \,+\, \|\epsilon^{k+1} - \epsilon^k\|_h^2 \,-\, \|\epsilon^k\|_h^2\,\right] = -2\frac{\delta t}{r}(\delta \bar{p}^{k+1}, \epsilon^{k+1})_h$$

Finally, taking $v = 2\,\delta t\, e^{k+1}$ in (II.4.15-(ii)), one obtains :

$$\|e^{k+1}\|_0^2 \,+\, \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 \,-\, \|\tilde{e}^{k+1}\|_0^2 \,= 0 \qquad \text{(II.4.17)}$$

Summing these three equations yields :

$$\|e^{k+1}\|_0^2 \,+\, \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 \,+\, \|\tilde{e}^{k+1} - e^k\|_0^2 \,-\, \|e^k\|_0^2 \,+ 2\delta t\,\|\nabla \tilde{e}^{k+1}\|_0^2$$

$$+\frac{\delta t}{r}\left[\,\|\epsilon^{k+1}\|_h^2 \,+\, \|\epsilon^{k+1} - \epsilon^k\|_h^2 \,-\, \|\epsilon^k\|_h^2\,\right] = \underbrace{-\,2\frac{\delta t}{r}(\delta \bar{p}^{k+1}, \epsilon^{k+1})_h}_{T_1} \qquad \text{(II.4.18)}$$

At this point, the proof separates into two branches, to obtain estimates for respectively low and high values of the penalty parameter.

a) Estimates independent of the penalty parameter

The proof is split in two steps : first, we prove that the pressure error is bounded independently of both the time step and the penalty parameter, then we use standard arguments from the analysis of the non-incremental projection method [21] to obtain estimates for the velocity as a function of the time step itself.

Using Cauchy-Schwarz, Young and (II.3.11) inequalities, the term T_1 in equation (II.4.18) may be bounded as follows :

$$|T_1| \leq \frac{\delta t}{r}\left[\frac{1+\delta t}{\delta t}\,\|\delta \bar{p}^{k+1}\|_h^2 \,+ \frac{\delta t}{1+\delta t}\,\|\epsilon^{k+1}\|_h^2\right] \leq c\,\frac{\delta t^2}{r} + \frac{\delta t}{r}\,\frac{\delta t}{1+\delta t}\,\|\epsilon^{k+1}\|_h^2$$

From equation (II.4.18), we thus infer :

$$\|e^{k+1}\|_0^2 \,+\, \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 \,+\, \|\tilde{e}^{k+1} - e^k\|_0^2 \,-\, \|e^k\|_0^2 \,+ 2\delta t\,\|\nabla \tilde{e}^{k+1}\|_0^2$$

$$+\frac{\delta t}{r}\left[\,\|\epsilon^{k+1}\|_h^2 \,+\, \|\epsilon^{k+1} - \epsilon^k\|_h^2 \,-\, \|\epsilon^k\|_h^2\,\right] \leq c\,\frac{\delta t^2}{r} + \frac{\delta t}{r}\,\frac{\delta t}{1+\delta t}\,\|\epsilon^{k+1}\|_h^2$$

and, summing up from $k=0$ to $k=n$ and using the fact that $e^0=0$ and $\epsilon^0=0$:

$$\|e^{n+1}\|_0^2 + \sum_{k=0}^{n} \|e^{k+1}-\tilde{e}^{k+1}\|_0^2 + \sum_{k=0}^{n} \|\tilde{e}^{k+1}-e^k\|_0^2 + 2\sum_{k=0}^{n} \delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2 + \frac{\delta t}{r}\left[\frac{1}{1+\delta t}\|\epsilon^{n+1}\|_h^2 + \sum_{k=0}^{n} \|\epsilon^{k+1}-\epsilon^k\|_h^2\right] \leq c\frac{\delta t}{r} + \sum_{k=0}^{n-1} \delta t \, \frac{\delta t}{r} \, \frac{1}{1+\delta t} \|\epsilon^{k+1}\|_h^2$$

The discrete Gronwall lemma II.3.4 thus yields in particular, for $1 \leq n \leq N$:

$$\|\epsilon^n\|_h^2 \leq c \tag{II.4.19}$$

We are now in position to derive velocity error estimates. Adding this time only the equations (II.4.16) and (II.4.17), we obtain with (II.4.15-(iii)) :

$$\|e^{k+1}\|_0^2 + \|e^{k+1}-\tilde{e}^{k+1}\|_0^2 + \|\tilde{e}^{k+1}-e^k\|_0^2 - \|e^k\|_0^2 + 2\delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2 = -2\delta t(\nabla \epsilon^{k+1}, \tilde{e}^{k+1})$$
$$= 2\delta t(\nabla \epsilon^{k+1}, e^{k+1}-\tilde{e}^{k+1}) \leq \frac{1}{2}\|e^{k+1}-\tilde{e}^{k+1}\|_0^2 + 2\delta t^2 \, \|\nabla \epsilon^{k+1}\|_0^2$$

As ϵ^{k+1} belongs to the discrete (finite dimensional) space M_h, we have $\|\nabla \epsilon^{k+1}\|_0 \leq l(h) \, \|\epsilon^{k+1}\|_0$. Summing up the preceding inequality from $k=0$ to n, using (II.4.19) and the fact that both e^0 and ϵ^0 vanish, we get from (II.3.14) :

$$\|e^{n+1}\|_0^2 + \frac{1}{2}\sum_{k=0}^{n} \|e^{k+1}-\tilde{e}^{k+1}\|_0^2 + \sum_{k=0}^{n} \|\tilde{e}^{k+1}-e^k\|_0^2 + 2\sum_{k=0}^{n} \delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2 \leq 2\,\gamma_h \, l(h)^2 \, \delta t$$

b) Penalty parameter dependent estimates

The following of the proof makes use of an idea issued from the penalty methods analysis [24, section 5.2]. By lemma II.3.2 and (II.3.11), we know that there exists w^{k+1} in V_h such that :

$$(\nabla \cdot w^{k+1}, q) = -(\delta \bar{p}^{k+1}, q)_h \quad \forall q \in M_h, \qquad \|w^{k+1}\|_1 \leq c\,\|\delta \bar{p}^{k+1}\|_0 \leq c\,\delta t$$

The right-hand side of equation (II.4.18) then reads, using (II.4.15-(i)) for $v = w^{k+1}$:

$$T_1 = 2\frac{\delta t}{r}(\nabla \cdot w^{k+1}, \epsilon^{k+1}) = 2\frac{\delta t}{r}\left[\frac{1}{\delta t}(\tilde{e}^{k+1}-e^k, w^{k+1}) + (\nabla \tilde{e}^{k+1}, \nabla w^{k+1})\right]$$

We then get :

$$\begin{aligned} |T_1| &\leq \frac{2}{r}|(\tilde{e}^{k+1}-e^k, w^{k+1})| + 2\frac{\delta t}{r}|(\nabla \tilde{e}^{k+1}, \nabla w^{k+1})| \\ &\leq \frac{1}{2}\|\tilde{e}^{k+1}-e^k\|_0^2 + \frac{2}{r^2}\|w^{k+1}\|_0^2 + \delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2 + \frac{\delta t}{r^2}\|\nabla w^{k+1}\|_0^2 \\ &\leq \frac{1}{2}\|\tilde{e}^{k+1}-e^k\|_0^2 + \delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2 + c\frac{\delta t^2+\delta t^3}{r^2} \end{aligned}$$

Substituting this bound in equation (II.4.18) and dropping the highest order term with respect to δt (or, equivalently, supposing that $\delta t \leq 1$), we get :

$$\|e^{k+1}\|_0^2 + \|e^{k+1}-\tilde{e}^{k+1}\|_0^2 + \frac{1}{2}\|\tilde{e}^{k+1}-e^k\|_0^2 - \|e^k\|_0^2 + \delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2$$
$$+\frac{\delta t}{r}\left[\, \|\epsilon^{k+1}\|_h^2 + \|\epsilon^{k+1}-\epsilon^k\|_h^2 - \|\epsilon^k\|_h^2 \,\right] \leq c\frac{\delta t^2}{r^2}$$

Summing these inequalities from $k = 0$ to $k = n$ and using the fact that $e^0 = 0$ and $\epsilon^0 = 0$, we obtain the following inequality :

$$\|e^{n+1}\|_0^2 + \sum_{k=0}^{n} \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 + \frac{1}{2}\sum_{k=0}^{n} \|\tilde{e}^{k+1} - e^k\|_0^2 + \sum_{k=0}^{n} \delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2 + \frac{\delta t}{r}\left[\|\epsilon^{n+1}\|_h^2 + \sum_{k=0}^{n} \|\epsilon^{k+1} - \epsilon^k\|_h^2 \right] \leq c\frac{\delta t}{r^2}$$

□

Remark II.4.3. In general, the function $l(h)$ in the inequality $\|\nabla \epsilon^{k+1}\|_0 \leq l(h) \, \|\epsilon^{k+1}\|_0$ tends to infinity when h tends to zero. For instance, for finite element spaces, $l(h)$ behaves as c/h. However, this inequality could be improved if ϵ^{k+1} was known to converge spatially, by regularity properties of the time semi-discrete systems, to a function ϵ belonging to $\mathrm{H}^1(\Omega)$. In this condition, taking once again the example of a conforming (in $\mathrm{H}^1(\Omega)$) finite element space including piecewise linear polynomials, we would obtain : $\|\nabla \epsilon^{k+1}\|_0 \leq c \, \|\epsilon\|_1$. However, the convergence of ϵ^{k+1} to a function with a H^1 norm uniformly bounded with respect to the time step and the penalty parameter does not seem to be clear in the present case, even if numerical experiments seem to show a h-uniform behaviour. Note also that, when r is set to zero, we recover the usual non-incremental scheme, of which the convergence has been studied in [21]. In this case, the pressure error committed in the prediction step is simply $\epsilon^{k+1} = -\bar{p}^{k+1}$ and its $\mathrm{H}^1(\Omega)$ stability is ensured.

Then we can prove the following stronger estimate for the velocity (although in a weaker norm).

Lemma II.4.4. *For any strictly positive value of the penalty parameter r, we have for $1 \leq n \leq N$:*

$$\left[\sum_{k=1}^{n} \delta t \, \|e^k\|_0^2\right]^{1/2} + \left[\sum_{k=1}^{n} \delta t \, \|\tilde{e}^k\|_0^2\right]^{1/2} \leq c \, \frac{\delta t}{r}$$

and, for low values of r, including the case $r = 0$:

$$\left[\sum_{k=1}^{n} \delta t \, \|e^k\|_0^2\right]^{1/2} + \left[\sum_{k=1}^{n} \delta t \, \|\tilde{e}^k\|_0^2\right]^{1/2} \leq c \, l(h) \, \delta t$$

Proof. We start from the sum of equation (II.4.15-(i)) written at time $k+1$ and equation (II.4.15-(ii)) written at time k :

$$\frac{1}{\delta t} \, (\tilde{e}^{k+1} - \tilde{e}^k, v) + (\nabla \tilde{e}^{k+1}, \nabla v) + (\nabla \xi, v) = 0 \qquad \forall v \in V_h$$

where ξ is an element of M_h ($\xi = \varphi + \epsilon^{k+1}$). Choosing $v = 2 \, \delta t \, S_h \tilde{e}^{k+1}$ in this equation yields :

$$\|\tilde{e}^{k+1}\|_s^2 + \|\tilde{e}^{k+1} - \tilde{e}^k\|_s^2 - \|\tilde{e}^k\|_s^2 + 2\delta t \, (\nabla \tilde{e}^{k+1}, \nabla S_h \tilde{e}^{k+1}) = 0$$

Making use of lemma II.3.1 then yields :

$$\|\tilde{e}^{k+1}\|_s^2 + \|\tilde{e}^{k+1} - \tilde{e}^k\|_s^2 - \|\tilde{e}^k\|_s^2 + 2\delta t \left[c_1 \, \|\tilde{e}^{k+1}\|_0^2 - c_2 \inf_{w \in H_h} \|\tilde{e}^{k+1} - w\|_0^2 \right] \leq 0$$

Finally, choosing $w = e^{k+1}$, we obtain :

$$\|\tilde{e}^{k+1}\|_s^2 \; + \; \|\tilde{e}^{k+1} - \tilde{e}^k\|_s^2 \; - \; \|\tilde{e}^k\|_s^2 \; + 2c_1\delta t \, \|\tilde{e}^{k+1}\|_0^2 \; \leq c_2\delta t \, \|\tilde{e}^{k+1} - e^{k+1}\|_0^2$$

The result then follows by summing up this inequality from $k = 0$ to $k = n$, using lemma II.4.2 and, finally, the triangle inequality. □

The pressure error estimate follows from lemma II.4.2, as stated hereafter.
Lemma II.4.5. *We suppose that the penalty parameter r is strictly positive. Then the following bound holds, for $1 \leq n \leq N$:*

$$\left[\sum_{k=1}^{n} \delta t \, \|\epsilon^k\|_0^2 \right]^{1/2} \leq \frac{c}{r}$$

Proof. From equation (II.4.15-(i)), we infer, using the inf-sup stability of the discretization :

$$\|\epsilon^{k+1}\|_0 \; \leq c \left[\frac{1}{\delta t} \, \|\tilde{e}^{k+1} - e^k\|_0 \; + \; \|\nabla \tilde{e}^{k+1}\|_0 \right]$$

Squaring this inequality, multiplying by δt and summing up from $k = 0$ to $k = n-1$, we get :

$$\sum_{k=1}^{n} \delta t \, \|\epsilon^k\|_0^2 \; \leq c \left[\frac{1}{\delta t^2} \sum_{k=1}^{n} \delta t \, \|\tilde{e}^k - e^{k-1}\|_0^2 \; + \sum_{k=1}^{n} \delta t \, \|\nabla \tilde{e}^k\|_0^2 \right]$$

and, from lemma II.4.2 :

$$\sum_{k=1}^{n} \delta t \, \|\epsilon^k\|_0^2 \; \leq c \left[\frac{1}{r^2} + \frac{\delta t}{r^2} \right]$$

□

II.5 The standard penalty-projection method

This section addresses the analysis of the standard penalty-projection method. We shall see that this method inherits the convergence features of both the so-called rotational pressure-correction method analysed in [16] (*i.e.* second-order convergence in time of the splitting error) at low values of the penalty parameter and the penalty method (*i.e.* convergence as $\delta t/r$ of the splitting error) for high values of the penalty parameter. These results are proven in two separate sub-sections.

II.5.1 Analysis for low values of the penalty parameter

The results are gathered in the following theorem :

Theorem II.5.1. *The following bounds hold for* $1 \leq n \leq N$ *:*

$$\left[\sum_{k=0}^{n} \delta t \, \|e^k\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \delta t \, \|\tilde{e}^k\|_0^2\right]^{1/2} \leq c \ \min(\delta t^2, \frac{\delta t^{3/2}}{r^{1/2}})$$
$$\left[\sum_{k=0}^{n} \delta t \, \|\nabla \tilde{e}^k\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \delta t \, \|\epsilon^k\|_0^2\right]^{1/2} \leq c \ \max(1, \frac{1}{r^{1/2}}) \ \delta t^{3/2}$$

The proof of this result follows closely the theory developed in [16]. We begin with stating the system of equations which controls the splitting errors, valid for $0 \leq k \leq N-1$:

$$\left|\begin{array}{lll} (i) & \frac{1}{\delta t} \, (\tilde{e}^{k+1} - e^k, v) + (\nabla \tilde{e}^{k+1}, \nabla v) + r \ c_h(\tilde{e}^{k+1}, v) + (\nabla \psi^k, v) = 0 & \forall v \in V_h \\ (ii) & \frac{1}{\delta t} \, (e^{k+1} - \tilde{e}^{k+1}, v) + (\nabla(\epsilon^{k+1} - \psi^k - r \ B_h \tilde{e}^{k+1}), v) = 0 & \forall v \in X_h \\ (iii) & (e^{k+1}, \nabla q) = 0 & \forall q \in M_h \end{array}\right. \tag{II.5.20}$$

where e^{k+1}, $\tilde{e}^{k+1}$, ϵ^{k+1} stand for the difference between, respectively, the end-of-step velocity, the predicted velocity and the pressure obtained with the standard penalty-projection method (II.2.9) and the velocity and pressure obtained by the coupled algorithm (II.2.6), and ψ^k is defined by $\psi^k = p^k - \bar{p}^{k+1} = \epsilon^k - \delta \bar{p}^{k+1}$. Note that the first equation is valid in particular because the bilinear form $c_h(\bar{u}^k, v)$ vanishes for any $k \in [1, N]$ and any v in V_h (see (II.2.7)), which is a consequence of the fact that we use an algebraic formulation of the penalty term. In the opposite case, an additional error, decreasing with the mesh size and growing with the penalty parameter, would appear.

By taking the difference of this system of equations at two consecutive time steps, we obtain the equations which control the splitting error increments, for $1 \leq k \leq N-1$:

$$\left|\begin{array}{lll} (i) & \frac{1}{\delta t} \, (\delta\tilde{e}^{k+1} - \delta e^k, v) + (\nabla \delta\tilde{e}^{k+1}, \nabla v) + r \ c_h(\delta\tilde{e}^{k+1}, v) + (\nabla \delta\psi^k, v) = 0 & \forall v \in V_h \\ (ii) & \frac{1}{\delta t} \, (\delta e^{k+1} - \delta\tilde{e}^{k+1}, v) + (\nabla(\delta\epsilon^{k+1} - \delta\psi^k - r \ B_h \delta\tilde{e}^{k+1}), v) = 0 & \forall v \in X_h \\ (iii) & (\delta e^{k+1}, \nabla q) = 0 & \forall q \in M_h \end{array}\right. \tag{II.5.21}$$

where $\delta\tilde{e}^{k+1} = \tilde{e}^{k+1} - \tilde{e}^k$, $\delta e^{k+1} = e^{k+1} - e^k$, $\delta\epsilon^{k+1} = \epsilon^{k+1} - \epsilon^k$ and $\delta\psi^k = \psi^k - \psi^{k-1}$.

As a first step, we prove the following estimate.

Lemma II.5.2. *The following bounds hold for* $1 \leq n \leq N$ *:*

$$\left[\sum_{k=1}^{n} \delta t \, \|\delta\tilde{e}^k - \delta e^k\|_0^2\right]^{1/2} \leq c \ \delta t^{5/2}$$
$$\|B_h \tilde{e}^n\|_0 \leq c \ \frac{\delta t^{3/2}}{r^{1/2}}$$
$$\|\nabla(\delta\epsilon^n - r \ B_h \tilde{e}^n)\|_0 \leq c \ \delta t$$

Proof. First, we take $v = 2\delta t\, \delta\tilde{e}^{k+1}$ in (II.5.21-(i)) to obtain, using the definition of the bilinear form $c_h(\cdot,\cdot)$:

$$\|\delta\tilde{e}^{k+1}\|_0^2 + \|\delta\tilde{e}^{k+1} - \delta e^k\|_0^2 - \|\delta e^k\|_0^2 + 2\delta t\, \|\nabla\delta\tilde{e}^{k+1}\|_0^2 + 2r\delta t\, \|B_h\delta\tilde{e}^{k+1}\|_0^2 + 2\delta t(\nabla\delta\psi^k, \delta\tilde{e}^{k+1}) = 0 \tag{II.5.22}$$

Then, reordering the terms in (II.5.21-(ii)), we get :

$$([\frac{\delta e^{k+1}}{\delta t} + \nabla(\delta\epsilon^{k+1} - r\ B_h\tilde{e}^{k+1})] - [\frac{\delta\tilde{e}^{k+1}}{\delta t} + \nabla(\delta\psi^k - r\ B_h\tilde{e}^k)], v) = 0 \qquad \forall v \in X_h$$

Choosing $v = \delta t^2\ ([\frac{\delta e^{k+1}}{\delta t} + \nabla(\delta\epsilon^{k+1} - r\ B_h\tilde{e}^{k+1})] + [\frac{\delta\tilde{e}^{k+1}}{\delta t} + \nabla(\delta\psi^k - r\ B_h\tilde{e}^k)]) \in X_h$ yields :

$$\|\delta e^{k+1}\|_0^2 + \delta t^2\, \|\nabla(\delta\epsilon^{k+1} - r\ B_h\tilde{e}^{k+1})\|_0^2 - \|\delta\tilde{e}^{k+1}\|_0^2 - \delta t^2\, \|\nabla(\delta\psi^k - r\ B_h\tilde{e}^k)\|_0^2 - 2\delta t(\nabla(\delta\psi^k - r\ B_h\tilde{e}^k), \delta\tilde{e}^{k+1}) = 0$$

Developing the last term yields :

$$\|\delta e^{k+1}\|_0^2 + \delta t^2\, \|\nabla(\delta\epsilon^{k+1} - r\ B_h\tilde{e}^{k+1})\|_0^2 - \|\delta\tilde{e}^{k+1}\|_0^2 - 2\delta t(\nabla\delta\psi^k, \delta\tilde{e}^{k+1}) + 2r\delta t(\nabla B_h\tilde{e}^k, \delta\tilde{e}^{k+1}) = \delta t^2\, \|\nabla(\delta\psi^k - r\ B_h\tilde{e}^k)\|_0^2 \tag{II.5.23}$$

As $B_h\tilde{e}^k$ belongs to M_h, by the definition of the operator B_h, we have :

$$(\nabla B_h\tilde{e}^k, \delta\tilde{e}^{k+1}) = (B_h\tilde{e}^k, B_h\delta\tilde{e}^{k+1})_h$$

Developing, we get :

$$-2\delta tr(\nabla\cdot\delta\tilde{e}^{k+1}, B_h\tilde{e}^k) = \delta tr\left[\, \|B_h\tilde{e}^{k+1}\|_h^2 - \|B_h\delta\tilde{e}^{k+1}\|_h^2 - \|B_h\tilde{e}^k\|_h^2 \,\right] \tag{II.5.24}$$

On the other hand, the right-hand side in (II.5.23) reads :

$$\begin{aligned}
\delta t^2\, \|\nabla(\delta\psi^k - r\ B_h\tilde{e}^k)\|_0^2 &= \delta t^2\, \|\nabla(\delta\epsilon^k - r\ B_h\tilde{e}^k) - \nabla\delta\delta\bar{p}^{k+1}\|_0^2 \\
&\le \delta t^2(1+\delta t)\, \|\nabla(\delta\epsilon^k - r\ B_h\tilde{e}^k)\|_0^2 + \delta t^2(1+\frac{1}{\delta t})\, \|\nabla\delta\delta\bar{p}^{k+1}\|_0^2 \\
&\le \delta t^2(1+\delta t)\, \|\nabla(\delta\epsilon^k - r\ B_h\tilde{e}^k)\|_0^2 + c\delta t^5
\end{aligned}$$

Combining these two latter estimates with the equations (II.5.22) and (II.5.23), we observe that the term $\delta t\ r\, \|B_h\delta\tilde{e}^{k+1}\|_h^2$ in (II.5.24) can be absorbed in the penalty term in (II.5.22) (see remark below), and we get for $1 \le k \le N-1$:

$$\begin{aligned}
&\|\delta e^{k+1}\|_0^2 + \|\delta\tilde{e}^{k+1} - \delta e^k\|_0^2 - \|\delta e^k\|_0^2 + 2\delta t\, \|\nabla\delta\tilde{e}^{k+1}\|_0^2 + \delta t\ r\, \|B_h\delta\tilde{e}^{k+1}\|_0^2 \\
&+\delta t^2\, \|\nabla(\delta\epsilon^{k+1} - r\ B_h\tilde{e}^{k+1})\|_0^2 + 2\delta t\ r\, \|B_h\tilde{e}^{k+1}\|_0^2 - 2\delta t\ r\, \|B_h\tilde{e}^k\|_0^2 \\
&\qquad \le \delta t^2(1+\delta t)\, \|\nabla(\delta\epsilon^k - r\ B_h\tilde{e}^k)\|_0^2 + c\delta t^5
\end{aligned}$$

To apply the Gronwall lemma, we need an estimate for $\|\delta e^1\|_0^2$, $\|B_h\tilde{e}^k\|_0^2$ and $\|\nabla(\delta\epsilon^1 - r\ B_h\tilde{e}^1)\|_0^2$, *i.e.* , as $e^0 = 0$ and $\epsilon^0 = 0$, $\|e^1\|_0^2$, $\|B_h\tilde{e}^k\|_0^2$ and

$\|\nabla(\epsilon^1 - r\ B_h\tilde{e}^1)\|_0^2$. Using once again $e^0 = 0$ and $\epsilon^0 = 0$, the system of equations controlling the splitting error at the first time step reads :

$$\left|\begin{array}{ll} (i) & \dfrac{1}{\delta t}\ (\tilde{e}^1, v) + (\nabla\tilde{e}^1, \nabla v) + r\ c_h(\tilde{e}^1, v) = (\nabla\delta\bar{p}^1, v) \qquad \forall v \in V_h \\ (ii) & \dfrac{1}{\delta t}\ (e^1 - \tilde{e}^1, v) + (\nabla(\epsilon^1 - r\ B_h\tilde{e}^1), v) = (\nabla\delta\bar{p}^1, v) \qquad \forall v \in X_h \\ (iii) & (e^1, \nabla q) = 0 \qquad \forall q \in M_h \end{array}\right.$$

Since, by assumption, $\|\nabla\delta\bar{p}^1\|_0 \leq c\delta t$, taking $v = \delta t\,\tilde{e}^1$ in the first relation yields using Young's inequality :

$$\frac{1}{2}\,\|\tilde{e}^1\|_0^2 + \delta t\,\|\nabla\tilde{e}^1\|_0^2 + r\delta t\,\|B_h\tilde{e}^1\|_h^2 \leq c\delta t^4$$

Remarking $\|e^1\|_0^2 = \|\tilde{e}^1\|_0^2 - \|e^1 - \tilde{e}^1\|_0^2 \leq \|\tilde{e}^1\|_0^2$, this relation gives the first two estimates. Then taking $v = \nabla(\epsilon^1 - r\ B_h\tilde{e}^1)$ in the second relation and using the third one, we obtain :

$$\|\nabla(\delta\epsilon^1 - r\ B_h\tilde{e}^1)\|_0^2 \leq c\delta t^2$$

The estimates of the lemma then follow by applying the discrete Gronwall lemma and remarking that, by orthogonality of the velocity correction with the discretely divergence free fields, we have $\|\delta\tilde{e}^{k+1} - \delta e^k\|_0^2 = \|\delta\tilde{e}^{k+1} - \delta e^{k+1}\|_0^2 + \|\delta e^{k+1} - \delta e^k\|_0^2$ and thus $\|\delta\tilde{e}^{k+1} - \delta e^{k+1}\|_0^2 \leq \|\delta\tilde{e}^{k+1} - \delta e^k\|_0^2$. □

Remark II.5.3. This proof is the analogue of the first step in the analysis of rotational pressure-correction methods [16, lemma 4.1]. By comparison, one may note that the penalization in the velocity prediction step yields two improvements. First, we do not need to absorb the term proportional to $\|B_h\delta\tilde{e}^{k+1}\|_h^2$ by the term proportional to $\|\nabla\delta\tilde{e}^{k+1}\|_0^2$ at the left-hand side; it means that the method is stable whatever the value of r may be and, thinking about Navier-Stokes equations, whatever the value of the viscosity may be. Second, as expected, we obtain a better control of the divergence of $\tilde{e}^{k+1}$ and $\delta\tilde{e}^{k+1}$ (division by $r^{1/2}$).

Remark II.5.4. By contrast with the penalty method analysed in the previous section, we may note that this bound (and the rest of the analysis) shows that the method will work with a pressure update of the form $p^{k+1} = p^k + \rho B_h\tilde{e}^{k+1} + \phi$, provided that $\rho \leq 2r$. However, $\rho = r$ seems to allow an optimal control of both the divergence of $\delta\tilde{e}^{k+1}$ and $\tilde{e}^{k+1}$.

We are now in position to derive error estimates for the velocity.

Lemma II.5.5. *The following bound holds for* $1 \leq n \leq N$:

$$\left[\sum_{k=0}^{n} \delta t\,\|\tilde{e}^k\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \delta t\,\|e^k\|_0^2\right]^{1/2} \leq c\ \min(\delta t^2, \frac{\delta t^{3/2}}{r^{1/2}})$$

Proof. Combining (II.5.20-(i)) and (II.5.20-(ii)) written at the previous time step, one gets :

$$\frac{1}{\delta t}\ (\tilde{e}^{k+1} - \tilde{e}^k, v) + (\nabla\tilde{e}^{k+1}, \nabla v) + (\nabla\xi, v) = 0 \quad \forall v \in V_h$$

where ξ is an element of M_h. Choosing $v = 2\,\delta t\,S_h\tilde{e}^{k+1}$ then yields, as in the proof of lemma II.4.4 :

$$\|\tilde{e}^{k+1}\|_s^2 + \|\tilde{e}^{k+1} - \tilde{e}^k\|_s^2 - \|\tilde{e}^k\|_s^2 + 2\alpha\delta t\,\|\tilde{e}^{k+1}\|_0^2 \leq 2\delta t c_\alpha\,\|\tilde{e}^{k+1} - e^{k+1}\|_0^2$$

For $1 \leq n \leq N$, summing up these equations from $k = 0$ to $k = n$ yields :

$$\sum_{k=0}^{n} \delta t \, \|\tilde{e}^{k+1}\|_0^2 \leq c \sum_{k=0}^{n} \delta t \, \|\tilde{e}^{k+1} - e^{k+1}\|_0^2 \tag{II.5.25}$$

Let ϕ be the pressure increment in equation (II.5.20-(ii)), *i.e.* :

$$\phi = \epsilon^{k+1} - \psi^k - r \; B_h \tilde{e}^{k+1} = \delta\epsilon^{k+1} + \delta\bar{p}^{k+1} - r \; B_h \tilde{e}^{k+1}$$

We have from (II.5.20-(ii)), for $0 \leq k \leq N-1$:

$$\|\tilde{e}^{k+1} - e^{k+1}\|_0 \leq \delta t \, \|\nabla\phi\|_0$$

Thanks to lemma II.5.2, we are going now to derive two different estimates for $\|\nabla\phi\|_0$. On one hand, we have :

$$\|\nabla\phi\|_0 \leq \|\nabla(\delta\epsilon^{k+1} - rB_h\tilde{e}^{k+1})\|_0 + \|\nabla\delta\bar{p}^{k+1}\|_0 \leq c\delta t \tag{II.5.26}$$

On the other hand, choosing $v = \nabla\phi$ in (II.5.20-(ii)) gives :

$$\|\nabla\phi\|_0^2 = \frac{1}{\delta t} \, (\tilde{e}^{k+1}, \nabla\phi) = \frac{1}{\delta t} \, (B_h\tilde{e}^{k+1}, \phi)_h$$

Thus, by a generalized Poincaré-Friedrichs inequality, since $\phi \in \mathrm{L}_0^2$:

$$\|\nabla\phi\|_0 \leq \frac{c}{\delta t} \; \|B_h\tilde{e}^{k+1}\|_h \leq c\frac{\delta t^{1/2}}{r^{1/2}} \tag{II.5.27}$$

Returning to the inequality (II.5.25), estimates (II.5.26) and (II.5.27) yield the first part of the result, namely the control of $\tilde{e}^{k+1}$. The second one is obtained by remarking that, by choosing $v = 2\,\delta t\, e^{k+1}$ in equation (II.5.20-(ii)), we get :

$$\|e^{k+1}\|_0^2 = \|\tilde{e}^{k+1}\|_0^2 - \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 \leq \|\tilde{e}^{k+1}\|_0^2$$

□

Lemma II.5.6. *The following bound holds for* $1 \leq n \leq N$ *:*

$$\left[\sum_{k=1}^{n} \delta t \, \|\delta\tilde{e}^k\|_0^2\right]^{1/2} + \left[\sum_{k=1}^{n} \delta t \, \|\delta e^k\|_0^2\right]^{1/2} \leq c \; \delta t^{5/2}$$

Proof. The starting point is the sum of equations (II.5.21-(ii)) and (II.5.21-(i)) respectively written at time k and $k-1$:

$$\frac{1}{\delta t} \, (\delta\tilde{e}^{k+1} - \delta\tilde{e}^k, v) + (\nabla\delta\tilde{e}^{k+1}, \nabla v) + (\nabla\xi, v) = 0 \quad \forall v \in V_h$$

where ξ is an element of M_h. Following the same line as in the proof of the preceding lemma, we get :

$$\sum_{k=0}^{n} \delta t \, \|\delta\tilde{e}^{k+1}\|_0^2 \leq c \sum_{k=0}^{n} \delta t \, \|\delta\tilde{e}^{k+1} - \delta e^{k+1}\|_0^2$$

The first part of the result (control of $\delta\tilde{e}^k$) then follows by lemma II.5.2. To obtain the second part, its sufficient to remark that taking $v = 2\,dt\,\delta e^{k+1}$ in (II.5.21-(ii)) yields :

$$\|\delta e^{k+1}\|_0^2 = \|\delta\tilde{e}^{k+1}\|_0^2 - \|\delta e^{k+1} - \delta\tilde{e}^{k+1}\|_0^2 \leq \|\delta\tilde{e}^{k+1}\|_0^2$$

□

Lemma II.5.7. *The following bound holds for* $1 \leq n \leq N$ *:*

$$\left[\sum_{k=0}^{n} \delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \delta t \, \|\epsilon^{k+1}\|_0^2\right]^{1/2} \leq c \ \max(1, \frac{1}{r^{1/2}}) \ \delta t^{3/2}$$

Proof. By lemma II.5.2 and the hypothesis on the bilinear form $(\cdot,\cdot)_h$, the following bound holds :

$$\forall q \in M_h, \qquad |(\nabla \cdot \tilde{e}^{k+1}, q)| = |(B_h \tilde{e}^{k+1}, q)_h| \leq \ \|B_h \tilde{e}^{k+1}\|_h \ \|q\|_h \ \leq c \ \frac{\delta t^{3/2}}{r^{1/2}} \ \|q\|_h$$
$$\leq c \ \frac{\delta t^{3/2}}{r^{1/2}} \ \|q\|_0$$

Summing equations (II.5.20-(i)) and (II.5.20-(ii)), we then obtain from the preceding estimate that $\tilde{e}^{k+1}$ and ϵ^{k+1} obeys the following system :

$$\begin{cases} (\nabla \tilde{e}^{k+1}, \nabla v) + (\nabla \epsilon^{k+1}, v) = -\dfrac{1}{\delta t} \ (e^{k+1} - e^k, v) & \forall v \in V_h \\ (\nabla \cdot \tilde{e}^{k+1}, q) = (g, q) & \forall q \in M_h \end{cases}$$

where $\|g\|_0 \ \leq c \ \dfrac{\delta t^{3/2}}{r^{1/2}}$. The results follow from lemma II.5.6 by stability of the Stokes problem. □

Remark II.5.8. The estimates of lemma II.5.7 explode for $r = 0$, which is clearly sub-optimal, as the standard variant of the penalty-projection method degenerates in this case into the classical incremental projection method, which is known to be stable and convergent. This is due to the fact that the techniques used in this proof, which are issued of the analysis of the rotational variant in [16], do not apply to the case $r = 0$. At this point, it is worth to note that a rotational version of the penalty-projection scheme can be defined by just adding $\mu B_h \tilde{u}^{k+1}$ to the pressure increment, where μ stands for the viscosity (here, $\mu = 1$). In other words, the rotational penalty-projection algorithm is obtained by replacing r by $r + \mu$ in (II.2.9-(ii)) and leaving (II.2.9-(i)) and (II.2.9-(iii)) unchanged. For this method, the present analysis yields the same estimates as in lemma II.5.7, with r changed to $r + \mu$, and the bound does not explode anymore for $r = 0$. This scheme has been tested numerically in [18] ; roughly speaking, the results for the velocity are left unchanged while, for low values of r, the pressure approximation is significantly improved.

II.5.2 Analysis for high values of the penalty parameter

In this section, we establish estimates which describe the behaviour of the standard penalty-projection method at high values of the penalty parameter r. The results of this section are gathered in the following theorem.

Theorem II.5.9. *For any strictly positive value of the penalty parameter r, the following bounds hold for $1 \leq n \leq N$:*

$$\|e^n\|_0 \ + \ \|\tilde{e}^n\|_0 \ + \left[\sum_{k=0}^{n} \delta t \, \|\nabla \tilde{e}^k\|_0^2\right]^{1/2} \leq c \ \frac{\delta t^{1/2}}{r}$$

$$\left[\sum_{k=0}^{n} \delta t \, \|e^k\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \delta t \, \|\tilde{e}^k\|_0^2\right]^{1/2} \leq c \ \frac{\delta t}{r}$$

$$\|\epsilon^n\|_h \ \leq c \ \frac{1}{r^{1/2}}$$

$$\left[\sum_{k=0}^{n} \delta t \, \|\epsilon^k\|_0^2\right]^{1/2} \leq c \ \frac{1}{r}$$

The starting point for this part of the analysis is now the system (II.2.10). By taking the difference with the variational formulation of the coupled system (II.2.6), we obtain a system of equations controlling the splitting errors :

$$\left|\begin{array}{lll} (i) & \dfrac{1}{\delta t}\,(\tilde{e}^{k+1} - e^k, v) + (\nabla \tilde{e}^{k+1}, \nabla v) + (\nabla \tilde{\epsilon}^{k+1}, v) = 0 & \forall v \in V_h \\ (ii) & -(\tilde{e}^{k+1}, \nabla q) + \dfrac{1}{r}(\tilde{\epsilon}^{k+1} - \epsilon^k, q)_h = -\dfrac{1}{r}(\delta \bar{p}^{k+1}, q)_h & \forall q \in M_h \\ (iii) & \dfrac{1}{\delta t}\,(e^{k+1} - \tilde{e}^{k+1}, v) + (\nabla(\epsilon^{k+1} - \tilde{\epsilon}^{k+1}), v) = 0 & \forall v \in X_h \\ (iv) & (e^{k+1}, \nabla q) = 0 & \forall q \in M_h \end{array}\right. \tag{II.5.28}$$

where $\tilde{\epsilon}^{k+1}$ stands for the difference between the intermediate pressure $\tilde{p}^{k+1}$ and the pressure given by the coupled algorithm : $\tilde{\epsilon}^{k+1} = \tilde{p}^{k+1} - \bar{p}^{k+1}$.

We then begin by proving the following set of estimates.

Lemma II.5.10. *For any strictly positive value of the penalty parameter r, the following bounds hold for $1 \leq n \leq N$:*

$$\|e^n\|_0 \ + \ \|\tilde{e}^n\|_0 \ + \left[\sum_{k=0}^{n} \delta t \, \|\nabla \tilde{e}^k\|_0^2\right]^{1/2} \leq c \ \frac{\delta t^{1/2}}{r}$$

$$\left[\sum_{k=0}^{n} \delta t \, \|e^k - \tilde{e}^k\|_0^2\right]^{1/2} + \left[\sum_{k=0}^{n} \delta t \, \|\tilde{e}^k - e^{k-1}\|_0^2\right]^{1/2} \leq c \ \frac{\delta t}{r}$$

$$\|\epsilon^n\|_h \ \leq c \ \frac{1}{r^{1/2}}$$

Proof. Choosing $v = 2\,\delta t\,\tilde{e}^{k+1}$ in (II.5.28-(i)), we get for $k = 0, \ \ldots, \ N-1$:

$$\|\tilde{e}^{k+1}\|_0^2 \ + \ \|\tilde{e}^{k+1} - e^k\|_0^2 \ - \ \|e^k\|_0^2 \ + 2\delta t \, \|\nabla \tilde{e}^{k+1}\|_0^2 \ + 2\delta t (\tilde{e}^{k+1}, \nabla \tilde{\epsilon}^{k+1}) = 0 \tag{II.5.29}$$

Taking $q = 2\,\delta t\,\tilde{\epsilon}^{k+1}$ in (II.5.28-(ii)) yields :

$$-2\delta t (\tilde{e}^{k+1}, \nabla \tilde{\epsilon}^{k+1}) + \frac{\delta t}{r}\left[\, \|\tilde{\epsilon}^{k+1}\|_h^2 \ + \ \|\tilde{\epsilon}^{k+1} - \epsilon^k\|_h^2 \ - \ \|\epsilon^k\|_h^2 \,\right] = -\frac{2\delta t}{r}(\delta \bar{p}^{k+1}, \tilde{\epsilon}^{k+1})_h \tag{II.5.30}$$

Then, taking $v = 2\,\delta t\, e^{k+1}$ in (II.5.28-(iii)), one obtains with (II.5.28-(iv)) :

$$\|e^{k+1}\|_0^2 \;+\; \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 \;-\; \|\tilde{e}^{k+1}\|_0^2 \;= 0 \qquad \text{(II.5.31)}$$

Finally, let $\xi \in M_h$ be defined by :

$$(\nabla\xi, \nabla q) = (\epsilon^{k+1}, q)_h \qquad \forall q \in M_h$$

Choosing $v = 2\,\nabla\xi$ in (II.5.28-(iii)) yields :

$$\frac{2}{\delta t}(e^{k+1}-\tilde{e}^{k+1}, \nabla\xi)+2(\nabla(\epsilon^{k+1}-\tilde{\epsilon}^{k+1}), \nabla\xi) = \frac{2}{\delta t}(e^{k+1}-\tilde{e}^{k+1}, \nabla\xi)+2(\epsilon^{k+1}-\tilde{\epsilon}^{k+1}, \epsilon^{k+1})_h = 0$$

and thus :

$$\frac{\delta t}{r}\left[\, \|\epsilon^{k+1}\|_h^2 \;+\; \|\epsilon^{k+1} - \tilde{\epsilon}^{k+1}\|_h^2 \;-\; \|\tilde{\epsilon}^{k+1}\|_h^2 \,\right] = -\frac{2}{r}(e^{k+1} - \tilde{e}^{k+1}, \nabla\xi)$$

By equation (II.5.20-(ii)), the right hand side of these latter equation reads :

$$-\frac{2}{r}(e^{k+1} - \tilde{e}^{k+1}, \nabla\xi) = \frac{2\delta t}{r}(\nabla\phi, \nabla\xi)$$

where ϕ is the pressure increment defined in the proof of lemma II.5.5, which is known from inequalities (II.5.26) and (II.5.27) and the Poincaré-Friedrichs inequality to satisfy :

$$\|\phi\|_h \;\leq c\delta t \quad , \qquad \|\phi\|_h \;\leq c\frac{\delta t^{1/2}}{r^{1/2}}$$

By the definition of ξ, we then get :

$$-\frac{2}{r}(e^{k+1} - \tilde{e}^{k+1}, \nabla\xi) = \frac{2\delta t}{r}(\epsilon^{k+1}, \phi)_h$$

and, finally :

$$\frac{\delta t}{r}\left[\, \|\epsilon^{k+1}\|_h^2 \;+\; \|\epsilon^{k+1} - \tilde{\epsilon}^{k+1}\|_h^2 \;-\; \|\tilde{\epsilon}^{k+1}\|_h^2 \,\right] = \frac{2\delta t}{r}(\epsilon^{k+1}, \phi)_h \qquad \text{(II.5.32)}$$

Summing up the four equations (II.5.29)-(II.5.32), we have :

$$\begin{aligned} &\|e^{k+1}\|_0^2 \;+\; \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 \;+\; \|\tilde{e}^{k+1} - e^k\|_0^2 \;-\; \|e^k\|_0^2 \;+ 2\delta t\, \|\nabla\tilde{e}^{k+1}\|_0^2 \\ &+\frac{\delta t}{r}\left[\, \|\epsilon^{k+1}\|_h^2 \;+\; \|\epsilon^{k+1} - \tilde{\epsilon}^{k+1}\|_h^2 \;+\; \|\tilde{\epsilon}^{k+1} - \epsilon^k\|_h^2 \;-\; \|\epsilon^k\|_h^2 \,\right] \\ &\qquad = -\frac{2\delta t}{r}(\delta\bar{p}^{k+1}, \tilde{\epsilon}^{k+1})_h + \frac{2\delta t}{r}(\epsilon^{k+1}, \phi)_h \end{aligned}$$

We decompose the last term of this equation to obtain :

$$\begin{aligned} &\|e^{k+1}\|_0^2 \;+\; \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 \;+\; \|\tilde{e}^{k+1} - e^k\|_0^2 \;-\; \|e^k\|_0^2 \;+ 2\delta t\, \|\nabla\tilde{e}^{k+1}\|_0^2 \\ &+\frac{\delta t}{r}\left[\, \|\epsilon^{k+1}\|_h^2 \;+\; \|\epsilon^{k+1} - \tilde{\epsilon}^{k+1}\|_h^2 \;+\; \|\tilde{\epsilon}^{k+1} - \epsilon^k\|_h^2 \;-\; \|\epsilon^k\|_h^2 \,\right] \\ &\qquad = \underbrace{\frac{2\delta t}{r}(-\delta\bar{p}^{k+1} + \phi, \tilde{\epsilon}^{k+1})_h}_{T_1} + \underbrace{\frac{2\delta t}{r}(\epsilon^{k+1} - \tilde{\epsilon}^{k+1}, \phi)_h}_{T_2} \end{aligned} \qquad \text{(II.5.33)}$$

Using the same trick as in the proof of lemma II.4.2 and taking benefit from the fact that both $\|\delta\bar{p}^{k+1}\|_0$ and $\|\phi\|_h$ are known to be bounded by $c\,\delta t$, we choose $w \in V_h$ such that :

$$(\nabla \cdot w, q) = (-\delta\bar{p}^{k+1} + \phi, q)_h \quad \forall q \in M_h, \qquad \|w\|_1 \leq c\,\delta t$$

By equation (II.5.28-(i)), we then have :

$$T_1 = \frac{2\delta t}{r}(\nabla \cdot w, \tilde{\epsilon}^{k+1}) = -\frac{2\delta t}{r}(w, \nabla\tilde{\epsilon}^{k+1}) = \frac{2\delta t}{r}\left[\frac{1}{\delta t}(\tilde{e}^{k+1} - e^k, w) + (\nabla\tilde{e}^{k+1}, \nabla w)\right]$$

and thus :

$$\begin{aligned} |T_1| &\leq \frac{1}{2}\|\tilde{e}^{k+1} - e^k\|_0^2 + \frac{2}{r^2}\|w\|_0^2 + \delta t\,\|\nabla\tilde{e}^{k+1}\|_0^2 + \frac{\delta t}{r^2}\|\nabla w\|_0^2 \\ &\leq \frac{1}{2}\|\tilde{e}^{k+1} - e^k\|_0^2 + \delta t\,\|\nabla\tilde{e}^{k+1}\|_0^2 + c\frac{\delta t^2}{r^2} \end{aligned}$$

In addition, using the fact that $\|\phi\|_h^2$ is known to be bounded by $c\,\delta t/r$:

$$|T_2| \leq \frac{\delta t}{2r}\|\epsilon^{k+1} - \tilde{\epsilon}^{k+1}\|_h^2 + \frac{2\delta t}{r}\|\phi\|_h^2 \leq \frac{\delta t}{2r}\|\epsilon^{k+1} - \tilde{\epsilon}^{k+1}\|_h^2 + c\frac{\delta t^2}{r^2}$$

Returning to equation (II.5.33) and absorbing terms, we then obtain :

$$\begin{aligned} &\|e^{k+1}\|_0^2 + \|e^{k+1} - \tilde{e}^{k+1}\|_0^2 + \frac{1}{2}\|\tilde{e}^{k+1} - e^k\|_0^2 - \|e^k\|_0^2 + \delta t\,\|\nabla\tilde{e}^{k+1}\|_0^2 \\ &+\frac{\delta t}{r}\left[\|\epsilon^{k+1}\|_h^2 + \frac{1}{2}\|\epsilon^{k+1} - \tilde{\epsilon}^{k+1}\|_h^2 + \|\tilde{\epsilon}^{k+1} - \epsilon^k\|_h^2 - \|\epsilon^k\|_h^2\right] \leq c\frac{\delta t^2}{r^2} \end{aligned}$$

The proof of the lemma then follows by summing up this inequality written for $k = 1$ up to n and using the fact that both $e^0 = 0$ and $\epsilon^0 = 0$. □

The proof of the theorem is then completed by using strictly the same line that for lemma II.4.4 and lemma II.4.5.

II.6 Numerical tests

The aim of this section is to check the validity of the theoretical analysis against a practical test case for which an analytic solution can be exhibited. This solution is built as follows. We choose a stream function and a geometrical domain such that homogeneous Dirichlet conditions holds :

$$\varphi = \frac{1}{4\pi}\,\left[\,\sin(2\pi x)\sin(2\pi y)\,\right]^2\exp(-t) \quad , \qquad \Omega =]0,1[\times]0,1[\quad , \qquad \bar{u} = \begin{bmatrix} \dfrac{\partial\varphi}{\partial y} \\ -\dfrac{\partial\varphi}{\partial x} \end{bmatrix}$$

then we pick an arbitrary pressure in $\mathrm{L}_0^2(\Omega)$:

$$p = \sin(2\pi x)\sin(2\pi y)\exp(-t)$$

and the right-hand side f is computed in order that the equations of the Stokes problem (II.1.1) are satisfied.

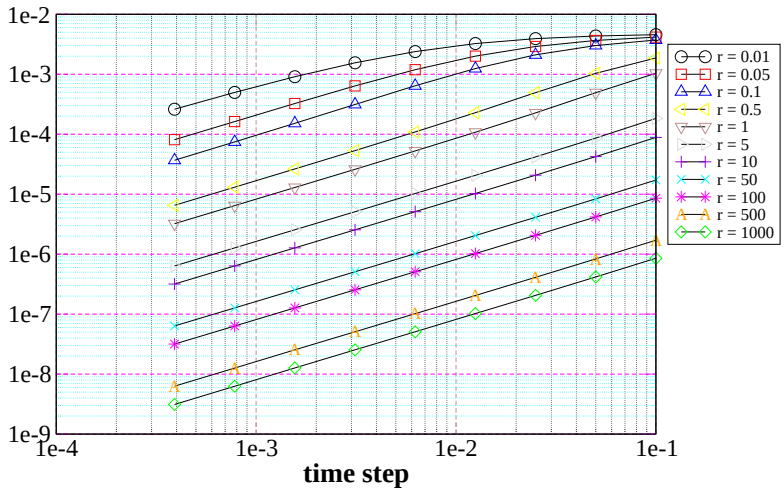


FIG. II.1 – Uzawa variant - L^2 norm of $\tilde{e}$ at $t = 1$, as a function of the time step.

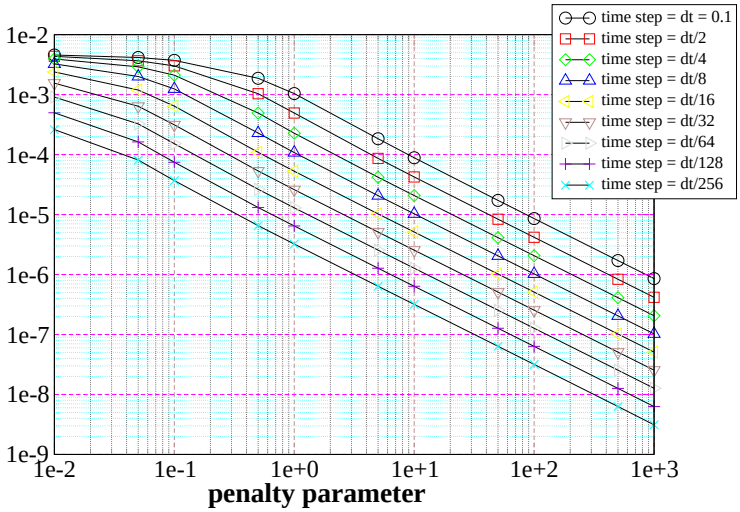


FIG. II.2 – Uzawa variant - L^2 norm of $\tilde{e}$ at $t = 1$, as a function of the penalty parameter.

Figures II.1, II.2, II.3 and II.4 show, for the Uzawa variant, the L^2 norm of the error of respectively the velocity then the pressure as a function of the time step and the penalty parameter. These results confirm the theoretical analysis : the velocity error behaves as $\delta t/r$ whereas the pressure error varies as $1/r$ as a function of the penalty parameter and, after a decrease for large time steps and penalty parameters, it does not decrease anymore with δt.

Figures II.5, II.6, II.7 and II.8 show the same curves for the standard penalty-projection scheme. For $r = 0$, we observe for the velocity splitting error the expected second-order convergence with respect to the time step, whereas the pressure splitting error convergence is also second-order, *i.e.* better than expected. This phenomenon is not new, and has received several explanations. First, in finite dimension, all the norms are equivalent and a numerical experiment with a fixed meshing cannot discriminate between a convergence in L^2 norm and a convergence in a weaker norm ; one will find an example of such a result in [15], for the pressure obtained with the standard incremental projection method (*i.e.* the method recovered here

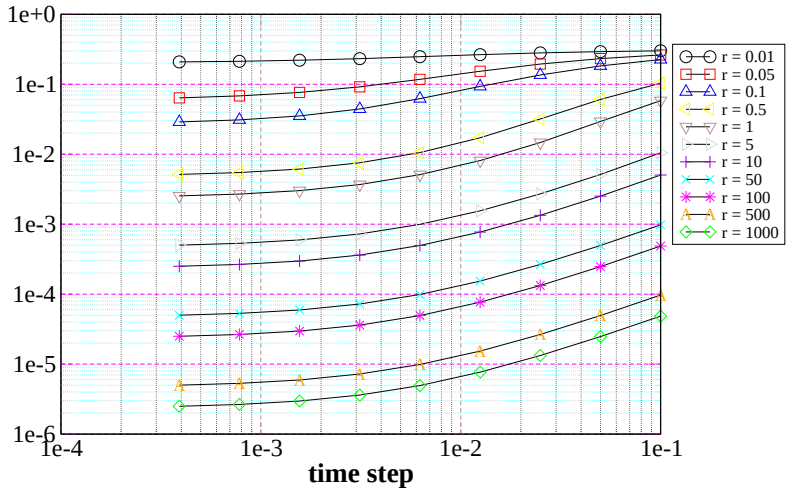


FIG. II.3 – Uzawa variant - L^2 norm of ϵ at $t = 1$, as a function of the time step.

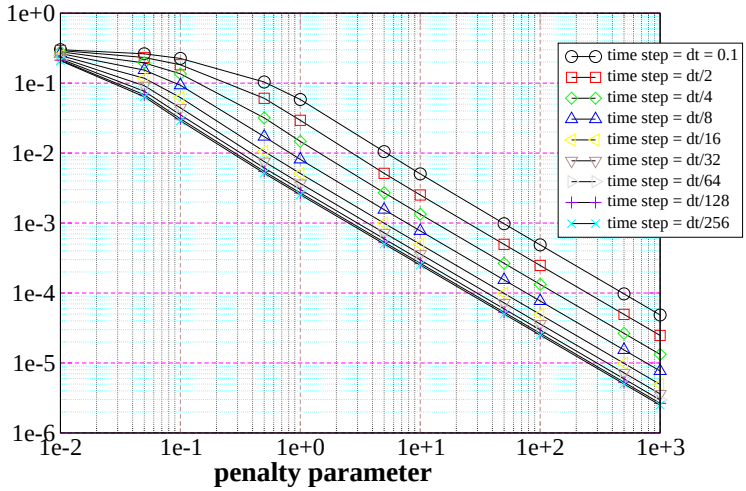


FIG. II.4 – Uzawa variant - L^2 norm of ϵ at $t = 1$, as a function of the penalty parameter.

for $r = 0$). Second, the pressure convergence rate has been observed to be strongly dependent on the regularity of the domain ; for a discussion on this point, see [16, section 5.3].

The comparison of both methods leads to the following observations. As far as the velocity is concerned, the standard penalty-projection method and the Uzawa variant are equivalent for large values of the penalty parameter ; for instance, the evolution of the error as a function of the time step for $r = 1000$ is very similar for both schemes. On the opposite, for small values of r (desirable in practice for reasons of conditioning of the algebraic operator in the prediction step), the standard method is far more accurate ; for instance, for $r = 1$, the errors ratio is about 5 for $\delta t = 0.1$ and increases for small values of the time step up to exceed 10^3 for $\delta t = 4.10^{-4}$. For the pressure, the results of both methods are similar for large values of the penalty parameter and of the time step, but for all the other cases, the standard method is still much more accurate ; the ratio between the errors is close to 10 for ($r = 1000,\ \delta t = 4.10^{-4}$) and ($r = 1,\ \delta t = 0.1$) and reaches 10^5 for ($r = 1,\ \delta t = 4.10^{-4}$).

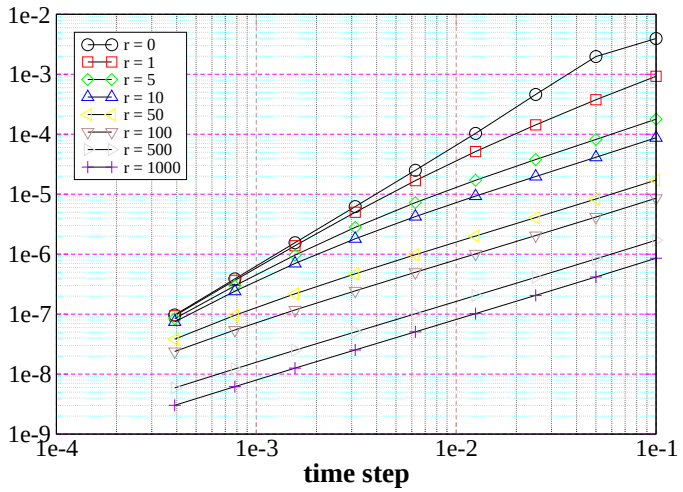


FIG. II.5 – Standard method - L^2 norm of $\tilde{e}$ at $t = 1$, as a function of the time step.

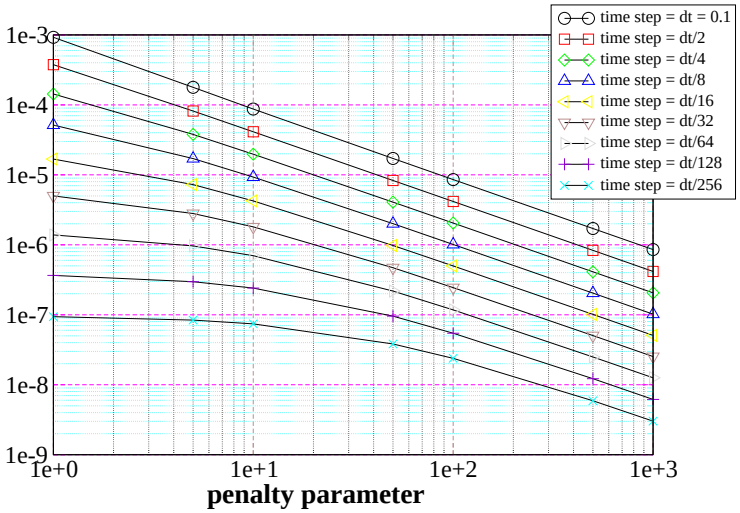


FIG. II.6 – Standard method - L^2 norm of $\tilde{e}$ at $t = 1$, as a function of the penalty parameter.

The general conclusion which can be drawn from this study is thus that the standard penalty-projection scheme appears, for a similar computational cost, to be much more accurate than the Uzawa variant. This scheme then seems to be preferred. An extensive computational study of this method for Navier-Stokes equations, with Dirichlet and open boundary conditions, can be found in [18]; these results are coherent with the present ones. In addition, we observe in [18] that, for second-order in time discretizations, the splitting errors become dominant for usual projection schemes, at values of the time step affordable in practical applications ; this reinforces the interest of the penalty-projection scheme.

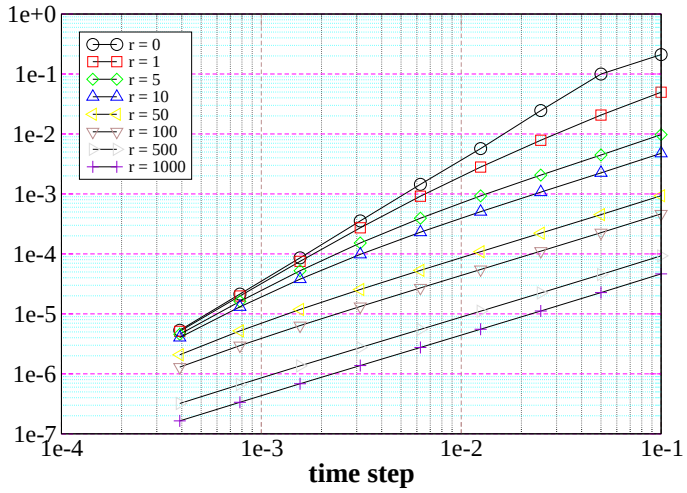


FIG. II.7 – Standard method - L^2 norm of ϵ at $t = 1$, as a function of the time step.

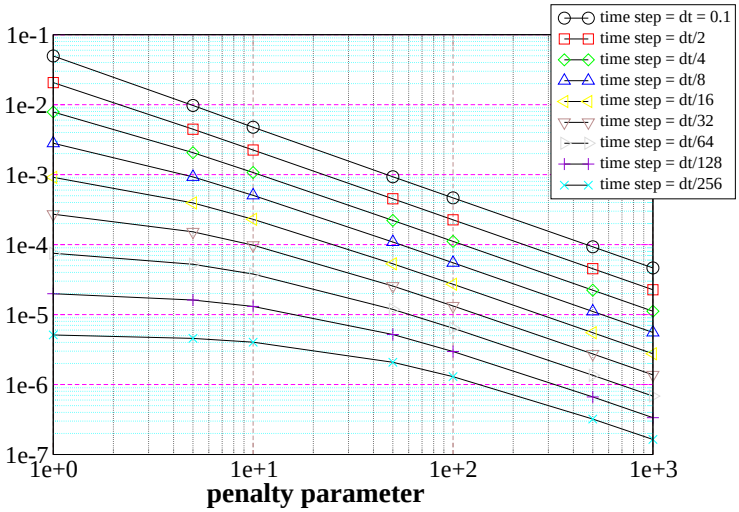


FIG. II.8 – Standard method - L^2 norm of ϵ at $t = 1$, as a function of the penalty parameter.

Bibliographie

[1] Robert A. Adams. *Sobolev Spaces.* Academic Press, 1975.

[2] Chérif Amrouche and Vivette Girault. On the existence and regularity of the solution of Stokes problem in arbitrary dimension. *Proc. Japan Acad., Série A*, 67 :171–175, 1991.

[3] David L. Brown, Ricardo Cortez, and Michael L. Minion. Accurate projection methods for the incompressible Navier-Stokes equations. *Journal of Computational Physics*, 168 :464–499, 2001.

[4] Jean-Paul Caltagirone and Jérôme Breil. Sur une méthode de projection vectorielle pour la résolution des équations de Navier-Stokes. *Comptes-Rendus de l'académie des Sciences, Paris – Série II*, 327 :1179–1184, 1999.

[5] Lamberto Cattabriga. Su un problema al contorno relativo al sistemo di equazioni di stokes. *Rend. Sem. Mat. Univ. Padova*, 31 :308–340, 1961.

[6] Alexandre Joel Chorin. Numerical solution of the Navier-Stokes equations. *Mathematics of Computation*, 22 :745–762, 1968.

[7] P. G. Ciarlet. *Handbook of Numerical Analysis Volume II : Finite Elements Methods – Basic Error Estimates for Elliptic Problems.* North-Holland, 1991.

[8] Alexandre Ern and Jean-Luc Guermond. *Eléments finis : théorie, applications, mise en œuvre*, volume 36 of *Mathématiques & Applications.* Springer, 2002.

[9] J.H. Ferziger and M. Perić. *Computational Methods for Fluid Dynamics.* Springer, third edition, 2002.

[10] M. Fortin and R. Glowinski. *Méthodes de Lagrangien Augmenté.* Dunod, Paris, 1982.

[11] Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms.*, volume 5 of *Springer Series in Computational Mathematics.* Springer-Verlag, 1986.

[12] Katuhiko Goda. A multistep technique with implicit difference schemes for calculating two- or three-dimensional cavity flows. *Journal of Computational Physics*, 30 :76–95, 1979.

[13] J.-L. Guermond and L. Quartapelle. On the approximation of the unsteady Navier-Stokes equations by finite element projection methods. *Numerische Mathematik*, 80 :207–238, 1998.

[14] Jean-Luc Guermond. Some implementations of projection methods for Navier-Stokes equations. *Mathematical Modelling and Numerical Analysis*, 30(5) :637–667, 1996.

[15] Jean-Luc Guermond. Un résultat de convergence d'ordre deux en temps pour l'approximation des équations de Navier-Stokes par une technique de projection

incrémentale. *Mathematical Modelling and Numerical Analysis*, 33(1) :169–189, 1999.

[16] J.L. Guermond and Jie Shen. On the error estimates for the rotational pressure-correction projection methods. *Mathematics of Computation*, 73(248) :1719–1737, 2003.

[17] John G. Heywood and Rolf Rannacher. Finite element approximation of the nonstationary Navier-Stokes problem. I. regularity of solutions and second-order error estimates for spatial discretization. *SIAM Journal on Numerical Analysis*, 19(2) :275–311, 1982.

[18] M. Jobelin, C. Lapuerta, J.-C. Latché, Ph Angot, and B. Piar. A finite element penalty-projection method for incompressible flows, 2005. Submitted to Journal of Computational Physics.

[19] Andreas Prohl. *Projection and Quasi-Compressibility Methods for Solving the Incompressible Navier-Stokes Equations*. Teubner, 1997.

[20] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, 1994.

[21] Jie Shen. On error estimates of projection methods for Navier-Stokes equations : First-order schemes. *SIAM Journal on Numerical Analysis*, 29(1) :57–77, 1992.

[22] Jie Shen. On error estimates of some higher order projection and penalty-projection methods for Navier-Stokes equations. *Numerische Mathematik*, 62 :49–73, 1992.

[23] Jie Shen. Remarks on the pressure estimates for the projection methods. *Numerische Mathematik*, 67 :513–520, 1994.

[24] Jie Shen. On error estimates of the penalty method for unsteady Navier-Stokes equations. *SIAM Journal on Numerical Analysis*, 32(2) :386–403, 1995.

[25] Jie Shen. On error estimates of projection methods for Navier-Stokes equations : Second-order schemes. *Mathematics of Computation*, 65(215) :1039–1065, 1996.

[26] R. Temam. Une méthode d'approximation de la solution des équations de Navier-Stokes. *Bull. Soc. Math. France*, 98 :115–152, 1968.

[27] R. Temam. Sur l'approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires II. *Arch. Rat. Mech. Anal.*, 33 :377–385, 1969.

[28] L.J.P. Timmermans, P.D. Minev, and F.N. Van de Vosse. An approximate projection scheme for incompressible flow using spectral elements. *International Journal for Numerical Methods in Fluids*, 22 :673–688, 1996.

Chapitre III

Une méthode élément fini de projection-pénalité pour des fluides incompressibles

The penalty-projection method for the solution of Navier-Stokes equations may be viewed as a projection scheme where an augmentation term is added in the first stage, namely the solution of the momentum balance equation, to constrain the divergence of the predicted velocity field. After a presentation of the scheme in the time semi-discrete formulation, then in fully discrete form for a finite element discretization, we assess its behaviour against a set of benchmark tests, including in particular prescribed velocity and open boundary conditions. The results demonstrate that the augmentation always produces beneficial effects. As soon as the augmentation parameter takes a significant value, the projection method splitting error is reduced, pressure boundary layers are suppressed and the loss of spatial convergence of the incremental projection scheme in case of open boundary conditions does not occur anymore. For high values of the augmentation parameter, the results of coupled solvers are recovered. Consequently, in contrast with standard penalty methods, there is no need for a dependence of the augmentation parameter with the time step, and this latter can be kept to reasonable values, to avoid to degrade too severely the conditionning of the linear operator associated to the velocity prediction step.

III.1 Introduction

Since the pioneering work of Chorin and Temam [5, 31] in the late sixties, pressure correction methods have gained a lot of popularity for the solution of transient incompressible flow problems. Indeed, schemes of this type have proved to be extremely efficient, essentially because, at each time step, they reduce the solution of a saddle point type problem to a cascade of decoupled elliptic equations for the velocity and the pressure. This feature makes them particularly attractive for industrial applications, as for instance nuclear safety studies, which are the context of this work.

The principle of pressure correction methods is to perform the advance in time in two steps : in a first stage, the momentum balance is solved to obtain an intermediate (or predicted) velocity field ; then this intermediate velocity is projected on a space of solenoidal vector fields. This process introduces a numerical error, often called

the "splitting error", which magnitude may be expected at least heuristically to be linked to the magnitude of the divergence of the predicted velocity. Hence the basic approach followed here is to constrain the divergence of the intermediate velocity field by adding in the first step of the scheme an augmentation term, of the same form as in Augmented Lagrangian methods [8]. This idea has been already suggested in the literature, first by Shen [26] and then, independently and with some additional variants, by Caltagirone and Breil [4]. In particular, a different projection step is proposed (called by the authors "vector projection step"), which is not used in the present work. In the paper by Shen, this novel family of numerical schemes was introduced as an improvement of both pressure correction and penalty methods (see [28] and references herein), and received the name of penalty-projection method. The numerical scheme presented here falls in this category, but is different from the original one, by changes in the algorithm itself, together with the fact that the augmentation parameter is totally independent from the time step Δt, whereas it was varying as Δt^{-2} in [26]. This point is discussed in the last remark of section III.2. To the best of our knowledge, no in-depth numerical study of any penalty-projection scheme is available in the literature ; this is the general purpose of the present paper, together with a derivation of the method, dealing with general boundary conditions and including some implementation issues.

From our experience of the use of Augmented Lagrangian methods in the finite element framework, it appears preferable to build the augmentation terms from the algebraic formulation of the discrete equations. This is due to the fact that the constraint *in the continuous sense* is usually not satisfied by the discrete solution : for the problem at hand, namely the solution of incompressible Navier-Stokes equations with a finite element method, the divergence constraint is only imposed in a weak sense. Consequently, an augmentation term built with the continuous expression of the constraint does not vanish for the solution of the discrete problem, and its presence may severely impact the results (see section III.3 for an explanation). However, to avoid an unnecessary restriction of the presentation to the finite element context, we choose to first carry out an introduction of the penalty-projection scheme in a space-continuous formulation. This is the goal of section III.2. The reader will keep in mind that, for practical implementation, the correct formulation of the method is given in the next section (section III.3), where the formulation of the scheme in the finite element context is detailed. In the last part of the paper, we compare the penalty-projection scheme to reference algorithms, namely the classical Euler semi-implicit method and frequently encountered projection schemes, for a set of benchmark tests involving a wide range of situations. In particular, besides prescribed velocity boundary conditions, we also consider the more challenging case (for pressure correction methods) of open boundary conditions.

III.2 The time semi-discrete algorithm

We are interested in the solution of the Navier-Stokes equations, governing an incompressible and isothermal flow of a Newtonian fluid, which read :

$$\varrho\Big[\frac{\partial u}{\partial t} + (u\cdot\nabla)u\Big] = -\nabla p + \nabla\cdot\tau(u) + \varrho g \qquad \text{in} \qquad [0,T]\times\Omega \qquad \text{(III.2.1a)}$$

$$\nabla\cdot u = 0 \qquad \text{in} \qquad [0,T]\times\Omega \qquad \text{(III.2.1b)}$$

$$u = u_{\partial\Omega_{\mathrm{D}}} \qquad \text{on} \qquad [0,T]\times\partial\Omega_{\mathrm{D}} \qquad \text{(III.2.1c)}$$

$$-pn + \tau(u)\cdot n = f_{\mathrm{N}} \qquad \text{on} \qquad [0,T]\times\partial\Omega_{\mathrm{N}} \qquad \text{(III.2.1d)}$$

$$u = u_0 \qquad \text{in} \qquad \{0\}\times\Omega \qquad \text{(III.2.1e)}$$

where u stands for the fluid velocity of initial value u_0, p for the pressure, g for the external body forces, and ϱ for the fluid density, supposed to be a positive constant. The computational domain Ω is an open bounded connected subset of $\mathbb{R}^d$ with $d = 2$ or $d = 3$. We suppose that the boundary $\partial\Omega$ of Ω is partitionned in two subsets $\partial\Omega_{\mathrm{D}}$ and $\partial\Omega_{\mathrm{N}}$ (with $\partial\Omega_{\mathrm{D}}\cap\partial\Omega_{\mathrm{N}} = \emptyset$), of outward normal vector n. On $\partial\Omega_{\mathrm{D}}$, the velocity is set to the value $u_{\partial\Omega_{\mathrm{D}}}$; the force per surface unit exerted at each point of the boundary $\partial\Omega_{\mathrm{N}}$ is given and equal to f_{N}. The tensor τ stands for the viscous part of the stress tensor, which divergence is given by one of the following expressions :

$$\nabla\cdot\tau(u) = \mu\Delta u \qquad \text{(III.2.2a)}$$

$$\text{or} \qquad \nabla\cdot\tau(u) = \nabla\cdot\mu(\nabla u + \nabla u^{\mathrm{T}}) \qquad \text{(III.2.2b)}$$

where μ stands for the dynamic viscosity of the fluid. As the equation (III.2.2a) is physically meaningful only if the parameter μ is a constant, this relation will be used only in this restrictive case in the rest of the paper.

Let $0 = t^0 < t^1 < \cdots < t^N = T$ be a partition of the time interval of computation $[0,T]$, which we suppose uniform for the sake of simplicity. We denote by Δt the time step, *i.e.* the constant difference between two successive times t^n and t^{n+1}, $0 \le n \le N-1$. A semi-implicit semi-discretization of the system of equations (III.2.1) with respect to the time variable is given by :

$$\varrho\Big[\frac{Du^{n+1}}{\Delta t} + (u^{\star,n+1}\cdot\nabla)u^{n+1}\Big] = -\nabla p^{n+1} + \nabla\cdot\tau(u^{n+1}) + \varrho g \quad \text{in } \Omega \qquad \text{(III.2.3a)}$$

$$\nabla\cdot u^{n+1} = 0 \qquad \text{in } \Omega \qquad \text{(III.2.3b)}$$

$$u^{n+1} = u^{n+1}_{\partial\Omega_{\mathrm{D}}} \qquad \text{on } \partial\Omega_{\mathrm{D}} \qquad \text{(III.2.3c)}$$

$$-p^{n+1}n + \tau(u^{n+1})\cdot n = f_{\mathrm{N}}^{n+1} \qquad \text{on } \partial\Omega_{\mathrm{N}} \qquad \text{(III.2.3d)}$$

where $u^{n+1}_{\partial\Omega_{\mathrm{D}}} = u_{\partial\Omega_{\mathrm{D}}}(t^{n+1})$, $f_{\mathrm{N}}^{n+1} = f_{\mathrm{N}}(t^{n+1})$, u^n and p^n stand for an approximation of respectively the velocity u and the pressure p at $t = t^n$, $u^{\star,n+1}$ is an extrapolation of the velocity at t^{n+1} and $\dfrac{Du^{n+1}}{\Delta t}$ provides an approximation of the time derivative of the velocity at t^{n+1} by a backward differentiation formula, which takes the general form :

$$Du^{n+1} \stackrel{\text{def}}{=} \beta_q u^{n+1} - \sum_{j=0}^{q-1}\beta_j u^{n-j}$$

In practice, the following choices are commonly used :

$$Du^{n+1} = u^{n+1} - u^n \qquad u^{\star,n+1} = u^n \qquad \text{(III.2.4)}$$

$$Du^{n+1} = \frac{3}{2}u^{n+1} - 2u^n + \frac{1}{2}u^{n-1} \qquad u^{\star,n+1} = 2u^n - u^{n-1} \qquad \text{(III.2.5)}$$

The first choice (III.2.4) leads to a first order scheme, the second one (III.2.5) to a formally second order numerical method.

Due to the saddle-point nature of the system (III.2.3), the solution is CPU-time consuming, which makes decoupled fractional step strategies attractive. The penalty-projection method belongs to this family of schemes. Classically, the first step consists in the solution of the momentum balance equation using a beginning-of-step value for the pressure. The specific feature of the proposed scheme lies in the addition at this stage of an augmentation term built from the divergence constraint. This yields the following elliptic problem for an intermediate velocity field at time t^{n+1}, $\tilde{u}^{n+1}$:

$$\varrho\Big[\frac{\beta_q\tilde{u}^{n+1}-\sum_{j=0}^{q-1}\beta_j u^{n-j}}{\Delta t}+(u^{\star,n+1}\cdot\nabla)\tilde{u}^{n+1}\Big] \qquad \text{in } \Omega \qquad \text{(III.2.6a)}$$
$$-r\nabla(\nabla\cdot\tilde{u}^{n+1})=-\nabla p^n+\nabla\cdot\tau(\tilde{u}^{n+1})+\varrho g$$

$$\tilde{u}^{n+1}=u^{n+1}_{\partial\Omega_{\mathrm{D}}} \qquad \text{on } \partial\Omega_{\mathrm{D}} \qquad \text{(III.2.6b)}$$

$$-p^n n+\tau(\tilde{u}^{n+1})\cdot n=f_{\mathrm{N}}^{n+1} \qquad \text{on } \partial\Omega_{\mathrm{N}} \qquad \text{(III.2.6c)}$$

where r is an augmentation parameter to be defined.

Let H be the following affine variety of divergence free vector fields :

$$\mathrm{H}=\{v\in[\mathrm{L}^2(\Omega)]^d,\ \nabla\cdot v=0,\ v\cdot n=u^{n+1}_{\partial\Omega_{\mathrm{D}}}\cdot n \text{ on } \partial\Omega_{\mathrm{D}}\}$$

The second step of a projection method is usually chosen to realize the L^2 orthogonal projection of $\tilde{u}^{n+1}$ onto H, which takes the general form :

$$\beta_q\varrho\,\frac{u^{n+1}-\tilde{u}^{n+1}}{\Delta t}+\nabla\phi=0 \qquad \text{in } \Omega \qquad \text{(III.2.7a)}$$

$$\nabla\cdot u^{n+1}=0 \qquad \text{in } \Omega \qquad \text{(III.2.7b)}$$

For computational efficiency reason, this Darcy system is then reformulated by taking the divergence of the first relation to obtain a Poisson problem for ϕ, which must be supplemented by boundary conditions on $\partial\Omega_{\mathrm{D}}$ and $\partial\Omega_{\mathrm{N}}$. The first one follows from the definition of H :

$$u^{n+1}\cdot n=\tilde{u}^{n+1}\cdot n=u^{n+1}_{\partial\Omega_{\mathrm{D}}}\cdot n \qquad \text{on } \partial\Omega_{\mathrm{D}}$$

and, consequently :

$$\nabla\phi\cdot n=0 \qquad \text{on } \partial\Omega_{\mathrm{D}}$$

The second one is derived from the L^2-orthogonality condition for the projection onto H, which reads :

$$\int_\Omega(u^{n+1}-\tilde{u}^{n+1})\cdot(u^{n+1}-v)=0 \qquad \forall v\in\mathrm{H}$$

By the relation (III.2.7a), integrating by parts and using the definition of H yields :

$$\begin{aligned}0&=\int_\Omega\nabla\phi\cdot(u^{n+1}-v)\\&=\int_{\partial\Omega}\phi\,(u^{n+1}-v)\cdot n-\int_\Omega\phi\,\nabla\cdot(u^{n+1}-v)\\&=\int_{\partial\Omega_{\mathrm{N}}}\phi\,(u^{n+1}-v)\cdot n \qquad\qquad \forall v\in\mathrm{H}\end{aligned}$$

which is satisfied if the following Dirichlet boundary condition for ϕ holds :

$$\phi = 0 \qquad \text{on } \partial\Omega_{\mathrm{N}}$$

Finally, ϕ is then the solution of the following elliptic problem :

$$\Delta\phi = \frac{\beta_q\ \varrho}{\Delta t}\ \nabla\cdot\tilde{u}^{n+1} \qquad \text{in } \Omega \tag{III.2.8a}$$

$$\nabla\phi\cdot n = 0 \qquad \text{on } \partial\Omega_{\mathrm{D}} \tag{III.2.8b}$$

$$\phi = 0 \qquad \text{on } \partial\Omega_{\mathrm{N}} \tag{III.2.8c}$$

The end-of-step velocity is computed afterwards by the relation (III.2.7a) :

$$u^{n+1} = \tilde{u}^{n+1} - \frac{\Delta t}{\beta_q\ \varrho}\ \nabla\phi \tag{III.2.9}$$

Finally, to obtain an expression for an approximation of the pressure at time t^{n+1}, we add the relations (III.2.6a) and (III.2.7a) to recover a discretization of the momentum balance equation at time t^{n+1} :

$$\varrho\Big[\frac{\beta_q u^{n+1} - \sum_{j=0}^{q-1}\beta_j u^{n-j}}{\Delta t} + (u^{\star,n+1}\cdot\nabla)\tilde{u}^{n+1}\Big] = -\nabla(p^n - r\nabla\cdot\tilde{u}^{n+1} + \phi) + \nabla\cdot\tau(\tilde{u}^{n+1}) + \varrho g \qquad \text{in } \Omega \tag{III.2.10}$$

This suggests the following expression for p^{n+1} :

$$p^{n+1} = p^n - r\nabla\cdot\tilde{u}^{n+1} + \phi \tag{III.2.11}$$

To summarize, the algorithm corresponding to one time step of the penalty-projection method consists in solving the sequence of equations (III.2.6) and (III.2.8) and computing the end-of step velocity and pressure by (III.2.9) and (III.2.11), respectively. For the reader's convenience, these four steps are gathered at the beginning of section III.4 (equation (III.4.22)), together with a synthetic presentation of some of the most popular projection schemes.

Remark [Construction of a rotational penalty-projection method]
Following ideas of [32] and [18], it is possible to build what is termed in [18] as a "rotational pressure correction method". We suppose temporarily that we are in the case where the viscosity is constant and $\nabla\cdot\tau(\tilde{u}^{n+1})$ is equal to $\mu\Delta\tilde{u}^{n+1}$. Using the identity $\Delta\tilde{u}^{n+1} = -\nabla\wedge(\nabla\wedge\tilde{u}^{n+1}) + \nabla(\nabla\cdot\tilde{u}^{n+1})$, equation (III.2.10) reads equivalently :

$$\varrho\Big[\frac{\beta_q u^{n+1} - \sum_{j=0}^{q-1}\beta_j u^{n-j}}{\Delta t} + (u^{\star,n+1}\cdot\nabla)\tilde{u}^{n+1}\Big] = -\nabla(p^n - r\nabla\cdot\tilde{u}^{n+1} + \phi) - \mu\nabla\wedge(\nabla\wedge\tilde{u}^{n+1}) + \mu\nabla(\nabla\cdot\tilde{u}^{n+1}) + \varrho g \tag{III.2.12}$$

Using the fact that the rotational of a gradient vanishes, the relation (III.2.7a) yields :

$$\nabla\wedge\tilde{u}^{n+1} = \nabla\wedge u^{n+1}$$

As the end-of-step velocity u^{n+1} is divergence free, we then get :

$$\nabla \wedge (\nabla \wedge \tilde{u}^{n+1}) = \nabla \wedge (\nabla \wedge u^{n+1}) = -\Delta u^{n+1}$$

and the time-discrete momentum balance equation finally reads :

$$\varrho\Big[\frac{\beta_q u^{n+1} - \sum_{j=0}^{q-1} \beta_j u^{n-j}}{\Delta t} + (u^{\star,n+1} \cdot \nabla)\tilde{u}^{n+1}\Big] = -\nabla(p^n - (r+\mu)\nabla \cdot \tilde{u}^{n+1} + \phi) + \mu\Delta u^{n+1} + \varrho g \quad \text{(III.2.13)}$$

We then have recovered an expression closer to the semi-implicit time discretization, with an additional term in the pressure increment :

$$p^{n+1} = p^n - (r+\mu)\nabla \cdot \tilde{u}^{n+1} + \phi$$

This variant will be named hereafter the rotational penalty-projection scheme.

Remark [From penalty methods to penalty-projection methods]
The penalty scheme for the Navier-Stokes equations takes, in its simplest version, the following form :

$$\varrho\Big[\frac{Du^{n+1}}{\Delta t} + (u^{\star,n+1} \cdot \nabla)u^{n+1}\Big] = -\nabla p^{n+1} + \nabla \cdot \tau(u^{n+1}) + \varrho g$$
$$\nabla \cdot u^{n+1} + \epsilon\, p^{n+1} = 0$$

where the positive penalty parameter ϵ is taken far smaller than 1. Since the pressure can be explicitely expressed as a function of the velocity divergence by the second equation, it can be eliminated from the system to yield :

$$\varrho\Big[\frac{Du^{n+1}}{\Delta t} + (u^{\star,n+1} \cdot \nabla)u^{n+1}\Big] - \frac{1}{\epsilon}\nabla(\nabla \cdot u^{n+1}) = \nabla \cdot \tau(u^{n+1}) + \varrho g \quad \text{(III.2.14)}$$

which introduces an augmentation-like term (with $r = 1/\epsilon$) in the momentum balance equation.

Remark [Comparison of the present scheme and the penalty-projection method proposed by Shen [26]]
As quoted in the introduction, the first penalty-projection scheme has been introduced by Shen in 92 [26, section 6]. The main difference between this scheme and the method presented here lies in the integration of the pressure in the algorithm. Indeed, J. Shen builds an auxiliary sequence ψ^n, $0 \leq n \leq N-1$, by the following induction relation :

$$\psi^{n+1} = \psi^n + \phi$$

where ϕ is given by the system (III.2.7). The quantity ψ^n is used in the prediction step (III.2.6) instead of p^n, and an approximation of the pressure is computed only as a "post-processing" of the outcomes of the computation, by a two possible relations in the spirit of (III.2.11), the most simple one reading :

$$p^{n+1} = \psi^{n+1} - r\nabla \cdot \tilde{u}^{n+1}$$

In addition, in [26], the augmentation parameter is varying with the time step, and the author shows that a second order Crank-Nicholson discretization of the momentum balance equation combined with an adjustment of r as $r \simeq \Delta t^{-2}$ leads to a second order scheme. We adopt here a different point of view, in trying to assess the properties of the method when the augmentation parameter is kept constant (and within reasonable values), to avoid to excessively degrade the conditionning of the linear system associated to the velocity prediction step.

III.3 A finite element implementation

We now turn to the finite element implementation of the penalty-projection method. As mentioned in the introduction, due to the fact that the discrete solution is only "discretely-divergence-free", it seems preferable to build the augmentation term at the algebraic level, to avoid a non-intentional perturbation of the solution growing with the augmentation parameter. To this purpose, we must proceed as follows : first, we discretize in space the time-discrete equations of the scheme keeping them free of any augmentation term, then this latter is added afterwards.

Let a finite element discretization of the velocity and pressure be given, or, equivalently, let two Lagrange finite element spaces V_h and M_h included respectively in $[\mathrm{H}^1(\Omega)]^d$ and $\mathrm{L}^2(\Omega)$ be chosen. In addition, the pressure approximation space is required to be included in $\mathrm{H}^1(\Omega)$. We note :

$$V_h = \mathrm{span}\{\varphi_i^u,\ 1 \leq i \leq N_{\mathrm{dof}}^u\}$$
$$M_h = \mathrm{span}\{\varphi_i^p,\ 1 \leq i \leq N_{\mathrm{dof}}^p\}$$

where the φ_i^u and φ_i^p are respectively vector and scalar functions defined on Ω. To simplify the subsequent notations, we suppose that, for the velocity approximation, the set of the indexes of the nodes located on $\partial\Omega_{\mathrm{D}}$ is the set of integer numbers i such that $N_{\mathrm{unk}}^u + 1 \leq i \leq N_{\mathrm{dof}}^u$. Consequently, we have :

$$V_h^{\partial\Omega_{\mathrm{D}}} = \mathrm{span}\{\varphi_i^u,\ 1 \leq i \leq N_{\mathrm{unk}}^u\} \subset \{v \in [\mathrm{H}^1(\Omega)]^d,\ v = 0 \text{ on } \partial\Omega_{\mathrm{D}}\}$$

We note u_{D}^{n+1} a function lying in the sub-space of V_h spanned by the basis functions φ_i^u, $N_{\mathrm{unk}}^u + 1 \leq i \leq N_{\mathrm{dof}}^u$ and interpolating the boundary condition $u_{\partial\Omega_{\mathrm{D}}}^{n+1}$ on $\partial\Omega_{\mathrm{D}}$, in a sense to be defined according to the regularity of $u_{\partial\Omega_{\mathrm{D}}}^{n+1}$ (in order for the problem to be well posed, this function must belong to $\mathrm{H}^{1/2}(\partial\Omega_{\mathrm{D}})$, in which case at least a quasi-interpolation operator *à la* Clément is well defined [6]). This function u_{D}^{n+1} may be viewed as a discrete lifting of the boundary condition.

At the time step t^n, the discrete solution u_h^n may then be decomposed as :

$$u_h^n = u_{\mathrm{D}}^n + u_{\mathrm{F}}^n$$

where u_{F}^n lies in the discrete space $V_h^{\partial\Omega_{\mathrm{D}}}$ and is the unknown velocity field of the problem.

The variational formulation for a time step of the semi-implicit scheme (III.2.3) reads :

Find $u_{\mathrm{F}}^{n+1} \in V_h^{\partial\Omega_{\mathrm{D}}}$ and $p_h^{n+1} \in M_h$ such that, $\forall v \in V_h^{\partial\Omega_{\mathrm{D}}}$, $\forall q \in M_h$:

$$\left\{\begin{array}{l} \displaystyle\int_\Omega \varrho\Big[\frac{\beta_q(u_{\mathrm{F}}^{n+1} + u_{\mathrm{D}}^{n+1}) - \sum_{j=0}^{q-1}\beta_j u_h^{n-j}}{\Delta t} + (u_h^{\star,n+1}\cdot\nabla)(u_{\mathrm{F}}^{n+1} + u_{\mathrm{D}}^{n+1})\Big]\cdot v \\ \qquad\qquad + \displaystyle\int_\Omega \tau(u_{\mathrm{F}}^{n+1} + u_{\mathrm{D}}^{n+1}) : \nabla v + \int_\Omega \nabla p_h^{n+1}\cdot v = \int_\Omega \varrho g\cdot v + \int_{\partial\Omega_{\mathrm{N}}} f_{\mathrm{N}}\cdot v \\ \\ \displaystyle -\int_\Omega q\ \nabla\cdot(u_{\mathrm{F}}^{n+1} + u_{\mathrm{D}}^{n+1}) = 0 \end{array}\right. \tag{III.3.15}$$

This discrete variational formulation is routinely converted into the following algebraic system :

$$\left\{\begin{array}{r} \dfrac{\beta_q}{\Delta t}\mathbf{M}\mathbf{U}_{\mathrm{F}} + \mathbf{A}\mathbf{U}_{\mathrm{F}} + \mathbf{B}^{\mathrm{T}}\mathbf{P} = \mathbf{F} \\ \mathbf{B}\mathbf{U}_{\mathrm{F}} = \mathbf{G} \end{array}\right. \tag{III.3.16}$$

where $\mathbf{U}_{\mathrm{F}}$ and $\mathbf{P}$ are vectors of respectively $\mathbb{R}^{N^u_{\mathrm{unk}}}$ and $\mathbb{R}^{N^p_{\mathrm{dof}}}$ gathering the degrees of freedom of respectively u_{F}^{n+1} and p_h^{n+1}, and the discrete operators and right-hand members are defined as follows :

$$\begin{aligned}
\mathbf{M}_{ij} &= \int_\Omega \varrho\ \varphi_j^u \cdot \varphi_i^u && 1 \le i,j \le N^u_{\mathrm{unk}} \\
\mathbf{A}_{ij} &= \int_\Omega \left[\varrho\ (u_h^{\star,n+1}\cdot\nabla)\varphi_j^u\right]\cdot\varphi_i^u + \int_\Omega \tau(\varphi_j^u) : \nabla\varphi_i^u && 1 \le i,j \le N^u_{\mathrm{unk}} \\
\mathbf{B}_{ij} &= -\int_\Omega \varphi_i^p\ \nabla\cdot\varphi_j^u && 1 \le i < N^p_{\mathrm{dof}} \\
& && 1 \le j \le N^u_{\mathrm{unk}} \\
\mathbf{F}_i &= -\int_\Omega \varrho\Big[\frac{\beta_q u_{\mathrm{D}}^{n+1} - \sum_{j=0}^{q-1}\beta_j u_h^{n-j}}{\Delta t} + (u_h^{\star,n+1}\cdot\nabla)u_{\mathrm{D}}^{n+1}\Big]\cdot\varphi_i^u \\
&\quad - \int_\Omega \tau(u_{\mathrm{D}}^{n+1}) : \nabla\varphi_i^u + \int_\Omega \varrho g\cdot\varphi_i^u + \int_{\partial\Omega_{\mathrm{N}}} f_{\mathrm{N}}\cdot\varphi_i^u && 1 \le i \le N^u_{\mathrm{unk}} \\
\mathbf{G}_i &= \int_\Omega \varphi_i^p\ \nabla\cdot u_{\mathrm{D}}^{n+1} && 1 \le i \le N^p_{\mathrm{dof}}
\end{aligned}$$

The augmentation term is built by pre-multiplying the discrete divergence constraint by a scaling (symmetrical definite positive) matrix, denoted $\mathbf{M}_{\mathbf{p}}^{-1}$ because a typical choice for this operator is the inverse of the lumped pressure mass matrix, then by the matrix $\mathbf{B}^{\mathrm{T}}$ to obtain a positive symmetrical operator :

$$\text{augmentation term} = r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}\mathbf{U}_{\mathrm{F}} - \mathbf{G})$$

where r is the (positive and constant) augmentation parameter. As in the time semi-discrete case, the first step of the penalty-projection method is obtained from the momentum balance equation by taking the pressure at the beginning of the time step and adding the augmentation term :

$$\frac{\beta_q}{\Delta t}\mathbf{M}\widetilde{\mathbf{U}}_{\mathrm{F}} + \mathbf{A}\widetilde{\mathbf{U}}_{\mathrm{F}} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathrm{exp}} = \mathbf{F} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{G} \tag{III.3.17}$$

where $\widetilde{\mathbf{U}}_{\mathrm{F}}$ and $\mathbf{P}_{\mathrm{exp}}$ are vectors of respectively $\mathbb{R}^{N^u_{\mathrm{unk}}}$ and $\mathbb{R}^{N^p_{\mathrm{dof}}}$ gathering the degrees of freedom of respectively the predicted velocity and p_h^n.

The algebraic analogue of the projection step (III.2.7) reads :

$$\left\{\begin{array}{l} \dfrac{\beta_q}{\Delta t}\mathbf{M}(\mathbf{U}_{\mathrm{F}} - \widetilde{\mathbf{U}}_{\mathrm{F}}) + \mathbf{B}^{\mathrm{T}}\mathbf{\Phi} = 0 \\ \mathbf{B}\mathbf{U}_{\mathrm{F}} = \mathbf{G} \end{array}\right. \tag{III.3.18}$$

Solving the first relation for $\mathbf{U}_{\mathrm{F}}$ and substituting the obtained expression in the second equation yields :

$$\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{\Phi} = \frac{\beta_q}{\Delta t}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G})$$

As, in the general case, the velocity mass matrix is not diagonal, handling the operator $\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^{\mathrm{T}}$ appears to be rather unefficient. Relation (III.2.8) suggests to replace it by :

$$\mathbf{L}\mathbf{\Phi} = \frac{\beta_q\,\varrho}{\Delta t}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \tag{III.3.19}$$

where $\mathbf{L}$ is the discrete operator associated to a Poisson problem with Dirichlet boundary conditions on $\partial\Omega_{\mathrm{N}}$ and natural Neumann conditions on $\partial\Omega_{\mathrm{D}}$. To avoid some complexity in the assembling process (for instance, eliminating in the linear systems the degrees of freedom of ϕ associated with the pressure nodes located on $\partial\Omega_{\mathrm{N}}$ would lead to vectors of different dimension for $\mathbf{\Phi}$ and $\mathbf{P}$), the Dirichlet boundary conditions may be simply imposed by penalization of a Robin boundary condition such that :

$$\mathbf{L}_{ij} = \int_{\Omega} \nabla\varphi_j^p \cdot \nabla\varphi_i^p + \frac{1}{\epsilon}\int_{\partial\Omega_{\mathrm{N}}} \varphi_j^p\varphi_i^p \qquad 1 \leq i,j \leq N_{\mathrm{dof}}^p$$

with a coefficient ϵ such as $\epsilon \ll 1$.

Finally, the end of step pressure is obtained by following the same technique than in the time semi-discrete case, that is building a discrete momentum equation by adding relation (III.3.17) to the first equation of (III.3.18) and comparing to the semi-implicit discretization of the momentum balance (III.3.16). This process yields :

$$\mathbf{P} = \mathbf{P}_{\mathrm{exp}} + \mathbf{\Phi} + r\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \tag{III.3.20}$$

To sum up the developments of this section, the algorithm for performing a time step of the penalty-projection projection method is obtained by gathering equations (III.3.17), (III.3.19), the first relation of (III.3.18) and (III.3.20) :

$$\begin{cases} \big[\dfrac{\beta_q}{\Delta t}\mathbf{M} + \mathbf{A} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{B}\big]\widetilde{\mathbf{U}}_{\mathrm{F}} + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathrm{exp}} = \mathbf{F} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{G} \\ \mathbf{L}\mathbf{\Phi} = \dfrac{\beta_q\,\varrho}{\Delta t}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \\ \mathbf{U}_{\mathrm{F}} = \widetilde{\mathbf{U}}_{\mathrm{F}} - \dfrac{\Delta t}{\beta_q}\mathbf{M}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{\Phi} \\ \mathbf{P} = \mathbf{P}_{\mathrm{exp}} + \mathbf{\Phi} + r\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}\widetilde{\mathbf{U}}_{\mathrm{F}} - \mathbf{G}) \end{cases} \tag{III.3.21}$$

Remark [On the derivation of the augmentation term]
A direct finite element space-discretization of the time semi-discrete equations of the preceding section leads to an augmentation term which appears in the variational form of the momentum balance equation as :

$$r\int_{\Omega} \nabla\cdot u_h\ \nabla\cdot v$$

As the divergence of the velocity test functions v does not lie in the pressure discrete space, this term does not vanish by the continuity equation for the discrete solution. Consequently, using this formulation leads to a dependence of the solution on the augmentation parameter. For the particular discretization used here (Taylor-Hood element), an excessive smearing of the solution is then observed at high values of r.

III.4 Numerical experiments

The objective of the present section is to perform a comparison between the projection method presented here and various pressure correction schemes widely used in the literature for the solution of unstationnary incompressible flow problems. In the time-discrete setting (which, as explained in the preceding section, is somewhat incorrect for the augmentation terms), the methods considered here can be recast under the form of a four stages fractional step scheme as follows :

$$\left|\begin{array}{l} \varrho\big[\dfrac{\beta_q\tilde{u}^{n+1}-\sum_{j=0}^{q-1}\beta_j u^{n-j}}{\Delta t}+(u^{\star,n+1}\cdot\nabla)\tilde{u}^{n+1}\big] \\ \quad -r_1\nabla(\nabla\cdot\tilde{u}^{n+1})=-\nabla p^n+\nabla\cdot\tau(\tilde{u}^{n+1})+\varrho g \qquad \text{in } \Omega \\ \tilde{u}^{n+1}=u^{n+1}_{\partial\Omega_{\rm D}} \qquad \text{on } \partial\Omega_{\rm D} \\ -p^n n+\tau(\tilde{u}^{n+1})\cdot n=f_{\rm N}^{n+1} \qquad \text{on } \partial\Omega_{\rm N} \end{array}\right. \tag{III.4.22a}$$

$$\left|\begin{array}{l} \Delta\phi=\dfrac{\beta_q\ \varrho}{\Delta t}\ \nabla\cdot\tilde{u}^{n+1} \qquad \text{in } \Omega \\ \nabla\phi\cdot n=0 \qquad \text{on } \partial\Omega_{\rm D} \\ \phi=0 \qquad \text{on } \partial\Omega_{\rm N} \end{array}\right. \tag{III.4.22b}$$

$$u^{n+1}=\tilde{u}^{n+1}-\frac{\Delta t}{\beta_q\ \varrho}\ \nabla\phi \tag{III.4.22c}$$

$$p^{n+1}=p^n+\alpha\phi-r_2\nabla\cdot\tilde{u}^{n+1} \tag{III.4.22d}$$

where r_1, r_2 and α are parameters to be specified.

The penalty-projection method is obtained by the choice of parameters $r_1=r_2=r$, $\alpha=1$. For the rotational penalty-projection variant, values of the parameters are $r_1=r$, $r_2=r+\mu$, $\alpha=1$.

The case $r_1=r_2=0$ and $\alpha=1$ corresponds to the so-called incremental projection method, which is probably today the most popular one. To the best of our knowledge, the origin of this scheme can be traced back to Goda [10], and, since that time, this method has received a considerable attention : see in particular [34] for the extension to (formally) second order in time, [25, 27, 29] for the analysis in the time semi-discrete case and [14, 13, 15, 16] for the analysis in the fully discrete case. The main identified drawback of this scheme is that, as the parameter r_2 is zero, the pressure inherits the boundary conditions of ϕ, *i.e.* spurious homogeneous Neumann boundary conditions on no-slip boundaries and homogeneous Dirichlet conditions on open boundaries.

The choice $r_1=0$, $r_2=\mu$ and $\alpha=1$ leads to the method introduced in [32] and analysed in [18, 16]. Following these last two papers, we will refer it as the rotational pressure correction scheme.

As detailed at the end of section III.2, making the particular choice $r_1=\Delta t^{-2}$ and $r_2=0$ yields a method which will give the same results for the velocity as the scheme studied in [26, section 6]. In this case, a post-processing will be necessary to recover an accurate approximation of the pressure.

Finally, taking $r_1=r_2=r$ and $\alpha=0$, one recovers a scheme which presents some similarities with the so-called vectorial projection method presented in the finite volume framework in [4] (see [1] for a discussion on this topic). The reason why

the parameter α is set to zero in the latter work is that the authors make use of a different projection step, which does not provide a natural pressure increment. On the opposite, with the projection step employed here, to include the variable ϕ in the pressure correction seems a reasonable option, and the case $\alpha = 0$ is not considered here.

Table III.1 gathers the choices of parameters corresponding to the various methods considered in numerical experiments. Of course, this choice of schemes does not pretend to be exhaustive (see e.g. [21, 2] for alternative pressure correction schemes).

Incremental projection method (standard form)	$r_1 = 0$, $r_2 = 0$, $\alpha = 1$
Incremental projection method (rotational form)	$r_1 = 0$, $r_2 = \mu$, $\alpha = 1$
Penalty-projection method (standard form)	$r_1 = r$, $r_2 = r$, $\alpha = 1$
Penalty-projection method (rotational form)	$r_1 = r$, $r_2 = r + \mu$, $\alpha = 1$

TAB. III.1 – Choice of parameters in (III.4.22) for the schemes used in the numerical experiments

The present section is organized as follows. First, we examine the accuracy of the method on a standard Navier-Stokes benchmark, namely the computation of Taylor-Green vortices. In a second time, we focus on the behaviour of the pressure near the boundaries for a Stokes flow with Dirichlet boundary conditions. Third, we deal with a case with open boundary conditions. Finally, we turn to a more complex situation, the two-dimensional flow past a cylinder.

All the simulations presented here are performed with a formally second order in time scheme, corresponding to equation (III.2.5), *i.e.* a second order Backward Differentiation Formula for the approximation of the time derivative and a second order Richardson extrapolation for the estimation of the advection field. This scheme is initialized with a first time step performed with a standard backward Euler differentiation (relations (III.2.4)). The pressure at $t = 0$ is obtained by interpolation of the solution when available (*i.e.* in the first three cases) or set to zero. The problem is discretized in space using $\mathbf{P}_2$-P_1 finite elements (the so-called Taylor-Hood mixed finite element, see e.g. [9], [7, chap. 5], [23, chap. 9]). The solution of the coupled system obtained when using the semi-implicit scheme is performed by a standard augmented Lagrangian technique (see [8]).

The practical implementation has been performed using the software object-oriented component library PELICANS, developed at IRSN ([22]).

III.4.1 Taylor-Green vortices

As a first benchmark for the proposed numerical method, we use a particular periodic flow widely used in the literature (e.g. [21]), which enjoys the property to obey the Navier-Stokes equations with a zero forcing term. It is a particular case of a class of flows which seems to have been first introduced by Taylor and Green [30]. The

velocity and pressure fields read :

$$u(x,y,t) = \begin{bmatrix} -\cos(2\pi x)\,\sin(2\pi y) \\ \sin(2\pi x)\,\cos(2\pi y) \end{bmatrix} \exp(-8\pi^2\mu t)$$

$$p(x,y,t) = -\frac{(\cos(4\pi x) + \cos(4\pi y))}{4} \exp(-16\pi^2\mu t)$$

The chosen computational domain is $\Omega = [1/8, 5/8] \times [1/8, 5/8]$ and the velocity is imposed on the whole boundary. Of course, initial and boundary conditions are taken in such a way to match the analytical solution. The density is set to $\varrho = 1$ and the viscosity to $\mu = 0.01$.

The meshes used in this study are obtained by cutting the domain in an $n \times n$ regular grid and splitting each squared cell along its diagonals to obtain four simplicial elements.

Figures III.1, III.2 and III.3 show the difference between the numerical solution and the analytical one at $t = 1$, measured in $[\mathrm{L}^2(\Omega)]^d$ and $[\mathrm{H}^1(\Omega)]^d$ norms for the velocity and in $\mathrm{L}^2(\Omega)$ norm for the pressure. These curves are drawn for the 80×80 mesh.

All the curves show an error decrease with the time step for large values of this latter, then a plateau is observed, which corresponds to the space discretization error. Values of the error on the plateau are gathered, for various meshes, in table III.2, together with the interpolation error (difference between the analytical solution and the element of the approximation space obtained by interpolating the solution at each node). One can check than an optimal spatial convergence is found in each case.

	20×20	40×40	80×80
$\|u - u_h\|_{[\mathrm{L}^2(\Omega)]^d}$	$1.09\ 10^{-4}$	$1.34\ 10^{-5}$	$1.66\ 10^{-6}$
$\|u - u_i\|_{[\mathrm{L}^2(\Omega)]^d}$	$2.07\ 10^{-5}$	$2.59\ 10^{-6}$	$3.23\ 10^{-7}$
$\|u - u_h\|_{[\mathrm{H}^1(\Omega)]^d}$	$3.75\ 10^{-3}$	$9.34\ 10^{-4}$	$2.32\ 10^{-4}$
$\|u - u_i\|_{[\mathrm{H}^1(\Omega)]^d}$	$9.13\ 10^{-4}$	$2.28\ 10^{-4}$	$5.71\ 10^{-5}$
$\|p - p_h\|_{\mathrm{L}^2(\Omega)}$	$2.86\ 10^{-3}$	$7.14\ 10^{-4}$	$1.78\ 10^{-4}$
$\|p - p_i\|_{\mathrm{L}^2(\Omega)}$	$5.04\ 10^{-3}$	$1.26\ 10^{-3}$	$3.15\ 10^{-4}$

TAB. III.2 – Difference between the exact solution and the numerical solution on the time-convergence plateau and between the exact solution and its interpolate, at $t = 1$.

As far as the time convergence is concerned, the first observation is the outstanding importance of the splitting error associated to the standard projection schemes at large time steps; for the incremental projection method, for instance, the error is about two decades higher than for the semi-implicit scheme, and this ratio even increases for the pressure when the time step decreases, since these schemes does not exhibit the same convergence rate.

The rates of convergence globally agree with theoretical studies. For the velocity, a second order convergence is observed in the $[\mathrm{L}^2(\Omega)]^d$ norm for the semi-implicit scheme (see [19] for an in-depth study of the Crank-Nicholson time stepping method), the incremental projection ([15]) and the rotational projection method ([18]). The pressure is first order accurate for the incremental projection method and second order accurate for the rotational projection scheme and the semi-implicit scheme,

which is slightly better than theoretical bounds : a 3/2 order convergence for the rotational projection scheme is proven in [18] (and seems to be optimal for general computational domains), a second order convergence for the pressure in a weaker norm than the $L^2(\Omega)$ one is given in [24] (see remark below).

Concerning the penalty-projection scheme, the first point is that any penalization leads to a decrease of the error. Keeping a constant value of the augmentation parameter r when the time-step decreases, the second order in time convergence is lost, but the error remains bounded from above by the incremental projection one. Second, as can be inferred from the literature of penalty methods [28], the splitting error vanishes when r raises. However, keeping this parameter within relatively moderate bounds seems to be sufficient ($r = 100$ for the greater value, having in mind the presence of the inverse of the pressure mass matrix as scaling operator in the definition of the penalty term). Finally, the results compare with the rotational projection scheme as follows. For the velocity, the results with $r = \mu$ are superimposed (the corresponding curves for the penalty-projection scheme have been omitted on figures III.1 and III.2) ; for this test, where the viscosity is rather small ($\mu = 0.01$), the difference with the incremental projection scheme is hardly visible. For the pressure, the rotational scheme is clearly better than the penalty-projection scheme with $r = \mu$, but the opposite conclusion holds for larger values of the augmentation parameter. The same behaviour has been confirmed in all the experiments presented in this paper.

Remark : [Convergence of the incremental projection scheme]

1) The evolution of the pressure error as a function of the time step for the incremental projection scheme shows a linear decrease at large time steps, then, for smaller values of the time step, the convergence seems to become faster. As pointed out in [17, section 7], this phenomenon has the following explanation. Higher order convergence results for the pressure are known to hold in weaker norms ([13, section 4.4]) than the $L^2(\Omega)$ one. As in finite dimensional spaces, all the norms are equivalent, the $L^2(\Omega)$ norm of the error is bounded by the weaker norm multiplied by a coefficient depending on the meshing. For large values of the time step, the L^2 estimate yields the sharpest bound and a first order convergence rate is observed ; then, when the time step falls below a threshold value (decreasing with the mesh size), the bound in the weaker norm becomes sharper and a faster decrease of the error occurs.

2) The pressure convergence is known to be limited by the prescription of spurious boundary conditions to the pressure increment in the projection step. This can be checked here by the following simple numerical experiment. Let us solve the same problem on a different computational domain $\Omega = [1/4, 5/4] \times [1/4, 5/4]$, such that (by chance) the exact pressure satisfies homogeneous Neumann boundary conditions over the whole boundary. Then a second order convergence rate is recovered.

III.4.2 A Stokes flow with Dirichlet boundary conditions

The incremental projection method is known to be plagued by the presence of boundary layers in the pressure solution, which results from artificial homegeneous Neu-

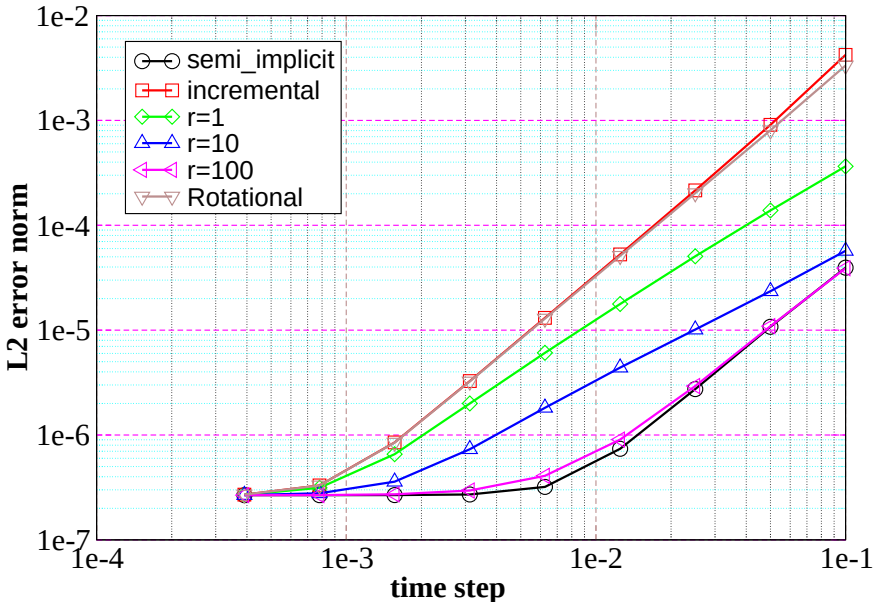


FIG. III.1 – Taylor-Green vortex - L^2 norm of the error for the velocity as a function of the time step, for the incremental, rotational, penalty projection ($r = 1$, $r = 10$, $r = 100$) schemes and for the semi-implicit scheme

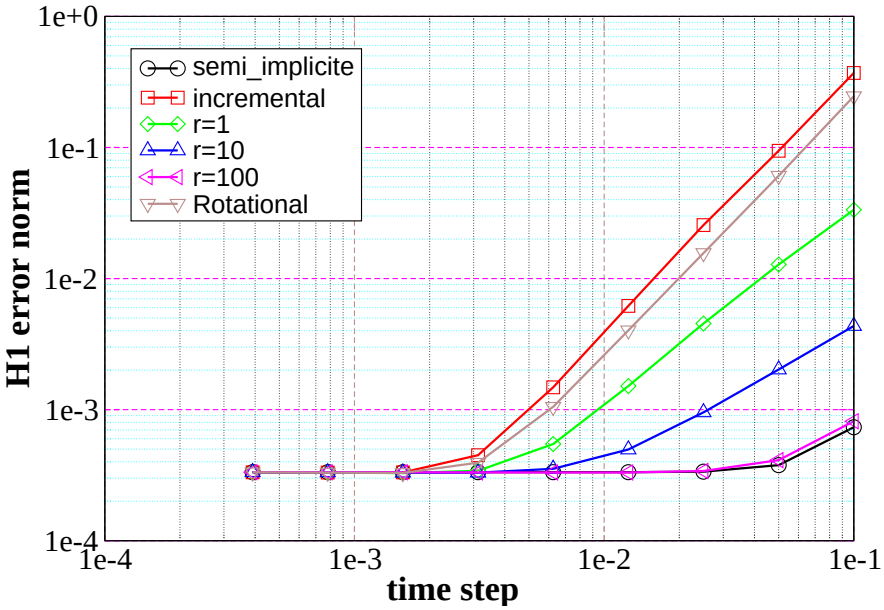


FIG. III.2 – Taylor-Green vortex - H^1 norm of the error for the velocity as a function of the time step, for the incremental, rotational, penalty-projection ($r = 1$, $r = 10$, $r = 100$) schemes and for the semi-implicit scheme

mann boundary conditions prescribed to the pressure increment on the part of the boundary where the velocity is prescribed. This phenomenon is highlighted in the setting used here, where a discrete pressure Laplacian operator with Neumann boundary conditions is explicitly built ; however, similar problems also affect projection methods derived via an inexact factorization of the algebraic discrete equations (e.g. [11, 12, 23, 33] for a presentation of such methods, [17, section 6] and [33, pp. 53-54], [23] for a discussion on this topic). The aim of this section is to assess the behaviour

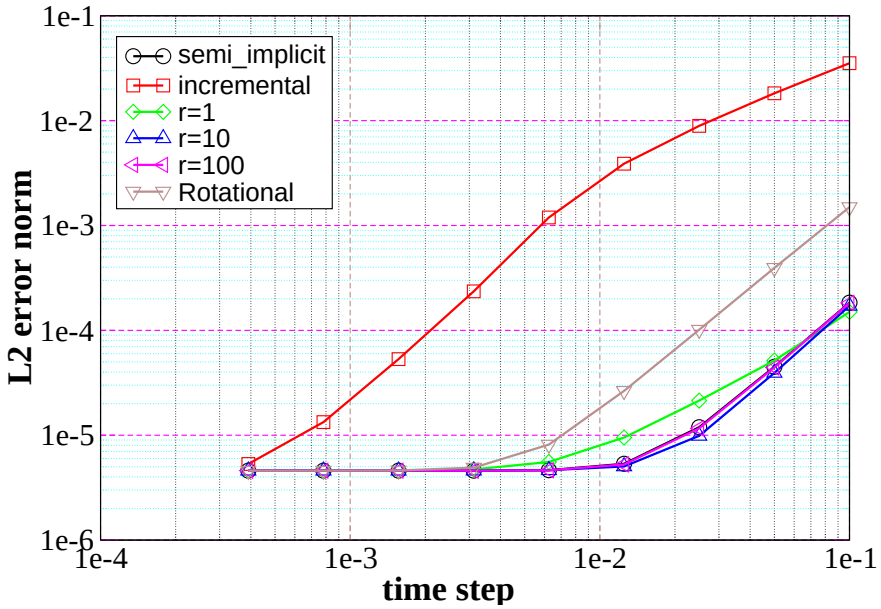


FIG. III.3 – Taylor-Green vortex - L^2 norm of the error for the pressure as a function of the time step, for the incremental, rotational, penalty-projection ($r = 1$, $r = 10$, $r = 100$) schemes and for the semi-implicit scheme

of the penalty-projection scheme with respect to this issue.

As the behaviour of the pressure error in the L^∞ norm is strongly affected by the regularity of the domain [18, 17], we chose for Ω a circle, of diameter 1. We only solve an unstationary Stokes problem (with $\mu = 1$, $\nabla \cdot \tau(u) = \mu \Delta u$), and we adjust the forcing term in such a way that the velocity and pressure fields read :

$$u(x, y, t) = \begin{bmatrix} \sin(x + t) & \sin(y + t) \\ \cos(x + t) & \cos(y + t) \end{bmatrix}$$

$$p(x, y, t) = \cos(x - y + t)$$

This test case is the same as studied in [18, 17], which allows a verification of our computations, as methods considered in these papers are part of the set of schemes tested in the present work.

We use for this test an unstructured mesh of 3600 elements, represented on figure III.4.

The distribution of the pressure error for the considered schemes at $t = 1$ is shown on figure III.5. The time step used for these computations is $\Delta t = 0.0125$.

As expected, a boundary layer is observed for the pressure error with the incremental scheme. This error measured in the L^∞ norm is, for this scheme, almost two orders of magnitude greater than for the implicit method.

This phenomenon disappears for the rotational variants of both the incremental projection scheme and the penalty-projection scheme, and for the standard penalty-projection method for high values of the augmentation parameter. The spatial distribution of the pressure error then becomes strongly influenced by the geometrical structure of the meshing. In particular, peaks appear along a diameter which has been used as a symmetry axis for the mesh generation.

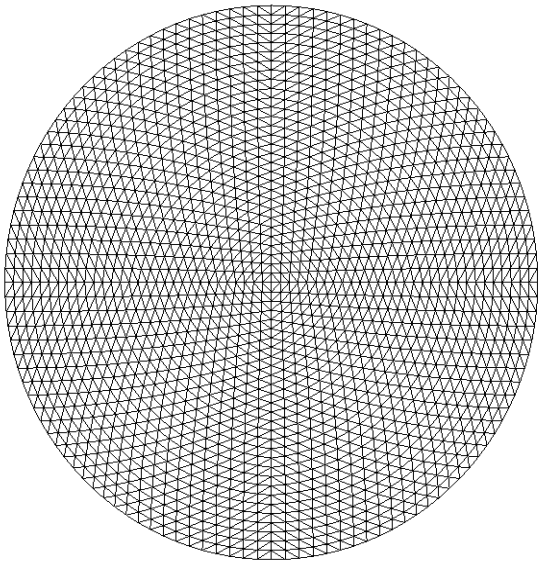


FIG. III.4 – Stokes flow with prescribed velocity boundary conditions - Meshing

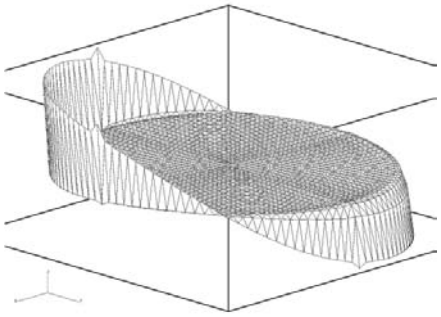
For the rotational projection scheme, the error is still ten times higher than for the implicit scheme. Equivalent results are obtained with the penalty-projection scheme for r ranging between 1 and 10, then, for higher values of r, the error becomes smaller.

III.4.3 A Stokes flow with open boundary conditions

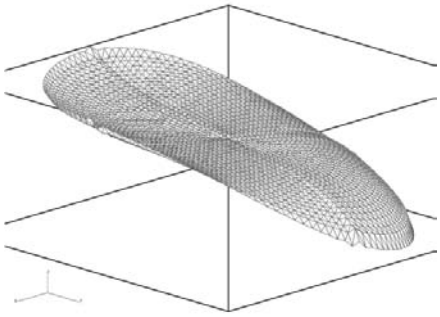
If the prescription of spurious Neumann boundary conditions for the pressure on Dirichlet parts of the boundary is an unquestionable drawback of the incremental projection method, things even go worse when open boundaries are to be dealt with. Indeed, homogeneous Dirichlet boundary conditions have to be artificially enforced to the pressure increment on that part of the boundary. This is done either explicitly (when a pressure Poisson problem is built at the continuous level) or implicitly (within the algebraic splitting approach, see [17, section 6] and [33, pp. 53-54] for a related discussion). Since the pressure approximation space is consequently modified, one may fear to lose even the spatial accuracy of the scheme ; both theoretical and experimental evidences that this indeed occurs are given in [16].

As before, the penalty-projection scheme and the rotational projection scheme share the property that the spurious boundary conditions are imposed only to the intermediate variable ϕ and not (by induction) to the pressure itself. Hence we can hope that these schemes enjoy significantly better convergence properties. The goal of the present section is to check this issue.

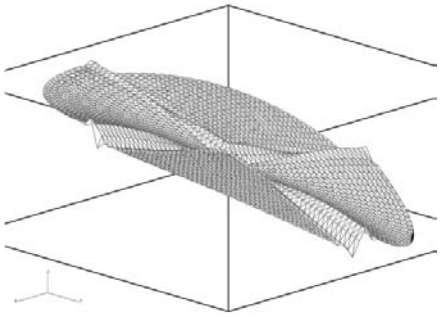
As in the preceding section, we choose a test case already used in the literature [16, 17]. It consists in an unstationary Stokes problem (with $\mu = 1$, $\nabla \cdot \tau(u) = \mu \Delta u$), with a forcing term and initial and boundary conditions corresponding to the



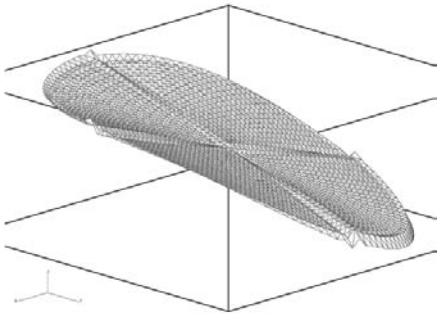
incremental projection method
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 1.53\,10^{-2}$



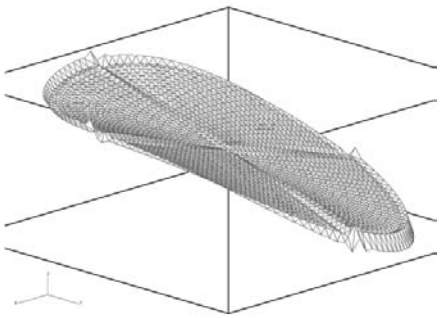
penalty-projection ($r = 1$)
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 2.82\,10^{-3}$



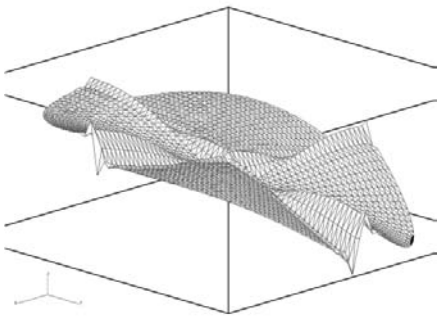
penalty-projection ($r = 100$)
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 2.84\,10^{-4}$



rotational penalty-proj. ($r = 1$)
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 9.51\,10^{-4}$



rotational projection method
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 1.13\,10^{-3}$



implicit method
$\|p_h - p\|_{\mathrm{L}^\infty(\Omega)} = 1.83\,10^{-4}$

FIG. III.5 – Stokes flow with prescribed velocity boundary conditions - distribution of the pressure error for the incremental, rotational, penalty-projection ($r = 1$, $r = 10$, $r = 100$) schemes and for the implicit scheme

following expression for the velocity and pressure fields :

$$u(x, y, t) = \begin{bmatrix} \sin(x)\ \sin(y + t) \\ \cos(x)\ \cos(y + t) \end{bmatrix}$$

$$p(x, y, t) = \cos(x) \sin(y + t)$$

The computational domain is $\Omega = [0,1] \times [0,1]$ and the velocity is prescribed on its boundary, except for the part included in the $y-$axis, where homogeneous natural Neumann conditions are imposed :

$$\mu \nabla u \cdot n - p\, n = 0$$

The meshes are obtained by cutting along its diagonals in four simplices each square of a $n \times n$ regular grid.

Figures III.6, III.7 and III.8 show, for the 80×80 meshing, the difference between the numerical solution and the analytical one at $t = 1$, measured in $[\mathrm{L}^2(\Omega)]^d$ and $[\mathrm{H}^1(\Omega)]^d$ norms for the velocity and in $\mathrm{L}^2(\Omega)$ norm for the pressure.

As expected, the incremental projection method converges for small time steps to a solution which differs from the plateau obtained with the other schemes, and corresponds to a much greater error for both velocity and pressure. Errors obtained at small time steps are gathered for various meshings in table III.3 for the incremental projection scheme and in table III.4 for the other schemes. A space convergence order of $1/2$ is observed for the incremental projection method, which is in agreement with the bound proven in [16], whereas an optimal convergence rate is obtained for the other schemes.

	20×20	40×40	80×80
$\|u - u_h\|_{[\mathrm{L}^2(\Omega)]^d}$	$7.76\ 10^{-4}$	$3.24\ 10^{-4}$	$1.46\ 10^{-4}$
$\|u - u_i\|_{[\mathrm{L}^2(\Omega)]^d}$	$9.66\ 10^{-7}$	$1.21\ 10^{-7}$	$1.51\ 10^{-8}$
$\|u - u_h\|_{[\mathrm{H}^1(\Omega)]^d}$	$3.64\ 10^{-2}$	$2.59\ 10^{-2}$	$1.84\ 10^{-2}$
$\|u - u_i\|_{[\mathrm{H}^1(\Omega)]^d}$	$8.39\ 10^{-5}$	$2.10\ 10^{-5}$	$5.26\ 10^{-6}$
$\|p - p_h\|_{\mathrm{L}^2(\Omega)}$	$6.14\ 10^{-2}$	$4.34\ 10^{-2}$	$3.07\ 10^{-2}$
$\|p - p_i\|_{\mathrm{L}^2(\Omega)}$	$2.19\ 10^{-4}$	$5.47\ 10^{-5}$	$1.37\ 10^{-5}$

TAB. III.3 – Difference between the exact solution and the numerical solution on the time-convergence plateau, for the incremental projection method, and difference between the exact solution and its interpolate. Values obtained at $t = 1$.

	20×20	40×40	80×80
$\|u - u_h\|_{[\mathrm{L}^2(\Omega)]^d}$	$9.12\ 10^{-7}$	$1.14\ 10^{-7}$	$1.55\ 10^{-8}$
$\|u - u_i\|_{[\mathrm{L}^2(\Omega)]^d}$	$9.66\ 10^{-7}$	$1.21\ 10^{-7}$	$1.51\ 10^{-8}$
$\|u - u_h\|_{[\mathrm{H}^1(\Omega)]^d}$	$7.77\ 10^{-5}$	$1.94\ 10^{-5}$	$4.85\ 10^{-6}$
$\|u - u_i\|_{[\mathrm{H}^1(\Omega)]^d}$	$8.39\ 10^{-5}$	$2.10\ 10^{-5}$	$5.26\ 10^{-6}$
$\|p - p_h\|_{\mathrm{L}^2(\Omega)}$	$6.01\ 10^{-5}$	$1.51\ 10^{-5}$	$3.76\ 10^{-6}$
$\|p - p_i\|_{\mathrm{L}^2(\Omega)}$	$2.19\ 10^{-4}$	$5.47\ 10^{-5}$	$1.37\ 10^{-5}$

TAB. III.4 – Difference between the exact solution and the numerical solution on the time-convergence plateau, for schemes other than the incremental projection method, and difference between the exact solution and its interpolate. Values obtained at $t = 1$.

A second order convergence with respect to the time step is observed for the implicit scheme for both the velocity and the pressure. The results of the rotational projection scheme conform to the numerical experiments reported in [16] : the convergence rate is respectively 1.65 and 1 for the velocity in $[\mathrm{L}^2(\Omega)]^d$ and $[\mathrm{H}^1(\Omega)]^d$ norm, and lies between 1 and $3/2$ for the pressure in $\mathrm{L}^2(\Omega)$ norm. As for the preceding numerical

experiments, the penalty-projection method with $r = \mu$ $(= 1)$ gives for the velocity the same results as the rotational projection scheme, then the results come closer and closer to the implicit scheme when the augmentation parameter increases. For the pressure, a higher value ($r = 10$) is necessary to recover the accuracy of the rotational projection method, this point being corrected by the use of the rotational penalty-projection variant.

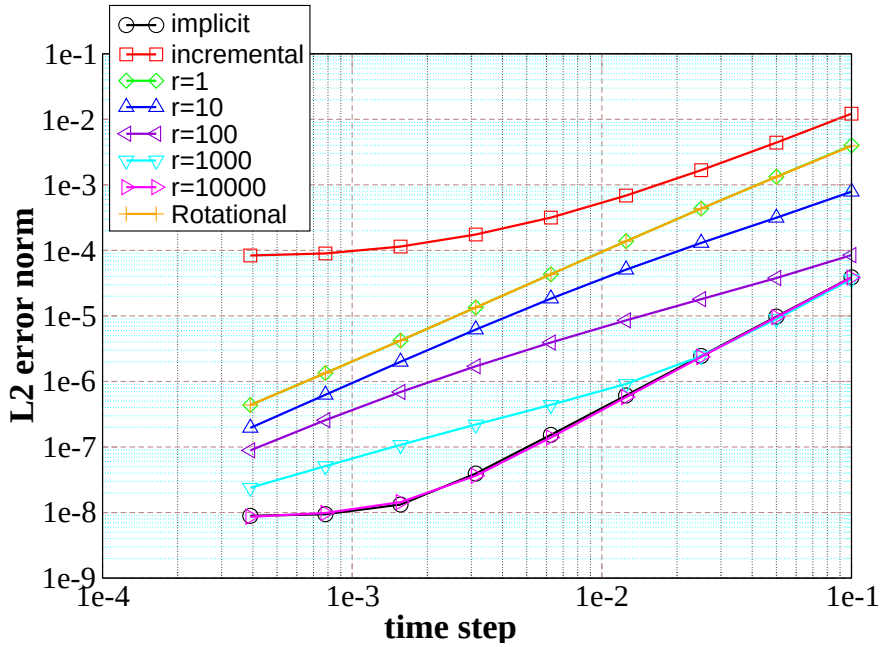


FIG. III.6 – Stokes flow with open boundary conditions - L^2 norm of the error for the velocity as a function of the time step, for the incremental, rotational, penalty-projection ($r = 1$, $r = 10$, $r = 100$) schemes and for the implicit scheme

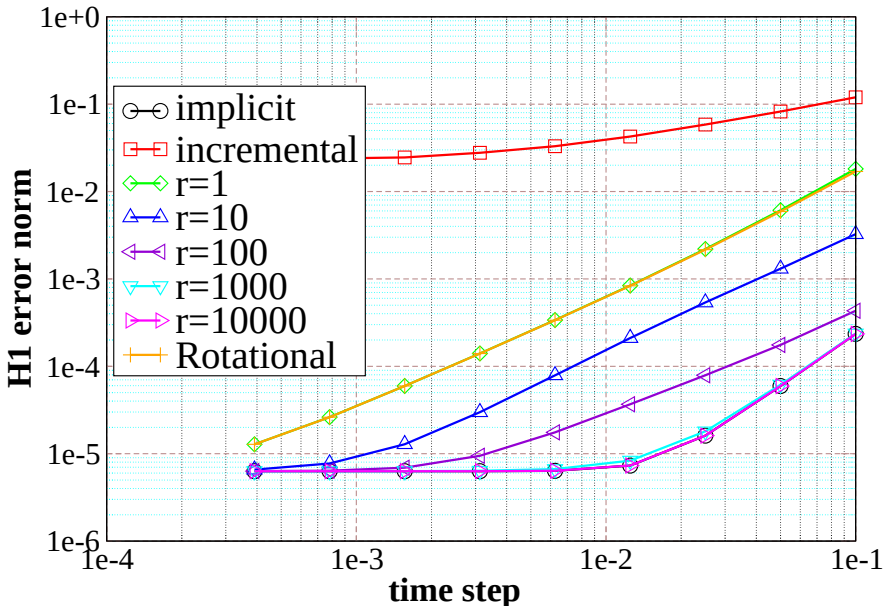


FIG. III.7 – Stokes flow with open boundary conditions - H^1 norm of the error for the velocity as a function of the time step, for the incremental, rotational, penalty-projection ($r = 1$, $r = 10$, $r = 100$) schemes and for the implicit scheme

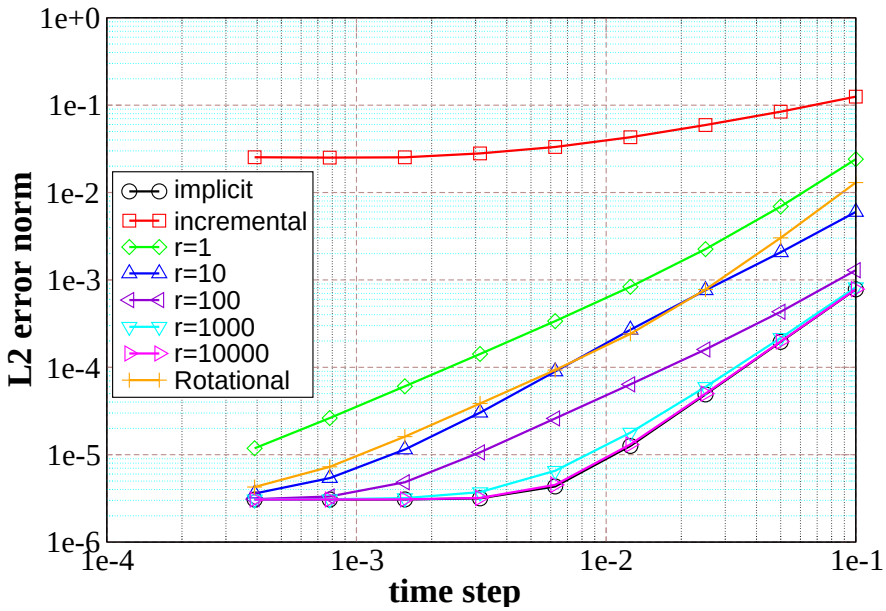


FIG. III.8 – Stokes flow with open boundary conditions - L^2 norm of the error for the pressure as a function of the time step, for the incremental, rotational, penalty-projection ($r = 1$, $r = 10$, $r = 100$) schemes and for the implicit scheme

III.4.4 Flow past a cylinder

In this section, we are concerned by a more physically meaningful problem, namely the two-dimensional, incompressible, laminar and unsteady flow of a viscous fluid around a solid circular cylinder, in the layout proposed in [20].

We use a rectangular domain $\Omega = [0, 20] \times [0, 5]$ with a solid obstacle represented by a disk C. Its center lies at the point (5,2.5) and its diameter is equal to 1.

The meshing of the computational domain is represented on figure III.9. The number of cells is roughly 4200, which leads to about 8500 degrees of freedom for each component of the velocity, 2200 for the pressure. The velocity is prescribed to zero on the boundary of C and to the value $(1, 0)$ at the inflow boundary of the computational domain. We impose a perfect slip condition to the velocity at the top and bottom of Ω and natural Neumann conditions at the exit.

The density is set to $\varrho = 1$ and the viscosity to $\mu = 0.01$, which leads to a cylinder diameter-based Reynolds number of 100, at which the flow is known to be unsteady, exhibiting the periodic shedding of vortices downstream the cylinder. This phenomenon is triggered in our computation by introducing a pertubation affecting the inlet velocity during one time unit, starting at time $t = 15$.

Results are clearly sensitive to the boundary conditions prescribed at the top and bottom of the computational domain, which shows than an accurate simulation of the physics of the flow would require a more appropriate treatment of the boundary conditions (either enlarging the computational domain, either using more sophisticated modelling of artificial boundaries as proposed in [3]). However, they are in reasonable agreement with the literature : drag coefficient varying in time from 1.69 to 1.72 and vortex shedding period of 4.92 time units.

The figure III.10 shows the evolution versus the augmentation parameter r of the

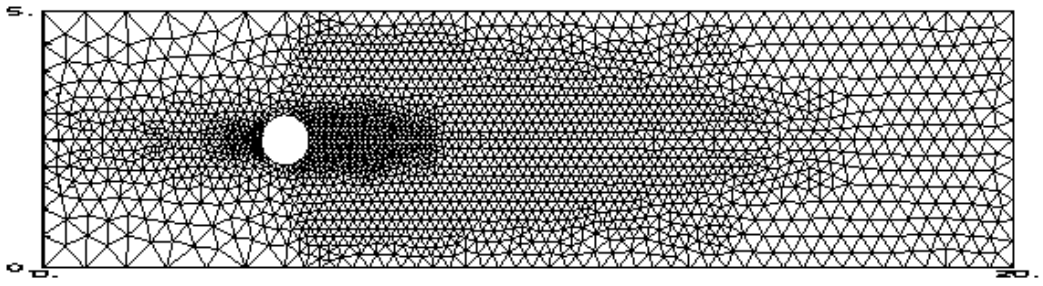


FIG. III.9 – Flow past a cylinder - Meshing of the computational domain

difference between the velocity fields obtained at time $t = 60$ by the semi-implicit coupled scheme and the penalty-projection one, measured in the $[L^2(\Omega)]^d$ norm. When r is large, the magnitude of the splitting error varies approximately as $1/r$. Moreover, we can see that it decreases with the time step, with an apparent order lying between 1 and 2. This behaviour can also be inferred from the variation of the total error as a function of the time step observed in the other test cases.

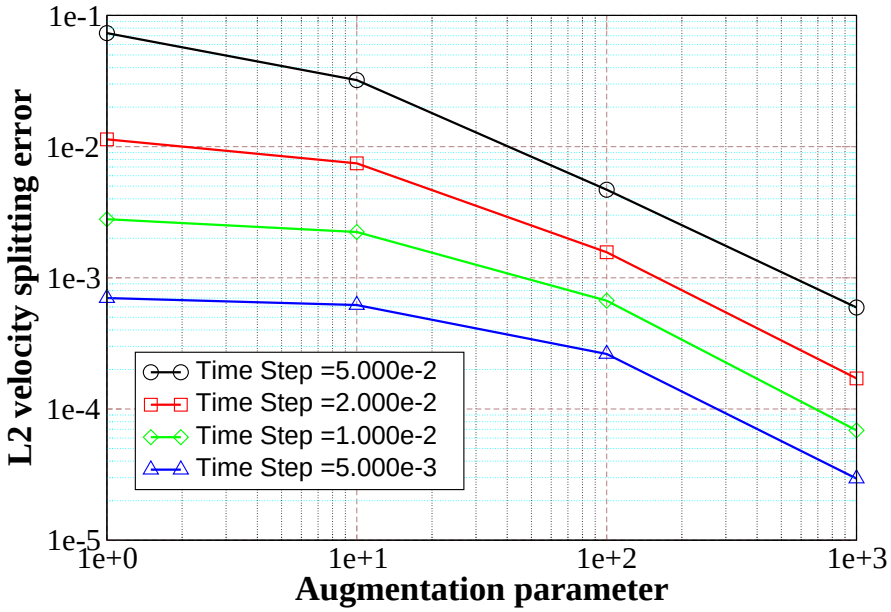


FIG. III.10 – Flow past a cylinder - L^2 norm of the splitting error for the velocity as a function of the augmentation parameter r, for the penalty-projection method

III.5 Discussion

The numerical experiments reported in this paper show that the penalty-projection method yields a considerable gain in accuracy compared to incremental projection schemes, implemented as well in standard form as in the rotational form (see [17] for a review). The splitting error is reduced as soon as the augmentation parameter r takes a significant value, and decreases down to negligible values as r is increased. Results of the semi-implicit method are thus recovered in the latter case. The pressure

layers suffered by the incremental projection scheme in the vicinity of the prescribed velocity boundaries are suppressed. Finally, the loss of spatial convergence of the standard incremental projection scheme in case of open boundary conditions does not occur anymore.

As a consequence, we observe that the scheme is convergent, or, more precisely speaking, enjoys convergence properties at least equivalent to the standard incremental projection method, whatever the augmentation parameter may be. This implies, in particular, that the rate of convergence for the velocity is bounded from above by the second order estimate valid for the incremental scheme with Dirichlet boundary conditions (even if the penalty-projection scheme is not second order), as soon as a formally second-order scheme is employed. We confirm this desirable feature by a theoretical analysis in energy norms in [1].

As classically observed when making use of penalty or augmented Lagrangian methods, the price to pay for this gain in accuracy is an ill-conditionning of the linear system associated to the prediction step, which may reinforce the importance of keeping a reasonable value for the augmentation parameter. Note also that the decoupling of the velocity components in the prediction step, when the divergence of the stress tensor can be expressed as the velocity laplacian (*i.e.* for constant viscosity and particular boundary conditions), is lost. However, even if the penalty-projection scheme is more time-consuming than the considered incremental projection methods for a given time-step, it has been observed in our tests to be cheaper to yield a given accuracy. Note nevertheless that this fact is reported here as a first information, but clearly needs additionnal assessment, with respect to at least two aspects. First, a numerical efficiency study must include a careful discussion concerning the preconditioning of the linear algebraic operators, specially for the velocity prediction step, in which the convergence of Krylov's subspaces methods highly depends on this issue; this point was disregarded in the present work. Second, the behaviour of linear solvers appears to be highly problem-dependent. For instance, at small time step and for the ILU preconditionned GMRES algorithm, taking a Richardson's time extrapolation of the velocity as starting point in the prediction stage, convergence was obtained within less than twenty iterations for the Taylor-Green vortices test, as it took a few hundred iterations for the computation of the flow past a cylinder. A careful design of a numerical efficiency study should then deal with this issue, which has not been the case for the choice of the tests investigated here. With this respect, note that the interest of the penalty-pojection scheme should grow for non-isothermal applications which involve temperature-dependent physical properties (and even, for some industrial applications, properties tabulated with respect to the temperature), for which the part of the CPU time devoted to the assembling of the discrete operators becomes important. In any case, testing the present scheme in a multi-grid framework seems appealing; this is the subject for a further work.

Bibliographie

[1] Ph. Angot, M. Jobelin, C. Lapuerta, and J.-C. Latché. On two variants of the penalty-projection method, 2005. in preparation.

[2] John B. Bell, Philip Colella, and Harland M. Glaz. A second-order projection method for the incompressible Navier-Stokes equations. *Journal of Computational Physics*, 85 :257–283, 1989.

[3] C.H. Bruneau, , and P. Fabrie. Effective downstream boundary conditions for incompressible Navier-Stokes equations. *International Journal for Numerical Methods in Fluids*, 19 :693–705, 1994.

[4] Jean-Paul Caltagirone and Jérôme Breil. Sur une méthode de projection vectorielle pour la résolution des équations de Navier-Stokes. *Comptes-Rendus de l'Académie des Sciences, Paris – Série II*, 327 :1179–1184, 1999.

[5] Alexandre Joel Chorin. Numerical solution of the Navier-Stokes equations. *Mathematics of Computation*, 22 :745–762, 1968.

[6] Ph. Clément. Approximation by finite element functions using local regularization. *Revue Française d'Automatique, Informatique et Recherche Opérationnelle*, R-2 :77–84, Août 1975.

[7] Alexandre Ern and Jean-Luc Guermond. *Eléments finis : théorie, applications, mise en œuvre*, volume 36 of *Mathématiques & Applications*. Springer, 2002.

[8] M. Fortin and R. Glowinski. *Méthodes de Lagrangien Augmenté*. Dunod, Paris, 1982.

[9] Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms.*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, 1986.

[10] Katuhiko Goda. A multistep technique with implicit difference schemes for calculating two- or three-dimensional cavity flows. *Journal of Computational Physics*, 30 :76–95, 1979.

[11] Philip M. Gresho. On the theory of semi-implicit projection methods for viscous incompressible flow and its implementation via a finite element method that also introduces a nearly consistent mass matrix. part 1 : Theory. *International Journal for Numerical Methods in Fluids*, 11 :587–620, 1990.

[12] Philip M. Gresho. On the theory of semi-implicit projection methods for viscous incompressible flow and its implementation via a finite element method that also introduces a nearly consistent mass matrix. part 2 : Implementation. *International Journal for Numerical Methods in Fluids*, 11 :621–659, 1990.

[13] J.-L. Guermond and L. Quartapelle. On the approximation of the unsteady Navier-Stokes equations by finite element projection methods. *Numerische Mathematik*, 80 :207–238, 1998.

[14] Jean-Luc Guermond. Some implementations of projection methods for Navier-Stokes equations. *Mathematical Modelling and Numerical Analysis*, 30(5) :637–667, 1996.

[15] Jean-Luc Guermond. Un résultat de convergence d'ordre deux en temps pour l'approximation des équations de Navier-Stokes par une technique de projection incrémentale. *Mathematical Modelling and Numerical Analysis*, 33(1) :169–189, 1999.

[16] J.L. Guermond, P. Minev, and Jie Shen. Error analysis of pressure-correction schemes for the Navier-Stokes equations with open boundary conditions. *submitted to SIAM Journal on Numerical Analysis*, 2004.

[17] J.L. Guermond, P. Minev, and Jie Shen. An overview of projection methods for incompressible flows, 2004. submitted to International Journal on Numerical Methods in Engineering.

[18] J.L. Guermond and Jie Shen. On the error estimates for the rotational pressure-correction projection methods. *Mathematics of Computation*, 73(248) :1719–1737, 2004.

[19] John G. Heywood and Rolf Rannacher. Finite element approximation of the nonstationary Navier-Stokes problem part IV : Error analysis for second-order time discretization. *SIAM Journal on Numerical Analysis*, 27(2) :353–384, 1990.

[20] Khodor Khadra, Philippe Angot, Sacha Parneix, and Jean-Paul Caltagirone. Fictitious domain approach for numerical modelling of Navier-Stokes equations. *International Journal for Numerical Methods in Fluids*, 34 :651–684, 2000.

[21] J. Kim and P. Moin. Application of a fractional-step method to incompressible Navier-Stokes equations. *Journal of Computational Physics*, 59 :308–323, 1985.

[22] Bruno Piar. PELICANS : Un outil d'implémentation de solveurs d'équations aux dérivées partielles. Note Technique 2004/33, IRSN, 2004. (in french).

[23] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, 1994.

[24] Rolf Rannacher. Finite element methods for the incompressible Navier-Stokes equations. In *Fundamental Directions in Mathematical Fluid Mechanics*, pages 191–293. Birkhauser, 2000.

[25] Jie Shen. On error estimates of projection methods for Navier-Stokes equations : First-order schemes. *SIAM Journal on Numerical Analysis*, 29(1) :57–77, 1992.

[26] Jie Shen. On error estimates of some higher order projection and penalty-projection methods for Navier-Stokes equations. *Numerische Mathematik*, 62 :49–73, 1992.

[27] Jie Shen. Remarks on the pressure estimates for the projection methods. *Numerische Mathematik*, 67 :513–520, 1994.

[28] Jie Shen. On error estimates of the penalty method for unsteady Navier-Stokes equations. *SIAM journal on Numerical Analysis*, 32(2) :386–403, 1995.

[29] Jie Shen. On error estimates of projection methods for Navier-Stokes equations : Second-order schemes. *Mathematics of Computation*, 65(215) :1039–1065, 1996.

[30] G.I. Taylor and B.A. Green. Mechanism of the production of small eddies from large ones. *Proceedings of the Royal Society of London A*, 158 :499–521, 1935.

[31] R. Temam. Sur l'approximation de la solution des Équations de Navier-Stokes par la méthode des pas fractionnaires (II). *Archive for Rational Mechanics and Analysis*, 33 :377–385, 1969.

[32] L.J.P. Timmermans, P.D. Minev, and F.N. Van de Vosse. An approximate projection scheme for incompressible flow using spectral elements. *International Journal for Numerical Methods in Fluids*, 22 :673–688, 1996.

[33] Stefan Turek. *Efficient Solvers for Incompressible Flow Problems : An Algorithmic Approach in View of Computational Aspects*. Springer, 1999.

[34] J. Van Kan. A second-order accurate pressure-correction scheme for viscous incompressible flow. *SIAM Journal on Scientific and Statistical Computing*, 7(3) :870–891, 1986.

Chapitre IV

Une méthode de projection-pénalité pour les écoulements dilatables

IV.1 Introduction

Dans cet article, nous nous intéressons à une classe de problèmes obtenus en adjoignant aux équations de bilan de masse et de quantité de mouvement une équation de transport diffusion d'une variable additionnelle, z, dont la masse volumique se déduit :

$$\left|\begin{array}{ll} \dfrac{\partial \varrho z}{\partial t} + \nabla \cdot \varrho z u = \nabla \cdot \mathcal{D} \nabla z & \text{dans } [0,T] \times \Omega \\ \varrho = \mathcal{G}(z) & \\ \dfrac{\partial \varrho u}{\partial t} + \nabla \cdot (\varrho u \otimes u) = \nabla \cdot \tau(u) - \nabla p + f & \text{dans } [0,T] \times \Omega \\ \dfrac{\partial \varrho}{\partial t} + \nabla \cdot \varrho u = 0 & \text{dans } [0,T] \times \Omega \end{array}\right. \tag{IV.1.1}$$

La variable ϱ désigne ici la masse volumique, t le temps, u la vitesse, p la pression, f une force volumique répartie, $\mathcal{D}$ est un coefficient de diffusion et τ est le tenseur des contraintes visqueuses. L'ensemble Ω est un domaine régulier de $\mathbb{R}^d$, $d = 2$ ou $d = 3$ et T est le temps final, $T < \infty$. Pour fixer les idées, on peut voir la fonction $\mathcal{G}(\cdot)$ comme une loi d'état. Pour que le problème soit complètement défini, ce système doit être complété par des conditions aux limites et initiales.

Des représentations mathématiques de ce type sont rencontrées lors de la modélisation d'une grande variété de phénomènes physiques. Les problèmes de convection naturelle à faible nombre de Mach, par exemple, rentrent dans ce cadre, lorsqu'ils sont traités en utilisant un modèle asymptotique, *i.e.* un système d'équations vérifié par les champs de vitesse et pression dans l'écoulement lorsque le nombre de Mach tend vers zéro, tel que décrit par Majda et Sethian [19]. La variable z représente alors la température et la loi $\mathcal{G}(\cdot)$ est déduite de la loi d'état, moyennant le calcul préalable de la pression thermodynamique lorsque le système physique considéré est clos. Si maintenant on identifie z à une concentration et $\varrho = \mathcal{G}(z)$ à une loi de mélange, on obtient les équations de la convection solutale. De la même manière, certains problèmes très simplifiés de combustion s'écrivent sous la forme du système (IV.1.1) ; la variable z prend alors la signification d'une variable d'avancement [3]. Ces problématiques physiques font partie des phénomènes d'intérêt dans le domaine de la sûreté nucléaire, telles que traitées à l'Institut de Radioprotection et de Sûreté Nucléaire (IRSN) ; c'est le cadre de la présente étude.

Du fait que la masse volumique du fluide est supposée indépendante de la pression, cette dernière joue d'un point de vue mathématique un rôle similaire à celui qu'elle tient dans les équations de Navier-Stokes incompressibles. Pour s'en convaincre, il suffit de réécrire l'équation de bilan de quantité de mouvement en fonction de la variable $q = \varrho u$ et les deux dernières équations du système (IV.1.1) retrouvent la structure classique d'un problème mixte.

En conséquence, il est naturel de mettre en œuvre pour la résolution numérique de ce problème des schémas initialement développés dans le contexte des écoulements incompressibles. Parmi ceux-ci, les méthodes de projection ont, depuis les travaux originels de Chorin et Témam [7, 22], acquis une popularité croissante. Ce succès tient dans le fait qu'elles découplent à chaque pas de temps les équations de bilan de quantité de mouvement et de bilan de masse, substituant ainsi à un problème mixte, de résolution difficile, une succession de problèmes elliptiques plus aisés à résoudre.

Le principe des méthodes de projection est le suivant. Dans une première étape, on obtient une prédiction de la vitesse par la résolution de l'équation de bilan de quantité de mouvement, dans laquelle la pression est ignorée (méthode originelle) ou approchée par une formule explicite (méthode dite incrémentale). La seconde étape consiste à projeter la vitesse prédite dans l'espace des fonctions solénoidales ; cette étape s'apparente à un problème de Darcy, qui est classiquement réécrit comme un problème elliptique pour la pression (méthode originelle) ou pour l'incrément de pression (méthode incrémentale). Sur cette idée de base sont venues, au fil des années, se greffer de multiples variantes. La méthode de projection incrémentale semble due à Goda [12], le premier schéma formellement de second ordre en temps à Van Kan [24]. Dans l'étape de projection, les conditions aux limites appliquées à l'incrément de pression sont artificielles, *i.e.* ne sont pas vérifiées par la solution du problème, ce qui induit des pertes de précision, particulièrement graves pour les écoulements visqueux obéissant à des conditions aux limites ouvertes sur une partie de la frontière [15]. Ce phénomène est corrigé dans une variante proposée par Timmermans *et al.* [23] puis analysée par Guermond et Shen [14], qui a reçu le nom de méthode rotationnelle. On trouvera une revue de ces différents schémas et de leurs propriétés de convergence respectives dans [16].

Si le découplage des équations de bilan de quantité de mouvement et de bilan de masse simplifie la résolution, il introduit également une erreur numérique, dite erreur de fractionnement, qui devient, pour des discrétisations en temps d'ordre deux, importante voire dominante à fort pas de temps [17]. Cette erreur de fractionnement disparaîtrait si, par un choix judicieux de la pression approchée, la vitesse prédite vérifiait la contrainte de divergence : la vitesse prédite serait alors la même que celle obtenue par un schéma couplé, et l'étape de projection serait sans objet. Bien sûr, ce comportement n'est pas accessible dans la pratique. Il peut toutefois être approché en ajoutant dans la première étape un terme de pénalisation associé à la contrainte de divergence, analogue à celui utilisé dans les techniques de Lagrangien augmenté : c'est le principe des méthodes de projection-pénalité. Le premier schéma de ce type semble avoir été suggéré par Shen [21, section 6], puis mis en œuvre indépendamment par Caltagirone et Breil [6], dans leur méthode dite de "projection vectorielle". Cette dernière, qui s'appuie sur une discrétisation spatiale en volumes finis, est l'un des ingrédients essentiels des schémas du code Aquilon [2]. Enfin, l'application de ces idées dans le contexte des éléments finis conduit à un schéma original expérimenté dans [17] et analysé dans [1].

Dans cet article, nous étendons dans un premier temps ces derniers travaux à des écoulements à masse volumique variable ; c'est l'objet de la section IV.2, qui traite donc du développement d'une méthode de projection-pénalité pour la résolution des équations de bilan de quantité de mouvement et de masse (deux dernières relations du système (IV.1.1)). Nous revenons ensuite au problème (IV.1.1) complet dans la dernière section, où nous décrivons la mise en œuvre d'un schéma à pas fractionnaires, basé sur une technique de projection, pour un écoulement de convection naturelle à faible nombre de Mach.

IV.2 Méthode de projection pour écoulement dilatable

L'objet de cette section est le développement d'une méthode de projection-pénalité pour la résolution du problème suivant :

$$\left|\begin{array}{ll} \dfrac{\partial \varrho u}{\partial t} + \nabla \cdot (\varrho u \otimes u) = \nabla \cdot \tau(u) - \nabla p + f & \text{dans } [0,T] \times \Omega \\ \nabla \cdot \varrho u + \dfrac{\partial \varrho}{\partial t} = 0 & \text{dans } [0,T] \times \Omega \\ u = u_{\mathrm{D}} & \text{sur } [0,T] \times \partial\Omega_{\mathrm{D}} \\ -pn + \tau(u) \cdot n = g_{\mathrm{N}} & \text{sur } [0,T] \times \partial\Omega_{\mathrm{N}} \end{array}\right.$$

où $\varrho = \varrho(x,t)$ est une fonction régulière donnée.

Les surfaces $\partial\Omega_{\mathrm{D}}$ et $\partial\Omega_{\mathrm{N}}$ forment une partition de la frontière $\partial\Omega$ du domaine de calcul de normale extérieure n, u_{D} et g_{N} désignent respectivement un champ donné de vitesse et de forces surfaciques définis sur la frontière. Le tenseur de cisaillement prend la forme classique caractéristique des écoulements newtoniens :

$$\tau(u) = \mu(\nabla u + \nabla u^{\mathrm{T}}) - \frac{2}{3}\mu \nabla \cdot u \, \mathrm{I}$$

μ étant la viscosité dynamique du fluide et I le tenseur identité dans $\mathbb{R}^d$.

Par souci de clarté, nous consacrons une première partie de cette section à la description sous forme semi-discrète en temps de la méthode de projection-pénalité proposée. Il conviendra toutefois de garder en mémoire que la formulation ainsi obtenue pour le terme de pénalisation n'est pas celle qui est utilisée *in fine* ; pour des raisons de précision, il est en effet préférable de construire ces derniers à partir de la formulation algébrique. Ce point est détaillé dans une remarque en fin de partie IV.2.3. Dans un second temps, la discrétisation par éléments finis est décrite, ce qui nous permet de donner exactement l'algorithme utilisé. Nous concluons cette section par des tests numériques.

IV.2.1 Méthode de projection-pénalité : formulation semi-discrète en temps

Notations : Soit $\phi : [0,T] \times \Omega \to \mathbb{R}$ une fonction arbitraire régulière. Notons $\phi^n = \phi(t^n)$ pour $0 \leq n \leq N$. Etant donnés $\phi^0, \ldots, \phi^n$, on obtient une approximation à l'ordre α de la valeur de ϕ à l'instant t^{n+1} par extrapolation de Richardson :

$$\phi(t^{n+1}) = \phi^{\star,n+1} + O(\Delta t^{\alpha}) \quad \text{avec} \quad \phi^{\star,n+1} \stackrel{\text{def}}{=} \sum_{j=0}^{\alpha-1} \gamma_j \phi^{n-j} \tag{IV.2.2}$$

De plus, une approximation à l'ordre α de la dérivée de ϕ à l'instant t^{n+1} est donnée par la formule de différentiation rétrograde suivante :

$$\frac{\partial \phi}{\partial t}(t^{n+1}) = \frac{D\phi^{n+1}}{\Delta t} + O(\Delta t^{q}) \quad \text{avec} \quad D\phi^{n+1} \stackrel{\text{def}}{=} \beta_q \phi^{n+1} - \sum_{j=0}^{q-1} \beta_j \phi^{n-j} \tag{IV.2.3}$$

Dans la pratique, nous utiliserons des approximations à l'ordre 1 et 2 où les coefficients β_j et γ_j sont donnés comme suit :

$$q = 1 \qquad D\phi^{n+1} = \phi^{n+1} - \phi^{n} \qquad \phi^{\star,n+1} = \phi^{n} \tag{IV.2.4}$$

$$q = 2 \quad D\phi^{n+1} = \frac{3}{2}\phi^{n+1} - 2\phi^{n} + \frac{1}{2}\phi^{n-1} \quad \phi^{\star,n+1} = 2\phi^{n} - \phi^{n-1} \tag{IV.2.5}$$

En introduisant la nouvelle variable $q = (\varrho u)$ représentant le débit massique, une semi-discrétisation en temps (formellement) à l'ordre q du problème traité dans cette section s'écrit :

$$\left|\begin{array}{ll}
\dfrac{Dq^{n+1}}{\Delta t} + \nabla \cdot (q^{\star,n+1} \otimes u^{n+1}) = \nabla \cdot \tau(u^{n+1}) - \nabla p^{n+1} + f^{n+1} & \text{dans } \Omega \\
\nabla \cdot q^{n+1} + \dfrac{D\varrho^{n+1}}{\Delta t} = 0 & \text{dans } \Omega \\
q^{n+1} = \varrho^{n+1} u^{n+1} & \\
u^{n+1} = u_{\mathrm{D}}^{n+1} & \text{sur } \partial\Omega_{\mathrm{D}} \\
-p^{n+1} n + \tau(u^{n+1}) \cdot n = g_{\mathrm{N}}^{n+1} & \text{sur } \partial\Omega_{\mathrm{N}}
\end{array}\right.$$

Les méthodes de projection ont pour principe d'utiliser une extrapolation $p^{\star,n+1}$ de la pression aux temps précédents dans l'équation de bilan de quantité de mouvement pour obtenir une prédiction de vitesse, avant de corriger cette dernière pour vérifier le bilan de masse dans une seconde étape. Nous choisirons simplement ici $p^{\star,n+1} = p^{n}$. L'étape de prédiction de la méthode de projection-pénalité décrite dans cet article s'écrit :

$$\left|\begin{array}{ll}
\dfrac{\beta_q \varrho^{n+1} \tilde{u}^{n+1} - \sum_{j=0}^{q-1} \beta_j q^{n-j}}{\Delta t} + \nabla \cdot (q^{\star,n+1} \otimes \tilde{u}^{n+1}) & \\
\quad -r\nabla(\nabla \cdot \varrho^{n+1} \tilde{u}^{n+1} + \dfrac{D\varrho^{n+1}}{\Delta t}) = \nabla \cdot \tau(\tilde{u}^{n+1}) - \nabla p^{n} + f^{n+1} & \text{dans } \Omega \\
\tilde{u}^{n+1} = u_{\mathrm{D}}^{n+1} & \text{sur } \partial\Omega_{\mathrm{D}} \\
-p^{n} n + \tau(\tilde{u}^{n+1}) \cdot n = g_{\mathrm{N}}^{n+1} & \text{sur } \partial\Omega_{\mathrm{N}}
\end{array}\right. \tag{IV.2.6}$$

Sa spécificité réside dans l'ajout du terme de pénalisation :

$$-r\nabla(\nabla \cdot \varrho^{n+1} \tilde{u}^{n+1} + \frac{D\varrho^{n+1}}{\Delta t}) \tag{IV.2.7}$$

où r est un coefficient positif que nous nommerons paramètre de pénalisation.

Ce terme est obtenu en appliquant l'opérateur gradient à l'équation de bilan de masse ; cette opération est bien sûr formelle pour le problème différentiel, mais ses analogues naturels variationnel et surtout, pour le système discret, algébrique peuvent, quant à eux, être précisément définis ; ce sera l'objet de la section IV.2.3.

Du fait notamment de la présence des termes de viscosité, l'inconnue naturelle de cette étape de prédiction (IV.2.6) est la vitesse $\tilde{u}^{n+1}$. Il est à noter alors que, contrairement à ce qui se passe lors d'une augmentation classique [10, 17], le terme de pénalisation a pour contrepartie variationnelle une forme bilinéaire qui n'est ni symétrique, ni positive. Cette caractéristique pourrait être corrigée en choisissant, à la place du terme proposé, l'expression suivante :

$$-r\varrho^{n+1}\nabla(\nabla\cdot\varrho^{n+1}\tilde{u}^{n+1}+\frac{D\varrho^{n+1}}{\Delta t})$$

Ce dernier choix présente l'inconvénient majeur que, ne pouvant plus s'écrire comme le gradient d'une quantité, il ne s'incorporera pas naturellement, par la suite, dans l'incrément de pression. En outre, nous prouvons dans la section suivante que, pour un pas de temps plus petit qu'une valeur seuil Δt_0 indépendante du paramètre de pénalisation r, l'équation aux dérivées partielles (IV.2.6) admet une solution et une seule, ce qui, d'une certaine manière, répond aux interrogations que soulève la non-positivité du terme de pénalisation choisi.

Pour compléter l'algorithme, il convient maintenant de construire l'étape de projection. Soit H l'espace affine défini comme suit :

$$H=\{q\in \mathrm{L}(\Omega)^d,\quad \nabla\cdot q=-\frac{D\varrho^{n+1}}{\Delta t},\quad q\cdot n=\varrho^{n+1}\ u_{\mathrm{D}}\cdot n \text{ sur } \partial\Omega_{\mathrm{D}}\}$$

La seconde étape revient alors à effectuer une projection L^2-orthogonale du débit prédit $\varrho^{n+1}\tilde{u}^{n+1}$ sur H ; par un argument de décomposition de Hodge, on obtient alors le système suivant :

$$\left|\begin{array}{ll} \beta_q\dfrac{q^{n+1}-\varrho^{n+1}\tilde{u}^{n+1}}{\Delta t}+\nabla\phi=0 & \text{dans } \Omega \\ \nabla\cdot q^{n+1}=-\dfrac{D\varrho^{n+1}}{\Delta t} & \text{dans } \Omega \\ q^{n+1}\cdot n=\varrho^{n+1}\ u_{\mathrm{D}}\cdot n & \text{sur } \partial\Omega_{\mathrm{D}} \end{array}\right. \qquad \text{(IV.2.8)}$$

En prenant la divergence de la première équation et en utilisant la seconde, on obtient le problème elliptique suivant pour ϕ :

$$\Delta\phi=\frac{\beta_q}{\Delta t}\left(\nabla\cdot\varrho^{n+1}\tilde{u}^{n+1}+\frac{D\varrho^{n+1}}{\Delta t}\right)$$

auquel il convient d'adjoindre des conditions aux limites. Sur les frontières où la vitesse est fixée, on a :

$$q^{n+1}\cdot n=\varrho^{n+1}u_{\mathrm{D}}^{n+1}\cdot n=\varrho^{n+1}\tilde{u}^{n+1}\cdot n$$

et par conséquent :

$$\nabla\phi\cdot n=0 \qquad \text{sur } \partial\Omega_{\mathrm{D}}$$

Sur les frontières de Neumann, la condition aux limites provient de la condition de L^2-orthogonalité de la projection sur H. Cette dernière s'écrit :

$$\int_\Omega (q^{n+1} - \varrho^{n+1}\tilde{u}^{n+1}) \cdot (q^{n+1} - v) = 0 \quad \forall v \in H$$

Grâce à la première relation de (IV.2.8) que l'on intègre par partie, puis par définition de H, on a :

$$\begin{aligned} \forall v \in H, \qquad 0 &= \int_\Omega \nabla\phi \cdot (q^{n+1} - v) \\ &= -\int_\Omega \phi \ \nabla \cdot (q^{n+1} - v) + \int_{\partial\Omega_N} \phi \ (q^{n+1} - v) \cdot n \\ &= \int_{\partial\Omega_N} \phi \ (q^{n+1} - v) \cdot n \end{aligned}$$

Cette relation nous donne la condition aux limites de ϕ sur la frontière $\partial\Omega_N$:

$$\phi = 0 \quad \text{sur } \partial\Omega_N$$

En rassemblant les relations obtenues, l'inconnue ϕ est solution du problème suivant :

$$\left| \begin{array}{ll} \Delta\phi = \dfrac{\beta_q}{\Delta t}\left(\nabla \cdot \varrho^{n+1}\tilde{u}^{n+1} + \dfrac{D\varrho^{n+1}}{\Delta t}\right) & \text{dans } \Omega \\ \nabla\phi \cdot n = 0 & \text{sur } \partial\Omega_D \\ \phi = 0 & \text{sur } \partial\Omega_N \end{array} \right. \tag{IV.2.9}$$

Une fois ϕ calculée, la première relation de (IV.2.8) permet de réactualiser le débit :

$$q^{n+1} = \varrho^{n+1}\tilde{u}^{n+1} - \frac{\Delta t}{\beta_q}\nabla\phi \quad \text{dans } \Omega \tag{IV.2.10}$$

Enfin, si l'on somme cette même relation avec l'équation de prédiction (IV.2.6), on obtient :

$$\begin{aligned} &\frac{\beta_q q^{n+1} - \sum_{j=0}^{q-1}\beta_j q^{n-j}}{\Delta t} + \nabla \cdot (q^{\star,n+1} \otimes \tilde{u}^{n+1}) \\ &\quad = \nabla \cdot \tau(\tilde{u}^{n+1}) - \nabla\left[p^n - r\left(\nabla \cdot \varrho^{n+1}\tilde{u}^{n+1} + \frac{D\varrho^{n+1}}{\Delta t}\right) + \phi\right] + f^{n+1} \end{aligned}$$

Cette relation n'est rien d'autre que la reconstitution de l'équation de quantité de mouvement, ce qui suggère l'expression suivante pour la pression en fin de pas :

$$p^{n+1} = p^n - r\left(\nabla \cdot \varrho^{n+1}\tilde{u}^{n+1} + \frac{D\varrho^{n+1}}{\Delta t}\right) + \phi \tag{IV.2.11}$$

En conclusion, effectuer un pas de temps consiste à résoudre en séquence les problèmes elliptiques (IV.2.6) et (IV.2.9) puis à réactualiser le débit et la pression par respectivement (IV.2.10) et (IV.2.11).

IV.2.2 Analyse de l'étape de prédiction

L'étude présentée dans cette partie est motivée par le fait qu'il est important dans la pratique de pouvoir choisir le paramètre de pénalisation r indépendamment du pas de temps ; nous confirmons que la méthode de projection-pénalité satisfait cette propriété, dans la mesure où l'étape de prédiction constitue un problème bien posé, pour un pas de temps inférieur à un pas de temps seuil qui est bien indépendant de r. Ce résultat n'est pas immédiat, du fait que des estimations *a priori* inhabituelles doivent être prouvées, pour deux raisons : d'une part, le champ advectif n'est pas à divergence nulle, et la forme bilinéaire associée au terme convectif n'est en conséquence pas antisymétrique ; d'autre part, le terme de pénalisation n'est pas symétrique.

A des fins de simplifications, nous supposerons que des conditions aux limites de Dirichlet homogènes sont imposées à la vitesse sur la totalité de la frontière $\partial\Omega$ (*i.e.* $\partial\Omega_{\rm D} = \partial\Omega$, $u_{\rm D} = 0$ et $\partial\Omega_{\rm N} = \emptyset$). Le problème que nous étudions ici prend alors la forme suivante :

$$\text{Trouver } u \in {\rm H}^1_0(\Omega)^d \text{ tel que} \qquad a(u,v) = (f,v), \quad \forall v \in {\rm H}^1_0(\Omega)^d$$

où, pour une discrétisation du premier ordre en temps :

$$\begin{aligned} a(u,v) \;&= \frac{1}{\Delta t}\int_\Omega \varrho u\cdot v + \int_\Omega \nabla\cdot(q\otimes u)\cdot v + c(u,v) \\ &\quad + r\int_\Omega \nabla\cdot\varrho u\;\nabla\cdot v \end{aligned}$$

Le passage à une discrétisation du second ordre de la dérivée temporelle n'aurait d'autre effet que de multiplier le premier terme par un coefficient positif constant, et l'adaptation des arguments développés ici ne nécessiterait que des modifications minimes.

Le débit q et la masse volumique ϱ sont connus ; nous supposons que q est borné dans ${\rm L}^\infty(\Omega)^d$ et que ϱ vérifie d'une part que $0 < \varrho_{\min} \leq \varrho \leq \varrho_{\max}$, d'autre part que $\nabla\varrho$ est borné dans ${\rm L}^\infty(\Omega)^d$. Il est à noter que ces deux hypothèses font que $\nabla\varrho/\varrho$ est également borné dans ${\rm L}^\infty(\Omega)^d$, propriété qui sera utilisée par la suite.

La forme bilinéaire $c(\cdot,\cdot)$ correspond à la dissipation visqueuse. Pour l'exprimer, nous supposons que la viscosité est constante, si bien que la divergence du tenseur de cisaillement s'écrit, pour une fonction régulière :

$$\nabla\cdot\tau(u) = \mu\,\Delta u + \frac{1}{3}\mu\,\nabla\nabla\cdot u$$

et, en conséquence :

$$c(u,v) = \mu\int_\Omega \nabla u : \nabla v + \frac{1}{3}\mu\int_\Omega \nabla\cdot u\;\nabla\cdot v$$

Nous débutons cette section par deux lemmes techniques.

Lemme 1 - L'opérateur de convection vérifie les résultats de continuité suivants :

$\forall u \in {\rm H}^1_0(\Omega)^d$,

$$\left|\int_\Omega \nabla\cdot(q\otimes u)\cdot\varrho u\right| \;\leq \sqrt{d}\;\|q\|_\infty\;\left(1+\sqrt{d}\,c_{\rm p}\;\|\frac{\nabla\varrho}{\varrho}\|_\infty\right)\;\|\nabla u\|_0\;\|\varrho u\|_0$$

$$\left|\int_\Omega \nabla\cdot(q\otimes\frac{1}{\varrho}u)\cdot u\right| \;\leq \sqrt{d}\;\|\frac{1}{\varrho}q\|_\infty\;\;\|\nabla u\|_0\;\|u\|_0$$

où $c_{\rm p} = c_{\rm p}(\Omega)$ désigne la constante de Poincaré.

Preuve - Du fait que le champ de vitesse u s'annule aux frontières du domaine, nous avons :

$$\begin{aligned}\int_\Omega \nabla\cdot(q\otimes u)\cdot\varrho u &= -\int_\Omega q\otimes u : \nabla(\varrho u)\\ &= -\int_\Omega \sum_{i,j=1}^d q_i\ u_j\ \frac{\partial}{\partial x_j}(\varrho u_i)\\ &= -\underbrace{\int_\Omega \sum_{i,j=1}^d q_i\ \varrho u_j\ \frac{\partial u_i}{\partial x_j}}_{(1)} - \underbrace{\int_\Omega \sum_{i,j=1}^d q_i\ \varrho u_j\ \frac{1}{\varrho}\frac{\partial \varrho}{\partial x_j}u_i}_{(2)}\end{aligned}$$

En appliquant l'inégalité de Cauchy-Schwarz d'abord dans $\mathbb{R}^{d\times d}$ puis dans L^2, la première de ces deux intégrales se majore comme suit :

$$\begin{aligned}|(1)| &\le \int_\Omega \left[\sum_{i,j=1}^d (q_i\ \varrho u_j)^2\right]^{1/2}\ \left[\sum_{i,j=1}^d (\frac{\partial u_i}{\partial x_j})^2\right]^{1/2}\\ &\le \left[\int_\Omega \sum_{i,j=1}^d (q_i\ \varrho u_j)^2\right]^{1/2}\left[\int_\Omega \sum_{i,j=1}^d (\frac{\partial u_i}{\partial x_j})^2\right]^{1/2}\\ &\le \sqrt{d}\ \|q\|_\infty\ \ \|\varrho u\|_0\ \|\nabla u\|_0\end{aligned}$$

De la même manière, nous avons pour la seconde intégrale :

$$\begin{aligned}|(2)| &\le \left[\int_\Omega \sum_{i,j=1}^d (q_i\ \varrho u_j)^2\right]^{1/2}\left[\int_\Omega \sum_{i,j=1}^d (\frac{1}{\varrho}\frac{\partial \varrho}{\partial x_j}u_i)^2\right]^{1/2}\\ &\le \sqrt{d}\ \|q\|_\infty\ \|\varrho u\|_0\ \sqrt{d}\ \|\frac{\nabla\varrho}{\varrho}\|_\infty\ \|u\|_0\\ &\le d\ c_{\rm p}\ \|q\|_\infty\ \|\frac{\nabla\varrho}{\varrho}\|_\infty\ \|\varrho u\|_0\ \|\nabla u\|_0\end{aligned}$$

En regroupant ces deux inégalités, on obtient la première relation à démontrer. Pour la seconde, on procède de même :

$$\int_\Omega \nabla\cdot(q\otimes\frac{1}{\varrho}u)\cdot u = -\int_\Omega (q\otimes\frac{1}{\varrho}u):\nabla u = \int_\Omega \sum_{i,j=1}^d \frac{1}{\varrho}\ q_i\ u_j\ \frac{\partial u_i}{\partial x_j}$$

Par l'inégalité de Cauchy-Schwarz, il vient :

$$\left|\int_\Omega \nabla\cdot(q\otimes\frac{1}{\varrho}u)\cdot u\right| \le \left[\int_\Omega \sum_{i,j=1}^d (\frac{q_i}{\varrho}u_j)^2\right]^{1/2}\ \left[\int_\Omega \sum_{i,j=1}^d (\frac{\partial u_i}{\partial x_j})^2\right]^{1/2}$$

et l'inégalité recherchée s'en déduit en faisant apparaître $\|\frac{1}{\varrho}q\|_\infty$ dans le premier terme.

Lemme 2 - L'opérateur de diffusion vérifie les inégalités de stabilité suivantes :

$$\forall u \in \mathrm{H}_0^1(\Omega)^d,$$

$$\begin{aligned} c(u,\varrho u) \geq \quad & \frac{1}{2}\,\mu\,\varrho_{\min}\,\|\nabla u\|_0^2 + \frac{1}{6}\,\mu\,\varrho_{\min}\,\|\nabla\cdot u\|_0^2 \\ & -\frac{2d}{3}\frac{\mu}{\varrho_{\min}}\,\|\frac{\nabla\varrho}{\varrho}\|_\infty^2 \quad \|\varrho u\|_0^2 \\ c(\frac{1}{\varrho}u,u) \geq \quad & \frac{1}{2}\,\frac{\mu}{\varrho_{\max}}\,\|\nabla u\|_0^2 + \frac{1}{6}\,\frac{\mu}{\varrho_{\max}}\,\|\nabla\cdot u\|_0^2 \\ & -\frac{2d}{3}\,\mu\varrho_{\max}\,\|\nabla(\frac{1}{\varrho})\|_\infty^2 \quad \|u\|_0^2 \end{aligned}$$

Preuve - Nous avons :

$$c(u,\varrho u) = \mu\int_\Omega \nabla u : \nabla(\varrho u) + \frac{\mu}{3}\int_\Omega \nabla\cdot u\,\nabla\cdot\varrho u$$

En développant les dérivées des produits, il vient :

$$c(u,\varrho u) = \mu\int_\Omega \varrho\ \|\nabla u\|^2 + T_1 + \frac{\mu}{3}\int_\Omega \varrho\ (\nabla\cdot u)^2 + T_2 \qquad \text{(IV.2.12)}$$

avec :

$$\begin{aligned} T_1 &= \mu\int_\Omega \sum_{i,j=1}^d \frac{1}{\varrho}\,\frac{\partial\varrho}{\partial x_j}\,\frac{\partial u_i}{\partial x_j}\,\varrho u_i \\ T_2 &= \frac{1}{3}\,\mu\int_\Omega \left(\sum_{i=1}^d \frac{\partial u_i}{\partial x_i}\right)\left(\sum_{i=1}^d \frac{1}{\varrho}\,\frac{\partial\varrho}{\partial x_i}\,\varrho u_i\right) \end{aligned}$$

Le premier de ces deux termes est majoré en utilisant l'inégalité de Cauchy-Schwarz successivement dans $\mathbb{R}^{d\times d}$ et dans L^2 comme suit :

$$\begin{aligned} |T_1| &\leq \mu\,\|\frac{\nabla\varrho}{\varrho}\|_\infty \int_\Omega \sum_{i,j=1}^d \frac{\partial u_i}{\partial x_j}\,\varrho u_i \\ &\leq \mu\,\|\frac{\nabla\varrho}{\varrho}\|_\infty \int_\Omega \left[\sum_{i,j=1}^d (\frac{\partial u_i}{\partial x_j})^2\right]^{1/2}\left[d\sum_{i=1}^d (\varrho u_i)^2\right]^{1/2} \\ &\leq \mu\,\|\frac{\nabla\varrho}{\varrho}\|_\infty \left[\int_\Omega \sum_{i,j=1}^d (\frac{\partial u_i}{\partial x_j})^2\right]^{1/2}\left[d\int_\Omega\sum_{i=1}^d (\varrho u_i)^2\right]^{1/2} \\ &\leq \sqrt{d}\,\mu\,\|\frac{\nabla\varrho}{\varrho}\|_\infty\ \|\nabla u\|_0\ \|\varrho u\|_0 \end{aligned}$$

De la même manière, par l'inégalité de Cauchy-Schwarz dans L^2, nous avons pour le second terme :

$$\begin{aligned} |T_2| &\leq \frac{1}{3}\mu\,\|\frac{\nabla\varrho}{\varrho}\|_\infty \left[\int_\Omega \left(\sum_{i=1}^d \frac{\partial u_i}{\partial x_i}\right)^2\right]^{1/2}\left[\int_\Omega\left(\sum_{i=1}^d \varrho u_i\right)^2\right]^{1/2} \\ &\leq \frac{\sqrt{d}}{3}\,\mu\,\|\frac{\nabla\varrho}{\varrho}\|_\infty\ \|\nabla\cdot u\|_0\ \|\varrho u\|_0 \end{aligned}$$

L'équation (IV.2.12) a donc pour conséquence :

$$c(u,\varrho u) \geq \mu\,\varrho_{\min}\,\|\nabla u\|_0^2 + \frac{\mu}{3}\,\varrho_{\min}\,\|\nabla\cdot u\|_0^2$$
$$-\sqrt{d}\,\mu\,\|\frac{\nabla\varrho}{\varrho}\|_\infty\;\|\varrho u\|_0\left(\;\|\nabla u\|_0 + \frac{1}{3}\,\|\nabla\cdot u\|_0\;\right)$$

et la première inégalité recherchée s'en déduit par l'inégalité de Young.

Pour la seconde inégalité, on a :

$$c(\frac{1}{\varrho}u,u) = \mu\int_\Omega \frac{1}{\varrho}\;\|\nabla u\|^2 + T_1 + \frac{\mu}{3}\int_\Omega\frac{1}{\varrho}\;(\nabla\cdot u)^2 + T_2 \qquad \text{(IV.2.13)}$$

avec :

$$T_1 = \mu\int_\Omega\sum_{i,j=1}^d \frac{\partial}{\partial x_j}(\frac{1}{\varrho})u_i\;\frac{\partial u_i}{\partial x_j}$$
$$T_2 = \frac{1}{3}\;\mu\int_\Omega\left[\nabla(\frac{1}{\varrho})\cdot u\right]\;\nabla\cdot u$$

Le premier de ces deux termes se majore comme suit :

$$\begin{aligned}|T_1| \;&\leq \mu\,\|\nabla(\frac{1}{\varrho})\|_\infty\;\left[\int_\Omega\sum_{i,j=1}^d(\frac{\partial u_i}{\partial x_j})^2\right]^{1/2}\left[\int_\Omega\sum_{i,j=1}^d u_i^2\right]^{1/2}\\ &\leq \sqrt{d}\,\mu\,\|\nabla(\frac{1}{\varrho})\|_\infty\;\|u\|_0\;\|\nabla u\|_0\end{aligned}$$

De même, pour le second :

$$|T_2| \leq \frac{1}{3}\;\sqrt{d}\,\mu\,\|\nabla(\frac{1}{\varrho})\|_\infty\;\|u\|_0\;\|\nabla\cdot u\|_0$$

Le résultat recherché s'en déduit en reportant ces deux relations dans (IV.2.13) et en utilisant l'inégalité de Young.

On rappelle le résultat général suivant (*c.f.*, par exemple, [9, théorème 3.2.3]) :

Théorème 1 - Soit V un espace de Hilbert muni du produit scalaire $(\cdot,\cdot)$ et de la norme associée $\|\cdot\|$, $a(\cdot,\cdot)$ une forme bilinéaire sur $V\times V$ et $f\in V$. On suppose que les deux hypothèses suivantes sont vérifiées :

$$\inf_{w\in V}\sup_{v\in V}\frac{a(w,v)}{\|w\|\;\|v\|} \geq \beta > 0$$
$$(\forall w\in V,\; a(w,v)=0) \Longrightarrow (v=0)$$

Alors le problème :

$$\text{Trouver } u\in V \text{ tel que } a(u,v) = (f,v)\quad \forall v\in V$$

admet une solution et une seule.

Ce théorème est exploité pour démontrer le résultat suivant :

Proposition - Si le pas de temps Δt est plus petit qu'une valeur seuil Δt_0 indépendante du paramètre de pénalisation r (plus précisément, tel que les inégalités (IV.2.15) et (IV.2.16) soient vérifiées), l'étape de prédiction admet une solution et une seule.

Preuve - Nous allons vérifier les deux hypothèses du théorème ci-dessus. Soit donc u une fonction de $\mathrm{H}_0^1(\Omega)^d$. Par les premières relations des deux lemmes ci-dessus, nous avons :

$$\begin{aligned} a(u,\varrho u) \;&\geq \frac{1}{\Delta t}\,\|\varrho u\|_0^2 \;-\sqrt{d}\;\|q\|_\infty\;\left(1+\sqrt{d}\,c_{\mathrm{p}}\;\|\frac{\nabla\varrho}{\varrho}\|_\infty\right)\;\|\nabla u\|_0\;\|\varrho u\|_0 \\ &+\frac{1}{2}\,\mu\,\varrho_{\min}\,\|\nabla u\|_0^2\;+\frac{1}{6}\,\mu\,\varrho_{\min}\,\|\nabla\cdot u\|_0^2\;-\frac{2d}{3}\;\frac{\mu}{\varrho_{\min}}\,\|\frac{\nabla\varrho}{\varrho}\|_\infty^2\;\;\|\varrho u\|_0^2 \\ &+r\,\|\nabla\cdot\varrho u\|_0^2 \end{aligned} \tag{IV.2.14}$$

En majorant le second terme du membre de droite par l'inégalité de Young, il vient :

$$\begin{aligned} a(u,\varrho u)\geq\;& \\ &\left[\frac{1}{\Delta t}-\frac{d}{\mu\,\varrho_{\min}}\;\|q\|_\infty^2\;\;\left(1+\sqrt{d}\,c_{\mathrm{p}}\;\|\frac{\nabla\varrho}{\varrho}\|_\infty\right)^2-\frac{2d}{3}\;\frac{\mu}{\varrho_{\min}}\,\|\frac{\nabla\varrho}{\varrho}\|_\infty^2\right]\;\|\varrho u\|_0^2 \\ &+\frac{1}{4}\,\mu\,\varrho_{\min}\,\|\nabla u\|_0^2\;+\frac{1}{6}\,\mu\,\varrho_{\min}\,\|\nabla\cdot u\|_0^2\;+r\,\|\nabla\cdot\varrho u\|_0^2 \end{aligned}$$

Compte tenu du fait que, par hypothèse sur ϱ :

$$\|\varrho u\|_0^2\;\geq\varrho_{\min}^2\,\|u\|_0^2\qquad\text{et}\qquad\|\varrho u\|_1\;\leq c(\varrho)\,\|u\|_1$$

l'inégalité ci-dessus fournit la première relation du théorème 1, pourvu que Δt soit suffisamment petit pour que la condition suivante soit vérifiée :

$$\frac{1}{\Delta t}-\frac{d}{\mu\,\varrho_{\min}}\;\|q\|_\infty^2\;\;\left(1+\sqrt{d}\,c_{\mathrm{p}}\;\|\frac{\nabla\varrho}{\varrho}\|_\infty\right)^2-\frac{2d}{3}\;\frac{\mu}{\varrho_{\min}}\,\|\frac{\nabla\varrho}{\varrho}\|_\infty^2\;>0 \tag{IV.2.15}$$

De la même manière, en utilisant cette fois-ci les secondes relations des lemmes 1 et 2, on a :

$$\begin{aligned} a(\frac{1}{\varrho}u,u)\;&\geq\frac{1}{\Delta t}\,\|u\|_0^2\;-\sqrt{d}\;\|\frac{1}{\varrho}\,q\|_\infty\;\;\|\nabla u\|_0\;\|u\|_0 \\ &+\frac{1}{2}\,\frac{\mu}{\varrho_{\max}}\,\|\nabla u\|_0^2\;+\frac{1}{6}\,\frac{\mu}{\varrho_{\max}}\,\|\nabla\cdot u\|_0^2\;-\frac{2d}{3}\,\mu\,\varrho_{\max}\,\|\nabla(\frac{1}{\varrho})\|_\infty^2\;\;\|u\|_0^2 \\ &+r\,\|\nabla\cdot u\|_0^2 \end{aligned}$$

Soit, toujours par l'inégalité de Young :

$$\begin{aligned} a(\frac{1}{\varrho}u,u)\;&\geq\left[\frac{1}{\Delta t}-d\,\frac{\varrho_{\max}}{\mu}\,\|\frac{1}{\varrho}\,q\|_\infty^2\;-\frac{2d}{3}\,\mu\,\varrho_{\max}\,\|\nabla(\frac{1}{\varrho})\|_\infty^2\right]\;\|u\|_0^2 \\ &+\frac{1}{4}\,\frac{\mu}{\varrho_{\max}}\,\|\nabla u\|_0^2\;+\frac{1}{6}\,\frac{\mu}{\varrho_{\max}}\,\|\nabla\cdot u\|_0^2\;+r\,\|\nabla\cdot u\|_0^2 \end{aligned}$$

ce qui fournit la deuxième relation du théorème 1, pourvu que :

$$\frac{1}{\Delta t}-d\,\frac{\varrho_{\max}}{\mu}\,\|\frac{1}{\varrho}\,q\|_\infty^2\;-\frac{2d}{3}\,\mu\,\varrho_{\max}\,\|\nabla(\frac{1}{\varrho})\|_\infty^2\;>0 \tag{IV.2.16}$$

Remarque - En fait, la seconde hypothèse du théorème 1 peut être démontrée directement à partir de la relation (IV.2.14), en changeant u par u/ϱ; on obtient alors, sous la condition (IV.2.15) :

$$\forall u \in \mathrm{H}_0^1(\Omega)^d, \quad a(\frac{1}{\varrho}u, u) \geq c \ \|u\|_0^2$$

où c est une constante strictement positive, ce qui permet de conclure. Demander au pas de temps de vérifier (IV.2.16) n'est donc pas nécessaire ; en revanche, nous obtenons sous cette hypothèse un résultat plus fort, à savoir une stabilité en norme H^1 :

$$\forall u \in \mathrm{H}_0^1(\Omega)^d, \quad a(\frac{1}{\varrho}u, u) \geq c \ (\ \|u\|_0^2 \ + \ \|\nabla u\|_0^2 \ )$$

IV.2.3 Une implémentation éléments finis

L'objet de cette section est d'effectuer une description des schémas introduits dans cet article sous forme algébrique. Nous présentons tout d'abord la formulation variationnelle du problème ontenu après une semi-discrétisation en temps semi-implicite couplant bilan de masse et de quantité de mouvement, ainsi qu'une méthode de résolution de ce problème ; dans un second temps, nous établissons la forme algébrique de la méthode de projection-pénalité qui, nous le rappelons, est celle utilisée en pratique et diffère légèrement, pour le terme de pénalisation, de ce que l'on obtiendrait si l'on discrétisait en espace les équations du schéma semi-discret en temps présenté dans la section précédente.

Le schéma linéairement implicite

Nous supposons donnés deux espaces éléments finis de Lagrange V_h et M_h, le premier étant utilisé pour discrétiser respectivement vitesse et débit, le second pour la pression, et inclus respectivement dans les espaces $\mathrm{H}^1(\Omega)^d$ et $\mathrm{L}^2(\Omega)$. On notera :

$$V_h^{\partial\Omega_\mathrm{D}} = \{v \in V_h, \ v = 0 \text{ sur } \partial\Omega_\mathrm{D}\}$$

Une fois la discrétisation en temps effectuée, le problème revient à rechercher à chaque pas de temps $n+1$, la vitesse $u_h^{n+1} \in V_h$, la pression $p_h^{n+1} \in M_h$ et le débit $q_h^{n+1} \in V_h$.

Nous décomposons alors la vitesse et le débit en deux parties telles que :

$$u_h = u_D + u_F$$
$$q_h = q_D + q_F$$

où u_F (respectivement q_F) appartient à l'espace $V_h^{\partial\Omega_\mathrm{D}}$ et est le champ de vitesse (respectivement débit) inconnu. Les fonctions u_D et q_D peuvent être considérés comme des relèvements discrets des conditions aux limites de Dirichlet non-homogènes.

La formulation variationnelle du précédent problème au pas de temps $n+1$ s'écrit :

Trouver $(u_F^{n+1}, q_F^{n+1}) \in (V_h^{\partial\Omega_\mathrm{D}})^2$ et $p_h^{n+1} \in M_h$ tels que, $\forall v \in V_h^{\partial\Omega_\mathrm{D}},\ \forall t \in M_h$:

$$\left|\begin{array}{l} \displaystyle\int_\Omega \frac{\varrho^{n+1}\beta_q(u_F^{n+1}+u_D^{n+1}) - \sum_{j=0}^{q-1}\beta_j q_h^{n-j}}{\Delta t}\cdot v \\ \quad \displaystyle + \int_\Omega \nabla\cdot(q_h^{\star,n+1}\otimes(u_F^{n+1}+u_D^{n+1}))\cdot v + \int_\Omega \tau(u_F^{n+1}+u_D^{n+1}) : \nabla v \\ \qquad\qquad \displaystyle - \int_\Omega p_h^{n+1}\nabla\cdot v = \int_\Omega f^{n+1}\cdot v + \int_{\partial\Omega_\mathrm{N}} g_\mathrm{N}^{n+1}\cdot v \\ \displaystyle\int_\Omega \left(\frac{D\varrho^{n+1}}{\partial t} + \nabla\cdot[\varrho^{n+1}(u_F^{n+1}+u_D^{n+1})]\right) t = 0 \\ \displaystyle\int_\Omega (q_F^{n+1}+q_D^{n+1})\cdot v = \int_\Omega \varrho^{n+1}(u_F^{n+1}+u_D^{n+1})\cdot v \end{array}\right.$$

En exploitant de la manière usuelle cette formulation variationnelle, nous obtenons le système algébrique suivant :

$$\left|\begin{array}{l} \dfrac{\beta_q}{\Delta t}\mathbf{M}_{\varrho^{n+1}}\mathbf{U}_\mathrm{F} + \mathbf{A}\mathbf{U}_\mathrm{F} + \mathbf{B}^\mathrm{T}\mathbf{P} = \mathbf{F} \\ \mathbf{B}_{\varrho^{n+1}}\mathbf{U}_\mathrm{F} = \mathbf{G} \\ \mathbf{M}\mathbf{Q}_\mathbf{F} = \mathbf{M}_{\varrho^{n+1}}\mathbf{U}_\mathrm{F} + \mathbf{H} \end{array}\right.$$

où, par souci de simplicité, les exposants relatifs à l'avancée en temps ont été omis. Les vecteurs d'inconnues correspondent aux vitesses et pression en fin de pas de temps. Les opérateurs discrets sont les suivants : $\mathbf{M}$ et $\mathbf{M}_{\varrho^{n+1}}$ désignent respectivement la matrice de masse de vitesse standard et pondérée par la masse volumique à t^{n+1} (*i.e.* ϱ^{n+1}), $\mathbf{A}$ regroupe convection et diffusion, $\mathbf{B}_{\varrho^{n+1}}$ correspond à l'opposée de la divergence pondérée par ϱ^{n+1} et $\mathbf{B}^\mathrm{T}$ désigne le gradient discret (donc, en conséquence, $\mathbf{B}$ correspond à l'opposé de la divergence discrète). Le second membre $\mathbf{F}$ regroupe les contributions des forces réparties, des vitesses aux pas de temps précédents et des conditions aux limites (vitesses ou contraintes imposées), $\mathbf{G}$ correspond à la dérivée en temps discrète de la masse volumique et aux conditions aux limites (vitesses imposées), tandis que $\mathbf{H}$ ne contient que la contribution de ces dernières.

Une des méthodes fréquemment utilisées pour la résolution du système correspondant dans le cas incompressible est la méthode de Lagrangien augmenté [10] ; nous l'extrapolons ici aux écoulements dilatables. Le principe de cet algorithme est d'ajouter un terme d'augmentation dans la première équation, puis de répéter jusqu'à convergence une séquence d'opérations consistant à résoudre cette première équation à pression fixée puis à corriger la pression.

Le terme de pénalisation est obtenu en pré-multipliant l'équation de bilan de masse par $\gamma\,\mathbf{B}^\mathrm{T}\mathbf{M}_\mathbf{p}^{-1}$, où γ est un réel strictement positif, appelé paramètre d'augmentation et ajusté suivant le problème, et $\mathbf{M}_\mathbf{p}$ est une matrice de normalisation, qui peut par exemple être obtenue à partir de la matrice de masse de pression par condensation sur la diagonale.

Cet algorithme s'écrit donc sous forme algébrique de la manière suivante, en notant k l'indice associé à chaque itération :

$$\left| \begin{array}{l} \dfrac{\beta_q}{\Delta t}\mathbf{M}_{\varrho^{n+1}}\mathbf{U}_{\mathrm{F}}^{k+1} + \gamma\ \mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}_{\varrho^{n+1}}\mathbf{U}_{\mathrm{F}}^{k+1} - \mathbf{G}) + \mathbf{A}\mathbf{U}_{\mathrm{F}}^{k+1} = \mathbf{F} - \mathbf{B}^{\mathrm{T}}\mathbf{P}^{k} \\ \mathbf{P}^{k+1} = \mathbf{P}^{k} + \gamma\ \mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}_{\varrho^{n+1}}\mathbf{U}_{\mathrm{F}}^{k+1} - \mathbf{G}) \end{array} \right.$$

La pression doit être initialisée, par exemple à la pression au pas de temps précédent ; la convergence est atteinte lorsque la pression n'évolue plus ou, de manière équivalente, lorsque le bilan de masse est satisfait, à une tolérance ϵ près donnée :

$$\|\mathbf{B}_{\varrho^{n+1}}\mathbf{U}_{\mathrm{F}}^{k+1} - \mathbf{G}\| < \epsilon$$

Une fois la vitesse de fin de pas $\mathbf{U}_{\mathrm{F}}$ obtenue, le débit est réactualisé :

$$\mathbf{M}\mathbf{Q}_{\mathbf{F}} = \mathbf{M}_{\varrho^{n+1}}\mathbf{U}_{\mathrm{F}} + \mathbf{H}$$

Méthode de projection-pénalité

Nous allons dans cette section construire la méthode de projection-pénalité en suivant pas à pas la démarche utilisée pour établir le schéma dans le formalisme semi-discret.

En premier lieu, comme dans toute méthode de projection, on découple les équations de bilan de masse et de quantité de mouvement en explicitant la pression dans cette dernière ; on obtient ainsi la relation suivante :

$$\left(\frac{\beta_q}{\Delta t}\mathbf{M}_{\varrho^{n+1}} + \mathbf{A}\right)\tilde{\mathbf{U}}_{\mathbf{F}} + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathrm{exp}} = \mathbf{F}$$

où $\mathbf{P}_{\mathrm{exp}}$ désigne la pression en début de pas de temps.

On obtient la première étape de l'algorithme en ajoutant à cette équation un terme de pénalisation, construit de la même manière que le terme d'augmentation introduit dans la section précédente :

$$\left(\frac{\beta_q}{\Delta t}\mathbf{M}_{\varrho^{n+1}} + \mathbf{A} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{B}_{\varrho^{n+1}}\right)\tilde{\mathbf{U}}_{\mathbf{F}} + \mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathrm{exp}} = \mathbf{F} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_{\mathbf{p}}^{-1}\mathbf{G}$$

Laissée sous forme de problème de Darcy, l'étape de projection s'écrit :

$$\left| \begin{array}{l} \dfrac{\beta_q}{\Delta t}(\mathbf{M}\mathbf{Q}_{\mathbf{F}} - \mathbf{M}_{\varrho^{n+1}}\tilde{\mathbf{U}}_{\mathbf{F}}) + \mathbf{B}^{\mathrm{T}}\Phi = 0 \\ \mathbf{B}\mathbf{Q}_{\mathbf{F}} = \mathbf{G} \end{array} \right.$$

En multipliant la première équation par $\mathbf{B}\,\mathbf{M}^{-1}$ puis en utilisant la contrainte $\mathbf{B}\mathbf{Q}_{\mathbf{F}} = \mathbf{G}$, il vient :

$$\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^{\mathrm{T}}\Phi = \frac{\beta_q}{\Delta t}\left(\mathbf{B}\mathbf{M}^{-1}\mathbf{M}_{\varrho^{n+1}}\tilde{\mathbf{U}}_{\mathbf{F}} - \mathbf{G}\right)$$

Malheureusement, l'utilisation de l'inverse de $\mathbf{M}$ est coûteuse en temps calcul car, dans le cas général, cette matrice de masse n'est pas diagonale. L'équation (IV.2.9)

suggère alors d'approcher $\mathbf{B}\mathbf{M}^{-1}\mathbf{M}_{\varrho^{n+1}}$ par $\mathbf{B}_{\varrho^{n+1}}$ et $\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^{\mathrm{T}}$ par $\mathbf{L}$ où $\mathbf{L}$ est l'opérateur associé à un problème de Poisson avec des conditions aux limites de Dirichlet homogènes sur $\partial\Omega_{\mathrm{N}}$ et de Neumann homogènes sur $\partial\Omega_{\mathrm{D}}$. Cette dernière substitution n'est possible, toutefois, que pour un choix restreint d'espaces d'approximation pour la pression, excluant ceux où aucune continuité n'est requise, même en moyenne ou ponctuellement, aux interfaces entre les éléments ; l'espace utilisé dans les tests numériques présentés ici, à savoir l'élément fini de Lagrange P1 classique, est inclus dans $\mathrm{H}^1(\Omega)$. Si l'on souhaite, pour des raisons de simplicité d'implémentation, garder les mêmes dimensions pour les vecteurs $\mathbf{P}$ et Φ, les conditions de Dirichlet peuvent être imposées par pénalisation :

$$\mathbf{L}_{ij} = \int_{\Omega} \nabla\varphi_i^p \cdot \nabla\varphi_j^p + \alpha \int_{\partial\Omega_N} \varphi_i^p \cdot \varphi_j^p, \qquad 1 \leq i,j \leq N_{\mathrm{dof}}^p$$

où les fonctions φ_i^p sont les fonctions de base de l'espace d'approximation de pression M_h, N_{dof}^p désigne le nombre de degrés de liberté de pression (soit la dimension de M_h) et le coefficient α est un réel positif vérifiant $\alpha \gg 1$. Enfin on réactualise la pression en fin de pas de temps :

$$\mathbf{P} = \mathbf{P}_{\mathrm{exp}} + \Phi + r\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}_{\varrho^{n+1}}\tilde{\mathbf{U}}_{\mathbf{F}} - \mathbf{G})$$

En résumé, l'algorithme décrivant un pas de temps de la méthode de projection-pénalité est le suivant :

$$\left|\begin{array}{l} \left(\dfrac{\beta_q}{\Delta t}\mathbf{M}_{\varrho^{n+1}} + \mathbf{A} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_p^{-1}\mathbf{B}_{\varrho^{n+1}}\right)\tilde{\mathbf{U}}_{\mathbf{F}} = -\mathbf{B}^{\mathrm{T}}\mathbf{P}_{\mathrm{exp}} + \mathbf{F} + r\mathbf{B}^{\mathrm{T}}\mathbf{M}_p^{-1}\mathbf{G} \\ \mathbf{L}\Phi = \dfrac{\beta_q}{\Delta t}\left(\mathbf{B}_{\varrho^{n+1}}\tilde{\mathbf{U}}_{\mathbf{F}} - \mathbf{G}\right) \\ \mathbf{M}\mathbf{Q}_{\mathbf{F}} = \mathbf{M}_{\varrho^{n+1}}\tilde{\mathbf{U}}_{\mathbf{F}} - \dfrac{\Delta t}{\beta_q}\mathbf{B}^{\mathrm{T}}\Phi \\ \mathbf{P} = \mathbf{P}_{\mathrm{exp}} + \Phi + r\mathbf{M}_{\mathbf{p}}^{-1}(\mathbf{B}_{\varrho^{n+1}}\tilde{\mathbf{U}}_{\mathbf{F}} - \mathbf{G}) \end{array}\right.$$

Remarque [De la nécessité de construire le terme de pénalisation de manière algébrique] - Une discrétisation "directe" du terme écrit en semi discret conduirait à la forme bilinéaire $c_{\mathrm{pen}}(\cdot,\cdot)$ sur $V_h \times V_h$ suivante :

$$c_{\mathrm{pen}}(u_h, v) = r\int_{\Omega}\left(\nabla\cdot\varrho u_h - \frac{D\varrho}{\Delta t}\right)\ \nabla\cdot v$$

Or la contrainte :

$$\nabla\cdot\varrho u_h - \frac{D\varrho}{\Delta t} = 0$$

n'est imposée à la solution discrète qu'au sens faible. Il semble donc naturel, pour ne pas voir la méthode perdre en précision pour les fortes valeurs du paramètre de pénalisation, de préférer une formulation qui utilise la contrainte réellement imposée, c'est à dire sa forme discrète ; c'est exactement ce qui est fait ici. Il est à noter que la formulation semi-discrète du terme de pénalisation :

$$\nabla\left(\nabla\cdot\varrho u_h - \frac{D\varrho}{\Delta t}\right)$$

est en outre incorrecte, du fait que l'intégration par partie conduisant à la forme $c_{\mathrm{pen}}(\cdot,\cdot)$ ferait apparaître des termes de bord sur $\partial\Omega_{\mathrm{N}}$ qui n'ont pas lieu d'être et qui, de fait, n'ont pas leurs pendants dans la pénalisation algébrique utilisée.

IV.2.4 Expérimentations numériques

Nous effectuons dans cette section une comparaison entre les méthodes introduites précédemment, à savoir la méthode linéairement implicite, la méthode de projection-pénalité et la méthode de projection incrémentale standard (obtenue à partir de la précédente en faisant $r = 0$), en les appliquant à un problème possédant une solution analytique.

Dans un premier temps, nous choisissons des conditions aux limites de type Dirichlet, c'est à dire que nous traitons le problème suivant :

$$\left|\begin{array}{ll} \dfrac{\partial \varrho u}{\partial t} + \nabla \cdot (\varrho u \otimes u) = \mu \Delta u - \nabla p + f & \text{dans } [0,1] \times \Omega \\ \nabla \cdot \varrho u + \dfrac{\partial \varrho}{\partial t} = 0 & \text{dans } [0,1] \times \Omega \\ u = u_{\mathrm{D}} & \text{sur } \partial\Omega, \ \forall t \in [0,1] \\ u = u_0 & \text{dans } \Omega, \ \text{à } t = 0 \end{array}\right.$$

où $\Omega =]0,1[\times]0,1[$ et f, u_{D} et u_0 sont donnés et tels que les champs de vitesse, pression et masse volumique suivants soient solution du problème :

$$u(x,y,t) = \begin{pmatrix} 0.5y\,(1-y)\,(2+\cos(2\pi t)) \\ 0 \end{pmatrix}$$

$$p(x,y,t) = -\mu\,(2+\cos(2\pi t))\,(x-0.5)$$

$$\varrho(x,y,t) = 1 + \frac{(X(x,y,t)-x_0)^2(-2X(x,y,t)+3x_1-x_0)}{(x_1-x_0)^3}$$

$$X(x,y,t) = x - 0.5\,y\,(1-y)\,(2t + \frac{\sin(2\pi t)}{2\pi}), \quad x_0 = 0.2 \text{ et } x_1 = 0.8$$

L'expression de la masse volumique est obtenue par la méthode des caractéristiques, en remarquant que le champ de vitesse choisi est solénoïdal, et que l'équation de bilan de masse dégénère donc en équation de transport ; cette propriété du champ de vitesse, spécifique à cette solution particulière, ne joue aucun rôle ici.

La viscosité est constante et fixée à $\mu = 1$ (*i.e.* $\mathrm{Re} \approx 1$) ; la diffusion est donc dominante, ce qui est la situation la plus pénalisante pour ce qui concerne le traitement par des méthodes de correction de pression de conditions aux limites ouvertes, cas que nous traiterons ultérieurement.

Nous utilisons des éléments finis P2 pour les champs de vitesse et de débit et des éléments finis P1 pour le champ de pression ; la stabilité de cette discrétisation, dite élément de Taylor-Hood, pour les problèmes incompressibles est classique (*c.f.* par exemple [11]), elle est étendue dans le cas dilatable dans [4]. Le maillage en triangles est obtenu en construisant tout d'abord une grille uniforme de pas $1/40$, puis en découpant chaque carré de la partition selon ses diagonales pour former 4 triangles isocèles. La discrétisation en temps est effectuée par une formule de différentiation rétrograde d'ordre deux.

Les différences entre la solution obtenue à $t = 1$ et la solution analytique en fonction du pas de temps sont tracées sur les figures IV.1, IV.2, IV.3, IV.4, IV.5. Ces

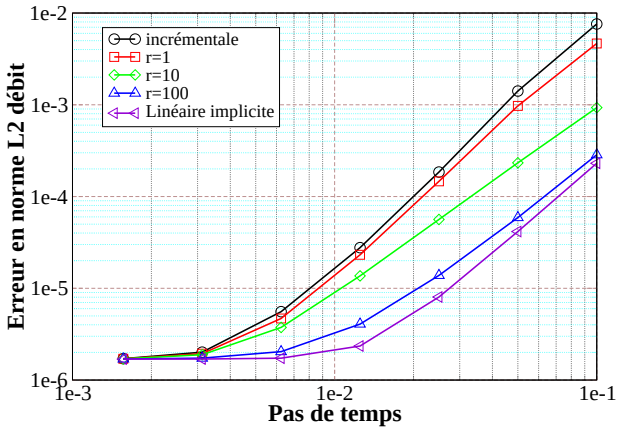


FIG. IV.1 – Cas test avec des conditions aux limites de type Dirichlet – Norme L^2 de l'erreur pour le débit à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

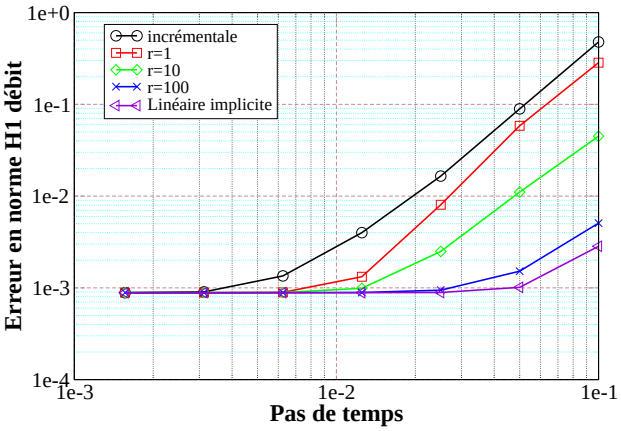


FIG. IV.2 – Cas test avec des conditions aux limites de type Dirichlet – Norme H^1 de l'erreur pour le débit à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

courbes ont l'allure usuelle : tout d'abord décroissance avec le pas de temps puis atteinte d'un plateau qui correspond à l'erreur d'approximation spatiale ; ce dernier phénomène n'est observé de manière marquée que pour le débit, ce qui est très probablement dû au fait que vitesse et pression appartiennent tous deux à leur espace d'approximation respectif. Tous les schémas étudiés montrent une convergence en temps approximativement d'ordre deux. Pour la méthode de projection usuelle, l'erreur de fractionnement (*i.e.* la différence entre la solution obtenue par la méthode de projection et la méthode linéairement implicite) est dominante à fort pas de temps : pour le débit par exemple, on constate que les erreurs associées à ces deux méthodes diffèrent d'un facteur 100. Cette erreur de fractionnement diminue avec le paramètre de pénalisation r, jusqu'à ce que les solutions coïncident pour r grand (typiquement, $r \geq 10^2$). Ces résultats sont en accord avec ce qui est observé [17] ou démontré [1]

pour le cas incompressible. Enfin, la comparaison des résultats obtenus pour divers maillages (non présentée ici) montre que l'erreur résiduelle obtenue pour le débit à faible pas de temps varie comme l'erreur d'approximation, à savoir comme h^2 pour la norme H^1 et h^3 pour la norme L^2, où h est le pas de maillage.

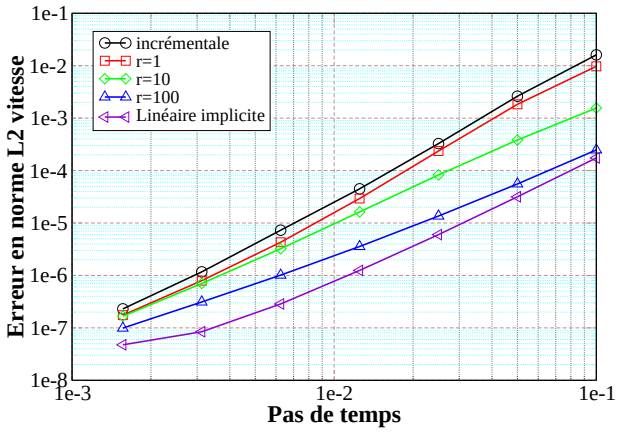


FIG. IV.3 – Cas test avec des conditions aux limites de type Dirichlet – Norme L^2 de l'erreur pour la vitesse à $t = 1$ en fonction du pas de temps pour les méhodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

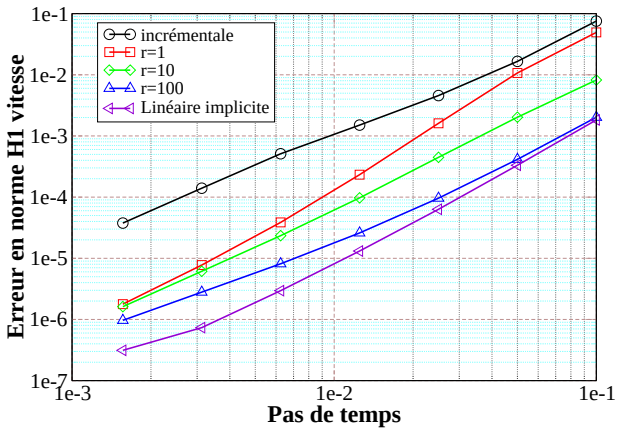


FIG. IV.4 – Cas test avec des conditions aux limites de type Dirichlet – Norme H^1 de l'erreur pour la vitesse à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

Nous changeons maintenant de conditions aux limites qui deviennent :

$$\left| \begin{array}{ll} \sigma \cdot \vec{n} = \begin{pmatrix} -p\,(1,y,t) \\ 0 \end{pmatrix} & \text{sur } \{1\} \times [0;1] \\ u = u(x,y,t) & \text{sur les autres frontières.} \end{array} \right.$$

Les résultats dans ce cas font l'objet des figures IV.6, IV.7, IV.8, IV.9, IV.10.

Dans le contexte des écoulements incompressibles, la méthode de projection incré-

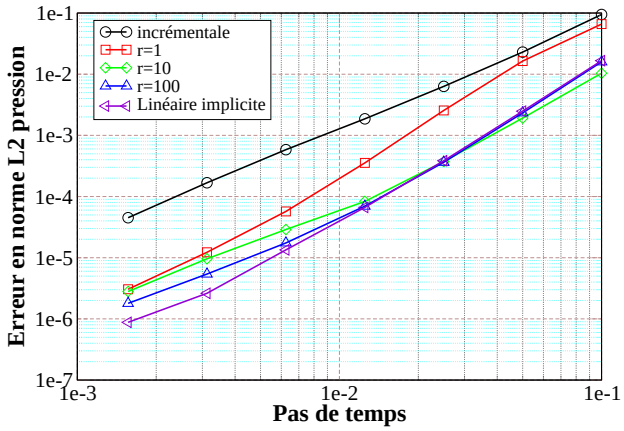


FIG. IV.5 – Cas test avec des conditions aux limites de type Dirichlet – Norme L^2 de l'erreur pour la pression à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

mentale est connue pour perdre, en présence de conditions aux limites de Neumann (dites également "conditions aux limites ouvertes") comme c'est le cas ici, ses propriétés de convergence. Ce phénomène est lié au fait que l'on impose, pour des raisons de stabilité, des conditions aux limites de Dirichlet à l'inconnue de l'étape de projection ϕ, et donc, par récurrence, compte-tenu de la relation donnant l'incrément de pression pour la méthode incrémentale, à la pression elle-même. On observe alors [15, 17] une réduction des ordres de convergence en temps et en espace, qui prennent, probablement par hasard, les mêmes valeurs : approximativement 1 pour la vitesse en norme L^2 et $1/2$ pour la vitesse en norme H^1 et la pression en norme L^2. Ce même comportement semble se reproduire ici puis, pour les valeurs les plus faibles du pas de temps, une accélération de la convergence survient. Une étude paramétrique au pas de maillage h semble montrer que le pas de temps où survient la rupture de pente dans la courbe de convergence décroît linéairement avec h, ce qui suggère l'explication suivante à ce phénomène. En exploitant des techniques d'inégalité inverse [8], il est possible de démontrer des relations du type :

$$\|u - u_h\|^\star \leq c(h) \, (\|u - u_h\|_s + I_{h\star}) + I_h^\star \tag{IV.2.17}$$

où u est une fonction de l'espace continu, u_h une fonction de l'espace d'approximation, $\|\cdot\|_s$ et $\|\cdot\|^\star$ sont deux normes, la première étant plus faible que l'autre (par exemple, norme H^{-1} et norme L^2), $I_{h\star}$ et $I_h^\star$ désignent l'erreur d'interpolation respectivement en norme $\|\cdot\|_s$ et en norme $\|\cdot\|^\star$. Imaginons que l'erreur en norme $\|\cdot\|_s$ décroisse plus vite avec le pas de temps que l'erreur en norme $\|\cdot\|^\star$, par exemple selon les estimations suivantes (avec $\alpha < \beta$) :

$$\|u - u_h\|^\star \leq c \, (\Delta t^\alpha + I_h^\star), \qquad \|u - u_h\|_s \leq c \, (\Delta t^\beta + I_{h\star}) \tag{IV.2.18}$$

avec $\beta > \alpha$. L'exploitation de l'inégalité [IV.2.17] fournit alors une relation de la forme :

$$\|u - u_h\|^\star \leq c'(h) \, (\Delta t^\beta + I_{h\star}) + I_h^\star$$

et, surtout lorsque l'erreur d'interpolation est faible, il peut exister un pas de temps seuil, dépendant de h, en deçà duquel cette dernière estimation est plus précise que

(IV.2.18) ; la courbe d'erreur pourrrait alors présenter une rupture de pente (pente α au dessus du pas de temps seuil, pente β au dessous).

Il est à noter que de tels résultats sont classiques pour les méthodes de projection (décroissance de l'erreur de vitesse en norme L^2 plus rapide qu'en norme H^1, estimation d'ordre supérieur pour la pression dans une norme dépendant du maillage [13], ...). C'est d'ailleurs ce que l'on constate ici si l'on s'en tient à la première partie de la courbe (*i.e.* à fort pas de temps), puisque débit et vitesse présentent une convergence d'ordre 2 en norme L^2, tandis qu'en norme H^1, l'ordre tombe à 1 pour le débit et 1/2 pour la vitesse ; l'ordre de convergence de la pression en norme L^2 est également 1/2.

Pour les méthodes autres que la projection incrémentale, les propriétés de convergence observées sont, avec des conditions aux limites ouvertes, les mêmes qu'avec des conditions de Dirichlet.

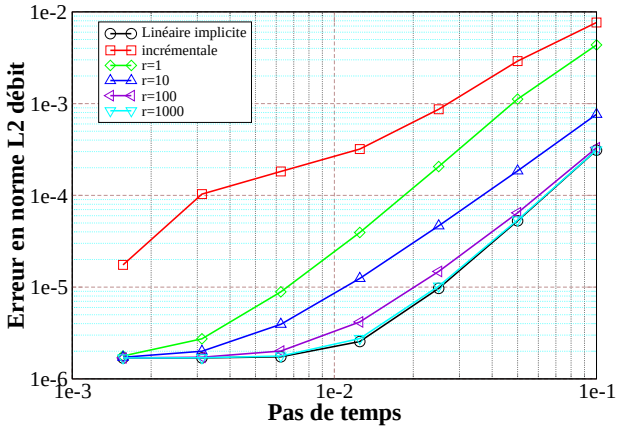


FIG. IV.6 – Cas test avec des conditions aux limites ouvertes – Norme L^2 de l'erreur pour le débit à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

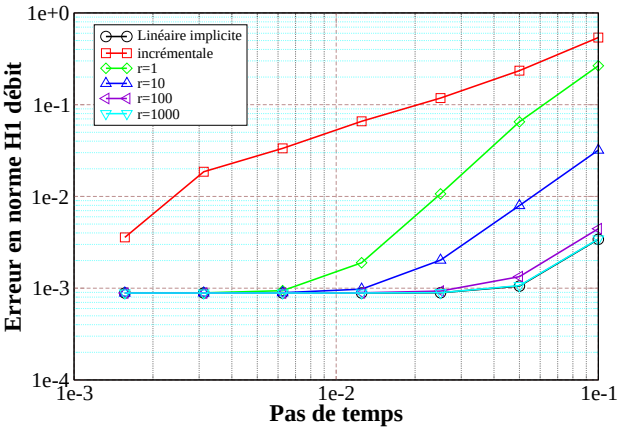


FIG. IV.7 – Cas test avec des conditions aux limites ouvertes – Norme H^1 de l'erreur pour le débit à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

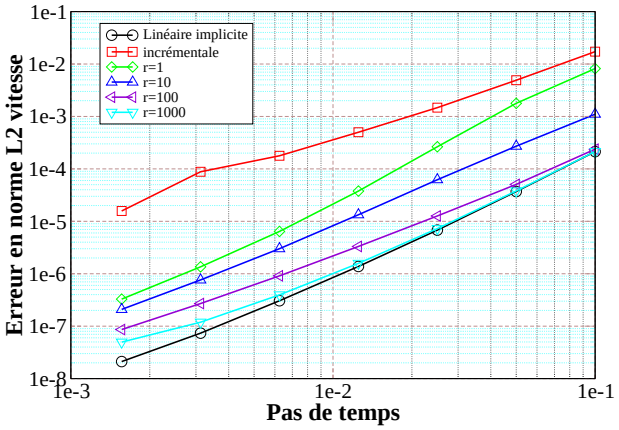


FIG. IV.8 – Cas test avec des conditions aux limites ouvertes – Norme L^2 de l'erreur pour la vitesse à $t = 1$ en fonction du pas de temps pour le méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

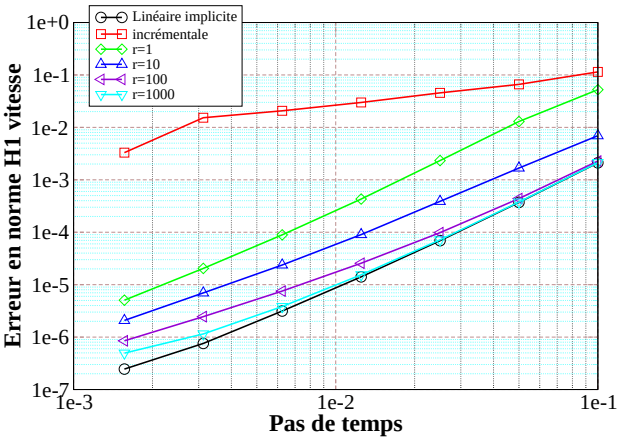


FIG. IV.9 – Cas test avec des conditions aux limites ouvertes – Norme H^1 de l'erreur pour la vitesse à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

En conclusion, le fait de pénaliser la première étape rend l'algorithme de projection nettement plus précis. Malheureusement, il semble que ce soit au prix d'une détérioration sensible du conditionnement de la matrice associée à l'étape de prédiction de vitesse ; ce comportement avait été également observé pour les écoulements incompressibles [17]. Il reste qu'à coût de calcul égal, dans les cas que nous avons étudiés, la méthode de projection-pénalité avec un paramètre de pénalisation raisonnable (typiquement, $r = 10$) reste plus précise que la méthode de projection incrémentale.

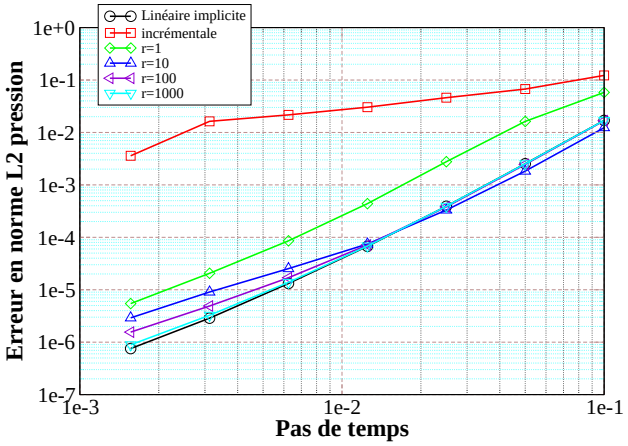


FIG. IV.10 – Cas test avec des conditions aux limites ouvertes – Norme L^2 de l'erreur pour la pression à $t = 1$ en fonction du pas de temps pour les méthodes de projection incrémentale, projection-pénalité ($r = 1$, $r = 10$, $r = 100$) et linéairement implicite.

IV.3 Un cas de convection naturelle à faible nombre de Mach

IV.3.1 Position du problème

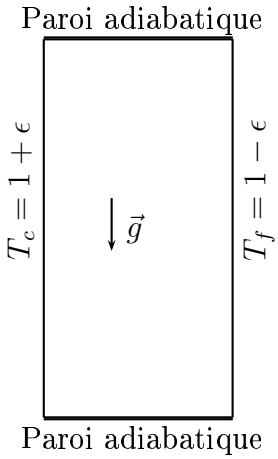
Le système d'équations aux dérivées partielles que nous étudions dans cette section est un modèle asymptotique établi à partir des équations de la dynamique des fluides compressibles en faisant l'hypothèse que la vitesse dans l'écoulement reste faible devant la vitesse des ondes de pression (vitesse du son) [20] ; le nombre de Mach est défini comme le rapport entre la vitesse matérielle et la vitesse du son et un tel écoulement est dit "à faible nombre de Mach". Ce système d'équations s'écrit sous forme adimensionnée de la manière suivante :

$$\left|\begin{array}{l} \dfrac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho u) = 0 \\ \dfrac{\partial \varrho u}{\partial t} + \nabla \cdot (\varrho u \otimes u) = \dfrac{1}{\mathrm{Re}} \left[\nabla \cdot (\mu(\nabla u + (\nabla u)^{\mathrm{T}})) - \dfrac{2}{3} \nabla (\mu \nabla \cdot u) \right] \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad -\nabla p + \dfrac{1}{\mathrm{Fr}^2} \varrho(-\vec{z}) \\ \varrho \left(\dfrac{\partial T}{\partial t} + u \cdot \nabla T \right) = \dfrac{1}{\mathrm{Re}\,\mathrm{Pr}} \nabla \cdot (\lambda \nabla T) + \dfrac{\gamma - 1}{\gamma} \dfrac{dP_{th}}{dt} \\ P_{th} = \varrho T \end{array}\right. \qquad \text{(IV.3.19)}$$

où T désigne la température, P_{th} est une quantité dépendant seulement du temps (*i.e.* constante en espace) dite pression thermodynamique, dont le calcul nécessite, dans un domaine fermé, une équation supplémentaire qui peut être simplement la conservation de la masse totale. Le nombres sans dimension intervenant dans ce système sont les nombres de Froude, de Reynolds, de Prandlt :

$$\mathrm{Fr} = \frac{u_0}{L_0 g} \qquad \mathrm{Re} = \frac{\varrho_0 u_0 L_0}{\mu_0} \qquad \mathrm{Pr} = \frac{\mu_0 C_p}{\lambda_0}$$

Du fait que la masse volumique ne dépend dans ce modèle que de la pression thermodynamique et plus de la pression dynamique (*i.e.* de la pression qui apparaît dans le bilan de quantité de mouvement), ce système d'équation aux dérivées partielles entre dans le cadre défini en introduction (système (IV.1.1)).



Le cas que nous traitons ici est un écoulement de convection naturelle dans une cavité différentiellement chauffée, de rapport d'aspect H/L égal à 4 (soit $\Omega = (0,1) \times (0,4)$).
Les nombres adimensionnels caractéristiques de l'écoulement prennent les valeurs suivantes :

$$\mathrm{Re} = 10^3,\ \mathrm{Fr} = 0.923,\ \mathrm{Pr} = 0.71,\ \gamma = 1.4$$

Des conditions aux limites de Dirichlet homogènes sont imposées à la vitesse tout le long de la frontière. En ce qui concerne le bilan d'énergie, les frontières inférieure et supérieure sont supposées adiabatiques (*i.e.* condition de Neumann homogène), tandis que la température est imposée sur les frontières verticales aux valeurs suivantes : $T_c = 1 + \epsilon$ sur la droite $x = 0$, $T_f = 1 - \epsilon$ sur la droite $x = 1$, ϵ étant un paramètre fixé à 0.6.

Le comportement de la viscosité dynamique et de la conductivité thermique adimensionnées est régi par la loi de Sutherland :

$$\mu(T) = \lambda(T) = T^{\frac{3}{2}} \left(\frac{1+S}{T+S} \right)$$

avec $S = 1.1/6$.

Ce problème reprend exactement les données du benchmark organisé par P. Le Queré et H. Paillère [18], à ceci près que la géométrie est modifiée ($\Omega = (0,1) \times (0,1)$ dans [18]). Par contre, alors que dans [18], l'objectif était le calcul de l'état stationnaire, nous nous intéressons ici essentiellement au transitoire d'établissement depuis l'état initial défini par $u = 0$ et $T = 1$.

IV.3.2 Implémentation éléments finis

Le schéma numérique utilisé pour la résolution de ce problème résulte du simple rajout d'une étape de résolution du bilan d'énergie aux algorithmes introduits et testés dans les sections précédentes. Il s'écrit de la manière suivante :

Trouver $T^{n+1},\ \tilde{u}^{n+1},\ \phi,\ p^{n+1},\ q^{n+1} \in T_h \times V_h \times M_h \times M_h \times V_h$, P_{th}^{n+1} et ϱ^{n+1} tels que :

$$
\left|
\begin{array}{ll}
(\dfrac{\varrho^n DT^{n+1}}{\Delta t} + q^{\star,n+1}\cdot\nabla T^{n+1}, t) + (\dfrac{\lambda}{\mathrm{RePr}}\nabla T^{n+1}, \nabla t) & \\
\qquad\qquad = (\dfrac{\gamma-1}{\gamma}\dfrac{DP_{th}^n}{\Delta t}, t) & \forall t \in T_h' \\
(\displaystyle\int_\Omega \frac{1}{T^{n+1}})\ P_{th}^{n+1} = (\int_\Omega \frac{1}{T^n})\ P_{th}^n \qquad \varrho^{n+1} = \dfrac{P_{th}^{n+1}}{T^{n+1}} & \\
(\dfrac{\varrho^{n+1}\tilde{u}^{n+1} - \sum_{j=0}^{q-1}\beta_j q_h^{n-j}}{\Delta t}, v) + (\tau(\tilde{u}^{n+1}), \nabla v) & \\
\qquad +(\nabla\cdot(q^{\star,n+1}\otimes\tilde{u}^{n+1}), v) = (\dfrac{1}{\mathrm{Fr}^2}\varrho^{n+1}(-\vec{z}), v) & \forall v \in V_h \\
(\nabla\phi, \nabla t) = (\nabla\cdot\varrho^{n+1}\tilde{u}^{n+1} + \dfrac{D\varrho^{n+1}}{\Delta t}, t) & \forall t \in M_h \\
(p^{n+1}, t) = (p^n + \phi, t) & \forall t \in M_h \\
(q^{n+1}, v) = (\varrho^{n+1}\tilde{u}^{n+1} + \dfrac{\Delta t}{\beta_q}\nabla\phi, v) & \forall v \in V_h
\end{array}
\right.
$$

où T_h et T_h' désignent respectivement l'espace d'approximation de la température et l'espace des fonctions tests associés, tous deux ne différant que par les conditions aux limites de Dirichlet prises en compte (non-homogènes pour le premier, homogènes pour l'autre).

La discrétisation en temps est d'ordre deux. Pour des raisons de précision et de stabilité, les calculs sont effectués à pas de temps suffisamment petit pour que l'erreur de fractionnement reste faible, si bien qu'il n'est pas nécessaire de pénaliser la première étape ; nous ne décrivons donc ici que les résultats obtenus avec la méthode de projection incrémentale.

Pour l'approximation de la vitesse et de la pression, nous utilisons l'élément fini de Taylor-Hood ; la température est approchée par des éléments finis simpliciaux de Lagrange de degré 2, identiques à ceux utilisés pour chacune des composantes de la vitesse, ce qui correspond au choix recommandé dans [5]. Dans les différentes étapes de méthode de projection, la masse volumique est évaluée à chaque point d'intégration en appliqant la loi d'état.

IV.3.3 Résultats

Le domaine de calcul est triangulé en construisant tout d'abord une grille structurée puis en découpant chacun des rectangles obtenus en deux le long de la diagonale montante. Le pas de la grille est uniforme dans la direction verticale. Dans la direction horizontale, il croît avec une raison géométrique jusqu'au centre et re-décroît

ensuite de manière symétrique ; les plus petites mailles, situées au contact de la frontière du domaine, sont deux fois plus fines que les plus grosses situées au centre.

Trois calculs sont effectués, en utilisant les maillages et pas de temps suivants :

	grille	$h_{\mathrm{x,min}}$	$h_{\mathrm{x,max}}$	ddl P_2	ddl P_1	Δt
F	240×320	0.003	0.006	305757	76719	$1.25\,10^{-3}$
M	160×320	0.0044	0.0088	205761	51681	$2\,10^{-3}$
G	120×240	0.006	0.012	115921	29161	$2.5\,10^{-3}$

où les nombres "ddl P_2" et "ddl P_1" désignent les nombres de degrés de liberté pour chacun des champs discrétisés respectivement avec des éléments de degré 2 (chacune des composantes de la vitesse et la température) et avec des éléments linéaires (pression). Pour le calcul le plus fin (calcul "F"), le nombre total de degrés de liberté avoisine ainsi 10^6.

Tous les calculs convergent vers un état permanent, après un transitoire d'établissement assez long (plus de 200 unités de temps), marqué par des instabilités au voisinage de la couche limite chaude. Ces instabilités sont illustrées par les figures IV.11 et IV.12, où l'on trace respectivement les courbes isothermes à quelques instants consécutifs et l'évolution de la température en trois points de la couche limite chaude. De manière qualitative, ces instabilités correspondent à des détachements de poches froides de la zone inférieure de la cavité. Les évolutions temporelles ne présentent aucun caractère de périodicité. On observe sur la figure IV.12 que les trois calculs se superposent, ce qui semble démontrer que la convergence au maillage et au pas de temps est atteinte.

La distribution de température à $t = 25.5$ est tracée sur la figure IV.13 ; on constate la finesse des couches limites ainsi que le caractère perturbé de la couche limite chaude, tandis que la couche limite froide reste régulière. Le nombre de Nusselt sur la paroi chaude à la même date est donné sur la figure IV.14.

Enfin, la figure IV.13 représente les lignes de courant à la même date, ou, plus précisément, les isovaleurs de la fonction ϕ solution de :

$$-\Delta\phi = \frac{\partial v_x}{\partial y} - \frac{\partial v_y}{\partial x}$$

avec des conditions aux limites homogènes, qui coinciderait avec le fonction de courant si le champs était à divergence nulle. Encore une fois, on constate l'extrême complexité de l'écoulement.

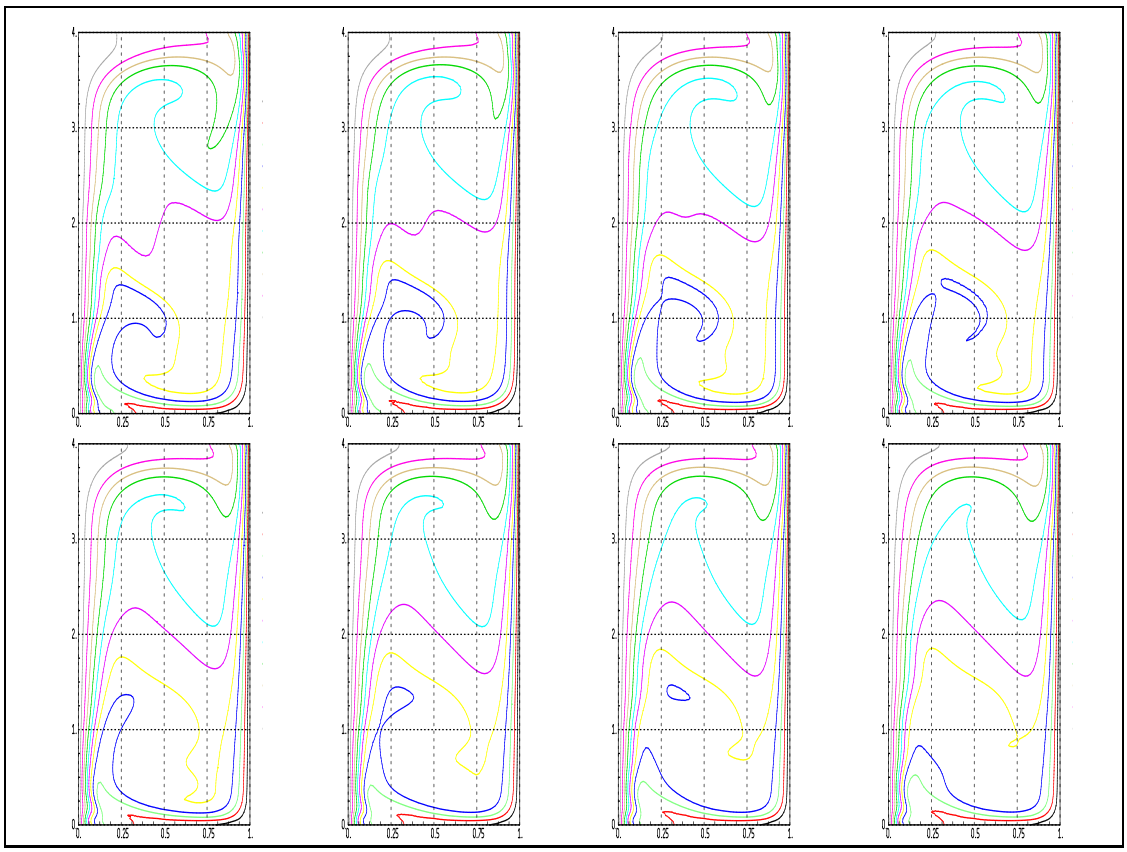


FIG. IV.11 – Problème de convection naturelle – Courbes isothermes de t=13 à t=20 par pas de 1.

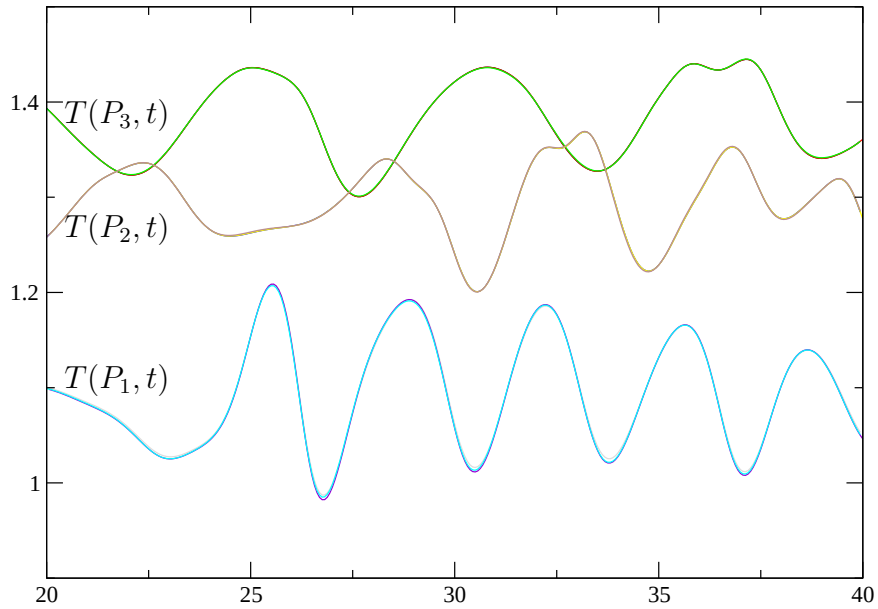


FIG. IV.12 – Problème de convection naturelle – Evolution de la température en trois points de la couche limite chaude pour les trois calculs – $P_1 = (0.065, 1)$, $P_2 = (0.065, 2)$, $P_3 = (0.065, 3)$.

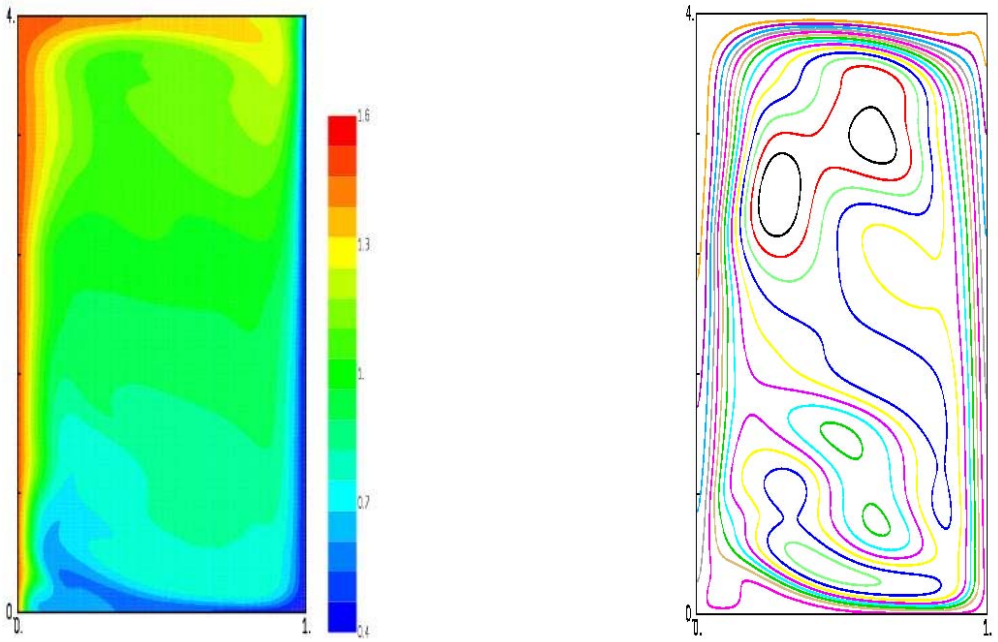


FIG. IV.13 – Problème de convection naturelle – Distribution de température (à gauche) et lignes de courant (à droite) à $t = 25.5$.

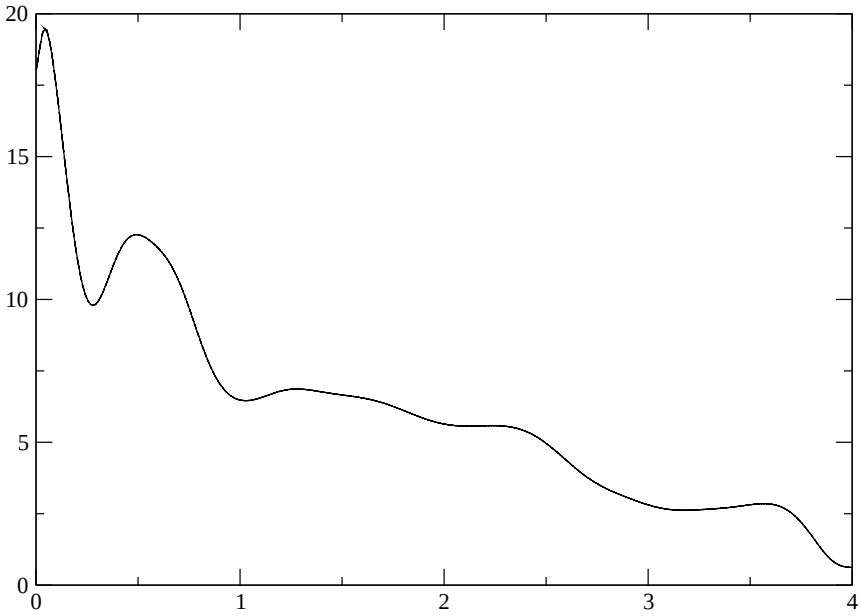


FIG. IV.14 – Problème de convection naturelle – Nombre de Nusselt local sur la paroi chaude à $t = 25.5$.

Bibliographie

[1] Ph. Angot, M. Jobelin, and J.-C. Latché. Error analysis of the penalty-projection method for the time-dependent Stokes equations, 2006.

[2] Aquilon. Aquilon, outil de modélisation en mécanique des fluides et transferts. http ://www.trefle.u-bordeaux1.fr/aquilon/index.html.

[3] F. Babik, T. Gallouët, J.-C. Latché, S. Suard, and D. Vola. On some fractional step schemes for combustion problems. In *Finite Volumes for Complex Applications IV (FVCA IV)*. Éditions Hermès, Paris, 2005.

[4] Christine Bernardi, Frédéric Laval, Brigitte Métivet, and Bernadette Pernaud-Thomas. Finite element approximation of viscous flows with varying density. *SIAM Journal on Numerical Analysis*, 29 :1203–1243, 1992.

[5] Christine Bernardi, Brigitte Métivet, and Bernadette Pernaud-Thomas. Couplage des équations de Navier-Stokes et de la chaleur : le modèle et son approximation par éléments finis. *Mathematical Modelling and Numerical Analysis*, 29(7) :871–921, 1995.

[6] Jean-Paul Caltagirone and Jérôme Breil. Sur une méthode de projection vectorielle pour la résolution des équations de Navier-Stokes. *Comptes-Rendus de l'académie des Sciences, Paris – Série II*, 327 :1179–1184, 1999.

[7] Alexandre Joel Chorin. Numerical solution of the Navier-Stokes equations. *Mathematics of Computation*, 22 :745–762, 1968.

[8] P.G. Ciarlet. *Finite Elements Methods – Basic Error Estimates for Elliptic Problems*, volume II of *Handbook of Numerical Analysis*. North-Holland, 1991.

[9] A. Ern and J.L. Guermond. *Eléments finis : théorie, applications, mise en œuvre*, volume 36 of *Mathématiques & Applications*. Springer, 2002.

[10] M. Fortin and R. Glowinski. *Méthodes de Lagrangien Augmenté*. Dunod, Paris, 1982.

[11] Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms.*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, 1986.

[12] Katuhiko Goda. A multistep technique with implicit difference schemes for calculating two- or three-dimensional cavity flows. *Journal of Computational Physics*, 30 :76–95, 1979.

[13] J.-L. Guermond. Un résultat de convergence d'ordre deux en temps pour l'approximation des équations de Navier-Stokes par une technique de projection incrémentale. *Mathematical Modelling and Numerical Analysis*, 33(1) :169–189, 1999.

[14] J.-L. Guermond and J. Shen. On the error estimates for the rotational pressure-correction projection methods. *Mathematics of Computation*, 73(248) :1719–1737, 2003.

[15] J.L. Guermond, P. Minev, and J. Shen. Error analysis of pressure-correction schemes for the time-dependent Stokes equations with open boundary conditions. *SIAM Journal on Numerical Analysis*, 43 :239–258, 2005.

[16] J.L. Guermond, P. Minev, and J. Shen. An overview of projection methods for incompressible flows, 2006.

[17] M. Jobelin, C. Lapuerta, J.-C. Latché, Ph. Angot, and B Piar. A finite element penalty-projection method for incompressible flows. *Journal of Computational Physics*, 217 :502–518, 2006.

[18] P. Le Quéré, C. Weisman, H. Paillère, J. Vierendeels, E. Dick, R. Becker, M. Braack, and J. Locke. Modelling of natural convection flows with large temperature differences : A benchmark problem for low mach number solvers. part 1. reference solutions. *Mathematical Modelling and Numerical Analysis*, 39(3) :609–616, 2005.

[19] A. Majda and J. Sethian. The derivation and numerical solution of the equations for zero Mach number solution. *Combustion Science and Techniques*, 42 :185–205, 1985.

[20] B. Müller. Low Mach number asymptotics of the Navier-Stokes equations and numerical implications. In *30th Computational Fluid Dynamics, March 8-12 1999*, number 1999-03 in Lecture Series. von Karman Institute for Fluid Dynamics, 1999.

[21] Jie Shen. On error estimates of some higher order projection and penalty-projection methods for Navier-Stokes equations. *Numerische Mathematik*, 62 :49–73, 1992.

[22] R. Temam. Sur l'approximation de la solution des Équations de Navier-Stokes par la méthode des pas fractionnaires (II). *Archive for Rational Mechanics and Analysis*, 33 :377–385, 1969.

[23] L.J.P. Timmermans, P.D. Minev, and F.N. Van de Vosse. An approximate projection scheme for incompressible flow using spectral elements. *International Journal for Numerical Methods in Fluids*, 22 :673–688, 1996.

[24] J. Van Kan. A second-order accurate pressure-correction scheme for viscous incompressible flow. *SIAM Journal on Scientific and Statistical Computing*, 7(3) :870–891, 1986.

Chapitre V

Méthode d'éléments finis joints

V.1 Introduction

La méthode d'éléments finis joints [5, 2, 3, 4, 10, 9, 6] est une méthode de décomposition de domaine sans recouvrement pour la résolution d'équations aux dérivées partielles. Cette méthode offre une large gamme de possibilités :
- les discrétisations des sous-domaines sont indépendantes ; en particulier, les maillages de part et d'autre de l'interface ne coïncident pas dans le cas général ;
- les problèmes différentiels sur chaque sous-domaine peuvent être, dans un même calcul, résolus par des techniques différentes (éléments finis, volumes finis, méthodes spectrales...) ;
- les problèmes différentiels posés sur chaque sous-domaine peuvent ne pas être identiques ;
- elle est adaptée au calcul parallèle, dans la mesure où la continuité au sens faible est suffisante aux interfaces des sous-domaines, ce qui limite les communications entre les sous-domaines ;
- elle peut être utilisée dans des résolutions avec raffinement local.

Parmi les potentialités offertes, nous pouvons citer par exemple le couplage sans modification de logiciels et/ou de discrétisations existants. Des gains importants peuvent également être attendus en matière d'optimisation de coût de calcul, du fait de la possibilité :
- de concaténer des maillages structurés (obtention des propriétés de superconvergence associées à ce type de discrétisation) et non-structurés (traitement des singularités géométriques) ;
- de raffiner certaines zones du maillage en conservant le maillage dans les autres zones.

Dans ce chapitre, nous débutons par une présentation rapide de la méthode, puis nous la mettons en œuvre :
- à des fins de raffinement local, pour le traitement, par l'assemblage de grilles structurées de pas différents, d'un problème elliptique présentant une singularité à l'origine,
- puis pour le traitement d'un problème multi-physique, à savoir un écoulement de convection naturelle couplé à de la conduction dans les parois, avec des maillages de la zone d'écoulement et de la paroi se raccordant de manière non conforme.

V.2 Présentation des éléments finis joints

Le choix de présentation effectué ici est de décrire la méthode des éléments finis joints sur un problème modèle, à savoir un problème de réaction-advection-diffusion. La formulation multi-domaines est tout d'abord introduite pour le problème continu ; la méthode d'éléments finis joints est ensuite obtenue par une discrétisation directe de cette formulation.

Le problème considéré ici s'écrit comme suit :

Trouver T tel que

$$\left|\begin{array}{ll} -\nabla\cdot\lambda(x)\nabla T(x)+\vec{u}_{adv}\cdot\nabla T(x)+\alpha T(x)=f(x) & \text{dans } \Omega \\ \lambda(x)\dfrac{\partial T(x)}{\partial n}=g(x) & \text{sur } \partial\Omega_{\mathrm{N}} \\ T(x)=0 & \text{sur } \partial\Omega_{\mathrm{D}} \end{array}\right.$$

Le domaine de calcul Ω est un ouvert polygonal de $\mathbb{R}^d$; sa frontière est découpée en deux parties disjointes $\partial\Omega_{\mathrm{D}}$ et $\partial\Omega_{\mathrm{N}}$, sur lesquelles sont appliquées respectivement des conditions aux limites de type Dirichlet et Neumann ; $\partial\Omega_{\mathrm{D}}$ est suposée de mesure non nulle. Le scalaire $\lambda(x)$ est un scalaire strictement positif, vérifiant $\lambda(x)\geq\lambda_0>0$, $\vec{u}_{adv}$ est une vitesse d'advection donnée, α un scalaire strictement positif, f un terme source et g un flux appliqué aux frontières de Neumann ; toutes ces quantités sont, pour le moment, supposées régulières et la divergence de la vitesse $\vec{u}_{adv}$ est nulle.

Soit une décomposition du domaine Ω sans recouvrement en K sous-domaines Ω_k, $k=1,\ldots,K$. Les sous domaines Ω_k sont tous supposés polygonaux. On a donc :

$$\overline{\Omega}=\bigcup_{k=1}^{K}\overline{\Omega}_k,\quad \Omega_k\cap\Omega_l=\varnothing \text{ si } k\neq l$$

Les interfaces communes à deux sous-domaines Ω_k et Ω_l sont notées $\overline{\Gamma}_{kl}=\overline{\Omega}_k\cap\overline{\Omega}_l$ et la frontière externe de Ω_k, *i.e.* la partie de $\overline{\Omega}_k$ commune avec le bord du domaine Ω est notée $\partial\Omega_{k,\mathrm{D}}$; la réunion des interfaces $\overline{\Gamma}_{kl}$ et des frontières externes de $\partial\Omega_{k,\mathrm{D}}$ forme la frontière de Ω_k, notée $\partial\Omega_k$. Le squelette associé à cette décomposition est défini comme la réunion des interfaces des sous-domaines :

$$S=\bigcup_{k,l}\Gamma_{kl}$$

L'étude est réalisée dans le cadre de décompositions de domaine conformes, ce qui signifie que l'intersection de deux sous-domaines $\Omega_k\cap\Omega_l$ est soit vide soit un sommet pour chacun des deux sous-domaines soit un coté commun aux deux sous-domaines. Un exemple d'une telle décomposition est donné par la figure V.1 ; à l'inverse, la figure V.2 représente une décomposition non-conforme, puisque Γ_{12} est une arète de Ω_1 mais seulement une partie d'une arête de Ω_2.

Nous introduisons ici les espaces fonctionnels nécessaires à la position du problème. Pour le problème initial, dans sa formulation "monodomaine", l'espace naturel est :

$$\mathrm{H}^1_{\partial\Omega_{\mathrm{D}}}(\Omega)=\{v\in\mathrm{H}^1(\Omega) : v=0 \text{ sur } \partial\Omega_{\mathrm{D}}\}$$

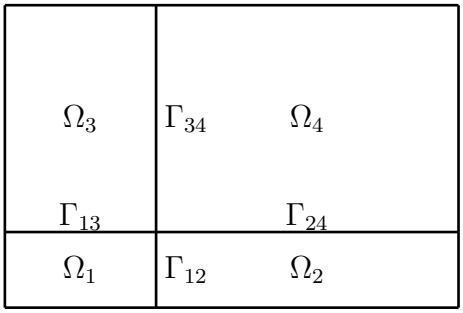


FIG. V.1 – Une décomposition de domaine conforme

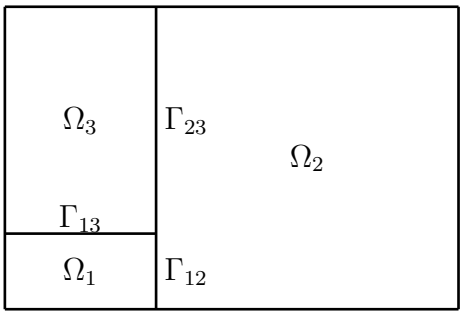


FIG. V.2 – Une décomposition de domaine non-conforme

Cet espace est inclus dans l'espace suivant, noté X_δ, qui est l'espace naturel pour la formulation multi-domaines :

$$\mathrm{X}_\delta = \{v \in \mathrm{L}^2(\Omega) : v_{|\Omega_k} \in \mathrm{H}^1(\Omega_k), k = 1, \ldots, K \text{ et } v = 0 \text{ sur } \partial\Omega_\mathrm{D}\}$$

où $v_{|\Omega_k}$ est la restriction à Ω_k de v. Cet espace est muni de la norme :

$$\|v\|_{1,\delta} = \left(\sum_{k=1}^{K} \|v\|^2_{1,\Omega_k} \right)^{\frac{1}{2}}$$

On admettra qu'il existe un espace de fonctions M portées par le squelette tel que l'on ait l'identité suivante :

$$\mathrm{H}^1_{\partial\Omega_\mathrm{D}}(\Omega) = \{v \in \mathrm{X}_\delta : (\mu, [v])_{0,\Gamma_{kl}} = 0 \ \forall \mu \in M, \ \Gamma_{kl} \subset S\}$$

où $[v]$ est une fonction définie sur chaque interface et qui désigne le saut de la fonction v à travers l'interface.

Soient les formes bilinéaires a et b :

$$a(\cdot,\cdot) \; : \left| \begin{array}{l} \mathrm{X}_\delta \times \mathrm{X}_\delta \longrightarrow \mathbb{R} \\ a(T,v) = \displaystyle\sum_k \int_{\Omega_k} \lambda \nabla T \cdot \nabla v + \vec{u}_{adv} \cdot \nabla T + \alpha T \cdot v \end{array} \right.$$

$$b(\cdot,\cdot) \; : \left| \begin{array}{l} \mathrm{X}_\delta \times \mathrm{M} \longrightarrow \mathbb{R} \\ b(v,\mu) = \displaystyle\sum_{\Gamma_{kl} \subset S} (\mu, [v])_{0,\Gamma_{kl}} = \sum_{\Gamma_{kl} \subset S} \int_{\Gamma_{kl}} \mu (v_{|\Omega_k} - v_{|\Omega_l}) \end{array} \right.$$

Le problème que nous considérons ici admet alors la formulation variationnelle suivante :

Trouver $(T, \lambda) \in \mathrm{X}_\delta \times \mathrm{M}$ tel que :

$$\left| \begin{array}{ll} a(T,v) + b(v,\lambda) = (f,v)_{0,\Omega} + (g,v)_{0,\partial\Omega_N} & \forall v \in \mathrm{X}_\delta \\ b(T,\mu) = 0 & \forall \mu \in \mathrm{M} \end{array} \right. \tag{V.2.1}$$

On démontre dans la littérature qu'il est bien posé [8, 5].

La méthode d'éléments finis joints peut maintenant être vue comme une approximation de Galerkin du problème de point selle (V.2.1). Soit donc une famille de triangulations $\mathcal{T}_{k,h},\ k = 1, \ldots, K$ de chaque sous-domaine $\Omega_k,\ k = 1, \ldots, K$, chacune étant choisie indépendemment des autres (en particulier, les noeuds appartenant à $\mathcal{T}_{k,h}$ situés sur l'interface Γ_{kl} ne sont pas forcément des noeuds de $\mathcal{T}_{l,h}$). L'espace correspondant aux éléments finis sur $\mathcal{T}_{k,h}$ est noté $S_h(\mathcal{T}_{k,h})$; l'approximation de l'espace X_δ s'écrit alors :

$$\mathrm{X_h} = \mathrm{X}_\delta \bigcap \prod_{k=1}^{K} S_h(\mathcal{T}_{k,h})$$

Il reste alors à choisir une discrétisation $\mathrm{M_h}$ de l'espace M ; pour ce faire, sur chaque interface, on choisit indifféremment l'espace des traces des fonctions de l'espace discret associé à l'un ou l'autre des domaines adjacents. Le problème discret s'écrit alors :

Trouver $(T_\delta, \lambda_\delta) \in \mathrm{X_h} \times \mathrm{M_h}$ tel que :

$$\left| \begin{array}{ll} a(T_\delta, v_\delta) + b(v_\delta, \lambda_\delta) = (f, v_\delta)_{0,\Omega} + (g, v_\delta)_{0,\partial\Omega_N} & \forall v_\delta \in \mathrm{X_h} \\ b(T_\delta, \psi_\delta) = 0 & \forall \psi_\delta \in \mathrm{M_h} \end{array} \right. \tag{V.2.2}$$

V.3 Un problème elliptique à donnée mesure

V.3.1 Position du problème

Le problème que nous traitons ici s'écrit :

$$\text{Trouver } T \text{ tel que} \qquad -\triangle T = \delta$$

où δ est la mesure de Dirac à l'origine. Ce problème posé dans $\mathbb{R}^d$ admet une solution analytique, que l'on peut trouver par exemple dans [1, théorèmes 7.55 et 7.56] :

$$\left| \begin{array}{ll} d=2 & T(x,y) = -\dfrac{1}{2\pi} \ln \sqrt{x^2+y^2} \\ \\ d=3 & T(x,y,z) = \dfrac{1}{4\pi} \dfrac{1}{\sqrt{x^2+y^2+z^2}} \end{array} \right.$$

Ce problème doit être posé ici sur un domaine fini, que nous choisissons comme $\Omega =]-1;1[^d$. Le problème avec conditions aux limites peut ainsi s'écrire en dimension deux :

$$\left\{ \begin{array}{ll} -\Delta T = \delta & \text{dans } \Omega =]-1;1[^2 \\ T(x,y) = -\dfrac{1}{2\pi} \ln \sqrt{x^2+y^2} & \text{sur } \partial\Omega \end{array} \right.$$

et en dimension trois :

$$\left| \begin{array}{ll} -\Delta T = \delta & \text{dans } \Omega =]-1;1[^3 \\ T(x,y,z) = \dfrac{1}{4\pi\sqrt{x^2+y^2+z^2}} & \text{sur } \partial\Omega \end{array} \right.$$

Dans la mesure où les espaces éléments finis utilisés ne contiendront que des fonctions continues à l'origine, la formulation variationnelle de ce problème posée avec les espaces discrets de manière usuelle a un sens et peut être conservée sans modification.

V.3.2 Expérimentations numériques avec une résolution monodomaine

Ces tests sont effectués en deux et trois dimensions d'espace. Le maillage utilisé est une grille régulière, avec une même subdivision dans chaque direction d'espace. Nous utilisons un espace d'éléments finis bilinéaires (éléments finis Q1).

En dimension deux, nous avons distingué deux cas dans nos résultats suivant que l'origine soit un noeud ou non du maillage ; l'erreur entre la solution exacte et la solution numérique est reportée dans les tableaux V.1 et V.2.

N. de	Norme $W^{1,1}$		Norme L^1		Norme L^2		Norme $L^{1.1}$		Norme $L^{1.2}$	
mailles	erreur	ordre	erreur	ordre	erreur	ordre	erreur	ordre	erreur	ordre
3^2	7.26e-1		4.76e-2		5.10e-2		4.66e-2		4.63e-2	
9^2	3.22e-1	0.74	7.62e-3	1.66	1.73e-2	0.98	8.11e-2	1.59	8.78e-2	1.51
27^2	1.35e-1	0.79	1.16e-3	1.71	5.78e-3	0.99	1.34e-3	1.63	1.58e-3	1.55
81^2	5.50e-2	0.81	1.65e-4	1.77	1.92e-3	1.00	2.07e-4	1.69	2.72e-4	1.60
243^2	2.16e-2	0.85	2.22e-5	1.82	6.43e-4	0.99	3.10e-5	1.73	4.54e-5	1.62

TAB. V.1 – Erreur et ordre de convergence lorsque l'origine est un noeud

Compte-tenu du peu d'influence que semble avoir la position de l'origine vis à vis de la position des noeuds sur l'ordre de convergence, nous n'avons pas continué à dupliquer les tests en trois dimensions : les maillages utilisés ont donc tous un noeud à l'origine. Les résultats obtenus sont reportés dans le tableau V.3.

N. de	Norme $W^{1,1}$		Norme L^1		Norme L^2		Norme $L^{1.1}$		Norme $L^{1.2}$	
mailles	erreur	ordre	erreur	ordre	erreur	ordre	erreur	ordre	erreur	ordre
6^2	3.21e-1		1.01e-2		8.48e-3		9.57e-3		9.17e-3	
18^2	1.50e-1	0.69	1.81e-3	1.56	2.87e-3	0.98	1.83e-3	1.50	1.89e-3	1.43
54^2	6.46e-2	0.76	2.81e-4	1.69	9.60e-4	0.99	3.10e-4	1.62	3.48e-4	1.54
162^2	2.64e-2	0.81	4.01e-5	1.77	3.2e-4	1	4.85e-5	1.62	6.04e-5	1.59
486^2	1.04e-2	0.84	5.40e-6	1.81	1.06e-4	1.00	7.29e-6	1.72	1.01e-5	1.62

TAB. V.2 – Erreur et ordre de convergence lorsque l'origine n'est pas un noeud

N. de	Norme $W^{1,1}$		Norme L^1		Norme L^2		Norme $L^{1.1}$		Norme $L^{1.2}$	
mailles	erreur	ordre	erreur	ordre	erreur	ordre	erreur	ordre	erreur	ordre
4^3	1.11		5.34e-2		4.10e-2		4.86e-2		4.55e-2	
8^3	6.21e-1	0.71	1.75e-2	1.60	2.88e-2	0.509	1.74e-2	1.47	1.79e-2	1.34
16^3	3.61e-1	0.78	5.44e-3	1.68	2.03e-2	0.504	6.00e-3	1.53	6.84e-3	1.39
32^3	2.06e-1	0.80	1.62e-3	1.74	1.44e-2	0.495	1.99e-3	1.59	2.53e-3	1.43
64^3	1.16e-1	0.82	4.72e-4	1.77	1.01e-2	0.511	6.44e-4	1.62	9.22e-4	1.45

TAB. V.3 – Erreur et ordre de convergence en dimension 3

Les résultats trouvés sont en accord avec l'estimation suivante, obtenue pour des conditions de régularité du problème que nous ne détaillerons pas ici (*c.f.* annexe consacrée à l'analyse de ce problème) :

$$||T - T_h||_{\mathrm{L}^p(\Omega)} \leq c\ h^{2-d+d/p} \qquad 1 < p < \frac{d}{d-2}$$

V.3.3 Expérimentations numériques en multi-domaines

Décompositions de domaine et discrétisation

Pour ce problème nous procéderons à deux décompositions de domaine :
- une décomposition en 5 sous-domaines telle que tracée sur la figure V.3,
- une décomposition en 9 sous-domaines, tracée sur la figure V.4.

Chaque sous-domaine est discrétisé par une méthode de Galerkin avec des éléments finis Q1 s'appuyant sur une grille uniforme. La stratégie de discrétisation spatiale adoptée est la suivante : plus le sous-domaine est proche de l'origine, plus son maillage est raffiné ; il existe un facteur 3 entre la taille des mailles des maillages de sous-domaine adjacents.

V.3.4 Résultats

La FIG.V.5 représente la norme L^2 de l'erreur *i.e.* la différence entre la solution exacte et la solution numérique obtenue pour des décompositions différentes. Le pas de maillage retenu est celui de la maille du domaine contenant l'origine.

Les résultats obtenus appellent les commentaires suivants :
- comme prévu par la théorie dans le cas d'une discrétisation conforme, le schéma est d'ordre un pour la norme L^2 en dimension deux ;

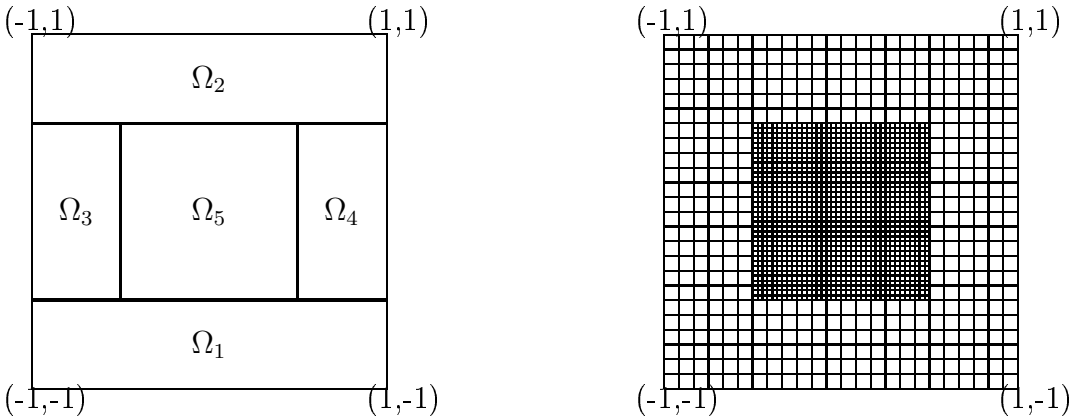


FIG. V.3 – Décomposition en 5 sous-domaines et maillage initial

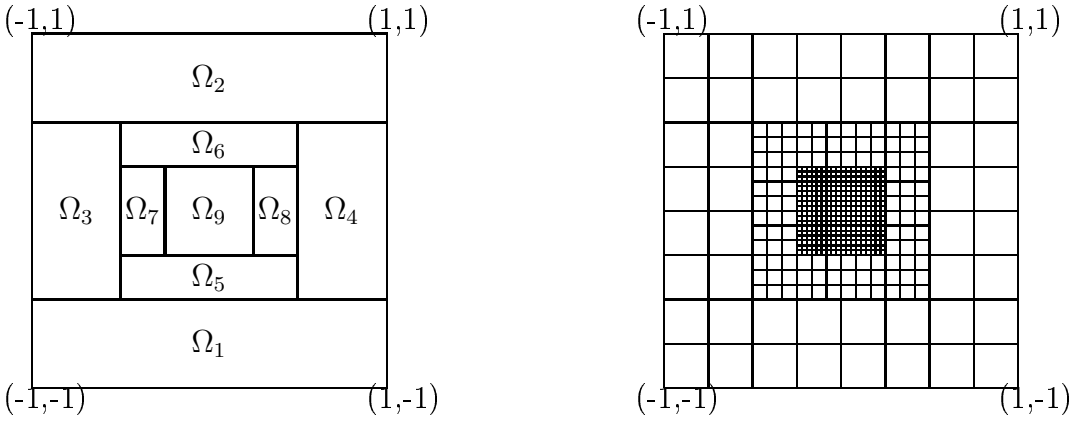


FIG. V.4 – Décomposition en 9 sous-domaines et maillage initial

– le schéma est d'ordre deux pour la norme L^2 sur chaque sous-domaine ne contenant pas l'origine *i.e.* ne contenant pas la discontinuité de la solution exacte ;
– quelle que soit la décomposition, l'erreur obtenue est pratiquement identique pour deux discrétisations dont le sous-domaine central possède le même pas de maillage,

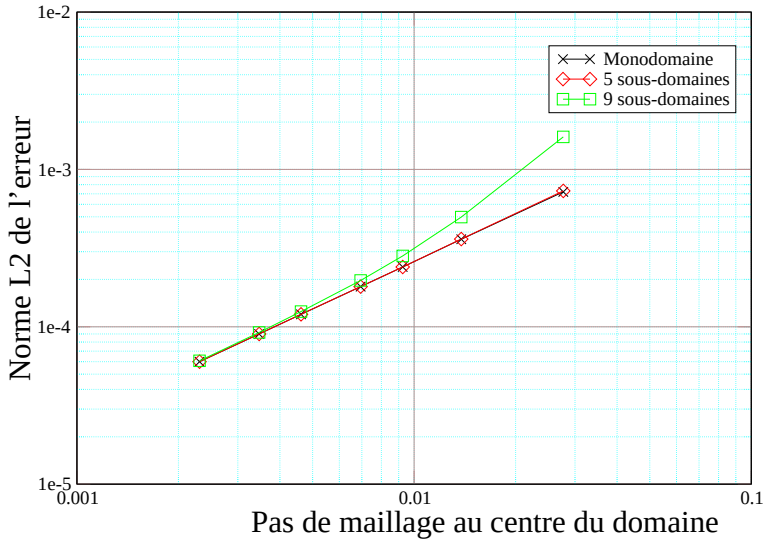


FIG. V.5 – Erreur en fonction du maillage pour différentes décompositions de domaine

ce qui montre la prépondérance du sous-domaine central sur l'erreur totale. En particulier, les résultats obtenus lorsque l'on ne raffine que la zone centrale sont très similaires à ceux obtenus lorsque l'on utilise le pas du domaine central pour l'ensemble du domaine. Un raffinement local du voisinage de l'origine est donc parfaitement approprié à ce cas.

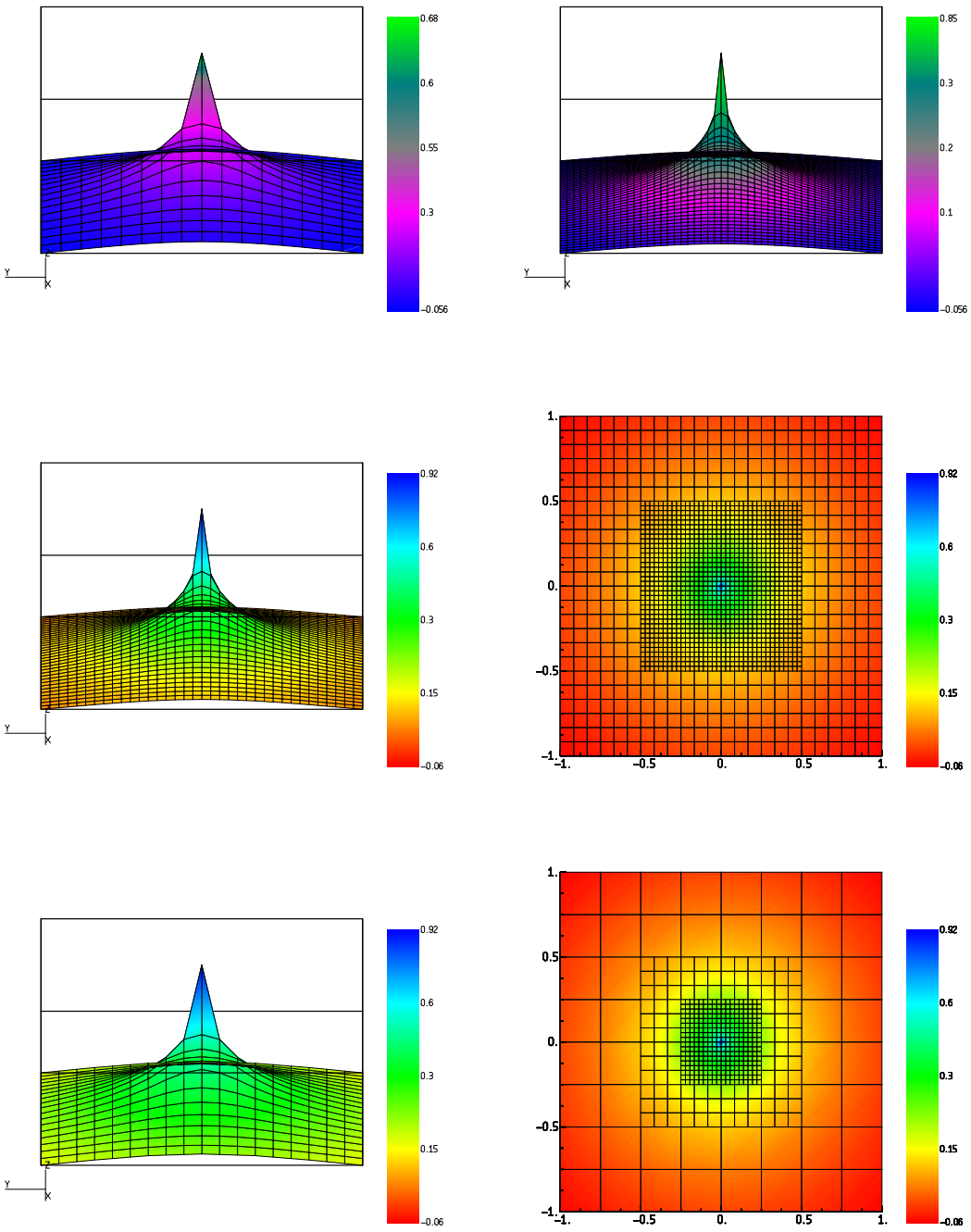


FIG. V.6 – Visualisation de la solution obtenue par une résolution monodomaine pour différents maillages (en haut), par une décomposition en 5 domaines (au milieu), par une décomposition en 9 sous-domaines (en bas)

V.4 Un problème multi-physiques

Dans cette section, nous appliquons la méthode des éléments finis joints à un problème couplant un problème de conduction thermique dans un solide à un problème de convection naturelle dans la cavité qu'il délimite. Nous souhaitons, par la mise en oeuvre de cette technique de décomposition de domaine, profiter de ce découpage naturel du domaine de calcul. Ce problème a déja fait l'objet d'une étude par Kim et Viskanta [7].

V.4.1 Position du problème

La géométrie du problème est représentée sur la figure V.7.

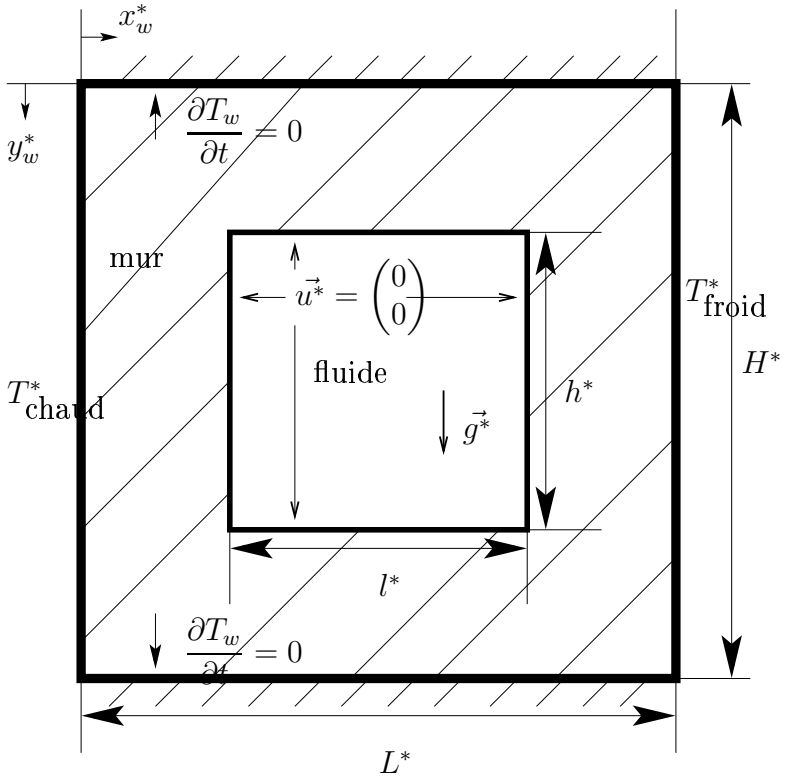


FIG. V.7 – Description physique du système

Les différentes notations sont regroupées dans le tableau V.4.

Les différences de température dans le système étant faibles, le problème physique étudié rentre dans le cadre de l'approximation de Boussinesq, où l'on pose $\varrho_c^* = \varrho_{c0}^*(1 - \beta(T_c^* - T_{c0}^*))$.

Les inconnues sont la température T_w^* dans le domaine solide Ω^*_{mur} et la température T_c^*, la vitesse $\mathbf{u}^*$ et la pression $p^* = p'^*$ dans le domaine fluide $\Omega^*_{\text{cavité}}$.

Données géométriques du problème	
Ω^*_{mur}	domaine du mur
$\Omega^*_{\text{cavité}}$	domaine de la cavité
Γ^*_{cw}	Interface mur cavité
$\mathbf{g}^* = \vec{g}$	gravité $(m^2.s^{-1})$
Quantités physiques dans le mur	
ϱ^*_w	masse volumique ($kg.m^{-3}$)
c^*_{pw}	chaleur spécifique(K^{-1})
λ^*_w	conductivité thermique ($J.s^{-1}.m^{-1}K^{-1}$)
a^*_w	diffusivité thermique ($m^2.s^{-1}$)
T^*_c	température (K)
Quantités physiques dans la cavité	
ϱ^*_c	masse volumique ($kg.m^{-3}$)
ϱ^*_{c0}	masse volumique moyenne ($kg.m^{-3}$)
μ^*_c	viscosité dynamique ($Pa.s$)
c^*_{pc}	chaleur spécifique(K^{-1})
λ^*_c	conductivité thermique ($J.s^{-1}.m^{-1}K^{-1}$)
a^*_c	diffusivité thermique ($m^2.s^{-1}$)
$\mathbf{u}^* = \vec{u^*}$	vecteur vitesse ($m.s^{-1}$)
p^*	pression totale du fluide (Pa)
p'^*	pression dynamique (Pa)
P^*	pression thermodynamique (Pa)
T^*_c	température (K)
T^*_{c0}	température moyenne (K)
β^*	coefficient d'expansion thermique (K^{-1})
Nombres adimensionnés	
$Ra = \dfrac{\varrho^*_c \Vert\mathbf{g}^*\Vert \beta^* L^{*3}(T^*_{\text{chaud}} - T^*_{\text{froid}})}{\mu^*_c a^*_c}$	nombre de Rayleigh
$Pr = \dfrac{\mu^*_c}{a^*_c \varrho^*_c}$	nombre de Prandtl

TAB. V.4 – Notations associées au problème d'interaction thermique fluide-structure

Elles vérifient le système d'équations suivant :

Equation bilan d'énergie dans le mur :

$$\varrho_w^* c_{pw}^* \frac{\partial T_w^*}{\partial t} - \nabla \cdot (\lambda_w^* \nabla T_w^*) = 0 \qquad \text{sur } [0;T^*] \times \Omega^*_{\text{mur}}$$

Equation bilan d'énergie dans la cavité :

$$\varrho_c^* c_{pc}^* (\frac{\partial T_c^*}{\partial t} + \mathbf{u}^* \cdot \nabla T_c^*) - \nabla \cdot (\lambda_c^* \nabla T_c^*) = 0 \qquad \text{sur } [0;T^*] \times \Omega^*_{\text{cavité}}$$

Continuité de la température à l'interface :

$$T_w^* = T_c^* \qquad \text{sur } [0;T^*] \times \Gamma^*_{cw}$$

Continuité du flux de chaleur à l'interface :

$$\lambda_w^* \frac{\partial T_w^*}{\partial n} = \lambda_c^* \frac{\partial T_c^*}{\partial n} \qquad \text{sur } [0;T^*] \times \Gamma^*_{cw}$$

Equation bilan de quantité de mouvement (cavité) :

$$\varrho_c^* (\frac{\partial \mathbf{u}^*}{\partial t} + \mathbf{u}^* \cdot \nabla \mathbf{u}^*) - \nabla \cdot (\tau^*(\mathbf{u}^*)) + \vec{\nabla} p^* = (\varrho_c^* - \varrho_{c0}^*)\mathbf{g}^* \qquad \text{sur } [0;T^*] \times \Omega^*_{\text{cavité}}$$

Equation bilan de masse (cavité) :

$$\nabla \cdot \mathbf{u}^* = 0 \qquad \text{sur } [0;T^*] \times \Omega^*_{\text{cavité}}$$

Le premier bloc d'équations correspond à l'équation de bilan d'énergie qui s'applique aux deux domaines, bien qu'avec des formulations différentes ; elle est posée ici sur chacun des domaines et complétée par des conditions de transmission à l'interface. A ce système vienne s'adjoindre les conditions aux limites spécifiées sur la figure V.7, à savoir $T_w^* = T^*_{\text{chaud}}$ et $T_w^* = T^*_{\text{froid}}$ sur les frontières verticales respectivement gauche et droite de la paroi solide, des conditions de Neumann homogène sur les frontières supérieure et inférieure, et des conditions d'adhérence sur l'ensemble des frontières de la cavité interne. L'état permanent est obtenu comme l'état stabilisé final après un transitoire, à partir de l'état initial défini par une température constante dans tout le système, prise à la moyenne de T^*_{froid} et T^*_{chaud}, et une vitesse nulle.

Le domaine complet ainsi que la cavité sont carrés, et le choix des longueurs de référence est donc sans équivoque. Les quantités adimensionnées sont ainsi définies par :

$$\varrho = \frac{\varrho_c^*}{\varrho_c^*} \ , \ c_p = \frac{c_p^*}{c_{pc}^*} \ , \ \lambda = \frac{\lambda^*}{\lambda_c^*} \ , \ a = \frac{a^*}{a_c^*} \ , \ \mathbf{u} = \frac{\mathbf{u}^*}{\frac{a_c^*}{L}} \ , \ p = \frac{p^*}{\varrho_c^* \left(\frac{a_c^*}{L}\right)^2}$$

$$x = \frac{x^*}{L^*} \ , \ y = \frac{y^*}{L^*} \ , \ l = \frac{l^*}{L^*} \ , \ t = \frac{a_c^* t^*}{L^{*2}}$$

$$T_c = \frac{T_c^* - T^*_{\text{froid}}}{T^*_{\text{chaud}} - T^*_{\text{froid}}} \ , \ T_w = \frac{T_w^* - T^*_{\text{froid}}}{T^*_{\text{chaud}} - T^*_{\text{froid}}}$$

$$\Omega_{\text{mur}} = [0,1] \times [0,1] \ , \ \Omega_{\text{cavité}} = [\frac{1-l}{2}, \frac{1+l}{2}] \times [\frac{1-l}{2}, \frac{1+l}{2}]$$

Nous introduisons également les deux rapports sans dimension suivants :

$$\lambda_r = \frac{\lambda_w^*}{\lambda_c^*} \text{ et } a_r = \frac{a_w^*}{a_c^*}$$

Les inconnues adimensionnées, soit la température T_w dans le domaine solide Ω_{mur} et la température T_c, la vitesse $\mathbf{u}$ et la pression $p = p'$ dans le domaine fluide $\Omega^*_{\text{cavité}}$, obéissent au système suivant :

Equation bilan d'énergie dans le mur :

$$\frac{\lambda_r}{a_r}\frac{\partial T_w}{\partial t} - \nabla \cdot (\lambda_r \nabla T_w) = 0 \qquad \text{sur } [0;T] \times \Omega_{\text{mur}}$$

Equation bilan d'énergie dans la cavité :

$$\frac{\partial T_c}{\partial t} + \mathbf{u} \cdot \nabla T_c - \nabla \cdot (\nabla T_c) = 0 \qquad \text{sur } [0;T] \times \Omega_{\text{cavité}}$$

Continuité de la température à l'interface :

$$T_w = T_c \qquad \text{sur } [0;T] \times \Gamma_{cw}$$

Continuité du flux de chaleur à l'interface :

$$\lambda_r \frac{\partial T_w}{\partial n} = \frac{\partial T_c}{\partial n} \qquad \text{sur } [0;T] \times \Gamma_{cw}$$

Equation bilan de quantité de mouvement (cavité) :

$$(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u}) - Pr \nabla \cdot (\nabla \mathbf{u} + \nabla^t \mathbf{u}) + \vec{\nabla} p = PrRa(T_c - 0.5) \qquad \text{sur } [0;T] \times \Omega_{\text{cavité}}$$

Equation bilan de masse (cavité) :

$$\nabla \cdot \mathbf{u} = 0 \qquad \text{sur } [0;T] \times \Omega_{\text{cavité}}$$

Les conditions aux limites de Dirichlet pour la température du solide deviennent $T_w = 1$ et $T_w = 0$ sur les frontières verticales respectivement gauche et droite de la paroi solide ; la condition initiale pour la température est $T_w = T_c = 0.5$.

V.4.2 Discrétisation temporelle

Pour l'application présentée ici, l'intervalle de temps est découpé de façon régulière, avec un pas de temps Δt fixé. La discrétisation temporelle de la dérivée en temps est effectuée par un schéma BDF2 (‘pour "Backward Differentiation Formula") *i.e.* une formule de différentiation rétrograde d'ordre 2. Le terme d'advection $\mathbf{u} \cdot \nabla \mathbf{u}$ est linéarisé en explicitant le champ advectif ; ce dernier est obtenu par une extrapolation de Richardson et a donc pour expression $2u^n - u^{n-1}$. Les équations de Navier-Stokes sont résolues par une méthode de projection. A chaque pas de temps, le problème consiste alors à calculer les quantités $T_w^{n+1} = T_w(t^{n+1})$, $T_c^{n+1} = T_c(t^{n+1})$, $\mathbf{u}^{n+1} = \mathbf{u}(t^{n+1})$ et $p^{n+1} = p(t^{n+1})$ solutions du système d'équations suivant :

$$
\left|
\begin{array}{ll}
\left[
\begin{array}{ll}
\dfrac{\lambda_r}{a_r}\,\dfrac{3T_w^{n+1}-4T_w^n+T_w^{n-1}}{2\Delta t}-\nabla\cdot(\lambda_r\nabla T_w^{n+1})=0 & \text{sur } [0;T]\times\Omega_{\text{mur}}\\
\dfrac{3T_c^{n+1}-4T_c^n+T_c^{n-1}}{2\Delta t}+(2\mathbf{u}^n-\mathbf{u}^{n-1})\cdot\nabla T_c^{n+1} & \\
\qquad -\nabla\cdot(\nabla T_c^{n+1})=0 & \text{sur } [0;T]\times\Omega_{\text{cavité}}
\end{array}
\right.\\
\left[
\begin{array}{ll}
T_w^{n+1}=T_c^{n+1} & \text{sur } [0;T]\times\Gamma_{cw}\\
\lambda_r\dfrac{\partial T_w^{n+1}}{\partial n}=\dfrac{\partial T_c^{n+1}}{\partial n} & \text{sur } [0;T]\times\Gamma_{cw}
\end{array}
\right.\\
\dfrac{3\tilde{\mathbf{u}}^{n+1}-4\mathbf{u}^n+\mathbf{u}^{n-1}}{2\Delta t}+(2\mathbf{u}^n-\mathbf{u}^{n-1})\cdot\nabla\tilde{\mathbf{u}}^{n+1}&\\
\quad -Pr\nabla\cdot(\nabla\tilde{\mathbf{u}}^{n+1}+\nabla^t\tilde{\mathbf{u}}^{n+1})+\nabla p^n=PrRa(T_c^{n+1}-0.5)\vec{y} & \text{sur } [0;T]\times\Omega_{\text{cavité}}\\
\left[
\begin{array}{ll}
\dfrac{3\mathbf{u}^{n+1}-3\tilde{\mathbf{u}}^{n+1}}{2\Delta t}+\nabla(p^{n+1}-p^n)=0 & \text{sur } [0;T]\times\Omega_{\text{cavité}}\\
\nabla\cdot\mathbf{u}^{n+1}=0 & \text{sur } [0;T]\times\Omega_{\text{cavité}}
\end{array}
\right.
\end{array}
\right.
$$

V.4.3 Discrétisation en espace

Nous nous attachons ici à préciser la discrétisation spatiale de l'équation de bilan d'énergie dans les domaines solide et fluide ; la discrétisation spatiale de l'étape de projection est effectuée de manière usuelle et n'est pas précisée ici.

Nous introduisons les formes bilinéaires suivantes :

$$
\begin{array}{lll}
m_{\text{w}}(\cdot,\cdot)\ : & (T,s)\mapsto\displaystyle\int_{\Omega_w}Ts\\
a_{\text{w}}(\cdot,\cdot)\ : & (T,s)\mapsto\displaystyle\int_{\Omega_w}\nabla T:\nabla v\\
m_{\text{c}}(\cdot,\cdot)\ : & (T,s)\mapsto\displaystyle\int_{\Omega_c}Ts\\
a_{\text{c}}(\cdot,\cdot)\ : & (T,s)\mapsto\displaystyle\int_{\Omega_c}(2\mathbf{u}^n-\mathbf{u}^{n-1})\cdot\nabla Ts+\nabla T:\nabla s\\
b_{\text{mortar}}(\cdot,\cdot)\ : & (T,\xi)\mapsto\displaystyle\int_{\Gamma_{cw}}[T]\ \xi
\end{array}
$$

où :

- la variable T appartient à un espace discret X_h de fonctions définies sur $\Omega_{\text{mur}}\cup\Omega_{\text{cavité}}$, régulières (*i.e.* appartenant à H^1) sur chacun des sous-domaines, vérifiant les conditions aux limites de Dirichlet du problème sur les frontières externes de Ω_{mur} et discontinues à travers l'interface,
- la notation $[T]$ désigne le saut de la variable T à travers l'interface Γ_{cw},
- et ξ appartient à une espace de fonctions M_h définies sur l'interface (les éléments joints), qui sera défini comme l'espace des traces des restrictions des fonctions de X_h à l'un ou l'autre, indifféremment, des sous-domaines.

Avec ces notations, le problème discret s'écrit :

Trouver T^{n+1} et $\chi^{n+1} \in \mathrm{X_h} \times \mathrm{M_h}$ tels que :

$$\left|\begin{array}{ll} \dfrac{\lambda_r}{a_r}\dfrac{1}{2\Delta t} m_{\mathrm{w}}(3T^{n+1} - 4T^n + T^{n-1}, s) + \dfrac{1}{2\Delta t} m_{\mathrm{c}}(3T^{n+1} - 4T^n + T^{n-1}, s) & \\ \qquad + \lambda_r a_{\mathrm{w}}(T^{n+1}, s) + a_{\mathrm{c}}(T^{n+1}, s) + b_{\mathrm{mortar}}(s, \chi^{n+1}) = 0 & \forall s \in \mathrm{X_h} \\ \\ b_{mortar}(T^{n+1}, \xi) = 0 & \forall \xi \in \mathrm{M_h} \end{array}\right.$$

V.4.4 Résultats

Le maillage utilisé pour effectuer ces calculs est représenté sur la FIG.V.8 ; les maillages de la cavité et de la paroi ne concordent pas à l'interface entre les deux domaines. Les températures et les vitesses sont discrétisées par des éléments finis de Lagrange simpliciaux quadratiques ; la pression est approximée par un élément linéaire. Le problème de point selle obtenu est résolu par une méthode de Lagrangien augmenté.

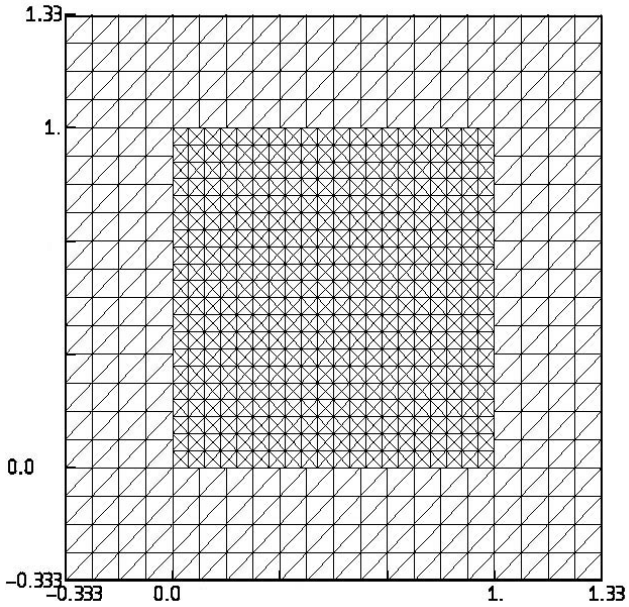


FIG. V.8 – Maillage utilisé pour l'expérimentation numérique

Rayleigh	Nusselt paroi chaude		Nusselt paroi chaude	
	KV	MEFJ	KV	MEFJ
10^4	1.310	1.264	1.311	1.264
10^5	2.195	2.380	2.184	2.381
10^6	3.801	4.645	3.781	4.647

TAB. V.5 – Comparaison du nombre de Nusselt moyen obtenu dans la présente étude et par Kim et Viskanta [7]

Nous définissons le nombre de Nusselt en un point de l'interface entre cavité et paroi par :

$$\mathrm{Nu} = -\lambda_r \frac{\partial T_c}{\partial n} / (T_c - T_{\mathrm{ref}})$$

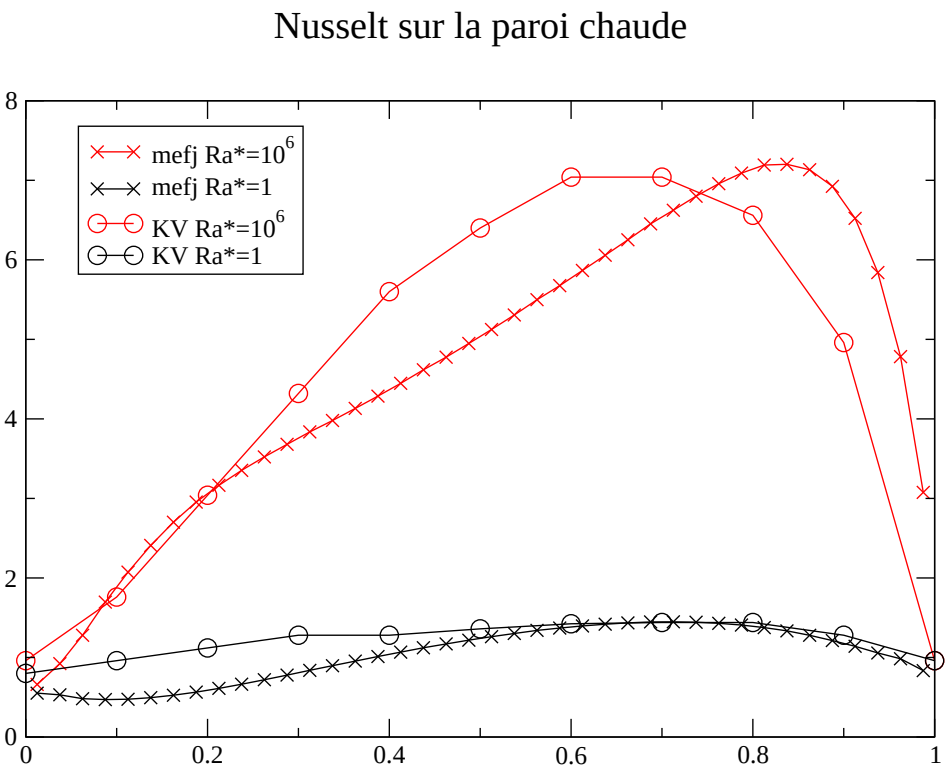


FIG. V.9 – Nusselt calculé sur la paroi chaude de la cavité une fois le régime permanent atteint, pour différents nombre de Rayleigh

Les nombres de Nusselt moyens à la paroi chaude et froide obtenus ici sont comparés avec ceux obtenus par Kim et Viskanta [7] dans le tableau V.5. Les résultats présentent une concordance acceptable. L'allure générale du champ de vitesse et celle du champ de température sont identiques comme le montrent les figures V.10 et V.11.

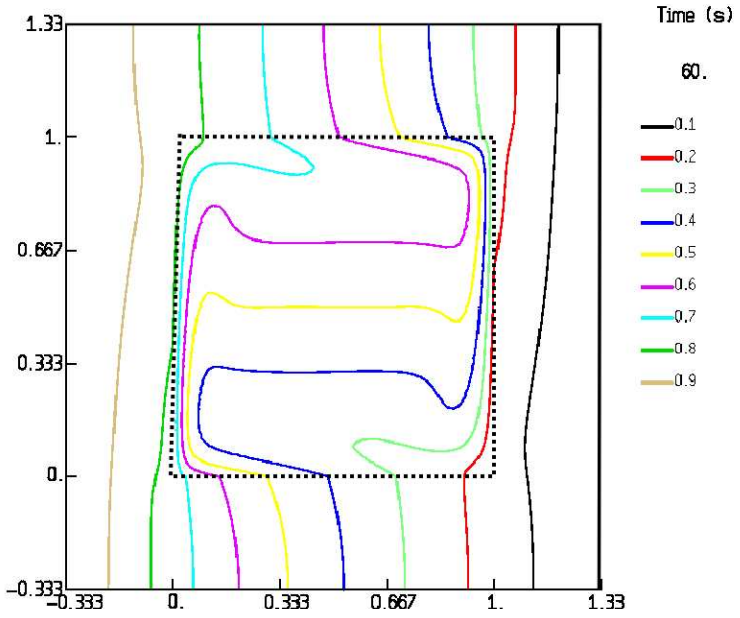


FIG. V.10 – Courbes isothermes en régime permanent

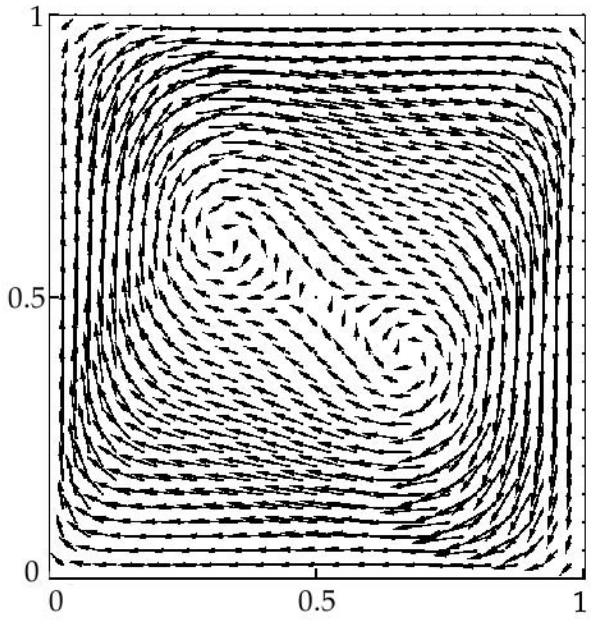


FIG. V.11 – Champ de vitesse dans la cavité en régime permanent

Bibliographie

[1] Walter Appel. *Mathématiques pour la physique et les physiciens*. H&K Editions, 2002.

[2] F. Ben Belgacem and Y. Maday. The mortar finite element method for three dimensional finite elements. *Mathematical Modelling and Numerical Analysis*, 31(2) :289–302, 1997.

[3] Faker Ben Belgacem. The mortar finite element method with Lagrange multipliers. *Numerische Mathematik*, 84 :173–197, 1999.

[4] Faker Ben Belgacem. The mixed mortar finite element method for the incompressible Stokes problem : Convergence analysis. *SIAM Journal on Numerical Analysis*, 37(4) :1085–1100, 2000.

[5] C. Bernardi, Y. Maday, and A.T. Patera. Domain decomposition by the mortar element method. In H. G. Kaper and M. Garbey, editors, *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters*, pages 269–286. Kluwer Academic Publishers, NATO ASI Series, 1993.

[6] David Braess, Wolfgang Dahmen, and Christian Wieners. A multigrid algorithm for the mortar finite element method. *SIAM Journal on Numerical Analysis*, 37(1) :48–69, 1999.

[7] D. M. Kim and R. Viskanta. Effect of wall heat conduction on natural convection heat transfer in a square enclosure. *Journal of Heat Transfer*, 107 :139–146, feb 1985.

[8] P.A. Raviart and J.M. Thomas. Primal hybrid finite element methods for 2nd order elliptic equations. *Mathematics of Computation*, 31(138) :391–413, 1977.

[9] Barbara I. Wohlmuth. Hierarchical a posteriori error estimators for mortar finite element methods with Lagrange multipliers. *SIAM Journal on Numerical Analysis*, 36(5) :1636–1658, 1999.

[10] Barbara I. Wohlmuth. A residual based error estimator for mortar finite element discretizations. *Numerische Mathematik*, 84 :143–171, 1999.

Annexes

Annexe A

Approximation par éléments finis P_1 de problèmes elliptiques du second ordre noncoercifs réguliers à donnée mesure

Nous traitons dans cette note les problèmes elliptiques du second ordre de la forme suivante :

$$\left| \begin{array}{ll} \mathbf{L}(u) \equiv -\nabla \cdot (A.\nabla u) + \vec{a}.\nabla u + du = f & \text{sur } \Omega \\ u = 0 \text{ sur } \partial\Omega & \end{array} \right.$$

où Ω est un domaine borné de $\mathbb{R}^n$, A est une matrice de $\mathbb{R}^{n\times n}$, $\vec{a}$ est un vecteur de $\mathbb{R}^n$, d et f sont des scalaires. On suppose vérifiées les conditions suivantes :

$$(H1) \quad \left| \begin{array}{ll} (A(x).\xi).\xi \geq \lambda\ |\xi|^2 & \forall \xi \in \mathbb{R}^n, \text{ presque partout dans } \Omega \\ |A(x).\xi| \leq \Lambda\ |\xi| & \forall \xi \in \mathbb{R}^n, \text{presque partout dans } \Omega \\ A_{i,j} \in \mathcal{C}^1(\bar{\Omega}) & \forall i,j \leq n \\ \vec{a}_i, d \in \mathrm{L}^\infty(\Omega) & \forall i \leq n \\ d \geq 0 & \text{presque partout dans } \Omega \end{array} \right.$$

Nous posons :

$$a(u,v) = \int_\Omega A.\nabla u.\nabla v + \vec{a}.\nabla u\ v + duv$$

$\mathrm{L}^p(\Omega)$, $\mathrm{H}^1_0(\Omega)$, $\mathrm{H}^{-1}(\Omega)$ et $\mathrm{W}^{1,p}_0(\Omega)$ désignent les espaces de Sobolev usuels, munis de leurs normes et semi-normes usuelles. On désignera par p' l'exposant conjugué de p, *i.e.* tel que $1/p + 1/p' = 1$.

Pour tout p vérifiant $1 < p < \infty$, la forme bilinéaire $a(.,.)$ est définie et continue sur $\mathrm{W}^{1,p}_0(\Omega) \times \mathrm{W}^{1,p'}_0(\Omega)$:

$$\forall (u,v) \in \mathrm{W}^{1,p}_0(\Omega) \times \mathrm{W}^{1,p'}_0(\Omega), \qquad a(u,v) \leq c\ ||u||_{1,p}\ ||v||_{1,p'} \tag{A.0.1}$$

où l'on choisit d'employer, ici et par la suite, la notation $\cdots \leq c\ \ldots$ en lieu et place de "il existe une constante c positive ne dépendant que de $\mathbf{L}(.)$ et de Ω telle que $\cdots \leq c\ \ldots$.

Le problème étudié s'écrit de manière différente selon la régularité du second membre f. Nous appelerons "problème à second membre régulier" le problème variationnel obtenu lorsque $f \in \mathrm{H}^{-1}(\Omega)$:

$$\text{Trouver } u \in \mathrm{H}_0^1(\Omega) \quad \text{tel que} \qquad a(u,v) = \int_\Omega f\ v \qquad \forall v \in \mathrm{H}_0^1(\Omega) \qquad \text{(A.0.2)}$$

Le "problème à données mesure" correspond au cas où f est une mesure de Radon ; il s'écrit sous forme variationnelle comme suit :

$$\text{Trouver } u \in \bigcap_{1\le p<n'} \mathrm{W}_0^{1,p}(\Omega) \quad \text{tel que} \qquad a(u,v) = \int_\Omega f\ v \qquad \forall v \in \bigcup_{n<p} \mathrm{W}_0^{1,p}(\Omega) \qquad \text{(A.0.3)}$$

où n' désigne l'exposant conjugué de n ($n' = n/(n-1)$).

Enfin, le problème adjoint au problème à second membre régulier s'écrit :

$$\text{Trouver } u \in \mathrm{H}_0^1(\Omega) \quad \text{tel que} \qquad a(v,u) = \int_\Omega f\ v \qquad \forall v \in \mathrm{H}_0^1(\Omega) \qquad \text{(A.0.4)}$$

Les développements effectués dans cette note reposent sur des techniques de dualité, dont l'utilisation nécessite l'hypothèse de régularité suivante :

(H2) $\forall p \le P$, si $f \in \mathrm{L}^p(\Omega)$, la solution de A.0.2 ou A.0.4, u, appartient à $\mathrm{W}^{2,p}$ et vérifie :

$$|u|_{2,p} \le c\ ||f||_{0,p}$$

où P est un réel strictement supérieur à n. Les implications de cette hypothèse de régularité sur le champ d'application des résultats présentés ici seront précisées en conclusion.

L'objectif de cette note est de présenter quelques résultats de convergence de la méthode de Galerkin standard avec des éléments finis de Lagrange linéaires pour les deux problèmes énoncés ci-dessus. Nous commençons par présenter le schéma numérique étudié puis nous démontrons les résultats d'existence et d'unicité de la solution du problème discret ainsi que de convergence pour le problème à second membre régulier ; enfin, pour le problème à second membre mesure, nous établissons la convergence du schéma en norme $\mathrm{W}^{1,p}(\Omega)$, $1 < p < n'$ et des bornes d'erreurs en norme $\mathrm{L}^p(\Omega)$, $1 < p < n/(n-2)$ qui, au vu des expérimentations numériques réalisées, semblent optimales.

Les résultats relatifs au problème régulier présentés ici sont connus ; on en trouvera par exemple une démonstration légèrement différente de celle effectuée ici mais équivalente dans [2, sections 5.6-5.8]. Les résultats de convergence pour le problème à donnée mesure généralisent, à notre connaissance, des résultats obtenus par Scott [8] (borne d'erreur en norme $\mathrm{L}^2(\Omega)$ pour le problème $-\Delta u = \delta$ où δ est la mesure de Dirac) puis Casas [3] (borne d'erreur en norme $\mathrm{L}^2(\Omega)$ pour un problème elliptique du second ordre coercif et une mesure générique). Pour ce faire, nous profitons d'un théorème de convergence en norme $\mathrm{W}^{1,p}(\Omega)$ de la projection elliptique sur un espace d'éléments finis obtenus par Rannacher et Scott [7] et généralisé dans [2].

A.1 Schéma numérique

Le schéma que nous étudions ici s'écrit :

$$\text{Trouver } u_{\rm h} \in V_h \quad \text{t.q.} \qquad a(u_{\rm h}, v_{\rm h}) = \int_\Omega f\ v_{\rm h} \qquad \forall v_{\rm h} \in V_h \tag{A.1.5}$$

où V_h est l'espace d'éléments finis P_1 inclus dans $\mathrm{H}^1_0(\Omega)$ usuel et la notation $\int_\Omega f\ v_{\rm h}$ est employée pour désigner un produit de dualité, dans l'espace correspondant à la régularité de f. Il est à noter à ce propos que V_h est inclus dans $\mathrm{W}^{1,p}(\Omega)$ quelque soit $p \in [1, \infty]$ ce qui donne un sens à l'expression $\int_\Omega f\ v_{\rm h}$ y compris pour f mesure.

On suppose que, pour la famille de triangulations du domaine utilisée, il existe un opérateur d'interpolation $r_{\rm h}$ tel que les inégalités suivantes soient vérifiées :

– les inégalités d'interpolation usuelles, pour $1 \leq p \leq \infty$:

$$\begin{aligned} \forall u \in \mathrm{W}^{2,p} \qquad & ||u - r_{\rm h}u||_{1,p} \leq c\ h\ |u|_{2,p} \\ & ||u - r_{\rm h}u||_{0,p} \leq c\ h^2\ |u|_{2,p} \end{aligned} \tag{A.1.6}$$

– l'inégalité d'interpolation suivante, pour $1 \leq q \leq p \leq \infty$:

$$\forall u \in \mathrm{W}^{2,q} \qquad ||u - r_{\rm h}u||_{1,p} \leq c\ h^{1-n(\frac{1}{q}-\frac{1}{p})}\ |u|_{2,q} \tag{A.1.7}$$

Cette relation est aisément déduite de [4], théorème 16.2.

– l'inégalité inverse :

$$\forall v_{\rm h} \in V_h, \quad \forall p, q \in [1, \infty] \qquad ||v_{\rm h}||_{0,p} \leq c\ h^{-n\max(0,\frac{1}{q}-\frac{1}{p})}\ ||v_{\rm h}||_{0,q} \tag{A.1.8}$$

Dans ces deux dernières relations, il convient de remplacer $1/p$ (respectivement $1/q$) par 0 lorsque p (respectivement q) est infini.

Pour que les inégalités A.1.7 et A.1.8 soient vérifiées, il faut que la famille de triangulations utilisée soit quasi-uniforme (voir [4]).

A.2 Une inégalité de stabilité

Soit f et g deux fonctions de, respectivement, $\mathrm{L}^2(\Omega)$ et $\mathrm{L}^2(\Omega)^n$ et (Pb) et $(Pb)^\star$ les problèmes suivants :

$$(Pb) \qquad \text{Trouver } u \in \mathrm{H}^1_0(\Omega) \text{ vérifiant :} \qquad a(u, v) = \int_\Omega fv + g.\nabla v \qquad \forall v \in \mathrm{H}^1_0(\Omega)$$

$$(Pb^\star) \qquad \text{Trouver } u \in \mathrm{H}^1_0(\Omega) \text{ vérifiant :} \qquad a(v, u) = \int_\Omega fv + g.\nabla v \qquad \forall v \in \mathrm{H}^1_0(\Omega)$$

On rappelle (voir [5], par exemple) que les problèmes (Pb) et $(Pb^\star)$ ont une solution et une seule, respectivement u et $u^\star$, vérifiant :

$$\exists c > 0 \text{ t.q.} \quad \left| \begin{aligned} \|u\|_1 &\leq c\ (\|f\|_{\rm o} + \|g\|_{\rm o}) \\ \|u^\star\|_1 &\leq c\ (\|f\|_{\rm o} + \|g\|_{\rm o}) \end{aligned} \right. \tag{A.2.9}$$

Le résultat principal de la présente section est le suivant :

Proposition A.2.1. Si le pas de discrétisation h est inférieur à une valeur h_0 ne dépendant que de $\mathbf{L}(.)$ et Ω, il existe β indépendant de h tel que :

$$\forall v_{\mathrm{h}} \in V_h, \qquad \sup_{w_{\mathrm{h}} \in V_h} \frac{a(v_{\mathrm{h}}, w_{\mathrm{h}})}{\|w_{\mathrm{h}}\|_1} \geq \beta \ \|v_{\mathrm{h}}\|_1 \tag{A.2.10}$$

Preuve

Soit v_{h} un élément de V_h. Nous allons construire explicitement un élément de V_h tel que l'inégalité A.2.10 soit vérifiée.

On définit $v^\star$ par :

$$v^\star \in \mathrm{H}_0^1(\Omega), \qquad a(w, v^\star) = \int_\Omega v_{\mathrm{h}} \ w + \int_\Omega \nabla v_{\mathrm{h}}.\nabla w \qquad \forall w \in \mathrm{H}_0^1(\Omega) \tag{A.2.11}$$

Par l'inégalité A.2.9, $v^\star \in \mathrm{H}_0^1(\Omega)$ et vérifie :

$$\|v^\star\|_1 \leq c \ \|v_{\mathrm{h}}\|_1$$

$v^\star$ n'est pas plus régulier : pour que $v^\star$ appartienne à $\mathrm{H}^2(\Omega)$, il faudrait que v_{h} soit également élément de $\mathrm{H}^2(\Omega)$, c'est à dire $V_h \subset \mathrm{H}^2(\Omega)$, ce qui n'est pas le cas.

Soit maintenant ${v_{\mathrm{h}}}^\star$ défini par :

$$\int_\Omega A.(\nabla {v_{\mathrm{h}}}^\star - \nabla v^\star).\nabla w_{\mathrm{h}} + d({v_{\mathrm{h}}}^\star - v^\star) \ w_{\mathrm{h}} = 0 \qquad \forall w_{\mathrm{h}} \in V_h \tag{A.2.12}$$

Cette relation définit ${v_{\mathrm{h}}}^\star$ par projection elliptique sur V_h, en utilisant un opérateur coercif. On a donc :

$$\|{v_{\mathrm{h}}}^\star\|_1 \leq c \ \|v^\star\|_1 \leq c \ \|v_{\mathrm{h}}\|_1 \tag{A.2.13}$$

Par ailleurs, du fait des propriétés de régularité du problème, on a par le lemme d'Aubin-Nitsche le résultat d'approximation suivant ([4, 2]) :

$$\|v^\star - {v_{\mathrm{h}}}^\star\|_{\mathrm{o}} \leq c \ h \ \|v^\star - {v_{\mathrm{h}}}^\star\|_1 \leq c \ h \ \|v_{\mathrm{h}}\|_1 \tag{A.2.14}$$

En combinant A.2.11 et A.2.12, il vient alors :

$$\begin{aligned} a(v_{\mathrm{h}}, {v_{\mathrm{h}}}^\star) &= a(v_{\mathrm{h}}, v^\star) + \int_\Omega \vec{a}.\nabla v_{\mathrm{h}} \ ({v_{\mathrm{h}}}^\star - v^\star) \\ &= \int_\Omega {v_{\mathrm{h}}}^2 + \int_\Omega \nabla v_{\mathrm{h}}.\nabla v_{\mathrm{h}} + \int_\Omega \vec{a}.\nabla v_{\mathrm{h}} \ ({v_{\mathrm{h}}}^\star - v^\star) \end{aligned}$$

D'où :

$$a(v_{\mathrm{h}}, {v_{\mathrm{h}}}^\star) \geq \int_\Omega {v_{\mathrm{h}}}^2 + \int_\Omega \nabla v_{\mathrm{h}}.\nabla v_{\mathrm{h}} - \ \|\vec{a}\|_\infty \ |v_{\mathrm{h}}|_1 \ \|{v_{\mathrm{h}}}^\star - v^\star\|_{\mathrm{o}}$$

Soit, en utilisant l'erreur d'interpolation A.2.14 :

$$a(v_{\rm h}, v_{\rm h}^{\star}) \geq \int_{\Omega} {v_{\rm h}}^2 + \int_{\Omega} \nabla v_{\rm h}.\nabla v_{\rm h} - c\ h\ \ \|\vec{a}\|_{\infty}\ \ |v_{\rm h}|_1^2 \geq (1 - c\ h\ \ \|\vec{a}\|_{\infty}\ )\ \|v_{\rm h}\|_1^2$$

ce qui, joint à l'inégalité A.2.13, conclut le démonstration.

□

En corollaire, on obtient :

Corollaire A.2.1. Si le pas de discrétisation h est inférieur à une valeur h_0 ne dépendant que de $\mathbf{L}(.)$ et Ω, il existe une solution $u_{\rm h}$ et une seule au problème discret A.1.5.

Remarque : Il est possible de construire des contre-exemples très simples - en dimension 1 - qui montrent qu'il n'est pas possible d'obtenir un résultat de stabilité valable quel que soit le pas de discrétisation ; sans autre hypothèse sur le champ d'advection, le schéma de Galerkin peut conduire pour un pas de discrétisation trop grossier à un opérateur discret singulier.

Par ailleurs, ce résultat de stabilité permet de démontrer la convergence optimale du schéma pour le problème régulier :

Corollaire A.2.2. Soit u la solution de A.0.2 et $u_{\rm h}$ celle de A.1.5. Pour tout pas de discrétisation h plus petit qu'une valeur h_0 ne dépendant que de Ω et $\mathbf{L}(.)$, la méthode de Galerkin vérifie la borne d'erreur optimale suivante :

$$\|u - u_{\rm h}\|_1 \leq c \inf_{v_{\rm h} \in V_h} \|u - v_{\rm h}\|_1 \tag{A.2.15}$$

Preuve

La démonstration de ce résultat est parfaitement classique ; nous ne la rappelons ici que par souci d'être complets.

Par l'inégalité triangulaire :

$$\|u - u_{\rm h}\|_1 \leq \|u - v_{\rm h}\|_1 + \|v_{\rm h} - u_{\rm h}\|_1$$

Supposons que h soit suffisamment petit pour que le résultat de stabilité précédent s'applique. On a alors :

$$\|v_{\rm h} - u_{\rm h}\|_1 \leq \frac{1}{\beta} \sup_{w_{\rm h} \in V_h} \frac{a(v_{\rm h} - u_{\rm h}, w_{\rm h})}{\|w_{\rm h}\|_1} \tag{A.2.16}$$

Du fait de la consistance de la méthode :

$$\forall w_{\rm h} \in V_h, \qquad a(u_{\rm h}, w_{\rm h}) = a(u, w_{\rm h})$$

et donc :

$$\forall w_{\rm h} \in V_h, \qquad a(v_{\rm h} - u_{\rm h}, w_{\rm h}) = a(v_{\rm h} - u, w_{\rm h})$$

Par continuité de $a(.,.)$, on a :

$$\forall w_{\rm h} \in V_h, \qquad a(v_{\rm h}-u, w_{\rm h}) \leq c \, \|v_{\rm h}-u\|_1 \, \|w_{\rm h}\|_1$$

L'inégalité A.2.16 donne donc :

$$\|v_{\rm h}-u_{\rm h}\|_1 \leq \frac{1}{\beta} \sup_{w_{\rm h}\in V_h} \frac{a(v_{\rm h}-u, w_{\rm h})}{\|w_{\rm h}\|_1} \leq c \, \|v_{\rm h}-u\|_1$$

ce qui, en reportant dans l'inégalité triangulaire initiale, fournit le résultat attendu.

□

A.3 Analyse d'erreur pour le problème à donnée mesure

Soit p compris strictement entre 1 et ∞, $u \in \mathrm{W}_0^{1,p}(\Omega)$ et $u_{\rm h}$ sa projection sur V_h donnée par :

$$a(u_{\rm h}, v_{\rm h}) = a(u, v_{\rm h}) \qquad \forall v_{\rm h} \in V_h$$

Nous commençons cette section par un résultat d'approximation en norme $\mathrm{L}^\infty(\Omega)$ relatif à cette projection.

On rappelle tout d'abord le résultat suivant, valable pour des éléments finis de Lagrange généraux :

Théorème A.3.1. Sous les hypothèses de régularité de l'opérateur $(H1)$ et du problème $(H2)$, pour un pas de discrétisation h inférieur à une valeur h_0 ne dépendant que de $\mathbf{L}(.)$ et Ω, la majoration suivante est vérifiée :

$$||u-u_{\rm h}||_{1,p} \leq c \inf_{v_{\rm h}\in V_h} ||u-v_{\rm h}||_{1,p} \tag{A.3.17}$$

On trouvera une preuve de cette majoration dans [7] pour $2 \leq p \leq \infty$ et des éléments finis de degré 1 ; la démonstration générale figure dans [2] (chapitre 7). Ce résultat difficile repose sur une majoration ponctuelle de $\nabla u_{\rm h}$ obtenue par une technique de norme discrète pondérée. Il sous-tend toute la présente analyse.

En corollaire, nous avons la majoration suivante :

Corollaire A.3.1. On suppose que $u \in \mathrm{W}_0^{1,p}(\Omega) \cap \mathrm{W}^{2,p}(\Omega)$ que p est tel que $\mathrm{W}^{2,p}(\Omega) \subset \mathrm{L}^\infty(\Omega)$ (soit $p > n/2$) et que $p' \leq P$ soit $p \geq P/(P-1)$. On a alors, sous les mêmes hypothèses de régularité que le théorème précédent :

$$\|u-u_{\rm h}\|_\infty \leq c \, h^{2-n/p} \, |u|_{2,p} \tag{A.3.18}$$

Preuve

Dans un premier temps, on utilise une technique de dualité analogue à celle employée pour la preuve du lemme d'Aubin-Nitsche pour majorer $||u-u_{\rm h}||_{0,p}$ puis, dans un second temps, une inégalité inverse permet de passer à la norme infinie. La

démonstration analogue pour $p = 2$ est classique ; on la trouvera par exemple dans [4].

On a :

$$||u - u_{\rm h}||_{0,p} = \sup_{g \in {\rm L}^{p'}(\Omega)} \frac{\displaystyle\int_\Omega g\ (u - u_{\rm h})}{||g||_{0,p'}}$$

Soit ϕ_g donnée par :

$$\phi_g \in {\rm H}^1_0(\Omega), \qquad a(v, \phi_g) = \int_\Omega g\ v \qquad \forall v \in {\rm H}^1_0(\Omega)$$

Du fait de la régularité du problème (hypothèse $(H2)$), ϕ_g vérifie :

$$\phi_g \in {\rm W}^{1,p'}_0(\Omega) \cap {\rm W}^{2,p'}(\Omega), \qquad |\phi_g|_{2,p'} \leq c\ ||g||_{0,p'} \tag{A.3.19}$$

L'égalité :

$$\int_\Omega g\ (u - u_{\rm h}) = a(u - u_{\rm h}, \phi_g)$$

découle de la définition de ϕ_g lorsque $p \leq 2$. Elle est vraie également pour $1 < p < 2$ pour les raisons suivantes : le membre de droite a un sens du fait que ϕ_g, par le résultat de régularité, est dans ${\rm W}^{1,p'}_0(\Omega)$; l'égalité s'en déduit par la continuité de $a(.,.)$ et la densité de fonctions suffisamment régulières dans ${\rm W}^{1,p}_0(\Omega)$ (par exemple, les fonctions continuement dérivables à support compact de Ω dans $\mathbb{R}$).

En conséquence :

$$\int_\Omega g\ (u - u_{\rm h}) = a(u - u_{\rm h}, \phi_g - v_{\rm h}) \qquad \forall v_{\rm h} \in V_h$$

et, par l'inégalité de Hölder (A.0.1) :

$$\int_\Omega g\ (u - u_{\rm h}) \leq c\ ||u - u_{\rm h}||_{1,p} \inf_{v_{\rm h} \in V_h} ||\phi_g - v_{\rm h}||_{1,p'}$$

Le résultat de convergence A.3.17, le résultat d'approximation A.1.6 et l'inégalité de régularité A.3.19 donnent :

$$||u - u_{\rm h}||_{0,p} \leq c\ h^2\ |u|_{2,p} \tag{A.3.20}$$

On a alors pour tout élément $v_{\rm h}$ de V_h, par l'inégalité triangulaire :

$$\|u - u_{\rm h}\|_\infty\ \leq\ \|u - v_{\rm h}\|_\infty\ +\ \|u_{\rm h} - v_{\rm h}\|_\infty$$

L'inégalité inverse A.1.8 permet d'écrire :

$$\|u_{\rm h} - v_{\rm h}\|_\infty\ \leq c\ h^{-n/p}\ ||u_{\rm h} - v_{\rm h}||_{0,p}$$

Et, à nouveau par l'inégalité triangulaire :

$$\|u_{\rm h} - v_{\rm h}\|_\infty\ \leq c\ h^{-n/p}\ (||u - u_{\rm h}||_{0,p} + ||u - v_{\rm h}||_{0,p})$$

En utilisant ces trois dernières inégalités, on obtient :

$$\|u - u_{\rm h}\|_\infty \leq c\ h^{-n/p}\ ||u - u_{\rm h}||_{0,p} + \inf_{v_{\rm h} \in V_h} \left(\|u - v_{\rm h}\|_\infty + c\ h^{-n/p}\ ||u - v_{\rm h}||_{0,p} \right)$$

Le résultat recherché en découle par l'estimation A.3.20 et les inégalités d'approximation A.1.6 et A.1.7.

□

Nous sommes maintenant en position de démontrer le résultat de convergence suivant :

Théorème A.3.2. Soit u solutions du problème à données mesure et $u_{\rm h}$ solution du problème discret A.1.5. Les estimation d'erreur suivantes sont vérifiées :

$$||u - u_{\rm h}||_{1,p} \leq c \inf_{v_{\rm h} \in V_h} ||u - v_{\rm h}||_{1,p} \qquad \text{pour } 1 < p < n/(n-1)$$

$$||u - u_{\rm h}||_{0,p} \leq c\ h^{2-n/p'}\ ||f||_{\rm m} \qquad \text{pour } P/(P-1) \leq p < n/(n-2)$$

En absence de résultat de régularité, la première de ces relations garantit uniquement la convergence dans $\mathrm{W}^{1,p}(\Omega)$ de la méthode.

Preuve

La première relation découle directement du résultat de convergence A.3.17, en remarquant que :

$$a(u_{\rm h}, v_{\rm h}) = a(u, v_{\rm h}) \qquad \forall v_{\rm h} \in V_h$$

Pour la seconde de ces relations, une fois encore, nous utilisons une technique de dualité, avec une variante due à Casas [3].

Soit $p < n/(n-2)$. Il existe alors $q < n/(n-1)$ tel que l'espace $\mathrm{W}^{1,q}(\Omega)$ s'injecte continuement dans $\mathrm{L}^p(\Omega)$ et l'on peut écrire :

$$||u - u_{\rm h}||_{0,p} = \sup_{g \in \mathrm{L}^{p'}(\Omega)} \frac{\displaystyle\int_\Omega g\ (u - u_{\rm h})}{||g||_{0,p'}}$$

Soit ϕ_g donnée par :

$$\phi_g \in \mathrm{H}^1_0(\Omega), \qquad a(v, \phi_g) = \int_\Omega g\ v \qquad \forall v \in \mathrm{H}^1_0(\Omega)$$

Du fait de la régularité du problème (hypothèse $(H2)$), ϕ_g vérifie :

$$\phi_g \in \mathrm{W}^{1,p'}_0(\Omega) \cap \mathrm{W}^{2,p'}(\Omega), \qquad |\phi_g|_{2,p'} \leq c\ ||g||_{0,p'}$$

Comme dans la démonstration précédente, l'égalité :

$$\int_\Omega g\ (u - u_{\rm h}) = a(u - u_{\rm h}, \phi_g)$$

est vraie dès que le second membre a un sens, c'est à dire lorsque ϕ_g est élément de $\mathrm{W}^{1,q}_0(\Omega)$ avec $q > n$, ce qui est vérifié, par la régularité de ϕ_g, dès que $p < n/(n-2)$.

On alors :

$$\int_\Omega g\ (u - u_\mathrm{h}) = a(u, \phi_g) - a(u_\mathrm{h}, \phi_g) = \int_\Omega f\ \phi_g - a(u_\mathrm{h}, \phi_g)$$

Soit $\phi_{g\mathrm{h}}$ défini par :

$$\phi_{g\mathrm{h}} \in V_h, \qquad a(v_\mathrm{h}, \phi_{g\mathrm{h}}) = a(v_\mathrm{h}, \phi_g) \qquad \forall v_\mathrm{h} \in V_h$$

On a alors du fait de la définition de $\phi_{g\mathrm{h}}$ puis de u_h :

$$a(u_\mathrm{h}, \phi_g) = a(u_\mathrm{h}, \phi_{g\mathrm{h}}) = \int_\Omega f\ \phi_{g\mathrm{h}}$$

Il vient donc :

$$\int_\Omega g\ (u_\mathrm{h} - u) = \int_\Omega f\ (\phi_{g\mathrm{h}} - \phi_g) \leq ||f||_\mathrm{m}\ \ \|\phi_{g\mathrm{h}} - \phi_g\|_\infty$$

ce qui, au vu du résultat de convergence A.3.18, fournit l'estimation recherchée.

□

Remarque : hypothèse de régularité du problème $(H2)$

Cette hypothèse revient, avec des conditions de Dirichlet homogènes et une fois supposée la régularité de l'opérateur $\mathbf{L}(.)$ $(H1)$ à une condition de régularité sur la frontière du domaine Ω. Nous pouvons distinguer deux cas :

- domaine polygonal (en 2D) ou union finie de polyhèdres (en 3D).
  Dans ce cas, la triangulation du domaine est, à moins d'y mettre de la mauvaise volonté, exacte. Par contre, P ne peut que rarement être choisi arbitrairement grand.
  Plus précisément, en 2D, P vérifie ([6], corollaire 5.3) :

  $$2 - \frac{2}{P} < \frac{\pi}{\omega}$$

  où ω désigne le plus grand angle intérieur entre deux segments consécutifs du polygone. Cette condition permet toutefois, en particulier, d'appliquer les résultats obtenus ici aux expérimentations numériques réalisées, pour lesquelles Ω est rectangulaire.
  En 3D, le résultat est plus difficile à énoncer, du fait que la borne sur P dépend des valeurs propres de certains problèmes aux limites associés à l'opérateur $\mathbf{L}(.)$, dont on n'a que rarement une estimation explicite. On sait toutefois ([6], corollaire 5.9) que P peut être pris égal à 2 si le domaine est convexe.
- domaine régulier.
  L'exposant P peut être pris arbitrairement grand si le domaine est de classe $\mathcal{C}^2$ ([1]). Dans ce cas, toutefois, la triangulation du domaine ne peut être exacte et ce qui induit une erreur supplémentaire.

Bibliographie

[1] S. Agmon, A. Douglis, and L. Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I. *Communications on Pure and Applied Mathematics*, XII :623–727, 1959.

[2] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Analysis*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, 1991.

[3] Eduardo Casas. L^2 estimates for the finite element method for the Dirichlet problem with singular data. *Numerische Mathematik*, 47 :627–632, 1985.

[4] P.G. Ciarlet. *Handbook of Numerical Analysis Volume II : Finite Elements Methods – Basic Error Estimates for Elliptic Problems*. North-Holland, 1991.

[5] David Gilbarg and Neil S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer, third edition, 2001.

[6] Pierre Grisvard. Behaviour of the solutions of an elliptic boundary value problem in a polygonal or polyhedral domain. In *Numerical Solution of Partial Differential Equations – III – Synspade 1975*, New York, San Francisco, London, 1976. Academic Press, INC.

[7] Rolf Rannacher and Ridgway Scott. Some optimal error estimates for piecewise linear finite element approximations. *Mathematics of Computation*, 38(158) :437–445, April 1982.

[8] Ridgway Scott. Finite element convergence for singular data. *Numerische Mathematik*, 21 :317–327, 1973.