Numéro d'ordre : 18/16

Université de Valenciennes et du Hainaut-Cambrésis

Thèse de doctorat

Pour obtenir le grade de Docteur de l'Université de

VALENCIENNES ET DU HAINAUT-CAMBRESIS

Mathématiques appliquées

présentée et soutenue par

BEKKOUCHE FATIHA

le 22/06/2018, à Valenciennes

École doctorale : Science Pour l'Ingénieur (SPI)

Équipe de recherche, Laboratoire : Laboratoire de Mathématiques et ses Applications de Valenciennes(LAMAV)

Étude théorique et numérique des équations non-linéaires de Sobolev

Membres du jury :

Directeurs :

Serge Nicaise. Université de Valenciennes et du Hainaut-Cambrésis

Wided Chikouche. Université Mohamed Sedik Ben Yahia de Jijel

Rapporteurs :

Chérif Amrouche. Université de Pau et des Pays de l'Adour

Stéphane Maingot. Université du Havre

Examinateur (Président du jury) :

Colette, De Coster. Université de Valenciennes et du Hainaut Cambrésis.

RÉSUMÉ

L'objectif de la thèse est l'étude mathématique et l'analyse numérique du problème non linéaire de Sobolev. Un premier chapitre est consacré à l'analyse a priori pour le problème de Sobolev où on utilise des méthodes de semi-discrétisation explicite en temps. Des estimations d'erreurs ont été obtenues assurant que les schémas numériques utilisés convergent lorsque le pas de discrétisation en temps et le pas de discrétisation en espace tendent vers zéro. Dans le second chapitre, on s'intéresse au problème de Sobolev singulièrement perturbé. En vue de la stabilité des schémas numériques, on utilise dans cette partie des méthodes numériques implicites (la méthode d'Euler et la méthode de Crank-Nicolson) pour discréti-ser le problème par rapport au temps. Dans le troisième chapitre, on présente des applications et des illustrations où on utilise le logiciel "FreeFem++". Dans le dernier chapitre, on considère une équation de type Sobolev et on s'intéresse à la dérivation d'estimations d'erreur a posteriori pour la discrétisation de cette équation par la méthode des éléments finis conforme en espace et un schéma d'Euler implicite en temps. La borne supérieure est globale en espace et en temps et permet le contrôle effectif de l'erreur globale. À la fin du chapitre, on propose un algorithme adaptatif qui permet d'atteindre une précision relative fixée par l'utilisateur en raffinant les maillages adaptativement et en équilibrant les contributions en espace et en temps de l'erreur. On présente également des essais numériques.

Mots-Clés

Équation de Sobolev, estimation d'erreur, estimation d'erreur a posteriori, algorithme adaptatif.

Abstract

The purpose of this work is the mathematical study and the numerical analysis of the nonlinear Sobolev problem. A first chapter is devoted to the a priori analysis for the Sobolev problem, where we use an explicit semidiscretization in time. A priori error estimates were obtained ensuring that the used numerical schemes converge when the time step discretization and the spatial step discretization tend to zero. In a second chapter, we are interested in the singularly perturbed Sobolev problem. For the stability of numerical schemes, we used in this part implicit semidiscretizations in time (the Euler method and the Crank-Nicolson method). Our estimates of Chapters 1 and 2 are confirmed in the third chapter by some numerical experiments. In the last chapter, we consider a Sobolev equation and we derive a posteriori error estimates for the discretization of this equation by a conforming finite element method in space and an implicit Euler scheme in time. The upper bound is global in space and time and allows effective control of the global error. At the end of the chapter, we propose an adaptive algorithm which ensures the control of the total error with respect to a user-defined relative precision by refining the meshes adaptively, equilibrating the time and space contributions of the error. We also present numerical experiments.

Keywords

Sobolev equation, error estimate, a posteriori error estimate, adaptive algorithm.

Remerciements

Ces remerciements clôturent mes travaux de thèse, financés par le Ministère de l'Enseignement Supérieur et de la Recherche Scientifique d'Algérie et réalisés au Laboratoire de Mathématiques et leurs Applications de Valenciennes (LMAV) dont je tiens à remercier chaleureusement tous les membres. À ce titre je tiens à exprimer ma gratitude à Felix Memmati, directeur de LAMAV, d'avoir accepté de m'accueillir au sein de son laboratoire. Je souhaite le remercier particulièrement pour sa gentillesse, sa disponibilité et de son soutien pour mes différentes démarches administratives. J'ai eu, au cours de cette thèse, le plaisir de rencontrer de nombreuses personnes, qui ont influencé d'une façon ou d'une autre mes activités de recherche et d'enseignement à l'Institut des Sciences et Techniques de Valenciennes (ISTV) dont je remercie également tous les collègues enseignants.

J'adresse mes sincères remerciements aux membres du jury qui ont accordé un temps précieux à l'étude de ce travail, particulièrement messieurs Chérif Amrouche et Stéphane Maingot pour l'honneur qu'ils m'ont fait en acceptant de relire cette thèse et d'en être rapporteurs malgré leurs emploi du temps plus que chargé. La version finale de ce mémoire a bénéficié de leur lecture très attentive et de leurs remarques précieuses. Je tiens à remercier Mme Colette De Coster, Professeur à l'Université de Valencienne, d'avoir accepté d'être présidente du jury et Mme Chikouche qui s'est déplacé depuis Jijel (ma ville natale) en Algérie.

Si cette thèse s'est accomplie dans de si bonnes conditions, c'est avant tout grâce à mon directeur de thèse à Valenciennes Mr Serge Nicaise. Je ne saurai trouver les mots pour le remercier de m'avoir dirigé tout au long de ces 42 mois de thèse. Je lui suis redevable non seulement pour sa disponibilité et à l'écoute de mes nombreuses questions, mais aussi il a indéniablement contribué à l'avancée de mes travaux de thèse. Les nombreuses discussions que nous avons eues ainsi que ses conseils, en particulier, en simulations numériques sont pour beaucoup dans le fruit de ce travail. Son énergie, sa façon de poser des hypothèses adéquates à des problèmes mathématiques particuliers, sa capacité d'analyse et son enthousiasme m'ont montré l'autre facette du monde de la recherche qui est un univers (non borné !) passionnant. Enfin, ses nombreuses relectures et corrections de cette thèse et des publications ont été très appréciables. Cette thèse lui doit beaucoup, merci encore une fois Serge !

Je remercie Mme Chikouche d'avoir co-encadré ce travail de thèse. Elle m'a tout d'abord permis d'intégrer l'équipe Analyse mathématiques à Jijel (en Algérie) en me proposant un sujet très intéressant avec Serge Nicaise et m'a laissé la liberté de le réorienter au cours du déroulement de ma thèse. C'est également grâce à sa collaboration avec le laboratoire LAMAV que j'ai eu la chance d'intégré le laboratoire LAMAV, ce qui c'est avéré une expérience très enrichissante. Au cours de ce travail, elle m'a consacré beaucoup de temps, notamment pour la relecture du manuscrit. Je salue également sa personnalité à la fois sympathique, sérieuse et modeste.

Je remercie Monsieur Emmanuel Creusé, Professeur à l'Université de Lille 1. En effet, il m'a consacré une partie de son précieux temps pour de nombreuses discussions sur les tests numériques.

Je remercie tous les thésards et docteurs du laboratoire pour la bonne ambiance de travail, je pense en particulier à Jérôme, Sadjya avec qui j'ai partagé mon bureau, merci pour ces bons moments passés ensembles. Entre autres Hawraa, Florent, Elsa, Tonia, Zainab, Thibault, Marilena, Alaa, etc. Je remercie tous les autres membres du laboratoire IEMN et LAMIH, notamment ceux avec qui j'ai eu l'occasion de discuter "recherche" et les autres simplement pour les bons moments partagés. Je remercie toutes les personnes qui m'ont permis d'oublier momentanément le travail dans des sorties, repas, ou autres, notamment Sadjya Arriche (à qui je souhaite un heureux mariage !), Maroua, Lynda et Kahina. Elles ont toujours répondu présentes, même dans les cas les plus désespérés. Merci.

Je voudrais remercier particulièrement la secrétaire Nabila pour sa sympathie et son efficacité dans l'organisation et la résolution des problèmes administratifs, mais aussi pour les bons moments lors des pauses café.

Je remercie Lynda Chehami et Abdelouahab Chikh Salah , qui ont pris le temps de relire une partie de ce manuscrit et revoir mes transparents, et qui se sont arrachés les cheveux pour essayer de comprendre le sens souvent tortueux de mes phrases en "math".

Enfin, mes plus grands remerciements s'adressent à ma chère famille, mes parents, mes sœurs et mes frères dont la patience et le soutiens sans faille m'ont permis d'aller jusqu'au bout de cette expérience. Je n'oublie pas ma petite princesse d'amour Serine qui a été une vraie bouffée d'oxygène durant les derniers mois de rédaction qui n'ont pas toujours été des plus agréables.

Table des matières

Introduction générale **1**

1 Analyse a priori par un schéma explicite en temps d'un problème de Sobolev **7**

1.1 Les équations semi-linéaires 9

1.1.1 Semi-discrétisation explicite en temps 10

1.1.2 Schéma totalement discrétisé 15

1.2 Les équations quasi-linéaires 22

1.2.1 Existence et unicité de la solution exacte 23

1.2.2 Semi-discrétisation explicite en temps 25

1.2.3 Schéma totalement discrétisé 29

2 Analyse a priori pour un schéma implicite en temps d'un problème singulièrement perturbé **35**

2.1 Le problème homogène 38

2.1.1 Etude du problème semi-discrétisé implicite 38

2.1.2 Etude du problème semi-discrétisé en espace 41

2.2 Le problème non homogène 56

2.2.1 Le problème semi-discrétisé en temps 56

2.2.2 Le problème totalement discrétisé 59

3 Applications et résultats numériques **69**

3.1 Applications pour des équations semi-linéaires particulières 69

3.1.1 Opérateurs elliptiques d'ordre deux : le cas régulier 69

3.1.2 Opérateurs elliptiques d'ordre deux : le cas singulier 74

3.1.3 Opérateurs elliptiques d'ordre deux à coefficients discontinus : le cas singulier 77

3.1.4 Résultats numériques 82

3.2 Applications pour des équations quasi-linéaires de Sobolev 92

3.2.1 Le cas d'une équation non autonome 92

3.2.2 Le cas d'une équation quasi-linéaire 95

3.3 Applications pour une équation de Sobolev singulièrement perturbée 105

4 Analyse a posteriori pour un problème de Sobolev 109

4.1 Le problème semi-discret et totalement discret 111
4.1.1 La discrétisation en temps par le schéma d'Euler implicite 111
4.1.2 La discrétisation totale 111
4.2 Quelques outils d'analyse 115
4.3 Analyse d'erreur a posteriori pour le problème semi-discrétisé 116
4.4 Analyse a posteriori de la discrétisation spatiale 119
4.4.1 Borne supérieure de l'erreur 120
4.4.2 Borne inférieure de l'erreur 123
4.5 Analyse a posteriori pour le problème totalement discrétisé 125
4.6 Résultats numériques 127
4.6.1 Dépendance de l'erreur 130
4.6.2 Algorithme adaptatif 132

Annexe A Les codes en FreeFem++ 141

A.1 Les codes pour une équation linéaire 141
A.2 Les codes pour une équation semi-linéaire 145
A.3 Les codes pour une équation quasi-linéaire 147

Bibliographie 148

Table des figures

3.1 Triangle de référence-Transformation affine. 73
3.2 Fonction de troncature η. 76
3.3 Triangulation régulière du carré $\Omega = (0,1)^2$ pour $h = 0.2$. 83
3.4 La solution approchée par le schéma d'Euler (3.8) du problème (3.2) au temps final $T = 1$ avec $\Delta t = h = 0.1$. 83
3.5 La solution exacte du problème (3.2) au temps final $T = 1$. 83
3.6 Erreur en norme $H_0^1(\Omega)$ par le schéma d'Euler explicite (3.8) en espace et en temps au temps final T. 84
3.7 Erreur en norme $H_0^1(\Omega)$ par le schéma de Runge-Kutta (3.9) en espace et en temps au temps final T. 85
3.8 La solution approchée par le schéma d'Euler (3.8) du problème (3.2) au temps final $T = 1$ avec $\Delta t = h = 0.1$. 87
3.9 La solution exacte du problème (3.2) au temps final $T = 1$. 87
3.10 Erreur en norme $H_0^1(\Omega)$ par le schéma d'Euler (3.8) en espace et en temps au temps final T. 88
3.11 A gauche : maillage initial uniforme avec $h = 0.2$, à droite : maillage initial raffiné près de 0. 89
3.12 Erreur en norme $H_0^1(\Omega)$ en espace au temps final $T = 0.1$ avec un maillage uniforme et un maillage raffiné pour $\Delta t = 0.0001$. 90
3.13 La solution numérique $U_{n,h}$ obtenue par le schéma d'Euler (3.8) avec $\Delta t = 0.1$ et $h = 0.1$. 91
3.14 La solution approchée par le schéma d'Euler (3.8) du problème (3.21) au temps final $T = 1$ avec $\Delta t = h = 0.1$. 94
3.15 La solution exacte du problème (3.21) au temps final $T = 1$. 94
3.16 Erreur en norme $H_0^1(\Omega)$ en espace au temps final $T = 0.1$ et en temps au temps final $T = 1$. 96
3.17 La solution numérique $U_{n,h}$ obtenue par le schéma d'Euler (3.8) avec $\Delta t = h = 0.2$. 97
3.18 L'élément HCT. 103

3.19 La solution approchée $U_{n,h}$ obtenue par le schéma d'Euler (3.8) avec $\Delta t = h = 0.1$ à l'instant final $T = 1$. 104
3.20 Le gradient de la solution approchée, avec $\Delta t = h = 0.1$ à l'instant final $T = 1$. 105
3.21 L'évolution de l'erreur en norme $L^2(\Omega)$ pour différentes valeurs de ϵ avec $\Delta t = 0.0001$ et $h = 0.00625$. 107
3.22 Erreur en norme $L^2(\Omega)$ en espace au temps final $T = 0.1$ avec un maillage uniforme et un maillage raffiné pour $\Delta t = 0.0001$. 108

4.1 à gauche : maillage non conforme en 2d, à droite : maillage conforme en 2d. 112
4.2 Exemple de deux éléments partageant le bord E et de normale n_E. 113
4.3 Patch d'un élément K. 114
4.4 L'erreur $\|u(t_N) - u_h^N\|_{1,\Omega}^2$ par rapport à Dof pour un maillage uniforme. . . 128
4.5 q_{up}^N et q_{low}^N par rapport à Dof pour un maillage uniforme. 129
4.6 Maillage non structuré avec $h = 0.2$. 130
4.7 q_{up}^N et q_{low}^N par rapport à Dof pour un maillage non structuré. 131
4.8 Schéma d'une boucle d'adaptation de maillage. 134
4.9 Raffinement R1 à gauche et R2 à droite. 135
4.10 La solution approchée après un raffinement adaptatif aux temps $t = 0.1$, $t = 0.5$ et $t = 1$ respectivement avec $\alpha = 0.5$ et $\delta = 0.25$. 137
4.11 Maillage après le raffinement adaptatif à $t = 0,1$, $t = 0.5$ et $t = 1$ avec $\alpha = 0.5$ et $\delta = 0.25$. 138
4.12 Maillage après le raffinement adaptatif à $t = 0,1$ avec $\alpha = 0.5$ et $\delta = 0.25$. . 139

Liste des tableaux

3.1 L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$. 84
3.2 L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 1$ pour différents Δt avec $h = 0.00625$. 84
3.3 L'évolution de l'erreur par le schéma de Runge-Kutta (3.9) et MEF$\mathbb{P}_2$ au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$. 85
3.4 L'évolution de l'erreur par le schéma de Runge-Kutta (3.9) et et MEF$\mathbb{P}_2$ au temps final $T = 20$ pour différents Δt avec $h = 0.00625$. 85
3.5 L'évolution de l'erreur par le schéma d'Euler (3.8) et MEF$\mathbb{P}_1$ au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$. 86
3.6 L'évolution de l'erreur par le schéma d'Euler (3.8) et MEF$\mathbb{P}_1$ au temps final $T = 1$ pour différents Δt avec $h = 0.00625$. 86
3.7 L'évolution de l'erreur par le schéma d'Euler (3.8) et et MEF$\mathbb{P}_1$ au temps final $T = 0.1$ pour différents h en utilisant un maillage initial uniforme ou raffiné près de 0 pour $\Delta t = 0.0001$ et $\epsilon = 3$. 89
3.8 L'ordre de convergence pour différents Δt avec $h = \frac{1}{160}$. 91
3.9 L'ordre de convergence pour différents h avec $\Delta t = \frac{1}{4000}$. 91
3.10 L'ordre de convergence pour différents h avec $\Delta t = \frac{1}{4000}$ pour le schéma d'Euler et EF$\mathbb{P}$2. 92
3.11 L'ordre de convergence pour différents Δt avec $h = \frac{1}{320}$ pour la méthode d'Euler et EF$\mathbb{P}$2. 92
3.12 L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$. 95
3.13 L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 1$ pour différents Δt avec $h = 0.00625$. 95
3.14 L'ordre de convergence pour différents Δt avec $h = \frac{1}{160}$. 95
3.15 L'ordre de convergence pour différents h avec $\Delta t = \frac{1}{4000}$. 96
3.16 L'ordre de convergence pour différents Δt avec $h = 0.003125$. 98
3.17 L'ordre de convergence pour différents h avec $\Delta t = 0.001$. 99
3.18 L'ordre de convergence pour différents Δt avec $h = 0.00625$. 104

3.19 L'ordre de convergence pour différents h avec $\Delta t = 0.001$. 104
3.20 L'ordre de convergence pour différents Δt avec $h = 0.00625$. 107
3.21 L'évolution de l'erreur en norme $L^2(\Omega)$ par le schéma d'Euler implicite (3.43) au temps final $T = 1$ pour différents Δt avec $h = 0.00625$ et $\epsilon = 10^{-3}$.107
3.22 L'évolution de l'erreur en norme $L^2(\Omega)$ par le schéma d'Euler (3.43) au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$ et $\epsilon = 10^{-3}$. 107
3.23 L'évolution de l'erreur en norme $L^2(\Omega)$ par le schéma de Crank-Nicolson (3.44) au temps final $T = 10$ pour différents Δt avec $h = 0.00625$ et $\epsilon = 10^{-3}$.108

4.1 q_{up}^N et q_{low}^N par rapport à Dof pour un maillage uniforme. 129
4.2 q_{up}^N et q_{low}^N par rapport à DoF pour un maillage non structuré. 130
4.3 Résultats de convergence où on a utilisé une triangulation uniforme et un pas de temps constant pour le premier exemple. 132
4.4 Résultats de convergence où on a utilisé une triangulation uniforme et un pas de temps constant pour le deuxième exemple. 133

Introduction générale

Notre compréhension des phénomènes du monde réel et les technologies d'aujourd'hui sont en grande partie basées sur les équations aux dérivées partielles. C'est en effet grâce à la modélisation de ces phénomènes au travers d'EDP que l'on a pu comprendre le rôle de tel ou tel paramètre, et surtout obtenir des prévisions parfois extrêmement précises.

Certaines de ces équations peuvent être résolues analytiquement et leurs solutions exactes sont connues. Toutefois, pour un nombre important d'équations, la solution exacte est inconnue. C'est dans cette optique que les recherches se sont penchées sur les méthodes numériques pour pouvoir approcher les solutions exactes de ces équations. Ces méthodes numériques nécessitent de transformer le problème continu en un problème discret en utilisant des méthodes de discrétisation spatiale et temporelle. Ainsi ces méthodes fournissent des solutions approchées qui sont le plus souvent dans un espace de dimension finie. La discrétisation de ces équations par ces méthodes numériques donne lieu en général à des systèmes linéaires. Une question importante se pose : est ce que la résolution de ce système fournit une solution approchée de bonne qualité ? Cette question est liée au problème d'estimer et de contrôler l'erreur entre la solution exacte et la solution approchée. En d'autres termes, il s'agit de déterminer la précision de la méthode numérique.

Dans cette thèse, nous nous intéressons à la discrétisation et à l'analyse d'erreur des équations non linéaires du type Sobolev suivantes

$$A_1 u_t + A_2 u = f, \tag{1}$$

où A_1 et A_2 sont des opérateurs différentiels pouvant dépendre de u et de t. Ce type d'équation n'est pas récent, il a été proposé pour la première fois en 1954 par Sobolev [68] d'où le nom. Il sont l'une des nombreuses équations aux dérivées partielles importantes qui modélisent, l'écoulement du fluide dans un milieu poreux fissuré [10], écoulement diphasique en milieu poreux avec une pression capillaire dynamique [23, 37], conduction thermique dans les systèmes à deux températures [18, 73] et le cisaillement à l'ordre 2 dans les fluides [22, 72].

Les résultats d'existence de tels problèmes sont prouvés pour des équations semi-

linéaires ou non-autonomes (i.e., le cas où A_1 et A_2 ne dépendent que de t) dans [53, 65, 66, 31, 16, 67] par exemple, où les auteurs exploitent le fait que A_1 est inversible. Ce qui leur a permis de réduire le problème à une équation d'évolution du premier ordre avec un opérateur borné (non autonome) et le résultat d'existence suit facilement. La même idée a été utilisée dans [43] pour montrer un résultat d'existence pour le cas des équations quasi-linéaires.

Concernant l'analyse de l'erreur pour diverses méthodes, de nombreux travaux sont publiés dans la littérature mathématique et qui traitent divers problèmes de Sobolev de type (1). Pour la méthode de Crank-Nicolson en temps combinée avec la méthode des éléments finis, Arnold, Douglas et Thomée [4] et Wahlbin [79] ont considéré une approximation totalement discrète d'un problème linéaire et semi-linéaire de Sobolev respectivement en dimension 1 avec une condition aux limites et ils ont établi l'ordre de convergence optimal dans diverses normes et la superconvergence nodale. Citons aussi l'approche proposée par Ewing et Lin [30, 52] des équations non-linéaires de Sobolev où ils ont obtenu un ordre de convergence optimal en norme H^1 et L^2 respectivement. Dans [5, 25], on trouve une approximation par la méthode de Petrov-Galerkin et dans [32, 58, 57, 70], les auteurs proposent une approximation totalement discrète par la méthode de Crank-Nicolson en temps combinée avec la méthode discontinue de Galerkin.

Les estimations d'erreur a priori étudiées dans les papiers précédents fournissent des bornes sur la différence entre la solution exacte (qu'on appelle u) et la solution approchée (qu'on appelle $u_{n,h}$) dans une certaine norme $\|\cdot\|$, sous la forme

$$\|u - u_h\| \leq C(\Delta t, h)\|u\|_*, \tag{2}$$

où $\|\cdot\|_*$ diffère en général de $\|\cdot\|$ et $C(\Delta t, h)$ est une fonction positive de Δt et h qui tend vers zéro lorsque Δt et h tendent vers zéro. Ces estimations sont utilisées afin de justifier théoriquement la convergence de la méthode numérique. Elles sont établies en supposant certaines régularités sur la solution exacte (ces régularités peuvent ne pas être vérifiées en pratique). De plus, la constante générique C qui apparaît dans l'estimation (2) est, soit inconnue, soit difficile à estimer. Cependant, la difficulté la plus importante est que la norme de la solution exacte $\|u\|_*$ n'est pas calculable, simplement parce que la solution exacte est inconnue. Donc, en pratique, la borne supérieure (2) ne permet pas le contrôle effectif de l'erreur et par suite ne peut pas servir comme critère d'arrêt pour un algorithme de résolution.

Pour pouvoir développer une méthode adaptative, on va avoir recours à des estimations d'erreur a posteriori. L'intérêt de ces estimations est de fournir des bornes sur l'erreur qui peuvent être évaluées dès que la solution approchée a été calculée. En effet, ces bornes sont évaluées uniquement en fonction de la solution approchée, des données du problème

et des données du maillage. Ainsi, $\eta(T_h, D, u_{n,h})$ est un estimateur d'erreur a posteriori si

$$\|u - u_{n,h}\| \leq \eta(T_h, D, u_{n,h}), \tag{3}$$

où T_h est le maillage et D est l'ensemble des données du problème.

Pour les problèmes de Sobolev, plusieurs travaux ont été effectués dans le cadre de la méthode des éléments finis conformes. On cite entre autres, [60], [78] et Bergam, Bernardi et Mghazli [11] pour l'équation de la chaleur et Verfürth [77] pour l'équation de convection-diffusion-réaction. Cascón, Ferragut, et Asensio [17] obtiennent une estimation d'erreur a posteriori pour l'équation de la chaleur discrétisée par des éléments finis mixtes et Nicaise et Soualem [56] pour la méthode des éléments finis non conformes. Dans le cadre des méthodes de Galerkin discontinues, on cite Sun et Wheeler [69], Ern et Proft [28], Yang et Chen [80], et Chen et Yang [19], qui dérivent des estimations d'erreur a posteriori pour le problème de convection-diffusion instationnaire, et Georgoulis, Lakkis, et Virtanen [33] qui présentent une estimation d'erreur a posteriori pour un problème de diffusion instationnaire. Houston et Süli [39] dérivent des estimations d'erreur a posteriori pour un problème de convection-diffusion instationnaire par la méthode de Lagrange-Galerkin.

Amara, Nadau et Trujillo [3] présentent une estimation d'erreur a posteriori pour l'équation de la chaleur discrétisée par la méthode des volumes finis centrés par maille sur un maillage rectangulaire en utilisant les schémas à 5 ou 9 points. Ces estimations sont établies en utilisant une interprétation du schéma volumes finis en termes d'une méthode d'éléments finis mixtes. L'étude est basée sur les travaux de Bergam, Bernardi et Mghazli [11] où l'erreur est décomposée en erreur temporelle et erreur spatiale et ensuite chacune de ces composantes est bornée par un indicateur. Cette technique de séparation des erreurs est également utilisée par Aboulaich, Achchab et Darouichi [2] qui considèrent un problème de transport unidimensionnel discrétisé par les volumes finis centrés par sommets. Ces estimations d'erreur a posteriori sont dérivées pour la norme d'énergie de l'erreur.

Ce travail comporte principalement quatre parties

- une première est consacrée à l'analyse a priori par un schéma explicite d'un problème quasi-linéaire de Sobolev,
- s'ensuit une seconde consacrée à l'analyse a priori par un schéma implicite d'un problème singulièrement perturbé de Sobolev,
- la troisième partie est orientée vers des applications et des illustrations numériques,
- la dernière partie est consacrée à l'analyse a posteriori pour une équation de Sobolev.

Dans le **Chapitre 1**, nous considérons une approximation numérique de l'équation

quasi-linéaire de type Sobolev suivante

$$\begin{cases} A_1(t,u)u_t + A_2(t,u)u = f(t,u), & \text{dans } V', \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{4}$$

dans un espace de Hilbert V où, pour tout $t \in [0,T]$ et pour tout $u \in V$, $A_1(t,u)$ est un isomorphisme de V sur son dual V', tandis que $A_2(t,u)$ est un opérateur borné de V dans V'. Ces hypothèses sur les opérateur $A_1(t,u)$ et $A_2(t,u)$ permettent de réécrire notre équation comme une équation différentielle ordinaire posée a priori dans V. En d'autres termes, le problème (4) peut s'écrire sous la forme

$$\begin{cases} u_t = F(t,u(t)), & \text{dans } V, \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{5}$$

où $F(t,u(t)) = A_1(t,u(t))^{-1} f(t,u(t)) - A_1(t,u(t))^{-1} A_2(t,u(t))u(t)$. Ce qui nous permet d'utiliser les schémas numériques standards pour les EDO pour la discrétisation temporelle de notre problème.

Notre objectif principal dans ce chapitre est d'effectuer une analyse générale pour des schémas totalement discrets. On utilise une discrétisation abstraite en espace basée sur une famille de sous espaces de dimension finie V_h de V (h joue le rôle de la taille de maille), et une discrétisation en temps basée sur les deux schémas d'Euler et Runge-Kutta d'ordre deux. Ces deux schémas numériques consistent à calculer la solution à l'instant courant en utilisant des valeurs calculées aux instants précédents (sont des schémas explicites !) et en partant de $u(t=0) = u_0$.

Ce chapitre est divisé en deux sections. La Section 1.1 est consacrée à l'étude du cas particulier des équations semi-linéaires de Sobolev, c'est-à-dire le cas où les opérateurs $A_1(t,u)$ et $A_2(t,u)$ ne dépendent ni de la variable temporelle t, ni de la variable spatiale u. Les résultats obtenus sont ensuite généralisés aux équations quasi-linéaires dans la Section 1.2. Dans les deux cas étudiés et pour démontrer l'estimation d'erreur entre la solution exacte de notre problème et la solution totalement discrète, on a d'abord démontré une estimation d'erreur entre la solution exacte et la solution semi-discrète en temps, ensuite on a démontré une estimation d'erreur entre la solution semi-discrète en temps et la solution totalement discrète (ces estimations sont obtenues sous certaines hypothèses de régularité). En combinant ces deux estimations, on trouve le résultat principal de ce chapitre, si $U_{n,h}$ est la solution totalement discrète de la solution exacte u au point t_n obtenue par le schéma d'Euler ou le schéma de Runge-Kutta d'ordre 2, on a l'estimation d'erreur

$$\|u(t_n) - U_{n,h}\| \leq C((\Delta t)^p + h^{q(s)}),$$

pour tout $n = 1, ..., N$, où Δt est le pas de temps, $p = 1$ (resp. $p = 2$) pour le schéma d'Euler (resp. Runge-Kutta), $q(s)$ est lié à des hypothèses abstraites qu'on a supposé (mais en

pratique cela dépend de la régularité de la donnée initiale et l'espace des éléments finis choisi), et C est une constate positive indépendante de h et Δt.

Dans le **Chapitre 2**, on considère une approximation numérique de l'équation linéaire de type Sobolev suivante

$$\begin{cases} u_{\epsilon,t} + \epsilon A_1 u_{\epsilon,t} + A_2 u_\epsilon = f(t), & \text{dans } V', \quad 0 < t \leq T, \\ u_\epsilon(0) = u_0, & \text{dans } V, \end{cases} \tag{6}$$

où ϵ est une constante positive et pour $i = 1$ ou 2, l'opérateur A_i est un isomorphisme de l'espace de Hilbert V sur son dual V'. On suppose de plus que les deux opérateurs sont auto-adjoints définis positifs dans un autre espace de Hilbert H ($V \hookrightarrow H$).

Un problème de type (6) avec $A_2 = A_1$ a également été utilisé pour approximer certains problèmes paraboliques en temps [29], [64]. S'il existe une solution u_ϵ du problème (6), alors lorsque ϵ tend vers 0^+, cette solution et sa dérivée $\frac{\partial u_\epsilon}{\partial t}$ convergent uniformément sur le compact $[0, T]$ vers u et u_t respectivement, où u est la solution du problème d'évolution

$$u_t + A_2 u = f, \tag{7}$$

avec la condition initiale $u(0) = u_0$. Le problème (7) est généralement instable et mal posé dans H dans le sens de Hadamard [35]. En 1960, Lattes et Lions [49] ont introduit la méthode de quasi-réversibilité (Q.R.) qui consiste à modifier légèrement les problèmes de type (7) pour les rendre bien posés et numériquement résolubles. L'idée générale de cette méthode est de remplacer l'opérateur A_2 dans (7) par une perturbation. Dans [64], Showalter a approximé le problème (7) avec une condition finale $u(T) = g$, où $g \in H$ par le problème (6) avec $A_1 = A_2$ et $f = 0$.

Comme dans le Chapitre 1, on peut considérer notre problème comme une équation différentielle ordinaire semi-linéaire à valeurs dans l'espace de Hilbert V. Ce qui permet d'utiliser les schémas numériques standards pour les EDO pour la discrétisation de notre problème en temps. L'opérateur $(I + \epsilon A_1)^{-1} A_2$ est borné de V dans V mais $\|(I + \epsilon A_1)^{-1} A_2\|_{\mathcal{L}(V)}$ tend vers l'infini lorsque ϵ tend vers zéro, et donc l'utilisation d'un schéma numérique explicite n'est pas appropriée dans ce cas, vu que la constante de stabilité dépend de la norme de cet opérateur. Pour cette raison, on a utilisé dans ce chapitre des schémas implicites puisqu'ils sont inconditionnellement stables.

Via le changement d'inconnue $v = (I + \epsilon A_1)^{\frac{1}{2}} u$, et sous quelques hypothèses de régularité, l'utilisation d'un résultat de [71] nous permet de démontrer l'estimation d'erreur suivante entre la solution exacte et la solution totalement discrète (noté $u_{n,h}$)

$$\|u(t_n) - u_{n,h}\|_* \leq C(\epsilon + h^2 + (\Delta t)^p),$$

pour tout $n = 1, ..., N, N \in \mathbb{N}^*$, où $p = 1$ (resp. $p = 2$) si on utilise le schéma d'Euler implicite (resp. de Crank-Nicolson) pour la discrétisation temporelle. Pour démontrer cette estimation, on a décomposé notre problème en un problème homogène avec une condition initiale $u_0 \in V$ et un problème non homogène avec une condition initiale nulle. On a commencé dans la Section 2.1 le problème homogène et ensuite dans la Section 2.2 le problème non homogène.

Dans le **Chapitre 3**, il s'agit de présenter quelques applications, et d'illustrer numériquement les estimations d'erreur obtenues dans les chapitres précédents à l'aide de Freefem++. À chaque application présentée dans ce chapitre et avant de passer aux illustrations numériques, on vérifie d'abord que le problème particulier qu'on a choisi rentre bien dans notre cadre abstrait. Pour la discrétisation en espace, on a utilisé la méthode des éléments finis qui est basée sur un maillage de Ω, et Ω étant un domaine borné de $\mathbb{R}^d, d = 2$ ou 3). Dans la Section 3.1, on présente les applications pour le cas des equations linéaires et semi-linéaires et ensuite dans la Section 3.2 on présente les applications pour le cas des équations quasi-linéaires, enfin, dans la Section 3.3, on présente une application pour le cas d'un problème perturbé singulièrement.

Dans le **Chapitre 4**, nous effectuons une analyse d'erreur a posteriori pour la discrétisation du problème de Sobolev suivant, par la méthode des éléments finis conforme et le schéma d'Euler implicite

$$\begin{cases} L_1 u_t + L_2 u = f & \text{dans } \Omega \times (0,T), \\ u(\cdot,t) = 0 & \text{sur } \partial\Omega \times (0,T), \\ u(\cdot,0) = u_0 & \text{dans } \Omega, \end{cases} \tag{8}$$

où L_i, $i = 1$ ou 2 est un opérateur elliptique d'ordre 2 et $\Omega \subset \mathbb{R}^d, d = 2$ ou 3, un domaine polygonal (polyédrique pour $d = 3$) ouvert et borné, de frontière $\partial\Omega$, et $(0,T)$ un intervalle de temps. Dans la Section 4.1, on présente le problème semi-discrétisé et le problème totalement discrétisé en utilisant le schéma d'Euler implicite et la méthode des éléments finis conforme, dans la Section 4.2, nous rassemblons les notions et les résultats que nous utilisons tout au long de ce chapitre, ensuite, la borne supérieure et inférieure de l'erreur en temps sont obtenues dans la Section 4.3 et la borne supérieure et inférieure de l'erreur en espace sont obtenues dans la Section 4.4. On présente dans la Section 4.5 la borne globale de l'erreur. Enfin, dans la section 4.6, on présente quelques tests numériques.

Chapitre 1

Analyse a priori par un schéma explicite en temps d'un problème de Sobolev

Résumé

Nous considérons un problème quasi-linéaire de type Sobolev dans un espace de Hilbert V. Pour approcher ce problème, nous proposons deux schémas numériques explicites et nous prouvons des estimations d'erreur avec des hypothèses minimales.

Sommaire

1.1 Les équations semi-linéaires **9**
1.1.1 Semi-discrétisation explicite en temps 10
1.1.2 Schéma totalement discrétisé 15
1.2 Les équations quasi-linéaires **22**
1.2.1 Existence et unicité de la solution exacte 23
1.2.2 Semi-discrétisation explicite en temps 25
1.2.3 Schéma totalement discrétisé 29

Soit H un espace de Hilbert réel et soit V un autre espace de Hilbert avec injection continue et dense dans H (on note $V \hookrightarrow H$). On notera par V' l'espace dual de V, par $(\cdot,\cdot)_H$ le produit scalaire de H et par $\langle\cdot,\cdot\rangle_{V',V}$ le produit de dualité entre V' et V. Par l'identification standard de H avec H' on obtient les injections continues et denses $V \hookrightarrow H \approx H' \hookrightarrow V'$. Les normes de V et H seront notées respectivement $\|\cdot\|$ et $\|\cdot\|_H$. Dans ce chapitre, on s'intéresse aux équations quasi-linéaires de la forme suivante

$$\begin{cases} A_1(t,u)u_t + A_2(t,u)u = f(t,u), & \text{dans } V', \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{1.1}$$

où chacun des opérateurs $A_i(t,u) : V \to V'$, $t \in [0,T], u \in V, i = 1,2$, est associé à une

forme bilinéaire $a_i(t;u;\cdot,\cdot)$, par la relation suivante

$$a_i(t;u;v,w) = \langle A_i(t,u)v, w\rangle_{V',V}, \forall v, w \in V. \tag{1.2}$$

Tout au long de ce chapitre, on suppose que

$\mathbf{H_1}$ (continuité uniforme de $a_i(t;u;\cdot,\cdot)$ par rapport à t et u) pour $i = 1$ ou 2, il existe une constante $M_i > 0$ indépendante de t et u telle que pour tout $t \in [0,T]$ et tout $u, v, w \in V$, on a

$$|a_i(t;u;v,w)| \leq M_i\|v\|\|w\|.$$

$\mathbf{H_2}$ (coercivité uniforme de $a_1(t;u;\cdot,\cdot)$ par rapport à t et u) il existe une constante $\alpha > 0$ indépendante de t et u telle que pour tout $t \in [0,T]$ et tout $u, v \in V$, on a

$$a_1(t;u;v,v) \geq \alpha\|v\|^2. \tag{1.3}$$

L'hypothèse $\mathbf{H_1}$ est équivalente à la continuité uniforme (par rapport à t et u) de $A_i(t,u)$ de V dans V', avec $\|A_i(t,u)\|_{\mathcal{L}(V,V')} \leq M_i$ pour tout $t \in [0,T]$ et $u \in V$; tandis que l'hypothèse $\mathbf{H_2}$ et le Lemme de Lax-Milgram assurent que l'opérateur $A_1(t,u)$, est un isomorphisme de V sur V' avec $\|A_1(t,u)^{-1}\|_{\mathcal{L}(V',V)} \leq \frac{1}{\alpha}$ pour tout $t \in [0,T]$ et $u \in V$.
On a aussi besoin d'introduire la restriction de $A_i(t,u)$ à H, à savoir l'opérateur non borné défini par

$$D(A_i(t,u)) = \{v \in V : \exists g \in H \text{ tel que } a_i(t;u;v,w) = (g,w)_H \text{ pour tout } w \in V\},$$

$$A_i(t,u) = g.$$

Comme l'opérateur $A_1(t,u), t \in [0,T], u \in V$ est inversible, on peut donc composer les deux membres de la première équation de (1.1) par $A_1(t,u)^{-1}$ et obtenir le problème équivalent

$$\begin{cases} u_t + \mathcal{A}(t,u)u = g(t,u), & \text{dans } V, \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{1.4}$$

où $\mathcal{A}(t,u) = A_1(t,u)^{-1}A_2(t,u)$ est un opérateur borné de V dans V (uniformément par rapport à t et u grâce aux hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$ avec $\|\mathcal{A}(t,u)\|_{\mathcal{L}(V)} \leq \frac{M_2}{\alpha}$, pour tout $t \in [0,T]$ et $u \in V$) et $g(t,u) = A_1(t,u)^{-1}f(t,u)$. Le problème (1.1) peut s'écrire alors sous la forme suivante

$$\begin{cases} u_t = F(t,u), & \text{dans } V, \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{1.5}$$

où $F(t,u) = g(t,u) - \mathcal{A}(t,u)u$. La solution exacte de ce problème (si elle existe) est donnée par

$$u(t) = u_0 + \int_0^t F(\tau, u(\tau))d\tau. \tag{1.6}$$

On utilisera souvent la définition suivante

Définition 1.0.1. *Soient E et F deux espaces de Hilbert. Une fonction $f : [0,T] \times E \longrightarrow F$ est dite (E,F)-continûment lipschitzienne par rapport à la deuxième variable uniformément en t, s'il existe une constante positive L telle que*

$$\|f(t,v) - f(t,w)\|_F \leq L\|v-w\|_E, \quad \forall v,w \in E, \quad \forall t \in [0,T].$$

Si $E = F$, on dira tout simplement f est E-continûment lipschitzienne par rapport à la deuxième variable uniformément en t.

1.1 Les équations semi-linéaires

Dans ce cas, le problème (1.1) se réduit à

$$\begin{cases} A_1 u_t + A_2 u = f(t,u), & \text{dans } V', \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{1.7}$$

alors que les hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$ sont tout simplement la continuité des formes bilinéaires (indépendantes de t et u) $a_i(\cdot,\cdot)$, $i = 1,2$, et la coercivité de $a_1(\cdot,\cdot)$. Le problème équivalent (1.4) s'écrit sous la forme

$$\begin{cases} u_t + \mathcal{A}u = g(t,u), & \text{dans } V, \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{1.8}$$

où $\mathcal{A} = A_1^{-1}A_2$ est un opérateur borné de V dans V et $g(t,u) = A_1^{-1}f(t,u)$. On a le résultat suivant d'existence et d'unicité pour le problème (1.8) (voir le Théorème 6.1.5 de [59]).

Théorème 1.1.1. *Sous les Hypothèses $\mathbf{H_1}$-$\mathbf{H_2}$, on suppose que $u_0 \in V$ et que f est continûment différentiable de $[0,T] \times V$ dans V'. Alors le problème (1.8) (par équivalence (1.7)) admet une unique solution $u \in C^1([0,T];V)$.*

Si la fonction $f \in C^k([0,T] \times V; V')$, avec $k = 1,2$, on peut en déduire un résultat sur l'ordre de dérivabilité de la solution u, et calculer les dérivées successives de u à l'aide de fonctions construites récursivement à partir de f. C'est utile pour obtenir des majorations, fondées sur la formule de Taylor, de l'erreur commise dans une approximation numérique de u.

Lemme 1.1.2. *Si la fonction $f \in C^k([0,T] \times V; V')$, avec $k = 1,2$, la solution $t \longmapsto u(t)$ de l'équation (1.7) est de classe C^{k+1} sur $[0,T]$ à valeurs dans V.*

Preuve : Pour $k = 1$, il s'agit simplement de l'égalité

$$u_t(t) = -\mathcal{A}u(t) + A_1^{-1}f(t,u(t)), \quad \forall t \in [0,T].$$

Du Théorème 1.1.1 et du fait que les opérateurs $\mathcal{A}$ et A_1^{-1} sont bornés, on voit facilement que la fonction $t \longmapsto u_t(t)$ est continûment dérivable et la formule de dérivation s'écrit

$$u''(t) = -\mathcal{A}u'(t) + A_1^{-1}\left(\frac{\partial f}{\partial t}(t, u(t)) + \frac{\partial f}{\partial u}(t, u(t))u'(t)\right).$$

Si on suppose de plus que $f \in C^2([0,T] \times V; V')$, la fonction $t \longmapsto u''(t)$ est continûment dérivable sur $[0,T]$ et on a

$$u^{(3)}(t) = -\mathcal{A}u''(t) + A_1^{-1}\Big\{\frac{\partial^2 f}{\partial t^2}(t, u(t)) + 2\frac{\partial^2 f}{\partial t \partial u}(t, u(t))u'(t) + \frac{\partial^2 f}{\partial u^2}(t, u(t))(u'(t), u'(t)) + \frac{\partial f}{\partial u}(t, u(t))u''(t)\Big\}.$$

■

1.1.1 Semi-discrétisation explicite en temps

On considère une subdivision régulière $(t_i = i\Delta t)_{i=0}^N$ de l'intervalle $[0,T]$, où $N \in \mathbb{N}^*$ et $\Delta t = \frac{T}{N}$ le pas de discrétisation en temps. Soit $\phi : [0,T] \times V \times [0,\Delta t] \longrightarrow V$ une fonction continue. On cherche à estimer la solution u de (1.5) aux points $(t_n), n = 1, ..., N-1$, en partant de $u(t_0 = 0) = u_0$ et en estimant étape par étape les valeurs de $u(t_{n+1})$ en utilisant la formule de variation des constantes

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} F(\tau, u(\tau))d\tau.$$

Dans la suite, on utilisera une méthode d'approximation à un pas qui consiste à approcher $\int_{t_n}^{t_{n+1}} F(\tau, u(\tau))d\tau$ par la fonction $\Delta t\phi(t_n, u(t_n), \Delta t)$, i.e., la solution approchée du problème (1.5) est donnée par le schéma suivant :

$$\begin{cases} U_0 = u_0, \\ U_{n+1} = U_n + \Delta t\phi(t_n, U_n, \Delta t), \ n = 0, ..., N-1. \end{cases} \tag{1.9}$$

La convergence de ce schéma numérique est basée sur l'estimation de l'erreur de consistance locale.

Définition 1.1.3. *L'erreur de consistance locale E_l relative à la solution exacte u de (1.5) est définie par*

$$E_l(t_{n+1}) = u(t_{n+1}) - u(t_n) - \Delta t\phi(t_n, u(t_n), \Delta t), \quad \forall n = 0, ..., N-1. \tag{1.10}$$

Le théorème suivant est une généralisation d'un résultat très connu pour les équations différentielles ordinaires de la forme $u_t = F(t,u)$, où F est à valeurs dans $\mathbb{R}^k$, $k \in \mathbb{N}^*$ (voir par exemple le Théorème 3.5 de [36]).

Théorème 1.1.4. *Sous les hypothèses du Théorème 1.1.1, soit u la solution exacte du problème (1.5). On suppose que ϕ est lipschitzienne par rapport à la deuxième variable uniformément en t, i.e., il existe une constante positive L telle que*

$$\|\phi(t,u,\Delta t)-\phi(t,v,\Delta t)\|\leq L\|u-v\|,\forall t\in[0,T],u,v\in V, \tag{1.11}$$

et on suppose que l'erreur locale vérifie l'estimation suivante pour un certain $p\in\mathbb{N}$ et pour une certaine constante positive C

$$\|E_l(t_{n+1})\|\leq C(\Delta t)^{p+1},\forall n=0,\ldots,N-1.$$

Alors, l'erreur globale $e_n=u(t_n)-U_n$ satisfait

$$\|e_n\|\lesssim C(\Delta t)^p,\ n=0,1,\ldots,N.$$

Preuve : On a d'après (1.10)

$$u(t_{n+1})=u(t_n)+\Delta t\phi(t_n,u(t_n),\Delta t)+E_l(t_{n+1}),\ n=0,\ldots,N-1.$$

Alors, par (1.9), et comme ϕ est lipschitzienne, on obtient

$$\begin{aligned}\|e_{n+1}\| &\leq \|e_n\|+\Delta t\|\phi(t_n,u(t_n),\Delta t)-\phi(t_n,U_n,\Delta t)\|+\|E_l(t_{n+1})\|\\ &\leq (1+L\Delta t)\|e_n\|+C(\Delta t)^{p+1}.\end{aligned} \tag{1.12}$$

Par itération sur n avec $1\leq n\leq N$, il est facile de démontrer que

$$\|e_n\|\leq(1+L\Delta t)^n\|e_0\|+C(\Delta t)^{p+1}\sum_{k=0}^{n-1}(1+L\Delta t)^{n-k-1}.$$

Comme $e_0=0$, on obtient

$$\begin{aligned}\|e_n\| &\leq C(\Delta t)^{p+1}\sum_{k=0}^{n-1}(1+L\Delta t)^{n-k-1}\\ &\lesssim n(\Delta t)^{p+1}(1+L\Delta t)^N\\ &\lesssim (\Delta t)^p(1+L\Delta t)^N\\ &= (\Delta t)^p\{(1+L\Delta t)^{\frac{1}{L\Delta t}}\}^{TL}\\ &\lesssim (\Delta t)^p(2e)^{TL}.\end{aligned}$$

■

Nous nous concentrons maintenant sur deux schémas particuliers.

Schéma d'Euler explicite

Ce schéma correspond au choix $\phi(t,u,\Delta t)=F(t,u)$, et donc prend la forme

$$\begin{cases} U_0 = u_0. \\ U_{n+1} = U_n + \Delta t F(t_n, U_n),\ n=0,\ \cdots,\ N-1. \end{cases} \tag{1.13}$$

Dans ce cas, on a l'estimation d'erreur suivante.

Proposition 1.1.5. *Sous les hypothèses* $\mathbf{H_1}$ *et* $\mathbf{H_2}$*, on suppose que* $u_0 \in V$ *et* $f \in C^1([0,T]\times V; V')$*. Soient* u *la solution exacte du problème* (1.7) *et* U_n *la solution approchée donnée par le schéma d'Euler explicite* (1.13)*. Alors l'erreur locale* E_l *définie par* (1.10) *vérifie l'estimation suivante*

$$\|E_l(t_{n+1})\| \lesssim (\Delta t)^2,\ \ \forall\, 0 \le n \le N-1. \tag{1.14}$$

En outre, l'erreur globale satisfait

$$\|u(t_n) - U_n\| \lesssim \Delta t,\ \ \forall\, 0 \le n \le N. \tag{1.15}$$

Preuve : Par le développement de Taylor avec reste intégral à l'ordre 1, on a

$$u(t_{n+1}) \ = \ u(t_n) + \Delta t u'(t_n) + \int_{t_n}^{t_{n+1}} (t_{n+1}-\tau) u''(\tau) d\tau.$$

On remplace cette valeur de $u(t_{n+1})$ dans (1.10), on obtient

$$E_l(t_{n+1}) = \int_{t_n}^{t_{n+1}} (t_{n+1}-\tau) u''(\tau) d\tau.$$

Or la fonction $t \longmapsto u''(t)$ est continue sur le compact $[0,T]$ d'après le Lemme 1.1.2, d'où

$$\begin{aligned} \|E_l(t_{n+1})\| &\le \sup_{t\in[0,T]} \|u''(t)\| \int_{t_n}^{t_{n+1}} (t_{n+1}-\tau) d\tau \\ &\lesssim \int_{t_n}^{t_{n+1}} (t_{n+1}-\tau) d\tau = \frac{1}{2}(\Delta t)^2. \end{aligned}$$

Pour la deuxième assertion, notons d'abord que la continuité différentiable de f sur $[0,T]\times V$ dans V' implique que f est (V,V')-continûment lipschitzienne par rapport à la deuxième variable uniformément en t. Alors du fait que A_1 est un isomorphisme de V dans V' et $\mathcal{A}$ est borné de V dans V, la fonction F est V-continûment lipschitzienne par rapport à la deuxième variable uniformément en t, i.e. $\phi(t,u,\Delta t)=F(t,u)$, satisfait l'inégalité (1.11). D'où le résultat en appliquant le Théorème 1.1.4. ∎

Schéma de Heun (Runge-Kutta d'ordre 2)

Cette méthode est un peu plus complexe à mettre en oeuvre car la méthode numérique utilisée pour intégrer la fonction F sur chaque intervalle $[t_n, t_{n+1}]$ est la méthode du Trapèze, i.e.,

$$\phi(t,u,\Delta t) = \frac{1}{2}\left(F(t,u) + F(t+\Delta t, u+\Delta t F(t,u))\right).$$

Cette méthode nécessite donc l'évaluation de F en t_n où la fonction $u(t)$ est déjà connue mais aussi en t_{n+1} où la fonction $u(t)$ n'est pas encore connue, d'où le besoin de trouver une approximation de u en ce point. Dans la méthode de Heun, une première estimation de U_{n+1} est obtenue en utilisant la méthode d'Euler (1.13). Cette estimation est utilisée pour calculer $F(t_{n+1}, u(t_{n+1}))$.
Plus précisément, l'approximation par la méthode de Heun consiste à réaliser le calcul de la récurrence suivante

$$\begin{cases} U_0 = u_0, \\ U^*_{n+1} = U_n + \Delta t F(t_n, U_n), \\ U_{n+1} = U_n + \frac{\Delta t}{2}[F(t_n,U_n) + F(t_{n+1}, U^*_{n+1})], n = 0, \ldots, N-1. \end{cases} \tag{1.16}$$

Proposition 1.1.6. *Sous les Hypothèses* $\mathbf{H_1}$-$\mathbf{H_2}$, *on suppose que* $u_0 \in V$ *et* $f \in C^2([0,T]\times V; V')$. *Alors l'erreur locale définie par* (1.10) *satisfait*

$$\|E_l(t_{n+1})\| \lesssim (\Delta t)^3, \forall\, 0 \le n \le N-1.$$

Par conséquent, on a

$$\|u(t_n) - U_n\| \lesssim (\Delta t)^2, \quad \forall 0 \le n \le N.$$

Preuve : Rappelons que $u \in C^3([0,T];V)$ grâce au Lemme 1.1.2, et alors par le développement de Taylor de u avec reste intégral à l'ordre 2, on a

$$\begin{aligned} u(t_{n+1}) &= u(t_n) + \Delta t u'(t_n) + \frac{(\Delta t)^2}{2}u''(t_n) + \frac{1}{2}\int_{t_n}^{t_{n+1}} (t_{n+1}-\tau)^2 u^{(3)}(\tau)d\tau \\ &= u(t_n) + \Delta t F(t_n, u(t_n)) + \frac{(\Delta t)^2}{2}\frac{\partial F}{\partial t}(t_n, u(t_n)) + \frac{(\Delta t)^2}{2}\frac{\partial F}{\partial u}(t_n, u(t_n))u'(t_n) \\ &\quad + \frac{1}{2}\int_{t_n}^{t_{n+1}} (t_{n+1}-\tau)^2 u^{(3)}(\tau)d\tau. \end{aligned}$$

De plus et comme $f \in C^2([0,T]\times V; V')$ il est facile de voir que $F \in C^2([0,T]\times V; V)$,

donc par le développement de Taylor de F avec reste intégral à l'ordre 1, on a

$$\begin{aligned}F\big(t_{n+1},u(t_n)+\Delta tF(t_n,u(t_n))\big)&=F\big(t_{n+1},u(t_n)+\Delta tu'(t_n)\big)\\&=F(t_n,u(t_n))+\frac{1}{2}DF(t_n,u(t_n))(\Delta t,\Delta tu'(t_n))\\&\quad+\int_0^1(1-\tau)D^2F(t_n+\tau\Delta t,u(t_n)+\tau\Delta tu'(t_n))\\&\qquad\big((\Delta t,\Delta tu'(t_n),(\Delta t,\Delta tu'(t_n))\big)d\tau,\end{aligned}$$

où $DF(t_n,u(t_n))$ est la différentielle au sens de Fréchet de F au point $(t_n,u(t_n))$, i.e.,

$$DF(t_n,u(t_n))(\Delta t,\Delta tu'(t_n))=\Delta t\frac{\partial F}{\partial t}(t_n,u(t_n))+\Delta t\frac{\partial F}{\partial u}(t_n,u(t_n))u'(t_n),$$

et pour $(t,u)\in[0,T]\times V$, $D^2F(t,u)\in\mathcal{L}(\mathbb{R}\times V,\mathbb{R}\times V;V))$ est la différentielle seconde au sens de Fréchet de F au point (t,u) ce qui donne

$$\begin{aligned}D^2F(t,u)\big((\Delta t,\Delta tu'(t_n)),(\Delta t,\Delta tu'(t_n))\big)\;=\;&(\Delta t)^2\frac{\partial^2F}{\partial t^2}(t,u)+2(\Delta t)^2\frac{\partial^2F}{\partial u\partial t}(t,u)u'(t_n)\\&+(\Delta t)^2\frac{\partial^2F}{\partial u^2}(t,u)(u'(t_n),u'(t_n)).\end{aligned}$$

Par conséquent

$$\begin{aligned}\|E_l(t_{n+1})\|&=\|u(t_{n+1})-u(t_n)-\frac{\Delta t}{2}\big[F(t_n,u(t_n))+F\big(t_{n+1},u(t_n)+\Delta tF(t_n,u(t_n))\big)\big]\|\\&=\frac{1}{2}\Big\|\int_0^{\Delta t}\Big\{\tau^2u^{(3)}(\tau)-(\Delta t)^2(1-\frac{\tau}{\Delta t})\Big(\frac{\partial^2F}{\partial t^2}(t_n+\tau,u(t_n)+\tau u'(t_n))\\&\qquad+2\frac{\partial^2F}{\partial t\partial u}(t_n+\tau,u(t_n)+\tau u'(t_n))u'(t_n)\\&\qquad+\frac{\partial^2F}{\partial u^2}\big(t_n+\tau,u(t_n)+\tau u'(t_n)\big)(u'(t_n),u'(t_n))\Big)\Big\}d\tau\Big\|,\end{aligned}$$

par l'inégalité triangulaire et comme la fonction u' est bornée sur $[0,T]$, on obtient

$$\begin{aligned}\|E_l(t_{n+1})\|\lesssim\int_0^{\Delta t}\Big\{&\tau^2\|u^{(3)}(\tau)\|+(\Delta t)^2(1-\frac{\tau}{\Delta t})\Big(\|\frac{\partial^2F}{\partial t^2}(t_n+\tau,u(t_n)+\tau u'(t_n))\|\\&+\|\frac{\partial^2F}{\partial t\partial u}(t_n+\tau\Delta t,u(t_n)+\tau\Delta tu'(t_n))\|\\&+\|\frac{\partial^2F}{\partial u^2}(t_n+\tau\Delta t,u(t_n)+\tau\Delta tu'(t_n))\|\Big)\Big\}d\tau.\end{aligned}$$

D'une part, on obtient grâce au Lemme 1.1.2

$$\begin{aligned}\int_0^{\Delta t} \tau^2 \|u^{(3)}(\tau)\| d\tau &\leq \sup_{t\in[0,T]} \|u^{(3)}(\tau)\| \int_0^{\Delta t} \tau^2 d\tau \\ &\lesssim \int_0^{\Delta t} \tau^2 \lesssim (\Delta t)^3,\end{aligned}$$

D'autre part et comme $F \in C^2([0,T]\times V; V)$, les termes restants de l'intégrale précédente sont majorés par

$$(\Delta t)^3 \sup_{\tau\in[0,T]} \Big\{ \|\frac{\partial^2 F}{\partial t^2}(t_n+\tau, u(t_n)+\tau u'(t_n))\| + \|\frac{\partial^2 F}{\partial t \partial u}(t_n+\tau, u(t_n)+\tau u'(t_n))\| + \|\frac{\partial^2 F}{\partial u^2}(t_n+\tau, u(t_n)+\tau u'(t_n))\| \Big\},$$

d'où la première assertion. Comme pour le schéma d'Euler, il facile de vérifier que

$$\|\phi(t,u,\Delta t) - \phi(t,v,\Delta t)\| \lesssim (1+\Delta t)\|u-v\|, \forall t \in [0,T], u, v \in V,$$

ce qui implique que ϕ est lipschitzienne par rapport à la deuxième variable uniformément en t. La deuxième assertion est alors une conséquence directe du Théorème 1.1.4 et de la proposition précédente. ∎

1.1.2 Schéma totalement discrétisé

Pour un paramètre positif h (qui joue le rôle de la taille de maille), on suppose donné un sous-espace de dimension finie V_h de V et on cherche à construire une approximation totalement discrète du problème (1.1). Pour cela, introduisons quelques notations utiles. Pour $i = 1$ ou 2, on considère une approximation discrète $A_{i,h}$ de A_i, définie par

$$\langle A_{i,h} u_h, v_h\rangle_{V_h', V_h} = a_i(u_h, v_h), \quad \forall u_h, v_h \in V_h,$$

où $a_i(.,.)$ est la forme bilinéaire continue définie sur $V \times V$ par (1.2), i.e.,

$$a_i(u,v) = \langle A_i u, v\rangle_{V',V}, \quad \forall u, v \in V.$$

On définit l'opérateur de projection P_h de V sur V_h associé à la forme bilinéaire $a_1(\cdot,\cdot)$, i.e., pour chaque $v \in V$, $P_h v \in V_h$ est l'unique solution de

$$a_1(P_h v, w_h) = a_1(v, w_h), \forall w_h \in V_h.$$

On introduit de même l'opérateur de projection orthogonale Q_h de V sur V_h associé au produit scalaire $(\cdot,\cdot)_V$. Si on note par

$$\|\cdot\|_1 = \sqrt{a_1(\cdot,\cdot)},$$

la norme sur V associée à la forme bilinéaire $a_1(\cdot,\cdot)$, par la continuité et la coercivité de $a_1(\cdot,\cdot)$ (les hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$), on a

$$\sqrt{\alpha}\|u\| \leq \|u\|_1 \leq \sqrt{M_1}\|u\| \ \forall u \in V. \tag{1.17}$$

On considère maintenant le problème discret (en espace) du problème (1.1)

$$\begin{cases} A_{1,h}u_{h,t} + A_{2,h}u_h = f_h(t,u_h), \text{ dans } V_h, 0 \leq t \leq T, \\ u_h(0) = P_h u_0, \qquad\qquad\qquad \text{dans } V_h, \end{cases} \tag{1.18}$$

où $f_h(t,u_h) = I_h f(t,u_h)$, et $I_h : V' \to V_h'$ est l'opérateur linéaire continu défini de sorte que

$$\langle f_h(t,u_h), v_h\rangle_{V_h',V_h} = \langle f(t,u_h), v_h\rangle_{V',V}, \quad \forall v_h \in V_h.$$

Comme dans le cas continu, les hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$ permettent de conclure grâce au Lemme de Lax-Milgram que $A_{1,h}$ est un isomorphisme de V_h sur V_h'. On peut alors composer la première équation de (1.18) par $A_{1,h}^{-1}$ et obtenir le problème discret équivalent

$$\begin{cases} u_{h,t} + \mathcal{A}_h u_h = g_h(t,u_h), \text{ dans } V_h, 0 \leq t \leq T, \\ u_h(0) = P_h u_0, \qquad\qquad \text{dans } V_h, \end{cases} \tag{1.19}$$

où $\mathcal{A}_h = A_{1,h}^{-1}A_{2,h}$ et $g_h(t,u_h) = A_{1,h}^{-1}f_h(t,u_h)$. Notons que l'opérateur $\mathcal{A}_h$ est borné (uniformément par rapport à h) de V_h dans lui même comme le montre le résultat suivant.

Lemme 1.1.7. *Sous les hypothèses* $\mathbf{H_1}$ *et* $\mathbf{H_2}$*, on a*

$$\|\mathcal{A}_h\|_{\mathcal{L}(V_h)} \leq \frac{M_2}{\alpha}. \tag{1.20}$$

Preuve : Par définition, on a

$$\|\mathcal{A}_h\|_{\mathcal{L}(V_h)} = \sup_{w_h \in V_h, w_h \neq 0} \frac{\|A_{1,h}^{-1}A_{2,h}w_h\|}{\|w_h\|}.$$

Pour un $w_h \in V_h, w_h \neq 0$ fixé, on pose $u_h = A_{1,h}^{-1}A_{2,h}w_h$. Alors u_h appartient à V_h et vérifie

$$a_1(u_h, v_h) = a_2(w_h, v_h), \forall v_h \in V_h,$$

par définition de $a_i, i = 1,2$. En posant $v_h = u_h$ dans cette identité et en tenant compte des hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$, on trouve

$$\alpha\|u_h\|^2 \leq a_2(w_h, u_h) \leq M_2\|w_h\|\|u_h\|,$$

ce qui implique (1.20).

■

Notons que le problème (1.19) peut être écrit de manière équivalente comme une EDO dans V_h :

$$\begin{cases} u_{h,t} = F_h(t, u_h), \text{ dans } V_h, 0 \leq t \leq T, \\ u_h(0) = P_h u_0, \quad \text{dans } V_h, \end{cases} \tag{1.21}$$

où $F_h(t, u_h) = g_h(t, u_h) - \mathcal{A}_h u_h$. Par le schéma d'Euler explicite et le schéma de Runge-Kutta d'ordre 2, la solution approchée du problème (1.21) est donnée par les récurrences suivantes respectivement :

$$\begin{cases} U_{0,h} = P_h u_0 \\ U_{n+1,h} = U_{n,h} + \Delta t F_h(t_n, U_{n,h}), \ n = 0, \ldots, N-1, \end{cases} \tag{1.22}$$

et

$$\begin{cases} U_{0,h} = P_h u_0 \\ U^*_{n+1,h} = U_{n,h} + \Delta t F_h(t_n, U_{n,h}), \\ U_{n+1,h} = U_{n,h} + \frac{\Delta t}{2}[F_h(t_n, U_{n,h}) + F_h(t_{n+1}, U^*_{n+1,h})], \ n = 0, \ldots, N-1. \end{cases} \tag{1.23}$$

Remarque 1.1.8. *Pour chaque $u_h \in V_h$, les hypothèses* $\mathbf{H_1}$ *et* $\mathbf{H_2}$ *impliquent que*

$$g_h(t, u_h) = P_h g(t, u_h), \ \forall t \in [0, T]. \tag{1.24}$$

En effet, comme $A_{1,h} g_h(t, u_h) = I_h f(t, u_h)$, on obtient

$$a_1(g_h(t, u_h), v_h) = \langle f(t, u_h), v_h \rangle_{V',V}, \quad \forall v_h \in V_h.$$

Par la dédinition de P_h, on a

$$a_1(P_h g(t, u_h), v_h) = a_1(g(t, u_h), v_h) = a_1(A_1^{-1} f(t, u_h), v_h),$$

et par la définition de la forme bilinéaire $a_1(\cdot, \cdot)$, on trouve

$$a_1(g_h(t, u_h), v_h) = a_1(P_h g(t, u_h), v_h), \quad \forall v_h \in V_h,$$

d'où (1.24) par le lemme de Lax-Milgram.

Notre analyse d'erreur du schéma totalement discret est basée sur les hypothèses suivantes.

$\mathbf{H_3}$ Il existe deux espaces de Hilbert D_s et $\tilde{D}_{s-1}$ ($s \geq 1$ est un paramètre) tels que $D_s \hookrightarrow V$ et que $A_1(t, u)$ est un isomorphisme de D_s dans $\tilde{D}_{s-1}$, et $A_2(t, u)$ est borné de D_s dans $\tilde{D}_{s-1}$, avec

$$\|A_1(t, u)^{-1}\|_{\mathcal{L}(\tilde{D}_{s-1}, D_s)} + \|A_2(t, u)\|_{\mathcal{L}(D_s, \tilde{D}_{s-1})} \lesssim 1,$$

pour tout $t \in [0, T]$ et $u \in V$.

$\mathbf{H_4}$ Pour tout paramètre s de l'hypothèse $\mathbf{H_3}$, il existe un nombre réel positif $q(s)$ tel que

$$\|\varphi - Q_h\varphi\| \lesssim h^{q(s)}\|\varphi\|_{D_s}, \quad \forall \varphi \in D_s. \tag{1.25}$$

En pratique, sous l'hypothèse $V \hookrightarrow H$, D_s correspond au domaine des puissances de $A_1(t,u)$ ou à un sous-domaine de D_s (ainsi $\tilde{D}_{s-1} = A_1(t,u)D_s$), tandis que l'estimation (1.25) provient d'une estimation d'erreur d'interpolation. En particulier pour $s = 1$, on peut choisir $D_1 = D(A_1(t;u))$ et $\tilde{D}_0 = H$, si $D(A_1(t;u))$ est indépendant de t et de u, et si on peut vérifier les hypothèses précédentes. On réfère au chapitre 3 pour quelques illustrations concrètes. L'hypothèse $\mathbf{H_3}$ est écrite dans le cas général (i.e. dans le cas où $A_i(t,u)$ dépend de t et u) pour pouvoir l'utiliser pour les équations quasi-linéaires.

Remarque 1.1.9. *Sous les hypothèses* $\mathbf{H_1}$-$\mathbf{H_4}$, *on a*

$$\|\varphi - P_h\varphi\| \lesssim h^{q(s)}\|\varphi\|_{D_s}, \quad \forall \varphi \in D_s. \tag{1.26}$$

En effet, pour $\varphi \in D_s$, *on a par* (1.17)

$$\|\varphi - P_h\varphi\| \leq \frac{1}{\sqrt{\alpha}}\|\varphi - P_h\varphi\|_1 \leq \frac{1}{\sqrt{\alpha}}\|\varphi - Q_h\varphi\|_1 \leq \sqrt{\frac{M_1}{\alpha}}\|\varphi - Q_h\varphi\|,$$

d'où (1.26) *grâce à* $\mathbf{H_4}$.

Dans l'analyse du problème semi-discrétisé en temps, on avait besoin de supposer que f est continûment différentiable sur $[0,T] \times V$ dans V' pour assurer l'existence de la solution exacte, mais pour démontrer les résultats de cette sous-section (à l'exception du Corollaire 1.1.14 où intervient la solution exacte), il suffit de supposer que

$\mathbf{H_5}$ f est (V, V')-continûment lipschitzienne par rapport à la deuxième variable uniformément en t.

Proposition 1.1.10. *Sous les hypothèses* $\mathbf{H_1}$-$\mathbf{H_5}$, *si on suppose de plus que* $f \in C([0,T] \times V; \tilde{D}_{s-1})$, *alors pour tout* $t \in [0,T]$ *et* $(v, v_h) \in D_s \times V_h$, *on a*

$$\|F_h(t,v_h) - F(t,v)\| \lesssim \|v - v_h\| + h^{q(s)}(\|v\|_{D_s} + \|f(t,v)\|_{\tilde{D}_{s-1}}). \tag{1.27}$$

Preuve : Par la définition de F_h et F et l'inégalité triangulaire, on a pour tout $t \in [0,T]$ et tout $(v, v_h) \in D_s \times V_h$,

$$\|F_h(t,v_h) - F(t,v)\| \leq \|g_h(t,v_h) - g(t,v)\| + \|\mathcal{A}v - \mathcal{A}_h v_h\|. \tag{1.28}$$

Commençons par estimer le premier terme du membre de droite de cette inégalité, on a par l'inégalité triangulaire

$$\|g_h(t,v_h) - g(t,v)\| \leq \|g_h(t,v_h) - P_h g(t,v)\| + \|P_h g(t,v) - g(t,v)\|.$$

D'une part, d'après la remarque précédente et le fait que A_1^{-1} est borné de V' dans V (l'hypothèse $\mathbf{H_1}$) et la continuité lipschitzienne de f par rapport à la deuxième variable uniformément en t (l'hypothèse $\mathbf{H_5}$), on a

$$\begin{aligned}\|g_h(t,v_h)-P_hg(t,v)\| = \|P_h\big(g(t,v_h)-g(t,v)\big)\| &\lesssim \|g(t,v_h)-g(t_n,v)\| \\ &= \|A_1^{-1}\big(f(t,v_h)-f(t,v)\big)\| \\ &\leq \|f(t,v_h)-f(t,v)\| \\ &\lesssim \|v_h-v\|.\end{aligned}$$

D'autre part et grâce à (1.17), l'hypothèse $\mathbf{H_4}$ et le fait que A_1^{-1} est borné de D_s dans $\tilde{D}_{s-1}$, (l'hypothèse $\mathbf{H_3}$) on obtient

$$\|P_hg(t,v)-g(t,v)\| \leq h^{q(s)}\|g(t,v)\|_{D_s} \lesssim h^{q(s)}\|f(t,v)_{\tilde{D}_{s-1}}.$$

Par conséquent

$$\|g_h(t,v_h)-g(t,v)\| \lesssim \|v_h-v\| + h^{q(s)}\|f(t,v)\|_{\tilde{D}_{s-1}}. \tag{1.29}$$

Reste alors à estimer le deuxième terme, on a

$$\|\mathcal{A}v-\mathcal{A}_hv_h\| \leq \|\mathcal{A}v-P_h\mathcal{A}v\| + \|P_h\mathcal{A}v-\mathcal{A}_hv_h\|.$$

Comme $\mathcal{A}v \in D_s$ du fait que $v \in D_s$, on en déduit des hypothèses $\mathbf{H_4}$ et $\mathbf{H_3}$ que

$$\|\mathcal{A}v-P_h\mathcal{A}v\| \lesssim h^{q(s)}\|\mathcal{A}v\|_{D_s} \lesssim h^{q(s)}\|v\|_{D_s}. \tag{1.30}$$

Posons maintenant $w=\mathcal{A}v$ et $w_h=\mathcal{A}_hv_h$, donc par définition de a_1 et a_2, on a

$$a_1(w,\chi)=a_2(v,\chi) \quad \text{et} \quad a_1(w_h,\chi)=a_2(v_h,\chi),$$

pour tout $\chi \in V_h$. On a aussi par définition de P_h,

$$a_1(P_hw,\chi)=a_1(w,\chi), \ \ \forall \chi \in V_h.$$

Par conséquent

$$a_1(P_hw-w_h,\chi)=a_2(v-v_h,\chi), \ \ \forall \chi \in V_h. \tag{1.31}$$

En choisissant dans (1.31), $\chi = P_hw-w$ on obtient grâce à la coercivité de a_1 et la continuité de a_2

$$\|P_hw-w_h\| \lesssim \|v-v_h\|. \tag{1.32}$$

De (1.30) et (1.32), on en déduit que

$$\|\mathcal{A}v-\mathcal{A}_hv_h\| \lesssim h^{q(s)}\|v\|_{D_s} + \|v-v_h\|,$$

d'où (1.27). ∎

Lemme 1.1.11. *Supposons que $u_0 \in D_s$, que $f \in C([0,T]\times V;\tilde{D}_{s-1})$ est $(V,\tilde{D}_{s-1})$-continûment lipschitzienne par rapport à la deuxième variable uniformément en t, et que l'hypothèse* $\mathbf{H_3}$ *est satisfaite. Alors la suite $(U_n)_{n=1}^N$ définie par (1.16) est bornée dans D_s, i.e.,*

$$\|U_n\|_{D_s} \lesssim 1,\ \forall n = 1,\dots,N.$$

Preuve : Par construction (voir (1.16)), on a

$$\begin{aligned}\|U_{n+1}\|_{D_s} &\le \|U_n\|_{D_s} + \frac{\Delta t}{2}\Big(\|F(t_n,U_n)\|_{D_s} + \|F\big(t_{n+1},U_n+\Delta t F(t_n,U_n)\big)\|_{D_s}\Big)\\ &\le \|U_n\|_{D_s} + \frac{\Delta t}{2}\Big(\|F(t_n,U_n)-F(t_n,0)\|_{D_s} + \|F(t_n,0)\|_{D_s}\\ &\qquad +\|F(t_{n+1},U_n+\Delta t F(t_n,U_n))-F(t_{n+1},0)\|_{D_s} + \|F(t_{n+1},0)\|_{D_s}\Big).\end{aligned}$$

Par la définition de F et le fait que A_1^{-1} est borné de $\tilde{D}_{s-1}$ dans D_s, on a :
D'une part,

$$\|F(t_n,0)\|_{D_s} = \|A_1^{-1}f(t_n,0)\|_{D_s} \lesssim \|f(t_n,0)\|_{\tilde{D}_{s-1}},\ \forall\, n = 0,1,\dots,N. \tag{1.33}$$

D'autre part, sachant que $\mathcal{A}$ est borné dans D_s et compte tenu de l'hypothèse de lipschitzianité de f, on trouve pour tout $u,v \in D_s$, et pour tout $t\in[0,T]$

$$\begin{aligned}\|F(t,u)-F(t,v)\|_{D_s} &\le \|\mathcal{A}(u-v)\|_{D_s} + \|g(t,u)-g(t,v)\|_{D_s}\\ &\le \|u-v\|_{D_s} + \|A_1^{-1}(f(t,u)-f(t,v))\|_{D_s}\\ &\lesssim \|u-v\|_{D_s} + \|f(t,u)-f(t,v)\|_{\tilde{D}_{s-1}}\\ &\lesssim \|u-v\|_{D_s} + \|u-v\|\\ &\lesssim \|u-v\|_{D_s}.\end{aligned} \tag{1.34}$$

Alors

$$\begin{aligned}\|U_{n+1}\|_{D_s} &\le \|U_n\|_{D_s} + \frac{\Delta t}{2}\Big(2\|U_n\|_{D_s} + \Delta t\|F(t_n,U_n)\|_{D_s} + \|f(t_n,0)\|_{\tilde{D}_{s-1}} + \|f(t_{n+1},0)\|_{\tilde{D}_{s-1}}\Big)\\ &\lesssim (1+\Delta t)\|U_n\|_{D_s} + \frac{(\Delta t)^2}{2}\|F(t_n,U_n)-F(t_n,0)\|_{D_s} + \frac{\Delta t}{2}(1+\Delta t)\|f(t_n,0)\|_{\tilde{D}_{s-1}}\\ &\qquad +\frac{\Delta t}{2}\|f(t_{n+1},0)\|_{\tilde{D}_{s-1}}\\ &\lesssim (1+\Delta t)\|U_n\|_{D_s} + \Delta t\|f(t_n,0)\|_{\tilde{D}_{s-1}} + \Delta t\|f(t_{n+1},0)\|_{\tilde{D}_{s-1}}.\end{aligned}$$

Par récurrence, on démontre que

$$\|U_n\|_{D_s} \lesssim (1+\Delta t)^n\|u_0\|_{D_s} + \Delta t\sum_{k=0}^{n-1}(1+\Delta t)^{n-1-k}\big(\|f(t_k,0)\|_{\tilde{D}_{s-1}} + \|f(t_{k+1},0)\|_{\tilde{D}_{s-1}}\big).$$

Comme pour tout $0\le k\le N$

$$(1+\Delta t)^k \le (1+\Delta t)^N \lesssim e^T,$$

le résultat découle puisque f est continue de $[0,T]\times V$ à valeurs dans $\tilde{D}_{s-1}$. ∎

Lemme 1.1.12. *Sous les hypothèses* $\mathbf{H_1}$-$\mathbf{H_5}$ *et les hypothèses du Lemme 1.1.11, soit* U_n *(resp.* $U_{n,h}$*) la solution approchée donnée par (1.16) (resp. (1.23)). Alors l'erreur* $e_{n,h} = U_n - U_{n,h}$ *satisfait l'estimation suivante*

$$\|e_{n,h}\| \lesssim h^{q(s)}.$$

Preuve : Par l'inégalité triangulaire, on a

$$\begin{aligned}\|e_{n+1,h}\| \leq \|e_{n,h}\| &+ \frac{\Delta t}{2}\|F_h(t_n, U_{n,h}) - F(t_n, U_n)\| \\ &+ \frac{\Delta t}{2}\|F_h\big(t_{n+1}, U_{n,h} + \Delta t F_h(t_n, U_{n,h})\big) - F\big(t_{n+1}, U_n + \Delta t F(t_n, U_n)\big)\|.\end{aligned}$$

Puisque $u_0 \in D_s$, on vérifie facilement par récurrence que $U_n \in D_s$, $\forall\ 1 \leq n \leq N$, donc grâce à la Proposition 1.1.10

$$\|F_h(t_n, U_{n,h}) - F(t_n, U_n)\| \lesssim \|e_{n,h}\| + h^{q(s)}\big(\|U_n\|_{D_s} + \|f(t_n, U_n)\|_{\tilde{D}_{s-1}}\big).$$

En appliquant aussi la même proposition et en utilisant l'inégalité précédente, on trouve

$$\begin{aligned}&\| F_h\big(t_{n+1}, U_{n,h} + \Delta t F_h(t_n, U_{n,h})\big) - F\big(t_{n+1}, U_n + \Delta t F(t_n, U_n)\big)\| \\ &\lesssim (1+\Delta t)\|e_{n,h}\| + h^{q(s)}\big(\|U_n\|_{D_s} + \|f(t_n, U_n)\|_{\tilde{D}_{s-1}} + \|f\big(t_{n+1}, U_n + \Delta t F(t_n, U_n)\big)\|_{\tilde{D}_{s-1}}\big).\end{aligned}$$

Par conséquent

$$\begin{aligned}\|e_{n+1,h}\| \lesssim (1+\Delta t)&\|e_{n,h}\| \\ &+h^{q(s)}\Delta t\big(\|U_n\|_{D_s} + \|f(t_n, U_n)\|_{\tilde{D}_{s-1}} + \|f(t_{n+1}, U_n + \Delta t F(t_n, U_n))\|_{\tilde{D}_{s-1}}\big).\end{aligned}$$

Par récurrence, il est facile de démontrer que

$$\begin{aligned}\|e_{n,h}\| \lesssim (1+\Delta t)^n\|e_{0,h}\| + h^{q(s)}\Delta t\sum_{k=0}^{n-1}(1+\Delta t)^{n-k-1}\Big(&\|U_k\|_{D_s} \\ &+\|f(t_k, U_k)\|_{\tilde{D}_{s-1}} + \|f(t_{k+1}, U_k + \Delta t F(t_k, U_k))\|_{\tilde{D}_{s-1}}\Big).\end{aligned} \tag{1.35}$$

Comme $u_0 \in D_s$, on a grâce à l'hypothèse $\mathbf{H_4}$

$$\|e_{0,h}\| = \|u_0 - P_h u_0\| \lesssim h^{q(s)}\|u_0\|_{D_s}. \tag{1.36}$$

De l'hypothèse f est $(V, \tilde{D}_{s-1})$-continûment lipschitzienne, et puisque D_s s'injecte continûment dans V, on obtient

$$\|f(t_k, U_k)\|_{\tilde{D}_{s-1}} \lesssim \|U_k\|_{D_s} + \|f(t_k, 0)\|_{\tilde{D}_{s-1}}. \tag{1.37}$$

De même,

$$\begin{aligned}\|f(t_{k+1}, U_k + \Delta t F(t_k, U_k))\|_{\tilde{D}_{s-1}} &\lesssim \|U_k\|_{D_s} + \Delta t\|F(t_k, U_k)\|_{D_s} + \|f(t_{k+1}, 0)\|_{\tilde{D}_{s-1}} \\ &\lesssim \|U_k\|_{D_s} + \|f(t_{k+1}, 0)\|_{\tilde{D}_{s-1}} \\ &+ \Delta t\big(\|F(t_k, U_k) - F(t_k, 0)\|_{D_s} + \|F(t_k, 0)\|_{D_s}\big).\end{aligned}$$

Utilisons les estimations (1.33) et (1.34) démontrées dans la preuve du lemme précédent, on obtient

$$\|f(t_{k+1}, U_k + \Delta t F(t_k, U_k)\|_{\tilde{D}_{s-1}} \lesssim (1+\Delta t)\|U_k\|_{D_s} + \|f(t_{k+1},0)\|_{\tilde{D}_{s-1}} + \Delta t\|f(t_k,0)\|_{\tilde{D}_{s-1}}.$$

Insérons cette dernière inégalité avec (1.36) et (1.37) dans (1.36), on en déduit grâce au Lemme 1.1.11 et la continuité de f de $[0,T]\times V$ dans $\tilde{D}_{s-1}$ que

$$\|e_{n,h}\| \lesssim h^{q(s)}\|u_0\|_{D_s} + h^{q(s)} n.\Delta t.$$

■

Remarque 1.1.13. *Sous les mêmes hypothèses du Lemme 1.1.12, si on approxime la solution du problème (1.7) par le schéma d'Euler (1.22), on peut démontrer comme dans le cas d'un schéma de Runge-Kutta que*

$$\|U_n - U_{n,h}\| \lesssim h^{q(s)}.$$

Corollaire 1.1.14. *Sous les hypothèses* $\mathbf{H_1}$-$\mathbf{H_4}$, *on suppose que* $u_0 \in D_s$, $f \in C([0,T]\times V;\tilde{D}_{s-1})$, *que* f *est* $(V,\tilde{D}_{s-1})$*-continûment lipschitzienne par rapport à* u *uniformément en* t, *et que* $f \in C^p([0,T]\times V;V')$, $p=1$ *ou* 2. *Soient* u *la solution exacte du problème (1.7) et* $U_{n,h}$ *la solution approchée donnée par (1.22) pour* $p=1$ *(resp. (1.23) pour* $p=2$*). Alors*

$$\|u(t_n) - U_{n,h}\| \lesssim (\Delta t)^p + h^{q(s)}. \tag{1.38}$$

Preuve : Conséquence directe :

- de la Remarque 1.1.13 et de la Proposition 1.1.5 pour le schéma d'Euler,
- du Lemme 1.1.12 et de la Proposition 1.1.6 pour le schéma de Runge-Kutta.

■

1.2 Les équations quasi-linéaires

Après avoir étudié à la section précédente des équations de Sobolev semi-linéaires, nous nous intéressons dans cette section aux équations de Sobolev quasi-linéaires (1.1), i.e., on considère le problème quasi-linéaire

$$\begin{cases} A_1(t,u)u_t + A_2(t,u)u = f(t,u), & \text{dans } V', \quad 0 < t \le T, \\ u(0) = u_0, & \text{dans } V. \end{cases}$$

Il est bien connu qu'il n'y a pas existence de solution classique globale de ce type de problèmes de Cauchy. T. Kato et A. Pazy ont montré un résultat d'existence locale de la solution (la solution existe pour un t suffisamment petit) sous certaines hypothèses sur la partie quasi-linéaire et la partie semi-linéaire de l'équation de Cauchy (voir [59, 43]).

1.2.1 Existence et unicité de la solution exacte

Afin d'établir l'existence et l'unicité de la solution locale du problème (1.7), on a besoin d'hypothèses supplémentaires. Soit W une boule ouverte dans V.

$\mathbf{H_6}$ Il existe $\gamma \in (0,1]$ tel que pour $i = 1$ ou 2

$$\big|a_i(t;y;u,v) - a_i(s;z;u,v)\big| \lesssim \big(|t-s|^\gamma + \|y-z\|\big)\|u\|\|v\|, \forall y,z,u,v \in V, t,s \in [0,T]. \tag{1.39}$$

$\mathbf{H_7}$ f est une fonction bornée sur $[0,T]\times V$ dans V' :

$$\|f(t,v)\|_{V'} \lesssim 1, \quad \forall\, t \in [0,T],\ \forall\, v \in V.$$

$\mathbf{H_8}$ Pour tout $v \in W$, la fonction $t \to f(t,v)$ est continue de $[0,T]$ dans V'.

Donnons maintenant quelques conséquences de ces hypothèses.

Lemme 1.2.1. *Sous les hypothèses* $\mathbf{H_1}$, $\mathbf{H_2}$ *et* $\mathbf{H_6}$, *la fonction* $(t,v) \to A_1(t,v)^{-1}$ *est continûment lipschitzienne sur* $[0,T]\times V$ *pour la norme de* $\mathcal{L}(V',V)$ *uniformément en* t *et en* v, *i.e.,*

$$\|A_1(t,v)^{-1} - A_1(t_0,v_0)^{-1}\|_{\mathcal{L}(V',V)} \lesssim |t-t_0|^\gamma + \|v-v_0\|,\ \forall t,t_0 \in [0,T], v,v_0 \in V. \tag{1.40}$$

Preuve : Soient $v, v_0 \in V$ et soient $t, t_0 \in [0,T]$ fixés arbitrairement. Alors par définition, on a

$$\|A_1(t,v)^{-1} - A_1(t_0,v_0)^{-1}\|_{\mathcal{L}(V',V)} = \sup_{h\in V', h\neq 0} \frac{\|A_1(t,v)^{-1}h - A_1(t_0,v_0)^{-1}h\|}{\|h\|_{V'}}.$$

Maintenant pour un $h \in V', h \neq 0$, si on pose $\phi_1 = A_1(t,v)^{-1}h$, et $\phi_2 = A_1(t_0,v_0)^{-1}h$, alors

$$\begin{cases} a_1(t;v;\phi_1,\psi) = \langle h,\psi\rangle_{V',V}, & \forall \psi \in V, \\ a_1(t_0;v;\phi_2,\psi) = \langle h,\psi\rangle_{V',V}, & \forall \psi \in V, \end{cases}$$

ce qui donne

$$a_1(t;v;\phi_1-\phi_2,\psi) = a_1(t_0;v_0;\phi_2,\psi) - a_1(t;v;\phi_2,\psi),\ \forall \psi \in V.$$

En choisissant $\psi = \phi_1 - \phi_2$, et en utilisant les hypothèses $\mathbf{H_2}$ et $\mathbf{H_6}$, on obtient

$$\begin{aligned} \|\phi_1-\phi_2\| \lesssim (|t-t_0|^\gamma + \|v-v_0\|)\|\phi_2\| &= (|t-t_0|^\gamma + \|v-v_0\|)\|A_1(t_0,v_0)^{-1}h\| \\ &\lesssim (|t-t_0|^\gamma + \|v-v_0\|)\|h\|_{V'}, \end{aligned}$$

ce qui prouve l'estimation (1.40). ∎

Corollaire 1.2.2. *Sous les hypothèses* $\mathbf{H_1}$, $\mathbf{H_2}$ *et* $\mathbf{H_6}$, *la fonction* $(t,v) \to \mathcal{A}(t,v)$ *est continûment lipschitzienne sur* $[0,T]\times V$ *pour la norme* $\mathcal{L}(V)$ *uniformément en* t, v, *i.e.,*

$$\|\mathcal{A}(t,v)-\mathcal{A}(t_0,v_0)\|_{\mathcal{L}(V)} \lesssim |t-t_0|^\gamma + \|v-v_0\|, \quad \forall t,t_0\in[0,T],\ \forall v,v_0\in V. \tag{1.41}$$

Et alors

$$\|\mathcal{A}(t,v)v-\mathcal{A}(t,v_0)v_0\| \lesssim (1+\|v_0\|)\|v-v_0\|, \quad \forall t\in[0,T],\ \forall v,v_0\in V. \tag{1.42}$$

Preuve : Par définition et pour $t,t_0\in[0,T], v,v_0\in V$ fixés, on a

$$\begin{aligned}
\|\mathcal{A}(t,v)-\mathcal{A}(t_0,v_0)\|_{\mathcal{L}(V)} &\leq \|A_1(t,v)^{-1}(A_2(t,v)-A_2(t_0,v_0))\|_{\mathcal{L}(V)} \\
&+ \|(A_1(t,v)^{-1}-A_1(t_0,v_0)^{-1})A_2(t_0,v_0)\|_{\mathcal{L}(V)} \\
&\leq \|(A_1(t,v)^{-1}\|_{\mathcal{L}(V',V)}\|\|A_2(t,v)-A_2(t_0,v_0)\|_{\mathcal{L}(V,V')} \\
&+ \|A_1(t,v)^{-1}-A_1(t_0,v_0)^{-1}\|_{\mathcal{L}(V',V)}\|A_2(t_0,v_0)\|_{\mathcal{L}(V,V')}.
\end{aligned}$$

Puisque l'hypothèse $\mathbf{H_6}$ pour $i=2$ est équivalente à

$$\|A_2(t,v)-A_2(t_0,v_0)\|_{\mathcal{L}(V',V)} \lesssim |t-t_0|^\gamma + \|v-v_0\|,$$

on conclut (1.41) par (1.40) et les hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$. Par l'inégalité triangulaire et (1.41), on a

$$\begin{aligned}
\|\mathcal{A}(t,v)v-\mathcal{A}(t,v_0)v_0\| &\leq \|\mathcal{A}(t,v)(v-v_0)\| + \|\mathcal{A}(t,v)v_0-\mathcal{A}(t,v_0)v_0\| \\
&\lesssim \|v-v_0\| + \|v-v_0\|\|v_0\|,
\end{aligned}$$

d'où (1.42). ∎

Corollaire 1.2.3. *Sous les hypothèses* $\mathbf{H_1}$,$\mathbf{H_2}$ *et* $\mathbf{H_5}$-$\mathbf{H_8}$, *la fonction* $g(t,u)=A_1(t,u)^{-1}f(t,u)$ *satisfait les propriétés suivantes*

1. *g est bornée sur $[0,T]\times V$ i.e.*

$$\|g(t,v)\| \lesssim 1, \quad \forall u,v\in V,\ \forall t\in[0,T],$$

2. *g est V-continûment lipschitzienne par rapport à la deuxième variable uniformément en t i.e.*

$$\|g(t,u)-g(t,v)\| \lesssim \|u-v\|, \quad \forall u,v\in V,\ \forall t\in[0,T]. \tag{1.43}$$

3. *Pour tout $v\in W$, la fonction $t\to g(t,v)$ est continue de $[0,T]$ dans V.*

Preuve :

1. Conséquence directe des hypothèses $\mathbf{H_2}$ et $\mathbf{H_7}$.
2. Soient $v, v_0 \in V$ et $t, t_0 \in [0,T]$, alors par l'hypothèse $\mathbf{H_2}$, on peut écrire

$$\begin{aligned}\|g(t,v) - & g(t_0,v_0)\| \\ &\leq \|A_1(t,v)^{-1}\left(f(t,v)-f(t_0,v_0)\right)\| + \|\left(A_1(t,v)^{-1}-A_1(t_0,v_0)^{-1}\right)f(t_0,v_0)\| \\ &\lesssim \|f(t,v)-f(t_0,v_0)\|_{V'} + \|A_1(t,v)^{-1}-A_1(t_0,v_0)^{-1}\|_{\mathcal{L}(V',V)}\|f(t_0,v)\|_{V'}.\end{aligned}$$

   Donc grâce à l'hypothèse $\mathbf{H_7}$ et (1.40), on obtient

$$\|g(t,v)-g(t_0,v_0)\| \leq \|f(t,v)-f(t_0,v_0)\|_{V'} + |t-t_0|^\gamma + \|v-v_0\|. \tag{1.44}$$

   En particulier, si $t_0 = t$, cette estimation et la propriété de lipschitzianité de f induisent (1.43).
3. Si $v_0 = v \in W$ dans l'estimation (1.44), on trouve

$$\|g(t,v)-g(t_0,v)\| \leq \|f(t,v)-f(t_0,v)\|_{V'} + |t-t_0|^\gamma,$$

   et la continuité de g découle de l'hypothèse $\mathbf{H_8}$.

Nous sommes maintenant en mesure de démontrer notre résultat d'existence.

Théorème 1.2.4. *On fixe une boule ouverte W de V et on suppose que $u_0 \in W$. Sous les hypothèses $\mathbf{H_1}$, $\mathbf{H_2}$ et $\mathbf{H_5}$-$\mathbf{H_8}$, il existe $T' \in (0,T]$ tel que le problème (1.4) (et alors (1.1) et (1.5)) admet une unique solution*

$$u \in C([0,T'];W) \cap C^1([0,T'];V).$$

Preuve : On applique le Théorème 6 de [43] avec l'espace de Hilbert $X = Y = V$ et la boule fixée W. Les hypothèses (A1), (A2) et (A4) de ce Théorème sont trivialement satisfaites car les opérateurs $\mathcal{A}(t,v)$ sont bornés dans V, l'hypothèse (A3) est satisfaite grâce au Corollaire 1.2.2, tandis que l'hypothèse (f1) est satisfaite grâce au Corollaire 1.2.3. ∎

1.2.2 Semi-discrétisation explicite en temps

On considère un entier $N > 0$ et on pose $\Delta t = \frac{T'}{N}$ le pas uniforme de discrétisation en temps, où T' est le temps d'existence de u. Pour tout $0 \leq n \leq N$, on note U_n l'approximation de $u(t_n)$, $t_n = n\Delta t$. Comme dans le cas d'un problème semi-linéaire, on utilisera ici les deux schémas d'Euler explicite et de Runge-Kutta qui consistent à : partant de $U_0 = u_0$, chercher U_{n+1} tel que

$$U_{n+1} = U_n + \Delta t F(t_n, U_n),\ n = 0, ..., N, \tag{1.45}$$

pour le schéma d'Euler, et

$$\begin{cases} U_{n+1}^* = U_n + \Delta t F(t_n, U_n), \\ U_{n+1} = U_n + \frac{\Delta t}{2}\big[F(t_n, U_n) + F(t_{n+1}, U_{n+1}^*)\big], \ n = 0, ..., N, \end{cases} \tag{1.46}$$

pour celui de Runge-Kutta.

Lemme 1.2.5. *On suppose que $u_0 \in W$. Sous les hypothèses* $\mathbf{H_1}$, $\mathbf{H_2}$ *et* $\mathbf{H_5}$-$\mathbf{H_7}$, *la fonction F est localement lipschitzienne par rapport à la deuxième variable uniformément en t, i.e. pour tout $t \in [0, T]$*

$$\|F(t, u) - F(t, v)\| \lesssim \|u - v\|, \ \forall u \in W, \forall v \in V.$$

Preuve : Par définition de F, il est facile de voir que

$$\|F(t, u) - F(t, v)\| \leq \|g(t, u) - g(t, v)\| + \|\mathcal{A}(t, v)v - \mathcal{A}(t, u)u\|.$$

Le résultat s'en suit grâce au Corollaire 1.2.3, 2., à l'inégalité (1.42) et au fait que $u \in W$. ∎

Comme la fonction F est seulement localement lipschitzienne, on a besoin d'une version légèrement modifiée du Théorème 1.1.4.

Théorème 1.2.6. *Sous les hypothèses du Théorème 1.2.4, soit u la solution exacte du problème (1.5). On suppose que ϕ est localement lipschitzienne par rapport à la deuxième variable uniformément par rapport à t, i.e., il existe une constante positive L indépendante de t et Δt telle que*

$$\|\phi(t, u, \Delta t) - \phi(t, v, \Delta t)\| \leq L\|u - v\|, \ \forall t \in [0, T], u \in W, v \in V, \tag{1.47}$$

et on suppose qu'il existe $p \in \mathbb{N}$ tel que l'erreur locale vérifie

$$\|E_l(t_{n+1})\| \lesssim (\Delta t)^{p+1}, \forall n = 0, \ldots, N - 1.$$

Alors, l'erreur globale $e_n = u(t_n) - U_n$ satisfait

$$\|e_n\| \lesssim (\Delta t)^p, \ \forall n = 0, 1, \ldots, N.$$

Preuve : D'après (1.10), on a

$$u(t_{n+1}) = u(t_n) + \Delta t \phi(t_n, u(t_n), \Delta t) + E_l(t_{n+1}), \ n = 0, \ldots, N - 1.$$

Alors, par (1.9), l'hypothèse de lipschitzianité sur ϕ et le fait que $u \in C([0, T']; W)$, on obtient

$$\begin{aligned} \|e_{n+1}\| &\leq \|e_n\| + \Delta t\|\phi(t_n, u(t_n), \Delta t) - \phi(t_n, U_n, \Delta t)\| + \|E_l(t_{n+1})\| \\ &\lesssim (1 + L\Delta t)\|e_n\| + (\Delta t)^{p+1}. \end{aligned} \tag{1.48}$$

On termine comme dans la preuve du Théorème 1.1.4. ∎

Proposition 1.2.7. *Sous les hypothèses du Théorème 1.2.4, on suppose que la solution exacte u du problème (1.1) a la régularité $C^2([0,T];V)$. Si U_n est la solution approchée donnée par le schéma d'Euler explicite (1.45), alors l'erreur $e_n = u(t_n) - U_n$ satisfait*

$$\|u(t_n) - U_n\| \lesssim \Delta t, \quad \forall\, 0 \leq n \leq N.$$

Preuve : Notons d'abord que le Lemme 1.2.5 assure que $\phi(t,u,\Delta t) = F(t,u)$ satisfait (1.47).
D'autre part, et comme dans la preuve de la Proposition 1.1.5, on obtient par un développement de Taylor avec reste intégral à l'ordre 1

$$E_l(t_{n+1}) = \int_{t_n}^{t_{n+1}} (t_{n+1} - \tau) u''(\tau) d\tau.$$

Par conséquent, on en déduit de notre hypothèse sur u que

$$\|E_l(t_{n+1})\| \lesssim (\Delta t)^2.$$

On conclut grâce au Théorème 1.2.6. ∎

Pour voir sous quelles hypothèses on peut assurer la régularité $C^2([0,T];V)$ pour la solution u, on montre d'abord le lemme suivant.

Lemme 1.2.8. *Si la fonction $A_1 : (t,v) \longmapsto A_1(t,v)$ est différentiable sur $[0,T] \times V$, alors la fonction $B : (t,v) \longmapsto A_1(t,v)^{-1}$ l'est aussi. Plus précisément, si $DA_1(t_0,v_0) \in \mathcal{L}(\mathbb{R} \times V, \mathcal{L}(V,V'))$ est la différentielle au sens de Fréchet de la fonction A_1 au point $(t_0,v_0) \in [0,T] \times V$, alors $DB(t_0,v_0) \in \mathcal{L}(\mathbb{R} \times V, \mathcal{L}(V',V))$, la différentielle au sens de Fréchet de la fonction B au point (t_0,v_0), est donnée par*

$$DB(t_0,v_0)(t,v) = -A_1(t_0,v_0)^{-1} DA_1(t_0,v_0)(t,v) A_1(t_0,v_0)^{-1}.$$

Preuve : On a

$$\lim_{\|(t,v)\|\to 0} \frac{1}{\|(t,v)\|} \left\| A_1(t+t_0, v+v_0)^{-1} - A_1(t_0,v_0)^{-1} + A_1(t_0,v_0)^{-1} DA_1(t_0,v_0)(t,v) A_1(t_0,v_0)^{-1} \right\|$$

$$= \lim_{\|(t,v)\|\to 0} \frac{1}{\|(t,v)\|} \Big\{ \Big\| A_1(t_0,v_0)^{-1} \big(A_1(t_0,v_0) - A_1(t+t_0, v+v_0) + DA_1(t_0,v_0)(t,v) A_1(t_0,v_0)^{-1} A_1(t+t_0,v+v_0) \big) A_1(t+t_0,v+v_0)^{-1} \Big\| \Big\}$$

$$\leq \lim_{\|(t,v)\|\to 0} \|A_1(t_0,v_0)^{-1}\| \frac{\|A_1(t+t_0,v+v_0) - A_1(t_0,v_0) - DA_1(t_0,v_0)(t,v)\|}{\|(t,v)\|} \|A_1(t+t_0,v+v_0)^{-1}\|$$

$$+ \lim_{\|(t,v)\|\to 0} \frac{\|DA_1(t_0,v_0)(t,v)\|}{\|(t,v)\|} (A_1(t_0,v_0)^{-1} - A_1(t+t_0,v+v_0)^{-1}).$$

Cette dernière limite est nulle par définition de $DA_1(t_0, v_0)$ et d'après la continuité de la fonction $(t, v) \mapsto A_1(t, v)$ (voir Lemme 1.2.1).
On vérifie de même que si $\frac{\partial A_1}{\partial t}(t_0, v_0) \in \mathcal{L}(\mathbb{R}, \mathcal{L}(V, V'))$ (resp. $\frac{\partial A_1}{\partial v}(t_0, v_0) \in \mathcal{L}(V, \mathcal{L}(V, V'))$) est la différentielle partielle de A_1 en (t_0, v_0) par rapport à t (resp. par rapport à v), alors

$$\frac{\partial B}{\partial t}(t_0, v_0)t = -A_1(t_0, v_0)^{-1}\frac{\partial A_1}{\partial t}(t_0, v_0)t\, A_1((t_0, v_0)^{-1}.$$
$$\text{resp. } \frac{\partial B}{\partial v}(t_0, v_0)v = -A_1(t_0, v_0)^{-1}\frac{\partial A_1}{\partial v}(t_0, v_0)v\, A_1(t_0, v_0)^{-1}.$$

■

Remarque 1.2.9. *Dans le cas semi-linéaire, l'hypothèse $f \in C^1([0, T'] \times V; V')$ suffit pour assurer la régularité $u \in C^2([0, T']; V)$ (voir Lemme 1.1.2). Mais dans le cas quasi-linéaire, vu que*

$$u_t = F(t, u) = g(t, u) - \mathcal{A}(t, u)u = A_1(t, u)^{-1}(f(t, u) - A_2(t, u)u),$$

et compte tenu du lemme précédent, on a besoin de supposer de plus que la fonction $(t, v) \to A_i(t, v), i = 1, 2$, est différentiable (au sens de Fréchet) sur $[0, T'] \times V$, et que

$$\|\tfrac{\partial A_i}{\partial t}(t, v)\|_{\mathcal{L}(\mathbb{R};\mathcal{L}(V;V'))} + \|\tfrac{\partial A_i}{\partial v}(t, v)\|_{\mathcal{L}(V;\mathcal{L}(V;V'))} \lesssim 1,\ \forall\, t \in [0, T], v \in V.$$

Proposition 1.2.10. *Sous les hypothèses du Théorème 1.2.4, on suppose que la solution exacte u du problème (1.1) a la régularité $C^3([0, T']; V)$. Si U_n est la solution approchée donnée par le schéma de Runge-Kutta (1.46), alors on a*

$$\|u(t_n) - U_n\| \lesssim (\Delta t)^2,\ \ \forall 0 \le n \le N.$$

Preuve : Grâce au Lemme 1.2.5, on vérifie facilement que

$$\|\phi(t, u, \Delta t) - \phi(t, v, \Delta t)\| \lesssim (1 + \Delta t)\|u - v\|,\ \forall t \in [0, T], u \in W, v \in V,$$

ce qui implique que $\phi(t, u, \Delta t) = \frac{1}{2}(F(t, u) + F(t + \Delta t, u + \Delta t F(t, u)))$ satisfait (1.47).
D'autre part, et comme dans la preuve de la Proposition 1.1.6, on obtient par un développement de Taylor avec reste intégral à l'ordre 2

$$\|E_l(t_{n+1})\| \lesssim (\Delta t)^3.$$

Le résultat découle du Théorème 1.2.6. ■

Remarque 1.2.11. *Dans le cas semi-linéaire, l'hypothèse $f \in C^2([0, T'] \times V; V')$ implique la régularité $u \in C^3([0, T']; V)$ d'après le Lemme 1.1.2. Dans le cas quasi-linéaire, on a besoin de supposer de plus que la fonction $(t, v) \to A_i(t, v), i = 1, 2$, est deux fois différentiable (au sens de Fréchet) sur $[0, T'] \times V$ avec*

$$\begin{aligned}\|\frac{\partial^2 A_i}{\partial t^2}(t, v)\|_{\mathcal{L}(\mathbb{R},\mathbb{R};\mathcal{L}(V;V'))} &+ \|\frac{\partial^2 A_i}{\partial v^2}(t, v)\|_{\mathcal{L}(V,V;\mathcal{L}(V;V'))} \\ &+ \|\frac{\partial^2 A_i}{\partial t \partial v}(t, v)\|_{\mathcal{L}(\mathbb{R},V;\mathcal{L}(V;V'))} \lesssim 1,\ \forall\, t \in [0, T], v \in V.\end{aligned}$$

1.2.3 Schéma totalement discrétisé

On suppose toujours qu'il existe un sous espace de dimension finie V_h de V, $h > 0$, et on considère l'approximation discrète $A_{i,h}(t, u_h)$ de $A_i(t, u_h), i = 1, 2$, définie par

$$\langle A_{i,h}(t, u_h)v_h, w_h\rangle_{V_h', V_h} = a_i(t; u_h; v_h, w_h), \quad \forall v_h, w_h \in V_h, \forall i = 1, 2, \tag{1.49}$$

où $a_i(t; u_h; \cdot, \cdot), i = 1, 2$ est la forme bilinéaire définie par (1.2).
Pour tout $t \in [0, T]$ et $u \in V$, on définit l'opérateur de projection $P_h(t, u)$ associé à la forme bilinéaire $a_1(t; u; \cdot, \cdot)$, i.e., pour chaque $v \in V, P_h(t, u)v \in V_h$ est l'unique solution de

$$a_1(t; u; P_h(t, u)v, w_h) = a_1(t; u; v, w_h), \ \forall w_h \in V_h.$$

Si on note par

$$\|\cdot\|_{t,v} = \sqrt{a_1(t; v; \cdot, \cdot)},$$

la norme associée à la forme bilinéaire $a_1(t; u; \cdot, \cdot)$, on trouve grâce à la continuité et la coercivité uniformes de $a_1(t; v; \cdot, \cdot)$ (hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$), que

$$\sqrt{\alpha}\|u\| \leq \|u\|_{t,v} \leq \sqrt{M_1}\|u\|, \quad \forall t \in [0, T], \forall u, v \in V. \tag{1.50}$$

On considère maintenant le problème semi-discret (en espace) du problème (1.51)

$$\begin{cases} A_{1,h}(t, u_h)u_{h,t} + A_{2,h}(t, u_h)u_h = f_h(t, u_h), & \text{dans } V_h, \quad 0 < t \leq T, \\ u_h(0) = P_h(t_0, u_0)u_0, & \text{dans } V_h, \end{cases} \tag{1.51}$$

où $f_h(t, u_h) = I_h(f(t, u_h))$ et I_h est l'opérateur défini dans le cas semi-linéaire.
Comme dans le cas continu, l'opérateur $A_{1,h}(t, u_h), t \in [0, T], u_h \in V_h$ étant inversible, ce problème est donc équivalent à

$$\begin{cases} u_{h,t} + \mathcal{A}_h(t, u_h)u_h = g_h(t, u_h), & \text{dans } V_h, \quad 0 < t \leq T, \\ u_h(0) = P_h(t_0, u_0)u_0, & \text{dans } V_h, \end{cases} \tag{1.52}$$

où $\mathcal{A}_h(t, u_h) = A_{1,h}(t, u_h)^{-1}A_{2,h}(t, u_h)$ et $g_h(t, u_h) = A_{1,h}(t, u_h)^{-1}f_h(t, u_h)$. Le Lemme suivant montre que l'opérateur $\mathcal{A}_h(t, u_h)$ est borné (uniformément par rapport à h) de V_h dans lui même.

Lemme 1.2.12. *Sous les hypothèses* $\mathbf{H_1}$ *et* $\mathbf{H_2}$, *on a pour chaque* $t \in [0, T]$ *et* $v_h \in V_h$

$$\|\mathcal{A}_h(t, v_h)\|_{\mathcal{L}(V_h)} \leq \frac{M_2}{\alpha}. \tag{1.53}$$

Preuve : Soient $t \in [0, T]$ et $v_h \in V_h$. On a

$$\|\mathcal{A}_h(t, v_h)\|_{\mathcal{L}(V_h)} = \sup_{w_h \in V_h, w_h \neq 0} \frac{\|\mathcal{A}_h(t, v_h)w_h\|}{\|w_h\|}.$$

Pour un $w_h \in V_h, w_h \neq 0$ fixé, on pose $v'_h = A_{1,h}(t, v_h)^{-1} A_{2,h}(t, v_h) w_h$, alors on a par (1.49)

$$a_1(t; v_h; v'_h, \varphi_h) = a_2(t; v_h; w_h, \varphi_h), \forall \varphi_h \in V_h.$$

En posant $\varphi_h = v'_h$ dans cette égalité et en utilisant les hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$, on trouve

$$\alpha \|v'_h\|^2 \leq a_2(t; v_h; w_h, v'_h) \leq M_2 \|w_h\| \|v'_h\|,$$

d'où (1.53). ∎

Notons que le problème (1.52) peut être écrit de manière équivalente comme une EDO dans V_h :

$$\begin{cases} u_{h,t} = F_h(t, u_h), & \text{dans } V_h, \quad 0 < t \leq T, \\ u_h(0) = P_h(t_0, u_0) u_0, & \text{dans } V_h, \end{cases} \tag{1.54}$$

où $F_h(t, u_h) = g_h(t, u_h) - \mathcal{A}_h(t, u_h) u_h$. Par conséquent, son approximation par un schéma d'Euler explicite ou par un schéma de Runge-Kutta d'ordre 2, prend respectivement les formes

$$\begin{cases} U_{0,h} = P_h(t_0, u_0) u_0 \\ U_{n+1,h} = U_{n,h} + \Delta t F_h(t_n, U_{n,h}), \ n = 0, \ldots, N-1, \end{cases} \tag{1.55}$$

ou

$$\begin{cases} U_{0,h} = P_h(t_0, u_0) u_0 \\ U^*_{n+1,h} = U_{n,h} + \Delta t F_h(t_n, U_{n,h}), \\ U_{n+1,h} = U_{n,h} + \frac{\Delta t}{2} [F_h(t_n, U_{n,h}) + F_h(t_{n+1}, U^*_{n+1,h})], \ n = 0, \ldots, N-1. \end{cases} \tag{1.56}$$

Remarque 1.2.13. *Pour chaque $u_h \in V_h$, il est facile de vérifier, sous les hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$, que*

$$g_h(t, u_h) = P_h(t, u_h) g(t, u_h), \ \forall t \in [0, T]. \tag{1.57}$$

En effet, comme $A_{1,h}(t, u_h) g_h(t, u_h) = I_h f(t, u_h)$, on obtient

$$a_1(t; u_h; g_h(t, u_h), v_h) = \langle f(t, u_h), v_h \rangle_{V',V}, \quad \forall v_h \in V_h.$$

D'une part, la définition de $P_h(t, u_h)$ implique que

$$a_1(t; u_h; P_h(t, u_h) g(t, u_h), v_h) = a_1(t; u_h; g(t, u_h), v_h) = a_1(t; u_h; A_1^{-1}(t, u_h) f(t, u_h), v_h).$$

D'autre part, par la définition de la forme bilinéaire $a_1(t; u_h; \cdot, \cdot)$, on trouve que

$$a_1(t; u_h; g_h(t, u_h), v_h) = a_1(t; u_h; P_h(t, u_h) g(t, u_h), v_h), \quad \forall v_h \in V_h,$$

d'où (1.57) par le lemme de Lax-Milgram.

Proposition 1.2.14. *Sous les hypothèses* $\mathbf{H_1}$-$\mathbf{H_8}$, *on suppose que* $f \in C([0,T]\times V;\tilde{D}_{s-1})$. *Alors pour tout* $t\in[0,T]$ *et* $(v,v_h)\in D_s\times V_h$, *on a*

$$\|F_h(t,v_h)-F(t,v)\| \lesssim (1+\|v\|)\|v-v_h\| + h^{q(s)}(\|v\|_{D_s}+\|f(t,v)\|_{\tilde{D}_{s-1}}). \tag{1.58}$$

Preuve : Fixons $t\in[0,T]$ et $(v,v_h)\in D_s\times V_h$. Par l'inégalité triangulaire on a

$$\|F_h(t,v_h)-F(t,v)\| \le \|g_h(t,v_h)-g(t,v)\| + \|\mathcal{A}(t,v)v-\mathcal{A}_h(t,v_h)v_h\|.$$

On estime d'abord le premier terme du membre de droite de cette inégalité. L'identité (1.57) donne

$$\begin{aligned}\|g_h(t,v_h)-g(t,v)\| &\le \|g_h(t,v_h)-P_h(t,v_h)g(t,v)\| + \|(P_h(t,v_h)-I)g(t,v)\|\\ &= \|P_h(t,v_h)(g(t,v_h)-g(t,v))\| + \|(P_h(t,v_h)-I)g(t,v)\|.\end{aligned}$$

D'une part, en utilisant (1.50), on obtient

$$\begin{aligned}\|P_h(t,v_h)(g(t,v_h)-g(t,v))\| &\le \frac{1}{\sqrt{\alpha}}\|P_h(t,v_h)(g(t,v_h)-g(t,v))\|_{t,v_h}\\ &\le \frac{1}{\sqrt{\alpha}}\|g(t,v_h)-g(t,v)\|_{t,v_h}\\ &\le \sqrt{\frac{M_1}{\alpha}}\|g(t,v_h)-g(t,v)\|\\ &\lesssim \sqrt{\frac{M_1}{\alpha}}\|v-v_h\|,\end{aligned}$$

où la dernière inégalité est due au Corollaire 1.2.3, 2.
D'autre part, on obtient en utilisant toujours la double inégalité (1.50)

$$\begin{aligned}\|(P_h(t,v_h)-I)g(t,v)\| &\le \frac{1}{\sqrt{\alpha}}\|(P_h(t,v_h)-I)g(t,v)\|_{t,v_h}\\ &\le \frac{1}{\sqrt{\alpha}}\|(Q_h-I)g(t,v)\|_{t,v_h}\\ &\le \sqrt{\frac{M_1}{\alpha}}\|(Q_h-I)g(t,v)\|.\end{aligned}$$

Grâce à l'hypothèse (1.25), on trouve

$$\|(P_h(t,v_h)-I)g(t,v)\| \lesssim h^{q(s)}\|g(t,v)\|_{D_s} \lesssim h^{q(s)}\|f(t,v)\|_{\tilde{D}_{s-1}}.$$

Pour le deuxième terme, on a

$$\begin{aligned}\|\mathcal{A}(t,v)v-\mathcal{A}_h(t,v_h)v_h\| &\le \|\mathcal{A}(t,v)v-\mathcal{A}(t,v_h)v\| + \|\mathcal{A}(t,v_h)v-P_h(t,v_h)\mathcal{A}(t,v_h)v\|\\ &\quad +\|P_h(t,v_h)\mathcal{A}(t,v_h)v-\mathcal{A}_h(t,v_h)v_h\|.\end{aligned}$$

Or, d'après le Corollaire 1.2.2

$$\|\mathcal{A}(t,v)v - \mathcal{A}(t,v_h)v\| \lesssim \|v - v_h\|\|v\|.$$

On a aussi grâce à (1.50)

$$\|(I - P_h(t,v_h))\mathcal{A}(t,v_h)v\| \leq \sqrt{\frac{M_1}{\alpha}}\|(I - Q_h)\mathcal{A}(t,v_h)v\|,$$

comme $\mathcal{A}(t,v_h)v \in D_s$ du fait que $v \in D_s$, on en déduit de l'hypothèse (1.25) que

$$\|(I - P_h(t,v_h))\mathcal{A}(t,v_h)v\| \lesssim h^{q(s)}\|\mathcal{A}(t,v_h)v\|_{D_s} \lesssim h^{q(s)}\|v\|_{D_s}.$$

Reste à estimer le dernier terme. Pour cela, on pose $w = \mathcal{A}(t,v_h)v$ et $w_h = \mathcal{A}_h(t,v_h)v_h$, donc par définition de a_1 et a_2, on a

$$a_1(t;v_h;w,\psi) = a_2(t;v_h;v,\psi),\ \forall\psi \in V,$$

et

$$a_1(t;v_h;w_h,\psi_h) = a_2(t;v_h;v_h,\psi_h),\ \forall\psi_h \in V_h,$$

et par définition de $P_h(t,v_h)$, on a

$$a_1(t;v_h;P_h(t,v_h)w,\psi_h) = a_1(t;v_h;w,\psi_h),\ \forall\psi_h \in V_h.$$

Par conséquent

$$a_1(t;v_h;P_h(t,v_h)w - w_h,\psi_h) = a_2(t;v_h;v - v_h,\psi_h),\ \forall\psi_h \in V_h. \tag{1.59}$$

En choisissant dans (1.59), $\psi_h = P_h(t,v_h)w - w_h$, on obtient grâce aux hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$

$$\|P_h(t,v_h)w - w_h\| \lesssim \|v - v_h\|.$$

Alors

$$\|P_h(t,v_h)\mathcal{A}(t,v_v)v - \mathcal{A}_h(t,v_h)v_h\| \lesssim h^{q(s)}\|v\|_{D_s} + \|v - v_h\|(1 + \|v\|_{D_s}).$$

Ce qui prouve que (1.58) a lieu. ∎

Remarque 1.2.15. *On vérifie facilement que le Lemme 1.1.11 reste vrai (sous les mêmes hypothèses) dans le cas quasi-linéaire. Il suffit d'utiliser le fait que les opérateurs $A_1(t,u)^{-1} : \tilde{D}_{s-1} \to D_s$ et $A_2(t,u) : D_s \to \tilde{D}_{s-1}$ sont bornés uniformément par rapport à t et u (hypothèse $\mathbf{H_3}$).*

Lemme 1.2.16. *Supposons que les hypothèses du Lemme 1.1.12 sont satisfaites en plus des hypothèses* $\mathbf{H_6}$-$\mathbf{H_8}$. *Soit* U_n *(resp.* $U_{n,h}$*) la solution approchée donnée par (1.16) (resp. (1.56)). Alors l'erreur* $e_{n,h} = U_n - U_{n,h}$ *est majorée comme suit*

$$\|e_{n,h}\| \lesssim h^{q(s)}. \tag{1.60}$$

Preuve : Par l'inégalité triangulaire, on a

$$\begin{aligned}\|e_{n+1,h}\| \leq \|e_{n,h}\| \quad + \quad & \frac{\Delta t}{2}\|F_h(t_n, U_{n,h}) - F(t_n, U_n)\| \\ & + \frac{\Delta t}{2}\|F_h\big(t_{n+1}, U_{n,h} + \Delta t F_h(t_n, U_{n,h})\big) - F\big(t_{n+1}, U_n + \Delta t F(t_n, U_n)\big)\|.\end{aligned}$$

La suite de la preuve est comme dans le cas semi-linéaire (Lemme 1.1.12), en utilisant la Proposition 1.2.14 au lieu de la Proposition 1.1.10 pour majorer les deux derniers termes du membre de droite de cette inégalité. ∎

Remarque 1.2.17. *Sous les mêmes hypothèses du Lemme 1.2.16, si le schéma d'Euler (1.55) est utilisé pour approcher la solution du problème (1.1), alors l'estimation d'erreur (1.60) reste valide.*

Corollaire 1.2.18. *Sous les Hypothèses* $\mathbf{H_1}$-$\mathbf{H_8}$, *on suppose que* $u_0 \in W \cap D_s$, $f \in C([0,T]\times V; \tilde{D}_{s-1})$, *que* f *est* $(V, \tilde{D}_{s-1})$*-continûment lipschitzienne par rapport à la deuxième variable uniformément en* t, *et que la solution exacte* u *du problème (1.1) appartient à* $C^{p+1}([0,T']; V)$, $p = 1$ *ou* 2. *Soit* $U_{n,h}$ *sa solution approchée donnée par (1.55) pour* $p = 1$ *(resp. (1.56) pour* $p = 2$*). Alors on a l'estimation d'erreur globale*

$$\|u(t_n) - U_{n,h}\| \lesssim (\Delta t)^p + h^{q(s)}, \quad \forall 1 \leq n \leq N.$$

Preuve : Conséquence directe :

- de la Proposition 1.2.7 et de la Remarque 1.2.17 pour le schéma d'Euler,
- de la Proposition 1.2.10 et du Lemme 1.2.16 pour le schéma de Runge-Kutta.

∎

Chapitre 2

Analyse a priori pour un schéma implicite en temps d'un problème singulièrement perturbé

Résumé

Nous considérons un problème de Sobolev singulièrement perturbé dans un espace de Hilbert V. Pour approcher ce problème, nous proposons deux schémas numériques implicites et nous prouvons des estimations d'erreur.

Sommaire

2.1 Le problème homogène **38**
2.1.1 Etude du problème semi-discrétisé implicite 38
2.1.2 Etude du problème semi-discrétisé en espace 41
2.2 Le problème non homogène **56**
2.2.1 Le problème semi-discrétisé en temps 56
2.2.2 Le problème totalement discrétisé 59

On considère deux espaces de Hilbert H et V muni de deux produits scalaires différents $(\cdot,\cdot)_H$, $(\cdot,\cdot)_V$ et on suppose que $V \hookrightarrow H$ avec injection continue et dense. L'identification de H avec son dual réalise une injection de cet espace dans V', le dual de V, ce qui donne le triple de Gelfand $V \hookrightarrow H \hookrightarrow V'$. On notera par $\langle\cdot,\cdot\rangle_{V',V}$ le produit de dualité entre V' et V et par $\|\cdot\|$ et $\|\cdot\|_H$ les normes de V et H respectivement. Ce chapitre est consacré à l'étude du problème de Sobolev suivant

$$\begin{cases} u_{\epsilon,t} + \epsilon A_1 u_{\epsilon,t} + A_2 u_\epsilon = f(t), & \text{dans } V', \quad 0 < t \leq T, \\ u_\epsilon(0) = u_0, & \text{dans } V, \end{cases} \tag{2.1}$$

où ϵ est une constante positive et pour $i = 1, 2$, l'opérateur A_i est un isomorphisme de V sur V'.

Nous nous intéressons dans ce chapitre à trouver une approximation totalement discrète du problème (2.1).

On associe à chacun des opérateurs $A_i : V \to V'$, $i = 1, 2$, une forme bilinéaire symétrique a_i par la relation

$$a_i(u, v) = \langle A_i u, v\rangle_{V',V}, \forall u, v \in V. \tag{2.2}$$

On suppose que les formes a_i sont continues et coercives. On notera $M_i > 0$ la constante de continuité et $\alpha_i > 0$ la constante de coercivité de a_i, $i = 1, 2$. On a alors

$$|a_i(u, v)| \leq M_i \|u\| \|v\|, \quad \text{et} \quad a_i(v, v) \geq \alpha_i \|v\|^2 \quad \forall u, v \in V.$$

Ceci assure grâce au lemme de Lax-Milgram que les opérateurs $(I + \epsilon A_1)$, $\epsilon > 0$, et A_2 sont des isomorphismes de V sur V'. Pour tout $i = 1, 2$, on introduit la restriction de A_i sur H (on la notera aussi A_i) qu'on peut considérer comme un opérateur non borné dans H défini par

$$D(A_i) = \{u \in V;\ \exists\, g_u \in H \text{ tel que } a_i(u, v) = (g_u, v)_H \text{ pour tout } v \in V\}$$

et

$$A_i u = g_u,\ \forall u \in D(A_i).$$

On munira $D(A_i)$ de la norme du graphe :

$$\|u\|_{D(A_i)} = \big(\|u\|_H^2 + \|A_i u\|_H^2\big)^{1/2},$$

qui fait de $D(A_i)$ un espace de Hilbert. L'opérateur non borné A_i, $i = 1, 2$, ainsi défini est auto-adjoint défini positif et est un isomorphisme de $D(A_i)$ sur H. Par conséquent, A_i est inversible et il existe une constate $\beta_i > 0$ telle que

$$\|A_i u\|_H \geq \beta_i \|u\|_H, \quad \text{pour tout } u \in D(A_i).$$

On suppose de plus que l'injection de $D(A_1)$ dans V est compacte. Comme conséquence de cette hypothèse, A_1 est à inverse compact et donc a un spectre discret. Soient $\{\nu_l\}_{l=1}^{+\infty}$ ses valeurs propres et $\{\psi_l\}_{l=1}^{+\infty}$ les vecteurs propres associés.

On a aussi besoin d'introduire $D(A_i^k)$, $k \geq 2$, domaines des puissances entières de A_i, $i = 1, 2$,

$$D(A_i^k) = \{v \in D(A_i^{k-1}); A_i v \in D(A_i^{k-1})\}.$$

$D(A_i^k)$ est un espace de Hilbert pour le produit scalaire

$$(u, v)_{D(A_i^k)} = \sum_{j=0}^{k} (A_i^j u, A_i^j v)_H,$$

et la norme correspondante

$$\|u\|_{D(A_i^k)} = \left(\sum_{j=0}^{k} \|A_i^j u\|_H^2 \right)^{\frac{1}{2}}.$$

Comme l'opérateur $(I + \epsilon A_1)$ est inversible, on peut donc composer les deux membres de la première équation de (2.1) par $(I + \epsilon A_1)^{-1}$ et obtenir le problème équivalent

$$\begin{cases} u_{\epsilon,t} + \mathcal{A}_\epsilon u_\epsilon = f_\epsilon(t), & \text{dans } V, \quad 0 < t \leq T, \\ u_\epsilon(0) = u_0, & \text{dans } V, \end{cases} \tag{2.3}$$

où $\mathcal{A}_\epsilon = (I + \epsilon A_1)^{-1} A_2$ et $f_\epsilon = (I + \epsilon A_1)^{-1} f$. L'opérateur $\mathcal{A}_\epsilon$ est borné de V dans lui même et on a

$$\|\mathcal{A}_\epsilon\|_{\mathcal{L}(V)} \leq \frac{M_2}{\alpha_1 \epsilon}. \tag{2.4}$$

En effet, on a par définition

$$\|\mathcal{A}_\epsilon\|_{\mathcal{L}(V)} = \sup_{u \in V, u \neq 0} \frac{\|\mathcal{A}_\epsilon u\|}{\|u\|},$$

alors pour $u \in V, u \neq 0$, si on pose $\varphi = \mathcal{A}_\epsilon u = (I + \epsilon A_1)^{-1} A_2 u$, on obtient

$$\|\varphi\|_H^2 + \epsilon a_1(\varphi, \varphi) = a_2(u, \varphi).$$

De la coercivité de a_1 et de la continuité de a_2, on trouve

$$\alpha_1 \epsilon \|\varphi\|^2 \leq M_2 \|u\| \, \|\varphi\|,$$

d'où (2.4).

On peut donc considérer notre problème comme une équation différentielle ordinaire semi-linéaire à valeurs dans l'espace de Hilbert V, i.e. le problème (2.3) (et alors (2.1)) est équivalent à

$$\begin{cases} u_{\epsilon,t} = F_\epsilon(t, u_\epsilon), & \text{dans } V, \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V, \end{cases} \tag{2.5}$$

où $F_\epsilon(t, u_\epsilon(t)) = f_\epsilon(t) - \mathcal{A}_\epsilon u_\epsilon(t)$.

Si on suppose que $u_0 \in V$ et $f \in C^0([0, T], V')$, alors par le Théorème de Cauchy-Lipschitz-Picard le problème (2.5) (et alors (2.1)) admet une unique solution $u_\epsilon \in C^1([0, T], V)$ (voir Théorème 7. 3 de [15]).

Nous pouvons utiliser, comme dans le chapitre précédent, les schémas d'approximation standards pour la discrétisation temporelle de notre problème. Mais de (2.4), on remarque que $\|\mathcal{A}_\epsilon\|_{\mathcal{L}(V)}$ tend vers $+\infty$ lorsque ϵ tend vers zéro, et donc l'utilisation d'un schéma numérique explicite n'est pas appropriée dans ce cas, vu que la constante de stabilité

dépend de $\|\mathcal{A}_\epsilon\|_{\mathcal{L}(V)}$. Pour cette raison, nous allons utiliser dans ce chapitre des schémas numériques implicites qui sont inconditionnellement stables.

Par la formule de variation des constantes, la solution exacte du problème (2.5) est donnée par

$$u_\epsilon(t) = u_0 + \int_0^t F_\epsilon(t, u_\epsilon(\tau))d\tau. \tag{2.6}$$

Pour étudier une approximation de u_ϵ, on décompose le problème (2.1) en un problème homogène avec une condition initiale égale u_0 et un problème non homogène avec condition initiale nulle.
On commence d'abord par étudier le problème homogène. Tout au long de ce chapitre, la solution exacte u_ϵ du problème (2.3) sera notée tout simplement u.

2.1 Le problème homogène

2.1.1 Etude du problème semi-discrétisé implicite

On considère le problème (2.3) avec $f_\epsilon = 0$, i.e.

$$\begin{cases} u_t + \mathcal{A}_\epsilon u = 0, & \text{dans } V, \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V. \end{cases} \tag{2.7}$$

Soit $(t_i = i\Delta t)_{i=0}^N$ une subdivision régulière de l'intervalle $[0, T]$, où $N \in \mathbb{N}^*$ et $\Delta t = \frac{T}{N}$ est le pas de discrétisation en temps. On cherche à approcher la solution exacte u de (2.7) aux points $(t_{n+1}), n = 0, ..., N-1$, en partant de $u(t_0 = 0) = u_0$ et en estimant étape par étape les valeurs de $u(t_{n+1})$ grâce à la formule

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} F_\epsilon(\tau, u(\tau))d\tau.$$

Par le schéma d'Euler implicite (qui consiste à approcher l'intégrale $\int_{t_n}^{t_{n+1}} F_\epsilon(\tau, u(\tau))d\tau$ par $\Delta t F_\epsilon(t_{n+1}, u(t_{n+1}))$), la solution semi-discrète du problème (2.7) est donnée par

$$u_{n+1} = u_n + \Delta t F_\epsilon(t_{n+1}, u_{n+1}), \quad n = 0, ..., N-1,$$

ou de manière équivalente

$$(I + \Delta t \mathcal{A}_\epsilon)u_{n+1} = u_n, \quad n = 0, ..., N-1. \tag{2.8}$$

Et, par le schéma de Crank-Nicolson (qui consiste à approcher l'intégrale $\int_{t_n}^{t_{n+1}} F_\epsilon(\tau, u(\tau))d\tau$ par $\frac{\Delta t}{2}(F_\epsilon(t_n, u(t_n)) + F(t_{n+1}, u(t_{n+1})))$), cette solution est donnée par

$$u_{n+1} = u_n + \frac{\Delta t}{2}(F_\epsilon(t_n, u_n) + F_\epsilon(t_{n+1}, u_{n+1})), \quad n = 0, ..., N-1,$$

ce qui équivaut à

$$(I + \frac{1}{2}\Delta t\mathcal{A}_\epsilon)u_{n+1} = (I - \frac{1}{2}\Delta t\mathcal{A}_\epsilon)u_n, \quad n = 0, ..., N-1. \tag{2.9}$$

Dans ce qui suit, on aura besoin du résultat technique suivant

Lemme 2.1.1. *Pour tout $y \in H$, on a*

$$\|(I + \epsilon A_1)^{-\frac{k}{2}} y\|_H \leq \|y\|_H, \quad k = 1, 2.$$

Preuve : Comme l'opérateur A_1 est auto-adjoint à résolvante compacte, ses vecteurs propres $\{\psi_l\}_{l=1}^{+\infty}$ forment une base hilbertienne de l'espace H. Pour tout $y \in H$, on a alors la décomposition suivante

$$y = \sum_{l=1}^{\infty} (y, \psi_l)_H \psi_l,$$

et donc

$$(I + \epsilon A_1)^{-\frac{k}{2}} y = \sum_{l=1}^{\infty} (1 + \epsilon\nu_l)^{-\frac{k}{2}} (y, \psi_l)_H \psi_l,$$

ce qui implique par la formule de Parseval que

$$\begin{aligned} \|(I + \epsilon A_1)^{-\frac{k}{2}} y\|_H^2 &= \sum_{l=1}^{\infty} (1 + \epsilon\nu_l)^{-k} |(y, \psi_l)_H|^2 \\ &\leq \sum_{l=1}^{\infty} |(y, \psi_l)_H|^2 = \|y\|_H^2, \end{aligned}$$

puisque $(1 + \epsilon\nu_l)^k \geq 1$, $\forall\, l \geq 1$. ∎

Pour des valeurs de $\epsilon > 0$, on montre des estimations d'erreur entre la solution exacte du problème (2.7) et la solution approchée du problème semi-discrétisé par la méthode d'Euler implicite (2.8) (ou par la méthode de Crank-Nicolson (2.9)) en utilisant d'une part, le Théorème 7.1 de [71] après un changement d'inconnue approprié et d'autre part le lemme précédent.

Lemme 2.1.2. *On suppose que $u_0 \in D(A_2)$. Soit u la solution exacte du problème (2.7) et soit u_n la solution approchée donnée par le schéma d'Euler implicite (2.8). Alors*

$$\|u(t_n) - u_n\|_H \lesssim \Delta t \|u_0\|_{D(A_2)}.$$

Preuve : L'opérateur $\mathcal{A}_\epsilon$ n'est pas auto-adjoint. Pour se ramener au cas auto-adjoint, on fait le changement d'inconnue suivant

$$\begin{cases} v = (I + \epsilon A_1)^{\frac{1}{2}} u, \\ v_0 = (I + \epsilon A_1)^{\frac{1}{2}} u_0, \end{cases} \tag{2.10}$$

le problème (2.1) est alors équivalent à

$$\begin{cases} v_t + \mathcal{B}_\epsilon v = 0, & \text{dans } H, \quad 0 < t \leq T, \ \epsilon > 0, \\ v(0) = v_0, & \text{dans } H, \end{cases} \tag{2.11}$$

avec $\mathcal{B}_\epsilon = (I+\epsilon A_1)^{-\frac{1}{2}} A_2 (I+\epsilon A_1)^{-\frac{1}{2}}$ qui est bien auto-adjoint défini positif. On considère le schéma d'Euler implicite, à partir de v_0, la solution approchée du problème (2.11) est donnée par

$$v_{n+1} = (I + \Delta t \mathcal{B}_\epsilon)^{-1} v_n, \quad 0 \leq n \leq N-1. \tag{2.12}$$

Comme l'opérateur $\mathcal{B}_\epsilon$ est auto-adjoint défini positif, alors par le Théorème 7.1 de [71, page 114] on a pour les solutions de (2.11) et (2.12)

$$\|v_n - v(t_n)\|_H \lesssim \Delta t \|\mathcal{B}_\epsilon v_0\|_H, \ \ \forall\, 1 \leq n \leq N.$$

Mais $\mathcal{B}_\epsilon v_0 = \mathcal{B}_\epsilon (I+\epsilon A_1)^{\frac{1}{2}} u_0 = (I+\epsilon A_1)^{-\frac{1}{2}} A_2 u_0$. Sachant que $A_2 u_0 \in H$, on conclut grâce au Lemme 2.1.1 que

$$\|\mathcal{B}_\epsilon v_0\|_H \lesssim \|A_2 u_0\|_H, \tag{2.13}$$

ce qui donne

$$\|v_n - v(t_n)\|_H \lesssim \Delta t \|u_0\|_{D(A_2)}.$$

Par récurrence sur n, $0 \leq n \leq N$, on peut vérifier facilement que

$$v_n = (I+\epsilon A_1)^{\frac{1}{2}} u_n. \tag{2.14}$$

En effet, cette identité est vérifiée pour $n = 0$ par définition de v_0. Si on suppose maintenant que (2.14) est vraie, on a grâce à (2.12)

$$(I + \Delta t \mathcal{B}_\epsilon) v_{n+1} = v_n = (I+\epsilon A_1)^{\frac{1}{2}} u_n,$$

en composant les deux membres de cette égalité par $(I+\epsilon A_1)^{-\frac{1}{2}}$, on trouve

$$(I + \Delta t \mathcal{A}_\epsilon)(I+\epsilon A_1)^{-\frac{1}{2}} v_{n+1} = u_n,$$

et grâce à (2.8), on obtient

$$v_{n+1} = (I+\epsilon A_1)^{\frac{1}{2}} u_{n+1}.$$

Par conséquent et grâce au Lemme 2.1.1, on obtient

$$\begin{aligned} \|u(t_n) - u_n\|_H &= \|(I+\epsilon A_1)^{-\frac{1}{2}} (v(t_n) - v_n)\|_H \\ &\lesssim \|v(t_n) - v_n\|_H, \end{aligned}$$

d'où le résultat. ∎

Pour montrer maintenant une estimation d'erreur entre la solution exacte et la solution semi-discrétisée par le schéma de Crank-Nicolson, on a besoin de supposer que

$\mathbf{H_1'}$ $D(A_1) = D(A_2)$ et $\|A_1 u\|_H \sim \|A_2 u\|_H$, pour tout $u \in D(A_1)$.

Lemme 2.1.3. *On suppose que l'hypothèse* $\mathbf{H_1'}$ *a lieu et* $u_0 \in D(A_2^2)$. *Soit* u *la solution exacte du problème* (2.1) *et soit* u_n *la solution approchée donnée par le schéma de Crank-Nicolson* (2.9). *Alors*

$$\|u(t_n) - u_n\|_H \lesssim (\Delta t)^2 \|u_0\|_{D(A_2^2)}, \quad \forall 1 \leq n \leq N.$$

Preuve : Toujours par le même changement d'inconnue (2.10) et en utilisant [71, Théorème 7.1, page 114], on obtient

$$\|v_n - v(t_n)\|_H \lesssim (\Delta t)^2 \|\mathcal{B}_\epsilon^2 v_0\|_H,$$

où v est la solution exacte du problème (2.11) et v_n est la solution approchée donnée par le schéma de Crank-Nicolson suivant

$$v_{n+1} = (I + \frac{1}{2}\Delta t \mathcal{B}_\epsilon)^{-1}(I - \frac{1}{2}\Delta t \mathcal{B}_\epsilon) v_n, \quad n = 0, ...N-1.$$

En appliquant le Lemme 2.1.1 et l'hypothèse $\mathbf{H_1'}$, on a pour tout $u_0 \in D(A_2^2)$

$$\begin{aligned} \|\mathcal{B}_\epsilon^2 v_0\|_H = \|\mathcal{B}_\epsilon^2 (I + \epsilon A_1)^{\frac{1}{2}} u_0\|_H &= \|(I + \epsilon A_1)^{-\frac{1}{2}} A_2 (I + \epsilon A_1)^{-1} A_2 u_0\|_H \\ &\lesssim \|A_1 (I + \epsilon A_1)^{-1} A_2 u_0\|_H. \end{aligned}$$

En tenant compte du fait que les opérateurs A_1 et $(I + \epsilon A_1)^{-1}$ commutent, on obtient en appliquant encore une fois le Lemme 2.1.1

$$\|\mathcal{B}_\epsilon^2 v_0\|_H \lesssim \|A_1 A_2 u_0\|_H,$$

le résultat se déduit donc de $\mathbf{H_1'}$. ∎

2.1.2 Etude du problème semi-discrétisé en espace

Pour un paramètre positif h, on suppose qu'il existe un sous-espace de dimension finie V_h de V, et on considère une approximation discrète $A_{i,h}$ de $A_i, i = 1, 2$ comme suit

$$(A_{i,h} u_h, v_h)_{V_h', V_h} = a_i(u_h, v_h), \quad \forall u_h, v_h \in V_h, \tag{2.15}$$

où $a_i(\cdot, \cdot)$ est la forme bilinéaire définie sur $V \times V$ par (2.2). On considère maintenant le problème discret (en espace) suivant

$$\begin{cases} (I + \epsilon A_{1,h}) u_{h,t} + A_{2,h} u_h = 0, & \text{dans } V_h', \quad 0 < t \leq T, \ \epsilon > 0, \\ u_h(0) = P_h u_0, & \text{dans } V_h, \end{cases} \tag{2.16}$$

où P_h est l'opérateur de projection orthogonale de V sur V_h associé au produit scalaire de H, i.e., pour tout $v \in V, P_h v$ est l'unique solution de

$$(P_h v, w_h)_H = (v, w_h)_H, \quad \forall w_h \in V_h.$$

La forme bilinéaire $a_1(\cdot,\cdot)$ étant continue et coercive, l'opérateur $(I+\epsilon A_{1,h})^{-1}$ est inversible de V_h dans V_h'. Donc on peut composer la première équation de (2.16) par $(I+\epsilon A_{1,h})^{-1}$ et obtenir le problème semi-discret équivalent

$$\begin{cases} u_{h,t} + \mathcal{A}_{\epsilon,h} u_h = 0, & \text{dans } V_h, \quad 0 < t \leq T, \ \epsilon > 0, \\ u_h(0) = P_h u_0, & \text{dans } V_h, \end{cases} \tag{2.17}$$

où $\mathcal{A}_{\epsilon,h} = (I + \epsilon A_{1,h})^{-1} A_{2,h}$.

Lemme 2.1.4. *On a*

$$\|\mathcal{A}_{\epsilon,h}\|_{\mathcal{L}(V_h)} \leq \frac{M_2}{\alpha_1 \epsilon}. \tag{2.18}$$

Preuve : Par définition, on a

$$\|\mathcal{A}_{\epsilon,h}\|_{\mathcal{L}(V_h)} = \sup_{w_h \in V_h, w_h \neq 0} \frac{\|(I + \epsilon A_{1,h})^{-1} A_{2,h} w_h\|}{\|w_h\|}.$$

Pour un $w_h \in V_h, w_h \neq 0$ fixé, on pose $u_h = (I + \epsilon A_{1,h})^{-1} A_{2,h} w_h$. Alors u_h appartient à V_h et vérifie

$$(u_h, v_h)_H + \epsilon a_1(u_h, v_h) = a_2(w_h, v_h), \quad \forall v_h \in V_h,$$

par définition de $a_i, i = 1, 2$. En posant $v_h = u_h$ dans cette identité et en tenant compte de la coercivité de la forme bilinéaire a_1 et de la continuité de la forme a_2, on obtient

$$\epsilon \alpha_1 \|u_h\|^2 \leq M_2 \|w_h\| \|u_h\|,$$

ce qui donne (2.18). ∎

Comme l'opérateur $\mathcal{A}_{\epsilon,h}$ est borné dans V_h, on peut alors considérer notre problème comme une EDO à valeurs dans V_h et par conséquent utiliser les schémas numériques standards pour discrétiser le problème (2.17) en temps. Mais puisque $\|\mathcal{A}_{\epsilon,h}\|_{\mathcal{L}(V_h)}$ tend vers $+\infty$ lorsque ϵ tend vers 0, alors comme dans la section précédente et en vue de la stabilité nous allons utiliser les deux schémas numériques implicites d'Euler et de Crank-Nicolson donnés par les récurrences suivantes respectivement

$$(I + \Delta t \mathcal{A}_{\epsilon,h}) u_{n+1,h} = u_{n,h}, \quad \forall\, 0 \leq n \leq N-1, \tag{2.19}$$

$$(I + \frac{1}{2}\Delta t \mathcal{A}_{\epsilon,h}) u_{n+1,h} = (I - \frac{1}{2}\mathcal{A}_{\epsilon,h}) u_{n,h}, \quad \forall\, 0 \leq n \leq N-1, \tag{2.20}$$

avec $u_{0,h} = P_h u_0$.

On cherche à estimer l'erreur entre la solution exacte du problème (2.1) et la solution approchée donnée par le schéma numérique totalement discret (2.19) (ou (2.20)). Pour cela on a besoin d'introduire les hypothèses suivantes.

$\mathbf{H'_2}$ l'inégalité inverse

$$\|\phi_h\| \lesssim h^{-1}\|\phi_h\|_H, \ \forall\, \phi_h \in V_h,$$

est vérifiée (c'est le cas pour des triangulations quasi-régulières).

$\mathbf{H'_3}$ Pour $i = 1$ ou 2,

$$\|A_{i,h}^{-1} P_h u - A_i^{-1} u\|_H \lesssim h^2 \|u\|_H, \ \forall\, u \in V.$$

$\mathbf{H'_4}$

$$\|f - P_h f\|_H \lesssim h^2 \|f\|_{D(A_1)}, \ \forall\, f \in D(A_1).$$

$\mathbf{H'_5}$

$$\|f - P_h f\| \lesssim h \|f\|_{D(A_1)}, \ \forall\, f \in D(A_1).$$

On commence par donner quelques résultats qui nous seront utiles pour montrer le résultat principal de cette section.

Lemme 2.1.5. *Sous l'hypothèse* $\mathbf{H'_1}$*, on a*

$$\|A_1^{-1} u\|_H \sim \|A_2^{-1} u\|_H, \ \ \forall u \in H.$$

Preuve : Par définition, on a pour tout $u \in H$

$$\|A_1^{-1} u\|_H = \sup_{v \in H, v \neq 0} \frac{|(A_1^{-1} u, v)_H|}{\|v\|_H}.$$

Comme les opérateurs A_1^{-1} et A_2^{-1} sont auto-adjoints, on a

$$\|A_1^{-1} u\|_H = \sup_{v \in H, v \neq 0} \frac{|(u, A_1^{-1} v)_H|}{\|v\|_H} = \sup_{v \in H, v \neq 0} \frac{|(u, A_2^{-1} A_2 A_1^{-1} v)_H|}{\|v\|_H} = \sup_{v \in H, v \neq 0} \frac{|(A_2^{-1} u, A_2 A_1^{-1} v)_H|}{\|v\|_H}.$$

D'après l'inégalité de Cauchy-Schwarz et $\mathbf{H'_1}$, on trouve

$$\|A_1^{-1} u\|_H \leq \|A_2^{-1} u\|_H \sup_{v \in H, v \neq 0} \frac{\|A_2 A_1^{-1} v\|_H}{\|v\|_H} \lesssim \|A_2^{-1} u\|_H.$$

En intervertissant A_1 et A_2, on obtient $\|A_2^{-1} u\|_H \lesssim \|A_1^{-1} u\|_H$, d'où l'équivalence cherchée.

■

Lemme 2.1.6. *Sous l'hypothèse* $\mathbf{H_2'}$*, on a pour* $i = 1$ *ou 2,*

$$\|A_{i,h}\phi_h\|_H \lesssim h^{-2}\|\phi_h\|_H, \forall \phi_h \in V_h, \tag{2.21}$$

par conséquent

$$\|A_{i,h}^{-1}\psi_h\|_H \gtrsim h^2\|\psi_h\|_H, \forall \psi_h \in V_h. \tag{2.22}$$

Preuve : Soit $\phi_h \in V_h$. D'après la définition de la forme bilinéaire a_i, $i = 1$ ou 2, on a

$$(A_{i,h}\phi_h, v_h)_H = a_i(\phi_h, v_h), \forall v_h \in V_h,$$

en prenant $v_h = A_{i,h}\phi_h$ et par la continuité de $a_i(\cdot,\cdot)$, on a

$$\|A_{i,h}\phi_h\|_H^2 = a_i(\phi_h, A_{i,h}\phi_h) \leq M_i\|\phi_h\|\|A_{i,h}\phi_h\|.$$

(2.21) se déduit alors de l'hypothèse $\mathbf{H_2'}$. Soit maintenant $\psi_h \in V_h$, en appliquant l'inégalité (2.21) à $\phi_h = A_{i,h}^{-1}\psi_h$, on trouve

$$\|\psi_h\|_H \lesssim h^{-2}\|A_{i,h}^{-1}\psi_h\|_H,$$

d'où (2.22). ∎

Proposition 2.1.7. *Sous les hypothèses* $\mathbf{H_1'}$*-*$\mathbf{H_3'}$*, on a*

$$\|A_{1,h}u_h\|_H \lesssim \|A_{2,h}u_h\|_H \lesssim \|A_{1,h}u_h\|_H, \quad \forall u_h \in V_h.$$

Preuve : Grâce à l'hypothèse $\mathbf{H_3'}$, on a pour $u_h \in V_h$

$$\|A_{i,h}^{-1}u_h - A_i^{-1}u_h\|_H \lesssim h^2\|u_h\|_H, \quad i = 1, 2. \tag{2.23}$$

Par (2.23) pour $i = 1$ et le Lemme 2.1.5, on a

$$\begin{aligned}\|A_{1,h}^{-1}u_h\|_H &\lesssim \|A_1^{-1}u_h\|_H + h^2\|u_h\|_H \\ &\lesssim \|A_2^{-1}u_h\|_H + h^2\|u_h\|_H.\end{aligned}$$

Donc par l'inégalité (2.23) pour $i = 2$, on obtient

$$\|A_{1,h}^{-1}u_h\|_H \lesssim \|A_{2,h}^{-1}u_h\|_H + h^2\|u_h\|_H,$$

et d'après le Lemme 2.1.6, on conclut que $\|A_{1,h}^{-1}u_h\|_H \lesssim \|A_{2,h}^{-1}u_h\|_H$. On vérifie de même que

$$\|A_{2,h}^{-1}u_h\|_H \lesssim \|A_{1,h}^{-1}u_h\|_H. \tag{2.24}$$

Dès lors,

$$\begin{aligned}\|A_{1,h}u_h\|_H &= \sup_{\varphi_h\in V_h,\varphi_h\neq 0}\frac{\left|\left(A_{1,h}u_h,\varphi_h\right)_H\right|}{\|\varphi_h\|_H}\\ &= \sup_{\varphi_h\in V_h,\varphi_h\neq 0}\frac{\left|\left(A_{1,h}A_{2,h}^{-1}A_{2,h}u_h,\varphi_h\right)_H\right|}{\|\varphi_h\|_H}.\end{aligned}$$

On obtient en utilisant l'inégalité de Cauchy-Schwarz

$$\|A_{1,h}u_h\|_H \leq \|A_{2,h}u_h\|_H \sup_{\varphi_h\in V_h,\varphi_h\neq 0}\frac{\|A_{2,h}^{-1}A_{1,h}\varphi_h\|_H}{\|\varphi_h\|_H},$$

mais par (2.24), on a

$$\|A_{2,h}^{-1}A_{1,h}\varphi_h\|_H \lesssim \|\varphi_h\|_H,$$

ce qui donne la première inégalité. La seconde s'obtient en intervertissant les rôles des indices 1 et 2. ∎

Comme dans le cas continu et afin de démontrer des estimations d'erreur, on utilise le changement d'inconnue suivant

$$\begin{cases} v_h = (I+\epsilon A_{1,h})^{\frac{1}{2}}u_h,\\ v_{0,h} = (I+\epsilon A_{1,h})^{\frac{1}{2}}u_{0,h}.\end{cases} \tag{2.25}$$

Par conséquent le problème (2.16) est équivalent à

$$\begin{cases} v_{h,t}+\mathcal{B}_{\epsilon,h}v_h = 0, & \text{dans } V_h, \quad 0<t\leq T,\ \epsilon>0,\\ v_h(0) = v_{0,h}, & \text{dans } V_h,\end{cases} \tag{2.26}$$

où $\mathcal{B}_{\epsilon,h} = (I+\epsilon A_{1,h})^{-\frac{1}{2}}A_{2,h}(I+\epsilon A_{1,h})^{-\frac{1}{2}}$. Par le schéma d'Euler implicite et le schéma de Crank-Nicolson, la solution approchée du problème (2.26) est donnée par les formules suivantes respectivement

$$v_{n+1,h} = (I+\Delta t\mathcal{B}_{\epsilon,h})^{-1}v_{n,h}, \quad 0\leq n\leq N-1, \tag{2.27}$$

$$v_{n+1,h} = (I+\frac{1}{2}\Delta t\mathcal{B}_{\epsilon,h})^{-1}(I-\frac{1}{2}\Delta t\mathcal{B}_{\epsilon,h})v_{n,h}, \quad 0\leq n\leq N-1. \tag{2.28}$$

Par récurrence et le fait que la solution exacte du problème (2.26) est donnée par

$$v_h(t_n) = v_{0,h}e^{-n\Delta t\ \mathcal{B}_{\epsilon,h}}, \quad \forall\, 0\leq n\leq N,$$

il est facile de voir que l'erreur entre la solution exacte du problème (2.26) et la solution approchée donnée par le schéma d'Euler implicite (2.27) (ou le schéma de Crank-Nicolson (2.28)) vérifie pour tout $0\leq n\leq N$

$$v_h(t_n) - v_{n,h} = E_n(\Delta t\mathcal{B}_{\epsilon,h})v_{0,h},$$

où pour un opérateur linéaire $\mathcal{B}$

$$E_n(\mathcal{B}) = \begin{cases} (I+\mathcal{B})^{-n} - e^{-n\mathcal{B}}, & \text{(pour Euler implicite)}, \\ (I+\frac{1}{2}\mathcal{B})^{-n}(I-\frac{1}{2}\mathcal{B})^n - e^{-n\mathcal{B}}, & \text{(pour Crank-Nicolson)}. \end{cases} \tag{2.29}$$

Lemme 2.1.8. *Sous les hypothèses* $\mathbf{H_1'}$ *à* $\mathbf{H_3'}$, *pour tout* $w \in V$, *on a l'estimation suivante*

$$\|A_2^{-1}w - \mathcal{B}_\epsilon^{-1}w\|_H + \|A_{2,h}^{-1}P_h w - \mathcal{B}_{\epsilon,h}^{-1}P_h w\|_H \lesssim \epsilon\|w\|_H. \tag{2.30}$$

Par conséquent

$$\|\mathcal{B}_\epsilon^{-1}w - \mathcal{B}_{\epsilon,h}^{-1}P_h w\|_H \lesssim (\epsilon + h^2)\|w\|_H. \tag{2.31}$$

Preuve : Soit $w \in V$. Par définition et l'inégalité triangulaire, on a

$$\begin{aligned} \|A_2^{-1}w - \mathcal{B}_\epsilon^{-1}w\|_H &= \|A_2^{-1}w - (I+\epsilon A_1)^{\frac{1}{2}}A_2^{-1}(I+\epsilon A_1)^{\frac{1}{2}}w\|_H \\ &\le \|\big(I-(I+\epsilon A_1)^{\frac{1}{2}}\big)A_2^{-1}w\|_H + \|A_2^{-1}(w-(I+\epsilon A_1)^{\frac{1}{2}}w)\|_H \\ &\quad + \|\big((I+\epsilon A_1)^{\frac{1}{2}} - I\big)A_2^{-1}\big(w-(I+\epsilon A_1)^{\frac{1}{2}}w\big)\|_H. \end{aligned}$$

Donc on estime les trois termes du membre de droite de cette inégalité. Pour tout $y \in D(A_1)$, on a par la formule de Parseval

$$\begin{aligned} \|\big(I-(I+\epsilon A_1)^{\frac{1}{2}}\big)y\|_H^2 &= \sum_{l=1}^{\infty}(1-(1+\epsilon\nu_l)^{\frac{1}{2}})^2|(y,\psi_l)_H|^2 \\ &= \sum_{l=1}^{\infty}\frac{\epsilon^2\nu_l^2}{(1+(1+\epsilon\nu_l)^{\frac{1}{2}})^2}|(y,\psi_l)_H|^2 \\ &\le \epsilon^2\sum_{l=1}^{\infty}\nu_l^2|(y,\psi_l)_H|^2 \\ &= \epsilon^2\|A_1 y\|_H^2. \end{aligned}$$

En utilisant l'hypothèse $\mathbf{H_1'}$, on obtient alors

$$\|\big(I-(I+\epsilon A_1)^{1/2}\big)A_2^{-1}w\|_H \lesssim \epsilon\|A_1A_2^{-1}w\|_H \lesssim \epsilon\|w\|_H. \tag{2.32}$$

Par la même formule de Parseval, on a

$$\begin{aligned} \|w-(I+\epsilon A_1)^{1/2}w\|_{V'}^2 &= \sum_{l=1}^{\infty}|1-(1+\epsilon\nu_l)^{1/2}|^2\frac{1}{\nu_l^2}|(w,\psi_l)_H|^2 \\ &\lesssim \sum_{l=1}^{\infty}\frac{\epsilon^2\nu_l^2}{\nu_l^2}|(w,\psi_l)_H|^2 \\ &\lesssim \epsilon^2\|w\|_H^2, \end{aligned}$$

par conséquent et comme l'opérateur A_2^{-1} est borné de V' dans V, on obtient

$$\|A_2^{-1}(w-(I+\epsilon A_1)^{1/2}w)\| \lesssim \epsilon\|w\|_H. \tag{2.33}$$

De la même façon et pour $\chi \in V$, on a

$$\begin{aligned}\|((I+\epsilon A_1)^{1/2}-I)\chi\|_H^2 &= \sum_{l=1}^{\infty}|(1+\epsilon\nu_l)^{1/2}-1|^2|(\chi,\psi_l)_H|^2\\ &\leq \sum_{l=1}^{\infty}\epsilon\nu_l|(\chi,\psi_l)_H|^2\\ &\lesssim \epsilon\|\chi\|^2_{D(A_1^{\frac{1}{2}})}\\ &= \epsilon\|\chi\|^2,\end{aligned}$$

ce qui donne par (2.33)

$$\begin{aligned}\|((I+\epsilon A_1)^{1/2}-I)A_2^{-1}(w-(I+\epsilon A_1)^{1/2}w)\|_H &\leq \sqrt{\epsilon}\|A_2^{-1}(w-(I+\epsilon A_1)^{1/2}w)\|\\ &\lesssim (\epsilon)^{3/2}\|w\|_H. \qquad (2.34)\end{aligned}$$

Des inégalités (2.32)-(2.34), on obtient alors

$$\|A_2^{-1}w-\mathcal{B}_\epsilon^{-1}w\|_H \lesssim \epsilon\|w\|_H.$$

On procède de même pour majorer $\|A_{2,h}^{-1}P_hw-\mathcal{B}_{\epsilon,h}^{-1}P_hw\|_H$ par $\epsilon\|w\|_H$. L'opérateur linéaire $A_{1,h}$ étant symétrique dans l'espace de dimension finie V_h, il admet un nombre fini de valeurs propres qu'on notera $\{\nu_{l,h}\}_{l=1}^{N_h}$ et les vecteurs propres associés $\{\psi_{l,h}\}_{l=1}^{N_h}$ forment une base orthonormale de V_h. L'hypothèse $\mathbf{H_1'}$ est remplacée par la Proposition 2.1.7. Ainsi, on obtient (2.30).
Pour (2.31), on a par l'inégalité triangulaire

$$\|\mathcal{B}_\epsilon^{-1}w-\mathcal{B}_{\epsilon,h}^{-1}P_hw\|_H \leq \|A_2^{-1}w-\mathcal{B}_\epsilon^{-1}w\|_H+\|A_{2,h}^{-1}P_hw-\mathcal{B}_{\epsilon,h}^{-1}P_hw\|_H+\|A_2^{-1}w-A_{2,h}^{-1}P_hw\|,$$

d'où le résultat grâce à (2.30) et l'hypothèse $\mathbf{H_3'}$. ∎

Lemme 2.1.9. *Pour tout $v_h \in V_h$, on a*

$$\|(I+\epsilon A_{1,h})^{-\frac{1}{2}}v_h\|_H \leq \|v_h\|_H.$$

Preuve : Soit $v_h \in V_h$, alors v_h s'écrit dans la base $\{\psi_{l,h}\}_{l=1}^{N_h}$

$$v_h = \sum_{l=1}^{N_h}(v_h,\psi_{l,h})_H\psi_{l,h},$$

ce qui entraîne

$$(I+\epsilon A_{1,h})^{-\frac{1}{2}}v_h = \sum_{l=1}^{N_h}(1+\epsilon\nu_{l,h})^{-\frac{1}{2}}(v_h,\psi_{l,h})_H\psi_{l,h},$$

et donc par la formule de Parseval

$$\begin{aligned}\|(I+\epsilon A_{1,h})^{-\frac{1}{2}}v_h\|_H^2 &= \sum_{l=1}^{N_h}(1+\epsilon\nu_{l,h})^{-1}|(v_h,\psi_{l,h})_H|^2\\ &\leq \sum_{l=1}^{N_h}|(v_h,\psi_{l,h})_H|^2=\|v_h\|_H^2.\end{aligned}$$

■

Proposition 2.1.10. *On suppose que les hypothèses* $\mathbf{H_1'}$ *à* $\mathbf{H_5'}$ *ont lieu et que* $u_0\in D(A_1)$. *Soient* v_h *la solution du problème* (2.26) *et* $v_{n,h}$ *la solution approchée donnée par le schéma d'Euler implicite* (2.27). *Alors*

$$\|v_h(t_n)-v_{n,h}\|_H\lesssim(\epsilon+\Delta t+h^2)\|u_0\|_{D(A_1)},\quad\forall\,1\leq n\leq N. \tag{2.35}$$

Et par conséquent

$$\|u_h(t_n)-u_{n,h}\|_H\lesssim(\epsilon+\Delta t+h^2)\|u_0\|_{D(A_1)},\quad\forall\,1\leq n\leq N, \tag{2.36}$$

où u_h *est la solution du problème* (2.17) *et* $u_{n,h}$ *la solution approchée donnée par le schéma d'Euler implicite* (2.19).

Preuve : Par l'inégalité triangulaire, on a

$$\|v_h(t_n)-v_{n,h}\|_H=\|E_n(\Delta t\mathcal{B}_{\epsilon,h})v_{0,h}\|_H\leq\|E_n(\Delta t\mathcal{B}_{\epsilon,h})P_hv_0\|_H+\|E_n(\Delta t\mathcal{B}_{\epsilon,h})(v_{0,h}-P_hv_0)\|_H.$$

Or pour tout $v\in V$, il est facile de vérifier par simple calcul que

$$P_hv=\sum_{j=0}^{q-1}\mathcal{B}_{\epsilon,h}^{-j}P_h(\mathcal{B}_\epsilon^{-1}-\mathcal{B}_{\epsilon,h}^{-1}P_h)\mathcal{B}_\epsilon^{j+1}v+\mathcal{B}_{\epsilon,h}^{-q}P_h\mathcal{B}_\epsilon^qv,\quad q=1\ \text{ ou }\ 2, \tag{2.37}$$

où $\mathcal{B}_{\epsilon,h}^0=I$. En appliquant cette égalité pour $v=v_0$ et $q=1$, on obtient

$$\|E_n(\Delta t\mathcal{B}_{\epsilon,h})P_hv_0\|_H\leq\|E_n(\Delta t\mathcal{B}_{\epsilon,h})P_h(\mathcal{B}_\epsilon^{-1}-\mathcal{B}_{\epsilon,h}^{-1}P_h)\mathcal{B}_\epsilon v_0\|_H+\|E_n(\Delta t\mathcal{B}_{\epsilon,h})\mathcal{B}_{\epsilon,h}^{-1}P_h\mathcal{B}_\epsilon v_0\|_H.$$

Par [71, Lemme 7.2, p. 124], on a

$$\|E_n(\Delta t\mathcal{B}_{\epsilon,h})P_h\mathcal{B}_{\epsilon,h}^{-j}\|_{\mathcal{L}(H)}\lesssim(\Delta t)^j,\quad 0\leq j\leq q,n\geq 0. \tag{2.38}$$

Appliquons cette inégalité au membre de droite de la précédente, d'abord avec $j=0$ pour le premier terme ensuite avec $j=1$ pour le second, on obtient

$$\|E_n(\Delta t\mathcal{B}_{\epsilon,h})P_hv_0\|_H\lesssim\|(\mathcal{B}_\epsilon^{-1}-\mathcal{B}_{\epsilon,h}^{-1}P_h)\mathcal{B}_\epsilon v_0\|_H+\Delta t\|\mathcal{B}_\epsilon v_0\|_H.$$

En utilisant le Lemme 2.1.8, l'inégalité (2.13) et l'hypothèse $\mathbf{H_1'}$, on en déduit

$$\|E_n(\Delta t\mathcal{B}_{\epsilon,h})P_hv_0\|_H \lesssim (\epsilon + h^2 + \Delta t)\|u_0\|_{D(A_1)}. \tag{2.39}$$

Pour l'autre terme, encore une fois par (2.38) avec $j = 0$, on a

$$\begin{aligned}\|E_n(\Delta t\mathcal{B}_{\epsilon,h})(v_{0,h} - P_hv_0)\|_H \lesssim \|v_{0,h} - P_hv_0\|_H &= \|P_h(v_{0,h} - v_0)\|_H \\ &\leq \|v_{0,h} - v_0\|_H.\end{aligned}$$

Or

$$\begin{aligned}\|v_{0,h} - v_0\|_H &= \|(I + \epsilon A_{1,h})^{\frac{1}{2}}P_hu_0 - (I + \epsilon A_1)^{\frac{1}{2}}u_0\|_H \\ &\leq \|(I + \epsilon A_{1,h})^{\frac{1}{2}}P_hu_0 - P_hu_0\|_H + \|u_0 - (I + \epsilon A_1)^{\frac{1}{2}}u_0\|_H + \|P_hu_0 - u_0\|_H.\end{aligned}$$

Par l'hypothèse $\mathbf{H_4'}$, on a

$$\|P_hu_0 - u_0\|_H \lesssim h^2\|u_0\|_{D(A_1)}. \tag{2.40}$$

Pour les autres termes, on a déjà démontré dans la preuve du Lemme 2.1.8 que

$$\|u_0 - (I + \epsilon A_1)^{\frac{1}{2}}u_0\|_H \lesssim \epsilon\|A_1u_0\|_H \leq \epsilon\|u_0\|_{D(A_1)},$$

et

$$\|(I + \epsilon A_{1,h})^{\frac{1}{2}}P_hu_0 - P_hu_0\|_H \lesssim \epsilon\|A_{1,h}P_hu_0\|_H.$$

Donc il suffit d'estimer $\|A_{1,h}P_hu_0\|_H$. En effet, on a par définition

$$\|A_{1,h}P_hu_0\|_H = \sup_{v_h\in V_h, v_h\neq 0} \frac{\left|(A_{1,h}P_hu_0, v_h)_H\right|}{\|v_h\|_H}.$$

Comme $A_{1,h}P_hu_0 \in H$, on a

$$\begin{aligned}\left|(A_{1,h}P_hu_0, v_h)_H\right| &= |a_1(P_hu_0, v_h)| \\ &\leq |a_1(P_hu_0 - u_0, v_h)| + |a_1(u_0, v_h)| \\ &= |a_1(P_hu_0 - u_0, v_h)| + |(A_1u_0, v_h)_H| \\ &\leq M_1\|P_hu_0 - u_0\|\|v_h\| + \|A_1u_0\|_H\|v_h\|_H.\end{aligned}$$

Et utilisant les hypothèses $\mathbf{H_2'}$ et $\mathbf{H_5'}$, on obtient

$$\left|(A_{1,h}P_hu_0, v_h)_H\right| \lesssim \|u_0\|_{D(A_1)}\|v_h\|_H,$$

ce qui donne

$$\|A_{1,h}P_hu_0\|_H \lesssim \|u_0\|_{D(A_1)}. \tag{2.41}$$

Par conséquent

$$\|v_{0,h} - v_0\|_H \lesssim (\epsilon + h^2)\|u_0\|_{D(A_1)}. \tag{2.42}$$

De (2.39) et (2.42), on conclut la majoration (2.35).
Finalement, puisque

$$\|u_h(t_n) - u_{n,h}\|_H = \|(I + \epsilon A_{1,h})^{-\frac{1}{2}}(v_h(t_n) - v_{n,h})\|_H,$$

l'estimation (2.44) se déduit alors du Lemme 2.1.9 et de (2.35). ∎

Remarque 2.1.11. *De la même façon, si on a $u_0 \in D(A_2^2)$, v_h la solution du problème (2.26) et $v_{n,h}$ la solution approchée donnée par le schéma de Crank-Nicolson (2.28), on peut démontrer que*

$$\|v_h(t_n) - v_{n,h}\|_H \lesssim (\epsilon + h^2 + (\Delta t)^2)\|u_0\|_{D(A_2^2)}, \quad \forall\, 1 \leq n \leq N. \tag{2.43}$$

Et par conséquent

$$\|u_h(t_n) - u_{n,h}\|_H \lesssim (\epsilon + h^2 + (\Delta t)^2)\|u_0\|_{D(A_2^2)}, \quad \forall\, 1 \leq n \leq N, \tag{2.44}$$

où u_h est la solution du problème (2.17) et $u_{n,h}$ la solution approchée donnée par le schéma Crank-Nicolson (2.20).

Proposition 2.1.12. *Sous les hypothèses $\mathbf{H_1'}$ à $\mathbf{H_5'}$, on suppose que $u_0 \in D(A_2^2)$. Soient v et v_h les solutions des problèmes (2.11) et (2.26) respectivement. Alors*

$$\|v_h(t) - v(t)\|_H \lesssim (\epsilon + h^2)\|u_0\|_{D(A_2^2)}.$$

Preuve : L'erreur $e_h = v_h - v$ satisfait l'identité suivante

$$\begin{aligned} \mathcal{B}_{\epsilon,h}^{-1} P_h e_{h,t} + e_h &= \mathcal{B}_{\epsilon,h}^{-1}(P_h \mathcal{B}_\epsilon v - \mathcal{B}_{\epsilon,h} v_h) + v_h - v \\ &= -(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h)\mathcal{B}_\epsilon v. \end{aligned}$$

Si on note $\rho = -(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h)\mathcal{B}_\epsilon v$, alors pour tout $s \in [0,T]$, et tout $\varphi \in H$

$$(\mathcal{B}_{\epsilon,h}^{-1} P_h e_{h,t}(s), \varphi)_H + (e_h(s), \varphi)_H = (\rho(s), \varphi)_H, \tag{2.45}$$

en choisissant $\varphi = 2e_{h,t}(s)$ dans (2.45), on obtient

$$2(\mathcal{B}_{\epsilon,h}^{-1} P_h e_{h,t}, e_{h,t})_H + 2(e_h, e_{h,t})_H = 2(\rho, e_{h,t})_H = 2\frac{d}{dt}(\rho, e_h)_H - 2(\rho_t, e_h)_H.$$

Or, $(\mathcal{B}_{\epsilon,h}^{-1} P_h e_{h,t}(s), e_{h,t}(s))_H = (\mathcal{B}_{\epsilon,h}^{-1} P_h e_{h,t}(s), P_h e_{h,t}(s))_H \geq 0$ puisque $\mathcal{B}_{\epsilon,h}$ est défini positif. Et alors, après une intégration par rapport à s, et en appliquant l'inégalité de Cauchy-Schwarz, on trouve

$$\begin{aligned} \|e_h(t)\|_H^2 &\leq \|e_h(0)\|_H^2 + 2\|\rho(t)\|_H\|e_h(t)\|_H + 2\|\rho(0)\|_H\|e_h(0)\|_H + 2\int_0^t \|\rho_t(s)\|_H\|e_h(s)\|_H ds \\ &\leq \sup_{s\leq t}\|e_h(s)\|_H\Big(\|e_h(0)\|_H + 4\sup_{s\leq t}\|\rho(s)\|_H + 2\int_0^t \|\rho_t(s)\|_H ds\Big). \end{aligned}$$

Comme v est la solution exacte du problème (2.11), alors

$$v(s) = v_0 + \int_0^s v_t(t)dt,$$

et grâce à la définition de ρ, on a

$$\rho(s) = \rho(0) + \int_0^s \rho_t(t)dt, \quad \forall s \in [0,T].$$

Alors

$$\sup_{s\in[0,T]} \|\rho(s)\|_H \leq \|\rho(0)\|_H + \int_0^T \|\rho_t(t)\|_H dt.$$

Par conséquent, si $\|e_h(\tau)\| = \sup_{s\in[0,T]} \|e_h(s)\|_H$, on obtient

$$\|e_h(t)\|_H \leq \|e_h(\tau)\|_H \lesssim \|e_h(0)\|_H + \|\rho(0)\|_H + \int_0^T \|\rho_t(t)\|_H dt.$$

D'une part, on a par (2.42)

$$\|e_h(0)\|_H = \|v_{0,h} - v_0\|_H \lesssim (\epsilon + h^2)\|u_0\|_{D(A_1)}.$$

D'autre part, en utilisant le Lemme 2.1.8 et (2.13), on conclut que

$$\|\rho(0)\|_H = \|(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1}P_h)\mathcal{B}_\epsilon v_0\|_H \lesssim (\epsilon + h^2)\|\mathcal{B}_\epsilon v_0\|_H \lesssim (\epsilon + h^2)\|u_0\|_{D(A_2)}.$$

Il nous reste donc à estimer $\int_0^T \|\rho_t(s)\|_H ds$. Par définition et le Lemme 2.1.8, on a

$$\begin{aligned}\int_0^T \|\rho_t(s)\|_H ds &= \int_0^T \|(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1}P_h)\mathcal{B}_\epsilon v_t(s)\|_H ds \\ &\leq (\epsilon + h^2)\int_0^T \|\mathcal{B}_\epsilon v_t(s)\|_H ds,\end{aligned}$$

ce qui donne le résultat si on démontre que

$$\int_0^T \|\mathcal{B}_\epsilon v_t(s)\|_H ds \lesssim \|u_0\|_{D(A_2^2)}. \tag{2.46}$$

En effet, comme v est la solution exacte du problème (2.11), on a

$$\big(v_t(s) + \mathcal{B}_\epsilon v(s), \mathcal{B}_\epsilon v(s)\big)_H = 0,$$

alors par l'inégalité de Cauchy-Schwarz et par intégration par rapport à t, on obtient

$$\int_0^T \|\mathcal{B}_\epsilon v(s)\|_H ds \leq \int_0^T \|v_t(s)\|_H ds. \tag{2.47}$$

On pose maintenant $w = v_t$, donc w vérifie l'équation : $w_t + \mathcal{B}_\epsilon w = 0$, dans $H, 0 < t \leq T$, $\epsilon > 0$, et $w(0) = -\mathcal{B}_\epsilon v_0$, et comme l'opérateur $\mathcal{B}_\epsilon$ est auto-adjoint défini positif, on en déduit que

$$\|v_t(s)\|_H = \|w(s)\|_H = \|e^{-s\mathcal{B}_\epsilon} w(0)\|_H \lesssim \|w(0)\|_H = \|\mathcal{B}_\epsilon v_0\|_H.$$

Alors, (2.47) implique

$$\int_0^T \|\mathcal{B}_\epsilon v(s)\|_H ds \lesssim \|\mathcal{B}_\epsilon v_0\|_H,$$

par conséquent

$$\begin{aligned} \int_0^T \|\mathcal{B}_\epsilon v_t(s)\|_H ds = \int_0^T \|\mathcal{B}_\epsilon w(s)\|_H ds &\lesssim \|\mathcal{B}_\epsilon w(0)\|_H \\ &= \|\mathcal{B}_\epsilon^2 v_0\|_H. \end{aligned}$$

Et à la fin en utilisant la dernière estimation dans la preuve du Lemme 2.1.3, on conclut (2.46). ∎

Dans le Corollaire suivant, $p = 1$ si on considère le schéma d'Euler et $p = 2$ si on considère celui de Crank-Nicolson.

Corollaire 2.1.13. *Sous les hypothèses* $\mathbf{H_1'}$ - $\mathbf{H_5'}$, *on suppose que* $u_0 \in D(A_2^2)$. *Soit* v *la solution exacte du problème* (2.11) *et* $v_{n,h}$ *la solution totalement discrète donnée par le schéma numérique* (2.27) (*ou* (2.28)). *Alors*

$$\|v(t_n) - v_{n,h}\|_H \lesssim (\epsilon + h^2 + (\Delta t)^p)\|u_0\|_{D(A_2^2)}. \tag{2.48}$$

Par conséquent, si u *la solution exacte du problème* (2.1) *et* $u_{n,h}$ *la solution totalement discrète donnée par le schéma* (2.8). *Alors*

$$\|u(t_n) - u_{n,h}\|_* \lesssim (\epsilon + h^2 + (\Delta t)^p)\|u_0\|_{D(A_2^2)}, \tag{2.49}$$

où $\|\cdot\|_*$ *est la norme définie par*

$$\|y\|_* = \|A_1^{-1} y\|_H, \quad \forall y \in H.$$

Preuve : La première estimation est une conséquence directe des Propositions 2.1.12 et 2.1.10 pour le schéma d'Euler, de la Proposition 2.1.12 et la Remarque 2.1.11 pour le schéma de Crank-Nicolson. Pour la deuxième estimation, on a par l'inégalité triangulaire

$$\begin{aligned} \|u(t_n) - u_{n,h}\|_* \leq\ & \|A_1^{-1}(u(t_n) - v(t_n))\|_H + \|A_1^{-1}(v(t_n) - v_{n,h})\|_H \\ & + \|(A_1^{-1} - A_{1,h}^{-1})(v_{n,h} - u_{n,h})\|_H + \|A_{1,h}^{-1}(v_{n,h} - u_{n,h})\|_H. \end{aligned}$$

Par la formule de Parseval, on a

$$\begin{aligned}\|A_1^{-1}(u(t_n)-v(t_n))\|_H^2 &= \|A_1^{-1}\big((I+\epsilon A_1)^{-1/2}v(t_n)-v(t_n)\big)\|_H^2\\ &= \sum_{l=1}^{\infty}\nu_l^{-2}((1+\epsilon\nu_l)^{-1/2}-1)^2|(v(t_n),\psi_l)_H|^2\\ &= \sum_{l=1}^{\infty}\nu_l^{-2}\frac{\epsilon^2\nu_l^2}{((1+\epsilon\nu_l)^{1/2}+1)^2(1+\epsilon\nu_l)}|(v(t_n),\psi_l)_H|^2\\ &\lesssim \epsilon^2\sum_{l=1}^{\infty}|(v(t_n),\psi_l)_H|^2=\epsilon^2\|v(t_n)\|_H^2.\end{aligned}$$

Pour le deuxième terme, on a grâce à (2.48)

$$\|A_1^{-1}(v(t_n)-v_{n,h})\|_H\lesssim\|v(t_n)-v_{n,h}\|_H\lesssim(\epsilon+h^2+(\Delta t)^p)\|u_0\|_{D(A_2^2)}.$$

Grâce à l'hypothèse $\mathbf{H_3'}$ et le Lemme 2.1.9, on a

$$\begin{aligned}\|(A_1^{-1}-A_{1,h}^{-1})(v_{n,h}-u_{n,h})\|_H &\lesssim h^2\|v_{n,h}-u_{n,h}\|_H\\ &= h^2\|v_{n,h}-(I+\epsilon A_{1,h})^{-1/2}v_{n,h}\|_H\\ &\le h^2\big(\|v_{n,h}\|_H+\|(I+\epsilon A_{1,h})^{-1/2}v_{n,h}\|_H\big)\\ &\lesssim h^2\|v_{n,h}\|_H.\end{aligned}$$

En utilisant toujours la formule de Parseval, on a pour le dernier terme

$$\begin{aligned}\|A_{1,h}^{-1}(v_{n,h}-u_{n,h})\|_H^2 &= \sum_{l=1}^{N_h}\nu_{l,h}^{-2}(1-(1+\epsilon\nu_{l,h})^{-1/2})^2|(v_{n,h},\psi_{l,h})_H|^2\\ &= \sum_{l=1}^{N_h}\nu_{l,h}^{-2}\frac{\epsilon^2\nu_{l,h}^2}{((1+\epsilon\nu_{l,h})^{1/2}+1)^2(1+\epsilon\nu_l)}|(v_{n,h},\psi_{l,h})_H|^2\\ &\lesssim \epsilon^2\sum_{l=1}^{N_h}|(v_{n,h},\psi_{l,h})_H|^2=\epsilon^2\|v_{n,h}\|_H^2.\end{aligned}$$

Pour terminer la preuve, il nous reste alors àmajorer $\|v_{n,h}\|_H$ et $\|v(t_n)\|_H$. Et comme d'après (2.48), on a

$$\begin{aligned}\|v_{n,h}\|_H &\le \|v_{n,h}-v(t_n)\|_H+\|v(t_n)\|_H\\ &\lesssim (\epsilon+h^2+(\Delta t)^p)\|u_0\|_{D(A_2^2)}+\|v(t_n)\|_H,\end{aligned}$$

ceci revient donc àestimer $\|v(t)\|_H$ pour $t\in[0,T]$. Or, v est la solution du problème (2.11)

et comme $\mathcal{B}_\epsilon$ est auto-adjoint défini positif, pour tout $t \in [0,T]$

$$\begin{aligned}
\|v(t)\|_H^2 \leq \|v_0\|_H^2 &= \|(I+\epsilon A_1)^{\frac{1}{2}} u_0\|_H^2 \\
&= \sum_{l=1}^{\infty} (1+\epsilon\nu_l)|(u_0,\psi_l)_H|^2 \\
&= \sum_{l=1}^{\infty} |(u_0,\psi_l)_H|^2 + \epsilon \sum_{l=1}^{\infty} \nu_l |(u_0,\psi_l)_H|^2 \\
&= \|u_0\|_H^2 + \epsilon \|A_1^{\frac{1}{2}} u_0\|_H^2 \\
&\lesssim \|u_0\|^2 \lesssim \|u_0\|_{D(A_2)},
\end{aligned}$$

ce qui permet de conclure. ∎

Si $A_1 = A_2 = A$, le problème (2.1) avec $f = 0$ se réduit à

$$\begin{cases} u_t + \epsilon A u_t + Au = 0, & \text{dans } V', \quad 0 < t \leq T, \\ u(0) = u_0, & \text{dans } V. \end{cases} \tag{2.50}$$

Ce type de problème est une approximation du problème (7) introduite pour la première fois par Yosida [81].
Dans ce cas, l'opérateur $\mathcal{A}_\epsilon = (I+\epsilon A)^{-1}A$ étant auto-adjoint défini positif dans l'espace de Hilbert H, on n'a pas besoin d'introduire l'opérateur $\mathcal{B}_\epsilon$. Les techniques utilisées dans cette section seront appliquées directement à l'opérateur $\mathcal{A}_\epsilon$ pour montrer le résultat suivant.

Théorème 2.1.14. *Sous les hypothèses* $\mathbf{H_3'}$ *et* $\mathbf{H_4'}$, *on suppose que* $u_0 \in D(A^3)$. *Soit* u *la solution exacte du problème* (2.50) *et soit* $u_{n,h}$ *la solution totalement discrète donnée par le schéma d'Euler implicite* (2.19). *Alors on a l'estimation suivante*

$$\|u(t_n) - u_{n,h}\|_H \lesssim (\Delta t + h^2)\|u_0\|_{D(A^3)}, \quad 0 \leq n \leq N. \tag{2.51}$$

Preuve : Par l'inégalité triangulaire, on a

$$\|u(t_n) - u_{n,h}\|_H \leq \|u_{n,h} - u_h(t_n)\|_H + \|u_h(t_n) - u(t_n)\|_H. \tag{2.52}$$

Pour estimer le premier terme du membre de droite de cette inégalité, on procède de même que dans la preuve de la Proposition 2.1.10, en remplaçant l'opérateur $\mathcal{B}_\epsilon$ par $\mathcal{A}_\epsilon$. Ainsi, on a

$$\begin{aligned}
\|u_{n,h} - u_h(t_n)\|_H &= \|E_n(\Delta t \mathcal{A}_{\epsilon,h})P_h u_0\|_H \\
&\leq \|E_n(\Delta t \mathcal{A}_{\epsilon,h})P_h(\mathcal{A}_\epsilon^{-1} - \mathcal{A}_{\epsilon,h}^{-1}P_h)\mathcal{A}_\epsilon u_0\|_H + \|E_n(\Delta t \mathcal{A}_{\epsilon,h})\mathcal{A}_{\epsilon,h}^{-1}P_h\mathcal{A}_\epsilon u_0\|_H.
\end{aligned}$$

Par le Lemme 7. 2 de [71], on obtient

$$\|u_{n,h} - u_h(t_n)\|_H \lesssim \|(\mathcal{A}_\epsilon^{-1} - \mathcal{A}_{\epsilon,h}^{-1}P_h)\mathcal{A}_\epsilon u_0\|_H + \Delta t \|\mathcal{A}_\epsilon u_0\|_H.$$

En utilisant maintenant les hypothèses $\mathbf{H_3'}$, $\mathbf{H_4'}$ et le Lemme 2.1.1, on trouve

$$\begin{aligned}\|(\mathcal{A}_\epsilon^{-1}-\mathcal{A}_{\epsilon,h}^{-1}P_h)\mathcal{A}_\epsilon u_0\|_H &= \|(A^{-1}+\epsilon I)\mathcal{A}_\epsilon u_0-(A_h^{-1}+\epsilon I)P_h\mathcal{A}_\epsilon u_0\|_H\\ &\leq \|A^{-1}\mathcal{A}_\epsilon u_0-A_h^{-1}P_h\mathcal{A}_\epsilon u_0\|_H+\epsilon\|\mathcal{A}_\epsilon u_0-P_h\mathcal{A}_\epsilon u_0\|_H\\ &\lesssim \epsilon h^2\|\mathcal{A}_\epsilon u_0\|_H+\epsilon h^2\|\mathcal{A}_\epsilon u_0\|_{D(A)} \qquad (2.53)\\ &\lesssim h^2\|u_0\|_{D(A^2)}.\end{aligned}$$

Par conséquent

$$\|u_{n,h}-u_h(t_n)\|_H\lesssim(\Delta t+h^2)\|u_0\|_{D(A^2)}.$$

Maintenant, et afin d'estimer le second terme du membre de droite de l'inégalité (2.52), on utilise des arguments similaires à ceux de la preuve de la Proposition 2.1.12. On obtient

$$\|u_h(t_n)-u(t_n)\|_H\leq\|u_0-P_hu_0\|_H+\|\rho(0)\|_H+\int_0^T\|\rho_t(s)\|_Hds,$$

où $\rho=-(\mathcal{A}_\epsilon^{-1}-\mathcal{A}_{\epsilon,h}^{-1}P_h)\mathcal{A}_\epsilon u$.
Par l'hypothèse $\mathbf{H_4'}$, on a directement

$$\|u_0-P_hu_0\|_H\lesssim h^2\|u_0\|_{D(A)}.$$

D'autre part, on a d'après l'hypothèse $\mathbf{H_3'}$ et le Lemme 2.1.1

$$\|\rho(0)\|_H=\|(\mathcal{A}_\epsilon^{-1}-\mathcal{A}_{\epsilon,h}^{-1}P_h)\mathcal{A}_\epsilon u_0\|_H\lesssim h^2\|u_0\|_{D(A^2)}.$$

Appliquons l'inégalité (2.53) à $u_t(s)$ au lieu de u_0, on obtient

$$\begin{aligned}\int_0^T\|\rho_t(s)\|_Hds &\lesssim \int_0^T\|(\mathcal{A}_\epsilon^{-1}-\mathcal{A}_{\epsilon,h}^{-1}P_h)\mathcal{A}_\epsilon u_t(s)\|_Hds\\ &\lesssim h^2\int_0^T\big(\|\mathcal{A}_\epsilon u_t(s)\|_H+\|\mathcal{A}_\epsilon u_t(s)\|_{D(A)}\big)ds.\end{aligned}$$

Comme dans la preuve de la Proposition 2.1.12, on peut démontrer que

$$\int_0^T\|\mathcal{A}_\epsilon u_t\|_Hdt\lesssim\|\mathcal{A}_\epsilon^2u_0\|_H\lesssim\|u_0\|_{D(A^2)},$$

et aussi de démontrer, comme l'opérateur $\mathcal{A}_\epsilon$ est auto-adjoint sur $D(A)$, que

$$\int_0^T\|\mathcal{A}_\epsilon u_t\|_{D(A)}dt\lesssim\|\mathcal{A}_\epsilon^2u_0\|_{D(A)}\lesssim\|u_0\|_{D(A^3)},$$

d'où le résultat. ∎

2.2 Le problème non homogène

2.2.1 Le problème semi-discrétisé en temps

Revenons maintenant au problème non homogène (2.1) avec la condition initiale $u(0) = 0$, i.e., on considère le problème suivant

$$\begin{cases} (I + \epsilon A_1)u_t + A_2 u = f(t), & \text{dans } V', \quad 0 < t \leq T, \ \epsilon > 0, \\ u(0) = 0, & \text{dans } V, \end{cases} \tag{2.54}$$

où f est continûment différentiable de $[0, T]$ dans V'. Comme dans la section précédente, nous allons utiliser des schémas numériques implicites pour la discrétisation en temps de ce problème. Pour cela, on considère toujours la subdivision régulière $(t_i \Delta t)_{i=0}^N$ de l'intervalle $[0, T]$, où $N \in \mathbb{N}^*$ et $\Delta t = \frac{T}{N}$ le pas de discrétisation en temps. Donc, par le schéma d'Euler implicite et le schéma de Crank-Nicolson, la solution approchée du problème (2.54) est donnée par les formules suivantes respectivement

$$(I + \Delta t \mathcal{A}_\epsilon)u_{n+1} = u_n + \Delta t f_\epsilon(t_{n+1}), \quad 0 \leq n \leq N - 1, \tag{2.55}$$

$$(I + \frac{1}{2}\Delta t \mathcal{A}_\epsilon)u_{n+1} = (I - \frac{1}{2}\Delta t \mathcal{A}_\epsilon)u_n + \frac{1}{2}(f_\epsilon(t_{n+1}) + f_\epsilon(t_n)), \quad 0 \leq n \leq N - 1, \tag{2.56}$$

où $u_0 = 0$, $\mathcal{A}_\epsilon = (I + \epsilon A_1)^{-1} A_2$ et $f_\epsilon(t) = (I + \epsilon A_1)^{-1} f(t)$.

Proposition 2.2.1. *Sous l'hypothèse* $\mathbf{H_1'}$, *on suppose que* $f \in C^1([0, T], H) \cap C([0, T], D(A_1))$. *Soit* u *la solution exacte du problème* (2.54) *et soit* u_n *la solution approchée donnée par le schéma d'Euler implicite* (2.55). *Alors*

$$\|u(t_n) - u_n\|_H \lesssim \Delta t \Big(\sup_{s \in [0,T]} \|f(s)\|_{D(A_1)} + \int_0^T \|f'(\tau)\|_H d\tau \Big), \quad \forall\, 1 \leq n \leq N. \tag{2.57}$$

Preuve : Toujours par le changement d'inconnue (2.10), on peut réécrire le problème (2.54) sous la forme suivante

$$\begin{cases} v_t + \mathcal{B}_\epsilon v = g_\epsilon(t), & \text{dans } H, \quad 0 < t \leq T, \ \epsilon > 0, \\ v(0) = 0, & \text{dans } H, \end{cases} \tag{2.58}$$

où $\mathcal{B}_\epsilon = (I + \epsilon A_1)^{-\frac{1}{2}} A_2 (I + \epsilon A_1)^{-\frac{1}{2}}$ et $g_\epsilon(t) = (I + \epsilon A_1)^{-\frac{1}{2}} f(t)$.

Par la formule de variation des constantes, la solution exacte au point t_{n+1}, $0 \leq n \leq N - 1$, est donnée par

$$v(t_{n+1}) = e^{-\Delta t \mathcal{B}_\epsilon} v(t_n) + \int_0^{\Delta t} e^{-(\Delta t - \tau)\mathcal{B}_\epsilon} g_\epsilon(t_n + \tau) d\tau.$$

Dans le but de montrer l'estimation d'erreur (2.57), on commence d'abord par estimer l'erreur entre v et la solution semi-discrète en temps par le schéma d'Euler implicite donnée par

$$v_{n+1} = (I + \Delta t \mathcal{B}_\epsilon)^{-1} v_n + \Delta t g_\epsilon(t_{n+1}), \quad 0 \leq n \leq N-1. \tag{2.59}$$

Donc par récurrence, l'erreur $e_n = v_n - v(t_n)$ satisfait

$$e_n = \Delta t \sum_{j=0}^{n-1} E_{n-j-1}(\Delta t \mathcal{B}_\epsilon) \Big(g_\epsilon(t_{j+1}) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_\epsilon} g_\epsilon(t_j + s\Delta t) ds \Big),$$

par conséquent, $e_n = e_n^1 + e_n^2$, où

$$e_n^1 = \Delta t \sum_{j=0}^{n-1} E_{n-j-1}(\Delta t \mathcal{B}_\epsilon) \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_\epsilon} g_\epsilon(t_j + s\Delta t) ds,$$

et

$$e_n^2 = \Delta t \sum_{j=0}^{n-1} (I + \Delta t \mathcal{B}_\epsilon)^{-(n-1-j)} \Big(g_\epsilon(t_{j+1}) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_\epsilon} g_\epsilon(t_j + s\Delta t) ds \Big),$$

où E_n est la fonction définie par (2.29). En utilisant le Théorème 7.1 de [71] et le fait que $e^{-(\Delta t - s\Delta t)\mathcal{B}_\epsilon}$ commute avec $E_{n-j-1}(\Delta t \mathcal{B}_\epsilon)$, on obtient

$$\begin{aligned}
\|e_n^1\|_H &\leq \Delta t \sum_{j=0}^{n-1} \int_0^1 \|e^{-(\Delta t - s\Delta t)\mathcal{B}_\epsilon} E_{n-j-1}(\Delta t \mathcal{B}_\epsilon) g_\epsilon(t_j + s\Delta t)\|_H ds \\
&\lesssim \Delta t \sum_{j=0}^{n-1} \int_0^1 \|E_{n-j-1}(\Delta t \mathcal{B}_\epsilon) g_\epsilon(t_j + s\Delta t)\|_H ds \\
&\lesssim (\Delta t)^2 \sum_{j=0}^{n-1} \int_0^1 \|\mathcal{B}_\epsilon g_\epsilon(t_j + s\Delta t)\|_H ds = \Delta t \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|\mathcal{B}_\epsilon g_\epsilon(s)\|_H ds.
\end{aligned}$$

Grâce au Lemme 2.1.1 et l'hypothèse $\mathbf{H_1'}$, on obtient

$$\begin{aligned}
\|\mathcal{B}_\epsilon g_\epsilon(s)\|_H &= \|(I + \epsilon A_1)^{-\frac{1}{2}} A_2 (I + \epsilon A_1)^{-1} f(s)\|_H \\
&\lesssim \|A_2 (I + \epsilon A_1)^{-1} f(s)\|_H \\
&\lesssim \|A_1 (I + \epsilon A_1)^{-1} f(s)\|_H \\
&= \|(I + \epsilon A_1)^{-1} A_1 f(s)\|_H \\
&\lesssim \|A_1 f(s)\|_H.
\end{aligned} \tag{2.60}$$

On conclut que

$$\|e_n^1\|_H \lesssim \Delta t \int_0^{t_n} \|f(s)\|_{D(A_1)} ds. \tag{2.61}$$

On estime maintenant $\|e_n^2\|_H$. Par le développement de Taylor avec reste intégral à l'ordre 1 appliqué à g_ϵ et l'inégalité triangulaire, on a

$$\begin{aligned}\|e_n^2\|_H &\leq \Delta t\sum_{j=0}^{n-1}\|(I+\Delta t\mathcal{B}_\epsilon)^{-(n-1-j)}\Big(g_\epsilon(t_j)-\int_0^1 e^{-(\Delta t-s\Delta t)\mathcal{B}_\epsilon}g_\epsilon(t_j)ds\Big)\|_H\\ &+ \Delta t\sum_{j=0}^{n-1}\|(I+\Delta t\mathcal{B}_\epsilon)^{-(n-1-j)}\Big(\int_{t_j}^{t_{j+1}}g_\epsilon'(\tau)d\tau-\int_0^1 e^{-(\Delta t-s\Delta t)\mathcal{B}_\epsilon}\int_{t_j}^{t_{j+1}}g_\epsilon'(\tau)d\tau ds\Big)\|_H.\end{aligned}$$

Comme l'opérateur $\mathcal{B}_\epsilon$ est auto-adjoint défini positif dans H, $\|e^{-(\Delta t-s\Delta t)\mathcal{B}_\epsilon}\|_{\mathcal{L}(H)}\lesssim 1$ et $\|(I+\Delta t\mathcal{B}_\epsilon)^{-(n-1-j)}\|_{\mathcal{L}(H)}\lesssim 1$, le deuxième terme du membre de droite de la dernière inégalité est majoré par

$$\Delta t\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\|g_\epsilon'(\tau)\|_H d\tau.$$

Par définition de g_ϵ et en utilisant le Lemme 2.1.1, on obtient

$$\Delta t\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\|g_\epsilon'(\tau)\|_H d\tau=\Delta t\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\|(I+\epsilon A_1)^{-\frac{1}{2}}f'(\tau)\|_H d\tau\lesssim\Delta t\int_0^{t_n}\|f'(\tau)\|_H d\tau.$$

Pour le premier terme, on a

$$\begin{aligned}&\Delta t\sum_{j=0}^{n-1}\|(I+\Delta t\mathcal{B}_\epsilon)^{-(n-1-j)}\Big(g_\epsilon(t_j)-\int_0^1 e^{-(\Delta t-s\Delta t)\mathcal{B}_\epsilon}g_\epsilon(t_j)ds\Big)\|_H\\ &\lesssim \Delta t\sum_{j=0}^{n-1}\|g_\epsilon(t_j)-\int_0^1 e^{-(\Delta t-s\Delta t)\mathcal{B}_\epsilon}g_\epsilon(t_j)ds\|_H.\end{aligned}$$

Or, d'après le Lemme 8. 1 de [71], on a

$$\|g_\epsilon(t_j)-\int_0^1 e^{-(\Delta t-s\Delta t)\mathcal{B}_\epsilon}g_\epsilon(t_j)ds\|_H\lesssim\Delta t\|\mathcal{B}_\epsilon g_\epsilon(t_j)\|_H,$$

ce qui permet d'en déduire grâce à (2.60)

$$\begin{aligned}\Delta t\sum_{j=0}^{n-1}\|(I+\Delta t\mathcal{B}_\epsilon)^{-(n-1-j)}\Big(g_\epsilon(t_j)-\int_0^1 e^{-(\Delta t-s\Delta t)\mathcal{B}_\epsilon}g_\epsilon(t_j)ds\Big)\|_H &\lesssim (\Delta t)^2\sum_{j=0}^{n-1}\|f(t_j)\|_{D(A_1)}\\ &\lesssim \Delta t\sup_{s\leq t_n}\|f(s)\|_{D(A_1)}.\end{aligned}$$

Par conséquent

$$\|e_n^2\|_H\lesssim\Delta t\big(\sup_{s\leq t_n}\|f(s)\|_{D(A_1)}+\int_0^{t_n}\|f'(\tau)\|_H d\tau\big).\tag{2.62}$$

On obtient ainsi de (2.61) et (2.62)

$$\|v_n-v(t_n)\|_H\lesssim\Delta t\big(\sup_{s\leq t_n}\|f(s)\|_{D(A_1)}+\int_0^{t_n}\|f'(\tau)\|_H d\tau\big),$$

ce qui prouve par le Lemme 2.1.1 que (2.57) a lieu. ∎

2.2.2 Le problème totalement discrétisé

Dans le même cadre que celui de la sous section 2.1.2, on passe maintenant à la semi-discrétisation en espace. On suppose toujours qu'il existe un sous espace de dimension finie V_h de V et on cherche à construire une approximation totalement discrète du problème (2.54). Donc on considère le problème discret suivant

$$\begin{cases} (I+\epsilon A_{1,h})u_{h,t}+A_{2,h}u_h=f_h(t), & \text{dans } V_h', \quad 0<t\leq T,\\ u_h(0)=0, & \text{dans } V_h, \end{cases} \tag{2.63}$$

où $A_{1,h}$, $A_{2,h}$ sont les approximations de A_1, A_2 dans V_h respectivement définis dans la section 2.1.2 par (2.15), et $f_h(t)=I_hf(t)$ où $I_h: V'\to V_h'$ est l'opérateur linéaire continu défini par

$$\langle f_h(t),v_h\rangle_{V_h',V_h}=\langle f(t),v_h\rangle_{V',V},\quad \forall v_h\in V_h.$$

Remarque 2.2.2. Comme dans la Proposition 2.2.1, on aura toujours besoin de l'hypothèse $f\in C([0,T],D(A_1))$ pour établir des estimations d'erreur. Ainsi la fonction f est aussi à valeurs dans V et H puisque $D(A_1)\hookrightarrow V\hookrightarrow H$. Comme conséquence, $f_h(t)=P_hf(t)$ pour tout $t\in[0,T]$. En effet, on sait que $P_hf(t)\in V_h$ qui s'identifie à un sous-espace de son dual V_h'.
D'après ce qui précède, on a alors

$$\begin{aligned}\langle P_hf(t),v_h\rangle_{V_h',V_h}=(P_hf(t),v_h)_H &= (f(t),v_h)_H\\ &= \langle f(t),v_h\rangle_{V',V}.\end{aligned}$$

Comme dans la section 2.1, le problème (2.63) est équivalent à

$$\begin{cases} u_{h,t}+\mathcal{A}_{\epsilon,h}u_h=f_{\epsilon,h}(t), & \text{dans } V_h, \quad 0<t\leq T,\ \epsilon>0,\\ u_h(0)=0, & \text{dans } V_h, \end{cases} \tag{2.64}$$

où $\mathcal{A}_{\epsilon,h}=(I+\epsilon A_{1,h})^{-1}A_{2,h}$ et $f_{\epsilon,h}(t)=(I+\epsilon A_{1,h})^{-1}f_h(t)$.

Par le schéma d'Euler implicite et le schéma de Crank-Nicolson, la solution totalement discrète du problème (2.64) (ou (2.63)) est donnée par les formules suivantes respectivement

$$(I+\Delta t\mathcal{A}_{\epsilon,h})u_{n+1,h}=u_n+f_{\epsilon,h}(t_{n+1}), \tag{2.65}$$

$$(1+\frac{1}{2}\Delta t\mathcal{A}_{\epsilon,h})u_{n+1,h}=(1-\frac{1}{2}\Delta t\mathcal{A}_{\epsilon,h})u_{n,h}+\frac{1}{2}\big(f_{\epsilon,h}(t_{n+1,h})+f_{\epsilon,h}(t_n)\big), \tag{2.66}$$

où $u_{0,h}=0$ et $0\leq n\leq N-1$.

Proposition 2.2.3. *Sous les hypothèses* $\mathbf{H_1'}$ *à* $\mathbf{H_5'}$, *on suppose que* $f\in C^1([0,T],H)\cap C([0,T],D(A_1))$. *Soit* u_h *la solution exacte du problème* (2.64) *et soit* $u_{n,h}$ *la solution totalement discrète donnée par le schéma* (2.65). *Alors, on a l'estimation d'erreur suivante*

$$\|u_h(t_n)-u_{n,h}\|_H\lesssim(\epsilon+\Delta t+h^2)\sup_{s\in[0,T]}\|f(s)\|_{D(A_1)}+\Delta t\int_0^T\|f'(\tau)\|_Hd\tau,$$

pour tout $1 \leq n \leq N$.

Preuve : En utilisant le changement d'inconnue (2.25), le problème (2.63) est équivalent au problème semi-discret du problème (2.58) suivant

$$\begin{cases} v_{h,t} + \mathcal{B}_{\epsilon,h} v_h = g_{\epsilon,h}(t), & \text{dans } V_h, \quad 0 < t \leq T, \\ v_h(0) = 0, & \text{dans } V_h, \end{cases} \tag{2.67}$$

où $\mathcal{B}_{\epsilon,h} = (I + \epsilon A_{1,h})^{-\frac{1}{2}} A_{2,h} (I + \epsilon A_{1,h})^{-\frac{1}{2}}$ et $g_{\epsilon,h}(t) = (I + \epsilon A_{1,h})^{-\frac{1}{2}} f_h(t)$.
On sait par la formule de variation des constantes que la solution de (2.67) au point t_{n+1} s'écrit comme suit

$$v_h(t_{n+1}) = e^{-\Delta t \mathcal{B}_{\epsilon,h}} v(t_n) + \int_0^{\Delta t} e^{-(\Delta t - \tau)\mathcal{B}_{\epsilon,h}} g_{\epsilon,h}(t_n + \tau) d\tau, \quad \forall\, 0 \leq n \leq N-1.$$

Par le schéma d'Euler implicite, la solution totalement discrète du problème (2.67) est donnée par

$$v_{n+1,h} = (I + \Delta t \mathcal{B}_{\epsilon,h})^{-1} v_{n,h} + \Delta t g_{\epsilon,h}(t_{n+1}), \quad \forall\, 0 \leq n \leq N-1. \tag{2.68}$$

Donc par récurrence l'erreur $e_{n,h} = v_{n,h} - v(t_n)$ satisfait $e_{n,h} = e^1_{n,h} + e^2_{n,h}$, où

$$e^1_{n,h} = \Delta t \sum_{j=0}^{n-1} E_{n-j-1}(\Delta t \mathcal{B}_{\epsilon,h}) \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} g_{\epsilon,h}(t_j + s\Delta t) ds,$$

et

$$e^2_{n,h} = \Delta t \sum_{j=0}^{n-1} (I + \Delta t \mathcal{B}_{\epsilon,h})^{-(n-1-j)} \left(g_{\epsilon,h}(t_{j+1}) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} g_{\epsilon,h}(t_j + s\Delta t) ds \right).$$

Par conséquent et dans le but d'estimer $\|e_{n,h}\|_H$, on estime les deux termes $\|e^1_{n,h}\|_H$ et $\|e^2_{n,h}\|_H$. Comme $e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}}$ commute avec $E_{n-j-1}(\Delta t \mathcal{B}_{\epsilon,h})$ et en utilisant le fait que $\|e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}}\|_{\mathcal{L}(H)} \lesssim 1$, car l'opérateur $\mathcal{B}_{\epsilon,h}$ est auto-adjoint dans H, on obtient

$$\begin{aligned} \|e^1_{n,h}\|_H &\leq \Delta t \sum_{j=0}^{n-1} \|E_{n-j-1}(\Delta t \mathcal{B}_{\epsilon,h}) \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} g_{\epsilon,h}(t_j + s\Delta t) ds\|_H \\ &\lesssim \Delta t \sum_{j=0}^{n-1} \int_0^1 \|E_{n-j-1}(\Delta t \mathcal{B}_{\epsilon,h}) g_{\epsilon,h}(t_j + s\Delta t)\|_H ds \\ &= \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|E_{n-1-j}(\Delta t \mathcal{B}_{\epsilon,h}) g_{\epsilon,h}(s)\|_H ds. \end{aligned}$$

Mais, par l'inégalité triangulaire

$$\begin{aligned} \|E_n(\Delta t \mathcal{B}_{\epsilon,h}) g_{\epsilon,h}(s)\|_H \leq \|E_n(\Delta t \mathcal{B}_{\epsilon,h}) P_h g_\epsilon(s)\|_H &+ \|E_n(\Delta t \mathcal{B}_{\epsilon,h}) \big(g_{\epsilon,h}(s) - P_h f(s)\big)\|_H \\ &+ \|E_n(\Delta t \mathcal{B}_{\epsilon,h}) \big(P_h f(s) - P_h g_\epsilon(s)\big)\|_H. \end{aligned}$$

En utilisant, comme dans la preuve de la Proposition 2.1.10, l'égalité (2.37) avec $v = g_\epsilon(s), s \in [0,T]$ et $q = 1$, et par application de (2.38) et du Lemme 2.1.8, on obtient

$$\begin{aligned}\|E_n(\Delta t\mathcal{B}_{\epsilon,h})P_h g_\epsilon(s)\|_H &\lesssim (\epsilon + h^2 + \Delta t)\|\mathcal{B}_\epsilon g_\epsilon(s)\|_H \\ &\lesssim (\epsilon + h^2 + \Delta t)\|f(s)\|_{D(A_1)},\end{aligned}$$

la dernière inégalité est due à (2.60). Pour les autres termes, en utilisant (2.38) avec $j = 0$ et la définition de g_ϵ, $g_{\epsilon,h}$, on obtient

$$\begin{aligned}\|E_n(\Delta t\mathcal{B}_{\epsilon,h})\big(g_{\epsilon,h}(s) &- P_h f(s)\big)\|_H + \|E_n(\Delta t\mathcal{B}_{\epsilon,h})\big(P_h f(s) - P_h g_\epsilon(s)\big)\|_H \\ &\lesssim \|g_{\epsilon,h}(s) - P_h f(s)\|_H + \|f(s) - g_\epsilon(s)\|_H \\ &= \|(I + \epsilon A_{1,h})^{-\frac{1}{2}} P_h f(s) - P_h f(s)\|_H + \|f(s) - (I + \epsilon A_1)^{-\frac{1}{2}} f(s)\|_H.\end{aligned} \tag{2.69}$$

Toujours par la formule de Parseval et comme $f(s) \in D(A_1)$, pour tout $s \in [0,T]$, on a

$$\begin{aligned}\|f(s) - (I + \epsilon A_1)^{-\frac{1}{2}} f(s)\|_H^2 &= \sum_{l=1}^{\infty}(1 - (1 + \epsilon\nu_l)^{-\frac{1}{2}})^2 |(f(s), \psi_l)_H|^2 \\ &\lesssim \sum_{l=1}^{\infty} \frac{\epsilon^2\nu_l^2}{(1 + \epsilon\nu_l)^{1/2} + (1 + \epsilon\nu_l)} |(f(s), \psi_l)_H|^2 \\ &\lesssim \epsilon^2 \sum_{l=1}^{\infty} \nu_l^2 |(f(s), \psi_l)_H|^2 \\ &\lesssim \epsilon^2 \|f(s)\|_{D(A_1)}^2.\end{aligned} \tag{2.70}$$

Ainsi que, pour tout $y_h \in V_h$, on a

$$\begin{aligned}\|y_h - (I + \epsilon A_{1,h})^{-\frac{1}{2}} y_h\|_H^2 &= \sum_{l=1}^{N_h}(1 - (1 + \epsilon\nu_{l,h})^{-\frac{1}{2}})^2 |(y_h, \psi_{l,h})_H|^2 \\ &\lesssim \sum_{l=1}^{N_h} \frac{\epsilon^2\nu_{l,h}^2}{(1 + \epsilon\nu_{l,h})^{\frac{1}{2}} + (1 + \epsilon\nu_{l,h})} |(y_h, \psi_{l,h})_H|^2 \\ &\lesssim \epsilon^2 \sum_{l=1}^{N_h} \nu_{l,h}^2 |(y_h, \psi_{l,h})_H|^2 \\ &= \epsilon^2 \|A_{1,h} y_h\|_H^2.\end{aligned} \tag{2.71}$$

Par conséquent, en utilisant (2.41) avec $f(s)$ au lieu de u_0, obtient

$$\|(I + \epsilon A_{1,h})^{-1/2} P_h f(s) - P_h f(s)\|_H \lesssim \epsilon \|A_{1,h} P_h f(s)\|_H \lesssim \epsilon \|f(s)\|_{D(A_1)}.$$

Finalement, on a

$$\|e_{n,h}^1\| \lesssim (\epsilon + h^2 + \Delta t) \int_0^{t_n} \|f(s)\|_{D(A_1)} ds.$$

Nous allons maintenant estimer $\|e^2_{n,h}\|_H$. Pour cela, en utilisant le développement de Taylor avec reste intégral à l'ordre 1 de $g_{\epsilon,h}$, on obtient

$$\|e^2_{n,h}\|_H \leq I_1 + I_2,$$

où

$$I_1 = \Delta t \sum_{j=0}^{n-1} \|(I + \Delta t \mathcal{B}_{\epsilon,h})^{-(n-1-j)} \big(g_{\epsilon,h}(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} g_{\epsilon,h}(t_j) ds\big)\|_H,$$

et

$$I_2 = \Delta t \sum_{j=0}^{n-1} \|(I + \Delta t \mathcal{B}_{\epsilon,h})^{-(n-1-j)} \big(\int_{t_j}^{t_{j+1}} g'_{\epsilon,h}(\tau) d\tau - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} \int_{t_j}^{t_{j+1}} g'_{\epsilon,h}(\tau) d\tau ds\big)\|_H.$$

En utilisant le fait que $\|e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}}\|_{\mathcal{L}(H)} \lesssim 1$, $\|(I + \Delta t \mathcal{B}_{\epsilon,h})^{-(n-1-j)}\|_{\mathcal{L}(H)} \lesssim 1$ (car l'opérateur $\mathcal{B}_{\epsilon,h}$ est auto-adjoint défini positif dans H), le Lemme 2.1.9 et le fait que $\|P_h\|_{\mathcal{L}(H)} \lesssim 1$, on trouve

$$\begin{aligned} I_2 &\lesssim \Delta t \int_0^{t_n} \|g'_{\epsilon,h}(\tau)\|_H d\tau. \\ &= \Delta t \int_0^{t_n} \|(I + \epsilon A_{1,h})^{-\frac{1}{2}} P_h f'(\tau)\|_H d\tau \\ &\lesssim \Delta t \int_0^{t_n} \|P_h f'(\tau)\|_H d\tau \\ &\lesssim \Delta t \int_0^{t_n} \|f'(\tau)\|_H d\tau. \end{aligned}$$

Pour I_1, on a

$$\begin{aligned} I_1 &\lesssim \Delta t \sum_{j=0}^{n-1} \|g_{\epsilon,h}(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} g_{\epsilon,h}(t_j) ds\|_H \\ &\leq \Delta t \sum_{j=0}^{n-1} \|g_{\epsilon,h}(t_j) - P_h g_\epsilon(t_j) + \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} (P_h g_\epsilon(t_j) - g_{\epsilon,h}(t_j)) ds\|_H \\ &\quad + \Delta t \sum_{j=0}^{n-1} \|P_h g_\epsilon(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} P_h g_\epsilon(t_j) ds\|_H. \end{aligned}$$

Toujours comme $\|e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}}\|_{\mathcal{L}(H)} \lesssim 1$, le premier terme du membre de droite de cette dernière inégalité est majoré par (voir la majoration de(2.69))

$$\begin{aligned} \Delta t \sum_{j=0}^{n-1} \|g_{\epsilon,h}(t_j) - P_h g_\epsilon(t_j)\|_H &\leq \Delta t \sum_{j=0}^{n-1} (\|g_{\epsilon,h}(t_j) - P_h f(t_j)\|_H + \|P_h f(t_j) - P_h g_\epsilon(t_j)\|_H) \\ &\lesssim \epsilon \Delta t \sum_{j=0}^{n-1} \|f(t_j)\|_{D(A_1)} \\ &= \epsilon \sup_{s\in[0,T]} \|f(s)\|_{D(A_1)}. \end{aligned}$$

Pour l'autre terme, en utilisant l'égalité (2.37) avec $v = g_\epsilon(t_j)$, on obtient

$$\begin{aligned}\Delta t \sum_{j=0}^{n-1} \|P_h g_\epsilon(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} P_h g_\epsilon(t_j) ds\|_H \\ \leq \Delta t \sum_{j=0}^{n-1} \|P_h(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h)\mathcal{B}_\epsilon g_\epsilon(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} P_h(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h)\mathcal{B}_\epsilon g_\epsilon(t_j) ds\|_H \\ + \Delta t \sum_{j=0}^{n-1} \|\mathcal{B}_{\epsilon,h}^{-1} P_h \mathcal{B}_\epsilon g_\epsilon(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} \mathcal{B}_{\epsilon,h}^{-1} P_h \mathcal{B}_\epsilon g_\epsilon(t_j) ds\|_H.\end{aligned}$$

D'après le Lemme 2.1.8 et (2.60), on a

$$\begin{aligned}\Delta t \sum_{j=0}^{n-1} \|P_h(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h)\mathcal{B}_\epsilon g_\epsilon(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} P_h(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h) g_\epsilon(t_j) ds\|_H \\ \lesssim \Delta t \sum_{j=0}^{n-1} \|(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h)\mathcal{B}_\epsilon g_\epsilon(t_j)\|_H \\ \lesssim (\epsilon + h^2) \sup_{s\in[0,T]} \|f(s)\|_{D(A_1)}.\end{aligned}$$

Et d'après le Lemme 8.1 de [71] et encore une fois de (2.60), on a

$$\begin{aligned}\Delta t \sum_{j=0}^{n-1} \|\mathcal{B}_{\epsilon,h}^{-1} P_h \mathcal{B}_\epsilon g_\epsilon(t_j) - \int_0^1 e^{-(\Delta t - s\Delta t)\mathcal{B}_{\epsilon,h}} \mathcal{B}_{\epsilon,h}^{-1} P_h \mathcal{B}_\epsilon g_\epsilon(t_j) ds\|_H \\ \lesssim (\Delta t)^2 \sum_{j=0}^{n-1} \|\mathcal{B}_{\epsilon,h}\mathcal{B}_{\epsilon,h}^{-1} P_h \mathcal{B}_\epsilon g_\epsilon(t_j)\|_H \\ \lesssim \Delta t \sup_{s\in[0,T]} \|f(s)\|_{D(A_1)}.\end{aligned}$$

Par conséquent

$$I_1 \leq (\epsilon + \Delta t + h^2) \sup_{s\in[0,T]} \|f(s)\|_{D(A_1)},$$

ce qui donne

$$\|e_{n,h}^2\|_H \lesssim (\epsilon + \Delta t + h^2) \sup_{s\in[0,T]} \|f(s)\|_{D(A_1)} + \Delta t \int_0^{t_n} \|f'(\tau)\|_H d\tau.$$

On en déduit que

$$\begin{aligned}\|v_{n,h} - v_h(t_n)\|_H \lesssim (\epsilon + \Delta t + h^2)\Big(\sup_{s\in[0,T]} \|f(s)\|_{D(A_1)} + \int_0^T \|f(\tau)\|_{D(A_1)} d\tau\Big) \\ + \Delta t \int_0^T \|f'(\tau)\|_H d\tau. \qquad (2.72)\end{aligned}$$

A la fin et pour conclure le résultat principal de ce théorème il suffit d'utiliser le fait que $u_h = (I + \epsilon A_{1,h})^{-\frac{1}{2}} v_h$, $u_{n,h} = (I + \epsilon A_{1,h})^{-\frac{1}{2}} v_{n,h}$, $0 \leq n \leq N$, et le Lemme 2.1.9. ∎

Remarque 2.2.4. *De la même façon et sous les hypothèses* $\mathbf{H_1'}$ *à* $\mathbf{H_5'}$*, si on suppose que* $f \in C([0,T], D(A_2^2)) \cap C^2([0,T], H)$ *et que* $f \in C^1([0,T], D(A_1))$*,* u_h *est la solution exacte du problème* (2.64) *et* $u_{n,h}$ *est la solution donnée par le schéma de Crank-Nicolson* (2.66)*, on peut démontrer que*

$$\|u_h(t_n)-u_{n,h}\|_H \lesssim (\epsilon+h^2+(\Delta t)^2)\Big(\sup_{s\in[0,T]} \|f(s)\|_{D(A_2^2)} + \sup_{s\in[0,T]} \|f'(s)\|_{D(A_1)} + \int_0^T \|f''(s)\|_H ds \Big),$$

pour tout $1 \leq n \leq N$.

Dans la proposition suivante, on montre une estimation d'erreur entre la solution exacte du problème (2.54) et la solution du problème semi-discrétisé en espace.

Proposition 2.2.5. *Sous les hypothèses* $\mathbf{H_1'}$-$\mathbf{H_4'}$*, on suppose que* $f \in C([0,T], D(A_1)) \cap C^2([0,T], H)$*. Soit* v *(resp.* v_h*) la solution du problème* (2.58) *(resp.* (2.67)*). Alors*

$$\|v_h(t_n)-v(t_n)\|_H \lesssim (\epsilon+h^2)\Big(\sup_{s\in[0,T]} \|f(s)\|_{D(A_1)} + \int_0^T \big(\|f(s)\|_H+\|f'(s)\|_H+\|f''(s)\|_H\big) ds \Big).$$

Preuve : On a l'erreur $e_h = v_h - v$ satisfait

$$\begin{aligned} \mathcal{B}_{\epsilon,h}^{-1} P_h e_{h,t} + e_h &= \mathcal{B}_{\epsilon,h}^{-1} v_{h,t} + v_h - \mathcal{B}_{\epsilon,h}^{-1} v_{h,t} - v \\ &= \mathcal{B}_{\epsilon,h}^{-1}(g_{\epsilon,h} - P_h g_\epsilon) + \rho, \end{aligned}$$

où $\rho = -(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1} P_h)\mathcal{B}_\epsilon v$. Alors, pour tout $s \in [0,T]$ et tout $\varphi \in H$

$$\big(\mathcal{B}_{\epsilon,h}^{-1}(P_h e_{h,t}(s) + P_h g_\epsilon(s) - g_{\epsilon,h}(s)), \varphi\big)_H + \big(e_h(s), \varphi\big)_H = \big(\rho(s), \varphi\big)_H,$$

en choisissant dans cette égalité $\varphi = 2(e_{h,t} + g_\epsilon - g_{\epsilon,h})(s)$, on obtient

$$\begin{aligned} 2\big(\mathcal{B}_{\epsilon,h}^{-1} P_h(e_{h,t} + g_\epsilon - g_{\epsilon,h})(s), (e_{h,t} + g_\epsilon &- g_{\epsilon,h})(s)\big)_H + 2\big(e_h(s), (e_{h,t} + g_\epsilon - g_{\epsilon,h})(s)\big)_H \\ &= 2\big(\rho(s), (e_{h,t} + g_\epsilon - g_{\epsilon,h})(s)\big)_H. \end{aligned}$$

En intégrant par rapport à s et en utilisant le fait que $\mathcal{B}_{\epsilon,h}$ est auto-adjoint défini positif, on obtient

$$\begin{aligned} \|e_h(t)\|_H^2 \leq 2(\rho(t), e_h(t))_H + 2\int_0^t (-\rho_t(s), e_h(s))_H ds &+ 2\int_0^t (\rho(s), g_\epsilon(s) - g_{\epsilon,h}(s))_H ds \\ &+ 2\int_0^t (e_h(s), g_{\epsilon,h}(s) - g_\epsilon(s))_H, \end{aligned}$$

ce qui donne d'après l'inégalité de Cauchy-Schwarz

$$\begin{aligned}\|e_h(t)\|_H^2 &\leq 2\|\rho(t)\|_H\|e_h(t)\|_H + 2\int_0^t \|\rho_t(s)\|_H\|e_h(s)\|_H ds \\ &\quad +2\int_0^t \|e_h(s)\|_H\|g_{\epsilon,h}(s)-g_\epsilon(s)\|_H ds + 2\int_0^t \|\rho(s)\|_H\|g_{\epsilon,h}(s)-g_\epsilon(s)\|_H ds \\ &\leq \sup_{s\leq t}\|e_h(s)\|_H\big(2\|\rho(t)\|_H + 2\int_0^t \|\rho_t(s)\|_H ds + 2\int_0^t \|g_{\epsilon,h}(s)-g_\epsilon(s)\|_H ds\big) \\ &\quad +2\int_0^t \|\rho(s)\|_H\|g_{\epsilon,h}(s)-g_\epsilon(s)\|_H ds.\end{aligned}$$

En tenant compte du fait que $\|\rho(t)\|_H \lesssim \int_0^t \|\rho_t(s)\|_H ds$, on obtient

$$\begin{aligned}\|e_h(t)\|_H^2 &\leq \sup_{s\leq t}\|e_h(s)\|_H(4\int_0^t \|\rho_t(s)\|_H ds + 2\int_0^t \|g_{\epsilon,h}(s)-g_\epsilon(s)\|_H ds\big) \\ &\quad +2\int_0^t \|\rho(s)\|_H\|g_{\epsilon,h}(s)-g_\epsilon(s)\|_H ds.\end{aligned}$$

D'une part, pour tout $s\in[0,T]$, on a grâce à (2.71), l'hypothèse $\mathbf{H_4'}$ et (2.70)

$$\begin{aligned}\|g_{\epsilon,h}(s)-g_\epsilon(s)\|_H &= \|(I+\epsilon A_{1,h})^{-1/2}P_h f(s) - (I+\epsilon A_1)^{-1/2}f(s)\|_H \\ &\leq \|(I+\epsilon A_{1,h})^{-1/2}P_h f(s) - P_h f(s)\|_H + \|P_h f(s) - f(s)\|_H \\ &\quad +\|f(s)-(I+\epsilon A_1)^{-1/2}f(s)\|_H \\ &\lesssim (\epsilon+h^2)\|f(s)\|_{D(A_1)},\end{aligned}$$

et par conséquent

$$\begin{aligned}\|e_h(t)\|_H^2 &\lesssim \|e_h(\tau)\|_H(\int_0^t \|\rho_t(s)\|_H ds + (\epsilon+h^2)\int_0^t \|f(s)\|_{D(A_1)} ds\big) \\ &\quad +(\epsilon+h^2)\sup_{s\in[0,T]}\|f(s)\|_{D(A_1)}\int_0^t \|\rho(s)\|_H ds,\end{aligned}$$

où $\|e_h(\tau)\|_H = \sup_{s\in[0,T]}\|e_h(s)\|_H$.
D'autre part, on a d'après le Lemme 2.1.8

$$\int_0^t \|\rho(s)\|_H ds = \int_0^t \|(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1}P_h)\mathcal{B}_\epsilon v(s)\|_H ds \lesssim (\epsilon+h^2)\int_0^t \|\mathcal{B}_\epsilon v(s)\|_H ds,$$

et

$$\int_0^t \|\rho_t(s)\|_H ds = \int_0^t \|(\mathcal{B}_\epsilon^{-1} - \mathcal{B}_{\epsilon,h}^{-1}P_h)\mathcal{B}_\epsilon v_t(s)\|_H ds \lesssim (\epsilon+h^2)\int_0^t \|\mathcal{B}_\epsilon v_t(s)\|_H ds.$$

Mais comme v est la solution du problème (2.58), on a

$$(v_t(s),\mathcal{B}_\epsilon v(s))_H + \|\mathcal{B}_\epsilon v(s)\|_H^2 = (g_\epsilon(s),\mathcal{B}_\epsilon v(s))_H,$$

alors grâce à l'inégalité de Cauchy-Schwarz, on obtient

$$\int_0^t \|\mathcal{B}_\epsilon v(s)\|_H ds \leq \int_0^t \|g_\epsilon(s)\|_H ds + \int_0^t \|v_t(s)\|_H ds.$$

On pose maintenant $w = v_t$, donc w vérifie l'équation suivante

$$\begin{cases} w_t + \mathcal{B}_\epsilon w = g_\epsilon', & \text{dans } H, \quad 0 < t \leq T, \\ w(0) = g_\epsilon(0), & \text{dans } H, \end{cases} \tag{2.73}$$

et comme l'opérateur $\mathcal{B}_\epsilon$ est auto-adjoint défini positif, on trouve

$$\begin{aligned} \|v_t\|_H = \|w(t)\|_H &\lesssim \|w(0)\|_H + \int_0^t \|g_\epsilon'(\tau)\|_H d\tau \\ &= \|g_\epsilon(0)\|_H + \int_0^t \|g_\epsilon'(\tau)\|_H d\tau. \end{aligned}$$

On conclut que, si v est la solution du problème (2.58), alors

$$\begin{aligned} \int_0^t \|\mathcal{B}_\epsilon v(s)\|_H ds &\lesssim \int_0^t \|g_\epsilon(s)\|_H ds + \|g_\epsilon(0)\|_H + \int_0^t \|g_\epsilon'(\tau)\|_H d\tau \\ &\lesssim \int_0^t \|f(s)\|_H ds + \|f(0)\|_H + \int_0^t \|f'(\tau)\|_H d\tau, \end{aligned}$$

en vertu du Lemme 2.1.1. En appliquant cette inégalité à w, la solution de problème (2.73), on obtient

$$\begin{aligned} \int_0^t \|\mathcal{B}_\epsilon v_t(s)\|_H ds &= \int_0^t \|\mathcal{B}_\epsilon w(s)\|_H ds \\ &\lesssim \int_0^t \|g_\epsilon'(s)\|_H ds + \|g_\epsilon'(0)\|_H + \int_0^t \|g''(\tau)\|_H d\tau \\ &\lesssim \int_0^t \|f'(s)\|_H ds + \|f'(0)\|_H + \int_0^t \|f''(\tau)\|_H d\tau. \end{aligned}$$

Par conséquent

$$\begin{aligned} \|e_h(t)\|_H^2 \lesssim\ & (\epsilon + h^2)\|e_h(\tau)\|_H \Big(\|f'(0)\|_H + \int_0^t \big(\|f(s)\|_{D(A_1)} + \|f'(s)\|_H + \|f''(s)\|_H \big) ds \Big) \\ & + (\epsilon + h^2)^2 \sup_{s\in[0,T]} \|f(s)\|_{D(A_1)} \Big(\|f(0)\|_H + \int_0^t \big(\|f(s)\|_H + \|f'(s)\|_H \big) ds \Big). \end{aligned}$$

On en déduit par l'inégalité $2ab \leq a^2 + b^2$ que

$$\begin{aligned} \|e_h(\tau)\|_H^2 \lesssim\ & (\epsilon + h^2)^2 \Bigg(\Big(\|f'(0)\|_H + \int_0^t \big(\|f(s)\|_{D(A_1)} + \|f'(s)\|_H + \|f''(s)\|_H \big) ds \Big)^2 \\ & + \Big(\sup_{s\in[0,T]} \|f(s)\|_{D(A_1)} \Big)^2 + \Big(\|f(0)\|_H + \int_0^t \big(\|f(s)\|_H + \|f'(s)\|_H \big) ds \Big)^2 \Bigg), \end{aligned}$$

d'où le résultat. ∎

Corollaire 2.2.6. *Sous les hypothèses* $\mathbf{H_1'}$ *à* $\mathbf{H_5'}$, *on suppose que* $f \in C([0,T], D(A_1)) \cap C^2([0,T], H)$. *Soit* v *la solution exacte du problème* (2.67) *et* $v_{n,h}$ *la solution totalement discrète donnée par le schéma d'Euler implicite* (2.68). *Alors on a l'estimation globale suivante*

$$\|v_h(t_n) - v(t_n)\|_H \lesssim (\epsilon + \Delta t + h^2)\left(\sup_{s\in[0,T]} \|f(s)\|_{D(A_1)} + \|f\|_{W^{2,1}([0,T],H)}\right).$$

Par conséquent, si u *est la solution exacte du problème* (2.54) *et* $u_{n,h}$ *la solution totalement discrète donnée par le schéma d'Euler implicite* (2.55), *on a*

$$\|u(t_n) - u_{n,h}\|_* \lesssim (\epsilon + \Delta t + h^2)\left(\sup_{s\in[0,T]} \|f(s)\|_{D(A_1)} + \|f\|_{W^{2,1}([0,T],H)}\right).$$

Preuve : La première estimation est une conséquence directe de l'inégalité (2.72) et de la Proposition 2.2.5. La deuxième estimation se démontre de manière similaire comme pour le cas homogène (Corollaire 2.1.13). ∎

Chapitre 3

Applications et résultats numériques

Résumé

On se propose dans cette partie, de détailler des applications et de présenter des illustrations numériques des problèmes modèles étudiés dans les chapitres précédents.

Sommaire

3.1 Applications pour des équations semi-linéaires particulières **69**
3.1.1 Opérateurs elliptiques d'ordre deux : le cas régulier 69
3.1.2 Opérateurs elliptiques d'ordre deux : le cas singulier 74
3.1.3 Opérateurs elliptiques d'ordre deux à coefficients discontinus : le cas singulier 77
3.1.4 Résultats numériques 82
3.2 Applications pour des équations quasi-linéaires de Sobolev **92**
3.2.1 Le cas d'une équation non autonome 92
3.2.2 Le cas d'une équation quasi-linéaire 95
3.3 Applications pour une équation de Sobolev singulièrement perturbée . **105**

3.1 Applications pour des équations semi-linéaires particulières

3.1.1 Opérateurs elliptiques d'ordre deux : le cas régulier

Soit Ω un domaine borné de $\mathbb{R}^d, d \geq 1$ à frontière lipschitzienne. Pour $i = 1, 2$, soit L_i l'opérateur elliptique d'ordre deux de la forme

$$L_i(x, D_x)u = -\sum_{k,\ell=1}^{d} \partial_k(a_{k,\ell}^{(i)}(x)\partial_\ell u) + \sum_{k=1}^{d} b_k^{(i)}(x)\partial_k u + c^{(i)}(x)u, \tag{3.1}$$

où $a_{k,\ell}^{(i)} = a_{\ell,k}^{(i)} \in C^{0,1}(\bar{\Omega})$, $b_k^{(i)}, c^{(i)} \in L^\infty(\Omega)$. On suppose de plus qu'il existe $\alpha_* > 0$ tel que

$$\sum_{k,\ell=1}^{d} a_{k,\ell}^{(1)}(x)\xi_\ell\xi_k \geq \alpha_*|\xi|^2, \forall \xi \in \mathbb{R}^d.$$

Dans ce cas, on introduit les formes bilinéaires continues a_i sur $H_0^1(\Omega) \times H_0^1(\Omega)$

$$a_i(u,v) = \int_\Omega \left(\sum_{k,\ell=1}^{d} a_{k,\ell}^{(i)}(x)\, \partial_\ell u\, \partial_k v + \sum_{k=1}^{d} b_k^{(i)}(x)\partial_k u\, v + c^{(i)}(x) u\, v \right) dx, \forall u, v \in H_0^1(\Omega).$$

Ainsi leurs opérateurs associés A_i

$$\langle A_i u, v\rangle_{V',V} = a_i(u,v), \forall u, v \in H_0^1(\Omega),$$

sont continus de $H_0^1(\Omega)$ dans son dual $H^{-1}(\Omega)$. Ces opérateurs satisfont les hypothèses $\mathbf{H_1}$ et $\mathbf{H_2}$ du Chapitre 1 avec $V = H_0^1(\Omega)$ et $H = L^2(\Omega)$, si on suppose que a_1 est coercive sur $H_0^1(\Omega)$ (qui est le cas si div $\mathbf{b}^{(1)} = \sum_{k=1}^{d} \partial_k b_k^{(1)} = 0$ et $c^{(1)} \geq 0$ par exemple).

Par conséquent le problème

$$\begin{cases} L_1 u_t + L_2 u = f(t,u), & \text{dans } \Omega \times (0,T), \\ u = 0, & \text{sur } \partial\Omega \times (0,T), \\ u(0) = u_0, & \text{dans } H_0^1(\Omega), \end{cases} \tag{3.2}$$

est bien posé pour une donnée initiale $u_0 \in H_0^1(\Omega)$ et f continûment différentiable de $[0,T] \times H_0^1(\Omega)$ dans $H^{-1}(\Omega)$. Ce système est une équation semi-linéaire de Sobolev sur Ω qui a été analysée dans [4, 30, 54, 57, 58, 66] dans certaines situations particulières avec des conditions aux limites de type Neumann. Ces conditions aux limites entrent aussi dans notre cadre en remplaçant simplement $H_0^1(\Omega)$ par $H^1(\Omega)$ (et en supposant que a_1 est coercive dans $H^1(\Omega)$).

Afin de vérifier les hypothèses $\mathbf{H_3}$ et $\mathbf{H_4}$, nous allons caractériser le domaine de A_1 (comme un opérateur non borné dans H) et de $A_1^{\frac{3}{2}}$ dans certaines situations particulières. Dans le premier cas, nous utilisons le résultat de Kadlec.

Lemme 3.1.1. *Sous les hypothèses précédentes sur les coefficients de $L_i, i = 1, 2$, et si Ω est convexe ou à bord de classe $C^{1,1}$, on a*

$$D(A_1) = H^2(\Omega) \cap H_0^1(\Omega) \hookrightarrow D(A_2). \tag{3.3}$$

Preuve : Pour $i = 1$, ou 2, rappelons que

$$D(A_i) = \{u \in H_0^1(\Omega) : A_i u \in L^2(\Omega)\}.$$

Donc $u \in H_0^1(\Omega)$ est dans $D(A_1)$ si et seulement si il existe $f \in L^2(\Omega)$ tel que

$$a_1(u,v) = \int_\Omega fv\,dx, \forall v \in H_0^1(\Omega). \tag{3.4}$$

Ceci est équivalent à

$$\int_\Omega \sum_{k,\ell=1}^{d} a_{k,\ell}^{(1)}(x)\partial_\ell u \partial_k v\,dx = \int_\Omega hv\,dx, \forall v \in H_0^1(\Omega),$$

où $h = f - \sum_{k=1}^{d} b_k^{(1)}\partial_k u - c^{(1)}u$ appartient à $L^2(\Omega)$ grâce aux hypothèses sur les coefficients et le fait que $u \in H_0^1(\Omega)$. Alors grâce au résultat de Kadlec [42] (voir aussi [34, Théorème 3.2.1.3]), on conclut que $u \in H^2(\Omega)$. Ceci prouve l'inclusion

$$D(A_1) \hookrightarrow H^2(\Omega) \cap H_0^1(\Omega).$$

L'inclusion inverse étant triviale, on a montré que

$$\|u\|_{D(A_1)} \sim \|u\|_{H^2(\Omega)}, \forall u \in D(A_1).$$

Il est clair que $D(A_1) = H^2(\Omega) \cap H_0^1(\Omega) \subset D(A_2)$, et donc pour $u \in D(A_1)$, $L_2 u$ appartient à $L^2(\Omega)$ avec

$$\|u\|_{D(A_2)} = \|u\|_{H_0^1(\Omega)} + \|L_2 u\|_{L^2(\Omega)} \lesssim \|u\|_{H^2(\Omega)},$$

ce qui prouve l'injection continue de $D(A_1)$ dans $D(A_2)$. ∎

A l'aide de ce résultat, comme premier exemple on peut prendre $D_1 = D(A_1)$ et $\tilde{D}_0 = H$, puisque A_1 est un isomorphisme de $D(A_1)$ dans H et A_2 est borné de $D(A_1)$ dans H. La caractérisation du domaine de $A_1^{\frac{3}{2}}$ et des hypothèses supplémentaires sur les coefficients de L_1 et L_2 permettent de construire un second choix de paires $D_s, \tilde{D}_{s-1}$.

Lemme 3.1.2. *Supposons que la frontière de Ω est de classe $C^{2,1}$, $a_{k,\ell}^{(i)} = a_{\ell,k}^{(i)} \in C^{1,1}(\bar{\Omega})$, et que $b_k^{(i)}, c^{(i)} \in C^{0,1}(\Omega)$. Alors*

$$D(A_1^{\frac{3}{2}}) = A_1^{-1}H_0^1(\Omega) = \{u \in H^3(\Omega) \cap H_0^1(\Omega) : A_1 u = 0 \ \text{ sur } \ \partial\Omega\}. \tag{3.5}$$

Par ailleurs, si nous avons

$$a_{k,l}^{(2)} = a_{k,l}^{(1)} \ \text{ et } d^{(2)} \cdot n = d^{(1)} \cdot n \ \text{ sur } \partial\Omega, \tag{3.6}$$

où n est le vecteur normal unitaire sortant le long de la frontière $\partial\Omega$ et $d^{(i)} = (d_l^{(i)})_{l=1}^d$ est le vecteur donné par

$$d_l^{(i)} = -\sum_k^d \partial_k a_{k,l}^{(i)} + b_l^{(i)}.$$

Alors A_2 est borné de $D(A_1^{\frac{3}{2}})$ dans $H_0^1(\Omega)$.

Preuve : Soit $u \in D(A_1^{\frac{3}{2}})$, donc il existe un $h \in H_0^1(\Omega)$ tel que

$$A_1 u = h,$$

alors $A_1 u = 0$ sur $\partial\Omega$. De plus grâce à [34, Théorème 2.5.1.1, p. 128], $u \in H^3(\Omega)$, d'où l'inclusion

$$D(A_1^{\frac{3}{2}}) \hookrightarrow \{u \in H^3(\Omega) \cap H_0^1(\Omega) : A_1 u = 0 \text{ sur } \partial\Omega\}.$$

L'inclusion inverse étant triviale, la première assertion est prouvée.
On passe maintenant à la deuxième assertion. Fixons $u \in D(A_1^{\frac{3}{2}})$, alors par la régularité des coefficients de L_2, on voit directement que $A_2 u = L_2 u$ appartient à $H^1(\Omega)$. Il reste donc à vérifier que c'est égale à zéro sur la frontière de Ω. Pour cela, nous remarquons que

$$L_i u = -\sum_{k,\ell=1}^{d} a_{k,\ell}^{(i)} \partial_k \partial_\ell u + d^{(i)} \cdot \nabla u + c^{(i)} u.$$

Ainsi sur la frontière, en décomposant le gradient de u en sa partie tangentielle et sa partie normale, et en rappelant que $u = 0$ sur la frontière, on obtient

$$L_i u = -\sum_{k,\ell=1}^{d} a_{k,\ell}^{(i)} \partial_k \partial_\ell u + (d^{(i)} \cdot n) \partial_n u \text{ sur } \partial\Omega.$$

L'hypothèse (3.6) implique alors que

$$L_2 u = L_1 u \text{ sur } \partial\Omega,$$

ce qui termine la preuve puisque $L_1 u = 0$ sur la frontière. ∎

Notons qu'un résultat similaire peut rester vrai pour des frontières moins régulières. En effet, c'est le cas par exemple pour un carré et pour L_1 réduit à l'opérateur de Laplace.

Lemme 3.1.3. *Si Ω est le carré $(0,1)^2$ du plan et $L_1 = \Delta$, alors (3.5) reste vrai.*

Preuve : On utilise un argument de la preuve du Lemme 2.4 dans [38]. Soit $v \in D(A_1^{\frac{3}{2}}) \hookrightarrow D(A_1) = H^2(\Omega) \cap H_0^1(\Omega)$, donc $g = \Delta v \in H_0^1(\Omega)$. Alors par la régularité elliptique, $v \in H^3(\Omega \setminus V)$, où V est un voisinage des coins. Il reste à montrer la régularité H^3 près des coins. Par symétrie, il suffit de montrer cette régularité près de 0. Fixons alors une fonction radiale de troncature η telle que $\eta = 1$ près de 0 et supp $\eta \subset B(0, \frac{1}{2})$. Par conséquent $u = \eta v$ (prolongée par 0 en dehors de son support) appartient à $H^2((0,\infty)^2) \cap H_0^1((0,\infty)^2)$ et vérifie

$$\begin{aligned} \Delta u &= \eta \Delta v + 2 \nabla \eta \nabla v + v \Delta \eta \\ &= \eta g + 2 \frac{\partial \eta}{\partial r} \frac{\partial v}{\partial r} + v \Delta \eta = \tilde{f} \in H_0^1((0,\infty)^2). \end{aligned}$$

On pose maintenant

$$U(x,y) = u(|x|,|y|) \text{ sign } xy \quad \text{et} \quad F(x,y) = \tilde{f}(|x|,|y|) \text{ sign } xy.$$

Comme $\tilde{f}|_{\{x=0 \text{ et } y>0\}} = \tilde{f}|_{\{y=0 \text{ et } x>0\}} = 0$, $F \in H^1(\mathbb{R}^2)$. Or d'après [63, p.85],

$$\Delta U = F,$$

au sens des distributions. Ainsi $U \in H^3(\mathbb{R}^2)$, par conséquent $u \in H^3((0,\infty)^2)$ et finalement $v \in H^3(\Omega)$. ∎

En conclusion, si (3.5) et (3.6) sont valides, on peut prendre $D_{3/2} = D(A_1^{\frac{3}{2}})$ et $\tilde{D}_{1/2} = D(A_1^{\frac{1}{2}}) = H_0^1(\Omega)$.

Maintenant pour construire le schéma totalement discret, nous allons utiliser une méthode des éléments finis basée sur un maillage de Ω. Pour cela, on considère une famille de maillages $\{\mathcal{T}_h\}_h$ de Ω avec $\mathcal{T}_h = (K_l)_{l=1,\dots,N}$, où les éléments finis K_l sont des tétraèdres (ou des triangles). Pour simplifier l'analyse, on suppose que le bord de Ω est exactement triangulé et on considère la méthode des éléments finis de Lagrange comme illustrée dans [12]. Aussi, pour chaque élément K, on note par $\mathcal{F}_K$ l'application qui envoie l'élément de référence $\hat{K}$ à K (voir Figure 3.1).

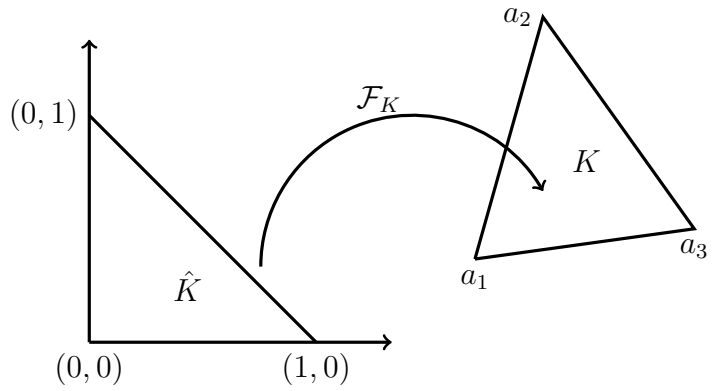


FIGURE 3.1 – Triangle de référence-Transformation affine.

À partir de cette triangulation $\mathcal{T}_h$, on définit l'espace d'approximation $V_h \subset H_0^1(\Omega)$ par

$$V_h = \left\{ v_h \in H_0^1(\Omega) : \ | \ v_h|_K \circ \mathcal{F}_K^{-1} \in \mathbb{P}_p(\hat{K}) \ \forall K \in \mathcal{T}_h \right\}, \tag{3.7}$$

où $\mathbb{P}_p(\hat{K})$ est l'ensemble des polynômes de degré inférieur ou égal à p.

Dans ce cadre, par le Corollaire 5.2 de [12] (voir aussi le Théorème 3.2.2 de [20]), l'hypothèse $\mathbf{H_4}$ est satisfaite pour $s = 1$ (sous l'hypothèse (3.3)) ou $s = \frac{3}{2}$ (sous les

hypothèses (3.5) et (3.6)), avec $q(s) = 2s - 1$ et le choix $p \geq q(s)$, i.e., pour tout $\varphi \in D_s$, on a

$$\|\varphi - Q_h\varphi\| \lesssim h^{q(s)}\|\varphi\|_{D_s}.$$

Enfin, les schémas totalement discrets du problème (3.2) peuvent être formulés comme suit :

- Le schéma d'Euler explicite consiste à trouver $U_{n+1,h} \in V_h$ solution de

$$a_1\big(U_{n+1,h}, \chi_h\big) = a_1(U_{n,h}, \chi_h) - \Delta t a_2(U_{n,h}, \chi_h) + \Delta t(f(t_n, U_{n,h}), \chi_h), \tag{3.8}$$

  pour tout $\chi_h \in V_h$, qui permet de calculer $U_{n+1,h}$ à partir de la connaissance de $U_{n,h}$ et de $f(t_n, U_{n,h})$.
- De même, par le schéma de Runge-Kutta d'ordre 2, on cherche $U^*_{n+1,h} \in V_h$ solution de

$$a_1(U^*_{n+1,h}, \chi_h) = a_1(U_{n,h}, \chi_h) + \Delta t(f_h(t_n, U_{n,h}), \chi_h) - \Delta t a_2(U_{n,h}, \chi_h),\ \forall \chi_h \in V_h,$$

  et ensuite $U_{n+1,h} \in V_h$ solution de

$$a_1(U_{n+1,h} - U_{n,h}, \chi_h) = \frac{\Delta t}{2}\big[\big(f(t_n, U_{n,h}) + f(t_{n+1}, U^*_{n+1,h}), \chi_h\big) - a_2(U_{n,h} + U^*_{n+1}, \chi_h)\big], \tag{3.9}$$

  pour tout $\chi_h \in V_h$.

3.1.2 Opérateurs elliptiques d'ordre deux : le cas singulier

On reste dans le même cadre que précédemment sauf qu'ici nous supposons que Ω est un polygone non convexe de $\mathbb{R}^2$. Dans ce cas, il est bien connu que le Lemme 3.1.1 n'est plus valable (voir [34, Chap. 4] par exemple), mais l'utilisation des espaces de Sobolev avec poids [61, 34] nous permettra de prouver un résultat approprié pour traiter le problème (3.2) dans notre cadre abstrait.

Nous commençons par la définition de ces espaces de Sobolev avec poids (voir [61] ou la Définition 8.4.1.1 de [34]). Puisque les résultats de régularité (et des stratégies de maillage raffiné) sont purement locaux, sans perte de généralité, on suppose que Ω a seulement un coin non convexe qu'on suppose être l'origine 0. Pour tout $x \in \Omega$, on note $r(x)$ la distance de x à 0.

Dans cette sous section, on suppose que pour un $l \in \mathbb{N}, l \geq 2$ fixé, les coefficients de L_i satisfont $a^{(i)}_{k,l} = a^{(i)}_{l,k} \in C^{l-2}(\Omega)$, $b^{(i)}_k$, $c^{(i)}_k \in C^{l-2}(\Omega)$.

Définition 3.1.4. *Pour $\alpha \in \mathbb{R}$ et $k \in \mathbb{N}^*$, on définit*

$$\begin{aligned} L^2_\alpha(\Omega) &= \{u \in L^2_{\mathrm{loc}}(\Omega) : r^\alpha u \in L^2(\Omega)\}, \\ V^k_\alpha(\Omega) &= \{u \in L^2_{\alpha-k}(\Omega) : r^{\alpha+|\beta|-k} D^\beta u \in L^2(\Omega), \forall \beta \in \mathbb{N}^2 : |\beta| \leq k\}. \end{aligned}$$

Ces espaces sont des espaces de Hilbert munis respectivement des normes naturelles

$$\begin{aligned} \|u\|_{L^2_\alpha(\Omega)} &= \|r^\alpha u\|_{L^2(\Omega)}, \\ \|u\|^2_{V^k_\alpha(\Omega)} &= \sum_{|\beta|\le k} \|r^{\alpha+|\beta|-k} D^\beta u\|^2_{L^2(\Omega)}. \end{aligned}$$

Lemme 3.1.5. *On a*

$$H^1_0(\Omega) \hookrightarrow V^1_0(\Omega).$$

Preuve : Par le Théorème 1.4.4.4 de [34], on a si $u \in H^1_0(\Omega)$, alors

$$\rho^{-1} u \in L^2(\Omega)$$

où $\rho(x)$ est la distance de x au bord, i.e. $\rho(x) = \min_{y\in\partial\Omega} |x-y| \le |x| = r(x)$, donc

$$|r^{-1}u| \le |\rho^{-1}u|.$$

■

On rappelle que par le Théorème 11.1.1 de [48], le plus petit exposant singulier $\lambda_a > 0$ vérifie

$$\lambda_a > \frac{1}{2}. \tag{3.10}$$

Notons que pour un $l \in \mathbb{N}, l \ge 2$ fixé et pour $i = 1$ ou 2, l'opérateur

$$A_i : V^l_\gamma(\Omega) \cap H^1_0(\Omega) \to V^{l-2}_\gamma(\Omega), u \mapsto L_i u,$$

est borné pour tout $\gamma \in \mathbb{R}$ tel que $l - 1 - \lambda_a \le \gamma \le l - \frac{3}{2}$. D'où la difficulté principale est de prouver que A_1 est un isomorphisme de $V^l_\gamma(\Omega) \cap H^1_0(\Omega)$ dans $V^{l-2}_\gamma(\Omega)$ pour un choix approprié de γ. Ceci est le but du résultat suivant.

Théorème 3.1.6. *Fixons $l \in \mathbb{N}, l \ge 2$. Alors pour tout $\gamma \in \mathbb{R}$ tel que*

$$l - 1 - \lambda_a < \gamma < l - \frac{3}{2}, \tag{3.11}$$

A_1 est un isomorphisme de $V^l_\gamma(\Omega) \cap H^1_0(\Omega)$ dans $V^{l-2}_\gamma(\Omega)$.

Preuve : Considérons $h \in V^{l-2}_\gamma(\Omega)$. Alors

$$h \in V^0_{\beta'}(\Omega), \quad \text{avec} \quad \beta' = \gamma - (l-2) < \frac{1}{2}. \tag{3.12}$$

Montrons qu'alors $h \in H^{-1}(\Omega)$. En effet pour $v \in H^1_0(\Omega)$ on écrit

$$\int_\Omega h\, v\, dx = \int_\Omega r^{\gamma-(l-2)}\, h\, r^{-1} v\, r^{l-1-\gamma}\, dx,$$

mais $r^{l-1-\gamma}$ est borné car $l-1-\gamma \geq 0$, $r^{-1}v \in L^2(\Omega)$ (par le Lemme 3.1.5) et $r^{\gamma-(l-2)}h \in L^2(\Omega)$ donc le produit est dans $L^1(\Omega)$, autrement dit par (3.11)

$$\Big|\int_\Omega hv\,dx\Big| \lesssim \|h\|_{L^2_{\beta'}(\Omega)}\|v\|_{L^2_{-1}(\Omega)} \lesssim \|h\|_{V_\gamma^{l-2}(\Omega)}\|v\|_{H_0^1(\Omega)},$$

ceci montre que

$$V_\gamma^{l-2}(\Omega) \hookrightarrow H^{-1}(\Omega).$$

Ainsi, il existe une et une seule solution $u \in H_0^1(\Omega)$ au problème

$$a_1(u,v) = \int_\Omega hv, \quad \forall v \in H_0^1(\Omega). \tag{3.13}$$

Maintenant on remarque que (3.12) implique que

$$h \in V_\beta^0(\Omega),$$

pour tout $\beta \in \mathbb{R}$ tel que $\frac{1}{2} - \epsilon \leq \beta \leq 1$, avec $\epsilon > 0$ suffisamment petit. En effet, on a par (3.11)

$$\gamma - (l-2) < \frac{1}{2},$$

donc grâce au Théorème 11.9.1 de [48], on obtient

$$u \in V_\beta^2(\Omega).$$

On utilise maintenant un argument de localisation et un théorème de comparaison. Fixons η une fonction de troncature telle que $\eta = 1$ près de 0. Alors $\eta u \in V_\beta^2(\Gamma)$ (Γ étant le cône près de 0) et vérifie

$$\begin{aligned} A_1(\eta u) &= \eta A_1 u + [A_1;\eta]u \\ &= \eta h + [A_1;\eta]u. \end{aligned}$$

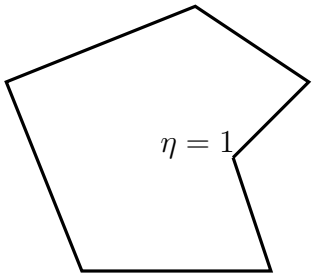


FIGURE 3.2 – Fonction de troncature η.

Mais on sait que $u \in H^l(\Omega)$ loin du coin, donc $[A_1;\eta]u \in H^{l-1}(\Omega)$ mais $[A_1;\eta]u = 0$ près de 0, donc $[A_1;\eta]u \in V_\gamma^{l-2}(\Gamma)$ et comme $\eta h \in V_\gamma^{l-2}(\Gamma)$, on a

$$A_1(\eta u) \in V_\gamma^{l-2}(\Gamma).$$

Par le Théorème 1.4.4 de [48], on obtient

$$\eta u \in V_\gamma^l(\Gamma),$$

tant que

$$l - 1 - \gamma < \lambda_a,$$

ce qui est vrai par hypothèse. ∎

Ce Théorème permet de vérifier que l'hypothèse $\mathbf{H_3}$ est satisfaite avec $D_s = V_\gamma^l(\Omega) \cap H_0^1(\Omega)$ et $\tilde{D}_{s-1} = V_\gamma^{l-2}(\Omega)$, pour tout $\gamma \in (1-\lambda_a, 1)$, et par conséquent le problème (3.2) est bien posé. Il est bien connu que la diminution de régularité implique que le taux de convergence, pour une MEF standard basée sur un maillage quasi-uniforme, est diminué mais en utilisant un maillage localement raffiné (près de 0), l'hypothèse $\mathbf{H_4}$ est satisfaite avec $q(s) = 1$ lorsqu'on choisit $l = 2$, voir [61] ou le Théorème 8.4.1.6 de [34].

3.1.3 Opérateurs elliptiques d'ordre deux à coefficients discontinus : le cas singulier

Nous étendons maintenant le cadre précédent au cas où les coefficients des parties principales de L_1 et L_2 sont constants par morceaux. Dans ce cas, le Lemme 3.1.1 n'est plus valable en général (voir [45, 46, 50, 55] par exemple). Comme précédemment l'utilisation des espaces de Sobolev avec poids nous permettra de faire entrer le problème (3.2) dans notre cadre abstrait.

Les hypothèses sur notre domaine et nos opérateurs sont les suivantes : On suppose que Ω est partitionné en sous-domaines Ω_j, $j = 1, \cdots, J$, avec J un entier positif tels que les Ω_j sont des polygones ouverts disjoints et que

$$\bar{\Omega} = \cup_{j=1}^J \bar{\Omega}_j.$$

Notons e_ℓ, $\ell = 1, \cdots, L$, l'ensemble des arêtes intérieures, i.e., l'ensemble des segments droits des intersections des $\bar{\Omega}_j \cap \bar{\Omega}_{j'}$ avec $j \neq j'$, qui ne sont pas inclus dans le bord de Ω. De même l'ensemble $\mathcal{S}$ des sommets de Ω est simplement l'ensemble de tous les sommets des Ω_j.

Dans la suite nous utiliserons les espaces à poids $\mathcal{V}_\alpha^k(\Omega)$ qui sont des espaces à poids par morceaux, plus précisément, si $r(x)$ est la distance aux sommets des Ω_j, on pose

$$\mathcal{V}_\alpha^k(\Omega) = \{v \in L_{\alpha-k}^p(\Omega) : v_j \in V_\alpha^k(\Omega_j), \forall j = 1, \cdots, J\},$$

où $v_j := v_{|\Omega_j}$ désigne la restriction de v à Ω_j. Ces espaces sont des espaces de Hilbert pour la norme naturelle.

On suppose maintenant que les opérateurs L_i sont elliptiques d'ordre 2 et sont de la forme précédente (3.1) mais avec des coefficients $a_{k,\ell}^{(i)} = a_{\ell,k}^{(i)}$ réguliers par morceaux, autrement dit la restriction de $a_{k,\ell}^{(i)}$ à Ω_j est régulière ($C^\infty(\bar{\Omega}_j)$). Comme précédemment on suppose également que la forme bilinéaire a_1 associée à A_1 est coercive de sorte que A_1 est un isomorphisme de $H_0^1(\Omega)$ dans $H^{-1}(\Omega)$.

Pour faciliter l'exposé, pour $i = 1$ ou 2, introduisons la matrice symétrique $M_i = (a_{k,\ell}^{(i)})_{k,\ell=1,2}$ et le saut du flux du gradient de u à travers une arête e_ℓ comme suit

$$[\![M_i \nabla u \cdot n]\!]_\ell = \Big((M_i \nabla u)_{|\Omega_j} - (M_i \nabla u)_{|\Omega_{j'}} \Big) \cdot n_\ell,$$

lorsque $e_\ell = \bar{\Omega}_j \cap \bar{\Omega}_{j'}$ et n_ℓ est le vecteur normal unitaire à e_ℓ orienté de Ω_j vers $\Omega_{j'}$.

Nous rappelons maintenant le Corollaire 4.4 de [55] qui est valable en dimension 2 sous l'hypothèse $\gamma < 1$ car il existe $r > 1$ tel que

$$L^2_\gamma(\Omega) \hookrightarrow L^r(\Omega). \tag{3.14}$$

Théorème 3.1.7. *Supposons que $\gamma \in (0,1)$ et que le segment $(0, 1-\gamma]$ ne contient pas d'exposant singulier de A_1 pour tous les sommets de Ω. Alors pour tout $f \in L^2_\gamma(\Omega)$ et $h_\ell \in V_\gamma^{\frac{1}{2}}(e_\ell)$, $\ell = 1, \cdots, L$, il existe une solution faible $u \in H^1_0(\Omega) \cap \mathcal{V}^2_\gamma(\Omega)$ au problème*

$$\begin{cases} L_1 u_j = f_j & \text{dans } \Omega_j, j = 1, \cdots, J, \\ \left[\!\left[M_1 \nabla u \cdot n \right]\!\right]_\ell = h_\ell & \text{sur } e_\ell, \ell = 1, \cdots, L, \end{cases} \tag{3.15}$$

dans le sens que

$$a_1(u,v) = F(v), \forall v \in H^1_0(\Omega),$$

où la forme linéaire F est donnée par

$$F(v) = \int_\Omega f(x)v(x)\,dx + \sum_{\ell=1}^{L} \int_{e_\ell} h_\ell \gamma v(x)\,d\sigma(x), \forall v \in H^1_0(\Omega).$$

Notons que la forme F définie précédemment est bien continue sur $H^1_0(\Omega)$ car pour tout $\gamma \in (0,1)$, il existe $r \in (1,2]$ tel que

$$V_\gamma^{\frac{1}{2}}(e_\ell) \hookrightarrow L^r(e_\ell). \tag{3.16}$$

En effet par définition, on a que

$$V_\gamma^{\frac{1}{2}}(e_\ell) \hookrightarrow V_0^{\gamma-\frac{1}{2}}(e_\ell),$$

donc tout $w \in V_\gamma^{\frac{1}{2}}(e_\ell)$ vérifiera

$$r^{\gamma-\frac{1}{2}} w \in L^2(e_\ell).$$

Si $\gamma \leq \frac{1}{2}$, on obtient directement que $w \in L^2(e_\ell)$, par contre si $\gamma \in (0, \frac{1}{2})$, comme

$$r^{\frac{1}{2}-\gamma} \in L^t(e_\ell),$$

pour tout $t < (\gamma - \frac{1}{2})^{-1}$, grâce à l'inégalité de Hölder on montre que $w \in L^r(e_\ell)$ pour un certain $r > 1$.

Ce théorème permet de vérifier que l'hypothèse $\mathbf{H_3}$ est satisfaite avec $D_1 = \mathcal{V}^2_\gamma(\Omega) \cap H^1_0(\Omega)$ et $\tilde{D}_0 = L^0_\gamma(\Omega) \times \prod_{\ell=1}^{L} V^{\frac{1}{2}}(e_\ell)$, pour tout $\gamma \in (1-\lambda_a, 1)$, lorsque λ_a est le plus petit exposant singulier positif de A_1. En effet, le résultat précédent affirme que A_1 est un isomorphisme de D_1 dans D_0, il nous reste à montrer que A_2 est borné de D_1 dans $\tilde{D}_0$.

Lemme 3.1.8. *A_2 est borné de D_s dans $\tilde{D}_{s-1}$.*

Preuve : Fixons $v \in V_\gamma^2(\Omega) \cap H_0^1(\Omega)$, alors $A_2 v$ est dans $H^{-1}(\Omega)$ et vaut

$$\begin{aligned} \langle A_2 v, w\rangle &= a_2(v,w) \\ &= \int_\Omega \left(\sum_{k,\ell=1}^{d} a_{k,\ell}^{(2)}(x)\ \partial_\ell v\ \partial_k w + \sum_{k=1}^{d} b_k^{(i)}(x) \partial_k v\ w + c^{(i)}(x) v\ w \right) dx, \end{aligned}$$

pour tout $w \in H_0^1(\Omega)$. Comme $b_k^{(i)} \partial_k v$ et $c^{(i)} v$ sont dans $L^2(\Omega)$, il reste à transformer le premier terme de ce second membre. Pour cela on fixe d'abord $w \in \mathcal{D}(\Omega)$. Par une application de l'inégalité de Hölder, il existe $r > 1$ tel que

$$\partial_\ell v_j \in W^{1,r}(\Omega_j), \forall j = 1, \cdots, J.$$

Ainsi une application de la formule de Green sur chaque Ω_j donne

$$\begin{aligned} \int_\Omega a_{k,\ell}^{(2)}(x)\ \partial_\ell v\ \partial_k w\, dx &= -\sum_{j=1}^{J} \int_\Omega \partial_k (a_{k,\ell}^{(2)}(x)\ \partial_\ell) v\ w\, dx \\ &+ \sum_{\ell=1}^{L} \int_{e_\ell} \left[\!\left[M_2 \nabla u \cdot n \right]\!\right]_\ell \gamma w(x)\, d\sigma(x), \end{aligned}$$

pour tout $w \in \mathcal{D}(\Omega)$. Comme $\partial_k(a_{k,\ell}^{(2)}\ \partial_\ell) v$ (resp. $\left[\!\left[M_2 \nabla u \cdot n \right]\!\right]_\ell$) appartient à $L_\gamma^2(\Omega)$, (resp. $V_\gamma^{\frac{1}{2}}(e_\ell)$) et que $H_0^1(\Omega)$ s'injecte dans $L^s(\Omega)$ pour tout $s > 1$ et en se rappelant (3.16) (3.14), grâce à l'inégalité de Hölder, l'égalité précédente reste vraie pour tout $w \in H_0^1(\Omega)$. Ainsi on en déduit que

$$\begin{aligned} \langle A_2 v, w\rangle &= \sum_{j=1}^{J} \int_\Omega L_2 v_j\ w_j\, dx \\ &+ \sum_{\ell=1}^{L} \int_{e_\ell} \left[\!\left[M_2 \nabla u \cdot n \right]\!\right]_\ell \gamma w(x)\, d\sigma(x), \end{aligned}$$

pour tout $w \in H_0^1(\Omega)$. Ceci termine le Lemme vu la régularité de $L_2 v_j$ et de $\left[\!\left[M_2 \nabla u \cdot n \right]\!\right]_\ell$. ∎

En conclusion le problème (3.2) est bien posé si on prend une donnée initiale u_0 dans D_1 et $f(t,u)$ continue à valeurs dans $\tilde{D}_0$. Comme à la sous-section précédente, en utilisant un maillage localement raffiné (près des coins), l'hypothèse $\mathbf{H_4}$ est vraie avec $q(s) = 1$, en appliquant [61] ou le Théorème 8.4.1.6 de [34] dans chaque sous-domaine Ω_j.

Remarque 3.1.9. Près d'un sommet extérieur où Ω est convexe ou si les coefficients $a_{k,\ell}^{(i)}$ sont continus en un sommet intérieur, alors le théorème du shift est valable dans les espaces de Sobolev standard et donc il n'est pas nécessaire de prendre les données dans des espaces à poids près de ce sommet et il suffit de les prendre dans H^2.

Remarque 3.1.10. Si L_2 est fortement elliptique, alors cet opérateur peut avoir une singularité S_μ près d'un sommet, avec $0 < \mu < 1$, autrement dit $S_\mu \in H_0^1(\Omega)$ se comporte comme r^μ près d'un sommet, est régulier par morceaux ailleurs et vérifie

$$L_2 S_\mu = g \in L^2(\Omega).$$

Alors la fonction $u(x,t) = S_\mu(x)$ est clairement solution de (3.2) avec $f = g$ et donnée initiale S_μ :

$$\begin{cases} L_1 u_t + L_2 u = g, & \text{dans } \Omega \times (0,T), \\ u = 0, & \text{sur } \partial\Omega \times (0,T), \\ u(0) = S_\mu, & \text{dans } H_0^1(\Omega). \end{cases}$$

Cette solution est bien fournie par notre cadre abstrait si on prend le paramètre γ de manière appropriée, en effet il suffit de prendre

$$\gamma > \max\{1-\mu, 1-\lambda_a\}.$$

Pour éviter de prendre des données initiales dans des espaces à poids, nous allons étendre le cadre précédent. Pour cela, pour chaque sommet intérieur $s \in \mathcal{S}_{\text{int}}$, on fixe une fonction de troncature régulière η_s qui vaut 1 près de s et est nulle sur un voisinage des autres sommets. On introduit alors

$$\hat{D}_1 = D_1 \oplus \text{ Span } \{\eta_s : s \in \mathcal{S}_{\text{int}}\},$$

autrement dit $v \in \hat{D}_1$ ssi il existe $v_w \in D_1$ et des coefficients $c_s \in \mathbb{R}$, $s \in \mathcal{S}_{\text{int}}$ tels que

$$v = v_w + \sum_{s \in \mathcal{S}_{\text{int}}} c_s \eta_s. \tag{3.17}$$

Cet espace est de Hilbert pour le produit scalaire naturel

$$(v, v') = (v_w, v'_w)_{D_1} + \sum_{s,s' \in \mathcal{S}_{\text{int}}} c_s c'_s,$$

lorsque $v = v_w + \sum_{s \in \mathcal{S}_{\text{int}}} c_s \eta_s$ et $v' = v'_w + \sum_{s \in \mathcal{S}_{\text{int}}} c'_s \eta_s$.

Nous pouvons alors démontrer le lemme suivant.

Lemme 3.1.11. *L'opérateur $A_1^{-1} A_2$ est borné de $\hat{D}_1$ dans lui-même.*

Preuve : Prenons $v \in \hat{D}_1$, alors il admet la décomposition (3.17) et donc

$$A_2 v = A_2 v_w + \sum_{s \in \mathcal{S}_{\text{int}}} c_s A_2 \eta_s.$$

Comme on a vu au Lemme 3.1.8 que $A_2 v_w$ est dans $\tilde{D}_0$, il reste à montrer que $A_2 \eta_s$ est aussi dans $\tilde{D}_0$. Si tel est le cas, alors $A_2 v$ sera dans $\tilde{D}_0$ et on conclut grâce au Théorème 3.1.7.

Fixons donc $s \in \mathcal{S}_{\text{int}}$ et calculons $A_2\eta_s$. Par définition on a

$$\begin{aligned}\langle A_2\eta_s, w\rangle &= a_2(v,w)\\ &= \int_\Omega \left(\sum_{k,\ell=1}^{d} a_{k,\ell}^{(2)}(x)\ \partial_\ell\eta_s\ \partial_k w + \sum_{k=1}^{d} b_k^{(i)}(x)\partial_k\eta_s\ w + c^{(i)}(x)\eta_s\ w\right) dx,\end{aligned}$$

pour tout $w \in H_0^1(\Omega)$. Mais comme η_s est régulière, on peut appliquer la formule de Green pour trouver

$$\begin{aligned}\langle A_2 v, w\rangle &= \sum_{j=1}^{J}\int_\Omega L_2\eta_{s,j}\ w_j\, dx\\ &+ \sum_{\ell=1}^{L}\int_{e_\ell} \left[\!\left[M_2\nabla\eta_s\cdot n\right]\!\right]_\ell \gamma w(x)\, d\sigma(x).\end{aligned}$$

Comme η_s est constant près des sommets de Ω, on en déduit que $L_2\eta_{s,j} \in L^2_\gamma(\Omega)$ et que $\left[\!\left[M_2\nabla\eta_s\cdot n\right]\!\right]_\ell \in V_\gamma^{\frac{1}{2}}(e_\ell)$, ce qui montre que $A_2\eta_s \in \tilde{D}_{s-1}$. ∎

Corollaire 3.1.12. *Si $f \in C([0,T];L^2(\Omega))$ et $u_0 \in \{v \in H_0^1(\Omega) : v_j \in H^2(\Omega)\}$, alors le problème*

$$\begin{cases} L_1u_t + L_2u = f, & \text{dans } \Omega\times(0,T),\\ u = 0, & \text{sur } \partial\Omega\times(0,T),\\ u(0) = u_0, & \text{dans } H_0^1(\Omega).\end{cases}$$

a une solution unique $u \in C^1([0,T];\hat{D}_1)$.

Preuve : Grâce aux inégalités de Hardy [34, p. 28], toute fonction $u_0 \in \{v \in H_0^1(\Omega) : v_j \in H^2(\Omega)\}$ admet la décomposition

$$u_0 = u_w + \sum_{s\in\mathcal{S}_{\text{int}}} u_0(s)\eta_s,$$

avec $u_w \in \mathcal{V}_\varepsilon^2(\Omega)$ pour tout $\varepsilon > 0$, ce qui montre que $u_0 \in \hat{D}_1$. Comme $L^2(\Omega) \hookrightarrow L^2_\gamma(\Omega)$, on aura de plus que $A_1^{-1}f \in C([0,T],D_1)$ et on conclut grâce à la continuité de $A_1^{-1}A_2$ de $\hat{D}_1$ dans lui-même. ∎

Dans le cadre de ce corollaire, nous trouvons une solution $u \in C^1([0,T];\hat{D}_1)$. Donc nos résultats de convergence seront garantis si on montre l'hypothèse $\mathbf{H_4}$ avec $q(s) = 1$ et des maillages raffinés mais pour φ dans $\hat{D}_1$. Mais une telle estimée est vérifiée car tout $\varphi \in \hat{D}_1$ s'écrit

$$\varphi = \varphi_w + \sum_{s\in\mathcal{S}_{\text{int}}} c_s\eta_s,$$

avec $\varphi_w \in D_1$ et des coefficients $c_s \in \mathbb{R}$. Pour le premier terme, on a par [61] ou le Théorème 8.4.1.6 de [34] que

$$\|\varphi_w - I_h\varphi_w\|_{1,\Omega} \lesssim h\|\varphi_w\|_{D_1},$$

où I_h est l'interpolé de Lagrange. Pour le deuxième terme, comme les fonctions η_s sont au moins dans $H^2(\Omega)$, une estimée d'interpolation standard donne

$$\|\eta_s - I_h\eta_s\|_{1,\Omega} \lesssim h\|\eta_s\|_{D_s} \lesssim h.$$

En conclusion la fonction $I_h\varphi$ vérifie

$$\|\varphi - I_h\varphi\|_{1,\Omega} \lesssim h\|\varphi\|_{\hat{D}_s},$$

ce qui prouve l'hypothèse $\mathbf{H_4}$ avec $q(s) = 1$.

3.1.4 Résultats numériques

Pour la validation de nos résultats du Chapitre 1 (voir le Corollaire 1.1.14), on propose des exemples tests. Dans chaque test proposé, on commence par choisir dans (3.2) les deux opérateurs L_1 et L_2 et le domaine Ω. Et ensuite on construit la courbe d'erreur en norme $H_0^1(\Omega)$ en espace et en temps. Dans les Tests 3.1.1 et 3.1.2 on présente les résultats numériques pour le cas d'une équation linéaire en prenant $L_1 = I - \Delta$ et $L_2 = -\Delta$ (Δ étant l'opérateur Laplace) avec Ω un polygone convexe de $\mathbb{R}^2$ dans le premier test et non convexe dans le deuxième. Toujours pour le cas d'une équation linéaire le Test 3.1.3 représente des résultats numériques en prenant dans (3.2), $L_2 = -\Delta$ et $L_1 = -\text{div } a\nabla$, avec a constante par morceaux. Comme la recherche de solution exacte pour ce type de problème n'est pas triviale, la vérification des résultats abstraits se fera en choisissant une solution exacte arbitraire, notée u, de façon à obtenir la fonction f (second membre) et u_0 (donnée initiale). Dans le Test 3.1.4, on présente les résultats numériques pour le cas d'une équation semi-linéaire. Comme la solution exacte est inconnue, on estimera les ordres de convergence en espace et en temps d'une manière différente par rapport aux premiers tests.
Les calculs numériques qui suivront ont été réalisés avec freefem++ [1]. Tout au long de cette sous section on note

$$\|U_n\|_\infty = \max_{0\le n\le N} \|U_n\|_{H^1(\Omega)}, \quad \forall\, (U_n)_n \in (H^1(\Omega))^N.$$

Test 3.1.1. (**Le cas régulier**) Le domaine Ω sera le carré $(0,1)^2 \subset \mathbb{R}^2$. Comme solution exacte, on prend

$$u(t,x,y) = x(1-x)y(1-y)\sin(t), \forall\, t \in [0,T],\ x,y \in (0,1).$$

u satisfait clairement la condition aux limites $u = 0$ sur la frontière du domaine, i.e. sur chaque côté du carré. En calculant directement f, on obtient

$$f(t,x,y) = x(1-x)y(1-y)\cos(t) + 2(\sin(t)+\cos(t))(x(1-x)+y(1-y)).$$

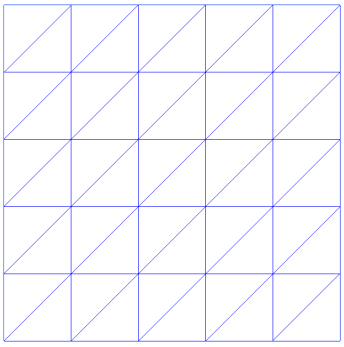


FIGURE 3.3 – Triangulation régulière du carré $\Omega = (0,1)^2$ pour $h = 0.2$.

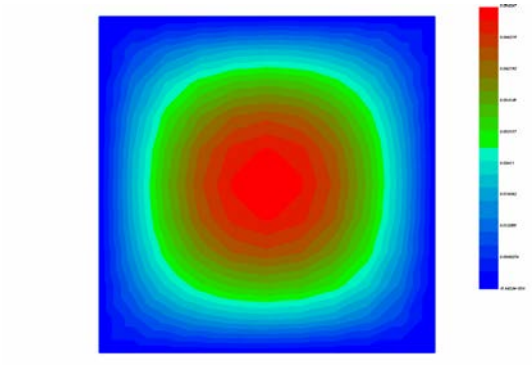


FIGURE 3.4 – La solution approchée par le schéma d'Euler (3.8) du problème (3.2) au temps final $T = 1$ avec $\Delta t = h = 0.1$.

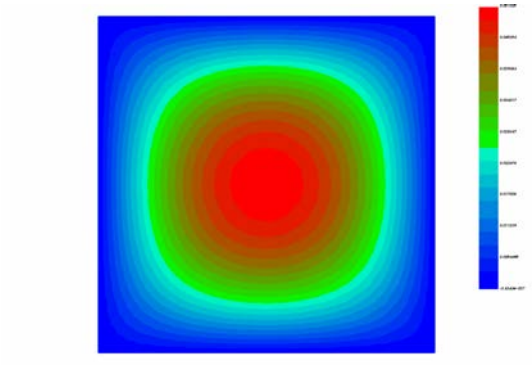


FIGURE 3.5 – La solution exacte du problème (3.2) au temps final $T = 1$.

Nous présentons premièrement les tests sur le schéma d'Euler explicite (3.8), i.e.

$$\Big(\frac{U_{n+1,h} - U_{n,h}}{\Delta t}, \chi_h\Big) + \Big(\nabla\big(\frac{U_{n+1,h} - U_{n,h}}{\Delta t}\big), \nabla\chi_h\Big) + \big(\nabla U_{n,h}, \nabla\chi_h\big) = \big(f(t_n), \chi_h\big), \quad \forall \chi_h \in V_h.$$

La Figure 3.3 montre la triangulation régulière du carré $(0,1)^2$. La solution approchée obtenue par ce schéma est illustrée dans la Figure 3.4 au temps final $T = 1$ avec $\Delta t = h = 0.1$. La Figure 3.5 montre la solution exacte au temps final $T = 1$. On construit la courbe d'erreur en norme $H_0^1(\Omega)$ en espace et en temps. Pour construire la courbe d'erreur à l'instant final en espace, on fixe le pas de temps en prenant $\Delta t = 0.001$ (i.e., on choisit le pas de temps Δt suffisamment petit de sorte que l'erreur due à la discrétisation en temps soit négligeable) et on calcule l'erreur avec différents pas d'espace. De même, pour construire la courbe d'erreur en temps à l'instant final, on fixe le pas d'espace en prenant $h = 0.00625$ de sorte que l'erreur due à la discrétisation en espace soit négligeable et on calcule l'erreur avec différents pas de temps. Les résultats numériques sont résumés dans

le Tableau 3.1 pour l'espace et dans le Tableau 3.2 pour le temps. En utilisant l'échelle

h	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
0.1	1.92 e-004
0.05	9.71 e-005
0.025	4.95 e-005
0.0125	2.65 e-005
0.00625	1.63 e-005

Tableau 3.1 – L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$.

Δt	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
0.1	3.98 e-002
0.05	1.96 e-002
0.025	9.83 e-003
0.0125	5.06 e-003
0.00625	2.85 e-003

Tableau 3.2 – L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 1$ pour différents Δt avec $h = 0.00625$.

logarithmique, on obtient la droite de régression d'équation $y = 0.94x - 6.394$ pour l'espace et la droite de régression d'équation $y = x - 0.9$ pour le temps. L'erreur en norme $H_0^1(\Omega)$ est donc d'ordre 1 en espace et d'ordre 1 en temps, ce qui est compatible avec l'estimation (1.38) du Corollaire 1.1.14. Les courbes d'erreurs sont illustrées dans la Figure 3.6.

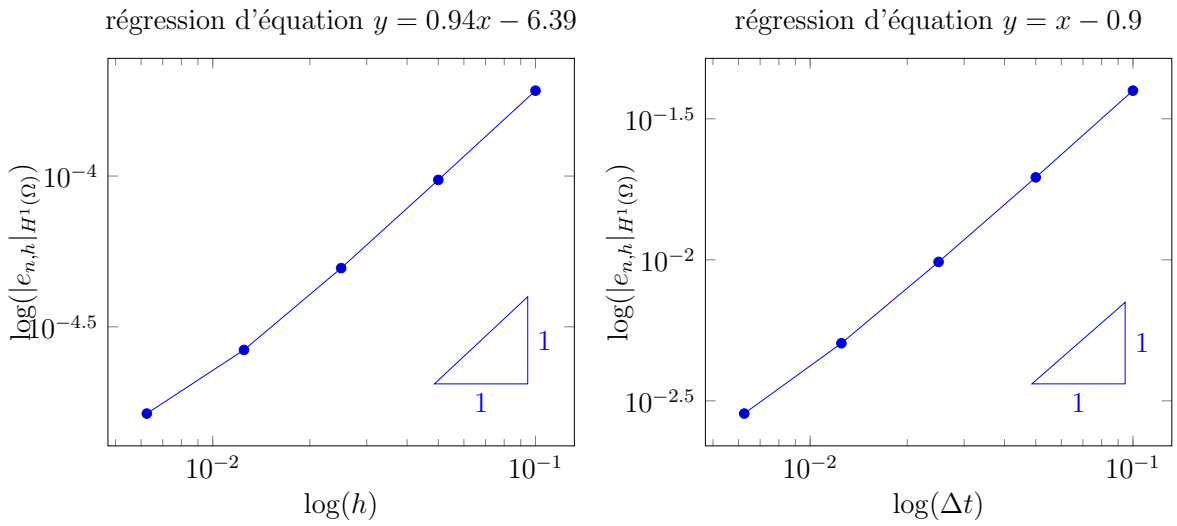


FIGURE 3.6 – Erreur en norme $H_0^1(\Omega)$ par le schéma d'Euler explicite (3.8) en espace et en temps au temps final T.

Nous présentons maintenant les résultats numériques relatifs au schéma numérique de Runge-Kutta d'ordre 2 (3.9) où l'espace d'approximation V_h est basé sur la méthode des éléments finis $\mathbb{P}_2$. Dans ce cas on prend comme solution exacte

$$u(t,x,y) = \big[x(1-x)y(1-y)\big]^3 \sin(t),$$

qui appartient alors à $C([0,T];D(A_1^{\frac{3}{2}}))$. En calculant f, on obtient

$$\begin{aligned} f(t,x,y) &= \cos(t)\big[x(1-x)y(1-y)\big]^3 \\ &\quad +6(\cos(t)+\sin(t))(y(1-y)+x(1-x))\big[(x(1-x)y(1-y))^2 \\ &\quad -((y-y^2)^2(1-2x)^2+(x-x^2)^2(1-2y)^2)\big]. \end{aligned}$$

h	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
0.1	4.98 e-006
0.05	1.44 e-006
0.025	3.78 e-007
0.0125	9.58 e-008
0.00625	2.40 e-008

Tableau 3.3 – L'évolution de l'erreur par le schéma de Runge-Kutta (3.9) et MEF$\mathbb{P}_2$ au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$.

Δt	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
1	0.000145
0.5	2.92 e-005
0.25	6.49 e-006
0.125	1.54 e-006
0.0625	4.36 e-007

Tableau 3.4 – L'évolution de l'erreur par le schéma de Runge-Kutta (3.9) et et MEF$\mathbb{P}_2$ au temps final $T = 20$ pour différents Δt avec $h = 0.00625$.

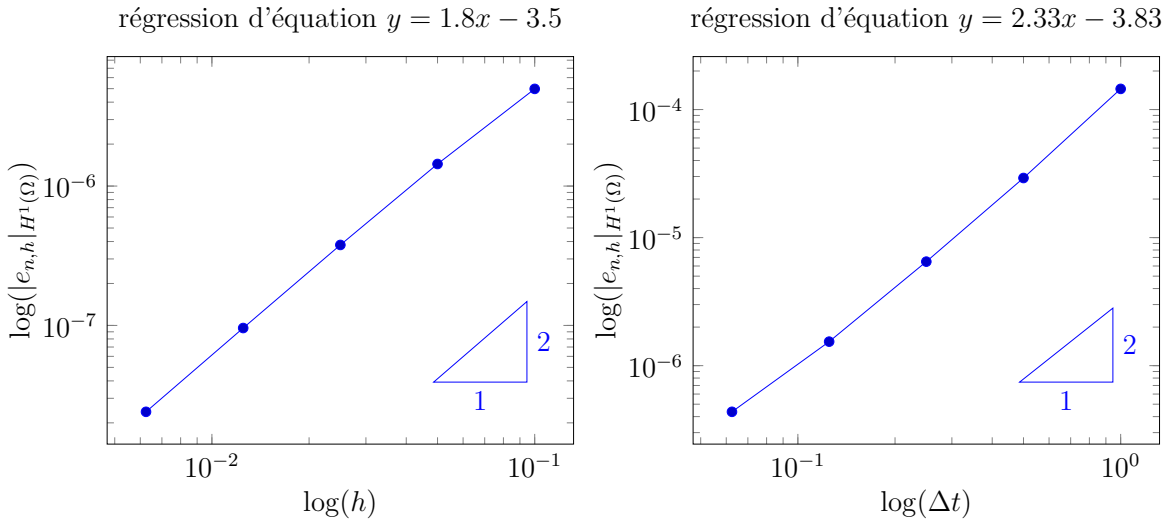


FIGURE 3.7 – Erreur en norme $H_0^1(\Omega)$ par le schéma de Runge-Kutta (3.9) en espace et en temps au temps final T.

A partir des Tableaux 3.3 et 3.4 (ou de la Figure 3.7), on voit que l'ordre de convergence en norme $H^1(\Omega)$ du schéma numérique (3.9) est 2 en espace et 2 en temps. Ainsi on retrouve exactement l'ordre de convergence théorique de (3.9) fourni par le Corollaire 1.1.14.

Test 3.1.2. (Le cas singulier) Passons maintenant au cas singulier, on choisit pour Ω le domaine non convexe $]0,2[^2\backslash[1,2[\times]0,1[\subset\mathbb{R}^2$.

Nous prendrons comme solution exacte

$$u(t,x,y)=r^{2/3}(x,y)\sin\big(\frac{2\theta}{3}\big)x(2-x)y(2-y)\sin(t),\forall t\geq 0,(x,y)\in\Omega,$$

ici, pour tout $(x,y)\in\Omega$, $r(x,y)$ est la distance entre (x,y) et le coin non convexe $(x=1,y=1)$ et $\theta\in[0,\frac{3}{2}\pi]$. Après quelques calculs, on trouve

$$\begin{aligned} f(t,x,y) &= r^{2/3}(x,y)\sin(\frac{2\theta}{3})x(2-x)y(2-y)\cos(t)+(\sin(t)+\cos(t))\\ &\Big(\frac{4}{3}r^{-1/3}(x,y)\big(-\sin(\frac{\theta}{3})x(2-2x)y(2-y)\big)+\cos(\frac{\theta}{3})x(2-x)(2-2y)\big)\\ &-2r^2 3(x,y)\sin\frac{2\theta}{3}(y(2-y)+x(2-x)\Big). \end{aligned}$$

La Figure 3.8 montre la solution approchée du problème (3.2) obtenue par le schéma numérique (3.8) au temps final $T=1$ avec $\Delta t=0.1$ et un pas de discrétisation en espace $h=0.1$. Et la Figure 3.9 montre la solution exacte au temps final $T=1$.

Le Tableau 3.5 (ou la Figure 3.10 à gauche) représente l'évolution de l'erreur au temps final $T=0.1$ pour différents h, où on a choisi le pas de temps Δt suffisamment petit de sorte que l'erreur due à la discrétisation en temps soit négligeable (ici on a choisi $\Delta t=\frac{1}{1000}$). De la même façon, le Tableau 3.6 (ou la Figure 3.10 à droite) représente l'évolution de l'erreur au temps final $T=1$ pour différents Δt où h est suffisamment petit de sorte que la discrétisation en espace soit négligeable (ici on a choisi $h=\frac{1}{80}$).

h	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
0.2	0.0136
0.1	6.83e-003
0.05	3.43e-003
0.025	1.74e-003
0.0125	9.05e-004

Tableau 3.5 – L'évolution de l'erreur par le schéma d'Euler (3.8) et MEF$\mathbb{P}_1$ au temps final $T=0.1$ pour différents h avec $\Delta t=0.001$.

Δt	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
0.2	0.7238
0.1	0.3527
0.05	0.1737
0.025	0.0868
0.0125	0.0447

Tableau 3.6 – L'évolution de l'erreur par le schéma d'Euler (3.8) et MEF$\mathbb{P}_1$ au temps final $T=1$ pour différents Δt avec $h=0.00625$.

Test 3.1.3. Dans ce cas, on choisit le domaine $\Omega=(-1,1)\times(0,1)\subset\mathbb{R}^2$, l'opérateur $L_2=-\Delta$, et l'opérateur $L_1=-\operatorname{div}a\nabla$, avec a constante par morceaux, i.e.,

$$a=\begin{cases}\epsilon & \text{dans } \Omega_2,\\ 1 & \text{dans } \Omega_1\cup\Omega_3,\end{cases}$$

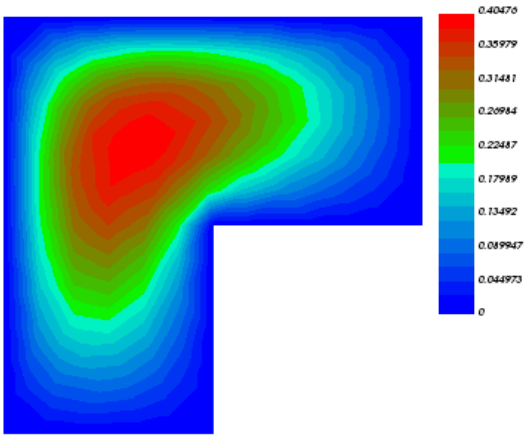


FIGURE 3.8 – La solution approchée par le schéma d'Euler (3.8) du problème (3.2) au temps final $T = 1$ avec $\Delta t = h = 0.1$.

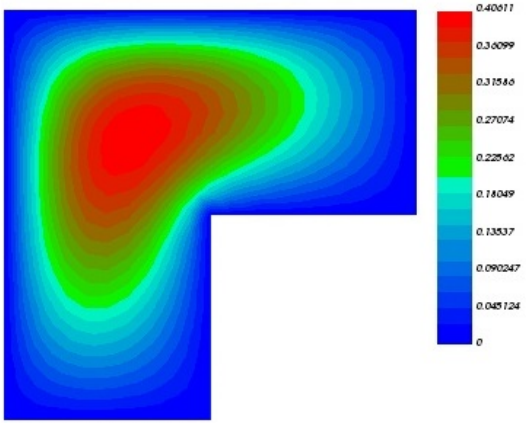


FIGURE 3.9 – La solution exacte du problème (3.2) au temps final $T = 1$.

où ϵ est un paramètre positif. Posons

$$\Omega_1 = \Omega \cap \{(r\cos\theta, r\sin\theta) : r > 0 \text{ et } 0 < \theta < \frac{\pi}{4}\},$$
$$\Omega_2 = \Omega \cap \{(r\cos\theta, r\sin\theta) : r > 0 \text{ et } \frac{\pi}{4} < \theta < \frac{3\pi}{4}\},$$
$$\Omega_3 = \Omega \cap \{(r\cos\theta, r\sin\theta) : r > 0 \text{ et } \frac{3\pi}{4} < \theta < \pi\},$$

et, comme d'habitude, (r, θ) sont les cordonnées polaires de (x, y) centrés à l'origine.

Dans ce cas, si $\epsilon \neq 1$, l'opérateur L_1 avec les conditions aux limites de Dirichlet a une singularité en $(0, 0)$ donnée par (voir [55])

$$S_\lambda = \begin{cases} r^\lambda \sin(\lambda\theta) & \text{dans } \Omega_1, \\ r^\lambda \left(\frac{2}{\epsilon+1}\sin(\lambda\theta) + \frac{\epsilon-1}{\sqrt{\epsilon}(\epsilon+1)}\cos(\lambda\theta)\right) & \text{dans } \Omega_1, \\ r^\lambda \sin(\lambda(\pi-\theta)) & \text{dans } \Omega_3, \end{cases}$$

avec $\lambda = \frac{4}{\pi}\arcsin\left(\sqrt{\frac{1}{\epsilon+1}}\right)$.

par conséquent on prend

$$u(t, x, y) = \sin t S_\lambda(x, y), \forall t \geq 0, x, y \in \Omega,$$

qui est considérée comme la solution exacte de

$$A_1 u_t + A_2 u = (\sin t)h,$$

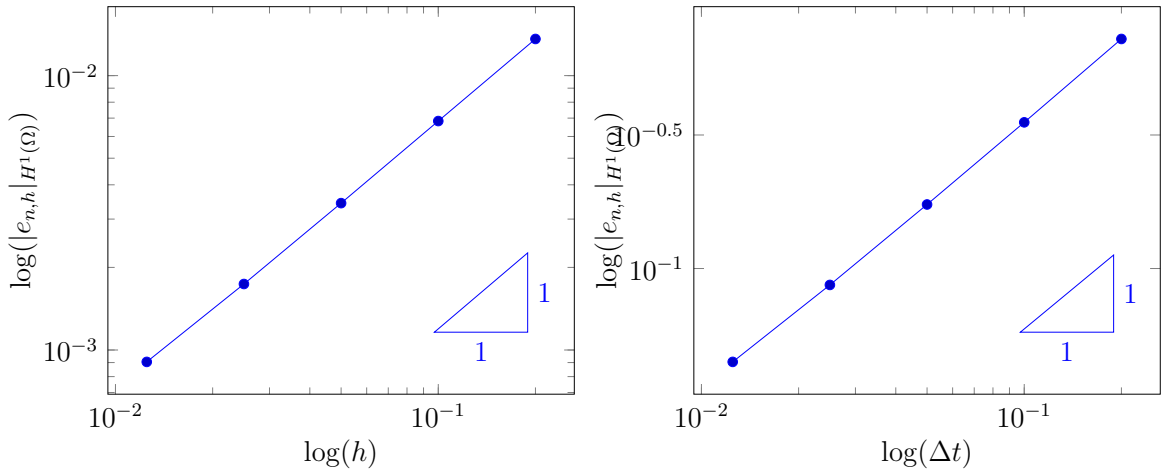


FIGURE 3.10 – Erreur en norme $H_0^1(\Omega)$ par le schéma d'Euler (3.8) en espace et en temps au temps final T.

où $h \in \tilde{D}_0$ est le saut de $\frac{\partial u}{\partial n}$ à travers les arêtes $e_1 = \bar{\Omega}_1 \cap \bar{\Omega}_2$ et $e_2 = \bar{\Omega}_2 \cap \bar{\Omega}_3$. Cela signifie que u est la solution de

$$\int_\Omega (a\nabla u_t \cdot \nabla v + \nabla u \cdot \nabla v)\, dxdy = \sum_{j=1,2} \int_{e_j} [\![\tfrac{\partial u}{\partial u}]\!]_\ell \gamma v d\sigma, \forall v \in H_0^1(\Omega).$$

Nous avons approché ce problème par le schéma d'Euler (3.8), où V_h est basé sur la méthode des éléments finis $\mathbb{P}_1$ sur des maillages uniformes ou raffinés (près de 0)(voir Figure 3.11) avec le choix $\epsilon = 3$ cela donne $\lambda = 2/3$. L'ordre de convergence de l'erreur en espace est présenté au Tableau 3.7 pour des maillages uniformes ou raffinés (voir Figure 3.11 pour $h = 0.2$) avec $\Delta t = 0.0001$ et un temps final $T = 0.1$.

A partir de Tableau 3.7, on voit que l'ordre de convergence en norme $H^1(\Omega)$ en espace est 2/3 (resp. 1) pour un maillage uniforme (resp. un maillage raffiné). Ici, par souci de brièveté, nous ne présentons pas l'ordre de convergence de l'erreur en temps parce que nous nous intéressons à l'influence des singularités spatiales.

Test 3.1.4. (Le cas d'une équation semi-linéaire) Considérons maintenant le problème (3.2) sur le carré unité $\Omega = (0,1)^2$ et une donnée initiale nulle avec

$$f(t,u) = \sqrt{1+t+u^2}. \tag{3.18}$$

Il est clair que la fonction f est continûment différentiable de $[0,T] \times H_0^1(\Omega)$ dans $\mathbb{R}$. Dans ce cas la solution exacte est inconnue, donc on ne peut pas calculer l'erreur entre la solution approchée et la solution exacte comme dans les exemples précédents, on va alors utiliser les solutions numériques $U_{n,2h}, U_{n,h}$, et $U_{n,h/2}$, et estimer l'ordre de convergence en

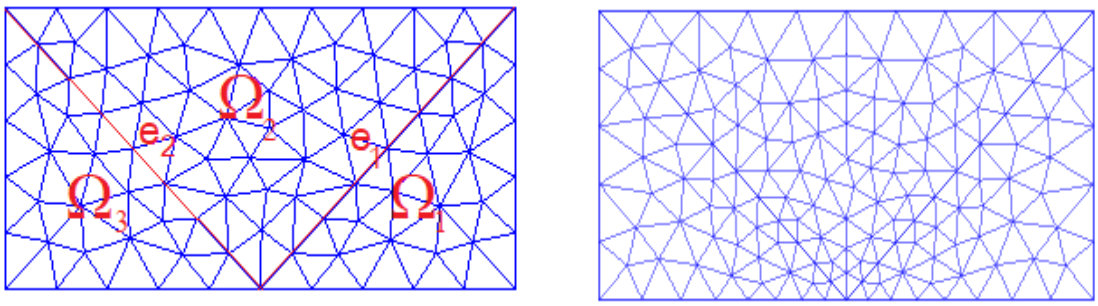


FIGURE 3.11 – A gauche : maillage initial uniforme avec $h = 0.2$, à droite : maillage initial raffiné près de 0.

	$\|u(t_n,\cdot) - U_{n,h}\|_\infty$	
h	maillage uniforme	maillage raffiné
0.2	0.0244	0.018
0.1	0.0164	0.0087
0.05	0.0105	0.0045
0.025	0.0064	0.0024
0.0125	0.0041	0.0013

Tableau 3.7 – L'évolution de l'erreur par le schéma d'Euler (3.8) et et MEF$\mathbb{P}_1$ au temps final $T = 0.1$ pour différents h en utilisant un maillage initial uniforme ou raffiné près de 0 pour $\Delta t = 0.0001$ et $\epsilon = 3$.

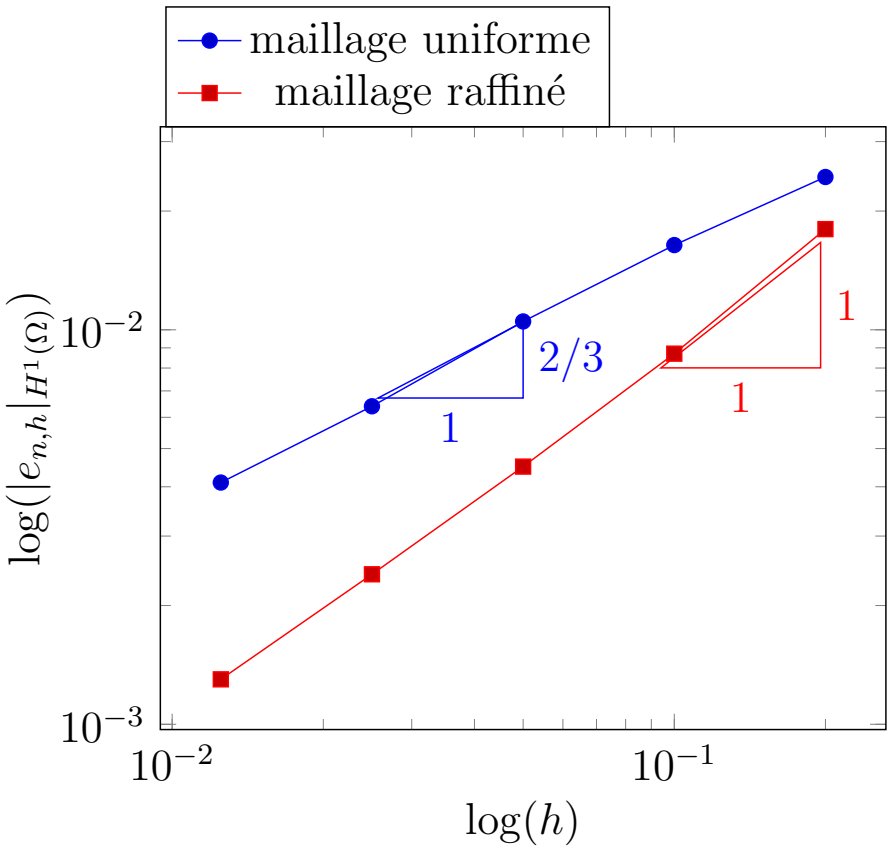


FIGURE 3.12 – Erreur en norme $H_0^1(\Omega)$ en espace au temps final $T = 0.1$ avec un maillage uniforme et un maillage raffiné pour $\Delta t = 0.0001$.

espace comme suit

$$\text{L'ordre de convergence en espace} = \log_2\left(\frac{\|U_{n,h} - U_{n,2h}\|_\infty}{\|U_{n,h/2} - U_{n,h}\|_\infty}\right), \tag{3.19}$$

où $U_{n,2h}$ et $U_{n,h/2}$ sont les solutions totalement discrètes pour les pas de maillage $2h$ et $h/2$, respectivement et un Δt suffisamment petit. De la même façon, on calcule l'ordre de convergence en temps

$$\text{L'ordre de convergence en temps} = \log_2\left(\frac{\|U_{n,h}^{\Delta t} - U_{n,h}^{2\Delta t}\|_\infty}{\|U_{n,h}^{\Delta t/2} - U_{n,h}^{\Delta t}\|_\infty}\right), \tag{3.20}$$

où $U_{n,h}^{2\Delta t}$ et $U_{n,h}^{\Delta t/2}$ sont les solutions totalement discrètes pour les pas de temps $2\Delta t$ et $\Delta t/2$, respectivement et avec un pas de maillage h suffisamment petit.

La Figure 3.13 montre la solution totalement discrète $U_{n,h}$ obtenue par le schéma d'Euler (3.8) et les éléments $\mathbb{P}_1$ au temps final $T = 1$ avec $\Delta t = 0.1$ et $h = 0.1$. Le Tableau 3.8 montre l'ordre de convergence pour différents Δt, où on a choisi h suffisamment petit de sorte que l'erreur due à la discrétisation en espace soit négligeable (ici $h = \frac{1}{160}$). De la

Δt	$\|U_{n,h}^{\Delta t} - U_{n,h}^{2\Delta t}\|_\infty$	L'ordre de convergence
0.1	0.035	1
0.05	0.017	1
0.025	0.0085	1
0.0125	0.0042	1
0.00625	0.0021	1
0.003125	0.0010	

Tableau 3.8 – L'ordre de convergence pour différents Δt avec $h = \frac{1}{160}$.

Le pas de maillage	$\|U_{n,h} - U_{n,2h}\|_\infty$	L'ordre de convergence
0.1	0.00582	0.94
0.05	0.00302	0.978
0.025	0.00153	1
0.0125	0.00076	1
0.00625	0.00038	1
0.003125	0.00019	

Tableau 3.9 – L'ordre de convergence pour différents h avec $\Delta t = \frac{1}{4000}$.

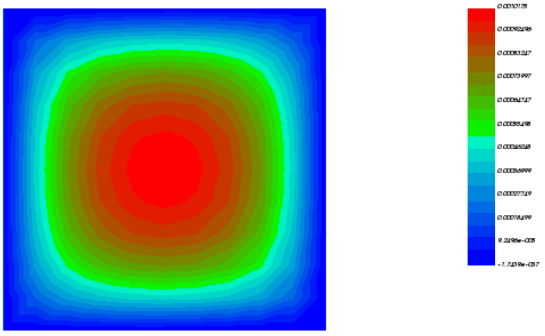


FIGURE 3.13 – La solution numérique $U_{n,h}$ obtenue par le schéma d'Euler (3.8) avec $\Delta t = 0.1$ et $h = 0.1$.

h	$\|U_{n,h} - U_{n,2h}\|_\infty$	L'ordre de convergence
0.1	0.00078	1.81
0.05	0.00022	1.85
0.025	6.18 e-005	1.87
0.0125	1.68 e-005	1.89
0.00625	4.53 e-006	

Tableau 3.10 – L'ordre de convergence pour différents h avec $\Delta t = \frac{1}{4000}$ pour le schéma d'Euler et EF$\mathbb{P}$2.

Δt	$\|U_{n,h}^{\Delta t} - U_{n,h}^{2\Delta t}\|_\infty$	L'ordre de convergence
0.1	0.032	0.962
0.05	0.0166	0.9816
0.025	0.0084	0.9908
0.0125	0.0042	0.9954
0.00625	0.0021	0.997
0.003125	0.001	

Tableau 3.11 – L'ordre de convergence pour différents Δt avec $h = \frac{1}{320}$ pour la méthode d'Euler et EF$\mathbb{P}$2.

même façon le Tableau 3.9 représente l'ordre de convergence pour différents h, où on a choisi Δt suffisamment petit ($\Delta t = \frac{1}{4000}$) de sorte que l'erreur due à la discrétisation en temps soit négligeable.
De même, le Tableau 3.10 (resp. 3.11) représente l'ordre de convergence en espace (resp. en temps) en utilisant la méthode des éléments finis $\mathbb{P}_2$.

3.2 Applications pour des équations quasi-linéaires de Sobolev

3.2.1 Le cas d'une équation non autonome

Dans ce paragraphe, nous nous concentrons sur le cas non autonome, i.e., on suppose que les opérateurs $A_1(t,u)$ et $A_2(t,u)$ dépendent seulement de la variable temporelle t, mais correspondent toujours aux opérateurs différentiels du second ordre. Plus précisément, dans un domaine borné Ω de $\mathbb{R}^d, d \geq 1$ à frontière lipschitzienne, pour

$i = 1, 2$, L_i est l'opérateur différentiel d'ordre deux de la forme

$$L_i(x, D_x, t)u = -\sum_{k,\ell=1}^{d} \partial_k(a_{k,\ell}^{(i)}(x,t)\partial_\ell u) + \sum_{k=1}^{d} b_k^{(i)}(x,t)\partial_k u + c^{(i)}(x,t)u,$$

où $a_{k,\ell}^{(i)} = a_{\ell,k}^{(i)} \in C([0,T]; C^{0,1}(\bar{\Omega})) \cap C^{0,\gamma}([0,T]; L^\infty(\Omega))$, $b_k^{(i)}, c^{(i)} \in C^{0,\gamma}([0,T]; L^\infty(\Omega))$, pour certains $\gamma \in (0,1]$. Supposons de plus que l'opérateur L_1 est uniformément elliptique, i.e., il existe un $\alpha_* > 0$ tel que

$$\sum_{k,\ell=1}^{d} a_{k,\ell}^{(1)}(x,t)\xi_\ell\xi_k \geq \alpha_*|\xi|^2, \forall \xi \in \mathbb{R}^d, t \in [0,T].$$

Dans ce cas, la forme bilinéaire $a_i(t;\cdot,\cdot), i = 1, 2$, est indépendante de u et est définie par

$$a_i(t; v, w) = \int_\Omega \left(\sum_{k,\ell=1}^{d} a_{k,\ell}^{(i)}(x,t)\partial_\ell v \partial_k w + \sum_{k=1}^{d} b_k^{(i)}(x,t)\partial_k v w + c^{(i)}(x,t) v w \right) dx,$$

pour tout $v, w \in H_0^1(\Omega)$, et pour tout $t \in [0,T]$, l'opérateur $A_i(t)$ défini par

$$\langle A_i(t)u, v\rangle_{V',V} = a_i(t, u, v), \forall u, v \in H_0^1(\Omega),$$

est continu de $H_0^1(\Omega)$ dans $H^{-1}(\Omega)$. Enfin, si on suppose que la forme bilinéaire a_1 est uniformément coercive sur $H_0^1(\Omega)$, alors les hypothèses $\mathbf{H_1}$, $\mathbf{H_2}$ et $\mathbf{H_6}$ seront satisfaites. Comme dans la sous-section 3.1.1, on peut montrer que $D(A_1(t)) = H^2(\Omega) \cap H_0^1(\Omega)$, pour tout $t \in [0,T]$ si la frontière de Ω est $C^{1,1}$ ou si Ω est convexe. Par conséquent, sous cette hypothèse supplémentaire, les hypothèses $\mathbf{H_3}$ et $\mathbf{H_4}$ avec $D_1 = H^2(\Omega) \cap H_0^1(\Omega)$ et $\tilde{D}_0 = L^2(\Omega)$ seront satisfaites si V_h est défini par (3.7) avec $p = 1$. Finalement, sous les hypothèses $\mathbf{H_5}$ et $\mathbf{H_7}$-$\mathbf{H_8}$ sur f, le problème suivant

$$\begin{cases} L_1(t)u_t + L_2(t)u = f(t,u), & \text{dans } \Omega \times (0,T), \\ u = 0, & \text{sur } \partial\Omega \times (0,T), \\ u(0) = u_0, & \text{dans } H_0^1(\Omega), \end{cases} \tag{3.21}$$

est bien posé et on peut l'approximer par le schéma totalement discret d'Euler (voir (1.22))

$$a_1(t_n; U_{n+1,h}, \chi_h) = a_1(t_n; U_{n,h}, \chi_h) - \Delta t a_2(t_n; U_{n,h}, \chi_h) + \Delta t(f(t_n, U_{n,h}), \chi_h), \quad \forall \chi_h \in V_h.$$

Test 3.2.1. Nous illustrons maintenant cette théorie en choisissant dans (3.21), $\Omega = (0,1)^2$, $L_1(t) = I - (1+t)\Delta$ et $L_2(t) = -(1+t)\Delta$. Il est clair que les formes bilinéaires $(a_i(t;\cdot,\cdot))_{t\in[0,T]}$ satisfont les hypothèses précédentes, en particulier on voit directement que

$$|a_i(t; u, v) - a_i(s; u, v)| \lesssim |t - s|\|u\|\|v\|, \quad \forall u, v \in H_0^1(\Omega), \forall\, t, s \in [0,T].$$

On commence d'abord par le cas d'une équation linéaire en considérant la solution exacte

$$u(t,x,y) = \sin(\pi x)\sin(\pi y)\sin t, \forall (x,y) \in \Omega, t > 0.$$

La solution approchée est illustrée dans la Figure 3.14 pour $\Delta t = 0.1$ et $h = 0.1$ au temps final $T = 1$ et la Figure 3.15 montre la solution exacte au temps final $T = 1$. On construit la courbe d'erreur en norme $H_0^1(\Omega)$ en espace à l'instant $T = 0.1$ et en temps à l'instant $T = 1$. Pour l'erreur en espace, on fixe le pas de temps en prenant à l'instant $T = 0.1, N = 1000$ i.e., $\Delta t = 0.001$ et on calcule l'erreur avec différents pas d'espace. De même, pour l'erreur en norme $H_0^1(\Omega)$ en temps, on fixe le pas d'espace en prenant $T = 1, N = 320$, i.e., $h = 0.003125$ et on calcule l'erreur avec différents pas de temps Δt. Les résultats numériques sont récapitulés dans le Tableau 3.12 pour l'espace et dans le Tableau 3.13 pour le temps. En utilisant une échelle logarithmique, on obtient la droite

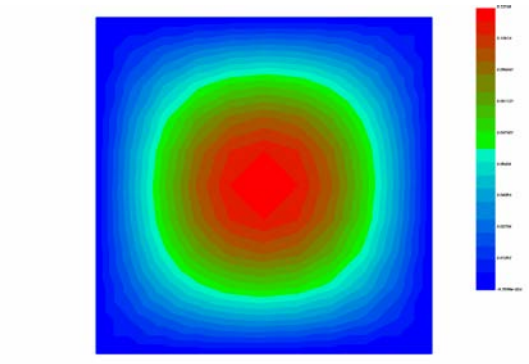


FIGURE 3.14 – La solution approchée par le schéma d'Euler (3.8) du problème (3.21) au temps final $T = 1$ avec $\Delta t = h = 0.1$.

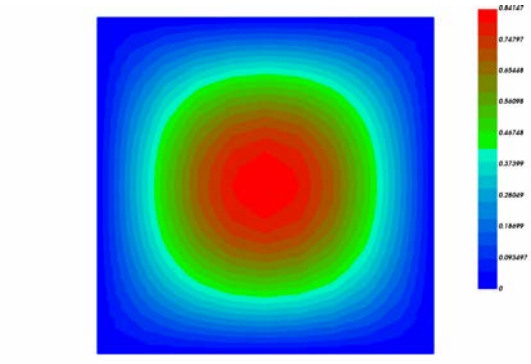


FIGURE 3.15 – La solution exacte du problème (3.21) au temps final $T = 1$.

de régression $y = 0.97x - 1.18$ pour l'espace et la droite de régression $y = 0.95x - 0.288$ pour le temps. Comme dans le cas linéaire, l'erreur en norme $H_0^1(\Omega)$ est donc d'ordre 1 en espace et en temps. Les courbes d'erreur sont données dans la Figure 3.16.

Test 3.2.2. On passe maintenant au cas d'une équation semi-linéaire en prenant le terme source $f(t,u)$ défini par (3.18) et la donnée initiale $u(0) = 0$.

La Figure 3.17 montre la solution totalement discrète $U_{n,h}$ au temps final $T = 1$ avec $\Delta t = h = 0.1$. L'ordre de convergence en temps (resp. en espace) est présenté au Tableau 3.14 avec $h = \frac{1}{160}$ (resp. 3.15 avec $\Delta t = \frac{1}{4000}$). Ces résultats sont en accord avec l'estimation d'erreur qu'on a déjà montré au Corollaire 1.1.14.

h	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
0.2	0.063
0.1	0.0325
0.05	0.0163
0.025	0.0082
0.0125	0.0042
0.00625	0.0023

Tableau 3.12 – L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$.

Δt	$\|u(t_n,x,y)-U_{n,h}\|_\infty$
0.2	0.168
0.1	0.082
0.05	0.040
0.025	0.021
0.0125	0.013
0.00625	0.0066

Tableau 3.13 – L'évolution de l'erreur par le schéma d'Euler (3.8) au temps final $T = 1$ pour différents Δt avec $h = 0.00625$.

Δt	$\|U_{n,h}^{\Delta t}-U_{n,h}^{2\Delta t}\|_\infty$	L'ordre de convergence
0.1	0.01632	0.94
0.05	0.00853	0.96
0.025	0.00437	0.98
0.0125	0.00221	0.99
0.00625	0.00111	1
0.003125	0.00055	

Tableau 3.14 – L'ordre de convergence pour différents Δt avec $h = \frac{1}{160}$.

3.2.2 Le cas d'une équation quasi-linéaire

Test 3.2.3. (Exemple en dimension 1) Dans cet exemple, on considère l'équation quasi-linéaire

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\big(\rho_1(x,u)\frac{\partial^2 u}{\partial x \partial t} + \rho_2(x,u)\frac{\partial u}{\partial x}\big) = f(t,u), & \text{dans } (0,1)\times(0,T), \\ u = 0, & \text{sur } \partial(0,1)\times(0,T), \\ u(0) = u_0, & \text{dans } (0,1), \end{cases} \tag{3.22}$$

où $\rho_1, \rho_2 : (0,1)\times\mathbb{R} \longmapsto \mathbb{R}$ sont deux fonctions continues vérifiant

- il existe deux constantes positive β, M telles que

$$\beta \le \rho_1(x,u) \le M, \quad \text{et} \quad |\rho_2(x,u)| \le M, \forall (x,u) \in \bar{\Omega}\times\mathbb{R}, \tag{3.23}$$

- la fonction $\rho_i, i = 1,2$, est globalement lipschitzienne, i.e., il existe une constante $L_i > 0$ telle que

$$|\rho_i(x,u) - \rho_i(x,\tilde{u})| \le L_i|u-\tilde{u}|, \quad \forall (x,u,\tilde{u}) \in (\bar{\Omega}\times\mathbb{R}^2). \tag{3.24}$$

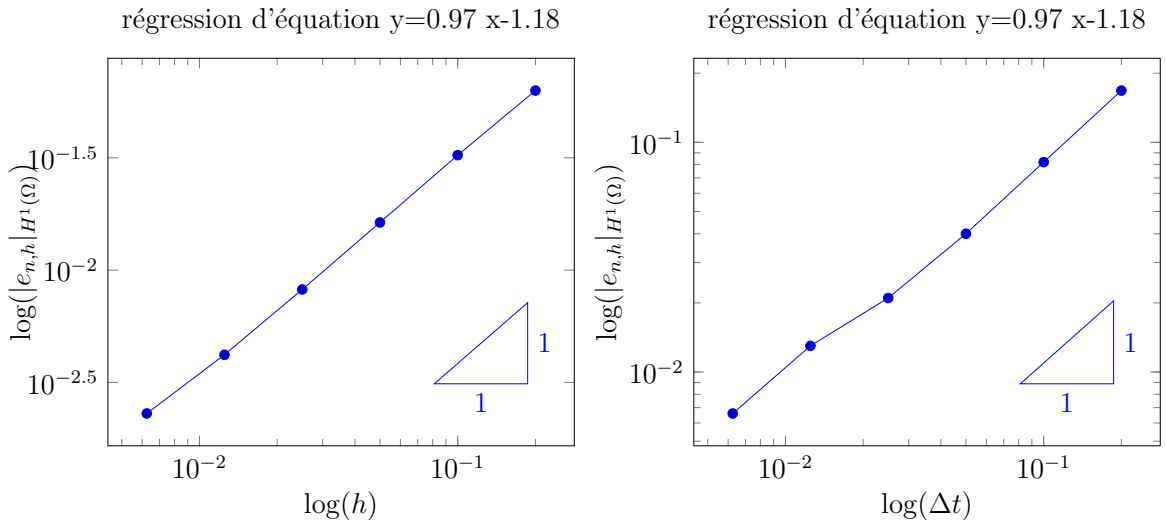


FIGURE 3.16 – Erreur en norme $H_0^1(\Omega)$ en espace au temps final $T = 0.1$ et en temps au temps final $T = 1$.

h	$\|U_{n,h} - U_{n,2h}\|_\infty$	L'ordre de convergence
0.1	0.00559	0.95
0.05	0.00290	0.98
0.025	0.00147	0.90
0.0125	0.00074	0.99
0.00625	0.000337	1
0.003125	0.00018	

Tableau 3.15 – L'ordre de convergence pour différents h avec $\Delta t = \frac{1}{4000}$.

Avec ces hypothèses, les formes bilinéaires $a_i(u;\cdot,\cdot), i = 1, 2$, définies sur $H_0^1(\Omega) \times H_0^1(\Omega)$, par

$$a_1(u;v,w) = \int_0^1 \big(\rho_1(x,u(x))\frac{\partial v}{\partial x}\frac{\partial w}{\partial x} + vw\big)dx \;\text{ et }\; a_2(u;v,w) = \int_0^1 \rho_2(x,u(x))\frac{\partial v}{\partial x}\frac{\partial w}{\partial x}dx,$$

sont indépendantes de t et satisfont les hypothèses $\mathbf{H_1}$, $\mathbf{H_2}$ et $\mathbf{H_6}$. En effet, par (3.23) combiné avec l'inégalité de Cauchy-Schwarz, on obtient pour tout $u, v, w \in H_0^1(\Omega)$

$$\begin{aligned} |a_1(u;v,w)| &\leq \int_0^1 \big(M\,|\tfrac{\partial v}{\partial x}|\,|\tfrac{\partial w}{\partial x}| + |v|\,|w|\big)dx \\ &\lesssim \|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)}, \end{aligned}$$

d'où la continuité de $a_1(u;\cdot,\cdot)$ sur $H_0^1(\Omega) \times H_0^1(\Omega)$, de même pour la forme bilinéaire $a_2(u;\cdot,\cdot)$. La coercivité de $a_1(u;\cdot,\cdot)$ découle directement de l'hypothèse $\rho_1(x,u) \geq \beta$ (voir

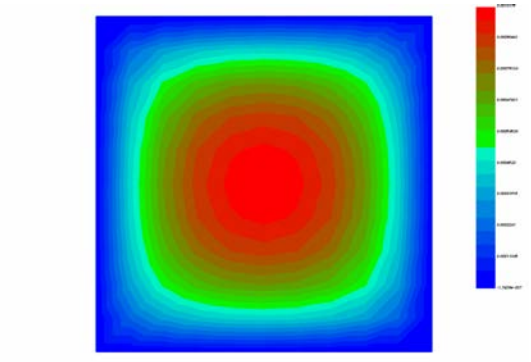


FIGURE 3.17 – La solution numérique $U_{n,h}$ obtenue par le schéma d'Euler (3.8) avec $\Delta t = h = 0.2$.

(3.23)). Enfin, pour tout $u, \tilde{u}, v, w \in H_0^1(\Omega)$, on a par (3.24)

$$\begin{aligned} |a_i(u; v, w) - a_i(\tilde{u}; v, w)| &\leq L_i \int_0^1 |u(x) - \tilde{u}(x)| \, |\tfrac{\partial v}{\partial x}| \, |\tfrac{\partial w}{\partial x}| \, dx \\ &\lesssim \sup_{x\in[0,1]} |u(x) - \tilde{u}(x)| \int_0^1 |\tfrac{\partial v}{\partial x}| \, |\tfrac{\partial w}{\partial x}| \, dx. \end{aligned}$$

Par l'inégalité de Cauchy-Schwarz et le fait que $H^1(\Omega) \hookrightarrow \mathcal{C}(\bar{\Omega})$ (par le théorème d'injection de Sobolev), on obtient

$$|a_i(u; v, w) - a_i(\tilde{u}; v, w)| \lesssim \|u - \tilde{u}\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \|w\|_{H^1(\Omega)},$$

d'où l'hypothèse $\mathbf{H_6}$.

Pour la discrétisation temporelle, on utilise le schéma d'Euler explicite. On note N le nombre de sous intervalles de temps. Le schéma explicite d'Euler du problème (3.22) est alors : chercher U_{n+1} connaissant U_n, tel que

$$\frac{U_{n+1} - U_n}{\Delta t} - \frac{\partial}{\partial x}\left[\rho_1(x, U_n)\frac{\partial}{\partial x}\left(\frac{U_{n+1} - U_n}{\Delta t}\right) - \rho_2(x, U_n)\frac{\partial U_n}{\partial x}\right] = f(t_n, U_n), \tag{3.25}$$

$$U_{n+1}(0) = U_{n+1}(1) = 0, n = 0, 1, ..., N-1. \tag{3.26}$$

Ainsi, la formulation variationnelle du problème (3.25)-(3.26) s'écrit : connaissant $U_n \in H_0^1(\Omega)$, on recherche une fonction $U_{n+1} \in H_0^1(\Omega)$ tel que pour tout $v \in H_0^1(\Omega)$

$$\begin{aligned} &\int_0^1 \frac{U_{n+1} - U_n}{\Delta t} v \, dx + \frac{1}{\Delta t}\int_0^1 \rho_1(x, U_n)\frac{\partial}{\partial x}(U_{n+1} - U_n)\frac{\partial v}{\partial x} \, dx + \int_0^1 \rho_2(x, U_n)\frac{\partial U_n}{\partial x}\frac{\partial v}{\partial x} \, dx \\ &\qquad = \int_0^1 f(t_n, U_n) v \, dx, \quad n = 0, 1, ..., N-1. \end{aligned} \tag{3.27}$$

On considère maintenant une subdivision uniforme de l'intervalle $[0,1]$ en N sous-intervalles

$$0 = x_0 \leq x_1 \leq ... \leq hN = 1.$$

On pose $h = x_{i+1} - x_i$ pour $0 \leq i \leq N-1$. Pour obtenir le schéma totalement discret, on construit une approximation par éléments finis $\mathbb{P}_1$ basée sur l'espace de dimension finie

$$V_h = \{v_h \in H_0^1(\Omega); v_h \mid_{[x_i,x_{i+1}]} \in \mathbb{P}_1, 0 \leq i \leq N, v_h(0) = v_h(1) = 0\}.$$

Le problème discret du problème (3.28) est alors : rechercher $U_{n+1,h}$ dans V_h, connaissant $U_{n,h} \in V_h$ tel que pour tout v_h de V_h, $n = 0, 1, ..., N-1$,

$$\int_0^1 \frac{U_{n+1,h} - U_{n,h}}{\Delta t} v_h \, dx + \frac{1}{\Delta t} \int_0^1 \rho_1(x, U_{n,h}) \frac{\partial}{\partial x}\big(U_{n+1,h} - U_{n,h}\big) \frac{\partial v_h}{\partial x} \, dx + \int_0^1 \rho_2(x, U_{n,h}) \frac{\partial U_{n,h}}{\partial x} \frac{\partial v_h}{\partial x} dx = \int_0^1 f(t_n, U_{n,h}) v_h \, dx. \quad (3.28)$$

Pour les illustrations numériques, on considère le terme source f défini par (3.18), une condition initiale nulle $u_0 = 0$ et

$$\rho_1(x,u) = \frac{1}{2} + \frac{u^2}{1+u^2}, \quad \rho_2(x,u) = \frac{u^2}{1+u^2}.$$

Il est clair que

$$\frac{1}{2} \leq \rho_1(x,u) \leq \frac{3}{2}, \quad \text{et} \quad \rho_2(x,u) \leq 1, \quad \forall x \in \bar{\Omega}, u \in \mathbb{R}.$$

Comme dans les exemple tests 3.1.4 et 3.2.2, on calcule l'ordre de convergence en espace (resp. en temps) par la formule (3.19) (resp. (3.20)). Le Tableau 3.16 montre l'ordre de

Δt	$\|u_{n,h}^{\Delta t} - u_{n,h}^{2\Delta t}\|_\infty$	l'ordre de convergence
0.1	0.0048	1
0.05	0.0024	1
0.025	0.0012	1
0.0125	0.0006	1
0.00625	0.0003	1
0.003125	0.00015	

Tableau 3.16 – L'ordre de convergence pour différents Δt avec $h = 0.003125$.

convergence pour différents Δt, où on a choisi h suffisamment petit ($h = \frac{1}{320}$) de sorte que l'erreur due à la discrétisation en espace soit négligeable. Tandis que le Tableau 3.17 représente l'ordre de convergence pour différents h, avec Δt suffisamment petit ($\Delta t = \frac{1}{1000}$).

h	$\|u_{n,h} - u_{n,2h}\|_\infty$	l'ordre de convergence
0.1	0.006	1
0.05	0.003	1
0.025	0.0015	1
0.0125	0.00073	1
0.00625	0.00036	1
0.003125	0.00018	

Tableau 3.17 – L'ordre de convergence pour différents h avec $\Delta t = 0.001$.

Test 3.2.4. (Exemple en dimension 2) Sur le carré unité $\Omega = (0,1)^2$ de $\mathbb{R}^2$, on considère le problème

$$\begin{cases} \Delta\big(\rho_1(x,u)\Delta u_t\big) + \Delta\big(\rho_2(x,u)\Delta u\big) = f(t,u), & \text{dans } \Omega \times (0,T), \\ u = \frac{\partial u}{\partial n} = 0, & \text{sur } \partial\Omega \times (0,T), \\ u(0) = u_0, & \text{dans } H_0^2(\Omega), \end{cases} \tag{3.29}$$

où $\frac{\partial u}{\partial n}$ désigne la dérivée normale extérieure de u sur $\partial\Omega$, ρ_1 et ρ_2 sont deux fonctions dans $C^2(\bar{\Omega} \times \mathbb{R}, \mathbb{R})$ vérifiant les hypothèses (3.23)-(3.24) avec des dérivées partielles secondes uniformément bornées en x et en u.

Pour tout $u \in H_0^2(\Omega)$, et $i = 1$ ou 2, on définit la forme bilinéaire $a_i(u,\cdot,\cdot)$ par

$$a_i(u;v,w) = \int_\Omega \big(\rho_i(x,u)\Delta v \Delta w\big)dx, \quad \forall v, w \in H_0^2(\Omega), \tag{3.30}$$

qui satisfait immédiatement les hypothèses $\mathbf{H_1}$ et $\mathbf{H_6}$, grâce à l'injection $H^2(\Omega) \hookrightarrow C(\bar{\Omega})$ (conséquence du théorème d'injection de Sobolev). Pour vérifier l'hypothèse $\mathbf{H_2}$, on remarque tout d'abord par (3.23) que

$$a_1(u;v,v) \geq \beta \|\Delta v\|^2_{L^2(\Omega)}, \forall v \in H_0^2(\Omega).$$

Deuxièmement, comme l'opérateur de Laplace est un isomorphisme de $H^2(\Omega) \cap H_0^1(\Omega)$ dans $L^2(\Omega)$, on a

$$\|v\|_{H^2(\Omega)} \lesssim \|\Delta v\|_{L^2(\Omega)}, \forall v \in H^2(\Omega) \cap H_0^1(\Omega).$$

Ces deux estimations montrent que $\mathbf{H_2}$ a lieu puisque $H_0^2(\Omega)$ est inclus dans $H^2(\Omega) \cap H_0^1(\Omega)$.

Afin de vérifier $\mathbf{H_3}$ et $\mathbf{H_4}$, nous utilisons le résultat suivant.

Lemme 3.2.1. *Pour tout $u \in H_0^2(\Omega)$, on a*

$$D(A_1(u)) = H^4(\Omega) \cap H_0^2(\Omega) \hookrightarrow D(A_2(u)). \tag{3.31}$$

Preuve : Pour un $u \in H_0^2(\Omega)$ fixé, rappelons que

$$D(A_1(u)) = \{v \in H_0^2(\Omega) : A_1(u)v \in L^2(\Omega)\}.$$

On voit donc que $v \in H_0^2(\Omega)$ appartient à $D(A_1(u))$ si et seulement si il existe $f \in L^2(\Omega)$ tel que

$$\Delta(\rho(x,u)\Delta v) = f,$$

au sens des distributions (par souci de brièveté, on écrit $\rho(x,u)$ pour $\rho_1(x,u)$). Par la règle du produit de Leibniz, on obtient de manière équivalente

$$\rho(x,u)\Delta^2 v = \tilde{f} := f - \Delta[\rho(x,u(x))]\Delta v - 2\nabla[\rho(x,u(x))]\cdot\nabla\Delta v. \tag{3.32}$$

La difficulté est de montrer que le membre de droite de cette égalité est dans $L^2(\Omega)$. D'abord par la règle de dérivation en chaîne, on a

$$\begin{aligned}
\partial_i[\rho(x,u(x))] &= \partial_i\rho(x,u(x)) + \partial_u\rho(x,u(x))\partial_i u(x),\\
\partial_{ij}^2[\rho(x,u(x))] &= \partial_{ij}^2\rho(x,u(x)) + \partial_u^2\rho(x,u(x))\partial_i u(x)\partial_j u(x)\\
&\quad + \partial_u\rho(x,u(x))\partial_{ij}^2 u(x),
\end{aligned}$$

pour tout $i,j \in \{1,2\}$. Notons que ces identités impliquent directement que

$$\rho(\cdot,u(\cdot)), \frac{1}{\rho(\cdot,u(\cdot))} \in H^2(\Omega). \tag{3.33}$$

Avec ces identités, pour $i = 1$ ou 2, on voit aussi que

$$\partial_i[\rho(x,u(x))]\partial_i\Delta v = \partial_i\rho(x,u)\partial_i\Delta v + \partial_u\rho(x,u)\partial_i u\partial_i\Delta v, \tag{3.34}$$

et que

$$\Delta[\rho(x,u(x))]\Delta v = \sum_{j=1}^{3} T_j(x,u)\Delta v, \tag{3.35}$$

où on a posé

$$\begin{aligned}
T_1(x,u) &= (\Delta\rho)(x,u),\\
T_2(x,u) &= \partial_u^2\rho(x,u)\{(\partial_1 u)^2 + (\partial_2 u)^2\},\\
T_3(x,u) &= \partial_u\rho(x,u)\Delta u.
\end{aligned}$$

Dans un premier temps, pour $i = 1$ ou 2, on montre que

$$\partial_i[\rho(x,u(x))]\partial_i\Delta v \in H^{-s}(\Omega), \forall s \in (1,2). \tag{3.36}$$

D'après l'identité (3.34), il suffit de montrer que chaque terme de son membre de droite appartient à $H^{-s}(\Omega)$, pour tout $s \in (1,2)$. Pour cela, on remarque que $\partial_i \Delta v \in H^{-1}(\Omega)$, et $\partial_i u \in H^1(\Omega)$. Par les hypothèses de régularité sur ρ, on voit que

$$\partial_i \rho(x,u), \partial_u \rho(x,u)\partial_i u \in H^1(\Omega). \tag{3.37}$$

Comme le Théorème 1.4.4.2 de [34] implique que le produit

$$uv \in H^{-s}(\Omega), \forall s \in (1,2), \tag{3.38}$$

si $u \in H^1(\Omega)$ et $v \in H^{-1}(\Omega)$, on conclut que (3.36) a lieu. De même pour $j = 1, 2$ ou 3, on montre que

$$T_j(x,u)\Delta v \in H^{-s}(\Omega), \forall s \in (1,2). \tag{3.39}$$

Comme $\Delta\rho$ est borné et par la régularité $v \in H^2(\Omega)$, on obtient directement $T_1(x,u)\Delta v \in L^2(\Omega)$, on en déduit donc (3.39) pour $j = 1$. Maintenant par le théorème d'injection de Sobolev, $H^1(\Omega) \hookrightarrow L^p(\Omega)$, pour tout $p \geq 1$, donc par l'inégalité de Hölder, on trouve

$$(\partial_i u)^2 \Delta v \in L^1(\Omega), \forall i = 1, 2,$$

puisque la condition $1 = \frac{2}{p} + \frac{1}{2}$ est vérifiée si $p = 4$. Comme $\partial_u^2 \rho(x,u)$ est borné, on en déduit que $T_2(x,u)\Delta v \in L^1(\Omega)$, ce qui implique (3.39) pour $j = 2$, car $H_0^s(\Omega) \hookrightarrow C(\bar{\Omega})$, grâce au théorème d'injection de Sobolev. Enfin, les régularités $u, v \in H^2(\Omega)$ impliquent que $\Delta u \Delta v \in L^1(\Omega)$ et donc $T_3(x,u)\Delta v \in L^1(\Omega)$, et on en conclut comme pour $j = 2$.

A ce stade de la preuve, par (3.36) et (3.39), on en déduit que $\tilde{f}$ (défini dans (3.32)) appartient à $H^{-s}(\Omega)$, pour tout $s \in (1,2)$. Par la propriété de régularité (3.33) et le Théorème 1.4.4.2 de [34], on conclut que

$$\Delta^2 v = \frac{\tilde{f}}{\rho(\cdot,u)} \in H^{-s}(\Omega), \forall s \in (1,2).$$

Par le Théorème 2 de [14] et le Corollaire 5.12 de [24], on obtient

$$v \in H^{4-s}(\Omega), \forall s \in (1,2),$$

ou de manière équivalente

$$v \in H^{3-\varepsilon}(\Omega), \forall \varepsilon \in (0,1).$$

Cette régularité supplémentaire permet de montrer que

$$\partial_i[\rho(x,u(x))]\partial_i \Delta v \in H^{-\varepsilon'}(\Omega), \forall \varepsilon' \in (0,1), \tag{3.40}$$

pou $i = 1$ ou 2 et

$$T_j(x,u)\Delta v \in H^{-\varepsilon'}(\Omega), \forall \varepsilon' \in (0,1), \tag{3.41}$$

pou $j = 1, 2$ ou 3. Pour la première assertion, par la régularité $\partial_i \Delta v \in H^{-\varepsilon}(\Omega)$, (3.37) et le Théorème 1.4.4.2 de [34], on conclut (3.40) pour $\varepsilon' > \varepsilon$. Pour la deuxième assertion, comme les dérivées secondes de ρ_1 sont bornées, on constate d'abord que

$$T_1(x, u)\Delta v \in L^2(\Omega),$$

d'où (3.41) pour $j = 1$. Pour $j = 2$ on remarque que $(\partial_1 u)^2 + (\partial_2 u)^2$ appartient à $L^2(\Omega)$ en vertu de l'injection $H^1(\Omega) \hookrightarrow L^p(\Omega)$, pour tout $p \geq 1$, et l'inégalité de Hölder. Comme on obtient directement $\Delta u \in L^2(\Omega)$, on conclut que

$$T_j(x, u) \in L^2(\Omega), \text{ pour } j = 2, 3.$$

Encore une fois, le Théorème 1.4.4.2 de [34] induit (3.41) pour $\varepsilon' > \varepsilon$, et $j = 2, 3$.

Par sa définition, on déduit que $\tilde{f}$ appartient à $H^{-\varepsilon'}(\Omega)$, pour tout $\varepsilon' \in (0, 1)$, et par (3.33) et le Théorème 1.4.4.2 de [34] on obtient

$$\Delta^2 v \in H^{-\varepsilon}(\Omega), \forall \varepsilon \in (0, 1).$$

Grâce au Théorème 2 de [14] et au Corollaire 5.12 de [24], on déduit que

$$v \in H^{4-\varepsilon}(\Omega), \forall \varepsilon \in (0, 1).$$

Cette régularité implique que $\partial_i \Delta v \in H^{1-\varepsilon}(\Omega)$, pour tout $\varepsilon \in (0, 1)$ ainsi $\partial_i \Delta v \in L^p(\Omega)$, pour tout $p > 1$, ce qui permet de montrer que

$$\partial_i[\rho(x, u(x))]\partial_i \Delta v \in L^2(\Omega).$$

Plus simplement comme Δv appartient à $H^{2-\varepsilon}(\Omega)$, pour tout $\varepsilon \in (0, 1)$, il est borné dans Ω et par conséquent

$$T_j(x, u)\Delta v \in L^2(\Omega).$$

Cela conduit à la propriété finale

$$\Delta^2 v \in L^2(\Omega),$$

et par le Théorème 2 de [14], on conclut que $v \in H^4(\Omega)$. ∎

Nous avons alors montrer que $D(A_1(u)) \hookrightarrow H^4(\Omega) \cap H_0^2(\Omega)$, ce qui termine la preuve puisque l'injection réciproque est triviale.
A l'aide de ce Lemme, l'hypothèse $\mathbf{H_3}$ est satisfaite avec le choix $D_1 = H^4(\Omega) \cap H_0^2(\Omega)$ et $\tilde{D}_0 = L^2(\Omega)$.

Comme l'espace variationnel est inclus dans $H^2(\Omega)$, l'utilisation d'élément finis de continuité seulement C^0 n'est pas adéquate, car la dérivée seconde de telles fonctions

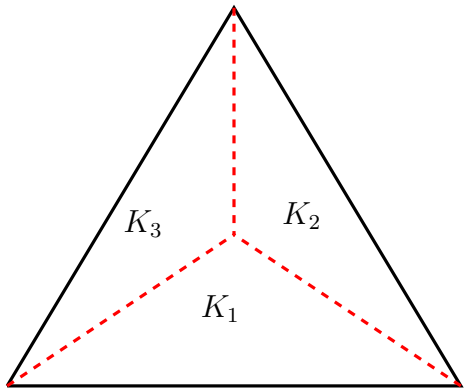


FIGURE 3.18 – L'élément HCT.

n'est en général pas $L^2(\Omega)$, ce qui n'est pas utilisable dans un cadre standard. Nous allons alors utiliser la méthode des éléments finis C^1 "HCT"(Hsieh, Chough et Tocher) décrite dans [20] par exemple, c'est une méthode d'approximation conforme et les éléments HCT sont des macro-éléments (voir Figure 3.18). En effet, chaque triangle K est subdivisé en trois sous-triangles en ajoutant un point interne au triangle (par exemple le barycentre), et l'espace d'approximation V_h est un sous-espace de polynômes de degré 3 sur chaque sous-triangle K_i, convenablement assemblés pour que la fonction discrète sur le triangle $K = \cup_{i=1}^{3} K_i$ appartienne à $C^1(K)$, i.e.,

$$P_K = \{v \in C^1(K) : v\mid_{K_i} \in \mathcal{P}_3(K_i), 1 \leq i \leq 3\},$$

et alors

$$V_h = \{v \in C^1(\Omega) : v\mid_K \in P_K, \forall K \in \mathcal{T}_h, v = \frac{\partial v}{\partial n} = 0 \text{ sur } \partial\Omega\} \subset H_0^2(\Omega).$$

Pour plus de détails, voir [20, p. 341]. Le Théorème 6.1.6 de [20] implique que

$$\|\varphi - P_h\varphi\|_{H^2(\Omega)} \lesssim h^2 \|\varphi\|_{H^4(\Omega)}, \quad \forall \varphi \in H^4(\Omega) \cap H_0^2(\Omega),$$

d'où l'hypothèse $\mathbf{H_4}$ avec $q(1) = 2$.

Le schéma explicite totalement discret d'Euler du problème (3.29) est alors donné par : trouver $U_{n+1,h} \in V_h$ solution de

$$\int_\Omega \{\rho_1(x, U_{n,h})\Delta\big(\frac{U_{n+1,h} - U_{n,h}}{\Delta t}\big)\Delta v_h + \rho_2(x, U_{n,h})\Delta U_{n,h}\Delta v_h\}dx = \int_\Omega f(t_n, U_{n,h})v_h \, dx,$$

pour tout $v_h \in V_h$.

Nous illustrons finalement ce cas en choisissant $\rho_1(x,u) = \frac{1}{2} + \frac{u^2}{1+u^2}$, $\rho_2(x,u) = \frac{u^2}{1+u^2}$, le terme source $f(t,u)$ défini par (3.18) et une donnée initiale nulle $u(0) = 0$. La Figure 3.19 représente la solution totalement discrète $U_{n,h}$ et la Figure 3.20 représente son gradient au

temps final $T = 1$ avec $\Delta t = h = 0.1$ L'ordre de convergence en temps (resp. en espace) est présenté au Tableau 3.18 avec $h = \frac{1}{160}$ (resp. 3.19 avec $\Delta t = \frac{1}{1000}$), où un ordre un en temps et deux en espace est obtenu.

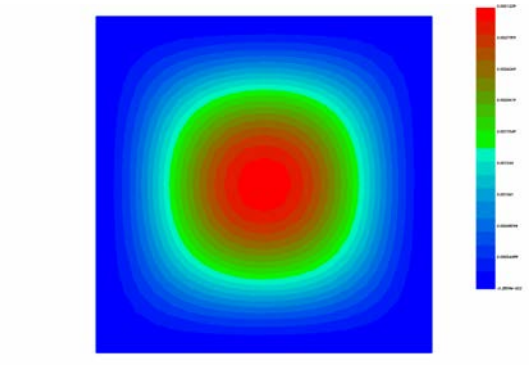


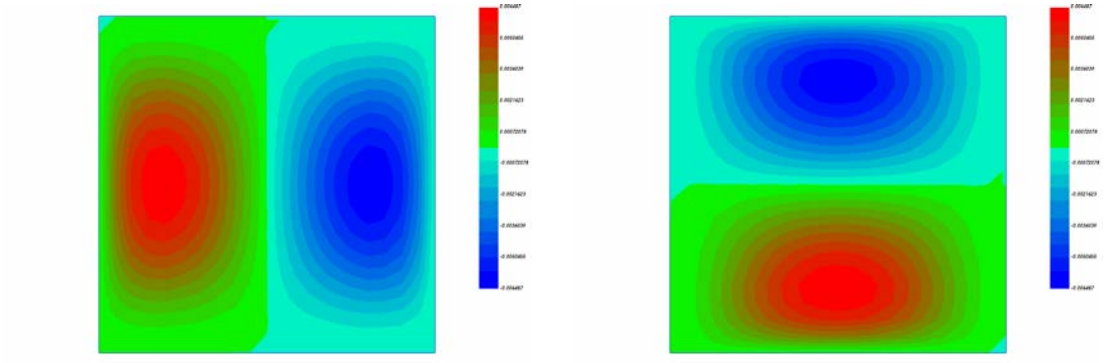
FIGURE 3.19 – La solution approchée $U_{n,h}$ obtenue par le schéma d'Euler (3.8) avec $\Delta t = h = 0.1$ à l'instant final $T = 1$.

Δt	$\|U_{n,h}^{\Delta t} - U_{n,h}^{2\Delta t}\|_\infty$	L'ordre de convergence
0.1	0.00375	1
0.05	0.00184	1
0.025	0.00091	1
0.0125	0.00044	1
0.00625	0.00022	1
0.003125	0.00011	

Tableau 3.18 – L'ordre de convergence pour différents Δt avec $h = 0.00625$.

h	$\|U_{n,h} - U_{n,2h}\|_\infty$	L'ordre de convergence
0.1	0.00106	1.82
0.05	0.0003	1.90
0.025	8.07 e-005	1.92
0.0125	2.12 e-005	1.98
0.00625	5.36 e-006	

Tableau 3.19 – L'ordre de convergence pour différents h avec $\Delta t = 0.001$.



(a) $\partial_1 u$ (b) $\partial_2 u$

FIGURE 3.20 – Le gradient de la solution approchée, avec $\Delta t = h = 0.1$ à l'instant final $T = 1$.

3.3 Applications pour une équation de Sobolev singulièrement perturbée

Dans cette section, nous nous intéressons à un cas particulier du problème étudié au Chapitre 2. Nous nous intéressons notamment au cas où $A_1 = A_2$. On considère alors le problème de Sobolev suivant

$$\begin{cases} u_{\epsilon,t} - \epsilon\Delta u_{\epsilon,t} - \Delta u_\epsilon = f, \text{ dans } \Omega \times (0,T), \\ u_\epsilon = 0, \qquad\qquad \text{sur } \partial\Omega \times (0,T), \\ u_\epsilon(0) = u_0, \qquad\qquad \text{dans } H_0^1(\Omega), \end{cases} \tag{3.42}$$

où ϵ est une constante positive et $\Omega = (0,1)^2$.
Dans [44, 51], les auteurs proposent l'étude de cette perturbation, via le changement d'inconnue $v = (I - \Delta)u_\epsilon$.

Dans ce cas, la forme bilinéaire symétrique a_i, $i = 1$ ou 2 est définie par

$$a_i(u,v) = \big(\nabla u, \nabla v\big)_{L^2(\Omega)}, \quad \forall u, v \in H_0^1(\Omega).$$

Cette forme bilinéaire symétrique est continue sur $H_0^1(\Omega) \times H_0^1(\Omega)$ et est coercive sur $H_0^1(\Omega)$. Par conséquent, pour $i = 1, 2$, l'opérateur A_i associé à la forme a_i

$$\langle A_i u, v\rangle_{V',V} = a_i(u,v), \quad \forall u, v \in H_0^1(\Omega),$$

est continu de $H_0^1(\Omega)$ dans son dual $H^{-1}(\Omega)$, et les hypothèses de l'introduction du Chapitre 2 sont satisfaites avec $V = H_0^1(\Omega)$ et $H = L^2(\Omega)$.

Pour le schéma totalement discret, on utilise d'une part, comme dans les sections précédentes, la méthode des éléments finis (qui est basée sur un maillage de Ω) pour la

discrétisation en espace et on choisit l'espace V_h défini par (3.7) comme espace d'approximation. D'autre part, pour la discrétisation temporelle, on utilise comme dans l'étude théorique du Chapitre 2, le schéma d'Euler implicite ou le schéma de Crank-Nicolson.

- Le schéma d'Euler implicite qui consiste à trouver $u_{n+1,h}$ par

$$\Big(\frac{u_{n+1,h}-u_{n,h}}{\Delta t},\varphi_h\Big)_{L^2(\Omega)}+\epsilon a_1\Big(\frac{u_{n+1,h}-u_{n,h}}{\Delta t},\varphi_h\Big)+a_2(u_{n+1,h},\varphi)=(f(t_n),\varphi_h)_{L^2(\Omega)}, \tag{3.43}$$

  pour tout $\varphi_h \in V_h$.

- Le schéma de Crank-Nicolson qui consiste à trouver $u_{n+1,h}$ par

$$\begin{aligned}\Big(\frac{u_{n+1,h}-u_{n,h}}{\Delta t},\varphi_h\Big)_{L^2(\Omega)}+\epsilon a_1\Big(\frac{u_{n+1,h}-u_{n,h}}{\Delta t},\varphi_h\Big)+\frac{1}{2}a_2(u_{n+1,h}-u_{n,h},\varphi_h)\\ =\frac{1}{2}(f(t_n)+f(t_{n+1}),\varphi_h)_{L^2(\Omega)},\end{aligned} \tag{3.44}$$

  pour tout $\varphi_h \in V_h$.

On sait bien que, si la famille $\{V_h\}_h$ basée sur une famille de triangulations uniformément régulière et V_h est l'espace composé de fonctions polynomiales par morceaux, alors il existe une constante positive C indépendante de h telle que

$$\|\chi\|_{H_0^1(\Omega)} \leq Ch^{-1}\|\chi\|_{L^2(\Omega)}, \quad \forall \chi \in V_h.$$

Pour plus de détails voir par exemple [20]. Ainsi l'hypothèse $\mathbf{H_2'}$ est satisfaite. Dans la preuve du Théorème 2. 2 de [71, page 30], on a

$$\|A_{i,h}^{-1}P_hu - A_i^{-1}u\|_{L^2(\Omega)} \lesssim h^2\|u\|_{L^2(\Omega)}, \quad \forall u \in L^2(\Omega),$$

d'où l'hypothèse $\mathbf{H_3'}$. Pour les illustrations numériques, on prend comme solution exacte

$$u_\epsilon(t,x,y) = (\sqrt{\epsilon}+1)\sin(t)\sin(\pi x)\sin(\pi y).$$

Le Tableau 3.20 (ou la Figure 3.21) montre l'évolution de l'erreur en norme $L^2(\Omega)$ pour différentes valeurs de ϵ avec un pas de temps Δt et un pas de maillage h suffisamment petits ($\Delta t = 0.001$ et $h = 0.00625$). Ces résultats montrent que l'erreur en norme $L^2(\Omega)$ ne dépend pas de ϵ. Pour un ϵ fixé, le Tableau 3.21 (resp. 3.22) représente l'évolution de l'erreur par le schéma d'Euler implicite pour différents pas de temps (resp. pour différents pas de maillage) avec un pas de maillage $h = 0.00625$ (resp. avec un pas de temps $\Delta t = 0.001$). Le Tableau 3.23 montre l'évolution de l'erreur par le schéma de Crank-Nicolson en norme $L^2(\Omega)$ pour différents pas de temps avec un pas de maillage $h = 0.00625$.

L'erreur en norme $L^2(\Omega)$ est donc d'ordre un (resp. deux) en temps pour le schéma d'Euler implicite (resp. de Crank-Nicolson) et deux en espace, ce qui est en accord avec les résultats théoriques du Chapitre 2.

ϵ	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
$\|u_\epsilon(t_N) - u_{N,h}\|_\infty$	0.0639	0.0529	0.0495	0.0484	0.0481	0.048	0.0479	0.0479

Tableau 3.20 – L'ordre de convergence pour différents Δt avec $h = 0.00625$.

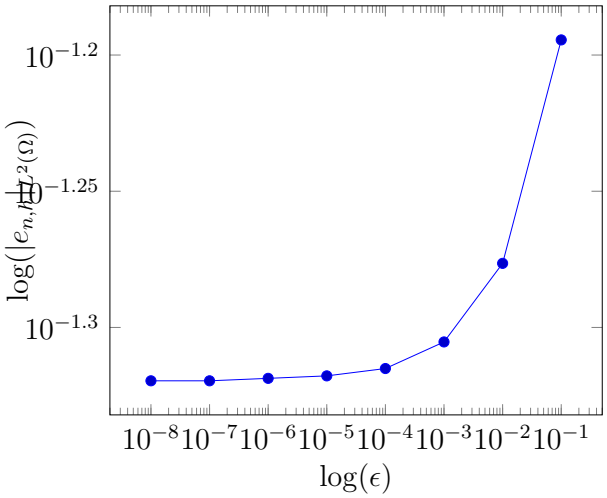


FIGURE 3.21 – L'évolution de l'erreur en norme $L^2(\Omega)$ pour différentes valeurs de ϵ avec $\Delta t = 0.0001$ et $h = 0.00625$.

Δt	$\|u_\epsilon(t_N) - u_{N,h}\|_{L^2(\Omega)}$
0.1	0.00103
0.05	0.00052
0.025	0.0002
0.0125	0.0001
0.00625	0.00005

Tableau 3.21 – L'évolution de l'erreur en norme $L^2(\Omega)$ par le schéma d'Euler implicite (3.43) au temps final $T = 1$ pour différents Δt avec $h = 0.00625$ et $\epsilon = 10^{-3}$.

h	$\|u_\epsilon(t_N) - u_{N,h}\|_{L^2(\Omega)}$
0.1	0.00091
0.05	0.00022
0.025	5.72 e-005
0.0125	1.43 e-005
0.00625	3.57 e-006

Tableau 3.22 – L'évolution de l'erreur en norme $L^2(\Omega)$ par le schéma d'Euler (3.43) au temps final $T = 0.1$ pour différents h avec $\Delta t = 0.001$ et $\epsilon = 10^{-3}$.

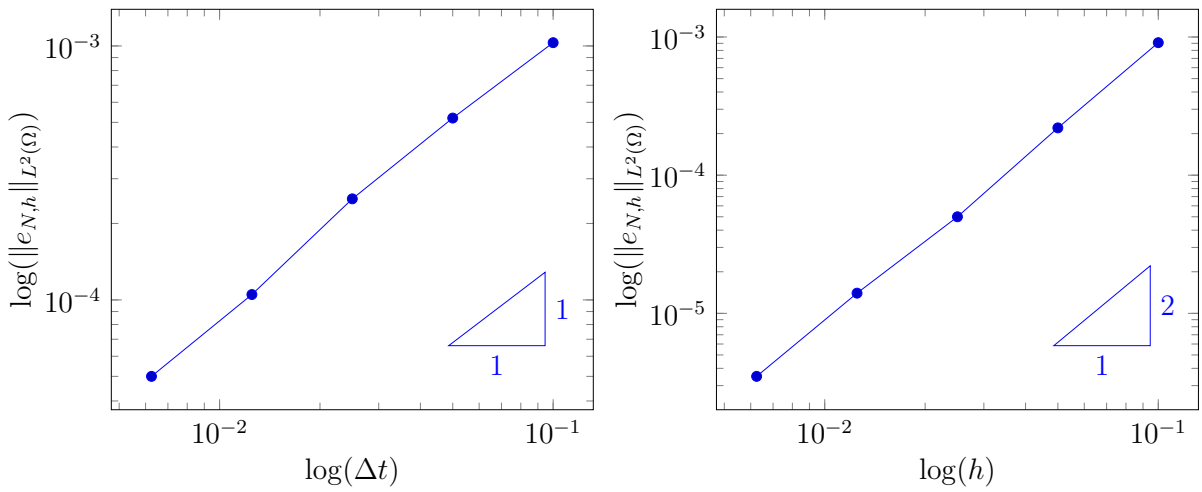


FIGURE 3.22 – Erreur en norme $L^2(\Omega)$ en espace au temps final $T = 0.1$ avec un maillage uniforme et un maillage raffiné pour $\Delta t = 0.0001$.

Δt	$\|u_\epsilon(t_N) - u_{N,h}\|_{L^2(\Omega)}$
1	0.00341
0.5	0.00087
0.25	0.000193
0.125	3.76e-005
0.0625	8.598e-006

Tableau 3.23 – L'évolution de l'erreur en norme $L^2(\Omega)$ par le schéma de Crank-Nicolson (3.44) au temps final $T = 10$ pour différents Δt avec $h = 0.00625$ et $\epsilon = 10^{-3}$.

Chapitre 4

Analyse a posteriori pour un problème de Sobolev

Résumé

On considère une équation de type Sobolev et on s'intéresse à la dérivation d'estimations d'erreur a posteriori pour la discrétisation de cette équation par la méthode des éléments finis conforme en espace et le schéma d'Euler implicite en temps. A la fin de ce chapitre, on présente quelques tests numériques.

Sommaire

4.1 Le problème semi-discret et totalement discret **111**
4.1.1 La discrétisation en temps par le schéma d'Euler implicite . . 111
4.1.2 La discrétisation totale 111
4.2 Quelques outils d'analyse **115**
4.3 Analyse d'erreur a posteriori pour le problème semi-discrétisé **116**
4.4 Analyse a posteriori de la discrétisation spatiale **119**
4.4.1 Borne supérieure de l'erreur 120
4.4.2 Borne inférieure de l'erreur 123
4.5 Analyse a posteriori pour le problème totalement discrétisé **125**
4.6 Résultats numériques **127**
4.6.1 Dépendance de l'erreur 130
4.6.2 Algorithme adaptatif 132

Soit Ω un ouvert borné de $\mathbb{R}^d$, $d = 2$ ou 3, à frontière Γ polygonale ($d = 2$) ou polyédrale ($d = 3$). Soit T un nombre réel positif fixé.

Pour $i = 1, 2$, soit L_i l'opérateur elliptique d'ordre 2

$$L_i(x, D_x)u = -\sum_{k,\ell=1}^{d} \partial_k(a_{k,\ell}^{(i)}(x)\partial_\ell u) + \sum_{k=1}^{d} b_k^{(i)}(x)\partial_k u + c^{(i)}(x)u,$$

où $a_{k,\ell}^{(i)} = a_{\ell,k}^{(i)}, b_k^{(i)}, c^{(i)} \in L^\infty(\Omega)$ et on introduit la forme bilinéaire

$$a_i(u,v) = \int_\Omega \left(\sum_{k,\ell=1}^{d} a_{k,\ell}^{(i)}(x)\partial_\ell u \partial_k v + \sum_{k=1}^{d} b_k^{(i)}(x)\partial_k uv + c^{(i)}(x)uv \right) dx, \forall u, v \in H_0^1(\Omega).$$

On suppose que a_1 et a_2 sont symétriques, et a_2 est non négative, i.e.,

$$a_2(u,u) \geq 0, \forall u \in H_0^1(\Omega), \tag{4.1}$$

et que a_1 est coercive sur $H_0^1(\Omega)$, c'est à dire il existe $\alpha > 0$ tel que

$$a_1(u,u) \geq \alpha \|u\|_{1,\Omega}^2, \forall u \in H_0^1(\Omega). \tag{4.2}$$

Dans ce chapitre, on considère l'équation de Sobolev suivante : trouver u solution du problème

$$\begin{cases} L_1 u_t + L_2 u = f \text{ dans } \Omega \times (0,T), \\ u(\cdot,t) = 0 \quad \text{sur } \partial\Omega \times (0,T), \\ u(\cdot,0) = u_0 \quad \text{dans } \Omega, \end{cases} \tag{4.3}$$

où u_t est la dérivée de u par rapport à t. On suppose que la fonction f satisfait $f \in L^2(0,T;H^{-1}(\Omega))$ et la donnée initiale $u_0 \in H_0^1(\Omega)$. Sous ces hypothèses, le problème (4.3) ou de manière équivalente

$$a_1(u_t(t),v) + a_2(u(t),v) = (f(t),v), \forall v \in H_0^1(\Omega), \forall \text{ a.e. } t \in (0,T), \tag{4.4}$$

admet une unique solution (faible) dans $C([0,T];H_0^1(\Omega))$, voir Théorème 1.1.1 du Chapitre 1. Sans perte de généralité, on peut supposer que a_2 est aussi coercive sur $H_0^1(\Omega)$, en effet par le changement d'inconnu

$$\tilde{u}(\cdot,t) = e^{-\lambda t}u(\cdot,t),$$

pour un nombre réel positif λ, il est clair que le problème (4.4) est équivalent à

$$a_1(\tilde{u}_t(t),v) + \tilde{a}_2(u(t),v) = (e^{-\lambda t}f(t),v), \forall v \in H_0^1(\Omega), \forall \text{ a.e. } t \in (0,T), \tag{4.5}$$

où

$$\tilde{a}_2(u,v) = a_2(u,v) + \lambda a_1(u,v), \forall u, v \in H_0^1(\Omega),$$

est clairement coercive sur $H_0^1(\Omega)$ grâce à (4.1)-(4.2). Ainsi, nous supposons aussi dans ce qui suit que a_2 est coercive et on note par

$$\|u\|_i = \sqrt{a_i(u,u)}, \forall u \in H_0^1(\Omega),$$

deux normes équivalentes de $H_0^1(\Omega)$. On munit $H^{-1}(\Omega)$ de la norme duale $\|\cdot\|_{-1}$ obtenue en utilisant la dualité avec la deuxième norme de $H_0^1(\Omega)$, en d'autres termes

$$\|g\|_{-1} = \sup_{v \in H_0^1(\Omega), v \neq 0} \frac{|\langle g; v \rangle|}{\|v\|_2}, \quad \forall\, g \in H^{-1}(\Omega).$$

4.1 Le problème semi-discret et totalement discret

4.1.1 La discrétisation en temps par le schéma d'Euler implicite

On suppose maintenant que $f \in C([0,T]; H^{-1}(\Omega))$. On considère une suite strictement croissante de temps discrets noté $\{t_p\}_{0\leq p\leq N}$ tel que $t_0 = 0$ et $t_N = T$ et on définit, pour tout $1 \leq p \leq N$, l'intervalle discret $I_p := (t_{p-1}, t_p]$ et le pas de temps $\tau_p := t_p - t_{p-1}$. On note par $\tau = \max_{1\leq p\leq N} \tau_p$ le pas global de discrétisation en temps.

L'approximation semi-discrète du problème continu (4.3) par le schéma d'Euler implicite consiste à trouver une suite $(u^p)_{0\leq p\leq N}$ solution de

$$\begin{cases} L_1\left(\dfrac{u^p - u^{p-1}}{\tau_p}\right) + L_2 u^p = f^p \text{ dans } \Omega & 1 \leq p \leq N, \\ u^p = 0 \qquad\qquad\qquad\qquad \text{sur } \partial\Omega & 1 \leq p \leq N, \\ u^0 = u_0 \qquad\qquad\qquad\qquad \text{dans } \Omega, \end{cases} \tag{4.6}$$

avec $f^p = f(\cdot, t_p)$. Ce problème admet une unique solution faible $u^p \in H_0^1(\Omega)$, dont la formulation variationnelle est

$$a_1(\frac{u^p - u^{p-1}}{\tau_p}, v) + a_2(u^p, v) = \int_\Omega f^p v, \forall v \in H_0^1(\Omega), \tag{4.7}$$

ou par équivalence

$$a_1(u^p, v) + \tau_n a_2(u^p, v) = a_1(u^{p-1}, v) + \tau_p \int_\Omega f^p v, \forall v \in H_0^1(\Omega). \tag{4.8}$$

L'existence et l'unicité de la solution du problème variationnel (4.8) est une conséquence directe du lemme de Lax-Milgram.

Remarque 4.1.1. *L'analyse d'erreur a posteriori que nous effectuons ci-dessous pour le système (4.8) est immédiatement applicable au schéma d'Euler explicite*

$$a_{1,p}(\frac{u^p - u^{p-1}}{\tau_p}, v) + a_2(u^p, v) = \int_\Omega f^{p-1} v, \forall v \in H_0^1(\Omega),$$

où

$$a_{1,p}(u, v) = a_1(u, v) - \tau_p a_2(u, v),$$

qui est uniformément coercive en p, si τ_p est assez petit.

4.1.2 La discrétisation totale

On discrétise maintenant le problème (4.8) en espace. Pour cela, on définit un ensemble de maillages $\{T_{ph}\}_{0\leq p\leq N}$ de Ω correspondant aux différents pas de temps $\{t_p\}_{0\leq p\leq N}$. Ces

maillages sont tels que, pour tout $0 \leq p \leq N$, $\bar{\Omega} = \bigcup_{K \in T_{ph}} K$. Les éléments du maillage sont par convention des fermés et ils sont tels que l'intersection des intérieurs de deux éléments distincts est vide. Pour simplifier, on suppose que les maillages sont triangulaires et conformes (dans le sens où ils ne contiennent pas des " nœuds orphelins", voir Figure 4.1).

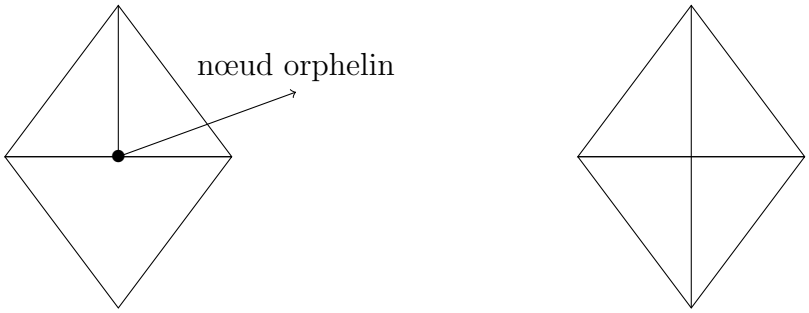


FIGURE 4.1 – à gauche : maillage non conforme en 2d, à droite : maillage conforme en 2d.

Ces maillages sont supposés réguliers. Cela signifie que, pour tout $0 \leq p \leq N$, si h_K désigne le diamètre d'un élément $K \in T_{ph}$ et ρ_k désigne le diamètre de la boule inscrite dans K, alors il existe une constante k indépendante des maillages $\{T_{ph}\}_{0 \leq p \leq N}$ telle que

$$\frac{h_K}{\rho_K} \leq k, \quad \forall K \in T_{ph}.$$

Le maillage initial T_{0h} désigne le maillage utilisé pour approcher la condition initiale alors que les maillages T_{ph}, pour tout $0 \leq p \leq N$, désignent les maillages utilisés pour passer du temps t_{p-1} au temps t_p. Pour chaque $0 \leq p \leq N$, on associe à chaque maillage T_{ph} son pas noté $h_p = \max_{K \in T_{ph}} h_K$.

Pour tout $K \in T_{ph}$, on désigne par $\mathcal{E}_K$ l'ensemble de ses arêtes/faces. Ensuite, on désigne par $\mathcal{E}_{ph}$ l'ensemble de toutes les arêtes/faces de T_{ph}, par $\mathcal{E}_{ph}^{\text{int}}$ l'ensemble des arêtes/faces situées à l'intérieur de T_{ph}. Enfin, pour deux éléments K et $L \in T_{ph}$ et pour une arête/face $E \in \mathcal{E}_K \cap \mathcal{E}_L$ on désigne par

$$h_E = \frac{d}{2}\left(\frac{|K|}{|E|} + \frac{|L|}{|E|}\right),$$

sa hauteur moyenne.

Pour tout $0 \leq p \leq N$, l'espace des éléments finis conforme associé au maillage T_{ph} est noté V_{ph} et est défini par :

$$V_{ph} = \{v \in H_0^1(\Omega) : v_{|K} \in \mathbb{P}_1, \forall K \in T_{ph}\} \subset H_0^1(\Omega).$$

L'approximation totalement discrète du problème (4.3) en utilisant le schéma d'Euler implicite et la méthode des éléments finis conforme est donnée par : étant donné une approximation $u_h^0 \in V_{0h}$ de u_0, trouver $u_h^p \in V_{ph}$, $1 \leq p \leq N$, tel que

$$a_1(\frac{u_h^p - u_h^{p-1}}{\tau_p}, v_h) + a_2(u_h^p, v_h) = \int_\Omega f^p v_h, \quad \forall v_h \in V_{ph}, \tag{4.9}$$

ou de manière équivalente

$$a_1(u_h^p, v_h) + \tau_p a_2(u_h^p, v_h) = a_1(u_h^{p-1}, v_h) + \tau_p \int_\Omega f^p v_h, \quad \forall v_h \in V_{ph}. \tag{4.10}$$

Définition 4.1.2. *Soit u^p la solution du problème (4.8) et u_h^p la solution de (4.10), alors on désigne l'erreur spatiale par*

$$e^p = u^p - u_h^p.$$

Rappelons maintenant quelques notations utiles et propriétés utilisées ci-dessous.

Pour une arête/face E, on désigne le vecteur normal orienté vers l'extérieur par n_E. Étant donné une arête/face intérieure E, nous choisissons une direction normale arbitraire n_E et on désigne par K_{int} et K_{ext} les deux éléments partageant cette arête/face. Dans toute la suite on suppose que n_E pointe vers K_{ext} comme dans la figure 4.2.

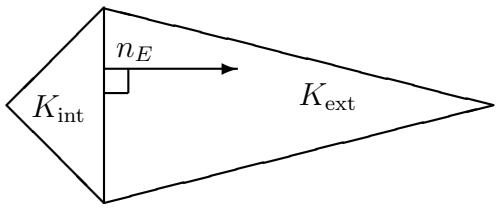


FIGURE 4.2 – Exemple de deux éléments partageant le bord E et de normale n_E.

Le saut d'une fonction v à travers une arête/face E à un point x est défini par

$$[\![v(x)]\!]_E = \begin{cases} \lim\limits_{\alpha \to 0^+} (v(x + \alpha n_E) - v(x - \alpha n_E) & \text{si } E \in \mathcal{E}_{ph}^{\text{int}}, \\ v(x) & \text{si } E \in \mathcal{E}_{ph} \setminus \mathcal{E}_{ph}^{\text{int}}. \end{cases}$$

Notons que le signe de $[\![v(x)]\!]_E$ dépend de l'orientation de n_E. Toutefois, une quantité comme un saut de gradient $[\![\nabla v \cdot n_E]\!]_E$ est indépendante de cette orientation.

Dans la suite, nous utiliserons les patchs locaux : pour un élément K, on définit le patch de K (noté ω_K) par l'ensemble de tous les éléments qui ont une arête/face commune avec K (voir la Figure 4.3). Pour une arête/face E, soit ω_E l'union des deux éléments ayant E comme arête/face et enfin pour un nœud x, on note par ω_x l'union des deux éléments ayant x comme nœud. De même, on note par $\tilde{\omega}_K$ (resp. $\tilde{\omega}_E$) l'union de tous les triangles partageant un nœud avec K (resp. E).

On définit maintenant l'opérateur d'interpolation de Clément que l'on note I_C^0. Pour cela on introduit les notations suivantes :

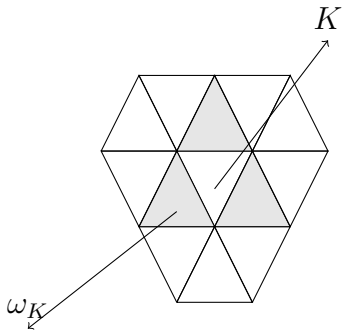


FIGURE 4.3 – Patch d'un élément K.

- On note par $\mathcal{N}_{ph}$ l'ensemble des nœuds de la triangulation T_{ph} et par $\mathcal{N}_{ph}^{\text{int}}$ l'ensemble des nœuds intérieurs de la triangulation T_{ph}. Pour chaque nœud $x \in \mathcal{N}_{ph}^{\text{int}}$

- on note plus loin par λ_x la fonction chapeau standard associée à x, c'est à dire $\lambda_x \in V_{ph}$ et satisfait

$$\lambda_x(y) = \delta_{x,y}, \forall y \in \mathcal{N}_{ph}^{\text{int}}.$$

Finalement, l'opérateur d'interpolation de Clément $I_C^0 : L^2(\Omega) \to V_{ph}$ est défini par :

$$I_C^0 w = \sum_{x \in \mathcal{N}_{nh}^{int}} |\omega_x|^{-1} \left(\int_{\omega_x} w \right) \lambda_x. \tag{4.11}$$

On rappelle quelques inégalités basées sur l'opérateur d'interpolation de Clément qui ont été démontrées dans [21].

Lemme 4.1.3. *Pour tout $w \in H_0^1(\Omega)$, on a*

$$\|w - I_C^0 w\|_K \lesssim h_K \|\nabla w\|_{\tilde{\omega}_K}, \forall K \in T_{ph}, \tag{4.12}$$

$$\|w - I_C^0 w\|_E \lesssim h_E^{1/2} \|\nabla w\|_{\tilde{\omega}_E}, \forall E \in \mathcal{E}_{ph}^{\text{int}}, \tag{4.13}$$

$$\|\nabla I_C^0 w\|_K \lesssim \|\nabla w\|_{\tilde{\omega}_K}, \forall K \in T_{ph}. \tag{4.14}$$

Soit $1 \leq p \leq N$, et $K \in T_{ph}$. On définit le résidu d'élément par

$$R_K^p = f(\cdot, t_p) - L_1 \left(\frac{u_h^p - u_h^{p-1}}{\tau_p} \right) - L_2 u_h^p,$$

tandis que si $E \in \mathcal{E}_{ph}^{\text{int}}$, alors le résidu d'arête (ou de face) est

$$J_{E,n}^p = \left[\!\left[\left(A_1 \nabla \left(\tfrac{u_h^p - u_h^{p-1}}{\tau_p} \right) + A_2 \nabla u_h^p \right) \cdot n_E \right]\!\right]_E,$$

où pour $i = 1$ ou 2, $A_i(d \times d)$ est une matrice symétrique donnée par

$$A_i = \left(a_{k,\ell}^{(i)} \right)_{1 \leq k,\ell \leq d}.$$

4.2 Quelques outils d'analyse

Dans cette section, on donne quelques propriétés satisfaites par l'erreur spatiale e^p que nous utiliserons dans la preuve des estimations d'erreur spatiale.

Lemme 4.2.1 (Orthogonalité de Galerkin)**.** *L'erreur e^p satisfait la relation d'orthogonalité de Galerkin*

$$a_1(\frac{e^p - e^{p-1}}{\tau_p}, v_h) + a_2(e^p, v_h) = 0, \forall v_h \in V_{ph}. \tag{4.15}$$

Preuve : Il suffit de soustraire (4.7) avec $v = v_h \in V_{ph}$ à l'identité (4.9).

Lemme 4.2.2. *L'erreur e^p satisfait*

$$a_1(\frac{e^p - e^{p-1}}{\tau_p}, v) + a_2(e^p, v) = \sum_{K\in T_{ph}} \int_K R_K^p v + \sum_{E\in\mathcal{E}_{ph}^{\text{int}}} \int_E J_{E,n}^p v, \forall v \in H_0^1(\Omega). \tag{4.16}$$

Preuve : Nous remarquons tout d'abord que

$$\begin{aligned} a_1(\frac{e^p - e^{p-1}}{\tau_p}, v) + a_2(e^p, v) = a_1(\frac{u^p - u^{p-1}}{\tau_p}, v) + a_2(u^p, v) \\ - \left(a_1(\frac{u_h^p - u_h^{p-1}}{\tau_p}, v) + a_2(u_h^p, v) \right). \end{aligned}$$

Nous transformons le premier terme en utilisant (4.7) et le second par intégration par parties élément par élément, rappelons que

$$a_i(u, v) = \sum_{K\in T_{ph}} \int_K \left((A_i \nabla u \cdot \nabla v + \sum_{k=1}^d b_k^{(i)} \partial_k u v + c^{(i)} u v \right).$$

Cela permet de conclure.

Les lemmes ci-dessus nous permettent de prouver le

Lemme 4.2.3. *On a l'identité suivante*

$$\begin{aligned} a_1(e^p, e^p) + \tau_p a_2(e^p, e^p) = a_1(e^{p-1}, e^p) + \tau_p \sum_{K\in T_{ph}} \int_K R_K^p (e^p - I_C^0 e^p) \\ + \tau_p \sum_{E\in\mathcal{E}_{ph}^{\text{int}}} \int_E J_{E,n}^p (e^p - I_C^0 e^p). \end{aligned} \tag{4.17}$$

Preuve : On écrit

$$a_2(e^p, e^p) = a_2(e^p, e^p - I_C^0 e^p) + a_2(e^p, I_C^0 e^p),$$

alors on transforme le premier terme en utilisant (4.16) avec $v = e^p - I_C^0 e^p$ et le deuxième terme en utilisant la relation d'orthogonalité de Galerkin (4.15) avec $v_h = I_C^0 e^p$.

4.3 Analyse d'erreur a posteriori pour le problème semi-discrétisé

Inspiré par [41, 60, 11, 13, 56], qui ont considéré l'équation de la chaleur, on définit l'indicateur d'erreur en temps par

$$\eta_t^p = \tau_p^{1/2} \|u_h^p - u_h^{p-1}\|_2, 1 \leq p \leq N. \tag{4.18}$$

La seule différence avec les papiers cités ci-dessus repose sur la norme choisie de $u_h^p - u_h^{p-1}$.

Par souci de brièveté, on introduit les notations suivantes :

- $\pi_\tau f$ désigne la fonction en escalier qui est constante et est égale $f(t_p)$ sur chaque intervalle (t_{p-1}, t_p), $1 \leq p \leq N$.
- Pour une suite $v^p \in V_{ph} \oplus H_0^1(\Omega)$, $0 \leq p \leq N$, on note v_τ l'interpolant de "Lagrange" qui est affine sur chaque intervalle $[t_{p-1}, t_p]$, $1 \leq p \leq N$, et égale à v^p en t_p, i.e., défini par,

$$v_\tau(t) = \frac{t_p - t}{\tau_p} v^{p-1} + \frac{t - t_{p-1}}{\tau_p} v^p, \forall t \in [t_{p-1}, t_p], 1 \leq p \leq N.$$

- Enfin, on note $e_\tau = u - u_\tau$, l'erreur temporelle.

Comme

$$\partial_t u_\tau = \frac{u^p - u^{p-1}}{\tau_p} \text{ sur } (t_{p-1}, t_p),$$

l'équation semi-discrète (4.8) est alors équivalente à

$$a_1(\partial_t u_\tau(t), v) + a_2(u^p, v) = (f^p, v), \forall v \in H_0^1(\Omega), \forall t \in (t_{p-1}, t_p). \tag{4.19}$$

En prenant la différence avec (4.4), on dérive l'équation résiduelle

$$\begin{aligned} a_1(\partial_t e_\tau(t), v) + a_2(e_\tau(t), v) &= ((f - \pi_\tau f)(t), v) \\ &+ a_2((u^p - u_\tau)(t), v), \forall v \in H_0^1(\Omega), \forall \text{ a.e. } t \in (t_{p-1}, t_p). \end{aligned} \tag{4.20}$$

Cette identité permet de prouver l'estimation d'erreur suivante

Théorème 4.3.1 (Borne supérieure de l'erreur en temps)**.** *On a l'estimation suivante*

$$\begin{aligned} \|e_\tau(t_n)\|_1^2 + \int_0^{t_n} \|e_\tau(s)\|_2^2 \, ds &\lesssim \sum_{p=1}^{n} (\eta_t^p)^2 \\ &+ \int_0^{t_n} \|(u_\tau - u_{h\tau})(s)\|_2^2 \, ds + \|f - \pi_\tau f\|^2_{L^2(0,t_n;H^{-1}(\Omega))}. \end{aligned} \tag{4.21}$$

Preuve : L'équation résiduelle (4.20) avec $v = e_\tau(t)$ donne

$$a_1(\partial_t e_\tau(t), e_\tau(t)) + a_2(e_\tau(t), e_\tau(t)) = ((f - \pi_\tau f)(t), e_\tau(t)) + a_2((u^p - u_\tau)(t), e_\tau(t)), \forall \text{ a.e. } t \in (t_{p-1}, t_p).$$

Comme a_1 est symétrique, on a

$$a_1(\partial_t e_\tau(t), e_\tau(t)) = \frac{1}{2}\partial_t a_1(e_\tau(t), e_\tau(t)),$$

En intégrant l'identité ci-dessus en temps sur l'intervalle (t_{n-1}, t_n), on trouve

$$\frac{1}{2}a_1(e_\tau(t_p), e_\tau(t_p)) - \frac{1}{2}a_1(e_\tau(t_{p-1}), e_\tau(t_{p-1})) + \int_{t_{p-1}}^{t_p} a_2(e_\tau(t), e_\tau(t)) = \int_{t_{p-1}}^{t_p} ((f - \pi_\tau f)(t), e_\tau(t))\, dt + \int_{t_{p-1}}^{t_p} a_2((u^p - u_\tau)(t), e_\tau(t))\, dt.$$

En utilisant l'inégalité de Young, on obtient

$$\frac{1}{2}a_1(e_\tau(t_p), e_\tau(t_p)) - \frac{1}{2}a_1(e_\tau(t_{p-1}), e_\tau(t_{p-1})) + \frac{1}{2}\int_{t_{p-1}}^{t_p} a_2(e_\tau(t), e_\tau(t)) \leq \int_{t_{p-1}}^{t_p} \|(f - \pi_\tau f)(t)\|_{-1}^2\, dt + \int_{t_{p-1}}^{t_p} \|(u^p - u_\tau)(t)\|_2^2\, dt. \tag{4.22}$$

On estime maintenant le second terme du membre de droite de cette inégalité. D'abord par définition de u_τ, il est clair que

$$\int_{t_{p-1}}^{t_p} \|(u^p - u_\tau)(t)\|_2^2\, dt = \frac{\tau_p}{3}\|u^p - u^{p-1}\|_2^2. \tag{4.23}$$

Deuxièmement, en utilisant l'inégalité triangulaire, on écrit simplement

$$\tau_p^{1/2}\|u^p - u^{p-1}\|_2 \leq \eta_t^p + \tau_p^{1/2}\|u^p - u_h^p\|_2 + \tau_p^{1/2}\|u_h^{p-1} - u^{p-1}\|_2. \tag{4.24}$$

Montrons maintenant que

$$\tau_p\|u^p - u_h^p\|^2 + \tau_p\|u^{p-1} - u_h^{p-1}\|_2^2 \leq 6\int_{t_{p-1}}^{t_p} \|(u_\tau - u_{h\tau})(t)\|_2^2\, dt. \tag{4.25}$$

En effet, par définition, on a

$$(u_\tau - u_{h\tau})(t) = \frac{t_p - t}{\tau_p}(u^{p-1} - u_h^{p-1}) + \frac{t - t_{p-1}}{\tau_p}(u^p - u_h^p), \forall t \in [t_{p-1}, t_p],$$

et alors

$$\|(u_\tau - u_{h\tau})(t)\|_2^2 = \left(\frac{t_p - t}{\tau_p}\right)^2 \|u^{p-1} - u_h^{p-1}\|_2^2 + \left(\frac{t - t_{p-1}}{\tau_p}\right)^2 \|u^p - u_h^p\|_2^2 + 2\frac{(t_p - t)(t - t_{p-1})}{\tau_p^2} a_2(u^{p-1} - u_h^{p-1}, u^p - u_h^p), \forall t \in (t_{p-1}, t_p).$$

En intégrant cette expression en $t \in (t_{p-1}, t_p)$, on trouve après des calculs simples

$$\int_{t_{p-1}}^{t_p} \|(u_\tau - u_{h\tau})(t)\|_2^2 \, dt = \frac{\tau_p}{3}(\|u^{p-1} - u_h^{p-1}\|_2^2 + \|u^p - u_h^p\|_2^2 + a_2(u^{p-1} - u_h^{p-1}, u^p - u_h^p)).$$

L'inégalité de Cauchy-Schwarz permet de conclure (4.25). Par conséquent, l'identité (4.23) et les estimations (4.24), (4.25) impliquent que

$$\int_{t_{p-1}}^{t_p} \|(u^p - u_\tau)(t)\|_2^2 \, dt \lesssim (\eta_t^p)^2 + \int_{t_{p-1}}^{t_p} \|(u_\tau - u_{h\tau})(t)\|_2^2 \, dt.$$

Cette estimation dans (4.22) donne

$$\begin{aligned} a_1(e_\tau(t_p), e_\tau(t_p)) - a_1(e_\tau(t_{p-1}), e_\tau(t_{p-1})) + \int_{t_{p-1}}^{t_p} a_2(e_\tau(t), e_\tau(t)) \\ \lesssim (\eta_t^p)^2 + \int_{t_{p-1}}^{t_p} \|(f - \pi_\tau f)(t)\|_{-1}^2 \, dt + \int_{t_{p-1}}^{t_p} \|(u_\tau - u_{h\tau})(t)\|_2^2 \, dt. \end{aligned}$$

En sommant cette estimation sur p pour $1 \le p \le n$, on conclut le résultat. ∎

Corollaire 4.3.2 (Seconde borne supérieure de l'erreur en temps)**.** *On a l'estimation suivante*

$$\begin{aligned} \|\partial_t e_\tau\|_{L^2(0,t_n;H_0^1(\Omega))}^2 &\lesssim \sum_{p=1}^{n} (\eta_t^p)^2 \\ + \int_0^{t_n} \|(u_\tau - u_{h\tau})(s)\|_2^2 \, ds &+ \|f - \pi_\tau f\|_{L^2(0,t_n;H^{-1}(\Omega))}^2. \end{aligned} \tag{4.26}$$

Preuve : L'équation résiduelle (4.20) donne directement

$$\begin{aligned} \|\partial_t e_\tau(t)\|_1 \lesssim \|(f - \pi_\tau f)(t)\|_{-1} \\ + \|e_\tau(t)\|_2 + \|(u^p - u_\tau)(t)\|_2, \forall t \in (t_{p-1}, t_p). \end{aligned}$$

En intégrant le carré de cette inégalité en $t \in (t_{p-1}, t_p)$ et en sommant sur p, on obtient

$$\begin{aligned} \int_0^{t_n} \|\partial_t e_\tau(t)\|_1^2 \, dt \lesssim \int_0^{t_n} \|(f - \pi_\tau f)(t)\|_{-1}^2 \, dt \\ + \int_0^{t_n} \|e_\tau(t)\|_2^2 \, dt + \int_0^{t_n} \|(u^p - u_\tau)(t)\|^2 \, dt. \end{aligned}$$

Le deuxième terme du membre de droite de cette inégalité est estimé dans (4.21), tandis que le troisième terme est estimé via (4.26). ∎

Passons maintenant à la borne inférieure de l'erreur en temps.

Théorème 4.3.3 (Borne inférieure de l'erreur en temps). *Pour tout $p = 1, \cdots, N$, on a l'estimation suivante*

$$\begin{aligned}(\eta_t^p)^2 \lesssim & \int_{t_{p-1}}^{t_p} (\|e_\tau(t)\|_2^2 + \|\partial_t e_\tau(t)\|_1^2)\, dt \\ & + \tau_p(\|u^p - u_h^p\|_2^2 + \|u^{p-1} - u_h^{p-1}\|_2^2) \\ & + \|f - \pi_\tau f\|^2_{L^2(t_{p-1},t_p;H^{-1}(\Omega))}.\end{aligned} \tag{4.27}$$

Preuve : Par l'inégalité triangulaire, on peut écrire

$$\eta_t^p \lesssim \tau_p^{1/2}(\|u^p - u^{p-1}\|_2 + \|u^p - u_h^p\|_2 + \|u^{p-1} - u_h^{p-1})\|_2). \tag{4.28}$$

Il reste donc à estimer le terme $\tau_p^{1/2}\|u^p - u^{p-1}\|_2$. Nous rappelons d'abord l'identité (4.23)

$$\frac{\tau_p}{3}\|u^p - u^{p-1}\|_2^2 = \int_{t_{p-1}}^{t_p} \|(u^p - u_\tau)(t)\|_2^2\, dt.$$

En prenant $v = (u^p - u_\tau)(t)$ comme fonction test dans (4.20), avec $t \in (t_{p-1}, t_p)$, on obtient

$$\begin{aligned}a_1(\partial_t e_\tau(t), (u^p - u_\tau)(t)) + a_2(e_\tau(t), (u^p - u_\tau)(t)) = & ((f - \pi_\tau f)(t), (u^p - u_\tau)(t)) \\ & + a_2((u^p - u_\tau)(t), (u^p - u_\tau)(t)), \forall \text{ a.e. } t \in (t_{p-1}, t_p).\end{aligned}$$

Grâce à l'inégalité de Cauchy-Schwarz et la continuité de la forme a_1 et la coercivité de a_2, on trouve

$$\|(u^p - u_\tau)(t)\|_2^2 \lesssim \|\partial_t e_\tau(t)\|_1^2 + \|e_\tau(t)\|_2^2 + \|(f - \pi_\tau f)(t)\|_{-1}^2, \forall \text{ a.e. } t \in (t_{p-1}, t_p).$$

En intégrant cette estimation en $t \in (t_{p-1}, t_p)$, on en déduit

$$\frac{\tau_p}{3}\|u^p - u^{p-1}\|_2^2 \lesssim \int_{t_{p-1}}^{t_p} (\|e_\tau(t)\|_2^2 + \|\partial_t e_\tau(t)\|_1^2 + \|f - \pi_\tau f\|_{-1}^2)\, dt.$$

Le résultat découle en insérant cette estimation dans (4.28). ∎

4.4 Analyse a posteriori de la discrétisation spatiale

Le but de cette section est de dériver une estimation d'erreur a posteriori pour la discrétisation du problème (4.3) avec la méthode des éléments finis en espace donnée par le schéma (4.9). Notre estimation borne l'erreur calculée en norme $H_0^1(\Omega)$ entre la solution exacte u et la solution u_h^p.

4.4.1 Borne supérieure de l'erreur

Comme d'habitude [76] le résidu d'élément exact R_K^p, est remplacé par un le résidu d'élément approché

$$r_K^p = f_h^p - L_1\left(\frac{u_h^p - u_h^{p-1}}{\tau_p}\right) - L_2 u_h^p, \tag{4.29}$$

où f_h^p est une approximation de $f(\cdot, t_p)$. On peut choisir

$$(f_h^p)_{|K} := \frac{1}{|K|}\int_K f(x, t_p)\, dx,$$

pour tout $K \in T_{ph}$.

Définition 4.4.1 (Définition de l'estimateur)**.** *Soit $p \geq 1$. L'estimateur d'erreur locale η_K^p est défini par*

$$\eta_K^p = h_K \|r_K^p\|_K + \sum_{E\in\mathcal{E}_K} h_E^{1/2} \left\|J_{E,n}^p\right\|_E,$$

tandis que l'estimateur d'erreur global η^p est défini par

$$(\eta^p)^2 = \sum_{K\in T_{ph}} (\eta_K^p)^2.$$

Les termes d'approximations locales et globales sont définis par

$$\xi_K^p = h_K \|f(\cdot, t_p) - f_h^p\|_{\omega_K}, \quad (\xi^p)^2 = \sum_{K\in T_{ph}} (\xi_K^p)^2.$$

Théorème 4.4.2 (Borne supérieure de l'erreur)**.** *On a l'estimation suivante*

$$\|e^n\|_1^2 + \sum_{p=1}^n \tau_p \|e^p\|_2^2 \lesssim \sum_{p=1}^n \tau_p((\eta^p)^2 + (\xi^p)^2) + \|e^0\|_1^2. \tag{4.30}$$

Preuve : Cette estimation est une conséquence du Lemme 4.2.3 en estimant de manière appropriée chaque terme du membre de droite de l'identité (4.17). On a par définition de R_K^p et r_K^p

$$\begin{aligned}\tau_p \sum_{K\in T_{ph}} \int_K R_K^p (e^p - I_C^0 e^p) &= \tau_p \sum_{K\in T_{ph}} \int_K \big(f_h^p - L_1(\frac{u_h^p - h_h^{p-1}}{\tau_p}) - L_2 u_h^p\big)(e^p - I_C^0 e^p) \\ &\quad + \tau_p \sum_{K\in T_{ph}} \int_K (f(\cdot, t_p) - f_h^p)(e^p - I_C^0 e^p) \\ &= \tau_p \sum_{K\in T_{ph}} \int_K r_K^p (e^p - I_C^0 e^p) + \tau_p \sum_{K\in T_{ph}} \int_K (f(\cdot, t_p) - f_h^p)(e^p - I_C^0 e^p).\end{aligned}$$

En utilisant successivement l'inégalité de Cauchy-Schwarz, l'inégalité (4.12) et la définition (4.4.1) de l'estimateur locale et le terme d'approximation, on obtient

$$\begin{aligned}\tau_p \sum_{K\in T_{ph}} \int_K R_K^p (e^p - I_C^0 e^p) &\lesssim \sum_{K\in T_{ph}} h_K (\|r_K^p\|_K + \|f(\cdot,t_p) - f_h^p\|_K) \, |e^p|_{1,\tilde{\omega}_K} \\ &\lesssim \sum_{K\in T_{ph}} (\eta_K^p + \xi_K^p) \, |e^p|_{1,\tilde{\omega}_K} .\end{aligned}$$

Par l'inégalité de Cauchy Schwarz discrète, on a

$$\begin{aligned}\sum_{K\in T_{ph}} (\eta_K^p + \xi_K^p) \, |e^p|_{1,\tilde{\omega}_K} &\leq \Big\{ (\sum_{K\in T_{ph}} (\eta_K^p)^2)^{1/2} + (\sum_{K\in T_{ph}} (\xi_K^p)^2)^{1/2} \Big\} (\sum_{K\in T_{ph}} |e^p|_{1,\tilde{\omega}_K}^2)^{1/2} \\ &= (\eta^p + \xi^p) |e^p|_{1,\omega}.\end{aligned}$$

Grâce à la coercivité de a_2, on obtient

$$\tau_p \sum_{K\in T_{ph}} \int_K R_K^p (e^p - I_C^0 e^p) \lesssim (\eta^p + \xi^p) \, |e^p|_{1,\Omega} \lesssim (\eta^p + \xi^p) \|e^p\|_2. \tag{4.31}$$

De même, en utilisant (4.13) on estime le terme résiduel de bord :

$$\begin{aligned}\sum_{E\in\mathcal{E}_{ph}^{\text{int}}} \int_E J_{E,n}^p (e^p - I_C^0 e^p) &\leq \sum_{E\in\mathcal{E}_{ph}^{\text{int}}} \left\| J_{E,n}^p \right\|_E \left\| e^p - I_C^0 e^p \right\|_E \\ &\lesssim \sum_{E\in\mathcal{E}_{ph}^{\text{int}}} \left\| J_{E,n}^p \right\|_E h_E^{1/2} \, |e^p|_{1,\tilde{\omega}_E} \\ &\lesssim \sum_{K\in T_{ph}} \eta_K^p \, |e^p|_{1,\tilde{\omega}_K} .\end{aligned}$$

Comme précédemment, l'inégalité de Cauchy-Schwarz discrète implique que

$$\sum_{E\in\mathcal{E}_{ph}^{\text{int}}} \int_E J_{E,n}^p (e^p - I_C^0 e^p) \lesssim \eta^p \, |e^p|_{1,\Omega} \lesssim \eta^p \|e^p\|_2. \tag{4.32}$$

Les inégalités de Cauchy-Schwarz et de Young impliquent que

$$\begin{aligned}a_1(e^{p-1}, e^p) &\leq \|e^{p-1}\|_1 \|e^p\|_1 \\ &\leq \frac{1}{2}(\|e^{p-1}\|_1^2 + \|e^p\|_1^2).\end{aligned}$$

Cette estimée et (4.31), (4.32) dans l'identité (4.17) donnent

$$\begin{aligned}\|e^p\|_1^2 + \tau_p \|e^p\|_2^2 &\leq \frac{1}{2}(\|e^{p-1}\|_1^2 + \|e^p\|_1^2) \\ &\quad + C\tau_p(\eta^p + \xi^p)\|e^p\|_2 \\ &\leq \frac{1}{2}(\|e^{p-1}\|_1^2 + \|e^p\|_1^2) \\ &\quad + \frac{C^2}{2}\tau_p(\eta^p + \xi^p)^2 + \frac{1}{2}\tau_p\|e^p\|_2^2,\end{aligned}$$

pour une certaine constante $C > 0$ qui dépend seulement de l'angle minimal de T_{ph}, où dans la deuxième étape, nous utilisons à nouveau l'inégalité de Young. Après quelques simplifications, cette estimation est équivalente à

$$\|e^p\|_1^2 + \tau_p \|e^p\|_2^2 \leq \|e^{p-1}\|_1^2 + C^2 \tau_p (\eta^p + \xi^p)^2,$$

le résultat se déduit donc en prenant la somme sur $p = 1, \ldots, n$. ∎

Corollaire 4.4.3 (Deuxième borne supérieure de l'erreur)**.** *On a l'estimation suivante*

$$\|\partial_t (u_\tau - u_{h\tau})\|^2_{L^2(0,t_n;H_0^1(\Omega))} \lesssim \sum_{p=1}^{n} \tau_p ((\eta^p)^2 + (\xi^p)^2) + \|e^0\|_1^2. \tag{4.33}$$

Preuve : Par la coercivité de a_1, on a

$$\|\partial_t (u_\tau - u_{h\tau})(t)\|_1 \leq \sup_{v \in H_0^1(\Omega)} \frac{a_1(\partial_t (u_\tau - u_{h\tau})(t), v)}{\|v\|_1}, \forall t \in (t_{p-1}, t_p). \tag{4.34}$$

En utilisant la propriété

$$\partial_t (u_\tau - u_{h\tau})(t) = \frac{e^p - e^{p-1}}{\tau_p}, \forall t \in (t_{p-1}, t_p),$$

et l'équation semi-discrète (4.8), pour tout $t \in (t_{p-1}, t_p)$ on peut écrire

$$a_1(\partial_t (u_\tau - u_{h\tau})(t), v) = R^p(v) - a_2(e^p, v),$$

où le résidu R^p est défini par

$$R^p(v) = (f(\cdot, t_p), v) - a_1(\frac{u_h^p - u_h^{p-1}}{\tau_p}, v) - a_2(u_h^p, v), \forall v \in H_0^1(\Omega).$$

Comme l'égalité (4.10) implique que

$$R^p(v_h) = 0, \forall v_h \in V_{ph},$$

l'identité ci-dessus devient

$$a_1(\partial_t (u_\tau - u_{h\tau})(t), v) = R^p(v - v_h) - a_2(e^p, v), \forall v_h \in V_{ph}, t \in (t_{p-1}, t_p).$$

Prenant $v_h = I_C v$, en appliquant la formule de Green (voir la preuve du Lemme 4.2.2), et en utilisant les estimations (4.12) et (4.13), on obtient

$$|R^p(v - v_h)| \lesssim (\eta^p + \xi^p)\|v\|_1,$$

et donc

$$|a_1(\partial_t (u_\tau - u_{h\tau})(t), v)| \lesssim (\eta^p + \xi^p + \|e^p\|_2)\|v\|_1, \forall t \in (t_{p-1}, t_p).$$

Cette estimation dans (4.34), donne

$$\|\partial_t (u_\tau - u_{h\tau})(t)\|_1 \lesssim \eta^p + \xi^p + \|e^p\|_2.$$

En intégrant le carré de cette estimation en $t \in (t_{p-1}, t_p)$ et en sommant sur $p = 1, \cdots, n$, le résultat découle de l'estimation (4.30). ∎

4.4.2 Borne inférieure de l'erreur

Nous établissons maintenant la borne de l'erreur inférieure de l'estimateur η_K^p de manière plus ou moins standard (voir [76]). Puisque nous considérons un problème non stationnaire, Nous avons besoin des hypothèses suivantes (voir [13, 78]), qui sont faciles à vérifier dans un contexte adaptatif :

Hypothèse 4.4.4. *Pour tout $1 \leq p \leq N$, il existe une triangulation conforme $\tilde{T}_{ph}$ telle que, chaque K de $T_{(p-1)h}$ ou de T_{ph} est l'union d'éléments $\tilde{K}$ de $\tilde{T}_{ph}$ tel que $h_K \sim h_{\tilde{K}}$.*

Nous avons également besoin de l'hypothèse suivante sur les coefficients des opérateurs L_i.

Hypothèse 4.4.5. *Pour tout $1 \leq p \leq N$, et $i = 1, 2$, les coefficients $a_{k,\ell}^{(i)}, b_k^{(i)}$ et $c^{(i)}$ sont constants sur chaque élément K de $\tilde{T}_{ph}$.*

Théorème 4.4.6 (Borne inférieure locale de l'erreur)**.** *Sous les Hypothèses 4.4.4 et 4.4.5, alors pour tout $1 \leq p \leq N$ et tout $K \in T_{ph}$, on a*

$$\eta_K^p \lesssim \|\frac{e^p - e^{p-1}}{\tau_p}\|_{1,\omega_K} + \|e^p\|_{1,\omega_K} + \sum_{K' \subset \omega_K} \xi_{K'}^p. \tag{4.35}$$

Preuve :

Résidu d'élément Fixons un élément arbitraire $K \in \tilde{T}_{nh}$ et en rappelant (4.29), on pose

$$w_K^p := b_K r_K^p,$$

où $b_K = \prod_{i=1}^{d+1} \lambda_i^K$ est la fonction bulle standard associée à K (voir par exemple [76]). Les inégalités inverses standards (cf. [76]) et le Lemme 4.2.2 avec $v = w_K^p$ donnent

$$\begin{aligned}\|r_K^p\|_K^2 \sim \int_K r_K^p w_K^p &= \int_K (f_h^p - f(\cdot, t_p)) w_K^p + \int_K R_K^p w_K^p \\ &= \int_K (f_h^p - f(\cdot, t_p)) w_K^p + a_1(\frac{e^p - e^{p-1}}{\tau_p}, w_K^p) + a_2(e^p, w_K^p).\end{aligned}$$

Par conséquent, par l'inégalité de Cauchy-Schwarz et en utilisant encore une fois les inégalités inverses standards (rappelons que u_h^{p-1} et u_h^p sont des polynômes d'ordre 1 en K), on obtient

$$\|r_K^p\|_K^2 \lesssim (\|\frac{e^p - e^{p-1}}{\tau_p}\|_{1,K} + h_K^{-1}\|e^p\|_{1,K} + \|f_h^p - f(\cdot, t_p)\|_K)\|r_K^p\|_K.$$

Cela prouve l'estimation

$$h_K\|r_K^p\|_K \lesssim \|\frac{e^p - e^{p-1}}{\tau_p}\|_{1,K} + \|e^p\|_{1,K} + h_K\|f_h^p - f(\cdot, t_p)\|_K. \tag{4.36}$$

Maintenant pour $K \in T_{ph}$, l'Hypothèse 4.4.4 implique que

$$h_K^2 \|r_K^p\|_K^2 \lesssim \sum_{\tilde{K} \in \tilde{T}_{ph} : \tilde{K} \subset K} h_{\tilde{K}}^2 \|r_{\tilde{K}}^p\|_{\tilde{K}}^2.$$

En utilisant l'estimation (4.36) et le fait que $h_{\tilde{K}} \leq h_K$ pour $\tilde{K} \subset K$, on obtient

$$h_K \|r_K^p\|_K \lesssim \|\frac{e^p - e^{p-1}}{\tau_p}\|_{1,K} + \|e^p\|_{1,K} + \xi_K^p. \tag{4.37}$$

Résidu d'arête

Ensuite, on considère une arête (ou une face) arbitraire E de $\tilde{T}_{ph}$ et on définit

$$w_E^p := b_E J_{E,n}^p.$$

En utilisant l'estimation inverse et le Lemme 4.2.2 avec $v = w_E^p$, on obtient

$$\|J_{E,n}^p\|^2 \lesssim \int_E J_{E,n}^p w_E^p = a_1(\frac{e^p - e^{p-1}}{\tau_p}, w_E^p) + a_2(e^p, w_E^p) - \sum_{K \in \tilde{T}_{ph}} \int_K R_K^p w_E^p.$$

Alors l'inégalité de Cauchy-Schwarz, les inégalités inverses et l'estimation (4.36) donnent

$$h_E^{1/2} \left\|J_{E,n}^p\right\|_E \lesssim \sum_{K \in \tilde{T}_{ph} : E \subset K} (\|\frac{e^p - e^{p-1}}{\tau_p}\|_{1,K} + \|e^p\|_{1,K} + h_K \|f_h^p - f(\cdot, t_p)\|_K).$$

Par l'Hypothèse 4.4.4, on conclut que

$$h_E^{1/2} \left\|J_{E,n}^p\right\|_E \lesssim \sum_{K \subset \omega_E} (\|\frac{e^p - e^{p-1}}{\tau_p}\|_{1,K} + \|e^p\|_{1,K} + \xi_K^p). \tag{4.38}$$

La conclusion découle des estimations (4.37) et (4.38). ∎

Corollaire 4.4.7 (Deuxième borne inférieure locale de l'erreur)**.** *Sous les Hypothèses 4.4.4 et 4.4.5, on a pour tout* $1 \leq p \leq N$ *et tout* $K \in T_{ph}$

$$\tau_p (\eta_K^p)^2 \lesssim \int_{t_{p-1}}^{t_p} \|\partial_t (u_\tau - u_{h\tau})(t)\|_{1,\omega_K}^2 \, dt + \tau_p \|e^p\|_{1,\omega_K}^2 + \tau_p \sum_{K' \subset \omega_K} (\xi_{K'}^p)^2. \tag{4.39}$$

Preuve : Conséquence directe de la propriété

$$\partial_t (u_\tau - u_{h\tau})(t) = \frac{e^p - e^{p-1}}{\tau_p}, \forall t \in (t_{p-1}, t_p),$$

et le Théorème 4.4.6. ∎

4.5 Analyse a posteriori pour le problème totalement discrétisé

Pour tout $n = 1, \cdots, N$, on note l'erreur global $E(t_n)$ au point t_n par

$$\begin{aligned} E(t_n)^2 &= \|u(t_n) - u_h^n\|_1^2 + \|u^n - u_h^n\|_1^2 \\ &\quad + \|\partial_t(u - u_\tau)\|^2_{L^2(0,t_n;H_0^1(\Omega))} + \|\partial_t(u_\tau - u_{h\tau})\|^2_{L^2(0,t_n;H_0^1(\Omega))} \\ &\quad + \int_0^{t_n} (\|(u - u_\tau)(\cdot, s)\|_2^2 + \|(u_\tau - u_{h\tau})(\cdot, s)\|_2^2)\, ds. \end{aligned}$$

En combinant les résultats des sections précédentes, on obtient les bornes supérieures et inférieures globales suivantes :

Théorème 4.5.1 (Bornes d'erreur globales)**.** *Pour tout $n = 1, \cdots, N$, on a la borne supérieure suivante*

$$\begin{aligned} E(t_n)^2 \lesssim \sum_{p=1}^{n} \left((\eta_t^p)^2 + \tau_p(\eta^p)^2\right) \\ + \|f - \pi_\tau f\|^2_{L^2(0,t_n;H^{-1}(\Omega))} + \sum_{p=1}^{n} \tau_p(\xi^p)^2 \\ + \|e^0\|_1^2 + \tau_1\|e^0\|_2^2. \end{aligned} \tag{4.40}$$

De plus, sous les hypothèses 4.4.4 et 4.4.5, pour tout $n = 1, \cdots, N$, on a la borne inférieure suivante

$$\begin{aligned} \sum_{p=1}^{n} \left((\eta_t^p)^2 + \tau_p(\eta^p)^2\right) \lesssim E(t_n)^2 \\ + \|f - \pi_\tau f\|^2_{L^2(0,t_n;H^{-1}(\Omega))} + \sum_{p=1}^{n} \tau_p(\xi^p)^2. \end{aligned} \tag{4.41}$$

Preuve : Commençons par la borne supérieure de l'erreur. D'abord par l'inégalité triangulaire, on a

$$\begin{aligned} E(t_n)^2 &\lesssim \|u(t_n) - u^n\|_1^2 + \|u^n - u_h^n\|_1^2 \\ &\quad + \|\partial_t e_\tau\|^2_{L^2(0,t_n;H_0^1(\Omega))} + \|\partial_t(u_\tau - u_{h\tau})\|^2_{L^2(0,t_n;H_0^1(\Omega))} \\ &\quad + \int_0^{t_n} (\|e_\tau(\cdot, s)\|_2^2 + \|(u_\tau - u_{h\tau})(\cdot, s)\|_2^2)\, ds. \end{aligned}$$

Par le Théorème 4.3.1 et le Corollaire 4.3.2, il est facile de vérifier que

$$\int_0^{t_n} \|(u_\tau - u_{h\tau})(\cdot, s)\|_2^2\, ds \lesssim \sum_{p=1}^{n} \tau_p(\|u^{p-1} - u_h^{p-1}\|_2^2 + \|u^p - u_h^p\|_2^2),$$

ce qui implique

$$\begin{aligned} E(t_n)^2 \lesssim & \sum_{p=1}^{n} (\eta_t^p)^2 + \sum_{p=1}^{n} \tau_p \|u^p - u_h^p\|_2^2 \\ & + \|f - \pi_\tau f\|^2_{L^2(0,t_n;H^{-1}(\Omega))} + \|u^n - u_h^n\|_1^2 \\ & + \|\partial_t(u_\tau - u_{h\tau})\|^2_{L^2(0,t_n;H_0^1(\Omega))} + \tau_1 \|e^0\|_2^2. \end{aligned}$$

On conclut par l'application du Théorème 4.4.2 et du Corollaire 4.4.3. On passe maintenant à la borne inférieure de l'erreur. En sommant (4.27) sur $p = 1, \cdots, n$, on trouve

$$\begin{aligned} \sum_{p=1}^{n} (\eta_t^p)^2 \lesssim & \int_0^{t_n} (\|e_\tau(t)\|_2^2 + \|\partial_t e_\tau(t)\|_1^2)\, dt \\ & + \sum_{p=1}^{n} \tau_p \|u^p - u_h^p\|_2^2 + \tau_1 \|e^0\|_2^2 \\ & + \|f - \pi_\tau f\|^2_{L^2(0,t_n;H^{-1}(\Omega))}. \end{aligned}$$

D'une part, par l'estimation (4.25), on obtient

$$\sum_{p=1}^{n} (\eta_t^p)^2 \lesssim E(t_n)^2 + \|f - \pi_\tau f\|^2_{L^2(0,t_n;H^{-1}(\Omega))}. \tag{4.42}$$

D'autre part, par le Corollaire 4.4.7, on a

$$\sum_{p=1}^{n} \tau_p (\eta^p)^2 \lesssim \int_0^{t_n} \|\partial_t(u_\tau - u_{h\tau})(t)\|^2_{1,\omega_K}\, dt + \sum_{p=1}^{n} \tau_p \|e^p\|_1^2 + \sum_{p=1}^{n} \tau_p (\xi_K^p)^2.$$

Grâce toujours à (4.25), on obtient

$$\sum_{p=1}^{n} \tau_p (\eta^p)^2 \lesssim E(t_n)^2 + \sum_{p=1}^{n} \tau_p (\xi^p)^2. \tag{4.43}$$

D'où l'estimation (4.41) grâce à (4.42) et (4.43). ∎

Remarque 4.5.2. Sous les Hypothèses 4.4.4 et 4.4.5, le Théorème 4.5.1 indique que l'erreur $E(t_n)$ est équivalente à l'estimateur d'erreur global

$$\left(\sum_{p=1}^{n} \left((\eta_t^p)^2 + \tau_p (\eta^p)^2 \right) \right)^{1/2},$$

à des termes approximatifs près. Dés lors, cet estimateur d'erreur global peut être utilisé pour un algorithme adaptatif.

4.6 Résultats numériques

Notre but est de confirmer notre analyse théorique par quelques tests numériques. Comme notre contribution principale concerne l'estimateur de l'erreur spatiale, nous concentrons uniquement nos efforts sur sa validité. Les deux premiers exemples servent à confirmer l'efficacité et la fiabilité de notre estimateur d'erreur. Les derniers exemples illustrent l'utilisation de notre estimateur pour présenter un algorithme de raffinement (ou déraffinement) adaptatif de maillage pour une solution ayant un comportement singulier en espace.

Test 4.6.1. Cet exemple consiste à résoudre l'équation bidimensionnelle de Sobolev sur le carré unité $\Omega =]0,1[\times]0,1[$ avec $L_1 = I - \Delta$ et $L_2 = -\Delta$ (Δ est l'opérateur standard de Laplace). Ici, nous utilisons l'élément de Lagrange sur un maillage régulier $T_{ph} = T_h$ obtenu en divisant chaque segment par p, $1 \leq p \leq N$, sous intervalles et en divisant chaque carré obtenu en deux triangles (voir Figure 3.3).

Les tests sont effectués avec $T = 1s$ et la solution exacte u suivante :

$$u(t,x,y) = e^{-t}xy(x-1)(y-1), \text{ sur } \Omega\times]0,1[,$$

avec la donnée initiale

$$u_0(x,y) = xy(x-1)(y-1), \text{ dans } \Omega,$$

et $u(.,t)_{|\Gamma} = 0$, pour tout $t \in]0,1[$. On fixe $\tau_p = 0.1s$, alors $N = T/\tau_p = 10$.

Tout d'abord, nous vérifions que la solution numérique u_h^N converge vers la solution exacte. Pour cela, nous avons tracé $|u(t_N) - u_h^N|^2_{1,\Omega}$ par rapport aux degrés de liberté (DoF$= 3n^2 - 4n + 2$ avec $h = 1/n$) à la Figure 4.4. Donc d'après les résultats obtenus, l'ordre de convergence est un en espace (avec $\Delta t = 0.001$) et en temps (avec $h = 0.00625$).

Maintenant, on passe à la confirmation des principaux résultats théoriques que sont les bornes supérieure et inférieure de l'erreur (4.26) et (4.35).

Fiabilité de l'estimateur spatial

On rappelle la borne supérieure de l'erreur obtenue dans le Théorème 4.4.2

$$\|e^n\|_1^2 + \sum_{p=1}^{n} \tau_p \|e^p\|_2^2 \lesssim \sum_{p=1}^{n} \tau_p((\eta^p)^2 + (\xi^p)^2) + \|e^0\|_1^2, \tag{4.44}$$

où, pour tout $1 \leq p \leq N$

$$(\eta^p)^2 = \sum_{K\in T_{ph}} (\eta_K^p)^2 = \sum_{K\in T_{ph}} \left(h_K\|r_K^p\|_K + \sum_{E\in\varepsilon_K} h_E^{1/2}\|J_{E,n}\|_E\right)^2,$$

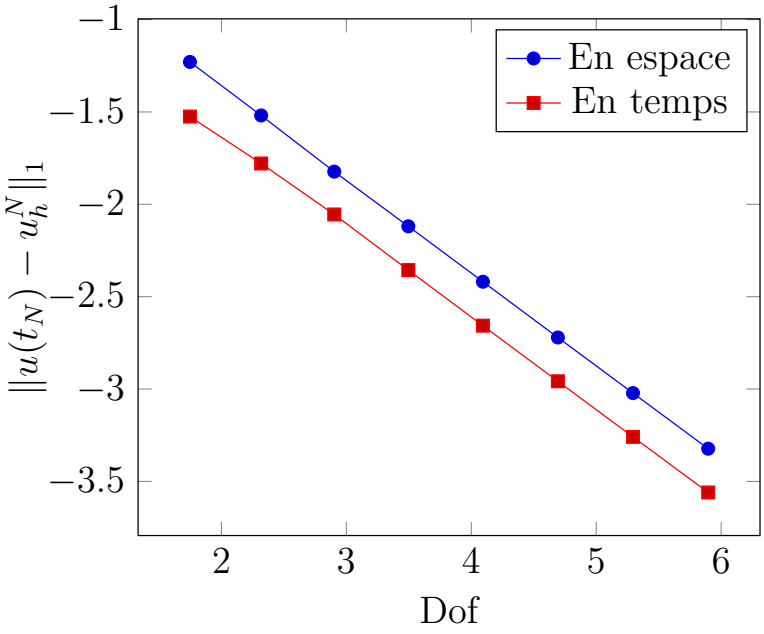


FIGURE 4.4 – L'erreur $\|u(t_N) - u_h^N\|_{1,\Omega}^2$ par rapport à Dof pour un maillage uniforme.

$$(\xi^p)^2 = \sum_{K \in T_{ph}} (\xi_K^p)^2 = \sum_{K \in T_{ph}} h_K^2 \|f(\cdot, t_p) - f_h^p\|_{w_K}^2.$$

On définit donc le rapport entre le membre de gauche et le membre de droite de l'inégalité (4.44) à l'instant final $T = 1s$:

$$q_{up}^N = \frac{\|e^N\|_1^2 + \sum_{p=1}^{N} \tau_p \|e^p\|_2^2}{\|e^0\|_1^2 + \sum_{p=1}^{N} \tau_p \sum_{K \in T_{ph}} ((\eta_K^p)^2 + h_K^2 \|f^p - f_h^p\|_K^2)},$$

qui est appelé l'indice d'efficacité. Il mesure la fiabilité de l'estimateur et est lié à l'estimation d'erreur supérieure globale. Le Théorème 4.4.2, montre que q_{up}^N est majoré par une constante. Cela est confirmé par nos résultats numériques présentés au Tableau 4.1 et à la Figure 4.5.

Efficacité de l'estimateur spatial

Nous rappelons maintenant la borne inférieure locale de l'erreur du Théorème 4.4.6

$$\eta_K^p \lesssim \|\frac{e^p - e^{p-1}}{\tau_p}\|_{1,\omega_K} + \|e^p\|_{1,\omega_K} + \sum_{K' \subset \omega_K} \xi_{K'}^p,$$

p	DoF	q_{up}^N	q_{low}^N
4	56	0.125	0.20
8	208	0.0759	0.16
16	800	0.0614	0.14
32	3136	0.0557	0.13
64	12416	0.0532	0.127
128	49408	0.0519	0.125
256	197120	0.0514	0.124
512	787456	0.0513	0.123

Tableau 4.1 – q_{up}^N et q_{low}^N par rapport à Dof pour un maillage uniforme.

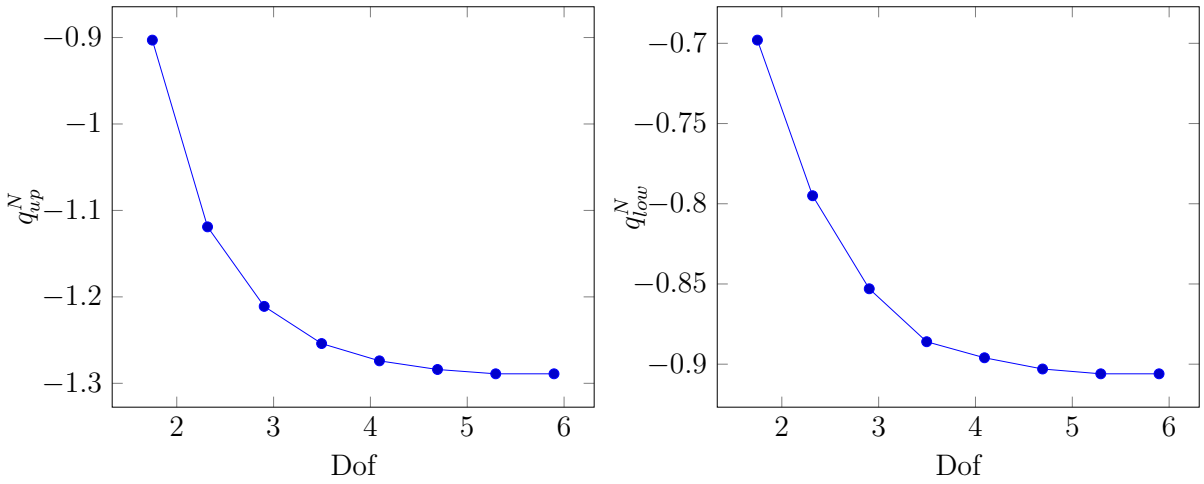


FIGURE 4.5 – q_{up}^N et q_{low}^N par rapport à Dof pour un maillage uniforme.

et on définit le rapport entre le membre de gauche et le membre de droite de cette inégalité à l'instant final $T = 1s$:

$$q_{low}^N = \max_{K \in T_{ph}} \frac{\eta_K^N}{h_k \| \frac{e^p - e^{p-1}}{\tau_p} \|_{1,\omega_K} + \| e^p \|_{1,\omega_K} + \sum_{K' \subset \omega_K} \xi_{K'}^p}.$$

qui est lié à la borne inférieure de l'erreur locale et mesure l'efficacité de l'estimateur. Selon la Figure 4.5 (voir aussi le Tableau 4.1), q_{low}^N est borné par une constante comme on l'a montré au Théorème 4.3.3. Donc notre estimateur spatial est également efficace.

Test 4.6.2. (Maillage non structuré)
Afin de valider la fiabilité et l'efficacité de notre estimateur d'erreur spatial, on approxime le même problème que précédemment avec les mêmes éléments mais sur différents

p	DoF	q^N_{up}	q^N_{low}
4	56	0.0819	0.1525
8	208	0.0630	0.13077
16	800	0.0556	0.1153
32	3136	0.0542	0.10899
64	12416	0.05398	0.1060
128	49408	0.05387	0.1051
256	197120	0.0534	0.105

Tableau 4.2 – q^N_{up} et q^N_{low} par rapport à DoF pour un maillage non structuré.

maillages non structurés en partant d'un maillage non structuré de 0.2 pas de maillage (voir la Figure 4.6) et en divisant chaque triangle en 4 triangles par un raffinement régulier [76]. La Figure 4.7 (voir aussi le Tableau 4.2) représente les indices q^N_{up} et q^N_{low} par rapport aux degrés de liberté. Les résultats obtenus sont cohérents comme dans l'exemple précédent.

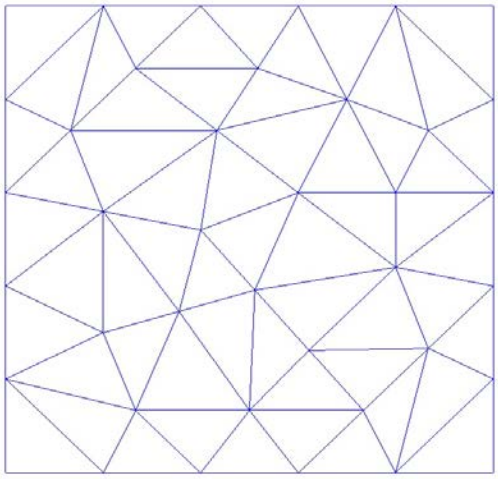


FIGURE 4.6 – Maillage non structuré avec $h = 0.2$.

4.6.1 Dépendance de l'erreur

De nos considérations précédentes, l'erreur entre la solution exacte et la solution approchée dépend du pas de maillage et/ou du pas de discrétisation en temps. Afin d'illustrer ce phénomène, comme dans [60], on considère un exemple où l'erreur due à la discrétisation en temps est plus importante que l'erreur due à la discrétisation en espace, et un autre exemple où le contraire apparaît. Pour cela, on considère le problème

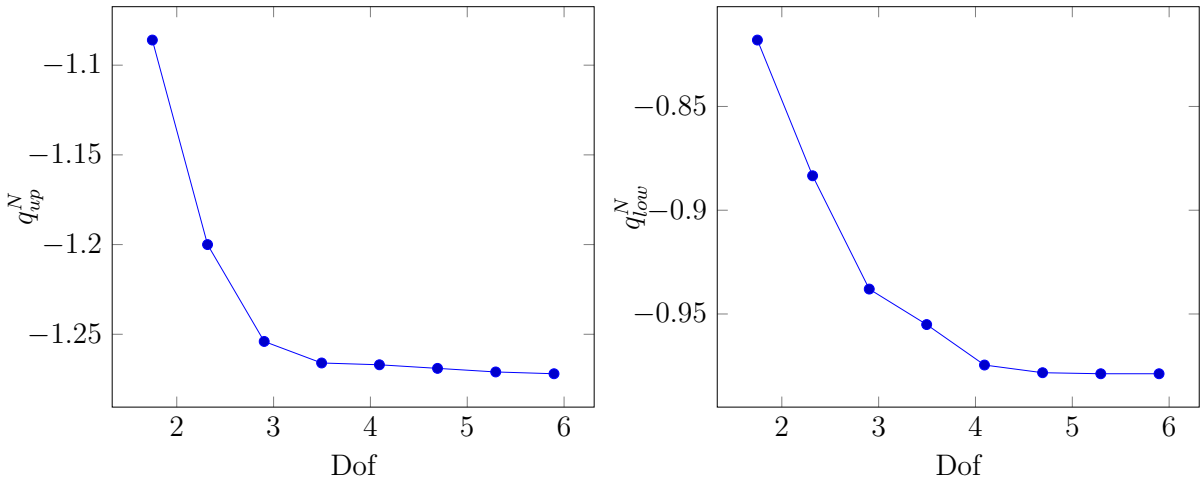


FIGURE 4.7 – q^N_{up} et q^N_{low} par rapport à Dof pour un maillage non structuré.

(4.3) pour un $\Omega =]0,1[\times]0,1[$ et $T = 1s$, avec les solutions exactes u_1 et u_2 définies par

$$u_1(t,x,y) = \sin(\pi t)\sin(\pi x)\sin(\pi y),$$

$$u_2(t,x,y) = \sin t \sin(\pi x)\sin(\pi y).$$

Les résultats numériques sont présentés dans les tableaux 4.3 et 4.4, où nous présentons les valeurs de l'indicateur spatial η, l'indicateur temporel η_t, l'erreur $\|e\| = \max\limits_{1\le p\le N} \|e^p\|_{H^1(\Omega)}$ et l'indice d'efficacité spatial pour différents pas de maillage et un pas de temps fixé.

Dans le premier exemple, le Tableau 4.3 montre que l'erreur est principalement due à la discrétisation en temps. En effet, pour un pas de temps fixé et différents pas de maillage, l'erreur est presque constante. Tandis que pour un pas de maillage fixé, l'erreur est diminuée pour différents pas de temps. Nous remarquons en outre une relation étroite entre l'erreur et l'indicateur temporel.
Pour le deuxième exemple et à partir du Tableau 4.4, l'erreur est principalement due à la discrétisation en espace, puisque nous voyons des relations inverses entre l'erreur et le pas de temps et la taille de maillage, alors qu'on détecte une relation étroite entre l'erreur et l'indicateur spatial. Dans le premier exemple q^N_{up} est corrélée à l'erreur, tandis que pour le second, la distorsion vient des termes d'approximation. Observons en outre que ces tests numériques montrent que l'indicateur temporel η_t est indépendant de h, tandis que l'indicateur spatial η est principalement indépendant de τ_p.
Cette propriété très importante de découplement des deux parties de l'erreur est utilisée efficacement dans notre algorithme adaptatif décrit ci-dessous, car les raffinements par rapport au temps (resp. par rapport à l'espace) ou les déraffinements sont basés sur η_t (resp. η).

$h = 1/n$	0.1	0.05	0.025	0.0125	0.1	0.05	0.025	0.0125
dt	0.1	0.1	0.1	0.1	0.05	0.05	0.05	0.05
η	4.62	2.32	1.16	0.58	4.34	2.16	1.08	0.54
η_t	0.22	0.22	0.22	0.22	0.077	0.077	0.077	0.077
$\|e\|$	0.53	0.53	0.53	0.53	0.26	0.26	0.26	0.26
q_{up}^N	1.77e-4	1.5e-4	1.5e-4	1.5e-4	4.77e-5	4.25e-5	4.25e-5	4.77e-5
$h = 1/n$	0.1	0.05	0.025	0.0125	0.1	0.05	0.025	0.0125
dt	0.025	0.025	0.025	0.025	0.0125	0.0125	0.0125	0.0125
η	4.48	2.24	1.12	0.56	4.60	2.30	1.15	0.57
η_t	0.027	0.027	0.027	0.027	0.009	0.009	0.009	0.009
$\|e\|$	0.13	0.13	0.13	0.13	0.06	0.06	0.06	0.06
q_{up}^N	1.57e-5	1.46e-5	1.02e-5	1.01e-5	3.70e-6	3.60e-006	3.54e-6	3.50e-6

Tableau 4.3 – Résultats de convergence où on a utilisé une triangulation uniforme et un pas de temps constant pour le premier exemple.

4.6.2 Algorithme adaptatif

A partir de nos considérations théoriques et des exemples de la sous-section précédente, un algorithme adaptatif doit utiliser de manière appropriée l'indicateur spatial η, l'indicateur temporel η_t et l'erreur d'approximation ξ. Pour concevoir cet algorithme, on définit l'indicateur d'erreur global $\bar{\eta}$ suivant :

$$\bar{\eta} := \left(\sum_{p=1}^{N} ((\eta_t^p)^2 + \tau_p(\eta^p)^2 + \tau_p(\xi)^2) \right)^{1/2}.$$

Pour la solution approchée $u_{h\tau}$, nous définissons l'estimateur d'erreur relative **Ind** par

$$\mathbf{Ind}^2 = \frac{\bar{\eta}}{\int_0^T \|\nabla u_{h\tau}(\cdot, t)\|^2 dt}. \tag{4.45}$$

Soient une tolérance δ et un paramètre $0 < \alpha < 1$ donnés. L'objectif de notre schéma adaptatif est de générer une séquence de sous-intervalles $[t_{p-1}, t_p]$ et de construire le maillage $T_{ph}, p = 1, ..., N$ de façon que **Ind** défini par (4.45), est proche de δ. Ceci se traduit par l'inégalité suivante :

$$(1 - \alpha)\delta \leq \mathbf{Ind} \leq (1 + \alpha)\delta. \tag{4.46}$$

Pour cela, pour $p = 1, ..., N$, on définit les deux bornes locales suivantes :

- Celle de gauche $\mathbf{Gb}^p$ définie par

$$\mathbf{Gb}^p := (1 - \alpha)^2 \delta^2 \int_{t_{p-1}}^{t_p} \|\nabla u_{h\tau}(\cdot, t)\|^2 dt. \tag{4.47}$$

$h = 1/n$	0.1	0.05	0.025	0.0125	0.1	0.05	0.025	0.0125
dt	0.1	0.1	0.1	0.1	0.05	0.05	0.05	0.05
η	2.56	1.29	0.65	0.32	2.52	1.27	0.065	0.32
η_t	0.039	0.039	0.039	0.039	0.014	0.014	0.014	0.014
$\|e\|$	0.30	0.16	0.08	0.04	0.30	0.16	0.08	0.04
q_{up}^N	1.36e-4	4.13e-5	2e-5	1.18e-5	1.36e-4	4.13e-5	2e-5	1.18e-5
$h = 1/n$	0.1	0.05	0.025	0.0125	0.1	0.05	0.025	0.0125
dt	0.25	0.0.25	0.025	0.025	0.025	0.0125	0.0125	0.0125
η	2.52	1.27	0.64	0.32	2.52	1.27	0.64	0.32
η_t	0.0049	0.0049	0.0049	0.0049	0.0017	0.0017	0.0017	0.0017
$\|e\|$	0.30	0.16	0.08	0.04	0.30	0.016	0.08	0.04
q_{up}^N	1.36e-4	4.13e-5	2e-5	1.18e-5	1.36e-4	4.13e-5	2e-5	1.18e-5

Tableau 4.4 – Résultats de convergence où on a utilisé une triangulation uniforme et un pas de temps constant pour le deuxième exemple.

- Celle de droite $\mathbf{Db}^p$ définie par

$$\mathbf{Db}^p := (1+\alpha)^2\delta^2 \int_{t_{p-1}}^{t_p} \|\nabla u_{h\tau}(\cdot,t)\|^2 dt. \tag{4.48}$$

Si, pour tout $p = 1, ..., N$, la condition

$$\mathbf{Gb}^p \leq (\eta_t^p)^2 + \tau_p(\eta^p)^2 + \tau_p(\xi^p)^2 \leq \mathbf{Db}^p, \tag{4.49}$$

est satisfaite, alors en sommant sur p, $p = 1, ..., N$, on obtient (4.46). Donc pour un temps discret t_{p-1}, notre algorithme consiste à choisir le pas de temps t_p et le maillage de façon à vérifier (4.49) pour tout p. Ceci sera réalisé en utilisant les éléments η^p et ξ^p pour contrôler la taille du maillage et en utilisant η_t^p et ξ^p pour contrôler le pas de temps. En d'autres termes, on cherche à construire le maillage T_{ph} et à choisir le pas de temps τ_p de façon à vérifier les inégalités suivantes :

$$\mathbf{Gb}^p \leq (\eta_t^p)^2 + \tau_p\frac{(\xi^p)^2}{2} \leq \mathbf{Db}^p, \tag{4.50}$$

$$\mathbf{Gb}^p \leq \tau_p\big((\eta^p)^2 + \frac{(\xi^p)^2}{2}\big) \leq \mathbf{Db}^p. \tag{4.51}$$

Structure globale de l'algorithme en espace

A un temps discret t_{p-1} donné, l'algorithme adaptatif en espace possède la structure générale suivante :

(e1) Considérer un maillage initial T_{nh}, poser $n := 1$.

(e2) Résoudre le problème sur T_{nh} et pour tout $K \in T_{nh}$, calculer les indicateurs d'espace locaux η_K^p, l'indicateur d'espace global η_p ainsi que ξ_p.

(e3) Si l'inégalité (4.51) est satisfaite, alors **arrêter** et poser $T_{ph} = T_{nh}$. Sinon, choisir les éléments à raffiner (ou déraffiner) et construire le nouveau maillage $T_{(n+1)h}$. Remplacer dans ce cas n par $(n+1)$ et reprendre à l'étape (e2).

Cette boucle est destinée à être répétée jusqu'à ce que la condition (4.51) soit assurée. Ce principe est présenté par la Figure 4.8.

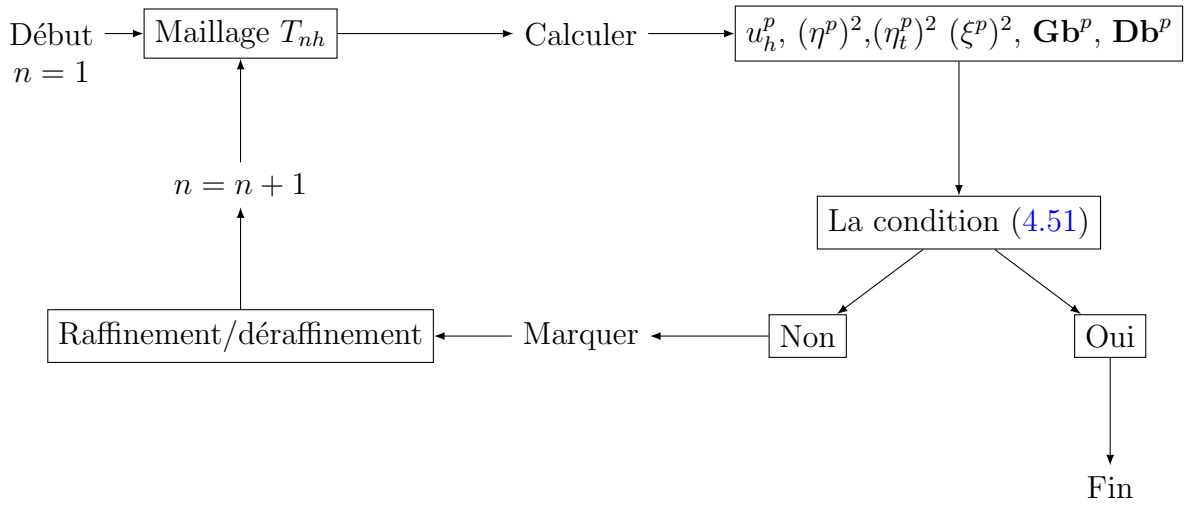


FIGURE 4.8 – Schéma d'une boucle d'adaptation de maillage.

Comme le problème est instationnaire, le processus de raffinement doit combiner le contrôle du pas d'espace et du pas de temps. On va alors considérer la stratégie globale suivante :

(t1) Pour un pas de temps donné, si (4.50) est vérifiée, raffiner (ou déraffiner) le maillage en utilisant les étapes (e1)-(e3). Sinon, raffiner (ou déraffiner) le pas de temps et reprendre les étapes (e1)-(e3). Cet algorithme est présenté dans le Programme 4.6.3. Notons qu'il est similaire à celui proposé dans [60, 56].

Remarque 4.6.1. (**Critères de marquage**) À l'étape (e3) et si la condition (4.51) n'est pas satisfaite, la question qui se pose est de décider quels éléments du maillage doivent être raffinés ou déraffinés : on parle de technique de marquage. On trouve dans la litérature plusieurs techniques de marquage [75, 26, 6, 27, 40]. Dans ce travail, on a utilisé la stratégie du **maximum** qui se décrit comme suit :

- Un élément $K \in T_{ph}$ est marqué pour raffinement si

$$\eta_K^p \geq \gamma_{\text{ref}} \max_{K \in T_{ph}} \eta_K^p,$$

où $\gamma_{\text{ref}} \in (0,1)$ est donné. Cette stratégie sélectionne les éléments dont la valeur de l'indicateur est plus grande que l'indicateur maximal multiplié par γ_{ref}.
Pour des valeurs élevées de γ_{ref} (proches de 1), la stratégie devient très sélective alors que les valeurs petites de γ_{ref} (proche de 0) correspondent au choix d'une grande partie des éléments du maillage. Typiquement, une valeur de $\gamma_{\text{ref}} = 0.5$ est utilisée ; voir par exemple, Verfürth [75].

- Inversement, un élément K est marqué pour déraffinement si $\eta_K^p \leq \gamma_{\text{deref}}\, \eta_{\max}^p$, avec $0 < \gamma_{\text{deref}} < \gamma_{\text{ref}}$.

Remarque 4.6.2. (Techniques de raffinement et de déraffinement) Pour des maillages triangulaires, il existe essentiellement deux stratégies de raffinement permettant de préserver la régularité du maillage [7, 8, 9, 62, 74, 47] :

- R1 : Diviser les triangles en joignant le milieu de l'arête la plus longue au sommet opposé à cette arête (longest edge bisection).
- R2 : Diviser les triangles en 4 en joignant les milieux des 3 arêtes (regular refinement).

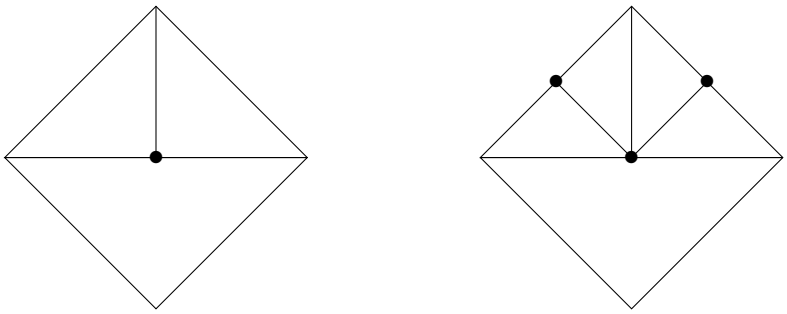


FIGURE 4.9 – Raffinement R1 à gauche et R2 à droite.

En ce qui concerne le déraffinement, les raffinements sont inversés. Dans ce travail on utilise la technique R2.

Programme 4.6.3. (**L'algorithme adaptatif**)

Poser $n = 1, T_{0h}, \tau$.... On initialise tous les paramètres utilisés.
while $t \leq T$ **do**
  Calculer $(\eta^n)^2, (\eta_t^n)^2,$
  $(\xi^n)^2, \mathbf{Rb}^n, \mathbf{Lb}^n$
  if $(\tau_n \frac{\xi^n}{2} + (\eta_t^n)^2) < \mathbf{Lb}^n$ **then**
    Le pas temps actuel est de taille petite

$\tau := 2\tau$ Déraffiner le pas de temps
else if $(\tau_n \frac{(\xi^n)^2}{2} + (\eta_t^n)^2) \leq \mathbf{Rb}^n$ **then**
if $\tau_n((\eta^n)^2 + \frac{(\xi^n)^2}{2}) < \mathbf{Lb}^n$ **then**
Continuer avec le critère La triangulation est trop fine
$\eta_K^n \leq 1.5 \min \eta_K^n$
else if $\tau_n((\eta^n)^2 + \frac{(\xi^n)^2}{2}) < \mathbf{Rb}^n$ **then**
La triangulation est correcte
$t := t + \tau$ passer au temps discret suivant
$n := n + 1$
else
Continuer avec le critère La triangulation est trop grossière
$\eta_K^n \geq 0.5 \min \eta_K^n$ Raffiner le maillage
end if
else
Le pas de temps est trop grand
$\tau := \tau/2$ Raffiner le pas de temps
end if
Construire le nouveau maillage
end while

Afin de tester notre schéma adaptatif, nous considérons deux exemples.

Test 4.6.4. Dans cet exemple, on considère le problème modèle sur le carré unité $\Omega =]0,1[\times]0,1[$ avec $L_1 = I - \Delta$, $L_2 = -\Delta$. On prend comme solution exacte

$$u(t,x,y) = \beta(t) * \exp(-50 * r^2(x,y,t)), \tag{4.52}$$

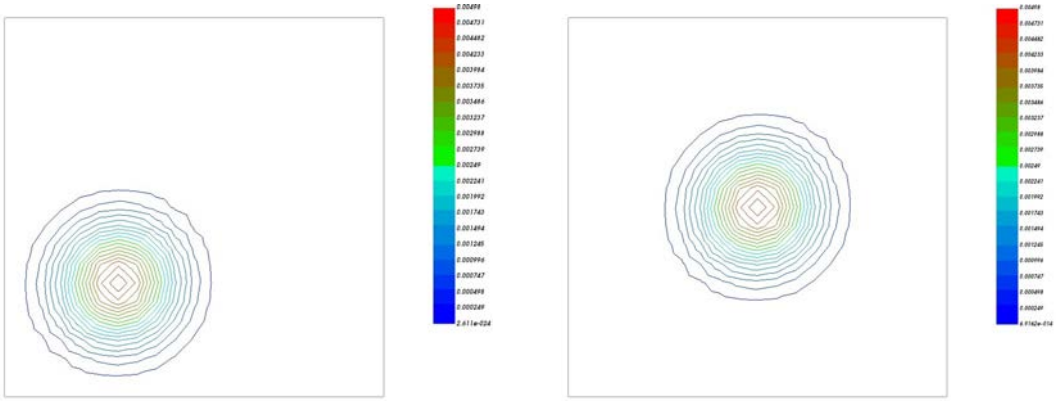
avec $r^2(t,x,y) = (x - 0.4 * t - 0.3)^2 + (y - 0.4 * t - 0.3)^2$, et

$$\beta(t) = \begin{cases} 1 - \exp(-50 * (0.98 * t + 0.01)^2) & \text{si } t < 1/1.96, \\ 1 - \exp(-50 * (1 - 0.98 * t + 0.01)^2) & \text{sinon.} \end{cases} \tag{4.53}$$

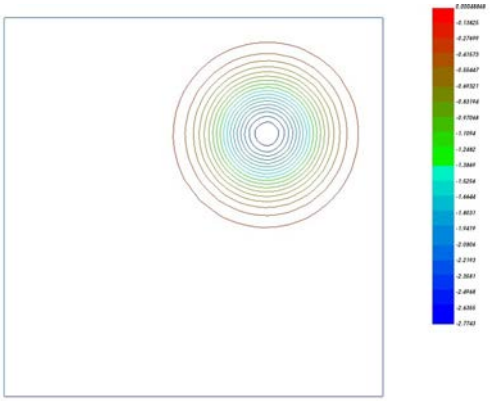
Cette solution correspond à une Gaussienne dont le centre se déplace du point $(0.3, 0.3)$ au temps $t = 0s$ au point $(0.7, 0.7)$ au temps $t = 1s$. La figure 4.10 représente la solution approchée aux temps $t = 0.1$, $t = 0.5$ et $t = 1$ et la figure 4.12 représente à ces même temps les maillages adaptatifs avec $\alpha = 0.5$ et $\delta = 0.25$. On observe que le raffinement adaptatif suit bien le déplacement de la solution.

Test 4.6.5. Dans cet exemple, on considère le problème (4.3) avec toujours $L_1 = I - \Delta$ et $L_2 = -\Delta$ sur le domaine $]-1,1[^2 \backslash]0,1[\times]0,-1[$. Comme solution exacte on prend

$$u(r,\theta) = e^{-t} * r^{2/3} \sin\left(\frac{2}{3}\theta\right),$$



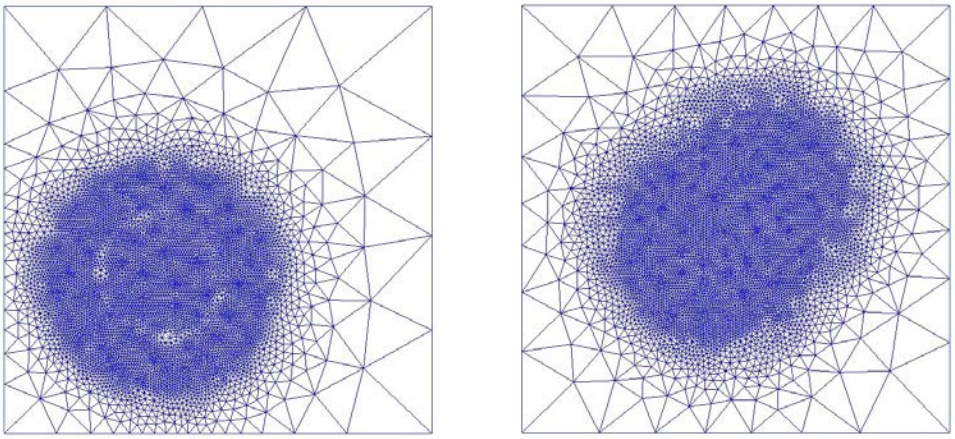
(a) $t = 0.1$ (b) $t = 0.5$



(c) $t = 1$

FIGURE 4.10 – La solution approchée après un raffinement adaptatif aux temps $t = 0.1$, $t = 0.5$ et $t = 1$ respectivement avec $\alpha = 0.5$ et $\delta = 0.25$.

où (r, θ) sont les coordonnées polaires centrées en $(0, 0)$. Dans ce cas, u a un comportement singulier le long de l'arête $(0, 0)\times]0, T[$. Ce comportement se traduit par un raffinement du maillage près de $(0, 0)$ lors de l'algorithme adaptatif, comme le montre la Figure 4.12 pour $t = 0.1$. Nous ne présentons pas les maillages pour d'autres pas de temps car ils ont une forme quasiment identique celui du cas $t = 0.1$.



(a) $n = 4,\ t_n = 0.1,\ N_v = 441$ (b) $n = 20,\ t_n = 0.5,\ N_v = 460$

(c) $n = 40,\ t_n = 1,\ N_v = 441$

FIGURE 4.11 – Maillage après le raffinement adaptatif à $t = 0,1$, $t = 0.5$ et $t = 1$ avec $\alpha = 0.5$ et $\delta = 0.25$.

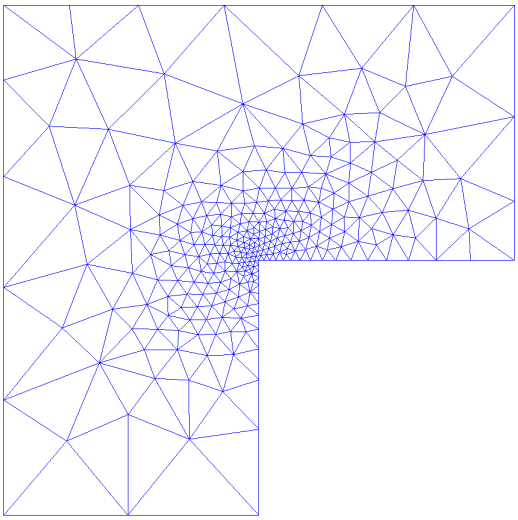


FIGURE 4.12 – Maillage après le raffinement adaptatif à $t = 0,1$ avec $\alpha = 0.5$ et $\delta = 0.25$.

Chapitre A

Les codes en FreeFem++

Sommaire

A.1 Les codes pour une équation linéaire 141
A.2 Les codes pour une équation semi-linéaire 145
A.3 Les codes pour une équation quasi-linéaire 147

FreeFem++ [1] est un logiciel destiné à la résolution numérique d'équations aux dérivées partielles, posées en dimension deux ou trois d'espace, par la méthode des éléments finis. Dans ce paragraphe, on donne les codes en freefem++ que l'on a modifié pour réaliser les calculs dans les chapitres précédents.

A.1 Les codes pour une équation linéaire

Le programme A.1.1 résout le problème linéaire de Sobolev suivant

$$\begin{cases} u_t - \Delta u_t - \Delta u = f(t), & \text{dans } \Omega \times (0,T), \\ u = 0, & \text{sur } \partial\Omega \times (0,T), \\ u(0) = u_0, & \text{dans } H_0^1(\Omega), \end{cases} \tag{A.1}$$

avec une discrétisation par le schéma d'Euler explicite en temps et la méthode des éléments finis $\mathbb{P}_1$. Il suffit de changer dans le code P_1 par P_2 ou P_3 si on veut utiliser la méthode des éléments finis $\mathbb{P}_2$ ou $\mathbb{P}_3$.

Programme A.1.1. (Pour le schéma d'Euler)

```
int n = ···;
mesh Th=square(n, n);                    //la définition du maillage
fespace Vh(Th, P1);                      //la définition de l'espace V_h
Vh uh, vh, uhn;
real t=0, T=1, dt=0.1, dt2=1, erreur=0;
real [int] erreur2(10);
erreur2=0;
```

```
func f=···;              //second membre
func u=···;              //solution exacte
func dxu=···; func dyu=···;;          // le gradient de la solution exacte
problem Eulerexplicite(uh,vh, solver=LU)=          //la formulation variationnelle
int2d(Th)( (dx(uh)*dx(vh)+dy(uh)*dy(vh))+uh*vh))
-int2d(Th)( (dt*f+uhn)*vh)
-int2d(Th)((1-dt)*(dx(uhn)*dx(vh)+dy(uhn)*dy(vh)))
+on(1, 2, 3, 4, uh=0);          //on peut modifier la condition sur le bord
int i=0;
for(i; i<3; i++){
dt=dt/2, t=dt, uhn=0, erreur=0;
for(t; t< T; i++){
Eulerexplicite;
uhn=uh;
erreur=max(erreur, sqrt(int2d(Th)(square(dxu-dx(uh))+square(dyu-dy(uh)))));
}
erreur2[i]=erreur;
}
i=0;
for(i; i<5; i++){
dt2=dt2/2;
cout<<"L'erreur vaut :"<<erreur2(i)<< "dt="<< dt2<< endl;
}
```

Programme A.1.2. (**Pour le schéma de Runge-Kutta**)

```
int n = 20;
mesh Th=square(n, n);
real t=0, t1=1 T=1, dt=0.1, erreur=0;
erreur2=0;
fespaceVh (Th, P1);
Vh uh, vh, uhn, uh1, uhn1;
func f1=x*(1-x)*y*(1-y)*cos(t1)+2*(sin(t1)+cos(t))*(x-x*x)*(y-y*y);
func f2=x*(1-x)*y*(1-y)*cos(t1+dt)+2*(sin(t1+dt)+cos(t+dt))*(x-x*x)*(y-y*y);
func u=x*(1-x)*y*(1-y)*sin(t1);
func dxu=(1-2*x)*y*(1-y)*sin(t1);
func dyu=x*(1-x)*(1-2*y)*sin(t1);
problem Eulerexplicite(uh1,vh, solver=LU)=
int2d(Th)( (dx(uh1)*dx(vh)+dy(uh1)*dy(vh))+uh1*vh))
```

```
-int2d(Th)( (dt*f1+uhn)*vh)
-int2d(Th)((1-dt)*(dx(uhn)*dx(vh)+dy(uhn)*dy(vh))
+on(1, 2, 3, 4, uh1=g) ;
problem RungeKutta(uh,vh, solver=LU)=
int2d(Th)( (dx(uh)*dx(vh)+dy(uh)*dy(vh))+uh*vh))
-int2d(Th)( (0.5*dt*f1+uhn)*vh)
-int2d(Th)((1-0.5*dt)*(dx(uhn)*dx(vh)+dy(uhn)*dy(vh))
-int2d(Th)(0.5*dt*(f2*vh-(dx(uh1)*dx(vh)+dy(uh1)*dy(vh)))
+on(1, 2, 3, 4, uh=g) ;
t=dt, uhn=0, t1=0 ;
for(t ; t< T ; i++){
Eulerexplicite ;
RungeKutta ;
uhn=uh ;
t1=t1+dt ;
erreur=max(erreur, sqrt(int2d(Th)(square(dxu-dx(uh))+square(dyu-dy(uh))))) ;
plot(uh, fill=1, value=true) ;
cout<<"L'erreur vaut :"<<erreur<< endl ;
}
```

Programme A.1.3. ((**Pour le cas singulier)**

```
//La définition de Ω = (−1, 1) × (0, 1) pour un maillage uniforme
int n=10 ;
real T=0.1, dt=0.01, t=0, epsilon=3, a1=1,a2=epsilon, a3=1 ;
border M(t=0,1) { x=-1+t ; y=0 ; label=1 ; } ;
border D1(t=0,1){ x=t ; y=0 ; label=2 ;} ;
border D2(t=0,1){ x=1 ; y=t ; label=3 ; } ;
border D31(t=0,1){ x=1-t ; y=1 ; label=4 ;} ;
border D32(t=1,2) { x=1-t ;y=1 ;label=5 ; } ;
border D4(t=0,1) { x=-1 ; y=1-t ; label=6 ; } ;
border E1(t=0,1){x=t ;y=t ;label=7 ;} ;
border E2(t=0,1){x=t-1 ; y=1-t ;label=8 ;} ;
border E3(t=0,1){x=0 ;y=1-t ;label=9 ;} ;
//La définition de Ω = (−1, 1) × (0, 1) pour un maillage raffiné
border M(t=0,1) { x=-(exp(z*t)-1)/(exp(z)-1) ; y=0 ; label=1 ; } ;
border D1(t=0,1){ x=-(exp(z*t)-1)/(exp(z)-1) ; y=0 ; label=2 ;} ;
border D2(t=0,1){ x=1 ; y=t ; label=3 ; } ;
border D31(t=0,1){ x=1-t ; y=1 ; label=4 ;} ;
```

```
border D32(t=1,2) { x=1-t ;y=1 ;label=5 ; } ;
border D4(t=0,1) { x=-1 ; y=1-t ; label=6 ; } ;
border E1(t=0,1){ x=(exp(z*t)-1)/(exp(z)-1) ;y=y=(exp(z*t)-1)/(exp(z)-1) ;label=7 ;} ;
border E2(t=0,1){x=-(exp(z*t)-1)/(exp(z)-1) ; y=y=(exp(z*t)-1)/(exp(z)-1) ;label=8 ;} ;
border E3(t=0,1){x=0 ;y=(exp(z*t)-1)/(exp(z)-1) ;label=9 ;} ;
mesh Th = buildmesh(M(n)+D1(n)+D2(n)+D31(n)+D32(n)+D4(n)+E1(n)
+E2(n)+E3(n)) ;
fespace Vh(Th, P1) ;
fespace V0(Th,P0) ;
V0 reg=region ;
Vh uh, vh, uhn ;
int Omega1=reg(0.80, 0.10) ;
int Omega21= reg(0.4, 0.80) ;
int Omega22=reg(-0.4, 0.80) ;
int Omega3=reg(-0.80,0.20) ;
func a=a1*(region==Omega1)+a2*(region==Omega21)+a2*(region==Omega22)
+a3*(region==Omega3) ;
func l=(4/pi)*asin(sqrt(1/(1+epsilon))) ;
func r=sqrt(x^2+y^2) ;
func theta1=atan2(y,x) ;
func theta2=x !=0 ? atan2(y,x) : pi/2 ;
func u1=sin(t)*(sqrt(x^2+y^2))^(l)*sin(l*theta1)*(region==Omega1) ;
func u21=sin(t)*(sqrt(x^2 +y^2))^(l)*(((2/(epsilon+1))*sin(l*theta1))
+((epsilon-1)/(sqrt(epsilon)*(epsilon+1)))*cos(l*theta1))*(region==Omega21) ;
func u22=sin(t)*(sqrt(x^2+y^2))^(l)*(((2/(epsilon+1))*sin(l*(theta2))
+((epsilon-1)/(sqrt(epsilon)*(epsilon+1)))*cos(l*(theta2))))*(region==Omega22) ;
func u3=sin(t)*(sqrt(x^2+y^2))^(l)*sin(l*(pi-theta2))*(region==Omega3) ;
problem Eulerexplicite(uh,vh,solver=LU) =
int2d(Th)( a*(dx(uh)*dx(vh) +dy(uh)*dy(vh)))
-int2d(Th)( (a-a*dt)* (dx(uhn)*dx(vh)+dy(uhn)*dy(vh)) )
-int2d(Th)( dt*( (l*r^(l-1)*(a2-1)^2)/(sqrt(a2)*(a2+1)^(3./2)) )*vh)
-int2d(Th)(dt*( (l*r^(l-1)*(a2-1)^3)/(sqrt(a2)*(a2+1)^(5./2)) )*vh)
+on(2,3, uh=u1)+on(4, uh=u21)+on(5,uh=u22)+on(1,6,uh=u3) ;
uhn=0 ; t=dt ; Vh truerror=0 ;
for(t ;t<T ; t=t+dt){
Eulerexplicite ;
uhn=uh ;
```

```
Vh uex2=u1+u21+u22+u3;
error = uh - (u1+u21+u22+u3);
real normerror = int2d(Th)( error^2+dx(error)^2+dy(error)^2 );
cout << "error=" <<sqrt(normerror) <<endl;
}
```

A.2 Les codes pour une équation semi-linéaire

On présente maintenant les codes que l'on a modifié pour calculer l'ordre de convergence en temps et en espace dans le cas d'une équation semi-linéaire. Comme on l'a déjà mentionné au chapitre 3,la solution exacte étant inconnue, l'erreur ne peut être calculée explicitement. Pour cette raison, on a utilisé une autre méthode qui consiste à déterminer l'ordre de convergence en espace et en temps par les relations suivantes :

$$\text{L'ordre de convergence en espace} = \log_2\left(\frac{\|U_{n,h} - U_{n,2h}\|_\infty}{\|U_{n,h/2} - U_{n,h}\|_\infty}\right),$$

$$\text{L'ordre de convergence en temps} = \log_2\left(\frac{\|U^{\Delta t}_{n,h} - U^{2\Delta t}_{n,h}\|_\infty}{\|U^{\Delta t/2}_{n,h} - U^{\Delta t}_{n,h}\|_\infty}\right),$$

Le code suivant consiste à calculer, pour un pas de maillage suffisamment petit, l'erreur commise entre la solution numérique pour un pas de temps Δt et la solution numérique pour le pas de temps suivant $2\Delta t$.

Programme A.2.1. (**L'ordre de convergence en temps**)

```
int n=20;
mesh Th=square(n,n);
real t=1, T=1, dt, dt1=0.003125,dt2=dt1/2;
fespace Vh(Th,P1);
Vh uh,vh,uhn=0, uh1=uhn, uh2=uhn;
func g=0;
func f=sqrt(1+t+uhn*uhn);
t=dt2, dt=dt2;
Eulerexplicite;
uh2=uh;
test = sqrt(int2d(Th)(square(dx(uh2)-dx(uh1))+square(dy(uh2)-dy(uh1))));
t=dt1;
for(t; t¡T; t=t+dt1){
uhn=uh2, dt=dt2;
Eulerexplicite;
```

```
uh2=uh, dt=dt1, uhn=uh1, uh1=uh;
test = max(test,sqrt(int2d(Th)((dx(uh2)-dx(uh1))+(dy(uh2)-dy(uh1)))));
t=t+dt2, dt=dt2, uhn=uh2;
prob;
uh2=uh;
test = max(test, sqrt(int2d(Th)((dx(uh2)-dx(uh1))+(dy(uh2)-dy(uh1)))));
t=t-dt2;
plot(uh,fill=1);
cout << "erreur = " << test << endl;
}
```

De même pour calculer l'ordre de convergence en espace. Le code suivant consiste à calculer, pour un pas de temps suffisamment petit, l'erreur commise entre la solution numérique pour un pas de maillage h et la solution numérique pour le pas de maillage suivant $2h$.

Programme A.2.2. (**L'ordre de convergence en temps**)

```
mesh Th=square(10,10);
mesh Th2=square(20,20);
real t=1, T=0.1,dt=0.001,dt2=dt;
fespace Vh(Th,P1);
fespace Vh2(Th2,P1);
Vh uh,vh,uhn=0;
Vh2 uh2,vh2, uhn2=0;
func g=0;
f=sqrt(1+t+uhn*uhn);
plot (uhn,fill=1);
problem Eulerexplicite1(uh,vh,solver=LU) =
int2d(Th)( (dx(uh)*dx(vh) +dy(uh)*dy(vh))+uh* vh)
- int2d(Th)((dt*f+uhn)*vh)
-int2d(Th)((dt-1)*(dx(uhn)*dx(vh)+dy(uhn)*dy(vh)))
+on(1,2,3,4,uh=g);
problem Eulerexplicite2(uh2,vh2,solver=LU) =
int2d(Th2)( (dx(uh2)*dx(vh2) +dy(uh2)*dy(vh2))+uh2*vh2)
- int2d(Th2)((dt*f+uhn2)*vh2)
-int2d(Th2)((dt-1)*(dx(uhn2)*dx(vh2)+dy(uhn2)*dy(vh2)))
+on(1,2,3,4,uh2=g);
real test;
t=dt;
```

```
for(t ; t<T ; t=t+dt){
prob ;
prob2 ;
uhn2=uh2 ;
uhn=uh ;
test = max(test,sqrt(int2d(Th)(square(dx(uh)-dx(uh2))+square(dy(uh)-dy(uh2))))) ;
plot(uh,fill=1) ;
}
cout << "erreur = " << test ¡¡ endl ;
```

A.3 Les codes pour une équation quasi-linéaire

Le code suivant consiste à résoudre sur un domaine $\Omega = (0,1)^2$, le problème quasi-linéaire suivant et calculer l'ordre de convergence par une discrétisation en temps par le schéma d'Euler et en espace par la méthode des éléments finis C^1 "HCT"

$$\begin{cases} \Delta\big(\rho_1(x,u)\Delta u_t\big) + \Delta\big(\rho_2(x,u)\Delta u\big) = f(t,u), & \text{dans } \Omega \times (0,T), \\ u = \frac{\partial u}{\partial n} = 0, & \text{sur } \partial\Omega \times (0,T), \\ u(0) = u_0, & \text{dans } H_0^2(\Omega), \end{cases} \tag{A.2}$$

avec $\rho_1(x,u) = \frac{1}{2} + \frac{u^2}{1+u^2}$, $\rho_2(x,u) = \frac{u^2}{1+u^2}$ et $f(t,u) = \sqrt{1+t+u^2}$.

Programme A.3.1. (**l'ordre de convergence en espace**)

```
load "Element HCT"
load "qf11to25"
real dt=0.001, T=0.1, t=dt ;
func g=0 ;
int n=10 ; mesh Th=square(n, n) ;
mesh Th2=square(2*n, 2*n) ;
fespace Vh(Th, HCT) ;
fespace Vh2(Th2, HCT) ;
macro bilaplacien(u, v) ( dxx(u)*dxx(v)+dyy(u)*dyy(v)+2*dxy(u)*dxy(v))
Vh[uh1,uh1x,uh1y], [vh, vhx,vhy], [uh1n, uh1nx, uh1ny] ;
Vh2 [uh2,uh2x, uh2y], [vh2,vh2x,vh2y], [uh2n,uh2nx,uh2ny] ;
QF2 qfHCT5 = tripleQF(qf5pT) ;
```

```
problem prob1([uh1, uh1x, uh1y], [vh,vhx,vhy], solver=LU) =
int2d(Th1,qft= qfHCT5)( (0.5+(uh1nˆ2/(1+uh1nˆ2)))*bilaplacien(uh1,vh))
−int2d(Th1)(dt*f*vh)
−int2d(Th1, qft=qfHCT5)( (0.5+(uh1nˆ2/(1+uh1nˆ2)))*bilaplacien(uh1n, vh))
+int2d(Th1, qft=qfHCT5)(dt*(uh1nˆ2/(1+uh1nˆ2))*bilaplacien(uh1n, vh))
+on(1, 2, 3, 4,uh1=0, uh1x=0,uh1y=0);

problem prob2([uh2, uh2x, uh2y], [vh,vhx,vhy], solver=LU) =
int2d(Th2,qft= qfHCT5)( (0.5+(uh2nˆ2/(1+uh2n*uh2n)))*bilaplacien(uh2,vh))
−int2d(Th2)(dt*f*vh)
−int2d(Th2, qft=qfHCT5)( (0.5+(uh2nˆ2/(1+uh2nˆ2)))*bilaplacien(uh2n, vh))
+int2d(Th2, qft=qfHCT5)(dt*(uh2nˆ2/(1+uh2nˆ2))*bilaplacien(uh2n, vh))
+on(1, 2, 3, 4,uh2=0, uh2x=0,uh2y=0);

[uh1n, uh1nx, uh1ny]=[0, 0, 0], [uh2n, uh2nx, uh2ny]=[0, 0, 0] ;
real test ;
for(t ; t<T ; t=t+dt){
prob1 ; prob2 ;
[uh2n,uh2nxuh2ny]=[uh2,uh2x,uh2y] ; [uh1n, uh1nx, uh1ny]=[uh1, uh1x, uh1y] ;
test = max(test, sqrt(int2d(Th)(square(dxx(uh1)- dxx(uh2))
+square(dyy(uh1)-dyy(uh2))+2*square(dxy(uh1)-dxy(uh2))))) ;
plot(uh1, cmm="u", fill=1, value=true) ;
cout << "erreur = " << test << endl ;
}
```

Bibliographie

[1] http ://www.freefem.org.

[2] R. Aboulaich, B. Achchab, and A. Darouichi. A posteriori error estimate for a one-dimensional pollution problem in porous media. *C. R. Math. Acad. Sci. Paris*, 347(19-20) :1217–1222, 2009.

[3] M. Amara, L. Nadau, and D. Trujillo. *A posteriori error estimator for finite volume schemes*, volume 31 of *Monogr. Semin. Mat. García Galdeano*. Prensas Univ. Zaragoza, Zaragoza, 2004.

[4] D. N. Arnold, J. Douglas, Jr., and V. Thomée. Superconvergence of a finite element approximation to the solution of a Sobolev equation in a single space variable. *Math. Comp.*, 36(153) :53–63, 1981.

[5] P. Avilez-Valente and F. J. Seabra-Santos. A Petrov-Galerkin finite element scheme for the regularized long wave equation. *Comput. Mech.*, 34(4) :256–270, 2004.

[6] I. Babuv ska and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15(4) :736–754, 1978.

[7] R. E. Bank. The efficient implementation of local mesh refinement algorithms. In *Adaptive computational methods for partial differential equations (College Park, Md., 1983)*, pages 74–81. SIAM, Philadelphia, PA, 1983.

[8] R. E. Bank. *PLTMG : a software package for solving elliptic partial differential equations*, volume 7 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990. Users' guide 6.0.

[9] R. E. Bank, A. H. Sherman, and A. Weiser. Refinement algorithms and data structures for regular local mesh refinement. In *Scientific computing (Montreal, Que., 1982)*, IMACS Trans. Sci. Comput., I, pages 3–17. IMACS, New Brunswick, NJ, 1983.

[10] G. I. Barenblatt, V. M. Entov, and V. M. Ryzhik. *Theory of fluid flow through natural rocks*. Dordrecht : Kluwer, 1990.

[11] A. Bergam, C. Bernardi, and Z. Mghazli. A posteriori analysis of the finite element discretization of a nonlinear parabolic equations. *Math. Comp.*, 74(251) :1117–1138, 2005.

[12] C. Bernardi. Optimal finite-element interpolation on curved domains. *SIAM J. Numer. Anal.*, 26(5) :1212–1240, 1989.

[13] C. Bernardi and R. Verfürth. A posteriori error analysis of the fully discretized time-dependent Stokes equations. *M2AN Math. Model. Numer. Anal.*, 38(3) :437–455, 2004.

[14] H. Blum and R. Rannacher. On the boundary value problem of the biharmonic operator on domains with angular corners. *Math. Methods Appl. Sci.*, 2(4) :556–581, 1980.

[15] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[16] R. W. Carroll and R. E. Showalter. *Singular and degenerate Cauchy problems*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1976. Mathematics in Science and Engineering, Vol. 127.

[17] J. M. Cascón, L. Ferragut, and M. I. Asensio. Space-time adaptive algorithm for the mixed parabolic problem. *Numer. Math.*, 103(3) :367–392, 2006.

[18] P. J. Chen and M. E. Gurtin. On a theory of heat conduction involving two temperatures. *Z. Angew. Math. Phys.*, 19 :614–627, 1968.

[19] Y. Chen and J. Yang. A posteriori error estimation for a fully discrete discontinuous Galerkin approximation to a kind of singularly perturbed problems. *Finite Elem. Anal. Des.*, 43(10) :757–770, 2007.

[20] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland, Amsterdam, 1978.

[21] P. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér.*, 9(R-2) :77–84, 1975.

[22] B. D. Coleman and W. Noll. An approximation theorem for functionals, with applications in continuum mechanics. *Arch. Rational Mech. Anal.*, 6 :355–370 (1960), 1960.

[23] C. Cuesta, C. J. van Duijn, and J. Hulshof. Infiltration in porous media with dynamic capillary pressure : travelling waves. *European J. Appl. Math.*, 11(4) :381–397, 2000.

[24] M. Dauge. *Elliptic boundary value problems on corner domains – smoothness and asymptotics of solutions*, volume 1341 of *Lecture Notes in Mathematics*. Springer, Berlin, 1988.

[25] A. Dogan. Numerical solution of regularized long wave equation using Petrov-Galerkin method. *Comm. Numer. Methods Engrg.*, 17(7) :485–494, 2001.

[26] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3) :1106–1124, 1996.

[27] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28(1) :43–77, 1991.

[28] A. Ern and J. Proft. A posteriori discontinuous Galerkin error estimates for transient convection-diffusion equations. *Appl. Math. Lett.*, 18(7) :833–841, 2005.

[29] R. E. Ewing. The approximation of certain parabolic equations backward in time by Sobolev equations. *SIAM J. Math. Anal.*, 6 :283–294, 1975.

[30] R. E. Ewing. Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations. *SIAM J. Numer. Anal.*, 15(6) :1125–1150, 1978.

[31] H. Gajewski and K. Zacharias. Über eine weitere Klasse nichtlinearer Differentialgleichungen im Hilbert-Raum. *Math. Nachr.*, 57 :127–140, 1973.

[32] F. Gao and X. Wang. A modified weak Galerkin finite element method for Sobolev equation. *J. Comput. Math.*, 33(3) :307–322, 2015.

[33] E. H. Georgoulis, O. Lakkis, and J. M. Virtanen. A posteriori error control for discontinuous Galerkin methods for parabolic problems. *SIAM J. Numer. Anal.*, 49(2) :427–458, 2011.

[34] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman, Boston–London–Melbourne, 1985.

[35] J. Hadamard. *Lectures on the Cauchy problem in Linear Partial Differential Equations*. Yale University Press, New Haven, Conn., 1923.

[36] E. Hairer, S. P. Nø rsett, and G. Wanner. *Solving ordinary differential equations. I*, volume 8 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1993. Nonstiff problems.

[37] S. M. Hassanizadeh and W. G. Gray. Thermodynamic basis of capillary pressure in porous media. *Water Resour. Res.*, 29 :858–879, 1993.

[38] T. Hell, A. Ostermann, and M. Sandbichler. Modification of dimension-splitting methods—overcoming the order reduction due to corner singularities. *IMA J. Numer. Anal.*, 35(3) :1078–1091, 2015.

[39] P. Houston and E. Süli. Adaptive Lagrange-Galerkin methods for unsteady convection-diffusion problems. *Math. Comp.*, 70(233) :77–106, 2001.

[40] H. Jarausch. On an adaptive grid refining technique for finite element approximations. *SIAM J. Sci. Statist. Comput.*, 7(4) :1105–1120, 1986.

[41] C. Johnson, Y. Y. Nie, and V. Thomée. An a posteriori error estimate and adaptive timestep control for a backward Euler discretization of a parabolic problem. *SIAM J. Numer. Anal.*, 27(2) :277–291, 1990.

[42] J. Kadlec. The regularity of the solution of the Poisson problem in a domain whose boundary is similar to that of a convex domain. *Czechoslovak Math. J.*, 14 (89) :386–393, 1964.

[43] T. Kato. Quasi-linear equations of evolution, with applications to partial differential equations. In *Spectral theory and differential equations (Proc. Sympos., Dundee, 1974 ; dedicated to Konrad Jörgens)*, pages 25–70. Lecture Notes in Math., Vol. 448. Springer, Berlin, 1975.

[44] D. N. Keck and M. A. McKibben. On a McKean-Vlasov stochastic integro-differential evolution equation of Sobolev-type. *Stochastic Anal. Appl.*, 21(5) :1115–1139, 2003.

[45] R. B. Kellogg. Singularities in interface problems. In B. Hubbard, editor, *Numerical solution of partial differential equations II*, pages 351–400. Academic Press, New York–London, 1971.

[46] R. B. Kellogg. On the Poisson equation with intersecting interfaces. *Appl. Analysis*, 4 :101–129, 1975.

[47] I. Kossaczký. A recursive approach to local mesh refinement in two and three dimensions. *J. Comput. Appl. Math.*, 55(3) :275–288, 1994.

[48] V. A. Kozlov, V. G. Maz'ya, and J. Rossmann. *Spectral Problems Associated with Corner Singularities of Solutions to Elliptic Equations*. Mathematical Surveys and Monographs, 85. American Mathematical Society, Providence, RI, 2001.

[49] R. Lattès and J.-L. Lions. *Méthode de quasi-réversibilité et applications*. Travaux et Recherches Mathématiques, No. 15. Dunod, Paris, 1967.

[50] K. Lemrabet. Régularité de la solution d'un problème de transmission. *J. Math. Pures et Appl.*, 56 :1–38, 1977.

[51] K. B. Liaskos, I. G. Stratis, and A. N. Yannacopoulos. Pseudoparabolic equations with additive noise and applications. *Math. Methods Appl. Sci.*, 32(8) :963–985, 2009.

[52] Y. P. Lin. Galerkin methods for nonlinear Sobolev equations. *Aequationes Math.*, 40(1) :54–66, 1990.

[53] J.-L. Lions. *Équations différentielles opérationnelles et problèmes aux limites*. Die Grundlehren der mathematischen Wissenschaften, Bd. 111. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1961.

[54] T. Liu, Y.-p. Lin, M. Rao, and J. R. Cannon. Finite element methods for Sobolev equations. *J. Comput. Math.*, 20(6) :627–642, 2002.

[55] S. Nicaise and A.-M. Sändig. General interface problems I. *Math. Methods in the Appl. Sc.*, 17 :395–429, 1994.

[56] S. Nicaise and N. Soualem. A posteriori error estimates for a nonconforming finite element discretization of the heat equation. *M2AN Math. Model. Numer. Anal.*, 39(2) :319–348, 2005.

[57] M. R. Ohm and H. Y. Lee. L^2-error analysis of fully discrete discontinuous Galerkin approximations for nonlinear Sobolev equations. *Bull. Korean Math. Soc.*, 48(5) :897–915, 2011.

[58] M. R. Ohm, H. Y. Lee, and J. Y. Shin. L^2-error estimates of the extrapolated Crank-Nicolson discontinuous Galerkin approximations for nonlinear Sobolev equations. *J. Inequal. Appl.*, pages Art. ID 895187, 17, 2010.

[59] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[60] M. Picasso. Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Engrg.*, 167(3-4) :223–237, 1998.

[61] G. Raugel. Résolution numérique par une méthode d'éléments finis du problème de Dirichlet pour le laplacien dans un polygone. *C. R. Acad. Sci. Paris Sér. A-B*, 286(18) :A791–A794, 1978.

[62] M. a.-C. Rivara. Mesh refinement processes based on the generalized bisection of simplices. *SIAM J. Numer. Anal.*, 21(3) :604–613, 1984.

[63] L. Schwartz. *Mathematics for the physical sciences*. Hermann, Paris ; Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1966.

[64] R. E. Showalter. The final value problem for evolution equations. *J. Math. Anal. Appl.*, 47 :563–572, 1974.

[65] R. E. Showalter. The Sobolev equation. I. *Applicable Anal.*, 5(1) :15–22, 1975.

[66] R. E. Showalter. The Sobolev equation. II. *Applicable Anal.*, 5(2) :81–99, 1975.

[67] R. E. Showalter. *Hilbert space methods for partial differential equations*. Pitman, London, 1977. Monographs and Studies in Mathematics, Vol. 1.

[68] S. L. Sobolev. On a new problem of mathematical physics. *Izv. Akad. Nauk SSSR. Ser. Mat.*, 18 :3–50, 1954.

[69] S. Sun and M. F. Wheeler. $L^2(H^1)$ norm a posteriori error estimation for discontinuous Galerkin approximations of reactive transport problems. *J. Sci. Comput.*, 22/23 :501–530, 2005.

[70] T. Sun and D. Yang. A priori error estimates for interior penalty discontinuous galerkin method applied to nonlinear sobolev equations. *Applied Mathematics and Computation*, 200(1) :147 – 159, 2008.

[71] V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[72] T. W. Ting. Certain non-steady flows of second-order fluids. *Arch. Rational Mech. Anal.*, 14 :1–26, 1963.

[73] T. W. Ting. A cooling process according to two-temperature theory of heat conduction. *J. Math. Anal. Appl.*, 45 :23–31, 1974.

[74] R. Verfürth. A posteriori error estimators for the Stokes equations. *Numer. Math.*, 55(3) :309–325, 1989.

[75] R. Verfürth. A posteriori error estimation and adaptive mesh-refinement techniques. In *Proceedings of the Fifth International Congress on Computational and Applied Mathematics (Leuven, 1992)*, volume 50, pages 67–83, 1994.

[76] R. Verfürth. A review of a posteriori error estimation and adaptive mesh-refinement techniques. *Wiley-Teubner, Chichester ; Stuttgart*, 1996.

[77] R. Verfürth. Error estimates for some quasi-interpolation operators. *M2AN Math. Model. Numer. Anal.*, 33(4) :695–713, 1999.

[78] R. Verfürth. A posteriori error estimates for finite element discretization of the heat equation. *Calcolo*, pages 40 :195–212, 2003.

[79] L. Wahlbin. Error estimates for a Galerkin method for a class of model equations for long waves. *Numer. Math.*, 23 :289–303, 1975. With an appendix by Lars Wahlbin, Jim Douglas, Jr. and Todd Dupont.

[80] J.-m. Yang and Y.-p. Chen. A unified a posteriori error analysis for discontinuous Galerkin approximations of reactive transport equations. *J. Comput. Math.*, 24(3) :425–434, 2006.

[81] K. Yosida. *Functional analysis*. Die Grundlehren der Mathematischen Wissenschaften, Band 123. Academic Press, Inc., New York ; Springer-Verlag, Berlin, 1965.