Notes de cours sur les méthodes variationnelles pour l'analyse et la résolution de problèmes non coercifs

Anne-Sophie Bonnet-Ben Dhia & Patrick Ciarlet
ENSTA Paris
828, boulevard des Maréchaux
91762 Palaiseau Cedex

©Anne-Sophie Bonnet-Ben Dhia & Patrick Ciarlet, 2024

©A.-S. Bonnet-Ben Dhia, P. Ciarlet 2024

Table des matières

1 Motivation et outils mathématiques **1**

1.1 Introduction 1

1.2 Outils mathématiques classiques 2

1.3 Distributions 5

1.4 Intégration par parties 9

1.5 Opérateur de trace normale, formule de Green généralisée 13

2 Etude théorique du modèle de la diffusion **19**

2.1 Modèle de la diffusion neutronique 19

2.2 Existence et unicité, théorème de Lax-Milgram 21

2.3 Existence et unicité, cadre général 23

3 Approximation des formulations variationnelles **33**

3.1 Formulation variationnelle discrète 33

3.2 Cas d'une forme coercive 35

3.3 Cas général 37

3.4 Application à la diffusion neutronique 40

4 Problèmes avec contraintes **49**

4.1 Modèle de Stokes 49

4.2 Outils mathématiques abstraits 52

4.3 Approximation par pénalisation 55

5 Problèmes avec contraintes - Compléments **59**

5.1 Approche tout-en-un 59

5.2 Approche un-plus-un : formulation variationnelle mixte 66

6 Introduction à la physique des matériaux électromagnétiques négatifs **73**

6.1 Les modèles de Drude et Drude-Lorentz 73

6.2 Deux effets du changement de signe des constantes électromagnétiques 78

7 Etude des problèmes de transmission avec changement de signe : premiers résultats **84**
7.1 Description du problème de transmission 84
7.2 La T-coercivité 86
7.3 Le cas des contrastes grands ou petits 88

8 Etude des problèmes de transmission avec changement de signe : compléments **93**
8.1 Le cas des domaines symétriques 93
8.2 Le cas des interfaces à coin 97
8.3 Une idée de la démarche générale 101

9 Synthèse et compléments **106**
9.1 Problèmes bien posés et T-coercivité 106
9.2 Approximation des formulations variationnelles 109
9.3 Bref retour sur le problème avec coefficient changeant de signe 117

Notes de cours 1

Motivation et outils mathématiques

1.1 Introduction

En règle générale, on part d'un *modèle*, qui consiste en :

— l'identification de la/des quantités à calculer, regroupée(s) sous le terme générique de *solution* ; de la/des *données* ;

— le choix de la *mesure* des solutions et des données ;

— la *modélisation* du phénomène physique (ou autre) les gouvernant, via par exemple des équations, des inéquations, des conditions au bord...

Le but du cours est de fournir les outils mathématiques permettant de prouver que ce modèle est *bien posé*, c'est-à-dire d'établir :

— l'*existence*, l'*unicité* de la solution pour un jeu de données fixé ;

— la *dépendance continue* de la solution par rapport aux données.

D'où l'importance cruciale de la mesure de la solution et des données.
On raisonnera essentiellement via l'analyse des formulations variationnelles. En outre, on construira des méthodes d'approximation de type éléments finis, permettant de résoudre numériquement le modèle.

La dimension spatiale est notée $\mathtt{d} \geq 1$, et on appelle $(\mathtt{e}_i)_{i=1,\mathtt{d}}$ la base orthonormale de $\mathbb{R}^{\mathtt{d}}$. Concernant l'étude des modèles proposés, on choisira habituellement $\mathtt{d} \leq 3$. Dans la suite, les équations/inéquations/conditions au bord font intervenir les *opérateurs de dérivation* "classiques" :

— le *gradient*, opérateur à valeurs vectorielles qui agit sur une fonction à valeurs scalaires :

$$\nabla v = \sum_{i=1,\mathtt{d}} \frac{\partial v}{\partial x_i} \mathtt{e}_i \, ;$$

— la *divergence*, opérateur à valeurs scalaires qui agit sur une fonction à valeurs vectorielles :

$$\operatorname{div} \boldsymbol{v} = \sum_{i=1,\mathtt{d}} \frac{\partial v_i}{\partial x_i}, \text{ où } \boldsymbol{v} = \sum_{i=1,\mathtt{d}} v_i \mathtt{e}_i \, ;$$

— le *Laplacien*, opérateur à valeurs scalaires qui agit sur une fonction à valeurs scalaires :

$$\Delta v = \sum_{i=1,\mathtt{d}} \frac{\partial^2 v}{\partial x_i^2} ;$$

— le *rotationnel* (cas $\mathtt{d} = 3$), opérateur à valeurs vectorielles qui agit sur une fonction à valeurs vectorielles :

$$\mathbf{rot}\, \boldsymbol{v} = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}\right) \mathtt{e}_1 + \left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}\right) \mathtt{e}_2 + \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) \mathtt{e}_3.$$

On a la relation suivante : $\Delta v = \mathrm{div}\,(\nabla v)$. On écrit parfois que $\mathrm{div}\, \boldsymbol{v} = \nabla \cdot \boldsymbol{v}$, et $\mathbf{rot}\, \boldsymbol{v} = \nabla \times \boldsymbol{v}$.

Le lecteur est supposé être familier avec les espaces de Lebesgue L^p, $p \in [1, \infty]$, ainsi qu'avec les bases de la théorie des distributions. Les espaces de Sobolev H^m, $m \in \mathbb{N}$ sont également supposé connus. On fait quelques rappels dans la suite (voir §1.2 et §1.3).
On se place dans des espaces fonctionnels dont les éléments sont à valeurs dans $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Le symbole $|\cdot|$ représente donc la valeur absolue ($\mathbb{K} = \mathbb{R}$) ou le module ($\mathbb{K} = \mathbb{C}$). Et, lorsque $\mathbb{K} = \mathbb{C}$, les produits scalaires sont hermitiens.
Enfin, lorsqu'on considère un espace vectoriel normé X, sa norme "naturelle" est notée $\|\cdot\|_X$. Parfois, on écrira $(X, \|\cdot\|_X)$ pour préciser les idées.

1.2 Outils mathématiques classiques

1.2.1 Autour des fonctions

Une fonction $f : X_0 \to Y$ (avec $X_0 \subset X$ et X, Y deux espaces vectoriels normés) est *lipschitzienne* si :

$$\exists \eta > 0,\ \forall x, x' \in X_0, \quad \|f(x) - f(x')\|_Y \leq \eta \, \|x - x'\|_X.$$

Soit f une fonction de $(x_1, \cdots, x_{\mathtt{d}})$. Sa *dérivée partielle d'ordre* α (si elle existe) est notée

$$\partial_\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_{\mathtt{d}}^{\alpha_{\mathtt{d}}}},$$

où $\alpha = (\alpha_1, \cdots, \alpha_{\mathtt{d}})$ de $\mathbb{N}^{\mathtt{d}}$ est un *multi-indice*, de longueur $|\alpha| = \sum_{j=1}^{\mathtt{d}} \alpha_j$.
Soit un ouvert Ω de $\mathbb{R}^{\mathtt{d}}$. Pour une fonction $v \in \mathcal{C}^0(\Omega)$, le *support* de v, noté $\mathrm{supp}(v)$, est le complémentaire dans Ω du plus grand ouvert Ω' tel que $v_{|\Omega'} = 0$.

1.2.2 Espaces de Banach, espaces de Hilbert

Soit $(X, \|\cdot\|_X)$ un espace vectoriel normé sur $\mathbb{K}$. On rappelle que :
— L'espace $(X, \|\cdot\|_X)$ est *séparable* s'il contient un sous-ensemble dénombrable et dense.

— L'espace $(X, \|\cdot\|_X)$ est un *espace de Banach* s'il est complet, c'est-à-dire que toute suite de Cauchy converge. Dans un espace de Banach séparable, il existe une famille dénombrable qui en constitue une base.

— L'espace $(X, \|\cdot\|_X)$ est un *espace de Hilbert* si c'est un espace de Banach, et si la norme $\|\cdot\|_X$ dérive d'un produit scalaire (hermitien si $\mathbb{K} = \mathbb{C}$), noté $(\cdot,\cdot)_X$.

— Soient X et Y deux espaces de Banach : une application *linéaire* A de X dans Y est *continue* si $\sup_{v\in X\setminus\{0\}} \frac{\|Av\|_Y}{\|v\|_X} < \infty$. Sa norme est

$$|||A||| = \sup_{v\in X\setminus\{0\}} \frac{\|Av\|_Y}{\|v\|_X}. \tag{1.1}$$

On note $\mathcal{L}(X,Y)$ l'ensemble des applications linéaires et continues de X dans Y, respectivement $\mathcal{L}(X)$ l'ensemble des applications linéaires et continues de X dans lui-même. Une application linéaire A de $\mathcal{L}(X,Y)$ est *compacte* si, de toute suite bornée de X, on peut extraire une sous-suite dont l'image par A converge dans Y. On vérifie que la composition d'une application linéaire continue par une application linéaire compacte (ou vice versa) est compacte.

— On note X' l'espace vectoriel des formes *linéaires et continues* sur X ($\mathbb{K} = \mathbb{R}$), respectivement des formes *antilinéaires et continues* sur X ($\mathbb{K} = \mathbb{C}$). On l'appelle le *dual* de X. Par définition, une forme *(anti)linéaire* $\ell \in X'$ est *continue* si $\sup_{v\in X\setminus\{0\}} |\ell(v)|/\|v\|_X < \infty$. On munit X' de la norme :

$$\|\ell\|_{X'} = \sup_{v\in X\setminus\{0\}} |\ell(v)|/\|v\|_X. \tag{1.2}$$

Pour $\ell(v)$, on utilise généralement la notation avec *crochets de dualité*, c'est-à-dire $\langle \ell, v\rangle_{X',X}$ ou $\langle \ell, v\rangle_X$. Si $(X, \|\cdot\|_X)$ est un espace de Banach, alors $(X', \|\cdot\|_{X'})$ est également un espace de Banach. On a la propriété $X \subset (X')'$, et on dit que l'espace X est *réflexif* si $(X')' = X$. Un espace de Hilbert est toujours réflexif.

— Soient X et Y deux espaces de Banach : une forme $a(\cdot,\cdot)$, *bilinéaire* sur $X \times Y$ ($\mathbb{K} = \mathbb{R}$), resp. *sesquilinéaire* sur $X \times Y$ ($\mathbb{K} = \mathbb{C}$), est *continue* si $\sup_{v\in X\setminus\{0\}, w\in Y\setminus\{0\}} \frac{|a(v,w)|}{\|v\|_X\|w\|_Y} < \infty$. Le module de continuité de a est égal à

$$|||a||| = \sup_{v\in X\setminus\{0\}, w\in Y\setminus\{0\}} \frac{|a(v,w)|}{\|v\|_X\|w\|_Y}. \tag{1.3}$$

Les principaux résultats que nous utiliserons sont les suivants.

Théorème 1.1 (Riesz) *Soit X un espace de Hilbert.*
Pour tout $\ell \in X'$, il existe une unique solution $u \in X$ de la formulation variationnelle

$$\begin{cases} \text{Trouver } u \in X \text{ telle que} \\ \forall v \in X, \quad (u,v)_X = \langle \ell, v\rangle_X. \end{cases}.$$

En outre, $\|u\|_X = \|\ell\|_{X'}$.

Théorème 1.2 (isomorphisme de Banach) *Soient X et Y deux espaces de Banach, et $A \in \mathcal{L}(X,Y)$ bijective. Alors $A^{-1} \in \mathcal{L}(Y,X)$.*

Ensuite, soient X et Y deux espaces de Banach tels que X est un sous-espace vectoriel de Y, dense dans Y, avec injection continue ($\exists C_{X\to Y} > 0$ telle que pour tout $v \in X$, $\|v\|_Y \leq C_{X\to Y}\|v\|_X$). Alors on a l'injection [1] "duale" $Y' \subset X'$ et, pour tout $\ell_Y \in Y'$, on a $\|\ell_Y\|_{X'} \leq C_{X\to Y}\|\ell_Y\|_{Y'}$, ce qui entraîne que l'injection "duale" est continue. Si de plus Y est réflexif, alors Y' est dense dans X'.
Enfin, soient X et Y deux espaces de Banach, et X^+ un sous-espace vectoriel de X, dense dans X. On suppose que A est une application linéaire de X^+ dans Y pour laquelle il existe une constante $C_A > 0$ telle que, pour tout $v^+ \in X^+$, on a $\|Av^+\|_Y \leq C_A \|v^+\|_X$. Alors on peut prolonger par continuité A en une application linéaire et continue de X dans Y, appelée *prolongement* de A, et ce prolongement est unique. En pratique, si on conserve la notation A pour le prolongement, on obtient Av pour tout $v \in X$ par passage à la limite dans Y : $Av = \lim_{k\to+\infty} Av_k^+$, où $(v_k^+)_k$ est une suite (quelconque) d'éléments de X^+ convergeant vers v dans X. De plus, on a l'inégalité $|||A||| \leq C_A$.

1.2.3 Espaces de Lebesgue

On se place dans un ouvert Ω de $\mathbb{R}^{\mathtt{d}}$. On renvoie à [8, 1] pour les démonstrations.

Définition 1.3 *L'espace $L^p(\Omega)$ est composé des fonctions de Ω dans $\mathbb{K}$, mesurables au sens de Lebesgue, et telles que :*

$$\left|\begin{array}{ll} \textit{pour } 1 \leq p < \infty & \|v\|_{L^p(\Omega)} = \left\{\displaystyle\int_\Omega |v|^p \, d\Omega\right\}^{1/p} < \infty \\ \textit{pour } p = \infty & \|v\|_{L^\infty(\Omega)} = \operatorname{supess}_{\boldsymbol{x}\in\Omega} |v(\boldsymbol{x})| < \infty \end{array}\right. .$$

Muni de la norme $\|\cdot\|_{L^p(\Omega)}$, $L^p(\Omega)$ est un espace de Banach. Lorsque $1 \leq p < \infty$, $L^p(\Omega)$ est de plus séparable.

Pour $p \in [1,\infty]$, on peut également définir les fonctions qui sont *localement* dans L^p de la façon suivante. Si [2] $v1_K$ appartient à $L^p(\Omega)$ pour tout sous-ensemble compact K de Ω, alors v est localement dans $L^p(\Omega)$, et on écrit $v \in L^p_{loc}(\Omega)$. Observons qu'on a les inclusions $L^p(\Omega) \subset L^p_{loc}(\Omega)$ pour tout $p \in [1,\infty]$.
L'égalité $v_1 = v_2$ dans $L^p_{loc}(\Omega)$ signifie que $v_1, v_2 \in L^p_{loc}(\Omega)$ et que $v_1 = v_2$ presque partout dans Ω. Si l'une des deux fonctions appartient à $L^p(\Omega)$, l'autre appartient aussi à $L^p(\Omega)$.

On a ensuite un résultat important de stabilité par la multiplication des éléments de $L^\infty(\Omega)$.

1. Lorsque X n'est pas dense dans Y, l'inclusion de Y' dans X' n'est pas injective.
2. Pour $S \subset \mathbb{R}^{\mathtt{d}}$, 1_S est la *fonction indicatrice* de S.

Proposition 1.4 *Soit $1 \leq p \leq \infty$. La multiplication est une application bilinéaire/sesquilinéaire et continue de $L^\infty(\Omega) \times L^p(\Omega)$ dans $L^p(\Omega)$.*

Plus généralement, soit $1 \leq p \leq \infty$: on définit son *exposant conjugué* $p' \in [1, \infty]$ par la relation $1/p + 1/p' = 1$. L'inégalité de Hölder [3] permet d'établir les résultats suivants.

Proposition 1.5 *Soit $1 \leq p \leq \infty$ et soit p' son exposant conjugué. La multiplication est une application bilinéaire/sesquilinéaire et continue de $L^p(\Omega) \times L^{p'}(\Omega)$ dans $L^1(\Omega)$. En outre,*

$$\forall (v, w) \in L^p(\Omega) \times L^{p'}(\Omega), \quad \|vw\|_{L^1(\Omega)} \leq \|v\|_{L^p(\Omega)} \|w\|_{L^{p'}(\Omega)}.$$

Lorsque Ω est borné, *c'est-à-dire que $vol(\Omega) = \int_\Omega 1 \, d\Omega < \infty$, pour tout $1 \leq p \leq q \leq \infty$, on a l'inclusion $L^q(\Omega) \subset L^p(\Omega)$, ainsi que l'estimation* [4]

$$\forall v \in L^q(\Omega), \quad \|v\|_{L^p(\Omega)} \leq \big(vol(\Omega)\big)^{1/p - 1/q} \|v\|_{L^q(\Omega)}.$$

On en déduit les inclusions

$$\forall p \in [1, \infty], \quad L^p(\Omega) \subset L^p_{loc}(\Omega) \subset L^1_{loc}(\Omega). \tag{1.4}$$

A l'aide du résultat de la proposition 1.5, on peut caractériser l'espace dual de l'espace $L^p(\Omega)$.

Proposition 1.6 *Soit $1 \leq p < \infty$ et soit p' son exposant conjugué. L'espace dual de $L^p(\Omega)$ peut être identifié à $L^{p'}(\Omega)$: $(L^p(\Omega))' = L^{p'}(\Omega)$. Par contre, $L^1(\Omega) \subset (L^\infty(\Omega))'$ avec inclusion stricte.*

Enfin, $L^2(\Omega)$ un espace de Hilbert.

Proposition 1.7 *L'espace $L^2(\Omega)$ est un espace de Hilbert séparable, muni du produit scalaire (hermitien si $\mathbb{K} = \mathbb{C}$)*

$$(v, w)_{L^2(\Omega)} = \int_\Omega v \, w \, d\Omega \ (\mathbb{K} = \mathbb{R}); \quad (v, w)_{L^2(\Omega)} = \int_\Omega v \, \overline{w} \, d\Omega \ (\mathbb{K} = \mathbb{C}).$$

Dans la suite, on identifiera donc l'espace de Hilbert $L^2(\Omega)$ à son dual, et on l'utilisera (ainsi que certains de ses sous-espaces) pour étudier nos modèles.

1.3 Distributions

On se place dans un ouvert Ω de $\mathbb{R}^{\mathtt{d}}$. On suppose que les fonctions sont à valeurs complexes (pour les fonctions à valeurs réelles, il suffira d'enlever la conjugaison).

3. Inégalité de Hölder : $\forall a, b \geq 0$, $ab \leq \frac{1}{p} a^p + \frac{1}{p'} b^{p'}$. Le résultat est évident si $a = 0$ ou $b = 0$. Lorsque $a, b > 0$, c'est une conséquence élémentaire de la concavité de la fonction log sur $]0, +\infty[$.

4. L'injection de $L^q(\Omega)$ dans $L^p(\Omega)$ est donc une application linéaire et continue lorsque $\mathrm{vol}(\Omega) < \infty$.

1.3.1 Définitions

Définition 1.8 *L'espace $\mathcal{D}(\Omega)$ est composé des fonctions $\mathcal{C}^\infty$ sur Ω et dont le support est compact dans Ω.*

Définition 1.9 *Une* distribution *sur Ω est une forme linéaire $T : \mathcal{D}(\Omega) \to \mathbb{R}$ qui possède la propriété suivante. Pour tout $K \subset \Omega$ compact, il existe $C > 0$ et $m \in \mathbb{N}$ tels que :*

$$\forall v \in \mathcal{D}(\Omega) \text{ avec } supp(v) \subset K, \quad |T(v)| \leq C \, \|v\|_{\mathcal{C}^m(K)}.$$

L'espace des distributions est noté $\mathcal{D}'(\Omega)$. Soient $T \in \mathcal{D}'(\Omega)$ et $v \in \mathcal{D}(\Omega)$: on utilise les crochets de dualité $\langle T, v\rangle$ pour dénoter l'action de T sur v.

On munit $\mathcal{D}(\Omega)$ de la topologie de la limite inductive stricte. En pratique, on peut utiliser la convergence des suites pour définir cette topologie. Soient $(v_k)_k$ une suite d'éléments de $\mathcal{D}(\Omega)$ et $v \in \mathcal{D}(\Omega)$. On dit que $(v_k)_k$ *converge dans $\mathcal{D}(\Omega)$ vers v* si, et seulement si :

(i) il existe $K \subset \Omega$ *compact* tel que $\mathrm{supp}(v_k) \subset K$ pour tout k, et $\mathrm{supp}(v) \subset K$;

(ii) pour tous les multi-indices $\alpha \in \mathbb{N}^{\mathsf{d}}$, la suite $(\partial_\alpha(v_k - v))_{k \geq k_0}$ converge vers 0 dans $\mathcal{C}^0(K)$.

A partir de là, on peut montrer que l'espace des distributions $\mathcal{D}'(\Omega)$ est l'espace dual de $\mathcal{D}(\Omega)$, par définition égal à l'ensemble des formes linéaires qui sont continues selon la topologie de la limite inductive stricte. En tant qu'espace dual, $\mathcal{D}'(\Omega)$ est muni de la topologie "naturelle" suivante.

Définition 1.10 *Soit $(T_k)_k$ une suite d'éléments de $\mathcal{D}'(\Omega)$: elle* converge dans $\mathcal{D}'(\Omega)$ vers T *si, et seulement si, pour tout $v \in \mathcal{D}(\Omega)$, on a $\langle T_k, v\rangle \to \langle T, v\rangle$.*

Par définition, une distribution T est nulle si, et seulement si,

$$\forall v \in \mathcal{D}(\Omega), \quad \langle T, v\rangle = 0.$$

On peut facilement prouver l'inclusion

$$L^1_{loc}(\Omega) \subset \mathcal{D}'(\Omega), \tag{1.5}$$

en *identifiant* tout $v \in L^1_{loc}(\Omega)$ à une distribution *toujours notée* v, via la correspondance

$$\forall w \in \mathcal{D}(\Omega), \quad \langle v, w\rangle = \int_\Omega v\, w \, d\Omega. \tag{1.6}$$

D'après (1.4), pour tout $p \in [1, \infty]$, tout élément de $L^p(\Omega)$ peut être considéré comme une distribution. En particulier, pour tout $v \in L^2(\Omega)$ et tout $w \in \mathcal{D}(\Omega)$, on a $\langle v, w\rangle = (v, \overline{w})_{L^2(\Omega)}$. D'après ce qui précède, on a pour finir le résultat élémentaire ci-dessous, important en "pratique"...

Proposition 1.11 *Soient $v \in L^1_{loc}(\Omega)$ et $T \in \mathcal{D}'(\Omega)$. Si on a $\langle v, w\rangle = \langle T, w\rangle$ pour tout $w \in \mathcal{D}(\Omega)$, alors $T \in L^1_{loc}(\Omega)$ avec $v = T$ presque partout dans Ω.*

Enfin, on appelle distribution vectorielle un élément de $\big(\mathcal{D}'(\Omega)\big)^{\mathsf{d}}$. Pour $\boldsymbol{T} = \sum_{i=1,\mathsf{d}} T_i \mathsf{e}_i$ appartenant à $\big(\mathcal{D}'(\Omega)\big)^{\mathsf{d}}$ et $\boldsymbol{v} = \sum_{i=1,\mathsf{d}} v_i \mathsf{e}_i$ appartenant à $\big(\mathcal{D}(\Omega)\big)^{\mathsf{d}}$, on définit l'action de $\boldsymbol{T}$ sur $\boldsymbol{v}$ par $\langle \boldsymbol{T}, \boldsymbol{v}\rangle = \sum_{i=1,\mathsf{d}} \langle T_i, v_i\rangle$.

1.3.2 Dérivation au sens des distributions

Définition 1.12 . *Soit $T \in \mathcal{D}'(\Omega)$: sa $j$ème dérivée partielle ($j = 1, \cdots, \mathtt{d}$) est définie par*

$$\forall v \in \mathcal{D}(\Omega), \quad \langle \frac{\partial T}{\partial x_j}, v \rangle = -\langle T, \frac{\partial v}{\partial x_j} \rangle.$$

On en déduit

Proposition 1.13 *L'application $T \mapsto \partial_j T$ est linéaire et continue de $\mathcal{D}'(\Omega)$ dans $\mathcal{D}'(\Omega)$.*

On a vu que $L^2(\Omega)$ est un sous-ensemble de $\mathcal{D}'(\Omega)$: on peut donc dériver ses éléments *au sens des distributions.*

Définition 1.14 *On appelle espace de Sobolev $H^1(\Omega)$ l'espace*

$$H^1(\Omega) = \{v \in L^2(\Omega) \text{ tel que } \partial_j v \in L^2(\Omega),\ j = 1, \cdots, \mathtt{d}\},$$

où la dérivation est comprise au sens des distributions.

Muni de la norme

$$\|v\|_{H^1(\Omega)} = \left\{ \int_\Omega (|v|^2 + |\nabla v|^2)\, d\Omega \right\}^{1/2},$$

c'est un espace de Banach séparable. Muni du produit scalaire hermitien

$$(v, w)_{H^1(\Omega)} = \int_\Omega (v\, \overline{w} + \nabla v \cdot \nabla \overline{w})\, d\Omega,$$

c'est un espace de Hilbert (séparable). On peut donner une définition équivalente de $H^1(\Omega)$.

Proposition 1.15 *Soit $v \in L^2(\Omega)$. Alors $v \in H^1(\Omega)$ si, et seulement si, pour tout $j = 1, \cdots, \mathtt{d}$,*

$$\exists C_j \geq 0, \quad \forall w \in \mathcal{D}(\Omega), \quad \left| (v, \frac{\partial w}{\partial x_j})_{L^2(\Omega)} \right| \leq C_j \|w\|_{L^2(\Omega)}.$$

Définition 1.16 *Soit $\alpha \in \mathbb{N}^{\mathtt{d}}$ un multi-indice. Soit $T \in \mathcal{D}'(\Omega)$: sa dérivée partielle d'ordre α est définie par*

$$\forall v \in \mathcal{D}(\Omega), \quad \langle \partial_\alpha T, v \rangle = (-1)^{|\alpha|} \langle T, \partial_\alpha v \rangle.$$

Si $\alpha = 0$, on a évidemment $\partial_\alpha T = T$! A partir de là, on peut définir les espace de Sobolev d'ordre entier.

Définition 1.17 *Soit $m \in \mathbb{N}$: on appelle espace de Sobolev $H^m(\Omega)$ l'espace*

$$H^m(\Omega) = \{v \in L^2(\Omega) \text{ tel que } \partial_\alpha v \in L^2(\Omega),\ \forall \alpha \in \mathbb{N}^{\mathtt{d}},\ |\alpha| \leq m\}.$$

Muni de la norme

$$\|v\|_{H^m(\Omega)} = \left\{ \int_\Omega \sum_{\alpha\in\mathbb{N}^{\mathtt{d}},\ |\alpha|\le m} |\partial_\alpha v|^2 \, d\Omega \right\}^{1/2}, \tag{1.7}$$

c'est un espace de Banach séparable. Muni du produit scalaire hermitien

$$(v,w)_{H^m(\Omega)} = \int_\Omega \sum_{\alpha\in\mathbb{N}^{\mathtt{d}},\ |\alpha|\le m} \partial_\alpha v \overline{\partial_\alpha w} \, d\Omega,$$

c'est un espace de Hilbert (séparable). Enfin, pour $m \ge 1$, on introduit la semi-norme

$$|v|_{m,\Omega} = \left\{ \int_\Omega \sum_{\alpha\in\mathbb{N}^{\mathtt{d}},\ |\alpha|=m} |\partial_\alpha v|^2 \, d\Omega \right\}^{1/2}, \tag{1.8}$$

pour laquelle on ne conserve que les dérivées d'ordre m.

Remarque 1.18 *On a $H^0(\Omega) = L^2(\Omega)$. Par construction, on a la suite d'injections linéaires et continues*

$$\cdots \subset H^{m+1}(\Omega) \subset H^m(\Omega) \subset \cdots \subset H^1(\Omega) \subset L^2(\Omega).$$

Définition 1.19 *Soit $m \in \mathbb{N}$, on appelle $\mathcal{C}_c^m(\overline{\Omega})$ l'ensemble composé des restrictions à $\overline{\Omega}$ de fonctions de $\mathcal{C}^m(\mathbb{R}^{\mathtt{d}})$ à support compact. On appelle $\mathcal{C}_c^\infty(\overline{\Omega})$ l'ensemble composé des restrictions à $\overline{\Omega}$ de fonctions de $\mathcal{D}(\mathbb{R}^{\mathtt{d}})$.*

On note que, par construction, $\mathcal{C}_c^\infty(\overline{\Omega}) \subset \mathcal{C}_c^m(\overline{\Omega})$ pour tout $m \ge 0$.
Les deux propriétés ci-dessous illustrent la différence entre les espaces $\mathcal{C}_c^m(\overline{\Omega})$ "classiques" et les espaces de Sobolev $H^m(\Omega)$:

— Pour tout $m \in \mathbb{N}$, $\mathcal{C}_c^m(\overline{\Omega}) \subset H^m(\Omega)$, avec injection continue si $\mathrm{vol}(\Omega) < \infty$;
— Si $\mathtt{d} \ge 2$, $H^1(\Omega) \not\subset \mathcal{C}_c^0(\overline{\Omega})$.

Pour conclure ces rappels, on introduit les espaces de fonctions que l'on peut atteindre (par passage à la limite) avec des éléments de $\mathcal{D}(\Omega)$.

Définition 1.20 *Soit $m \in \mathbb{N}$: on appelle $H_0^m(\Omega)$ l'adhérence de $\mathcal{D}(\Omega)$ dans $H^m(\Omega)$:*

$$\forall v \in H_0^m(\Omega),\ \exists (v_k)_k \in \big(\mathcal{D}(\Omega)\big)^{\mathbb{N}},\ \lim_{k\to\infty} \|v - v_k\|_{H^m(\Omega)} = 0.$$

Notons qu'on peut appliquer les *opérateurs de dérivation* aux distributions. On rappelle ci-dessous comment on procède pour le gradient et la divergence.

— Comme l'opérateur gradient est à valeurs vectorielles, ie. $\nabla\cdot = \sum_{i=1,\mathtt{d}} \frac{\partial\cdot}{\partial x_i}\mathtt{e}_i$, le gradient d'une distribution est une distribution vectorielle agissant sur les fonctions $\boldsymbol{v} = \sum_{i=1,\mathtt{d}} v_i \mathtt{e}_i$ de $\big(\mathcal{D}(\Omega)\big)^{\mathtt{d}}$. Si on utilise la définition 1.12 terme à terme, alors le *gradient* de $T \in \mathcal{D}'(\Omega)$ est caractérisé par :

$$\forall \boldsymbol{v} \in \big(\mathcal{D}(\Omega)\big)^{\mathtt{d}}, \quad \langle \nabla T, \boldsymbol{v}\rangle = \sum_{i=1,\mathtt{d}} \langle \frac{\partial T}{\partial x_i}, v_i\rangle = -\sum_{i=1,\mathtt{d}} \langle T, \frac{\partial v_i}{\partial x_i}\rangle = -\langle T, \sum_{i=1,\mathtt{d}} \frac{\partial v_i}{\partial x_i}\rangle = -\langle T, \mathrm{div}\, \boldsymbol{v}\rangle.$$

Ci-dessus, l'opérateur divergence agissant sur $\boldsymbol{v}$ est l'opérateur "classique".

— D'autre part, comme l'opérateur divergence agit sur des champs à valeurs vectorielles, on l'applique à une distribution vectorielle $\boldsymbol{T} = \sum_{i=1,\mathtt{d}} T_i \mathtt{e}_i$ de $\left(\mathcal{D}'(\Omega)\right)^{\mathtt{d}}$: sa *divergence* est la distribution $\operatorname{div} \boldsymbol{T} = \sum_{i=1,\mathtt{d}} \frac{\partial T_i}{\partial x_i}$ et on a la caractérisation

$$\forall v \in \mathcal{D}(\Omega), \quad \langle \operatorname{div} \boldsymbol{T}, v \rangle = \langle \sum_{i=1,\mathtt{d}} \frac{\partial T_i}{\partial x_i}, v \rangle = \sum_{i=1,\mathtt{d}} \langle \frac{\partial T_i}{\partial x_i}, v \rangle = - \sum_{i=1,\mathtt{d}} \langle T_i, \frac{\partial v}{\partial x_i} \rangle = -\langle \boldsymbol{T}, \nabla v \rangle.$$

L'opérateur gradient agissant sur v est l'opérateur "classique".

Proposition 1.21 *L'application $T \mapsto \nabla T$ est linéaire et continue de $\mathcal{D}'(\Omega)$ dans $\left(\mathcal{D}'(\Omega)\right)^{\mathtt{d}}$. L'application $\boldsymbol{T} \mapsto \operatorname{div} \boldsymbol{T}$ est linéaire et continue de $\left(\mathcal{D}'(\Omega)\right)^{\mathtt{d}}$ dans $\mathcal{D}'(\Omega)$.*

Exercice : En utilisant la dérivation au sens des distributions et l'inégalité de Cauchy-Schwarz dans $L^2(\Omega)$, $\left(L^2(\Omega)\right)^{\mathtt{d}}$ ou $\mathbb{R}^{\mathtt{d}}$, montrer que :

1. l'application divergence est telle que

$$\operatorname{div} \in \mathcal{L}(\left(H^1(\Omega)\right)^{\mathtt{d}}, L^2(\Omega)) \text{ avec } |||\operatorname{div}||| \leq \sqrt{\mathtt{d}}\,; \tag{1.9}$$

2. l'application $j$ème dérivée partielle ($j = 1, \cdots, \mathtt{d}$) est telle que

$$\partial_j \in \mathcal{L}\left(L^2(\Omega), (H_0^1(\Omega))'\right) \text{ avec } |||\partial_j||| \leq 1\,; \tag{1.10}$$

3. l'application gradient est telle que

$$\nabla \in \mathcal{L}\left(L^2(\Omega), \left((H_0^1(\Omega))'\right)^{\mathtt{d}}\right) \text{ avec } |||\nabla||| \leq \sqrt{\mathtt{d}}\,; \tag{1.11}$$

4. l'application Laplacien est telle que

$$\Delta \in \mathcal{L}\left(H^1(\Omega), (H_0^1(\Omega))'\right) \text{ avec } |||\Delta||| \leq 1. \tag{1.12}$$

1.4 Intégration par parties

On suppose que les fonctions sont à valeurs complexes (pour les fonctions à valeurs réelles, il suffira d'enlever la conjugaison).

1.4.1 Notion de domaine

On rappelle la notion de *frontière lipschitzienne ou régulière*, cf. la définition 1.2.1.1 de [19].

Définition 1.22 *Soit Ω un ouvert de $\mathbb{R}^{\mathtt{d}}$, et $k \in \mathbb{N} \setminus \{0\}$.*
Sa frontière $\partial\Omega$ est lipschitzienne (respectivement de classe $C^{k,1}$) si, et seulement si, en tout point $\boldsymbol{x}$ de $\partial\Omega$, pour un système de coordonnées $(y_1, \cdots, y_{\mathtt{d}})$ bien choisi, il existe une application lipschitzienne (resp. de classe $C^{k,1}$) $\varphi : V^- \to \mathbb{R}$ définie sur un $(\mathtt{d}-1)$-hypercube ouvert $V^- \subset \mathbb{R}^{\mathtt{d}-1}_y \times \{0\}$, et un voisinage V ouvert de $\boldsymbol{x}$, tels que :

$$V \cap \partial\Omega = \{\boldsymbol{x}' = (\mathbf{y}, y_{\mathtt{d}}) \in V \text{ tel que } y_{\mathtt{d}} = \varphi(\mathbf{y})\} \quad \textit{(le graphe de } \varphi \textit{ représente localement } \partial\Omega);$$
$$V \cap \Omega = \{\boldsymbol{x}' = (\mathbf{y}, y_{\mathtt{d}}) \in V \text{ tel que } y_{\mathtt{d}} < \varphi(\mathbf{y})\} \quad (\Omega \textit{ est localement d'un seul côté de } \partial\Omega).$$

En un point $\boldsymbol{x}$ de $\partial\Omega$, l'application φ est appelée *carte locale*. En reprenant les notations précédentes, on peut définir le vecteur normal unitaire sortant à $\partial\Omega$, noté $\boldsymbol{n}$, dès que $\partial\Omega$ est lipschitzienne. Précisément, au voisinage d'un point $\boldsymbol{x}$ de $\partial\Omega$, on a

$$\boldsymbol{n}(\mathbf{y}, \varphi(\mathbf{y})) = \frac{(-\partial_1\varphi(\mathbf{y}), \cdots, -\partial_{\mathtt{d}-1}\varphi(\mathbf{y}), 1)}{(\sum_{i=1,\mathtt{d}-1}(\partial_i\varphi(\mathbf{y}))^2 + 1)^{1/2}}, \text{ presque pour tout } \mathbf{y} \in V^-.$$

Sur une frontière $\partial\Omega$ lipschitzienne, on peut définir les espaces de Lebesgue $L^p(\partial\Omega)$ pour tout $p \in [1, \infty]$, avec des propriétés similaires à celles de §1.2 ; voir §1.4.3 pour un exemple.

Proposition 1.23 *Soit Ω un ouvert de $\mathbb{R}^{\mathtt{d}}$ à frontière lipschitzienne, alors $\boldsymbol{n}$ appartient à $(L^\infty(\partial\Omega))^{\mathtt{d}}$. Si de plus la frontière est de classe $C^{k,1}$ pour $k \geq 1$, alors $\boldsymbol{n}$ appartient à $(C^{k-1}(\partial\Omega))^{\mathtt{d}}$.*

Dans la suite du cours, on raisonnera généralement dans un *domaine*, dont on donne la définition ci-dessous.

Définition 1.24 *Un ouvert borné connexe de $\mathbb{R}^{\mathtt{d}}$ à frontière lipschitzienne est appelé un domaine.*

1.4.2 Formules de Green classiques

On se place dans un ouvert Ω de $\mathbb{R}^{\mathtt{d}}$ à frontière lipschitzienne. On rappelle les formules d'intégration par parties ci-dessous, où $(n_i)_{i=1,\mathtt{d}}$ sont les composantes de $\boldsymbol{n}$.

Proposition 1.25 *Soit $v \in \mathcal{C}^1(\Omega) \cap \mathcal{C}^0(\overline{\Omega})$ à support borné, on a pour tout $i = 1, \mathtt{d}$:*

$$\int_\Omega \frac{\partial v}{\partial x_i}\, d\Omega = \int_{\partial\Omega} v n_i\, d\Gamma. \tag{1.13}$$

Soient $v, w \in \mathcal{C}^1(\Omega) \cap \mathcal{C}^0(\overline{\Omega})$ à support borné, on a pour tout $i = 1, \mathtt{d}$:

$$\int_\Omega \left(\frac{\partial v}{\partial x_i}\,\overline{w} + v\,\frac{\partial \overline{w}}{\partial x_i}\right) d\Omega = \int_{\partial\Omega} v\overline{w} n_i\, d\Gamma. \tag{1.14}$$

Soient $\boldsymbol{v} \in \big(\mathcal{C}^1(\Omega) \cap \mathcal{C}^0(\overline{\Omega})\big)^{\mathtt{d}}$ et $w \in \mathcal{C}^1(\Omega) \cap \mathcal{C}^0(\overline{\Omega})$ à support borné, on a :

$$\int_\Omega \big((\operatorname{div} \boldsymbol{v})\overline{w} + \boldsymbol{v} \cdot \nabla\overline{w}\big) d\Omega = \int_{\partial\Omega} (\boldsymbol{v} \cdot \boldsymbol{n})\overline{w}\, d\Gamma. \tag{1.15}$$

A partir d'une des trois formules, il est très simple de retrouver les deux autres.
Dans (1.15), la quantité $\boldsymbol{v}\cdot\boldsymbol{n}_{|\partial\Omega}$ définie sur $\partial\Omega$ est appelée *flux* de $\boldsymbol{v}$ à travers $\partial\Omega$. Typiquement, les formules (1.14)-(1.15) permettent de passer du modèle à résoudre à la formulation variationnelle équivalente, en intégrant les équations/inéquations, et en utilisant les conditions de bord. Dans (1.14), v joue le rôle de la solution (ou d'une de ses composantes), et w est une fonction-test. Dans (1.15), $\boldsymbol{v}$ joue le rôle du gradient de la solution ou d'une quantité proportionnelle au gradient, et w est à nouveau une fonction-test. Ceci étant dit, la régularité pré-supposée de la solution et des fonctions-test dans (1.14)-(1.15) est trop restrictive, c'est pourquoi on va généraliser ces formules...

1.4.3 Théorème de trace, formule de Green généralisée

Le but est maintenant de généraliser ces formules d'intégration par parties à des objets plus généraux, à savoir appartenant à des espaces de Sobolev (d'ordre le plus petit possible). On rappelle un résultat de densité qui permettra de remplacer ces objets "peu réguliers" par des fonctions régulières, beaucoup plus simples à manipuler. En particulier, ce résultat permet d'étendre des résultats connus pour des fonctions régulières, tels que la valeur, ou trace, sur la frontière. On suit [19].
On a tout d'abord un résultat de densité des fonctions régulières dans les espaces de Sobolev.

Proposition 1.26 *Soit $m\in\mathbb{N}$. Dans tout ouvert Ω à frontière lipschitzienne, $\mathcal{C}_c^\infty(\overline{\Omega})$ est dense dans $H^m(\Omega)$:*

$$\forall v\in H^m(\Omega),\ \exists (v_k)_k\in\left(\mathcal{C}_c^\infty(\overline{\Omega})\right)^{\mathbb{N}},\ \lim_{k\to\infty}\|v-v_k\|_{H^m(\Omega)}=0.$$

Soit un ouvert Ω de frontière lipschitzienne. Pour tout $v\in\mathcal{C}_c^\infty(\overline{\Omega})$, on peut définir sa valeur sur la frontière $\partial\Omega$, ou *trace*, notée $v_{|\partial\Omega}$. En outre,

$$v_{|\partial\Omega}\in L^2(\partial\Omega)=\left\{\lambda \text{ mesurable sur } \partial\Omega \text{ tel que } \int_{\partial\Omega}|\lambda|^2 d\Gamma<+\infty\right\},$$

où $d\Gamma$ désigne l'élément de "surface" porté par $\partial\Omega$; $d\Gamma_{\boldsymbol{x}}$ est défini autour d'un point $\boldsymbol{x}$ à l'aide d'une carte locale, en partant d'un hypercube de $\mathbb{R}^{\mathtt{d}-1}$ (voir la définition 1.22). Muni du produit scalaire hermitien

$$(\lambda_0,\lambda_1)_{L^2(\partial\Omega)}=\int_{\partial\Omega}\lambda_0\overline{\lambda_1}\,d\Gamma,$$

$L^2(\partial\Omega)$ est un espace de Hilbert, qu'on identifie à son dual. On dit que $\partial\Omega$ est *bornée* si $\mathrm{vol}(\partial\Omega)=\int_{\partial\Omega}1\,d\Gamma<\infty$. En particulier, la frontière d'un domaine est toujours bornée.

Théorème 1.27 *Soit Ω un ouvert de $\mathbb{R}^{\mathtt{d}}$ à frontière lipschitzienne et bornée. Alors, l'application trace*

$$\gamma_0:\left|\begin{array}{ccc}\mathcal{C}_c^\infty\left(\overline{\Omega}\right) & \to & L^2(\partial\Omega)\\ v & \mapsto & \gamma_0 v=v_{|\partial\Omega}\end{array}\right.$$

se prolonge par continuité en une application linéaire continue, toujours notée γ_0, *de* $H^1(\Omega)$ *dans* $L^2(\partial\Omega)$.

On introduit

$$H^{1/2}(\partial\Omega) = \left\{\lambda \in L^2(\partial\Omega) \text{ tel que } \exists v \in H^1(\Omega),\ \lambda = \gamma_0 v\right\}, \tag{1.16}$$

qu'on munit de la norme

$$\|\lambda\|_{H^{1/2}(\partial\Omega)} = \inf_{v\in H^1(\Omega)\ tq\ \gamma_0 v=\lambda} \|v\|_{H^1(\Omega)}. \tag{1.17}$$

Proposition 1.28 *Soit* Ω *un ouvert de* $\mathbb{R}^{\mathtt{d}}$ *à frontière lipschitzienne et bornée. Muni de la norme (1.17), l'ensemble des traces* $H^{1/2}(\partial\Omega)$ *est un espace de Banach. En outre,* $H^{1/2}(\partial\Omega)$ *est un sous-ensemble strict, dense dans* $L^2(\partial\Omega)$ *(avec injection continue).*

Par définition, l'application trace γ_0 est continue de $H^1(\Omega)$ dans $H^{1/2}(\partial\Omega)$, de norme égale à 1. Il suit de la proposition 1.28 que $L^2(\partial\Omega) \subset (H^{1/2}(\partial\Omega))'$ (avec injection continue).

A l'aide de l'application trace γ_0 et de la densité de $\mathcal{C}_c^\infty(\overline{\Omega})$ dans $H^1(\Omega)$, on peut généraliser la formule d'intégration par parties (1.14). Donnons d'abord une idée intuitive de l'énoncé du résultat. On rappelle qu'on a défini des espaces de Sobolev $(H^m(\Omega))_{m\in\mathbb{N}}$ inclus les uns dans les autres, voir la définition 1.17. Si on choisit *a priori* v dans l'espace le plus grand, c'est-à-dire dans $H^0(\Omega) = L^2(\Omega)$, on sait que l'on doit choisir $(\partial_i w)_{i=1,\mathtt{d}}$ dans l'espace dual de $L^2(\Omega)$ (proposition 1.5), c'est-à-dire lui-même, pour que toutes les intégrales

$$\forall i = 1, \mathtt{d}, \quad \int_\Omega v\, \frac{\partial \overline{w}}{\partial x_i}\, d\Omega,$$

aient un sens. Et symétriquement (en inversant le rôle de v et w), on choisit $w \in L^2(\Omega)$ et $(\partial_i v)_{i=1,\mathtt{d}} \in \left(L^2(\Omega)\right)^{\mathtt{d}}$ pour que les autres intégrales volumiques

$$\forall i = 1, \mathtt{d}, \quad \int_\Omega \frac{\partial v}{\partial x_i}\, \overline{w}\, d\Omega$$

aient un sens. Ainsi, on doit choisir v, w dans $H^1(\Omega)$. Et, si les hypothèses du théorème 1.27 sont vérifiées, on peut utiliser les traces $\gamma_0 v$ et $\gamma_0 w$ pour donner un sens à l'intégrale de bord

$$\int_{\partial\Omega} v\overline{w} n_i\, d\Gamma.$$

Le résultat et sa démonstration sont proposés ci-dessous.

Proposition 1.29 *Soit* Ω *un ouvert de* $\mathbb{R}^{\mathtt{d}}$ *à frontière lipschitzienne et bornée. Soient* $v, w \in H^1(\Omega)$, *on a pour tout* $i = 1, \mathtt{d}$ *:*

$$\int_\Omega \left(\frac{\partial v}{\partial x_i}\, \overline{w} + v\, \frac{\partial \overline{w}}{\partial x_i}\right) d\Omega = \int_{\partial\Omega} \gamma_0 v\, \overline{\gamma_0 w}\, n_i\, d\Gamma. \tag{1.18}$$

Démonstration : D'après la densité de $\mathcal{C}_c^\infty(\overline{\Omega})$ dans $H^1(\Omega)$, il existe deux suites de $\left(\mathcal{C}_c^\infty(\overline{\Omega})\right)^{\mathbb{N}}$, notées respectivement $(v_k)_k$ et $(w_k)_k$, telles que

$$\lim_{k\to\infty} \|v - v_k\|_{H^1(\Omega)} = \lim_{k\to\infty} \|w - w_k\|_{H^1(\Omega)} = 0.$$

Pour chaque $k \in \mathbb{N}$, par définition de l'application trace γ_0, on a d'après (1.14) la formule d'intégration par parties :

$$\int_\Omega \left(\frac{\partial v_k}{\partial x_i}\, \overline{w_k} + v_k\, \frac{\partial \overline{w_k}}{\partial x_i} \right) d\Omega = \int_{\partial\Omega} \gamma_0 v_k\, \overline{\gamma_0 w_k}\, n_i \, d\Gamma.$$

Puisque $(\partial_i v_k)_k$ converge vers $\partial_i v$ dans $L^2(\Omega)$, et que $(\overline{w_k})_k$ converge vers $\overline{w}$ dans $L^2(\Omega)$, on déduit de la proposition 1.5 que $(\partial_i v_k \overline{w_k})_k$ converge vers $\partial_i v \overline{w}$ dans $L^1(\Omega)$. De même, $(v_k \partial_i \overline{w_k})_k$ converge vers $v\partial_i \overline{w}$ dans $L^1(\Omega)$. Il suit que

$$\lim_{k\to\infty} \left(\int_\Omega \left(\frac{\partial v_k}{\partial x_i}\, \overline{w_k} + v_k\, \frac{\partial \overline{w_k}}{\partial x_i} \right) d\Omega \right) = \int_\Omega \left(\frac{\partial v}{\partial x_i}\, \overline{w} + v\, \frac{\partial \overline{w}}{\partial x_i} \right) d\Omega.$$

Puis, par continuité de l'application trace γ_0, on note que $(\gamma_0 v_k)_k$ converge vers $\gamma_0 v$ dans $L^2(\partial\Omega)$, et de même que $(\gamma_0 w_k)_k$ converge vers $\gamma_0 w$ dans $L^2(\partial\Omega)$. Il suit que $(\gamma_0 v_k\, \overline{\gamma_0 w_k})_k$ converge vers $\gamma_0 v\, \overline{\gamma_0 w}$ dans $L^1(\partial\Omega)$. Et comme $n_i \in L^\infty(\partial\Omega)$ (voir la proposition 1.23), on a finalement que $(\gamma_0 v_k\, \overline{\gamma_0 w_k}\, n_i)_k$ converge vers $\gamma_0 v\, \overline{\gamma_0 w}\, n_i$ dans $L^1(\partial\Omega)$. En d'autres termes,

$$\lim_{k\to\infty} \left(\int_{\partial\Omega} \gamma_0 v_k\, \overline{\gamma_0 w_k}\, n_i \, d\Gamma \right) = \int_{\partial\Omega} \gamma_0 v\, \overline{\gamma_0 w}\, n_i \, d\Gamma,$$

ce qui démontre (1.18). ∎

Théorème 1.30 *Soit Ω un ouvert de $\mathbb{R}^{\mathtt{d}}$ à frontière lipschitzienne et bornée. On a l'identification :*

$$H_0^1(\Omega) = \left\{ v \in H^1(\Omega) \text{ tel que } \gamma_0 v = 0 \right\}.$$

1.5 Opérateur de trace normale, formule de Green généralisée

Il reste maintenant à généraliser la formule d'intégration par parties (1.15). Après quelques considérations intuitives, on suivra [18].

1.5.1 Un peu d'intuition...

Si on cherche à nouveau l'énoncé intuitif du résultat, on constate cette fois une dissymétrie entre l'espace des $\boldsymbol{v}$, et l'espace des w. Partant de $w \in L^2(\Omega)$, on constate qu'on doit choisir $\operatorname{div} \boldsymbol{v} \in L^2(\Omega)$ pour que

$$\int_\Omega (\operatorname{div} \boldsymbol{v}) \overline{w}\, d\Omega$$

ait un sens. Si maintenant on part de $\boldsymbol{v} \in \big(L^2(\Omega)\big)^{\mathtt{d}}$, on doit choisir $(\partial_i w)_{i=1,\mathtt{d}} \in \big(L^2(\Omega)\big)^{\mathtt{d}}$ pour que

$$\int_\Omega \boldsymbol{v} \cdot \nabla \overline{w}\, d\Omega$$

ait un sens. Ainsi, w appartient à $H^1(\Omega)$, alors que $\boldsymbol{v}$ appartient à l'espace de Sobolev

$$\boldsymbol{H}(\mathrm{div}\,;\Omega) = \Big\{ \boldsymbol{v} \in \big(L^2(\Omega)\big)^{\mathtt{d}} \text{ tel que } \mathrm{div}\,\boldsymbol{v} \in L^2(\Omega) \Big\},$$

où la divergence est comprise au sens des distributions.
Muni de la norme

$$\|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} = \Big(\|\boldsymbol{v}\|^2_{(L^2(\Omega))^{\mathtt{d}}} + \|\mathrm{div}\,\boldsymbol{v}\|^2_{L^2(\Omega)} \Big)^{1/2},$$

$\boldsymbol{H}(\mathrm{div}\,;\Omega)$ est un espace de Banach.
Muni du produit scalaire associé

$$(\boldsymbol{v},\boldsymbol{w})_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} = (\boldsymbol{v},\boldsymbol{w})_{(L^2(\Omega))^{\mathtt{d}}} + (\mathrm{div}\,\boldsymbol{v}, \mathrm{div}\,\boldsymbol{w})_{L^2(\Omega)},$$

$\boldsymbol{H}(\mathrm{div}\,;\Omega)$ est un espace de Hilbert.
Ensuite, il reste à considérer l'intégrale de bord

$$\int_{\partial\Omega} (\boldsymbol{v}\cdot\boldsymbol{n})\overline{w}\, d\Gamma.$$

Si les hypothèses du théorème 1.27 sont vérifiées, on a vu que $\gamma_0 w$ appartient à $L^2(\partial\Omega)$, l'idée naturelle consiste donc à établir que le flux $\boldsymbol{v}\cdot\boldsymbol{n}_{|\partial\Omega}$ appartient également à $L^2(\partial\Omega)$. Malheureusement, ce résultat est faux si on considère un élément quelconque de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$. [5] Néanmoins, on sait que $\gamma_0 w$ appartient à $H^{1/2}(\partial\Omega)$, qui est un sous-espace strict de $L^2(\partial\Omega)$. On va donc procéder *par dualité*, c'est-à-dire démontrer que le flux $\boldsymbol{v}\cdot\boldsymbol{n}_{|\partial\Omega}$ appartient à $\big(H^{1/2}(\partial\Omega)\big)'$, l'espace dual de $H^{1/2}(\partial\Omega)$. On notera l'action d'un élément de $(H^{1/2}(\partial\Omega))'$ sur un élément de $H^{1/2}(\partial\Omega)$ à l'aide de crochets de dualité : $\langle\cdot,\cdot\rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)}$.
Puisque $L^2(\partial\Omega)$ est identifié à son dual, on rappelle qu'on a les inclusions strictes (avec injections continues)

$$H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega) \subset \big(H^{1/2}(\partial\Omega)\big)'. \tag{1.19}$$

Pour un couple $(\lambda_0,\lambda_1) \in L^2(\partial\Omega)\times H^{1/2}(\partial\Omega)$, on peut écrire

$$\langle\lambda_0,\lambda_1\rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)} = (\lambda_0,\lambda_1)_{L^2(\partial\Omega)} = \int_{\partial\Omega} \lambda_0\, \overline{\lambda_1}\, d\Gamma.$$

1.5.2 Résultat de densité

On a un nouveau résultat de densité des fonctions régulières.

5. Le résultat est bien sûr vrai si $\boldsymbol{v}$ est "régulière jusqu'au bord", par exemple $\boldsymbol{v} \in \mathcal{C}^0(\overline{\Omega})$.

Proposition 1.31 *Dans tout ouvert Ω à frontière lipschitzienne, $(\mathcal{C}_c^\infty(\overline{\Omega}))^{\mathtt{d}}$ est dense dans $\boldsymbol{H}(\mathrm{div}\,;\Omega)$:*

$$\forall \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}\,;\Omega),\ \exists (\boldsymbol{v}_k)_k \in \left((\mathcal{C}_c^\infty(\overline{\Omega}))^{\mathtt{d}}\right)^{\mathbb{N}},\ \lim_{k\to\infty} \|\boldsymbol{v}-\boldsymbol{v}_k\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} = 0.$$

A partir de là, on va étudier la trace normale, qui apparaît dans la formule d'intégration par parties (1.15), pour des éléments de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$.

1.5.3 Opérateur de trace normale γ_n

Soit Ω un ouvert, à frontière $\partial\Omega$ lipschitzienne et bornée. Pour tout $\boldsymbol{v} \in (\mathcal{C}_c^\infty(\overline{\Omega}))^{\mathtt{d}}$ on peut définir sa trace normale sur $\partial\Omega$: $(\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}$. Lorsqu'on prolonge par continuité aux éléments de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$, comme espéré on aboutit cette fois à une application trace normale à valeurs dans le dual de $H^{1/2}(\partial\Omega)$. Enfin, on peut généraliser la formule d'intégration par parties (1.15).

Théorème 1.32 *Soit Ω un ouvert de $\mathbb{R}^{\mathtt{d}}$, à frontière lipschitzienne et bornée. Alors, on peut définir l'application trace normale*

$$\gamma_n : \left| \begin{array}{ccc} \left(\mathcal{C}_c^\infty\left(\overline{\Omega}\right)\right)^{\mathtt{d}} & \to & (H^{1/2}(\partial\Omega))' \\ \boldsymbol{v} & \mapsto & \gamma_n\boldsymbol{v} = (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega} \end{array} \right. .$$

De plus, elle se prolonge par continuité en une application linéaire continue et surjective, encore notée γ_n, de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ dans $(H^{1/2}(\partial\Omega))'$.
Soient $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}\,;\Omega)$ et $w \in H^1(\Omega)$, on a la formule d'intégration par parties :

$$\int_\Omega ((\mathrm{div}\,\boldsymbol{v})\overline{w} + \boldsymbol{v}\cdot\nabla\overline{w})\,d\Omega = \langle \gamma_n\boldsymbol{v}, \gamma_0 w\rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)}. \tag{1.20}$$

Démonstration : On raisonne par étapes.

— **Trace normale pour $\boldsymbol{v} \in \left(\mathcal{C}_c^\infty(\overline{\Omega})\right)^{\mathtt{d}}$.** Comme l'espace $H^{1/2}(\partial\Omega)$ est égal à l'espace des traces de $H^1(\Omega)$, on choisit tout d'abord $w \in H^1(\Omega)$. Si on raisonne comme pour démontrer la proposition 1.29 à l'aide d'une suite de fonctions de $\left(\mathcal{C}_c^\infty(\overline{\Omega})\right)^{\mathbb{N}}$ tendant vers w dans $H^1(\Omega)$, on montre que

$$\int_{\partial\Omega} (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}\,\overline{\gamma_0 w}\,d\Gamma = \int_\Omega (\boldsymbol{v}\cdot\nabla\overline{w} + (\mathrm{div}\,\boldsymbol{v})\overline{w})\,d\Omega.$$

En appliquant l'inégalité de Cauchy-Schwarz à deux reprises, on en déduit que

$$\begin{aligned} \left|\int_{\partial\Omega} (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}\,\overline{\gamma_0 w}\,d\Gamma\right| &\le \|\boldsymbol{v}\|_{(L^2(\Omega))^{\mathtt{d}}}\|\nabla w\|_{(L^2(\Omega))^{\mathtt{d}}} + \|\mathrm{div}\,\boldsymbol{v}\|_{L^2(\Omega)}\|w\|_{L^2(\Omega)} \\ &\le \|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)}\|w\|_{H^1(\Omega)}. \end{aligned}$$

L'application trace γ_0 est linéaire, continue et surjective de $H^1(\Omega)$ dans $H^{1/2}(\partial\Omega)$: pour tout $\lambda \in H^{1/2}(\partial\Omega)$, on applique la majoration ci-dessus aux $w \in H^1(\Omega)$ tels que $\gamma_0 w = \lambda$. Il suit

$$\left|\int_{\partial\Omega} (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}\,\overline{\lambda}\,d\Gamma\right| \le \|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} \inf_{w\in H^1(\Omega)\ tq\ \gamma_0 w=\lambda} \|w\|_{H^1(\Omega)}.$$

En utilisant la définition (1.17) de la norme dans $H^{1/2}(\partial\Omega)$, on trouve

$$\forall \lambda \in H^{1/2}(\partial\Omega), \quad \Big| \int_{\partial\Omega} (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}\, \overline{\lambda}\, d\Gamma \Big| \leq \|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} \|\lambda\|_{H^{1/2}(\partial\Omega)}.$$

Ainsi, on a bien $(\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega} \in (H^{1/2}(\partial\Omega))'$, avec l'identification

$$\langle (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}, \lambda \rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)} = \int_{\partial\Omega} (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}\, \overline{\lambda}\, d\Gamma.$$

— **Application trace normale dans $\mathcal{L}\big(\boldsymbol{H}(\mathrm{div}\,;\Omega), (H^{1/2}(\partial\Omega))'\big)$.** D'après ce qui précède, on a pour tout $\boldsymbol{v} \in (\mathcal{C}_c^\infty(\overline{\Omega}))^{\mathtt{d}}$:

$$\|\gamma_n \boldsymbol{v}\|_{(H^{1/2}(\partial\Omega))'} = \sup_{\lambda \in H^{1/2}(\partial\Omega)\setminus\{0\}} \frac{\Big| \displaystyle\int_{\partial\Omega} (\boldsymbol{v}\cdot\boldsymbol{n})_{|\partial\Omega}\, \overline{\lambda}\, d\Gamma \Big|}{\|\lambda\|_{H^{1/2}(\partial\Omega)}} \leq \|\boldsymbol{v}\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)}.$$

L'application γ_n est donc continue de $(\mathcal{C}_c^\infty(\overline{\Omega}))^{\mathtt{d}}$ dans $(H^{1/2}(\partial\Omega))'$ *par rapport à la norme de* $\boldsymbol{H}(\mathrm{div}\,;\Omega)$. D'après la proposition 1.31, on peut prolonger par continuité γ_n en une application linéaire et continue de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ dans $(H^{1/2}(\partial\Omega))'$, toujours notée γ_n. On note que $|||\gamma_n||| \leq 1$ dans $\mathcal{L}\big(\boldsymbol{H}(\mathrm{div}\,;\Omega), (H^{1/2}(\partial\Omega))'\big)$.

— **Intégration par parties (1.20)**. Il suffit maintenant de raisonner par densité sur $\boldsymbol{v}$ et sur w, puisque le terme de bord aura un sens par dualité, sous la forme $\langle \gamma_n \boldsymbol{v}, \gamma_0 w \rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)}$.

— **Surjectivité de $\gamma_n \in \mathcal{L}\big(\boldsymbol{H}(\mathrm{div}\,;\Omega), (H^{1/2}(\partial\Omega))'\big)$.** Soit $g \in (H^{1/2}(\partial\Omega))'$, on considère la formulation variationnelle

$$\begin{cases} \text{Trouver } u \in H^1(\Omega) \text{ telle que} \\ (u,w)_{H^1(\Omega)} = \langle g, \gamma_0 w \rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)}, \quad \forall w \in H^1(\Omega) \end{cases}.$$

D'après le théorème de représentation de Riesz, cette formulation variationnelle admet une solution, et une seule. En outre, si on choisit une fonction-test $v \in \mathcal{D}(\Omega)$ quelconque, on en déduit facilement que $-u + \Delta u = 0$ dans $\mathcal{D}'(\Omega)$. Si on pose $\boldsymbol{v} = \nabla u$, on a $\boldsymbol{v} \in (L^2(\Omega))^{\mathtt{d}}$ puisque $u \in H^1(\Omega)$ et, en outre, $\mathrm{div}\,\boldsymbol{v} = \Delta u = u$ appartient à $L^2(\Omega)$: ainsi $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}\,;\Omega)$. A partir de là, si on remplace u par $\mathrm{div}\,\boldsymbol{v}$ et ∇u par $\boldsymbol{v}$ dans la formulation variationnelle, on trouve que pour tout $w \in H^1(\Omega)$,

$$\langle g, \gamma_0 w \rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)} = (u,w)_{L^2(\Omega)} + (\nabla u, \nabla w)_{(L^2(\Omega))^{\mathtt{d}}} = (\mathrm{div}\,\boldsymbol{v}, w)_{L^2(\Omega)} + (\boldsymbol{v}, \nabla w)_{(L^2(\Omega))^{\mathtt{d}}}.$$

Et si on utilise la formule d'intégration par parties (1.20), on a donc établi que

$$\forall w \in H^1(\Omega), \quad \langle g, \gamma_0 w \rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)} = \langle \gamma_n \boldsymbol{v}, \gamma_0 w \rangle_{(H^{1/2}(\partial\Omega))',H^{1/2}(\partial\Omega)}.$$

Et comme l'application trace γ_0 est surjective de $H^1(\Omega)$ dans $H^{1/2}(\partial\Omega)$, on a démontré que $\gamma_n \boldsymbol{v} = g$ dans $(H^{1/2}(\partial\Omega))'$, d'où la propriété de surjectivité.

Ceci achève la démonstration. ∎

Remarque 1.33 *On note que si $w \in H_0^1(\Omega)$, le terme de bord disparaît dans (1.20). Ce résultat s'obtient trivialement si on raisonne par densité, en approchant w dans $H^1(\Omega)$ avec des éléments de $\mathcal{D}(\Omega)$.*

Enfin, l'espace de Hilbert $\boldsymbol{H}_0(\mathrm{div}\,;\Omega)$ est défini comme l'adhérence de $(\mathcal{D}(\Omega))^{\mathtt{d}}$ dans $\boldsymbol{H}(\mathrm{div}\,;\Omega)$:

$$\forall \boldsymbol{v} \in \boldsymbol{H}_0(\mathrm{div}\,;\Omega),\ \exists (\boldsymbol{v}_k)_k \in \left((\mathcal{D}(\Omega))^{\mathtt{d}}\right)^{\mathbb{N}},\ \lim_{k\to\infty} \|\boldsymbol{v}-\boldsymbol{v}_k\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} = 0.$$

On a le

Théorème 1.34 *Soit Ω est un ouvert de $\mathbb{R}^{\mathtt{d}}$ à frontière lipschitzienne et bornée. Alors on a l'identification :*

$$\boldsymbol{H}_0(\mathrm{div}\,;\Omega) = \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}\,;\Omega) \text{ tel que } \gamma_n \boldsymbol{v} = 0\}\,.$$

Remarque 1.35 *On note que si $\boldsymbol{v} \in \boldsymbol{H}_0(\mathrm{div}\,;\Omega)$, le terme de bord disparaît dans (1.20). Ce résultat s'obtient trivialement si on raisonne par densité, en approchant $\boldsymbol{v}$ dans $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ avec des éléments de $(\mathcal{D}(\Omega))^{\mathtt{d}}$.*

1.5.4 Propriétés de saut

Selon la régularité du champ considéré, on a des propriétés différentes de saut au franchissement d'une interface. Un résultat simplifié peut être obtenu facilement pour des champs réguliers (au sens classique) "par morceaux", voir l'exercice ci-après.

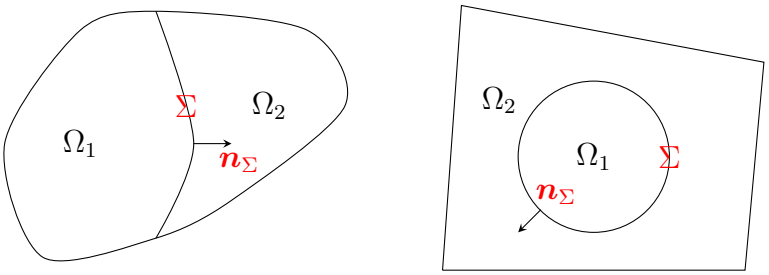

FIGURE 1.1 – Exemples de partition d'un domaine de $\mathbb{R}^2$

Exercice : Soit Ω un domaine de $\mathbb{R}^{\mathtt{d}}$. On suppose qu'il existe deux domaines disjoints Ω_1 et Ω_2 tels que $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$. On note $\Sigma = \partial\Omega_1 \cap \partial\Omega_2$ l'interface [6] et $\boldsymbol{n}_\Sigma$ un vecteur unitaire normal à Σ dirigé de Ω_1 vers Ω_2. Voir des exemples sur la figure 1.1.

1. Soit $v \in L^2(\Omega)$ telle que $v_{|\overline{\Omega_j}} \in \mathcal{C}^1(\overline{\Omega_j})$ pour $j = 1, 2$. Montrer que si le saut [7] de v s'annule, c'est-à-dire si $[v]_\Sigma = 0$, alors on a $v \in H^1(\Omega)$. Que pensez-vous de la réciproque ?

6. En pratique l'interface est une courbe si $\mathtt{d} = 2$, et une surface si $\mathtt{d} = 3$.

7. Pour w régulière de part et d'autre de l'interface Σ séparant Ω_1 et Ω_2, le *saut* de w au travers de Σ (conforme à l'orientation du vecteur unitaire normal $\boldsymbol{n}_\Sigma$ dirigé de Ω_1 vers Ω_2) est égal à

$$[w]_\Sigma := w_{|\partial\Omega_2} - w_{|\partial\Omega_1}.$$

2. Soit $\boldsymbol{v} \in \big(L^2(\Omega)\big)^{\mathtt{d}}$ telle que $\boldsymbol{v}_{|\overline{\Omega_j}} \in \big(\mathcal{C}^1(\overline{\Omega_j})\big)^{\mathtt{d}}$ pour $j = 1, 2$. Montrer que si le saut de la trace normale de $\boldsymbol{v}$ s'annule, c'est-à-dire si $[\boldsymbol{v} \cdot \boldsymbol{n}_\Sigma]_\Sigma = 0$, alors on a $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}\,;\Omega)$. Que pensez-vous de la réciproque ?

Plus généralement, on a les résultats ci-dessous sur les sauts.

Théorème 1.36 *Soit Ω un domaine de $\mathbb{R}^{\mathtt{d}}$. On suppose qu'il existe deux domaines disjoints Ω_1 et Ω_2 tels que $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$. On note $\Sigma = \partial\Omega_1 \cap \partial\Omega_2$ l'interface. Alors :*

— *pour $v \in L^2(\Omega)$ telle que $v_{|\Omega_j} \in H^1(\Omega_j)$ pour $j = 1, 2$, on a l'équivalence*

$$v \in H^1(\Omega) \quad \Longleftrightarrow \quad [v]_\Sigma = 0. \tag{1.21}$$

— *pour $\boldsymbol{v} \in \big(L^2(\Omega)\big)^{\mathtt{d}}$ telle que $\boldsymbol{v}_{|\Omega_j} \in \boldsymbol{H}(\mathrm{div}\,;\Omega_j)$ pour $j = 1, 2$, on a l'équivalence*

$$\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}\,;\Omega) \quad \Longleftrightarrow \quad [\boldsymbol{v} \cdot \boldsymbol{n}_\Sigma]_\Sigma = 0. \tag{1.22}$$

Notes de cours 2

Etude théorique du modèle de la diffusion

L'équation de la diffusion permet de modéliser différents phénomènes physiques tels que la loi de Darcy, la loi de Fick ou la diffusion des neutrons. Pour la diffusion neutronique par exemple (voir par exemple [9]), elle permet de déterminer le *flux neutronique* (grandeur à valeurs scalaires réelles). On peut choisir soit de l'exprimer comme une équation aux dérivées partielles (EDP) du second ordre, soit comme deux équations aux dérivées partielles du premier ordre couplées entre elles. Dans ce second cas, on détermine, outre le flux neutronique, le *courant neutronique* (grandeur à valeurs vectorielles réelles proportionnelle au gradient du flux neutronique). Ainsi, d'un point de vue variationnel, deux approches coexistent. On utilise soit la formulation variationnelle à une inconnue pour se concentrer sur le flux neutronique ; soit la formulation variationnelle à deux inconnues pour se concentrer à la fois sur le flux et le courant neutroniques.

On se place dans un domaine Ω de $\mathbb{R}^{\mathtt{d}}$, pour $1 \leq \mathtt{d} \leq 3$. Le modèle étudié est basé sur des données et des solutions à valeurs réelles. On choisit donc $\mathbb{K} = \mathbb{R}$ lorsqu'on l'étudie, en particulier pour les formulations variationnelles. Concernant la théorie "abstraite", on se placera a priori dans $\mathbb{K} = \mathbb{C}$, et on considèrera le cas $\mathbb{K} = \mathbb{R}$ plus brièvement. On note $(\cdot|\cdot)$ le produit scalaire (hermitien) de $L^2(\Omega)$ ou de $(L^2(\Omega))^{\mathtt{d}}$, et $\|\cdot\|$ la norme associée. Enfin, on note $\boldsymbol{L}^2(\Omega) = (L^2(\Omega))^{\mathtt{d}}$.

2.1 Modèle de la diffusion neutronique

2.1.1 Mesure, équation et condition aux limites

On reprend le formalisme énoncé au §1.1. Résoudre le modèle de la diffusion revient à résoudre :

$$\begin{cases} \text{Trouver } u \in H^1(\Omega) \text{ tel que} \\ -\operatorname{div} \mathbb{D}\nabla u + \sigma u = S_f \text{ dans } \Omega \\ u = 0 \text{ sur } \partial\Omega. \end{cases} \tag{2.1}$$

La solution u cherchée est le *flux neutronique*, $\mathbb{D}$ est le *tenseur de diffusion*, σ la *section efficace*, et la donnée S_f la *source*.[1] Toutes les grandeurs sont à valeurs réelles. De plus, la modélisation conduit à faire les hypothèses suivantes :

— Le coefficient $\mathbb{D}$ est un champ de matrices de $\mathbb{R}^{\mathtt{d}\times\mathtt{d}}$ mesurable sur Ω tel que

$$\left|\begin{array}{l} \exists D_{min}>0,\ \forall \boldsymbol{z}\in\mathbb{R}^{\mathtt{d}},\ \text{presque pour tout } \boldsymbol{x}\in\Omega,\ D_{min}|\boldsymbol{z}|^2\le \mathbb{D}(\boldsymbol{x})\boldsymbol{z}\cdot\boldsymbol{z}\,; \\ \exists D_{max}>0,\ \forall \boldsymbol{z}\in\mathbb{R}^{\mathtt{d}},\ \text{presque pour tout } \boldsymbol{x}\in\Omega,\ |\mathbb{D}(\boldsymbol{x})\boldsymbol{z}|\le D_{max}|\boldsymbol{z}|. \end{array}\right. \tag{2.2}$$

— Le coefficient σ est un champ de scalaires mesurable sur Ω tel que

$$\exists \sigma_{min},\sigma_{max}>0,\ \text{presque pour tout } \boldsymbol{x}\in\Omega,\ \sigma_{min}\le\sigma(\boldsymbol{x})\le\sigma_{max}. \tag{2.3}$$

— La source S_f appartient à $L^2(\Omega)$.

Sous les hypothèses sur u et sur $\mathbb{D}$, on note que le *courant neutronique* $\boldsymbol{p}=-\mathbb{D}\nabla u$ est tel que $\boldsymbol{p}\in\boldsymbol{L}^2(\Omega)$ (voir la proposition 1.4). De plus, sous les hypothèses sur u, σ et S_f, on a $\operatorname{div}\boldsymbol{p}=S_f-\sigma u\in L^2(\Omega)$. Ainsi, le courant neutronique est tel que

$$\boldsymbol{p}\in\boldsymbol{H}(\operatorname{div};\Omega).$$

2.1.2 Formulation variationnelle équivalente

La condition aux limites sur u est une condition aux limites de Dirichlet. Classiquement [15], c'est une condition aux limites *essentielle*, et on choisit des fonctions-test $w\in H^1_0(\Omega)$.
Chaque terme de l'EDP apparaissant dans le modèle (2.1) appartient à $L^2(\Omega)$. On réalise donc le "produit scalaire dans $L^2(\Omega)$" de l'EDP avec $w\in H^1_0(\Omega)$, pour trouver[2] :

$$\int_\Omega S_f w\,d\Omega=\int_\Omega(-\operatorname{div}\mathbb{D}\nabla u+\sigma u)w\,d\Omega\overset{(1.20)}{=}\int_\Omega\Big(\mathbb{D}\nabla u\cdot\nabla w+\sigma uw\Big)d\Omega.$$

On en conclut que si u est solution du modèle (2.1), alors u est solution de la formulation variationnelle

$$\left\{\begin{array}{l} \text{Trouver } u\in H^1_0(\Omega) \text{ tel que} \\ \forall w\in H^1_0(\Omega),\quad \displaystyle\int_\Omega\Big(\mathbb{D}\nabla u\cdot\nabla w+\sigma uw\Big)d\Omega=\int_\Omega S_f w\,d\Omega. \end{array}\right. \tag{2.4}$$

Réciproquement, en raisonnant au sens des distributions, on montre facilement que si u est solution de la formulation variationnelle (2.4), alors u est solution du modèle (2.1). On a donc bien construit *une formulation variationnelle équivalente à notre modèle*. Nous prouvons son caractère bien posé ci-dessous à l'aide du théorème de Lax-Milgram.

1. On peut également modéliser la diffusion neutronique avec une autre condition aux limites que celle de flux neutronique nul, telle que $\mathbb{D}\nabla u\cdot\boldsymbol{n}_{|\partial\Omega}=0$ (réflexion), ou $(\mathbb{D}\nabla u\cdot\boldsymbol{n}+\alpha u)_{|\partial\Omega}=0$ (albédo, avec $\alpha>0$) ; ou bien des conditions aux limites distinctes sur différentes parties de la frontière...

2. Au lieu de la formule d'intégration par parties (1.20), on peut utiliser la remarque 1.33.

2.2 Existence et unicité, théorème de Lax-Milgram

On rappelle les fondements de la théorie "abstraite" (voir par exemple [8, 15]). Soient :

— V un espace de Hilbert défini sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, de produit scalaire (hermitien) $(\cdot,\cdot)_V$. On note $\|\cdot\|_V$ la norme associée ;

— $a(\cdot,\cdot)$ une forme, bilinéaire si $\mathbb{K} = \mathbb{R}$, sesquilinéaire si $\mathbb{K} = \mathbb{C}$, et continue sur $V \times V$.

Pour $\ell \in V'$ donné, on étudie le problème écrit sous forme variationnelle,

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ \forall w \in V, \quad a(u,w) = \langle \ell, w \rangle_V. \end{cases} \tag{2.5}$$

Pour le modèle de la diffusion, on a respectivement (avec $\mathbb{K} = \mathbb{R}$) :

— $V = H_0^1(\Omega)$;

— $a(v,w) = (\mathbb{D}\nabla v | \nabla w) + (\sigma v | w)$;

— $\langle \ell, w \rangle_V = (S_f | w)$.

2.2.1 Théorème de Lax-Milgram

Définition 2.1 (coercivité) *On dit qu'une forme a, sesquilinéaire si $\mathbb{K} = \mathbb{C}$, bilinéaire si $\mathbb{K} = \mathbb{R}$, est coercive sur V si, et seulement si :*

$$\exists \alpha > 0, \quad \forall v \in V, \quad |a(v,v)| \geq \alpha \|v\|_V^2. \tag{2.6}$$

Théorème 2.2 (Lax-Milgram) *Soit V un espace de Hilbert sur $\mathbb{K}$, a une forme sesquilinéaire si $\mathbb{K} = \mathbb{C}$, bilinéaire si $\mathbb{K} = \mathbb{R}$, continue et coercive sur V.*
Pour tout $\ell \in V'$, il existe une unique solution $u \in V$ de la formulation variationnelle (2.5). En outre, la solution u dépend continûment de la forme linéaire ℓ : il existe $C > 0$ indépendante de u et ℓ telle que

$$\|u\|_V \leq C \, \|\ell\|_{V'}.$$

Lorsque $\mathbb{K} = \mathbb{C}$, on peut établir une caractérisation *équivalente* de la coercivité, voir [8].

Lemme 2.3 *[$\mathbb{K} = \mathbb{C}$] Soit a une forme sesquilinéaire et continue sur $V \times V$. Alors a est coercive sur V si, et seulement si :*

$$\exists \alpha > 0, \quad \exists \theta \in \mathbb{R}, \quad \forall v \in V, \quad \Re\big(\exp(\imath\theta) a(v,v)\big) \geq \alpha \|v\|_V^2. \tag{2.7}$$

Si la forme est de plus *hermitienne*, c'est-à-dire si

$$\forall v, w \in V, \quad a(v,w) = \overline{a(w,v)},$$

on a le

Lemme 2.4 */$\mathbb{K} = \mathbb{C}$/ Soit a une forme sesquilinéaire, continue et hermitienne sur $V \times V$. Alors a est coercive sur V si, et seulement si, on a une des deux alternatives :*

$$\begin{aligned} &\exists \alpha > 0,\ \forall v \in V,\ +a(v,v) \geq \alpha \|v\|_V^2 \,; \\ &\exists \alpha > 0,\ \forall v \in V,\ -a(v,v) \geq \alpha \|v\|_V^2. \end{aligned} \tag{2.8}$$

Démonstration : Si on a l'alternative (2.8), il est clair que a est coercive.
Réciproquement, supposons a coercive. On note tout d'abord que, puisque a est hermitienne, pour tout $v \in V$, on a $a(v,v) = \overline{a(v,v)}$. Ainsi, pour tout $v \in V$, on a $a(v,v) \in \mathbb{R}$. Supposons maintenant qu'il existe $v_1, v_2 \in V$ tels que $a(v_1,v_1) > 0$ et $a(v_2,v_2) < 0$. Notons que v_1 et v_2 sont non-nuls puisque a est sesquilinéaire ; de plus, v_1 et v_2 ne sont pas colinéaires : sinon, $\exists \beta \in \mathbb{C}$ tel que $v_1 = \beta v_2$ et dans ce cas par linéarité par rapport au 1er argument, resp. antilinéarité par rapport au 2ème argument,

$$a(v_1,v_1) = a(\beta v_2, \beta v_2) = |\beta|^2 a(v_2,v_2) \leq 0,$$

ce qui contredit le fait que $a(v_1,v_1) > 0$.
On introduit la fonction continue de $[0,1]$ dans $\mathbb{R}$

$$f \ : \ t \mapsto a(tv_1 + (1-t)v_2, tv_1 + (1-t)v_2).$$

Par définition, $f(0) < 0$ et $f(1) > 0$. Puisque f est continue, $\exists t^\star \in]0,1[$ tel que $f(t^\star) = 0$. Si on pose $v^\star = t^\star v_1 + (1-t^\star)v_2$, on conclut de ce qui précède que $v^\star \neq 0$ (v_1 et v_2 ne sont pas colinéaires), et que $a(v^\star, v^\star) = 0$, ce qui contredit le fait que a est coercive. Ainsi, l'alternative (2.8) est vérifiée. ■

Lorsque $\mathbb{K} = \mathbb{R}$, on peut facilement démontrer l'équivalent du lemme 2.4, sans hypothèse de symétrie de la forme.

Lemme 2.5 */$\mathbb{K} = \mathbb{R}$/ Soit a une forme bilinéaire et continue sur $V \times V$. Alors a est coercive sur V si, et seulement si, on a une des deux alternatives :*

$$\begin{aligned} &\exists \alpha > 0,\ \forall v \in V,\ +a(v,v) \geq \alpha \|v\|_V^2 \,; \\ &\exists \alpha > 0,\ \forall v \in V,\ -a(v,v) \geq \alpha \|v\|_V^2. \end{aligned} \tag{2.9}$$

Si a est de plus *symétrique*, c'est-à-dire que :

$$\forall v, w \in V, \quad a(v,w) = a(w,v),$$

alors on peut montrer que la solution u de (2.5) est l'unique point de minimum de la fonctionnelle :

$$v \ \mapsto \ \frac{1}{2} a(v,v) - \langle \ell, v \rangle_V,$$

définie de V dans $\mathbb{R}$.

2.2.2 Diffusion à une inconnue

Pour le modèle de la diffusion posé sur $\mathbb{R}$, on doit donc vérifier que la forme bilinéaire

$$a(v,w) = \int_\Omega \Big(\mathbb{D} \nabla v \cdot \nabla w + \sigma v w \Big) d\Omega$$

est continue et coercive sur $H_0^1(\Omega)$ (sous les hypothèses (2.2)-(2.3) sur les coefficients).

— Continuité : soient $v, w \in H^1_0(\Omega)$, on a

$$\begin{aligned}
|a(v,w)| &\leq |(\mathbb{D}\nabla v|\nabla w)| + |(\sigma v|w)| \\
\text{(Cauchy-Schwarz ds } \boldsymbol{L}^2(\Omega),\, L^2(\Omega)) &\leq \|\mathbb{D}\nabla v\|\,\|\nabla w\| + \|\sigma v\|\,\|w\| \\
\text{(2.2)-(2.3)} &\leq D_{max}\,\|\nabla v\|\,\|\nabla w\| + \sigma_{max}\,\|v\|\,\|w\| \\
&\leq \max(D_{max},\sigma_{max})\,(\|\nabla v\|\,\|\nabla w\| + \|v\|\,\|w\|) \\
\text{(Cauchy-Schwarz ds } \mathbb{R}^2) &\leq \max(D_{max},\sigma_{max})\,\left(\|\nabla v\|^2 + \|v\|^2\right)^{1/2}\left(\|\nabla w\|^2 + \|w\|^2\right)^{1/2} \\
&= \max(D_{max},\sigma_{max})\,\|v\|_{H^1(\Omega)}\,\|w\|_{H^1(\Omega)}\,;
\end{aligned}$$

d'où la continuité de a avec $|||a||| \leq \max(D_{max}, \sigma_{max})$.

— Coercivité : soit $v \in H^1_0(\Omega)$, on a

$$\begin{aligned}
a(v,v) &= (\mathbb{D}\nabla v|\nabla v) + (\sigma v|v) \\
\text{(2.2)-(2.3)} &\geq D_{min}\,\|\nabla v\|^2 + \sigma_{min}\,\|v\|^2 \\
&\geq \min(D_{min},\sigma_{min})\,\|v\|^2_{H^1(\Omega)}\,;
\end{aligned}$$

d'où la coercivité de a, avec une constante de coercivité $\alpha = \min(D_{min}, \sigma_{min}) > 0$.

Par ailleurs, d'après la proposition 1.6, $L^2(\Omega)$ est son propre dual. Comme $H^1_0(\Omega)$ est un sous-espace strict, dense dans $L^2(\Omega)$ avec injection continue, on a l'injection continue "duale" $L^2(\Omega) \subset (H^1_0(\Omega))'$, et en particulier $w \mapsto (S_f|w)$ appartient à $(H^1_0(\Omega))'$ puisque $S_f \in L^2(\Omega)$. D'après le théorème de Lax-Milgram, on en déduit que la formulation variationnelle (2.4) admet une solution et une seule. De plus, il existe une constante $C > 0$ indépendante de S_f telle que : $\|u\|_{H^1(\Omega)} \leq C\,\|S_f\|$.

2.3 Existence et unicité, cadre général

Nous proposons ci-dessous une approche plus générale que celle du théorème de Lax-Milgram pour résoudre les formulations variationnelles. Entre autres, les formulations variationnelles peuvent mettre en jeu un espace des solutions différent de l'espace des fonctions-test. On se concentre sur le cas $\mathbb{K} = \mathbb{C}$, le cas $\mathbb{K} = \mathbb{R}$ est complètement similaire. Précisément, soient :

— V et W deux espaces de Hilbert définis sur $\mathbb{C}$;

— $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V \times W$.

Pour $\ell \in W'$ donné, on étudie la formulation variationnelle,

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ \forall w \in W, \quad a(u,w) = \langle \ell, w\rangle_W. \end{cases} \tag{2.10}$$

2.3.1 Caractère bien posé

Nous commençons par une définition générale.

Définition 2.6 (Hadamard) *La formulation variationnelle (2.10) est* bien posée *si, et seulement si, pour tout $\ell \in W'$, (2.10) admet une solution et une seule $u \in V$, avec dépendance continue :*

$$\exists C > 0,\ \forall \ell \in W',\ \|u\|_V \leq C\|\ell\|_{W'}.$$

Remarque 2.7 *Par stabilité, on cherche à déterminer dans quelle mesure une "petite perturbation" de ℓ induit une "petite perturbation" de u. Par abus de langage, on dira que plus C est petite, plus la formulation variationnelle est stable.*

On introduit l'application linéaire et continue $\mathtt{A} \in \mathcal{L}(V, W)$ associée à $a(\cdot,\cdot)$, et définie par :

$$\forall (v, w) \in V \times W, \quad (\mathtt{A}v, w)_W = a(v, w). \tag{2.11}$$

Pour tout $v \in V \setminus \{0\}$, on rappelle le résultat élémentaire :

$$\|\mathtt{A}v\|_W = \sup_{w \in W\setminus\{0\}} \frac{|(\mathtt{A}v, w)_W|}{\|w\|_W} = \frac{|a(v, w)|}{\|w\|_W}.$$

On en déduit que :

$$|||\mathtt{A}||| = \sup_{v\in V\setminus\{0\}} \frac{\|\mathtt{A}v\|_W}{\|v\|_V} = \sup_{v\in V\setminus\{0\}, w\in W\setminus\{0\}} \frac{|a(v, w)|}{\|v\|_V\, \|w\|_W} = |||a|||. \tag{2.12}$$

Proposition 2.8 *Soient V, W deux espaces de Hilbert, et $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V \times W$. Alors la formulation variationnelle (2.10) est bien posée si, et seulement si, l'inverse de $\mathtt{A}$ défini par (2.11) existe et est continu : $\mathtt{A}^{-1} \in \mathcal{L}(W, V)$.*

Remarque 2.9 *Ainsi, (2.10) est bien posée si, et seulement si, $\mathtt{A}$ est un isomorphisme de V dans W.*

Démonstration : Supposons que (2.10) est bien posée. Soit $\mathtt{f} \in W$, alors il existe une solution et une seule à la formulation variationnelle (2.10) avec pour donnée la forme $w \mapsto (\mathtt{f}, w)_W$, antilinéaire et continue sur W, qu'on note $u_\mathtt{f}$. En outre, $\|u_\mathtt{f}\|_V \leq C\,\|\mathtt{f}\|_W$ avec $C > 0$ indépendante de $\mathtt{f}$. On introduit $\mathtt{B} : \mathtt{f} \mapsto u_\mathtt{f}$, dont on vérifie facilement que c'est une application linéaire (par linéarité de a par rapport au premier argument) et continue de W dans V : $\mathtt{B} \in \mathcal{L}(W, V)$.
Par ailleurs, pour $\mathtt{f} \in W$:

$$\forall w \in W, \quad (\mathtt{ABf}, w)_W = a(\mathtt{Bf}, w) = (\mathtt{f}, w)_W,$$

et ainsi $\mathtt{ABf} = \mathtt{f}$. on a donc $\mathtt{AB} = \mathtt{I}_W$.
De même, pour $u \in V$:

$$\forall w \in W, \quad a(\mathtt{BA}u, w) = a(\mathtt{B}(\mathtt{A}u), w) = (\mathtt{A}u, w)_W = a(u, w)\,;$$

il suit que $\mathtt{B}\mathtt{A}u = u$, par unicité de la solution de la formulation variationnelle (2.10). On en conclut que $\mathtt{B}\mathtt{A} = \mathtt{I}_V$. Ainsi, $\mathtt{B}$ est l'inverse (continu) de $\mathtt{A}$.

Réciproquement, si l'inverse de $\mathtt{A}$ existe et est continu, soit $\ell \in W'$: d'après le théorème de Riesz, il existe $\mathtt{f} \in W$ unique tel que $(\mathtt{f}, w)_W = \ell(w)$ pour tout $w \in W$, avec $\|\mathtt{f}\|_W = \|\ell\|_{W'}$. Notons $u_\mathtt{f} = \mathtt{A}^{-1}\mathtt{f}$, par construction $\mathtt{A}u_\mathtt{f} = \mathtt{A}(\mathtt{A}^{-1}\mathtt{f}) = \mathtt{f}$ dans W, soit

$$\forall w \in W, \quad a(u_\mathtt{f}, w) = (\mathtt{f}, w)_W = \ell(w).$$

En d'autres termes, $u_\mathtt{f}$ résout la formulation variationnelle (2.10) avec la donnée ℓ, et

$$\|u_\mathtt{f}\|_V = \|\mathtt{A}^{-1}\mathtt{f}\|_V \leq \left|\left|\left|\mathtt{A}^{-1}\right|\right|\right| \|\mathtt{f}\|_W = \left|\left|\left|\mathtt{A}^{-1}\right|\right|\right| \|\ell\|_{W'}.$$

On a donc dépendance continue de la solution $u_\mathtt{f}$ par rapport à la donnée ℓ. Vérifions enfin l'unicité de la solution. Pour ℓ donnée, si u_1, u_2 sont deux solutions, alors $\mathtt{A}(u_1 - u_2) = 0$ et donc $u_1 - u_2 = \mathtt{A}^{-1}\mathtt{A}(u_1 - u_2) = 0$, ce qui conclut la démonstration. ∎

2.3.2 Théorie de Banach-Necas-Babuska

Nous allons énoncer des conditions nécessaires et suffisantes pour que la formulation variationnelle (2.10) soit bien posée. Précisément, il s'agit de la *condition de stabilité*, également appelée la *condition inf-sup*, et de la *condition de solvabilité*.

Définition 2.10 *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire sur $V \times W$. Alors $a(\cdot,\cdot)$ vérifie une* condition inf-sup *si, et seulement si,*

$$\exists \alpha' > 0, \ \forall v \in V, \ \sup_{w \in W\setminus\{0\}} \frac{|a(v,w)|}{\|w\|_W} \geq \alpha' \|v\|_V. \tag{2.13}$$

Dans ce cas, on appelle constante inf-sup *la plus grande valeur de α' dans (2.13). Elle est égale à*

$$\inf_{v \in V\setminus\{0\}} \sup_{w \in W\setminus\{0\}} \frac{|a(v,w)|}{\|v\|_V \, \|w\|_W}.$$

Définition 2.11 *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire sur $V \times W$. Alors $a(\cdot,\cdot)$ vérifie une* condition de solvabilité *si, et seulement si,*

$$\{w \in W \text{ tel que } a(v,w) = 0, \ \forall v \in V\} = \{0\}. \tag{2.14}$$

Intuitivement, la condition de solvabilité exprime le fait qu'on n'a pas "trop" de fonctions-test. Nous allons maintenant énoncer et démontrer le résultat général, le théorème de Banach-Necas-Babuska, et son corollaire (on utilisera en général la notation BNB dans la suite).

Théorème 2.12 (Banach-Necas-Babuska) *Soient V et W deux espaces de Hilbert, $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V \times W$. La formulation variationnelle (2.10) est bien posée si, et seulement si, les conditions (2.13) et (2.14) sont satisfaites.*

Démonstration : Supposons la formulation variationnelle (2.10) bien posée. D'après la proposition 2.8, l'application linéaire $\mathtt{A}$ associée à a par (2.11) admet un inverse continu de W dans V.
Soit $v \in V \setminus \{0\}$, on a :

— d'une part, $a(v, \mathtt{A}v) = (\mathtt{A}v, \mathtt{A}v)_W = \|\mathtt{A}v\|_W^2$;

— d'autre part, $\|v\|_V = \|\mathtt{A}^{-1}(\mathtt{A}v)\|_V \leq |||\mathtt{A}^{-1}||| \, \|\mathtt{A}v\|_W$ (en particulier, $\mathtt{A}v \neq 0$).

Ainsi,

$$\sup_{w \in W \setminus \{0\}} \frac{|a(v,w)|}{\|w\|_W} \geq \frac{|a(v,\mathtt{A}v)|}{\|\mathtt{A}v\|_W} = \|\mathtt{A}v\|_W \geq \frac{1}{|||\mathtt{A}^{-1}|||}\|v\|_V.$$

La condition inf-sup (2.13) est vérifiée avec $\alpha' = 1/|||\mathtt{A}^{-1}|||$.
Soit $w \in W$ tel que $a(v,w) = 0$ pour tout $v \in V$. En particulier,

$$0 = a(\mathtt{A}^{-1}w, w) = (\mathtt{A}(\mathtt{A}^{-1}w), w)_W = \|w\|_W^2, \text{ soit } w = 0.$$

La condition de solvabilité (2.14) est vérifiée elle aussi.

Réciproquement, supposons que les conditions (2.13) et (2.14) sont satisfaites. Nous allons prouver que l'application linéaire $\mathtt{A}$ associée à la forme a est bijective, d'inverse continu. La proposition 2.8 permettra alors de conclure.

Soit $u \in V$ tel que $\mathtt{A}u = 0$: alors, $a(u,w) = (\mathtt{A}u, w)_W = 0$ pour tout $w \in W$ et, d'après (2.13), il suit $\|u\|_V = 0$, c'est-à-dire $u = 0$. $\mathtt{A}$ est donc injective.

Ensuite, nous allons prouver que l'image de $\mathtt{A}$, $\mathrm{Im}(\mathtt{A}) = \{w \in W \text{ tel que } \exists v \in V, \ w = \mathtt{A}v\}$, est un sous-espace vectoriel fermé de W, d'orthogonal réduit à $\{0\}$. Pour commencer, $\mathrm{Im}(\mathtt{A})$ est un sous-espace vectoriel de W. Pour le caractère fermé, on va établir que toute suite d'éléments $(w^k)_k$ de $\mathrm{Im}(\mathtt{A})$ qui converge dans W a sa limite dans $\mathrm{Im}(\mathtt{A})$. Par définition, pour tout k il existe $v^k \in V$ tel que $w^k = \mathtt{A}v^k$. A l'aide de l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned} \forall k, m, \ \forall w \in W, \quad |a(v^k - v^m, w)| &= |(\mathtt{A}v^k - \mathtt{A}v^m, w)_W| = |(w^k - w^m, w)_W| \\ &\leq \|w^k - w^m\|_W \|w\|_W. \end{aligned}$$

En utilisant (2.13), on en déduit que

$$\forall k, m, \quad \|v^k - v^m\|_V \leq \frac{1}{\alpha'} \sup_{w \in W \setminus \{0\}} \frac{|a(v^k - v^m, w)|}{\|w\|_W} \leq \frac{1}{\alpha'}\|w^k - w^m\|_W.$$

La suite $(w^k)_k$ étant convergente dans W, elle est a fortiori de Cauchy : $\lim_{k,m\to\infty} \|w^k - w^m\|_W = 0$. Il en est donc de même pour $(v^k)_k$ dans V. Et, comme V est un espace complet, $(v^k)_k$ converge dans V, c'est-à-dire qu'il existe $v \in V$ tel que $\lim_{k\to\infty} \|v^k - v\|_V = 0$. On en déduit :

$$\|w^k - \mathtt{A}v\|_W = \|\mathtt{A}(v^k - v)\|_W \leq |||\mathtt{A}||| \, \|v^k - v\|_V \to 0 \text{ quand } k \to \infty,$$

et ainsi $(w^k)_k$ converge vers $\mathtt{A}v \in \mathrm{Im}(\mathtt{A})$; $\mathrm{Im}(\mathtt{A})$ est donc fermée dans W.
Soit maintenant $w \in (\mathrm{Im}(\mathtt{A}))^\perp$, l'orthogonal de $\mathrm{Im}(\mathtt{A})$ dans W : par définition, ça signifie que, pour tout $v \in V$, $0 = (\mathtt{A}v, w)_W = a(v,w)$. D'après (2.14), on a $w = 0$. On en conclut [3] que $\mathrm{Im}(\mathtt{A}) = W$: $\mathtt{A}$ est surjective.

3. Comme W est un espace de Hilbert, pour tout sous-espace vectoriel fermé Y de W, on a $W = Y \overset{\perp}{\oplus} Y^\perp$.

Comme $\mathtt{A}$ est bijective, pour tout $\mathtt{f} \in W$, on sait qu'il existe $u_{\mathtt{f}} = \mathtt{A}^{-1}\mathtt{f} \in V$ unique tel que $\mathtt{A}u_{\mathtt{f}} = \mathtt{f}$. Et, d'après (2.13) et l'inégalité de Cauchy-Schwarz, on trouve

$$\|\mathtt{A}^{-1}\mathtt{f}\|_V = \|u_{\mathtt{f}}\|_V \leq \frac{1}{\alpha'} \sup_{w\in W\setminus\{0\}} \frac{|a(u_{\mathtt{f}}, w)|}{\|w\|_W} \leq \frac{1}{\alpha'}\|\mathtt{A}u_{\mathtt{f}}\|_W = \frac{1}{\alpha'}\|\mathtt{f}\|_W,$$

c'est-à-dire que $|||\mathtt{A}^{-1}||| \leq 1/\alpha'$. On en conclut comme annoncé que la formulation variationnelle (2.10) est bien posée au sens de Hadamard. ■

Remarque 2.13 *Pour une forme sesquilinéaire définie sur $V \times V$, on a donc le choix entre coercivité d'une part, et stabilité plus solvabilité d'autre part. On vérifie facilement que, si la forme est coercive (avec une constante $\alpha > 0$), alors elle satisfait à la fois une condition de stabilité (avec $\alpha' = \alpha$) et la condition de solvabilité. La réciproque est* fausse, *voir le §2.3.3.*

Notons que, lorsque la théorie de Banach-Necas-Babuska s'applique, on a une *estimation de stabilité* (voir la remarque 2.7), puisqu'on peut calculer une constante C apparaissant de la définition 2.6. On peut préciser, voir le

Corollaire 2.14 `(Stabilité)` *Lorsque la formulation variationnelle (2.10) est bien posée, on a, pour tout $\alpha' > 0$ réalisant la condition inf-sup (2.13), l'estimation de stabilité :*

$$\forall \ell \in W',\ \|u\|_V \leq \frac{1}{\alpha'}\,\|\ell\|_{W'}.$$

En outre, la plus grande valeur de α' (la constante inf-sup), est égale à $1/|||\mathtt{A}^{-1}|||$.

Remarque 2.15 *En général, on ne dispose que d'une* estimation *de la meilleure constante C dans la définition 2.6. En effet on ne connaît pas la valeur exacte de la constante inf-sup (pour cela il faut connaître $|||\mathtt{A}^{-1}|||$), mais plutôt des nombres α' réalisant (2.13).*

Démonstration : Tout d'abord, à l'aide de (2.13), on obtient la borne :

$$\|u\|_V \leq \frac{1}{\alpha'} \sup_{w\in W\setminus\{0\}} \frac{|a(u,w)|}{\|w\|_W} \overset{(2.10)}{=} \frac{1}{\alpha'} \sup_{w\in W\setminus\{0\}} \frac{|\langle \ell, w\rangle_W|}{\|w\|_W} = \frac{1}{\alpha'}\,\|\ell\|_{W'}.$$

Ensuite, puisque la formulation variationnelle (2.10) est bien posée, on sait que $\mathtt{A}$ est bijective. On déduit que

$$\begin{aligned}
\inf_{v\in V\setminus\{0\}} \sup_{w\in W\setminus\{0\}} \frac{|a(v,w)|}{\|v\|_V\,\|w\|_W} &= \inf_{v\in V\setminus\{0\}} \frac{\|\mathtt{A}v\|_W}{\|v\|_V} \\
&= \inf_{w\in W\setminus\{0\}} \frac{\|w\|_W}{\|\mathtt{A}^{-1}w\|_V} \\
&= \left(\sup_{w\in W\setminus\{0\}} \frac{\|\mathtt{A}^{-1}w\|_V}{\|w\|_W}\right)^{-1} = \frac{1}{|||\mathtt{A}^{-1}|||},
\end{aligned}$$

ce qui est le résultat demandé. ■

Si la forme a est de plus hermitienne, on a le

Corollaire 2.16 (Banach-Necas-Babuska) *Soit V un espace de Hilbert, $a(\cdot,\cdot)$ une forme sesquilinéaire, continue et hermitienne sur $V \times V$. La formulation variationnelle (2.10) avec $W = V$ est bien posée si, et seulement si, la condition inf-sup (2.13) est satisfaite.*

Démonstration : D'après le théorème BNB, lorsque a est hermitienne, il suffit de prouver que la condition inf-sup (2.13) implique la condition de solvabilité (2.14). Soit $w \in V$ tel que $a(v,w) = 0$ pour tout $v \in V$: on en déduit que $a(w,v) = 0$ pour tout $v \in V$. D'après (2.13), on a $w = 0$: la condition de solvabilité (2.14) est satisfaite, et le théorème précédent nous permet de conclure. ∎

2.3.3 Diffusion à deux inconnues

A quoi sert le théorème BNB pour le modèle de la diffusion ? En préambule, on note que, si on considère des espaces de Hilbert sur $\mathbb{R}$ avec des formes bilinéaires et continues, voire symétriques, et la valeur absolue $|\cdot|$, on a exactement les mêmes définitions et résultats qu'aux §2.3.1-§2.3.2...

On choisit d'utiliser la variable auxiliaire, c'est-à-dire le courant neutronique $\boldsymbol{p}$. Grâce aux les hypothèses sur les coefficients $\mathbb{D}$ et σ et sur la donnée S_f, on note que le modèle (2.1) est équivalent à :

$$\left\{\begin{array}{l} \text{Trouver } u \in H^1_0(\Omega),\ \boldsymbol{p} \in \boldsymbol{H}(\mathrm{div}\,;\Omega) \text{ tels que} \\ \mathrm{div}\,\boldsymbol{p} + \sigma u = S_f \text{ dans } \Omega \\ \mathbb{D}^{-1}\boldsymbol{p} + \nabla u = 0 \text{ dans } \Omega. \end{array}\right. \tag{2.15}$$

Dans la suite, nous allons mettre ce modèle sous forme variationnelle, avec les *deux inconnues* u et $\boldsymbol{p}$. D'après l'hypothèse (2.2) sur $\mathbb{D}$, on en déduit facilement que

$$\left|\begin{array}{l} \exists D_{min}^{-1} > 0,\ \forall \boldsymbol{z} \in \mathbb{R}^{\mathtt{d}},\ \text{presque pour tout } \boldsymbol{x} \in \Omega,\ D_{min}^{-1}|\boldsymbol{z}|^2 \le \mathbb{D}^{-1}(\boldsymbol{x})\boldsymbol{z}\cdot\boldsymbol{z}\,; \\ \exists D_{max}^{-1} > 0,\ \forall \boldsymbol{z} \in \mathbb{R}^{\mathtt{d}},\ \text{presque pour tout } \boldsymbol{x} \in \Omega,\ |\mathbb{D}^{-1}(\boldsymbol{x})\boldsymbol{z}| \le D_{max}^{-1}|\boldsymbol{z}|. \end{array}\right. \tag{2.16}$$

Démonstration : Soit $\boldsymbol{x} \in \Omega$ tel que $\mathbb{D}(\boldsymbol{x})$ existe. D'après (2.2a), $\mathbb{D}(\boldsymbol{x})$ est inversible. Soit $\boldsymbol{z} \in \mathbb{R}^{\mathtt{d}}$ non-nul, on pose $\boldsymbol{y} = \mathbb{D}^{-1}(\boldsymbol{x})\boldsymbol{z}$.
Tout d'abord, d'après (2.2a) pour $\boldsymbol{y}$, on a $D_{min}|\boldsymbol{y}|^2 \le \mathbb{D}(\boldsymbol{x})\boldsymbol{y}\cdot\boldsymbol{y} = \boldsymbol{z}\cdot\boldsymbol{y} \le |\boldsymbol{z}|\,|\boldsymbol{y}|$. Après division par $|\boldsymbol{y}| \neq 0$, on en déduit que $|\mathbb{D}^{-1}(\boldsymbol{x})\boldsymbol{z}| = |\boldsymbol{y}| \le (D_{min})^{-1}|\boldsymbol{z}|$, c'est-à-dire (2.16b) avec $D_{max}^{-1} = (D_{min})^{-1}$. Puis, d'après (2.2b) pour $\boldsymbol{y}$, on a $|\mathbb{D}(\boldsymbol{x})\boldsymbol{y}| \le D_{max}|\boldsymbol{y}|$, c'est-à-dire que $|\boldsymbol{z}| \le D_{max}|\mathbb{D}^{-1}(\boldsymbol{x})\boldsymbol{z}|$. D'où, si on réutilise (2.2a) pour $\boldsymbol{y}$, et si on se souvient que le produit scalaire est symétrique :

$$\frac{D_{min}}{(D_{max})^2}|\boldsymbol{z}|^2 \le D_{min}|\mathbb{D}^{-1}(\boldsymbol{x})\boldsymbol{z}|^2 = D_{min}|\boldsymbol{y}|^2 \le \mathbb{D}(\boldsymbol{x})\boldsymbol{y}\cdot\boldsymbol{y} = \boldsymbol{y}\cdot\mathbb{D}(\boldsymbol{x})\boldsymbol{y} = \mathbb{D}^{-1}(\boldsymbol{x})\boldsymbol{z}\cdot\boldsymbol{z},$$

et on a obtenu (2.16a) avec $D_{min}^{-1} = D_{min}/(D_{max})^2$. ∎

Pour simplifier l'analyse, nous supposons en outre que le coefficient $\mathbb{D}$ est un champ de *matrices symétriques*. Pour construire une formulation variationnelle équivalente au modèle

(2.15), on commence par réaliser donc le "produit scalaire dans $\boldsymbol{L}^2(\Omega)$" de la seconde EDP avec $\boldsymbol{r} \in \boldsymbol{H}(\mathrm{div}\,;\Omega)$, pour trouver [4] :

$$0 = -\int_\Omega (\mathbb{D}^{-1}\boldsymbol{p} + \nabla u)\cdot \boldsymbol{r}\, d\Omega \overset{(1.20)}{=} \int_\Omega \Big(-\mathbb{D}^{-1}\boldsymbol{p}\cdot\boldsymbol{r} + u(\mathrm{div}\,\boldsymbol{r})\Big) d\Omega.$$

On note qu'après intégration par parties, l'inconnue u n'apparaît plus que comme appartenant à $L^2(\Omega)$, et de même dans la première EDP. En outre, la condition aux limites sur u est explicitement utilisée dans l'intégration par parties... Pour poursuivre, on réalise donc le "produit scalaire dans $L^2(\Omega)$" de la première EDP avec $w \in L^2(\Omega)$, pour former

$$\int_\Omega S_f w\, d\Omega = \int_\Omega (\mathrm{div}\,\boldsymbol{p} + \sigma u) w\, d\Omega = \int_\Omega \Big(w(\mathrm{div}\,\boldsymbol{p}) + \sigma u w\Big)\, d\Omega.$$

On en conclut que si $(u,\boldsymbol{p})$ est solution du modèle (2.15), alors $(u,\boldsymbol{p})$ est solution de la formulation variationnelle

$$\left\{\begin{array}{l} \text{Trouver } (u,\boldsymbol{p}) \in L^2(\Omega)\times \boldsymbol{H}(\mathrm{div}\,;\Omega) \text{ tel que} \\ \forall (w,\boldsymbol{r}) \in L^2(\Omega)\times \boldsymbol{H}(\mathrm{div}\,;\Omega), \\ \qquad \displaystyle\int_\Omega \Big(-\mathbb{D}^{-1}\boldsymbol{p}\cdot\boldsymbol{r} + u(\mathrm{div}\,\boldsymbol{r}) + w(\mathrm{div}\,\boldsymbol{p}) + \sigma u w\Big) d\Omega = \int_\Omega S_f w\, d\Omega. \end{array}\right. \tag{2.17}$$

Réciproquement, en raisonnant au sens des distributions et en utilisant la surjectivité de l'application trace normale γ_n de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ dans $(H^{1/2}(\partial\Omega))'$ (voir le théorème 1.32), on montre que si $(u,\boldsymbol{p})$ est solution de la formulation variationnelle (2.17), alors $(u,\boldsymbol{p})$ est solution du modèle (2.15) (on retrouve au passage que $u \in H^1(\Omega)$, puis que $u_{|\partial\Omega} = 0$). On a donc bien construit *une seconde formulation variationnelle équivalente à notre modèle.*

Par rapport à la formulation variationnelle "abstraite" (2.5), on a respectivement :

— $V = L^2(\Omega)\times \boldsymbol{H}(\mathrm{div}\,;\Omega)$, muni de la norme $\|(v,\boldsymbol{q})\|_V = (\|v\|^2 + \|\boldsymbol{q}\|^2_{\boldsymbol{H}(\mathrm{div}\,;\Omega)})^{1/2}$;

— $a((v,\boldsymbol{q}),(w,\boldsymbol{r})) = -(\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{r}) + (v|\mathrm{div}\,\boldsymbol{r}) + (w|\mathrm{div}\,\boldsymbol{q}) + (\sigma v|w)$;

— $\langle \ell, (w,\boldsymbol{r})\rangle_V = (S_f|w)$.

On vérifie facilement que la forme bilinéaire a est continue sur $(L^2(\Omega)\times \boldsymbol{H}(\mathrm{div}\,;\Omega))\times(L^2(\Omega)\times \boldsymbol{H}(\mathrm{div}\,;\Omega))$, et que la forme linéaire ℓ est continue sur $L^2(\Omega)\times \boldsymbol{H}(\mathrm{div}\,;\Omega)$. Nous prouvons son caractère bien posé ci-dessous à l'aide du théorème BNB. En effet, le théorème de Lax-Milgram ne s'applique pas, puisque la forme a n'est pas coercive ! Si on examine $|a((0,\boldsymbol{q}),(0,\boldsymbol{q}))| = |(\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{q})|$, on note que $D_{min}^{-1}\|\boldsymbol{q}\|^2 \le |a((0,\boldsymbol{q}),(0,\boldsymbol{q}))| \le D_{max}^{-1}\|\boldsymbol{q}\|^2$. Il suit que $|a((0,\boldsymbol{q}),(0,\boldsymbol{q}))| = |(\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{q})|$ ne peut pas majorer $\|\boldsymbol{q}\|^2_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} = \|\boldsymbol{q}\|^2 + \|\mathrm{div}\,\boldsymbol{q}\|^2$ uniformément (c'est-à-dire avec une constante de coercivité indépendante de $\boldsymbol{q}$), car on peut construire une suite $(\boldsymbol{q}_k)_{k\ge1}$ d'éléments de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ telle que $\|\boldsymbol{q}_k\|$ reste bornée, mais pas $\|\mathrm{div}\,\boldsymbol{q}_k\|$. Par exemple, on peut choisir $\boldsymbol{q}_k = \sin(kx_1)\mathbf{e}_1$, pour $k \ge 1$: pour cette suite, il n'existe pas de constante $c > 0$ telle que, pour tout $k \ge 1$, on ait $|a((0,\boldsymbol{q}_k),(0,\boldsymbol{q}_k))| \ge c\,\|\mathrm{div}\,\boldsymbol{q}_k\|^2$.

Par contre, la forme a étant symétrique, pour obtenir le caractère bien posé, on peut passer par

4. Au lieu de la formule d'intégration par parties (1.20), on peut encore une fois utiliser la remarque 1.33.

le corollaire BNB, ce qui simplifie quelque peu l'analyse. Il suffit d'établir la condition inf-sup (2.13). Nous allons la construire explicitement : pour cela, pour $(v, \boldsymbol{q}) \in L^2(\Omega) \times \boldsymbol{H}(\mathrm{div}\,;\Omega)$ non-nul, nous allons chercher $(w^\star, \boldsymbol{r}^\star) \in L^2(\Omega) \times \boldsymbol{H}(\mathrm{div}\,;\Omega)$ non-nul, dépendant linéairement de $(v, \boldsymbol{q})$, et tel que

$$|a((v,\boldsymbol{q}),(w^\star,\boldsymbol{r}^\star))| \geq \alpha' \|(v,\boldsymbol{q})\|_V \, \|(w^\star,\boldsymbol{r}^\star)\|_V,$$

avec une constante $\alpha' > 0$ indépendante de $(v, \boldsymbol{q})$. L'idée de cette approche est que l'on pourra, lors de la discrétisation, reproduire ce calcul afin d'obtenir une estimation "similaire". Voir les notes de cours 3. On raisonne par étapes.

1. **Cas particulier $\boldsymbol{q} = 0$.** On a

$$a((v,0),(w,\boldsymbol{r})) = (v|\mathrm{div}\,\boldsymbol{r}) + (\sigma v|w) \text{ et } \|(v,0)\|_V = \|v\|.$$

   Quel $(w^\star, \boldsymbol{r}^\star)$ choisir ? Avec $w^\star = v$ et $\boldsymbol{r}^\star = 0$, on a $\|(w^\star, \boldsymbol{r}^\star)\|_V = \|v\|$, et ce choix donne

$$a((v,0),(w^\star,\boldsymbol{r}^\star)) = (\sigma v|v) \geq \sigma_{min} \|v\|^2 = \sigma_{min} \|(v,0)\|_V \, \|(w^\star,\boldsymbol{r}^\star)\|_V.$$

2. **Cas particulier** $\mathrm{div}\,\boldsymbol{q} = 0$**.** On a

$$a((v,\boldsymbol{q}),(w,\boldsymbol{r})) = (-\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{r}) + (v|\mathrm{div}\,\boldsymbol{r}) + (\sigma v|w) \text{ et } \|(v,\boldsymbol{q})\|_V = (\|v\|^2 + \|\boldsymbol{q}\|^2)^{1/2}.$$

   Quel $(w^\star, \boldsymbol{r}^\star)$ choisir ?
   — $\boldsymbol{r}^\star = -\boldsymbol{q}$ donne : $a((v,\boldsymbol{q}),(w,\boldsymbol{r}^\star)) = (\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{q}) + (\sigma v|w)$, puisque $\mathrm{div}\,\boldsymbol{q} = 0$;
   — $w^\star = v$ donne : $a((v,\boldsymbol{q}),(w^\star,\boldsymbol{r}^\star)) = (\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{q}) + (\sigma v|v)$.
   Avec ce choix, $\|(w^\star, \boldsymbol{r}^\star)\|_V = (\|v\|^2 + \|\boldsymbol{q}\|^2)^{1/2}$, puisque $\mathrm{div}\,\boldsymbol{q} = 0$. On a donc

$$\begin{aligned} a((v,\boldsymbol{q}),(w^\star,\boldsymbol{r}^\star)) &\geq D_{min}^{-1} \|\boldsymbol{q}\|^2 + \sigma_{min} \|v\|^2 \\ &\geq \min(D_{min}^{-1}, \sigma_{min}) \, (\|\boldsymbol{q}\|^2 + \|v\|^2) \\ &= \min(D_{min}^{-1}, \sigma_{min}) \, \|(v,\boldsymbol{q})\|_V \, \|(w^\star,\boldsymbol{r}^\star)\|_V. \end{aligned}$$

3. **Cas général** $\mathrm{div}\,\boldsymbol{q} \neq 0$**.** Quel $(w^\star, \boldsymbol{r}^\star)$ choisir ?
   — On commence comme pour "la divergence nulle" avec $\boldsymbol{r}^\star = -\boldsymbol{q}$, et on a

$$a((v,\boldsymbol{q}),(w,\boldsymbol{r}^\star)) = (\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{q}) + (w - v|\mathrm{div}\,\boldsymbol{q}) + (\sigma v|w).$$

   — Ensuite, on considère $w^\star = \alpha(v + \sigma^{-1}\mathrm{div}\,\boldsymbol{q})$, pour $\alpha > 0$ à déterminer. Le choix αv reprend "la divergence nulle". En plus, on a introduit $\sigma^{-1}\mathrm{div}\,\boldsymbol{q}$ pour le "contrôle" de $\mathrm{div}\,\boldsymbol{q}$; avec le coefficient σ^{-1} devant $\mathrm{div}\,\boldsymbol{q}$ par *homogénéité*, cf. la première EDP dans le modèle (2.15). On trouve

$$a((v,\boldsymbol{q}),(w^\star,\boldsymbol{r}^\star)) = (\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{q}) + (2\alpha - 1)(v|\mathrm{div}\,\boldsymbol{q}) + \alpha(\sigma^{-1}\mathrm{div}\,\boldsymbol{q}|\mathrm{div}\,\boldsymbol{q}) + \alpha(\sigma v|v).$$

   Si on fixe $\alpha = \frac{1}{2}$, le terme croisé en $(v|\mathrm{div}\,\boldsymbol{q})$ disparaît ! Il reste

$$a((v,\boldsymbol{q}),(w^\star,\boldsymbol{r}^\star)) = (\mathbb{D}^{-1}\boldsymbol{q}|\boldsymbol{q}) + \frac{1}{2}(\sigma^{-1}\mathrm{div}\,\boldsymbol{q}|\mathrm{div}\,\boldsymbol{q}) + \frac{1}{2}(\sigma v|v),$$

   pour $(w^\star, \boldsymbol{r}^\star) = (\frac{1}{2}(v + \sigma^{-1}\mathrm{div}\,\boldsymbol{q}), -\boldsymbol{q})$. On a bien dépendance linéaire.

Avec ce choix, on a maintenant

$$\begin{aligned} a((v,\boldsymbol{q}),(w^\star,\boldsymbol{r}^\star)) &\geq D_{min}^{-1}\,\|\boldsymbol{q}\|^2 + \frac{1}{2}(\sigma_{max})^{-1}\|\mathrm{div}\,\boldsymbol{q}\|^2 + \frac{1}{2}\sigma_{min}\,\|v\|^2 \\ &\geq \min\left(D_{min}^{-1}, \frac{1}{2}(\sigma_{max})^{-1}, \frac{1}{2}\sigma_{min}\right)(\|\boldsymbol{q}\|^2 + \|\mathrm{div}\,\boldsymbol{q}\|^2 + \|v\|^2) \\ &= \min\left(D_{min}^{-1}, \frac{1}{2}(\sigma_{max})^{-1}, \frac{1}{2}\sigma_{min}\right)\|(v,\boldsymbol{q})\|_V^2. \end{aligned} \tag{2.18}$$

Pour conclure, il suffit de majorer [5] $\|(w^\star,\boldsymbol{r}^\star)\|_V$ par $\|(v,\boldsymbol{q})\|_V$. On a

$$\begin{aligned} \|(w^\star,\boldsymbol{r}^\star)\|_V^2 &= \frac{1}{4}\|(v+\sigma^{-1}\mathrm{div}\,\boldsymbol{q})\|^2 + \|\boldsymbol{q}\|^2 + \|\mathrm{div}\,\boldsymbol{q}\|^2 \\ &\leq \|v\|^2 + \frac{1}{3}\|\sigma^{-1}\mathrm{div}\,\boldsymbol{q}\|^2 + \|\boldsymbol{q}\|^2 + \|\mathrm{div}\,\boldsymbol{q}\|^2 \\ &\leq \max\left(1, \frac{1}{3}(\sigma_{min})^{-2}+1\right)(\|\boldsymbol{q}\|^2 + \|\mathrm{div}\,\boldsymbol{q}\|^2 + \|v\|^2) \\ &= \left(\frac{1}{3}(\sigma_{min})^{-2}+1\right)(\|\boldsymbol{q}\|^2 + \|\mathrm{div}\,\boldsymbol{q}\|^2 + \|v\|^2) \\ \Longrightarrow \|(w^\star,\boldsymbol{r}^\star)\|_V &\leq \left(\frac{1}{3}(\sigma_{min})^{-2}+1\right)^{1/2}\|(v,\boldsymbol{q})\|_V. \end{aligned} \tag{2.19}$$

Finalement, on a établi que, pour tout $(v,\boldsymbol{q})$ non-nul, si on choisit $(w^\star,\boldsymbol{r}^\star) = (\frac{1}{2}(v+\sigma^{-1}\mathrm{div}\,\boldsymbol{q}), -\boldsymbol{q})$, on a

$$a((v,\boldsymbol{q}),(w^\star,\boldsymbol{r}^\star)) \geq \frac{\min\left(D_{min}^{-1}, \frac{1}{2}(\sigma_{max})^{-1}, \frac{1}{2}\sigma_{min}\right)}{\left(\frac{1}{3}(\sigma_{min})^{-2}+1\right)^{1/2}}\|(v,\boldsymbol{q})\|_V\,\|(w^\star,\boldsymbol{r}^\star)\|_V.$$

Et ainsi la condition inf-sup (2.13) est valable avec

$$\alpha' = \frac{\min\left(D_{min}^{-1}, \frac{1}{2}(\sigma_{max})^{-1}, \frac{1}{2}\sigma_{min}\right)}{\left(\frac{1}{3}(\sigma_{min})^{-2}+1\right)^{1/2}}. \tag{2.20}$$

Qui plus est, pour tout $(v,\boldsymbol{q})$ on dispose d'un représentant "explicite" $(w^\star,\boldsymbol{r}^\star))$ réalisant (2.13). Comme annoncé plus haut, on peut appliquer le corollaire BNB et en conclure que le modèle de la diffusion est bien posé, en étant cette fois passé par la formulation variationnelle (2.17) à deux inconnues.

2.3.4 Introduction à la T-coercivité

Reprenons l'étude de (2.17) avec les résultats obtenus précédemment.
Explicitons la note de bas de page 5 page 31, et introduisons $\mathtt{T} \in \mathcal{L}(V)$ définie par :

$$\mathtt{T}((v,\boldsymbol{q})) = (w^\star,\boldsymbol{r}^\star) = (\frac{1}{2}(v+\sigma^{-1}\mathrm{div}\,\boldsymbol{q}), -\boldsymbol{q})\,;$$

5. Si on n'est pas intéressé par une majoration *explicite*, il suffit de noter que $(v,\boldsymbol{q}) \mapsto (w^\star,\boldsymbol{r}^\star)$ est linéaire, et de plus appartient à $\mathcal{L}(V)$, puisque : $\|w^\star\| \leq \frac{1}{2}\|v\| + \frac{1}{2}\sigma_{min}^{-1}\|\mathrm{div}\,\boldsymbol{q}\|$, et $\|\boldsymbol{r}^\star\|_{\boldsymbol{H}(\mathrm{div};\Omega)} = \|\boldsymbol{q}\|_{\boldsymbol{H}(\mathrm{div};\Omega)}$.

on a $|||\mathtt{T}||| \leq \left(\frac{1}{3}(\sigma_{min})^{-2}+1\right)^{1/2}$ d'après (2.19).
La forme $a'((v,\boldsymbol{q}),(w,\boldsymbol{r})) = a((v,\boldsymbol{q}),\mathtt{T}((w,\boldsymbol{r})))$, définie sur $V\times V$, est bilinéaire et continue. En outre, si on se souvient de (2.18), on a prouvé que :

$$\forall (v,\boldsymbol{q})\in V,\quad a'((v,\boldsymbol{q}),(v,\boldsymbol{q})) \geq \min\left(D_{min}^{-1},\frac{1}{2}(\sigma_{max})^{-1},\frac{1}{2}\sigma_{min}\right)\|(v,\boldsymbol{q})\|_V^2,$$

ce qui signifie que la forme a' est coercive !
De même, la forme $\langle \ell',(w,\boldsymbol{r})\rangle_V = \langle \ell,\mathtt{T}((w,\boldsymbol{r}))\rangle_V$ est telle que $\ell'\in V'$.
Par ailleurs $\mathtt{T}$ est bijective, en effet :

- $\mathtt{T}((v,\boldsymbol{q})) = (0,0)$ signifie que $v+\sigma^{-1}\mathrm{div}\,\boldsymbol{q}=0$ et $\boldsymbol{q}=0$, c'est-à-dire que $(v,\boldsymbol{q})=(0,0)$, d'où l'injectivité ;
- pour $(w,\boldsymbol{r})\in V$, si on pose $\boldsymbol{q}=-\boldsymbol{r}$ et $v=2w+\sigma^{-1}\mathrm{div}\,\boldsymbol{r}$, alors $(v,\boldsymbol{q})\in V$ est tel que $\mathtt{T}((v,\boldsymbol{q})) = (w,\boldsymbol{r})$, d'où la surjectivité.

Si on récapitule, puisque $\mathtt{T}$ est bijective, on a établi que $(u,\boldsymbol{p})$ est solution de (2.17) si, et seulement si, $(u,\boldsymbol{p})$ est solution de

$$\left\{\begin{array}{l}\text{Trouver } (u,\boldsymbol{p})\in V \text{ tel que}\\ \forall (w,\boldsymbol{r})\in V,\quad a'((u,\boldsymbol{p}),(w,\boldsymbol{r})) = \langle \ell',(w,\boldsymbol{r})\rangle_V.\end{array}\right. \tag{2.21}$$

D'après le théorème de Lax-Milgram, la formulation variationnelle (2.21) est bien posée.
Pour résumer, on a utilisé la notion de $\mathtt{T}$-coercivité, qui permet d'établir le caractère bien posé de la formulation variationnelle (2.17). Voici une définition abstraite (cf. [6]).

Définition 2.17 *Soient V un espace de Hilbert défini sur $\mathbb{R}$, et $a(\cdot,\cdot)$ une forme bilinéaire sur $V\times V$. $a(\cdot,\cdot)$ est* $\mathtt{T}$-coercive *si, et seulement si,*

$$\exists \mathtt{T}\in\mathcal{L}(V),\ \textit{bijective},\ \exists \underline{\alpha}>0,\ \forall v\in V,\ a(v,\mathtt{T}v)\geq \underline{\alpha}\|v\|_V^2. \tag{2.22}$$

Ici, $\underline{\alpha} = \min\left(D_{min}^{-1},\frac{1}{2}(\sigma_{max})^{-1},\frac{1}{2}\sigma_{min}\right)$, et on a pour finir l'estimation de stabilité

$$\underline{\alpha}\|(u,\boldsymbol{p})\|_V^2 \ \leq\ a((u,\boldsymbol{p}),\mathtt{T}(u,\boldsymbol{p})) = \langle \ell,\mathtt{T}((u,\boldsymbol{p}))\rangle_V \leq \|\ell\|_{V'}\|\mathtt{T}(u,\boldsymbol{p})\|_V \leq \|\ell\|_{V'}|||\mathtt{T}|||\|(u,\boldsymbol{p})\|_V,$$

c'est-à-dire que

$$\|(u,\boldsymbol{p})\|_V \leq \frac{|||\mathtt{T}|||}{\underline{\alpha}}\|\ell\|_{V'} = \frac{1}{\alpha'}\|\ell\|_{V'},$$

avec α' défini en (2.20).

Exercice : On revient sur les autres conditions aux limites pour le modèle de la diffusion, voir la note de bas de page 1 page 20.

1. Construire la formulation variationnelle à deux inconnues pour la diffusion avec condition aux limites de type réflexion $\mathbb{D}\nabla u\cdot\boldsymbol{n}_{|\partial\Omega}=0$. Etablir que le modèle est bien posé.
2. Construire la formulation variationnelle à une inconnue pour la diffusion avec condition aux limites de type albedo $(\mathbb{D}\nabla u\cdot\boldsymbol{n}+\alpha u)_{|\partial\Omega}=0$, avec $\alpha>0$. Etablir que le modèle est bien posé.

Notes de cours 3

Approximation des formulations variationnelles

On s'intéresse à l'approximation des formulations variationnelles, dont on rappelle la forme générale (espaces de solutions et de fonctions-test non nécessairement égaux). Soient

— V et W deux espaces de Hilbert définis sur $\mathbb{C}$, de dimension infinie ;

— $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V \times W$.

Pour $\ell \in W'$ donné, on étudie la formulation variationnelle,

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ \forall w \in W, \quad a(u,w) = \langle \ell, w \rangle_W. \end{cases} \tag{3.1}$$

On veut approcher la solution de la formulation variationnelle (3.1), que nous supposons bien posée au sens d'Hadamard (voir la définition 2.6). On introduit $(V_\delta)_\delta$ et $(W_\delta)_\delta$ deux suites d'espaces vectoriels de dimensions finies, paramétrées par δ, avec $V_\delta \subset V$ et $W_\delta \subset W$ pour tout δ (*approximation conforme*). Par convention, le paramètre δ prend des valeurs strictement positives, et on a $\lim_{\delta\to 0}(\dim(V_\delta)) = +\infty$. De même pour $(W_\delta)_\delta$.

Comme précédemment, on considérera des espaces de Hilbert V et W définis sur $\mathbb{R}$, avec $a(\cdot,\cdot)$ une forme bilinéaire et continue sur $V \times W$, pour l'approximation du modèle de la diffusion.

3.1 Formulation variationnelle discrète

On va résoudre des problèmes approchés, ou discrets, posés dans les espaces V_δ (espace de solutions discrètes) et W_δ (espace de fonctions-test discrètes), pour un $\delta > 0$. L'approximation de (3.1) s'écrit simplement

$$\begin{cases} \text{Trouver } u_\delta \in V_\delta \text{ tel que} \\ \forall w_\delta \in W_\delta,\ a(u_\delta, w_\delta) = \langle \ell, w_\delta \rangle_W. \end{cases} \tag{3.2}$$

Après introduction de l'application linéaire $\mathtt{A}_\delta \in \mathcal{L}(V_\delta, W_\delta)$ et de $\mathtt{f}_\delta \in W_\delta$ tels que[1]

$$\forall (v_\delta, w_\delta) \in V_\delta \times W_\delta,\ (\mathtt{A}_\delta v_\delta, w_\delta)_W = a(v_\delta, w_\delta),\ \text{ et } \forall w_\delta \in W_\delta,\ (\mathtt{f}_\delta, w_\delta)_W = \langle \ell, w_\delta \rangle_{W_\delta}, \tag{3.3}$$

on peut reformuler (3.2) sous une première forme *équivalente* :

$$\left\{ \begin{array}{l} \text{Trouver } u_\delta \in V_\delta \text{ tel que} \\ \mathtt{A}_\delta u_\delta = \mathtt{f}_\delta \text{ dans } W_\delta. \end{array} \right. \tag{3.4}$$

La formulation variationnelle discrète (3.2) peut aussi être reformulée sous la forme d'un système linéaire. Ci-dessous, on omet la dépendance par rapport à δ...

Lemme 3.1 *Soit $(\varphi_j)_{1\le j\le N}$ une base de l'espace vectoriel V_δ, respectivement $(\psi_i)_{1\le i\le M}$ une base de l'espace vectoriel W_δ. Alors la formulation variationnelle discrète (3.2) est* équivalente *au système linéaire posé dans $\mathbb{C}^N$:*

$$\textit{Trouver } \vec{X} \in \mathbb{C}^N \textit{ tel que } \mathbb{A}\vec{X} = \vec{L}, \tag{3.5}$$

où

$$\forall 1 \le j \le N,\ \forall 1 \le i \le M, \quad \mathbb{A}_{ij} = a(\varphi_j, \psi_i) \textit{ et } L_i = \langle \ell, \psi_i \rangle, \tag{3.6}$$

et on a la relation suivante entre les solutions

$$u_\delta = \sum_{j=1,N} X_j \varphi_j. \tag{3.7}$$

Démonstration : On établit le résultat intermédiaire suivant (à connaître !)

$$\forall w_\delta = \sum_{i=1,M} Y_i \psi_i \in W_\delta, \quad a(u_\delta, w_\delta) = (\mathbb{A}\vec{X}|\vec{Y})_{\mathbb{C}^M}. \tag{3.8}$$

En effet, on a

$$\begin{aligned} (\mathbb{A}\vec{X}|\vec{Y})_{\mathbb{C}^M} &= \sum_{i=1,M} \Big(\sum_{j=1,N} \mathbb{A}_{ij} X_j \Big) \overline{Y_i} = \sum_{i=1,M} \Big(\sum_{j=1,N} a(\varphi_j, \psi_i) X_j \Big) \overline{Y_i} \\ &\overset{a(\cdot,\psi_i)\ \text{linéaire}}{=} \sum_{i=1,M} a\Big(\sum_{j=1,N} X_j \varphi_j, \psi_i \Big) \overline{Y_i} = \sum_{i=1,M} a(u_\delta, \psi_i) \overline{Y_i} \\ &\overset{a(u_\delta,\cdot)\ \text{antilinéaire}}{=} a\Big(u_\delta, \sum_{i=1,M} Y_i \psi_i \Big) = a(u_\delta, w_\delta). \end{aligned}$$

Pour tout $\vec{Y} \in \mathbb{C}^M$, on a donc, en choisissant $w_\delta = \sum_{i=1,M} Y_i \psi_i$:

$$(\mathbb{A}\vec{X}|\vec{Y})_{\mathbb{C}^M} \overset{(3.8)}{=} a(u_\delta, w_\delta) \overset{(3.2)}{=} \langle \ell, w_\delta \rangle_W \overset{\ell\ \text{antilinéaire}}{=} \sum_{i=1,M} \langle \ell, \psi_i \rangle_W \overline{Y_i} \overset{\text{def. } \vec{L}}{=} (\vec{L}|\vec{Y})_{\mathbb{C}^M}.$$

Ainsi, u_δ solution de (3.2) implique que $\vec{X}$ solution de (3.5). La réciproque est évidente. ∎

Enfin, il est clair qu'une condition *nécessaire* pour que la formulation variationnelle discrète admette une solution et une seule est que $\dim(V_\delta) = \dim(W_\delta)$.

1. En dimension finie, toute application linéaire est continue.

3.2 Cas d'une forme coercive

Dans ce cas, on doit résoudre une formulation avec $V = W$. On note $\alpha > 0$ la constante de coercivité apparaissant dans (2.6), ie.

$$\forall v \in V, \quad |a(v,v)| \geq \alpha \|v\|_V^2.$$

On choisit $V_\delta = W_\delta$ pour tout δ pour assurer en particulier que $\dim(V_\delta) = \dim(W_\delta)$. On note $(\varphi_i)_{1\leq i\leq N}$ une base de l'espace vectoriel V_δ.

Proposition 3.2 *Si la forme $a(\cdot,\cdot)$ est coercive, alors pour tout δ la matrice $\mathbb{A}$ est inversible.*

Démonstration : On note que la matrice $\mathbb{A}$ est carrée dans le cas présent. Ainsi $\mathbb{A}$ est inversible si, et seulement si, $\ker(\mathbb{A}) = \{0\}$. Soit donc $\vec{Y} \in \mathbb{C}^M$ tel que $\mathbb{A}\vec{Y} = 0$. On a $(\mathbb{A}\vec{Y}|\vec{Y})_{\mathbb{C}^M} = 0$, et on déduit de (3.8) que $a(w_\delta, w_\delta) = 0$, avec $w_\delta = \sum_{i=1,N} Y_i\varphi_i$. Comme la forme a est coercive, on a nécessairement $w_\delta = 0$, et également $\vec{Y} = 0$. ∎

Remarque 3.3 *Lorsque V est un espace de Hilbert défini sur $\mathbb{R}$, on peut montrer que la matrice est soit* définie-positive*, soit* définie-négative*, pour tout δ (idem lemme 2.5).*

Ainsi, u_δ existe et est unique pour tout δ. On peut alors étudier l'*erreur* entre la solution exacte et la solution approchée, $u - u_\delta$. Et plus précisément, si on a *convergence*, c'est-à-dire si $\lim_{\delta\to 0} \|u - u_\delta\|_V = 0$.

Théorème 3.4 (`Lemme de Céa`) *Si la forme $a(\cdot,\cdot)$ est coercive avec une constante de coercivité $\alpha > 0$, cf. (2.6), on a l'estimation d'erreur*

$$\forall \ell \in V',\ \forall \delta, \quad \|u - u_\delta\|_V \leq \frac{|||a|||}{\alpha} \inf_{v_\delta \in V_\delta} \|u - v_\delta\|_V. \tag{3.9}$$

Si de plus la forme $a(\cdot,\cdot)$ est hermitienne, on a l'estimation d'erreur

$$\forall \ell \in V',\ \forall \delta, \quad \|u - u_\delta\|_V \leq \left(\frac{|||a|||}{\alpha}\right)^{1/2} \inf_{v_\delta \in V_\delta} \|u - v_\delta\|_V. \tag{3.10}$$

Démonstration : On suppose que $u \neq u_\delta$ (dans le cas contraire (3.9)-(3.10) sont vérifiées !).
D'après la coercivité de a, on a en particulier $\alpha\|u - u_\delta\|_V^2 \leq |a(u - u_\delta, u - u_\delta)|$.
Soit maintenant $w_\delta \in V_\delta$: puisque u résout (3.1) et que u_δ résout (3.2), on a

$$a(u, w_\delta - u_\delta) = \langle \ell, w_\delta - u_\delta\rangle_W = a(u_\delta, w_\delta - u_\delta).$$

Ainsi, $a(u - u_\delta, u_\delta - w_\delta) = 0$. On en déduit que

$$\alpha\|u - u_\delta\|_V^2 \leq |a(u - u_\delta, u - u_\delta) + a(u - u_\delta, u_\delta - w_\delta)| = |a(u - u_\delta, u - w_\delta)|.$$

Et, d'après la continuité de a, on a $|a(u - u_\delta, u - w_\delta)| \leq |||a|||\, \|u - u_\delta\|_V\, \|u - w_\delta\|_V$. Si on rassemble les estimations, on trouve, après division par $\|u - u_\delta\|_V \neq 0$,

$$\alpha\|u - u_\delta\|_V \leq |||a|||\, \|u - w_\delta\|_V.$$

Le résultat étant vrai pour tout $w_\delta \in V_\delta$, on en conclut que (3.9) est vraie.
Si de plus $a(\cdot,\cdot)$ est hermitienne, on raisonne différemment. Tout d'abord, on écrit pour $w_\delta \in V_\delta$:

$$\begin{aligned} a(u-w_\delta, u-w_\delta) &= a\big((u-u_\delta)+(u_\delta-w_\delta),(u-u_\delta)+(u_\delta-w_\delta)\big) \\ &= a(u-u_\delta,u-u_\delta)+a(u-u_\delta,u_\delta-w_\delta) \\ &\qquad +a(u_\delta-w_\delta,u-u_\delta)+a(u_\delta-w_\delta,u_\delta-w_\delta) \\ a \text{ hermitienne} &= a(u-u_\delta,u-u_\delta)+2\Re\big(a(u-u_\delta,u_\delta-w_\delta)\big)+a(u_\delta-w_\delta,u_\delta-w_\delta) \\ u \text{ sol. (3.1)}, u_\delta \text{ sol. (3.2)} &= a(u-u_\delta,u-u_\delta)+a(u_\delta-w_\delta,u_\delta-w_\delta). \end{aligned}$$

Ensuite, d'après le lemme 2.4, on a l'alternative (2.8) :

$$\exists \epsilon \in \{-1,+1\},\ \exists \alpha>0,\ \forall v \in V,\ \epsilon a(v,v) \geq \alpha \|v\|_V^2.$$

En particulier, pour tout $v \in V$, on a $|a(v,v)| = \epsilon a(v,v)$. On en conclut que, pour $w_\delta \in V_\delta$:

$$\begin{aligned} |a(u-w_\delta,u-w_\delta)| &= \epsilon a(u-w_\delta,u-w_\delta) \\ &= \epsilon a(u-u_\delta,u-u_\delta)+\epsilon a(u_\delta-w_\delta,u_\delta-w_\delta) \\ &\geq \epsilon a(u-u_\delta,u-u_\delta) = |a(u-u_\delta,u-u_\delta)|. \end{aligned}$$

A l'aide enfin de la coercivité et de la continuité de a, on a obtenu

$$\alpha \|u-u_\delta\|_V^2 \leq |||a|||\,\|u-w_\delta\|_V^2,$$

et (3.10) suit. ∎

Remarque 3.5 *Lorsque V est un espace de Hilbert défini sur $\mathbb{R}$, on a l'estimation (3.9) lorsque la forme est coercive, et l'estimation (3.10) lorsque la forme est de plus symétrique.*

On note que, dans les estimations ci-dessus, la constante $C = |||a|||/\alpha$ ou $C = (|||a|||/\alpha)^{1/2}$ est *indépendante de δ et de u*. On en déduit que la norme de l'erreur est *comparable à* l'erreur d'approximation de u par des éléments de V_δ. En effet,

$$\inf_{v_\delta \in V_\delta} \|u-v_\delta\|_V \leq \|u-u_\delta\|_V \leq C \inf_{v_\delta \in V_\delta} \|u-v_\delta\|_V.$$

Enfin, puisque V_δ est de dimension finie, c'est un sous-espace vectoriel fermé de V, et on a $\inf_{v_\delta \in V_\delta} \|u-v_\delta\|_V = \min_{v_\delta \in V_\delta} \|u-v_\delta\|_V = \|u-P_\delta u\|_V$, où P_δ est la projection orthogonale sur V_δ.

Une *condition suffisante* pour que $\lim_{\delta\to 0} \|u-u_\delta\|_V = 0$ est la propriété d'approximabilité.

Définition 3.6 *La famille $(V_\delta)_\delta$ vérifie la* propriété d'approximabilité *de V si, et seulement si,*

$$\begin{aligned} &\exists V^\star \text{ s.e.v. dense de } V, \quad \forall \delta,\ \exists r_\delta\ :\ V^\star \to V_\delta \text{ tels que} \\ &\forall v^\star \in V^\star, \quad \lim_{\delta\to 0} \|v^\star - r_\delta v^\star\|_V = 0. \end{aligned} \tag{3.11}$$

Proposition 3.7 *Si la propriété d'approximabilité (3.11) est vérifiée, alors*

$$\forall v \in V, \quad \lim_{\delta \to 0} \left(\inf_{v_\delta \in V_\delta} \|v - v_\delta\|_V \right) = 0. \tag{3.12}$$

Démonstration : Soient $v \in V$, et $\eta > 0$ donnés.
D'après (3.11) 1ère ligne, comme $V^\star$ est dense dans V, il existe $v^\star \in V^\star$ tel que $\|v - v^\star\|_V \leq \eta/2$. Et, d'après (3.11) 2ème ligne, il existe $\delta_\eta > 0$ tel que, pour tout $\delta \in]0, \delta_\eta[$, on a $\|v^\star - r_\delta v^\star\|_V \leq \eta/2$. A l'aide de l'inégalité triangulaire, on en déduit que

$$\forall \delta \in]0, \delta_\eta[, \quad \|v - r_\delta v^\star\|_V \leq \|v - v^\star\|_V + \|v^\star - r_\delta v^\star\|_V \leq \eta,$$

soit

$$\forall \delta \in]0, \delta_\eta[, \quad \inf_{v_\delta \in V_\delta} \|v - v_\delta\|_V \leq \eta,$$

ce qui permet d'atteindre le résultat cherché. ■

Théorème 3.8 *Si la forme $a(\cdot, \cdot)$ est coercive, et si la propriété d'approximabilité (3.11) est vérifiée, alors la suite des solutions approchées $(u_\delta)_{\delta > 0}$ converge vers u dans V :*

$$\lim_{\delta \to 0} \|u - u_\delta\|_V = 0. \tag{3.13}$$

Démonstration : Il suffit d'appliquer le lemme de Céa et $(3.12)_{v=u}$. ■

Une fois la convergence de $(u_\delta)_{\delta>0}$ vers u obtenue, on pourra se poser la question de la qualité de l'approximation, ou de la *vitesse de convergence*.

Définition 3.9 (vitesse de convergence) *On dit que l'approximation $(u_\delta)_{\delta>0}$ est convergente à l'ordre $s > 0$ s'il existe une constante $C > 0$, indépendante de δ telle que :*

$$\|u - u_\delta\|_V \leq C\, \delta^s.$$

3.3 Cas général

On se pose maintenant la question de la convergence d'une approximation conforme dans le cas général, c'est-à-dire dans les cas : $V \neq W$; ou bien $V = W$ avec une forme $a(\cdot, \cdot)$ qui n'est pas coercive. Rappelons que, comme la formulation variationnelle (3.1) est *a priori* bien posée au sens d'Hadamard, alors d'après le théorème BNB (théorème 2.12), ceci est équivalent au fait que la forme a vérifie une condition inf-sup (2.13) et une condition de solvabilité (2.14). Pour que les formulations variationnelles discrètes (3.2) soient bien posées, on peut également appliquer le théorème BNB : il faut donc, pour la forme a, une *condition inf-sup discrète* (on dit parfois *stabilité discrète*) ainsi qu'une *condition de solvabilité discrète*, pour tout $\delta > 0$.

Définition 3.10 *La forme a vérifie une* condition inf-sup discrète *si, et seulement si*

$$\forall \delta > 0, \ \exists \alpha_\delta > 0, \ \forall v_\delta \in V_\delta, \quad \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a(v_\delta, w_\delta)|}{\|w_\delta\|_W} \geq \alpha_\delta \|v_\delta\|_V. \tag{3.14}$$

Définition 3.11 *La forme a vérifie une* condition de solvabilité discrète *si, et seulement si,*

$$\forall \delta > 0,\ \{w_\delta \in W_\delta \text{ tel que } a(v_\delta, w_\delta) = 0,\ \forall v_\delta \in V_\delta\} = \{0\}. \tag{3.15}$$

Pour $\delta > 0$, on a introduit l'application linéaire $\mathtt{A}_\delta \in \mathcal{L}(V_\delta, W_\delta)$ associée à $a(\cdot,\cdot)$ sur $V_\delta \times W_\delta$, voir (3.3). On vérifie facilement que la condition (3.14) est équivalente à $\mathtt{A}_\delta$ injective, et que la condition (3.15) est équivalente à $\mathtt{A}_\delta$ surjective. Ces deux conditions sont donc équivalentes à $\mathtt{A}_\delta$ bijective, ce qui est bien sûr équivalent au fait que la formulation variationnelle discrète (3.2) est bien posée. Pour des raisons pratiques, on va remplacer ces deux conditions et définir un nouveau *jeu de conditions équivalentes*. En effet, si $\dim(V_\delta) = \dim(W_\delta)$, alors disposer d'une des deux conditions est nécessaire et suffisant.

Corollaire 3.12 (Banach-Necas-Babuska discret) *Soient V_δ et W_δ deux espaces de Hilbert de dimension finie, $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V_\delta \times W_\delta$. La formulation variationnelle discrète (3.2) est bien posée si, et seulement si, la condition inf-sup discrète (3.14) est satisfaite, et* $\dim(V_\delta) = \dim(W_\delta)$.

Dans la suite on supposera donc que

$$\boxed{\dim(V_\delta) = \dim(W_\delta) \text{ pour tout } \delta > 0}.$$

Mais, contrairement à la coercivité qui se transmet automatiquement aux formulations variationnelles discrètes pour une approximation conforme, la condition inf-sup *ne se transmet pas* ! ? On rappelle la condition inf-sup pour la formulation variationnelle (3.1) ci-dessous :

$$\exists \alpha' > 0,\ \forall v \in V,\ \sup_{w \in W\setminus\{0\}} \frac{|a(v,w)|}{\|w\|_W} \geq \alpha' \|v\|_V.$$

Or, cette condition n'implique pas (3.14), car on n'a pas :

$$\sup_{w_\delta \in W_\delta\setminus\{0\}} \frac{|a(v,w_\delta)|}{\|w_\delta\|_W} \geq \sup_{w \in W\setminus\{0\}} \frac{|a(v,w)|}{\|w\|_W}.$$

En effet, $W_\delta \subset W$ entraîne au contraire que

$$\sup_{w_\delta \in W_\delta\setminus\{0\}} \frac{|a(v,w_\delta)|}{\|w_\delta\|_W} \leq \sup_{w \in W\setminus\{0\}} \frac{|a(v,w)|}{\|w\|_W}.$$

Ainsi, pour obtenir le caractère bien posé des formulations variationnelles discrètes on devra *explicitement* établir la condition inf-sup discrète (3.14).

Théorème 3.13 `(Lemme de Céa)` *Soit $\delta > 0$. Si la forme $a(\cdot,\cdot)$ vérifie la condition inf-sup discrète (3.14), on a l'estimation d'erreur*

$$\forall \ell \in W',\quad \|u - u_\delta\|_V \leq \left(1 + \frac{|||a|||}{\alpha_\delta}\right) \inf_{v_\delta \in V_\delta} \|u - v_\delta\|_V. \tag{3.16}$$

Démonstration : Soit $v_\delta \in V_\delta$. D'après (3.14), on a

$$\exists w'_\delta \in V_\delta \setminus \{0\}, \quad \|u_\delta - v_\delta\|_V \leq \frac{1}{\alpha_\delta} \frac{|a(u_\delta - v_\delta, w'_\delta)|}{\|w'_\delta\|_W}.$$

Et, puisque u résout (3.1) et que u_δ résout (3.2), on a $a(u, w'_\delta) = \langle \ell, w'_\delta \rangle_W = a(u_\delta, w'_\delta)$, et ainsi $a(u - u_\delta, w'_\delta) = 0$. On a donc

$$|a(u_\delta - v_\delta, w'_\delta)| = |a(u_\delta - v_\delta, w'_\delta) + a(u - u_\delta, w'_\delta)| = a(u - v_\delta, w'_\delta)|.$$

Par conséquent,

$$\frac{|a(u_\delta - v_\delta, w'_\delta)|}{\|w'_\delta\|_W} = \frac{|a(u - v_\delta, w'_\delta)|}{\|w'_\delta\|_W} \leq |||a||| \, \|u - v_\delta\|_V.$$

Si on revient à l'estimation initiale, on a finalement

$$\|u_\delta - v_\delta\|_V \leq \frac{|||a|||}{\alpha_\delta} \|u - v_\delta\|_V.$$

Par inégalité triangulaire, on trouve

$$\|u - u_\delta\|_V \leq \|u - v_\delta\|_V + \|v_\delta - u_\delta\|_V \leq \left(1 + \frac{|||a|||}{\alpha_\delta}\right) \|u - v_\delta\|_V,$$

et, comme c'est valable pour tout $v_\delta \in V_\delta$, on peut prendre l'infimum et en déduire (3.9). ∎

Pour que la norme de l'erreur soit *comparable à* l'erreur d'approximation entre V_δ et V, il faut que la constante apparaissant dans (3.9) soit *uniformément majorée* par rapport à δ ou, ce qui est équivalent, que les constantes $(\alpha_\delta)_\delta$ soient *uniformément minorées* par un nombre strictement positif : c'est ce qu'on appelle la *condition inf-sup discrète uniforme* (ou de *stabilité discrète uniforme*).

Définition 3.14 *La forme a vérifie une* condition inf-sup discrète uniforme *si, et seulement si,*

$$\exists \alpha_\dagger > 0, \ \forall \delta > 0, \ \forall v_\delta \in V_\delta, \ \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a(v_\delta, w_\delta)|}{\|w_\delta\|_W} \geq \alpha_\dagger \|v_\delta\|_V. \tag{3.17}$$

Si cette condition est vérifiée (encore une fois, c'est une autre condition que la condition inf-sup (2.13) vérifiée dans $V \times W$), on a convergence à l'aide de la propriété d'approximabilité.

Théorème 3.15 *Si la forme $a(\cdot, \cdot)$ vérifie une condition inf-sup discrète uniforme, et si la propriété d'approximabilité (3.11) est vérifiée, alors la suite des solutions approchées $(u_\delta)_{\delta>0}$ converge vers u dans V :*

$$\lim_{\delta \to 0} \|u - u_\delta\|_V = 0. \tag{3.18}$$

Démonstration : Il suffit d'appliquer le lemme de Céa (théorème 3.13) et $(3.12)_{v=u}$. ∎

Une fois la convergence de $(u_\delta)_{\delta>0}$ vers u obtenue, on pourra se poser la question de la vitesse de convergence, cf. la définition 3.9.

3.4 Application à la diffusion neutronique

3.4.1 Diffusion à une inconnue

On commence par étudier la discrétisation de la formulation variationnelle (2.4), posée dans $H^1_0(\Omega)$, où Ω est un domaine, qu'on réécrit ici :

$$\left\{\begin{array}{l} \text{Trouver } u \in H^1_0(\Omega) \text{ tel que} \\ \forall w \in H^1_0(\Omega), \quad \displaystyle\int_\Omega \Big(\mathbb{D}\nabla u \cdot \nabla w + \sigma u w\Big) d\Omega = \int_\Omega S_f w \, d\Omega, \end{array}\right. \tag{3.19}$$

avec un $\mathbb{D}$ un champ de matrices *symétriques* vérifiant (2.2), et σ vérifiant (2.3). Les hypothèses (2.2)-(2.3) sur les coefficients $\mathbb{D}$ et σ garantissent que la forme bilinéaire est coercive, d'où le caractère bien posé. On va définir des sous-espaces discrets à l'aide de la méthode des éléments finis (voir [15]). Pour simplifier l'exposé, on suppose que le domaine Ω est polygonal si $\mathtt{d} = 2$, ou polyédrique si $\mathtt{d} = 3$. Dans ce cadre, le paramètre de discrétisation est égal au pas du maillage (voir plus bas), et est noté h. On note $(V^0_h)_h$ la suite des sous-espaces discrets de $H^1_0(\Omega)$. Pour $h > 0$ donné, la formulation variationnelle discrète s'écrit :

$$\left\{\begin{array}{l} \text{Trouver } u_h \in V^0_h \text{ tel que} \\ \forall w_h \in V^0_h, \quad \displaystyle\int_\Omega \Big(\mathbb{D}\nabla u_h \cdot \nabla w_h + \sigma u_h w_h\Big) d\Omega = \int_\Omega S_f w_h \, d\Omega. \end{array}\right. \tag{3.20}$$

Le caractère bien posé découle du fait que la forme bilinéaire est coercive (cf. §3.2).

Pour construire les sous-espaces discrets $(V^0_h)_h$, on s'appuie sur des maillages $(\mathcal{T}_h)_h$ de Ω, formés de simplexes (triangles si $\mathtt{d} = 2$; tétraèdres si $\mathtt{d} = 3$). Soit un maillage [2] $\mathcal{T}_h$ donné : on appelle $(T_\ell)_{\ell=1,L}$ l'ensemble des simplexes le composant (par convention, ce sont des fermés de $\mathbb{R}^{\mathtt{d}}$). Le *pas du maillage* $\mathcal{T}_h$ est égal à $h = \max_\ell h_\ell$, où h_ℓ est le rayon de la plus petite boule contenant T_ℓ.
Pour un simplexe T quelconque et $k \in \mathbb{N}$, on définit [3] :

$$P^k(T) = \{q \in P(T) \text{ tel que } d^\circ(q) \le k\}.$$

A partir de là, on définit classiquement V^0_h en deux étapes. Soit $k \ge 1$.

1. On introduit d'abord [4] $V_h \subset H^1(\Omega)$:

$$V_h = \{v_h \in H^1(\Omega) \text{ tel que } v_{h|T_\ell} \in P^k(T_\ell),\ \forall \ell = 1, L\}. \tag{3.21}$$

2. Un maillage de Ω formé de simplexes est un ensemble de simplexes fermés $(T_\ell)_{\ell=1,L}$ tels que :
– $\overline{\Omega} = \cup_\ell T_\ell$;
– $\forall \ell$, $\mathrm{int}(T_\ell) \neq \emptyset$;
– $\forall \ell_1 \neq \ell_2$, $\mathrm{int}(T_{\ell_1}) \cap \mathrm{int}(T_{\ell_2}) = \emptyset$;
– $\forall \ell$, toute facette de T_ℓ est soit incluse dans $\partial\Omega$, soit une facette d'un autre simplexe.
3. Lorsque $k = 1$, on a aussi $P^1(T) = \{q \in P(T) \text{ tel que } q(\boldsymbol{x}) = a + \boldsymbol{b} \cdot \boldsymbol{x},\ \forall \boldsymbol{x} \in T\}$.
4. On a la *définition équivalente* : $V_h = \{v_h \in \mathcal{C}^0(\overline{\Omega}) \text{ tel que } v_{h|T_\ell} \in P^k(T_\ell),\ \forall \ell = 1, L\}$.

2. Puis on introduit $V_h^0 \subset H_0^1(\Omega)$:

$$V_h^0 = V_h \cap H_0^1(\Omega) = \{v_h \in V_h \text{ tel que } v_{h|\partial\Omega} = 0\}. \tag{3.22}$$

Ceci correspond à la discrétisation conforme par *éléments finis de Lagrange d'ordre k*.
On fait en outre l'hypothèse que la famille de maillages $(\mathcal{T}_h)_h$ est régulière. Pour h donné, et pour un simplexe $T_\ell \in \mathcal{T}_h$, on introduit ρ_ℓ le rayon de la plus grande boule incluse dans T_ℓ. On dit que la famille de maillages $(\mathcal{T}_h)_h$ est *régulière* si

$$\exists \varsigma > 0,\ \forall h,\ \forall T_\ell \in \mathcal{T}_h, \quad \frac{h_\ell}{\rho_\ell} \leq \varsigma.$$

Remarque 3.16 *Pour la discrétisation par éléments finis de Lagrange, nous donnons quelques détails pour réaliser les calculs élémentaires, c'est-à-dire les intégrations sur les simplexes du maillage. Pour cela, on peut effectuer un changement de variables et passer par le* simplexe de référence*, noté $\hat{T}$, dont les sommets sont l'origine, et les $\mathtt{d}$ points dont les coordonnées sont $(\mathtt{d}-1)$ 0, et un 1... Pour un simplexe T_ℓ, on note $F_\ell : \hat{T} \to T_\ell$ la transformation affine transformant $\hat{T}$ en T_ℓ, et on utilise le changement de variables $\boldsymbol{x} = F_\ell(\hat{\boldsymbol{x}}) = \mathbb{A}_\ell \hat{\boldsymbol{x}} + \boldsymbol{b}_\ell$, avec $\mathbb{A}_\ell \in \mathbb{R}^{\mathtt{d}\times\mathtt{d}}$ et $\boldsymbol{b}_\ell \in \mathbb{R}^{\mathtt{d}}$. Le jacobien de la transformation vaut $J_{F_\ell}(\hat{\boldsymbol{x}}) = det(dF_\ell(\hat{\boldsymbol{x}})) = det(\mathbb{A}_\ell)$, et on a $d\boldsymbol{x} = |J_{F_\ell}(\hat{\boldsymbol{x}})|\, d\hat{\boldsymbol{x}}$; notons que $det(\mathbb{A}_\ell) = \pm 2mes(T_\ell)$, selon l'orientation des sommets de T_ℓ. A une fonction v définie sur T_ℓ, on associe alors $\hat{v}$ définie sur $\hat{T}$ telle que $\hat{v}(\hat{\boldsymbol{x}}) = v(F_\ell(\hat{\boldsymbol{x}})) = v(\boldsymbol{x})$ pour tout $\hat{\boldsymbol{x}} \in \hat{T}$, et on a $\nabla v(\boldsymbol{x}) = (\mathbb{A}_\ell^T)^{-1}\hat{\nabla}\hat{v}(\hat{\boldsymbol{x}})$.*

On montre que les fonctions de V_h peuvent être caractérisées par des valeurs ponctuelles (ce qui semble raisonnable, puisque qu'elles appartiennent à $\mathcal{C}^0(\overline{\Omega})$ et sont donc définies partout).

Nous détaillons ci-dessous pour les éléments finis de Lagrange d'ordre $k = 1$. On note $(M_i)_{i=1,N}$ les sommets du maillage $\mathcal{T}_h$, et $(w_i)_{i=1,N}$ les fonctions de base associées aux sommets, caractérisées par : pour tout $1 \leq i \leq N$, on a $w_i \in V_h$ avec pour tout $1 \leq j \leq N$, $w_i(M_j) = \delta_{ij}$. Pour un élément $v_h \in V_h$, on vérifie que $v_h(\boldsymbol{x}) = \sum_{i=1,N} v_h(M_i) w_i(\boldsymbol{x})$ pour tout $\boldsymbol{x} \in \overline{\Omega}$. Ainsi, les *degrés de liberté* associés, c'est-à-dire les quantités caractérisant v_h, sont les quantités $(v_h(M_i))_{i=1,N}$. Ensuite, si on note $\mathcal{I}$ l'ensemble des indices correspondant aux sommets n'appartenant pas à la frontière, c'est-à-dire $\mathcal{I} = \{i \in \{1, \cdots, N\} \text{ tels que } M_i \notin \partial\Omega\}$, on note que $(w_i)_{i\in\mathcal{I}}$ est une base de V_h^0.

Remarque 3.17 *En suivant le formalisme du lemme 3.1 avec $(w_i)_{i\in\mathcal{I}}$ comme base de V_h^0, on obtient un système linéaire équivalent à (3.20), et on peut montrer que la matrice $\mathbb{A} \in \mathbb{R}^{|\mathcal{I}|\times|\mathcal{I}|}$ est creuse, au sens suivant. Si on indice par le pas du maillage, alors le nombre d'éléments non-nuls de $\mathbb{A}_h$, noté $nnz(\mathbb{A}_h)$, peut en principe croître comme $|\mathcal{I}_h|^2$ lorsque h tend vers 0. On montre en fait que $nnz(\mathbb{A}_h)$ est borné par $C|\mathcal{I}_h|$, avec $C > 0$ une constante qui ne dépend pas de h. Ce résultat est toujours vrai lorsque $\mathtt{d} = 2$. Lorsque $\mathtt{d} = 3$, il faut que la famille des maillages $(\mathcal{T}_h)_h$ soit régulière.*

La propriété d'approximabilité (3.11) qui implique la convergence de l'erreur vers 0 est obtenue grâce à un opérateur d'interpolation. Nous rappelons le plus "classique" ci-dessous (à h donné).

$$\pi_h \ : \ \left| \begin{array}{lcl} \mathcal{C}^0(\overline{\Omega}) & \to & V_h \\ v & \mapsto & \pi_h v = \sum_{i=1,N} v(M_i) w_i \end{array} \right. . \tag{3.23}$$

On utilise la définition de $H^1_0(\Omega)$: $\mathcal{D}(\Omega)$ est dense dans $H^1_0(\Omega)$, et on introduit les opérateurs $r_h = \pi_{h|\mathcal{D}(\Omega)}$, à valeurs dans V_h^0. On peut prouver par le calcul que, pour tout $v^\star \in \mathcal{D}(\Omega)$, on a $\lim_{h\to 0} \|v^\star - \pi_h v^\star\|_{H^1(\Omega)} = 0$ (voir par exemple le chapitre 3 de [13]). [5] D'après le théorème 3.8, la convergence suit :

$$\lim_{h\to 0} \|u - u_h\|_{H^1(\Omega)} = 0.$$

Pour les éléments finis de Lagrange d'ordre $k \geq 2$, les degrés de liberté sont les valeurs ponctuelles aux sommets, ainsi qu'en certains points des arêtes, des faces (si $\mathtt{d} = 3$) et de l'intérieur des simplexes. Concernant le nombre d'éléments non-nuls de $\mathbb{A}_h$, il est toujours borné par $C_k|\mathcal{I}_h|$, où $\mathcal{I}_h$ est l'ensemble des indices de degrés de liberté ne se trouvant pas sur $\partial\Omega$, avec $C_k > 0$ une constante qui dépend uniquement de k, mais pas de h. La propriété d'approximabilité (3.11) est obtenue à l'aide de l'opérateur d'interpolation π_h défini en (3.23).

Concernant la vitesse de convergence, on a un résultat "classique" en théorie des éléments finis, que nous admettons (voir [15] pour des éléments de preuve).

Proposition 3.18 *Supposons que la solution u est de régularité H^{1+s} "par morceaux" sur une partition de Ω, où s appartient à $\mathbb{R}^*_+$,* [6] *alors si on utilise les éléments finis de Lagrange d'ordre k sur une famille régulière de maillages qui respectent la partition, il existe une constante $C > 0$ indépendante de h telle que*

$$\|u - u_h\|_{H^1(\Omega)} \leq C\, h^{\min(s,k)}.$$

Ainsi, la vitesse de convergence dépend d'une part de la régularité "par morceaux" de la solution, et d'autre part de l'ordre des éléments finis.

On peut également obtenir une vitesse de convergence *minimale*. Pour cela, on utilise le fait que, si la donnée est "régulière", alors la solution l'est également. Précisément, on admet (voir par exemple [14]) que la formulation variationnelle (3.19) possède la *propriété de décalage* suivante (S_f est la donnée, et u la solution) :

$$\begin{array}{l} \text{il existe un } \textit{exposant de décalage } r_{max} \in]0,1] \text{ tel que, pour tout } S_f \in L^2(\Omega) : \\ \text{- si } r_{max} < 1,\ u \text{ appartient à } H^{1+r} \text{ "par morceaux",}^{6} \text{ pour tout } r \in [0, r_{max}[, \\ \text{- si } r_{max} = 1,\ u \text{ appartient à } H^2 \text{ "par morceaux".}^{6} \end{array} \tag{3.24}$$

5. On a aussi une propriété d'approximabilité pour une discrétisation dans $H^1(\Omega)$, à l'aide de $r_h = \pi_{h|\mathcal{C}^\infty_c(\overline{\Omega})}$.

6. Intuitivement, ceci signifie qu'il existe une *partition* de Ω en sous-domaines disjoints $(\Omega_p)_{p=1,P}$, telle que la solution restreinte à chaque sous-domaine Ω_p appartient à $H^{1+s}(\Omega_p)$. Si $s \in \mathbb{N}^*$, ce sont les espaces de Sobolev "usuels" de la définition 1.17. Si $s \in \mathbb{R}_+ \setminus \mathbb{N}$, on a une appartenance "par morceaux" à des espaces de Sobolev d'ordre fractionnaire... Principale propriété des espaces de Sobolev d'ordre positif : pour tous $s_1, s_2 \in \mathbb{R}_+$ tels que $s_1 < s_2$, $H^{1+s_2}(\Omega_p)$ est un sous-ensemble strict de $H^{1+s_1}(\Omega_p)$, avec injection continue et compacte.

De plus, la norme H^{1+r} "par morceaux" de u, pour tout $r \in [0, r_{max}[$ (si $r_{max} < 1$), ou la norme H^2 "par morceaux" de u (si $r_{max} = 1$) dépend continûment de $\|S_f\|_{L^2(\Omega)}$.
Si la donnée S_f est dans $L^2(\Omega)$, on en déduit donc le résultat *minimal* de convergence pour la diffusion à une inconnue .[7]

Théorème 3.19 *Soit $r_{max} \in]0, 1]$ l'exposant de décalage défini comme ci-dessus. Alors on a l'estimation suivante en norme $\|\cdot\|_{H^1(\Omega)}$:*

- *si $r_{max} < 1$: $\forall r \in [0, r_{max}[$, $\exists C'_r > 0$ telle que $\forall h$, $\forall S_f \in L^2(\Omega)$, $\|u - u_h\|_{H^1(\Omega)} \leq C'_r\, h^r \|S_f\|_{L^2(\Omega)}$,*
- *si $r_{max} = 1$: $\exists C'_1 > 0$ telle que $\forall h$, $\forall S_f \in L^2(\Omega)$, $\|u - u_h\|_{H^1(\Omega)} \leq C'_1\, h \|S_f\|_{L^2(\Omega)}$.*

Si maintenant on s'intéresse à la convergence en norme $\|\cdot\|_{L^2(\Omega)}$, qui est une *norme plus faible* que la norme $\|\cdot\|_{H^1(\Omega)}$, on a en général une *meilleure vitesse de convergence*. En effet, à partir de (3.24), on établit le résultat de convergence en norme $\|\cdot\|_{L^2(\Omega)}$ à l'aide du lemme d'Aubin-Nitsche, qui garantit, entre la vitesse de convergence en norme $\|\cdot\|_{H^1(\Omega)}$ et celle en norme $\|\cdot\|_{L^2(\Omega)}$, un gain de $O(h^r)$ pour tout $r < r_{max}$ si $r_{max} < 1$, et de $O(h)$ si $r_{max} = 1$.
Pour la diffusion à une inconnue avec donnée S_f dans $L^2(\Omega)$, on en déduit donc le second résultat *minimal* de convergence ci-dessous.

Corollaire 3.20 *Soit $r_{max} \in]0, 1]$ l'exposant de décalage défini comme ci-dessus. Alors on a l'estimation suivante en norme $\|\cdot\|_{L^2(\Omega)}$:*

- *si $r_{max} < 1$: $\forall r \in [0, r_{max}[$, $\exists C''_r > 0$ telle que $\forall h$, $\forall S_f \in L^2(\Omega)$, $\|u - u_h\|_{L^2(\Omega)} \leq C''_r\, h^{2r} \|S_f\|_{L^2(\Omega)}$,*
- *si $r_{max} = 1$: $\exists C''_1 > 0$ telle que $\forall h$, $\forall S_f \in L^2(\Omega)$, $\|u - u_h\|_{L^2(\Omega)} \leq C''_1\, h^2 \|S_f\|_{L^2(\Omega)}$.*

Situation "typique" pour la résolution numérique de la diffusion à une inconnue : avec des éléments finis de Lagrange d'ordre $k = 1$, et un exposant de décalage $r_{max} = 1$, l'erreur $u - u_h$ converge vers 0 à l'ordre 1 en norme $\|\cdot\|_{H^1(\Omega)}$, et à l'ordre 2 en norme $\|\cdot\|_{L^2(\Omega)}$.

3.4.2 Diffusion à deux inconnues

On veut maintenant discrétiser la formulation variationnelle (2.17), posée dans $L^2(\Omega) \times \boldsymbol{H}(\mathrm{div}\,;\Omega)$, où Ω est un domaine, qu'on réécrit :

$$\left\{ \begin{array}{l} \text{Trouver } (u, \boldsymbol{p}) \in L^2(\Omega) \times \boldsymbol{H}(\mathrm{div}\,;\Omega) \text{ tel que} \\ \forall (w, \boldsymbol{r}) \in L^2(\Omega) \times \boldsymbol{H}(\mathrm{div}\,;\Omega), \\ \qquad \displaystyle\int_\Omega \Big(-\mathbb{D}^{-1}\boldsymbol{p}\cdot\boldsymbol{r} + u(\mathrm{div}\,\boldsymbol{r}) + w(\mathrm{div}\,\boldsymbol{p}) + \sigma u w \Big) d\Omega = \int_\Omega S_f w \, d\Omega. \end{array} \right.$$

On suppose toujours que le coefficient $\mathbb{D}$ est un champ de matrices *symétriques* vérifiant (2.2), et que le coefficient σ est un champ de scalaires vérifiant (2.3), qui est de plus *constant par*

7. La vitesse de convergence sera meilleure, sous réserve que $\min(s, k) > r_{max}$. Pour cela, il faut et il suffit que : (1) la solution u soit extra-régulière ($s > r_{max}$) ; (2) l'ordre k vérifie $k > r_{max}$.

morceaux, voir plus bas (nouvelle hypothèse simplificatrice ; on peut s'en abstraire, voir le §5.1 de [14]). On utilise à nouveau la méthode des éléments finis, dans un domaine Ω supposé polygonal si $\mathtt{d} = 2$, ou polyédrique si $\mathtt{d} = 3$. Dans ce cadre, le paramètre de discrétisation est égal au pas du maillage.

On note cette fois $(M_h)_h$ la suite des sous-espaces discrets de $L^2(\Omega)$, et $(\boldsymbol{Q}_h)_h$ la suite des sous-espaces discrets de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$. Pour $h > 0$ donné, la formulation variationnelle discrète s'écrit :

$$\left\{\begin{array}{l} \text{Trouver } (u_h, \boldsymbol{p}_h) \in M_h \times \boldsymbol{Q}_h \text{ tel que} \\ \forall (w_h, \boldsymbol{r}_h) \in M_h \times \boldsymbol{Q}_h, \\ \qquad \displaystyle\int_\Omega \Big(-\mathbb{D}^{-1}\boldsymbol{p}_h \cdot \boldsymbol{r}_h + u_h(\mathrm{div}\,\boldsymbol{r}_h) + w_h(\mathrm{div}\,\boldsymbol{p}_h) + \sigma u_h w_h \Big) d\Omega = \int_\Omega S_f w_h \, d\Omega. \end{array}\right. \tag{3.25}$$

Comme on l'a noté au §3.3, il faut disposer d'une condition inf-sup discrète (uniforme) dans $(M_h \times \boldsymbol{Q}_h)_h$. Ceci signifie qu'*on ne peut pas choisir M_h et $\boldsymbol{Q}_h$ indépendamment l'un de l'autre.* Un choix possible est celui des *éléments finis de Raviart-Thomas d'ordre k*. Les maillages $(\mathcal{T}_h)_h$ sont similaires, et possèdent les mêmes propriétés, que ceux précédemment utilisés, avec la *contrainte supplémentaire* que le coefficient σ restreint à tout simplexe est *constant* sur ce simplexe (de façon plus imagée, si on partitionne Ω en sous-domaines sur lesquels σ est constant, alors les maillages respectent cette partition).

Etape 1 : définition de $\boldsymbol{Q}_h \subset \boldsymbol{H}(\mathrm{div}\,;\Omega)$.

En préambule, on note que, comme on discrétise des champs à valeurs vectorielles, on considère des approximations à l'aide de polynômes à coefficients vectoriels, c'est-à-dire qu'on se place localement dans des sous-espaces vectoriels de $\boldsymbol{P}(T) = (P(T))^{\mathtt{d}}$, où T est un simplexe.

Si on a un champ $\boldsymbol{q}_h$ régulier (polynomial) par simplexe, à quelles conditions a-t-on $\boldsymbol{q}_h \in \boldsymbol{H}(\mathrm{div}\,;\Omega)$? Pour répondre à cette question, soient deux simplexes distincts T, T' de $\mathcal{T}_h$ ayant une arête a en commun (si $\mathtt{d} = 2$) ; une face f en commun (si $\mathtt{d} = 3$). On rappelle la condition sur le saut de la composante normale énoncée au théorème 1.36 :

$$\boldsymbol{q}_h \in \boldsymbol{H}(\mathrm{div}\,; int(T \cup T')) \quad \Longleftrightarrow \quad \left| \begin{array}{ll} [\boldsymbol{q}_h \cdot \boldsymbol{n}_a]_a = 0 & \text{si } \mathtt{d} = 2\,; \\ [\boldsymbol{q}_h \cdot \boldsymbol{n}_f]_f = 0 & \text{si } \mathtt{d} = 3. \end{array}\right.$$

Et, plus globalement, l'appartenance de $\boldsymbol{q}_h$ à $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ est garantie si, et seulement si, le saut de la composante normale sur toutes les arêtes communes à deux triangles de $\mathcal{T}_h$ est nul (si $\mathtt{d} = 2$) ; le saut de la composante normale sur toutes les faces communes à deux tétraèdres de $\mathcal{T}_h$ est nul (si $\mathtt{d} = 3$). On parle d'arêtes internes, respectivement de faces internes, de $\mathcal{T}_h$. Les degrés de liberté associés à cette composante normale seront donc identiques des deux côtés des arêtes ($d = 2$) ou faces ($d = 3$) internes.

Remarque 3.21 *La continuité des autres composantes de la trace de $\boldsymbol{q}_h$ n'est pas requise. Par conséquent, les éléments de $\boldsymbol{Q}_h$ ne sont pas obligatoirement continus, c'est-à-dire que $\boldsymbol{Q}_h \not\subset \boldsymbol{C}^0(\overline{\Omega})$ est possible.*

Nous introduisons, pour $k \in \mathbb{N}$:[8] [9]

$$\boldsymbol{Q}_h = \{\boldsymbol{q}_h \in \boldsymbol{H}(\mathrm{div}\,;\Omega) \text{ tel que } \boldsymbol{q}_{h|T_\ell} \in \boldsymbol{P}^k(T_\ell) + P^k(T_\ell)\boldsymbol{x},\ \forall \ell = 1, L\}. \tag{3.26}$$

Nous détaillons ci-dessous pour les éléments finis de Raviart-Thomas d'ordre $k = 0$. Notons que, si $\mathtt{d} = 2$, dans tout triangle T_ℓ de $\mathcal{T}_h$ il y a 3 arêtes, et que, si $\mathtt{d} = 3$, dans tout tétraèdre T_ℓ de $\mathcal{T}_h$ il y a 4 faces. Or $\dim(\boldsymbol{P}^0(T_\ell) + P^0(T_\ell)\boldsymbol{x}) = \mathtt{d}+1$! Intuitivement, il semble raisonnable de considérer un degré de liberté par arête si $\mathtt{d} = 2$, par face si $\mathtt{d} = 3$: on peut directement vérifier par le calcul que cette définition est possible (voir à nouveau le chapitre 3 de [13]). Globalement, on note $(a_i)_{i=1,A}$ les arêtes du maillage $\mathcal{T}_h$ (si $\mathtt{d} = 2$) ; $(f_i)_{i=1,F}$ les faces du maillage $\mathcal{T}_h$ (si $\mathtt{d} = 3$). Dans ce cas, les *degrés de liberté* associés, c'est-à-dire les quantités caractérisant $\boldsymbol{q}_h$, sont les quantités :

$$\left| \begin{array}{ll} \displaystyle\left(\int_{a_i} \boldsymbol{q}_h \cdot \boldsymbol{n}_a \, dl\right)_{i=1,A} & \text{si } \mathtt{d} = 2\,; \\ \displaystyle\left(\int_{f_i} \boldsymbol{q}_h \cdot \boldsymbol{n}_f \, dS\right)_{i=1,F} & \text{si } \mathtt{d} = 3. \end{array} \right. \tag{3.27}$$

Observons que ce ne sont plus des valeurs ponctuelles, mais des moments.
On peut construire une base de $\boldsymbol{Q}_h$ avec, pour chaque arête interne (si $\mathtt{d} = 2$), ou pour chaque face interne (si $\mathtt{d} = 3$), commune à deux simplexes T et T', une fonction de base associée, de support égal à $T \cup T'$, dont le degré de liberté vaut 1 sur ladite arête, ou face, et 0 ailleurs. Et pour chaque arête de la frontière (si $\mathtt{d} = 2$), ou pour chaque face de la frontière (si $\mathtt{d} = 3$), appartenant à un simplexe T, une fonction de base associée, de support égal à T, dont le degré de liberté vaut 1 sur ladite arête, ou face, et 0 ailleurs.

Pour les éléments finis de Raviart-Thomas d'ordre $k \geq 1$. Le processus de construction est similaire, nous renvoyons au chapitre 2 de [3]. Notons qu'on doit définir des degrés de liberté supplémentaires, toujours de type moment, avec notamment des intégrales sur les simplexes (et des fonctions de bases associées dont le support est égal à ce simplexe).

Etape 2 : définition de $M_h \subset L^2(\Omega)$. Le point de départ est la définition de l'espace discret $\boldsymbol{Q}_h$ par (3.26). On observe que, pour tout $\boldsymbol{q}_h \in \boldsymbol{Q}_h$, on a d'une part $\mathrm{div}\,\boldsymbol{q}_h \in L^2(\Omega)$, et de plus, pour tout $T_\ell \in \mathcal{T}_h$, on a $\mathrm{div}\,\boldsymbol{q}_{h|T_\ell} \in P^k(T_\ell)$. Ceci suggère de définir

$$M_h = \{v_h \in L^2(\Omega) \text{ tel que } v_{h|T_\ell} \in P^k(T_\ell),\ \forall \ell = 1, L\}. \tag{3.28}$$

8. Lorsque $k = 0$, on a aussi $\boldsymbol{P}^0(T) + P^0(T)\boldsymbol{x} = \{q \in \boldsymbol{P}(T) \text{ tel que } q(\boldsymbol{x}) = \boldsymbol{a} + b\,\boldsymbol{x},\ \forall \boldsymbol{x} \in T\}$.
9. Par construction, on observe que :
   — Si $\mathtt{d} = 2$: alors, pour tout h, pour tout $T_\ell \in \mathcal{T}_h$, pour toutes les arêtes $(a_\alpha)_{\alpha=1,3}$ de ∂T_ℓ, et pour tout $\boldsymbol{q}_h \in \boldsymbol{Q}_h$, on a $\boldsymbol{q}_h \cdot \boldsymbol{n}_{|a_\alpha} \in P^k(a_\alpha)$;
   — Si $\mathtt{d} = 3$: alors, pour tout h, pour tout $T_\ell \in \mathcal{T}_h$, pour toutes les faces $(f_\alpha)_{\alpha=1,4}$ de ∂T_ℓ, et pour tout $\boldsymbol{q}_h \in \boldsymbol{Q}_h$, on a $\boldsymbol{q}_h \cdot \boldsymbol{n}_{|f_\alpha} \in P^k(f_\alpha)$.

De cette façon, on a $\operatorname{div}[\boldsymbol{Q}_h] \subset M_h$. *L'intérêt de la définition de M_h à partir de celle de $\boldsymbol{Q}_h$ est qu'on va pouvoir directement transposer, pour la diffusion à deux inconnues, la démonstration de la condition inf-sup (cf. §2.3.3), ou celle de la la* T*-coercivité (cf. §2.3.4), au cas discret.*
Pour cela, pour $(v_h, \boldsymbol{q}_h) \in M_h \times \boldsymbol{Q}_h$ non-nul, nous allons chercher $(w_h^\star, \boldsymbol{r}_h^\star) \in M_h \times \boldsymbol{Q}_h$ non-nul, dépendant linéairement de $(v_h, \boldsymbol{q}_h)$, et tel que

$$|a((v_h, \boldsymbol{q}_h), (w_h^\star, \boldsymbol{r}_h^\star))| \geq \alpha_\dagger \|(v_h, \boldsymbol{q}_h)\|_V \, \|(w_h^\star, \boldsymbol{r}_h^\star)\|_V,$$

avec $\|(v, \boldsymbol{q})\|_V = (\|v\|^2_{L^2(\Omega)} + \|\boldsymbol{q}\|^2_{\boldsymbol{H}(\operatorname{div};\Omega)})^{1/2}$, et une constante $\alpha_\dagger > 0$ indépendante de $(v_h, \boldsymbol{q}_h)$ et de h. En s'inspirant des calculs du §2.3.3, on choisit $(w_h^\star, \boldsymbol{r}_h^\star) = (\frac{1}{2}(v_h + \sigma^{-1}\operatorname{div}\boldsymbol{q}_h), -\boldsymbol{q}_h)$. Par construction, et puisque par hypothèse σ^{-1} est constant par morceaux, on a, pour tout $T_\ell \in \mathcal{T}_h$, $\sigma^{-1}\operatorname{div}\boldsymbol{q}_{h|T_\ell} \in P^k(T_\ell)$, et ainsi $w_h^\star \in M_h$. En reprenant *exactement* les calculs du §2.3.3, avec des indices $_h$ en plus, on trouve que

$$a((v_h, \boldsymbol{q}_h), (w_h^\star, \boldsymbol{r}_h^\star)) \geq \frac{\min\left(D_{min}^{-1}, \frac{1}{2}(\sigma_{max})^{-1}, \frac{1}{2}\sigma_{min}\right)}{\left(\frac{1}{2}(\sigma_{min})^{-2} + 1\right)^{1/2}} \|(v_h, \boldsymbol{q}_h)\|_V \, \|(w_h^\star, \boldsymbol{r}_h^\star)\|_V,$$

ce qui est le résultat cherché.
Il est important de remarquer qu'on a utilisé une preuve "adaptant" la T-coercivité introduite au §2.3.4. En effet, si on considère $\mathtt{T}_h \in \mathcal{L}(M_h \times \boldsymbol{Q}_h)$ définie par

$$\mathtt{T}_h((v_h, \boldsymbol{q}_h)) = (\frac{1}{2}(v_h + \sigma^{-1}\operatorname{div}\boldsymbol{q}_h), -\boldsymbol{q}_h),$$

on note que, pour tout h, la forme a restreinte à $(M_h \times \boldsymbol{Q}_h) \times (M_h \times \boldsymbol{Q}_h)$ est T-coercive au sens de la définition 2.17, avec une constante $\underline{\alpha}$ indépendante de h. Qui plus est, la norme de l'application linéaire $\mathtt{T}_h$ est également indépendante de h. On précisera cette notion de "T-coercivité discrète" au chapitre 9 (voir la définition 9.15).

Avant de conclure à la convergence de l'erreur vers 0 dans $L^2(\Omega) \times \boldsymbol{H}(\operatorname{div};\Omega)$, on note que la propriété d'approximabilité (3.11) est obtenue :

- pour $(M_h)_h$ dans $L^2(\Omega)$, grâce aux opérateurs d'interpolation $(\pi_h)_h$ définis en (3.23), avec $V^\star = \mathcal{D}(\Omega)$, où la convergence en norme $\|\cdot\|_{L^2(\Omega)}$ est une conséquence de celle déjà énoncée au §3.4.1 ;
- pour $(\boldsymbol{Q}_h)_h$ dans $\boldsymbol{H}(\operatorname{div};\Omega)$, grâce à un opérateur d'interpolation agissant sur des fonctions à valeurs vectorielles, avec $V^\star = \boldsymbol{C}^\infty(\overline{\Omega})$, nous renvoyons p. 110 de [3].

D'après le théorème 3.15, la convergence suit : $\lim_{h\to 0} \|(u - u_h, \boldsymbol{p} - \boldsymbol{p}_h)\|_V = 0$, ou bien

$$\lim_{h\to 0} \left(\|u - u_h\|_{L^2(\Omega)} + \|\boldsymbol{p} - \boldsymbol{p}_h\|_{\boldsymbol{H}(\operatorname{div};\Omega)}\right) = 0.$$

Concernant la vitesse de convergence, on a le résultat suivant, que nous admettons.

Proposition 3.22 *Supposons que la solution $(u, \boldsymbol{p})$ est de régularité $H^1 \times \boldsymbol{H}^1$ "par morceaux", alors si on utilise les éléments finis de Raviart-Thomas d'ordre $k \geq 0$ sur une famille régulière*

de maillages qui respectent la partition, il existe une constante $C > 0$ indépendante de h telle que

$$\|u - u_h\|_{L^2(\Omega)} + \|\boldsymbol{p} - \boldsymbol{p}_h\|_{\boldsymbol{H}(\mathrm{div}\,;\Omega)} \leq C\,h.$$

Si la solution est plus régulière, la vitesse de convergence est meilleure, si bien sûr on augmente l'ordre de l'élément fini de façon adéquate (voir à nouveau p. 110 de [3]).

Remarque 3.23 *Pour la discrétisation par éléments finis de Raviart-Thomas, nous donnons quelques détails pour réaliser les calculs élémentaires (intégrations sur les simplexes du maillage), en reprenant les notations de la remarque 3.16. Pour un simplexe T_ℓ, si on effectue le changement de variables $F_\ell : \hat{T} \to T_\ell$ avec $\boldsymbol{x} = F_\ell(\hat{\boldsymbol{x}}) = \mathbb{A}_\ell\hat{\boldsymbol{x}} + \boldsymbol{b}_\ell$, alors, à une fonction vectorielle $\hat{\boldsymbol{q}}$ définie sur $\hat{T}$, on associe $\boldsymbol{q}$ définie sur T_ℓ telle que :*

$$\forall \hat{\boldsymbol{x}} \in \hat{T}, \quad \boldsymbol{q}(\boldsymbol{x}) = \frac{1}{|det(\mathbb{A}_\ell)|}\mathbb{A}_\ell\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}}), \quad \mathrm{div}\,\boldsymbol{q}(\boldsymbol{x}) = \frac{1}{|det(\mathbb{A}_\ell)|}\hat{\mathrm{div}}\,\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}}).$$

C'est la transformation de Piola*, utilisée pour transformer des éléments de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ (voir par ex. p. 59 de [3]). Pour les éléments de M_h, on procède comme pour les éléments finis de Lagrange, cf. la remarque 3.16. Le choix de la transformation de Piola pour les éléments de $\boldsymbol{H}(\mathrm{div}\,;\Omega)$ peut-être "justifié" via les changements de variables dans les intégrales, mettant en jeu des éléments de $\boldsymbol{Q}_h$ et M_h :*

$$\begin{aligned}
\int_{T_\ell} v_h(\boldsymbol{x})\mathrm{div}\,\boldsymbol{q}_h(\boldsymbol{x})\,d\boldsymbol{x} &= \int_{\hat{T}} \hat{v}(\hat{\boldsymbol{x}})\frac{1}{|det(\mathbb{A}_\ell)|}\hat{\mathrm{div}}\,\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}})\,|det(\mathbb{A}_\ell)|d\hat{\boldsymbol{x}} = \int_{\hat{T}} \hat{v}(\hat{\boldsymbol{x}})\hat{\mathrm{div}}\,\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}})\,d\hat{\boldsymbol{x}},\\
\int_{T_\ell} \nabla v_h(\boldsymbol{x})\cdot\boldsymbol{q}_h(\boldsymbol{x})\,d\boldsymbol{x} &= \int_{\hat{T}} (\mathbb{A}_\ell^T)^{-1}\hat{\nabla}\hat{v}(\hat{\boldsymbol{x}})\cdot\frac{1}{|det(\mathbb{A}_\ell)|}\mathbb{A}_\ell\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}})\,|det(\mathbb{A}_\ell)|d\hat{\boldsymbol{x}} = \int_{\hat{T}} \hat{\nabla}\hat{v}(\hat{\boldsymbol{x}})\cdot\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}})\,d\hat{\boldsymbol{x}}.
\end{aligned}$$

On peut également montrer que :

$$\int_{\partial T_\ell} v_h(\boldsymbol{x})\boldsymbol{q}_h(\boldsymbol{x})\cdot\boldsymbol{n}\,d\Gamma_{\boldsymbol{x}} = \int_{\partial\hat{T}} \hat{v}(\hat{\boldsymbol{x}})\hat{\boldsymbol{q}}(\hat{\boldsymbol{x}})\cdot\hat{\boldsymbol{n}}\,d\Gamma_{\hat{\boldsymbol{x}}}.$$

Remarque 3.24 *En suivant le formalisme du lemme 3.1, on peut montrer, mais c'est un peu fastidieux (voir le chapitre 3 de [13]), que la matrice associée à la formulation variationnelle discrète (3.25) est creuse.*

3.4.3 Comparaison des discrétisations

Donnons quelques éléments de comparaison entre la discrétisation u_h^L par élément fini de Lagrange d'ordre 1 (utilisée pour la diffusion à une inconnue), et la discrétisation $(u_h^{RT}, \boldsymbol{p}_h)$ par élément fini de Raviart-Thomas d'ordre 0 (utilisée pour la diffusion à deux inconnues), en supposant que l'exposant de décalage r_{max} est égal à 1 (la solution u est donc H^2 "par morceaux"), et que $\boldsymbol{p} = -\mathbb{D}\nabla u$ est $\boldsymbol{H}^1(\Omega)$ "par morceaux" :

— D'après la proposition 3.18, le théorème 3.19 et le corollaire 3.20, on a

$$\|u - u_h^L\|_{H^1(\Omega)} \leq C\,h, \quad \|u - u_h^L\|_{L^2(\Omega)} \leq C\,h^2.$$

— D'après la proposition 3.22, on a

$$\|u - u_h^{RT}\|_{L^2(\Omega)} \leq C\, h, \quad \|\boldsymbol{p} - \boldsymbol{p}_h\|_{\boldsymbol{H}(\mathrm{div};\Omega)} \leq C\, h.$$

En général $u_h^L \neq u_h^{RT}$: en effet, u_h^L est un champ continu par définition, ce qui n'est pas le cas de u_h^{RT}. L'approximation est *a priori* meilleure pour u_h^L que pour u_h^{RT} en norme $\|\cdot\|_{L^2(\Omega)}$; elle est du même ordre en norme $\|\cdot\|_{H^1(\Omega)}$ où, pour l'approximation par élément fini de Raviart-Thomas, on passe par la convergence en norme $\|\cdot\|_{\boldsymbol{L}^2(\Omega)}$ de $\nabla u = -\mathbb{D}^{-1}\boldsymbol{p}$...

Notes de cours 4

Problèmes avec contraintes

On s'intéresse à la résolution de problèmes pour lesquels les solutions doivent vérifier certaines contraintes, qui peuvent être indépendantes (ou pas...) des données. Un exemple classique est le modèle de Stokes, qui modélise l'écoulement stationnaire de fluides incompressibles.

On se place dans un domaine Ω de $\mathbb{R}^{\mathtt{d}}$, pour $2 \leq \mathtt{d} \leq 3$. Comme pour la diffusion, le modèle étudié ici est basé sur des données et des solutions à valeurs réelles. On choisit donc $\mathbb{K} = \mathbb{R}$ lorsqu'on l'étudie. Concernant la théorie "abstraite", on se placera dans $\mathbb{K} = \mathbb{C}$ (pour l'appliquer au modèle, il suffira d'enlever la conjugaison). On note $(\cdot|\cdot)$ le produit scalaire (hermitien) de $L^2(\Omega)$ ou de $(L^2(\Omega))^{\mathtt{d}}$, et $\|\cdot\|$ la norme associée. On pose $\boldsymbol{L}^2(\Omega) = (L^2(\Omega))^{\mathtt{d}}$, et $\boldsymbol{H}^1_0(\Omega) = (H^1_0(\Omega))^{\mathtt{d}}$. Lorsqu'on considére les dérivées d'un champ $\boldsymbol{v}$ à valeurs vectorielles, on note $\boldsymbol{\nabla v}$ le tenseur $(\partial_j v_i)_{i,j=1,\mathtt{d}}$; et pour $\boldsymbol{v}, \boldsymbol{w}$ à valeurs vectorielles, on note $\boldsymbol{\nabla v} : \boldsymbol{\nabla w} = \sum_{i,j} \partial_j v_i \partial_j w_i$.

4.1 Modèle de Stokes

Résoudre ce modèle revient à résoudre les équations et conditions au bord :

$$\left\{ \begin{array}{l} \text{Trouver } (\boldsymbol{u}, p) \text{ telles que} \\ -\nu\, \Delta \boldsymbol{u} + \nabla p = \boldsymbol{f} \text{ dans } \Omega \\ \operatorname{div} \boldsymbol{u} = 0 \text{ dans } \Omega \\ \boldsymbol{u} = 0 \text{ sur } \partial\Omega. \end{array} \right. \tag{4.1}$$

La partie de la solution à valeurs vectorielles $\boldsymbol{u}$ est la *vitesse du fluide*, et la partie à valeurs scalaires p est la *pression*, ν est la *viscosité cinématique*, et la donnée $\boldsymbol{f}$ est la *densité des forces extérieures*. La relation $\operatorname{div} \boldsymbol{u} = 0$ traduit la contrainte d'incompressibilité, et la condition aux limites signifie que la vitesse du fluide est nulle sur les parois délimitant le domaine. Toutes les grandeurs sont à valeurs réelles. On observe que, par construction, $\Delta \boldsymbol{u}$, ∇p et $\boldsymbol{f}$ sont des

champs à valeurs vectorielles dans $\mathbb{R}^{\mathtt{d}}$. On peut donc réécrire (4.1) sous la forme équivalente

$$\left\{\begin{array}{l} \text{Trouver } ((u_i)_{i=1,\mathtt{d}}, p) \text{ telles que} \\ -\nu\,\Delta u_i + \partial_i p = f_i \text{ dans } \Omega, \text{ pour } i = 1, \mathtt{d} \\ \sum_{i=1,\mathtt{d}} \partial_i u_i = 0 \text{ dans } \Omega \\ u_i = 0 \text{ sur } \partial\Omega, \text{ pour } i = 1, \mathtt{d}. \end{array}\right. \tag{4.2}$$

Dans la suite, on fait l'hypothèse suivante :

$$\text{Le coefficient } \nu \text{ est un nombre réel strictement positif.} \tag{4.3}$$

Avant de définir une formulation variationnelle équivalente à ces modèles, on va proposer des espaces fonctionnels auxquels la solution, et la donnée, appartiennent.
Tout d'abord, comme on a une condition aux limites pour chaque composante u_i de la vitesse dans (4.2), on suppose *a priori* que $u_i \in H^1(\Omega)$ pour $i = 1, \mathtt{d}$. Ceci permet de donner un sens à la condition aux limites via l'application trace γ_0, et plus précisément on a $u_i \in H^1_0(\Omega)$ pour $i = 1, \mathtt{d}$. Dans la suite, on munit $H^1_0(\Omega)$ de la semi-norme $|\cdot|_{1,\Omega} : v \mapsto \|\nabla v\| = (\sum_{j=1,\mathtt{d}} \|\partial_j v\|^2)^{1/2}$, voir (1.8), qui définit une norme équivalente à la norme $\|\cdot\|_{H^1(\Omega)}$ d'après l'inégalité de Poincaré. Pour le modèle (4.1), on a donc $\boldsymbol{u} \in \boldsymbol{H}^1_0(\Omega)$, et on munit $\boldsymbol{H}^1_0(\Omega)$ de la semi-norme

$$\boldsymbol{v} \mapsto |\boldsymbol{v}|_{1,\Omega} = \left(\sum_{i=1,\mathtt{d}} |v_i|^2_{1,\Omega}\right)^{1/2},$$

qui définit une norme équivalente à la norme $\|\cdot\|_{\boldsymbol{H}^1(\Omega)}$ d'après l'inégalité de Poincaré.

Puisque $u_i \in H^1_0(\Omega)$, on en déduit que Δu_i appartient à $(H^1_0(\Omega))'$, pour $i = 1, \mathtt{d}$, avec dépendance continue, cf. (1.12). De même, si $p \in L^2(\Omega)$, alors $\partial_i p \in (H^1_0(\Omega))'$ pour $i = 1, \mathtt{d}$, avec dépendance continue, cf. (1.10). Comme on a modifié la norme utilisée sur $H^1_0(\Omega)$, le module de continuité est lui aussi modifié. Néanmoins, on peut vérifier que les bornes respectivement obtenues en (1.12) et (1.10) restent valables. Les résultats sont résumés ci-dessous, et complétés de la transposition aux grandeurs vectorielles.

Proposition 4.1 *Dans un domaine Ω, on a les résultats :*
- *soit $u \in H^1(\Omega)$, alors $-\Delta u \in (H^1_0(\Omega))'$;*
  de plus, on a $-\Delta \in \mathcal{L}\,(H^1(\Omega), (H^1_0(\Omega))')$, avec $|||-\Delta||| \leq 1$;
- *soit $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$, alors $-\Delta \boldsymbol{u} \in (\boldsymbol{H}^1_0(\Omega))'$;*
  de plus, on a $-\Delta \in \mathcal{L}\,\big(\boldsymbol{H}^1(\Omega), (\boldsymbol{H}^1_0(\Omega))'\big)$, avec $|||-\Delta||| \leq 1$;
- *soit $p \in L^2(\Omega)$, alors $\nabla p \in (\boldsymbol{H}^1_0(\Omega))'$ et $\partial_i p \in (H^1_0(\Omega))'$ pour $i = 1, \mathtt{d}$;*
  de plus, on a $\nabla \in \mathcal{L}\,\big(L^2(\Omega), (\boldsymbol{H}^1_0(\Omega))'\big)$, avec $|||\nabla||| \leq \sqrt{\mathtt{d}}$, ainsi que $\partial_i \in \mathcal{L}\,(L^2(\Omega), (H^1_0(\Omega))')$, avec $|||\partial_i||| \leq 1$ pour $i = 1, \mathtt{d}$.

Ainsi, lorsque $u_i \in H^1_0(\Omega)$ pour $i = 1, \mathtt{d}$ et $p \in L^2(\Omega)$, la donnée f_i apparaissant dans la première équation de (4.2) étant égale à $-\nu\,\Delta u_i + \partial_i p$, doit appartenir à $(H^1_0(\Omega))'$. De même, pour la

donnée $\boldsymbol{f}$ apparaissant dans la première équation de (4.1), on la choisit telle que

$$\boldsymbol{f} \in (\boldsymbol{H}_0^1(\Omega))'. \tag{4.4}$$

Par ailleurs, on note que la pression p est connue à une constante près : en effet, p apparaît uniquement via son gradient, or l'équation $\nabla p = 0$ admet pour solution les constantes, puisque le domaine Ω est connexe. Pour lever cette indétermination, on va chercher p dans le sous-espace de $L^2(\Omega)$ orthogonal au sous-espace formé des constantes, c'est-à-dire dans

$$L^2_{vmn}(\Omega) = \{q \in L^2(\Omega) \text{ tel que } (q|1) = 0\}.$$

En d'autres termes, $L^2(\Omega) = L^2_{vmn}(\Omega) \overset{\perp}{\oplus} \mathbb{R}$. On note que $L^2_{vmn}(\Omega)$ est le sous-espace vectoriel de $L^2(\Omega)$ formé des fonctions à moyenne nulle, et $p \in L^2_{vmn}(\Omega)$.
A partir de là, on va résoudre, sous les hypothèses (4.3) et (4.4), le modèle de Stokes "mathématique"

$$\left\{ \begin{array}{l} \text{Trouver } (\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\Omega) \times L^2_{vmn}(\Omega) \text{ telles que} \\ -\nu\, \Delta \boldsymbol{u} + \nabla p = \boldsymbol{f} \text{ dans } (\boldsymbol{H}_0^1(\Omega))' \\ \operatorname{div} \boldsymbol{u} = 0 \text{ dans } \Omega. \end{array} \right. \tag{4.5}$$

On en construit une formulation variationnelle équivalente. Tout d'abord, on déduit facilement des résultats précédents que la première équation de (4.5) est équivalente à

$$\forall \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega), \quad \nu \int_\Omega \boldsymbol{\nabla u} : \boldsymbol{\nabla v}\, d\Omega - \int_\Omega p \operatorname{div} \boldsymbol{v}\, d\Omega = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{(\boldsymbol{H}_0^1(\Omega))', \boldsymbol{H}_0^1(\Omega)}.$$

Il reste à écrire la seconde équation de (4.5) sous forme variationnelle. Comme $\operatorname{div} \boldsymbol{u}$ appartient à $L^2(\Omega)$ puisque toutes ses dérivées partielles premières sont dans $L^2(\Omega)$, on pourrait écrire la formulation variationnelle équivalente

$$\forall q \in L^2(\Omega), \quad \int_\Omega q \operatorname{div} \boldsymbol{u}\, d\Omega = 0.$$

Mais on peut caractériser un peu plus précisément $\operatorname{div} \boldsymbol{u}$: en effet, si on choisit $q = 1$ ci-dessus, on a par intégration par parties (1.20)

$$\int_\Omega \operatorname{div} \boldsymbol{u}\, d\Omega = \langle \gamma_n \boldsymbol{u}, 1 \rangle_{(H^{1/2}(\partial\Omega))', H^{1/2}(\partial\Omega)} = (\boldsymbol{u} \cdot \boldsymbol{n}_{|\partial\Omega}, 1)_{L^2(\partial\Omega)} = 0,$$

puisque $\boldsymbol{u}_{|\partial\Omega} = 0$. Ainsi $\operatorname{div} \boldsymbol{u} \in L^2_{vmn}(\Omega)$, et la formulation variationnelle équivalente à la seconde équation de (4.5) finalement retenue est

$$\forall q \in L^2_{vmn}(\Omega), \quad \int_\Omega q \operatorname{div} \boldsymbol{u}\, d\Omega = 0.$$

Pour résumer, on a montré qu'une formulation variationnelle équivalente à (4.5) est

$$\left\{ \begin{array}{l} \text{Trouver } (\boldsymbol{u}, p) \in \boldsymbol{H}_0^1(\Omega) \times L^2_{vmn}(\Omega) \text{ telles que} \\ \forall (\boldsymbol{v}, q) \in \boldsymbol{H}_0^1(\Omega) \times L^2_{vmn}(\Omega), \\ \qquad \nu \displaystyle\int_\Omega \boldsymbol{\nabla u} : \boldsymbol{\nabla v}\, d\Omega - \int_\Omega p \operatorname{div} \boldsymbol{v}\, d\Omega - \int_\Omega q \operatorname{div} \boldsymbol{u}\, d\Omega = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{(\boldsymbol{H}_0^1(\Omega))', \boldsymbol{H}_0^1(\Omega)}. \end{array} \right. \tag{4.6}$$

Dans la suite, on va également résoudre le modèle de Stokes "généralisé", où la contrainte d'incompressibilité est remplacée par $\operatorname{div} \boldsymbol{u} = g$ dans Ω où, comme on l'a vu plus haut, on choisit une donnée $g \in L^2_{vmn}(\Omega)$. On résout donc

$$\left\{ \begin{array}{l} \text{Trouver } (\boldsymbol{u}, p) \in \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega) \text{ telles que} \\ -\nu\, \Delta \boldsymbol{u} + \nabla p = \boldsymbol{f} \text{ dans } \Omega \\ \operatorname{div} \boldsymbol{u} = g \text{ dans } \Omega. \end{array} \right. \tag{4.7}$$

C'est un modèle "mathématique", dont l'étude se révèlera utile pour la suite. On vérifie facilement qu'une formulation variationnelle équivalente est

$$\left\{ \begin{array}{l} \text{Trouver } (\boldsymbol{u}, p) \in \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega) \text{ telles que} \\ \forall (\boldsymbol{v}, q) \in \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega), \\ \qquad \nu \displaystyle\int_\Omega \boldsymbol{\nabla u} : \boldsymbol{\nabla v}\, d\Omega - \int_\Omega p \operatorname{div} \boldsymbol{v}\, d\Omega - \int_\Omega q \operatorname{div} \boldsymbol{u}\, d\Omega \\ \qquad\qquad = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)} - \displaystyle\int_\Omega q\, g\, d\Omega. \end{array} \right. \tag{4.8}$$

Notons que, si on revient au formalisme de la formulation variationnelle "abstraite" (2.5), on a pour le modèle de Stokes, et pour le modèle de Stokes "généralisé" :

— $\mathcal{V} = \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega)$, muni de la norme $\|(\boldsymbol{v}, q)\|_{\mathcal{V}} = (|\boldsymbol{v}|^2_{1,\Omega} + \|q\|^2)^{1/2}$;

— $a((\boldsymbol{v}, q), (\boldsymbol{w}, r)) = \nu \displaystyle\int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w}\, d\Omega - \int_\Omega q \operatorname{div} \boldsymbol{w}\, d\Omega - \int_\Omega r \operatorname{div} \boldsymbol{v}\, d\Omega$;

— $\langle \ell_{St}, (\boldsymbol{w}, r) \rangle_{\mathcal{V}} = \langle \boldsymbol{f}, \boldsymbol{w} \rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)}$; $\langle \ell_{St-gen}, (\boldsymbol{w}, r) \rangle_{\mathcal{V}} = \langle \boldsymbol{f}, \boldsymbol{w} \rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)} - \displaystyle\int_\Omega r\, g\, d\Omega.$

Or, on sait que $\operatorname{div} \in \mathcal{L}(\boldsymbol{H}^1(\Omega), L^2(\Omega))$, avec $|||\operatorname{div}||| \leq \sqrt{\mathtt{d}}$ (cf. (1.9). On en déduit que la forme a est bilinéaire, continue et symétrique sur $\mathcal{V} \times \mathcal{V}$. Par contre, a n'est pas coercive ! En effet, on a, pour tout $(0, q) \in \mathcal{V}$, $a((0, q), (0, q)) = 0$. Enfin , il est clair que, pour toutes les données $\boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))'$ et $g \in L^2_{vmn}(\Omega)$, on a $\ell_{St}, \ell_{St-gen} \in \mathcal{V}'$.

4.2 Outils mathématiques abstraits

Afin de pouvoir analyser mathématiquement les modèles de Stokes et de Stokes "généralisé", nous allons maintenant énoncer trois théorèmes abstraits, tirés de [2], dont la démonstration sort du cadre de ce cours.

On rappelle qu'on a la décomposition orthogonale $L^2(\Omega) = \mathbb{R} \overset{\perp}{\oplus} L^2_{vmn}(\Omega)$ et que d'après la proposition 4.1, $\nabla \in \mathcal{L}\left(L^2(\Omega), (\boldsymbol{H}^1_0(\Omega))'\right)$.

Théorème 4.2 *Soit Ω un domaine de $\mathbb{R}^{\mathtt{d}}$. Alors,* $\ker(\nabla) = \mathbb{R}$ *et l'image $\nabla[L^2_{vmn}(\Omega)]$ est fermée dans $(\boldsymbol{H}^1_0(\Omega))'$.*

Ainsi, $\nabla[L^2_{vmn}(\Omega)]$ muni de la norme $\|\cdot\|_{(\boldsymbol{H}^1_0(\Omega))'}$ est un espace de Banach. En outre, comme $\nabla \in \mathcal{L}\left(L^2_{vmn}(\Omega), \nabla[L^2_{vmn}(\Omega)]\right)$ est bijective, on en déduit à l'aide du théorème de l'isomorphisme de Banach que $\nabla^{-1} \in \mathcal{L}\left(\nabla[L^2_{vmn}(\Omega)], L^2_{vmn}(\Omega)\right)$. On a donc

$$\exists C_\nabla > 0,\ \forall \boldsymbol{w} \in \nabla[L^2_{vmn}(\Omega)], \quad \|\nabla^{-1} \boldsymbol{w}\| \leq C_\nabla \|\boldsymbol{w}\|_{(\boldsymbol{H}^1_0(\Omega))'}.$$

Or, à tout $\boldsymbol{w} \in \nabla[L^2_{vmn}(\Omega)]$, on peut bien sûr associer $p \in L^2_{vmn}(\Omega)$ unique tel que $\boldsymbol{w} = \nabla p$, ou $p = \nabla^{-1}\boldsymbol{w}$. Ce que l'on résume sous la forme

$$\exists C_\nabla > 0,\ \forall p \in L^2_{vmn}(\Omega), \quad \|p\| \leq C_\nabla \|\nabla p\|_{(\boldsymbol{H}^1_0(\Omega))'}. \tag{4.9}$$

Ensuite, on introduit le sous-espace vectoriel de $\boldsymbol{H}^1_0(\Omega)$ contenant la contrainte

$$\boldsymbol{V}_0(\Omega) = \{\boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega) \text{ tel que } \operatorname{div} \boldsymbol{v} = 0\}.$$

$\boldsymbol{V}_0(\Omega)$ est fermé dans $\boldsymbol{H}^1_0(\Omega)$, car il est égal au noyau de la restriction de l'application (linéaire) continue $\operatorname{div} \in \mathcal{L}\left(\boldsymbol{H}^1_0(\Omega), L^2(\Omega)\right)$ (voir (1.9)). Ainsi, $(\boldsymbol{V}_0(\Omega), |\cdot|_{1,\Omega})$ est un espace de Hilbert. On note $\boldsymbol{V}_0(\Omega)^\perp$ son orthogonal dans $\boldsymbol{H}^1_0(\Omega)$.

Théorème 4.3 *Soit Ω un domaine de $\mathbb{R}^{\mathtt{d}}$. Alors,* $\ker(\operatorname{div}) = \boldsymbol{V}_0(\Omega)$ *et* $\operatorname{div}[\boldsymbol{V}_0(\Omega)^\perp] = L^2_{vmn}(\Omega)$.

L'application linéaire $\operatorname{div} \in \mathcal{L}\left(\boldsymbol{V}_0(\Omega)^\perp, L^2_{vmn}(\Omega)\right)$ est donc bijective et, d'après le théorème de l'isomorphisme de Banach, on en déduit cette fois que

$$\exists C_{\text{div}} > 0,\ \forall p \in L^2_{vmn}(\Omega),\ \ \exists! \boldsymbol{v}_p \in \boldsymbol{V}_0(\Omega)^\perp \text{ tel que } \operatorname{div} \boldsymbol{v}_p = p,\ \text{avec } |\boldsymbol{v}_p|_{1,\Omega} \leq C_{\text{div}}\, \|p\|. \tag{4.10}$$

Enfin, on a le

Théorème 4.4 (de Rham) *Soit Ω un domaine de $\mathbb{R}^{\mathtt{d}}$. Soit $\boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))'$. On a l'équivalence :*

$$\forall \boldsymbol{v}_0 \in \boldsymbol{V}_0(\Omega),\ \ \langle \boldsymbol{f}, \boldsymbol{v}_0\rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)} = 0 \quad \Longleftrightarrow \quad \exists p^\star \in L^2_{vmn}(\Omega),\ \ \boldsymbol{f} = \nabla p^\star \text{ dans } (\boldsymbol{H}^1_0(\Omega))'.$$

A l'aide de ces résultats, nous allons voir comment on peut résoudre le modèle de Stokes (4.5)-(4.6), ainsi que le modèle de Stokes "généralisé" (4.7)-(4.8).

Commençons par le modèle de Stokes (4.5)-(4.6) : soit $\boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))'$. Par définition, on cherche $\boldsymbol{u} \in \boldsymbol{V}_0(\Omega)$. En outre, puisque tout élément de $\boldsymbol{V}_0(\Omega)$ est à divergence nulle, $\boldsymbol{u}$ est en particulier solution de

$$\left\{ \begin{array}{l} \text{Trouver } \boldsymbol{u} \in \boldsymbol{V}_0(\Omega) \text{ telle que} \\ \forall \boldsymbol{v}_0 \in \boldsymbol{V}_0(\Omega), \quad \nu \displaystyle\int_\Omega \boldsymbol{\nabla u} : \boldsymbol{\nabla v}_0 \, d\Omega = \langle \boldsymbol{f}, \boldsymbol{v}_0\rangle_{(\boldsymbol{V}_0(\Omega))', \boldsymbol{V}_0(\Omega)}. \end{array} \right. \tag{4.11}$$

Ici, la donnée $\boldsymbol{f}$ est considérée comme appartenant à $(\boldsymbol{V}_0(\Omega))'$.[1] D'après le théorème de représentation de Riesz, cette formulation variationnelle admet une solution, et une seule, qui dépend continûment de la donnée, car on a $|\boldsymbol{u}|_{1,\Omega} = \nu^{-1}\|\boldsymbol{f}\|_{(\boldsymbol{V}_0(\Omega))'}$. On a donc :

$$\forall \boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))', \quad \exists! \boldsymbol{u} \in \boldsymbol{V}_0(\Omega) \text{ solution de (4.11), avec } |\boldsymbol{u}|_{1,\Omega} \leq \nu^{-1}\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}.$$

1. Puisque $\boldsymbol{V}_0(\Omega) \subset \boldsymbol{H}^1_0(\Omega)$, on a $(\boldsymbol{H}^1_0(\Omega))' \subset (\boldsymbol{V}_0(\Omega))'$. Précisément, pour tout $\boldsymbol{f}$ dans $(\boldsymbol{H}^1_0(\Omega))'$, on a $\boldsymbol{f} \in (\boldsymbol{V}_0(\Omega))'$, et

$$\|\boldsymbol{f}\|_{(\boldsymbol{V}_0(\Omega))'} = \sup_{\boldsymbol{v}_0 \in \boldsymbol{V}_0(\Omega)\setminus\{0\}} \frac{|\langle \boldsymbol{f}, \boldsymbol{v}_0\rangle_{(\boldsymbol{V}_0(\Omega))', \boldsymbol{V}_0(\Omega)}|}{|\boldsymbol{v}_0|_{1,\Omega}} \leq \sup_{\boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega)\setminus\{0\}} \frac{|\langle \boldsymbol{f}, \boldsymbol{v}\rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)}|}{|\boldsymbol{v}|_{1,\Omega}} = \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}.$$

D'après la proposition 4.1, on a $(\boldsymbol{f}+\nu\,\Delta\boldsymbol{u}) \in (\boldsymbol{H}^1_0(\Omega))'$. Et d'après (4.11), $\boldsymbol{f}+\nu\,\Delta\boldsymbol{u}$ est tel que

$$\forall \boldsymbol{v}_0 \in \boldsymbol{V}_0(\Omega), \quad \langle \boldsymbol{f}+\nu\,\Delta\boldsymbol{u}, \boldsymbol{v}_0\rangle_{(\boldsymbol{V}_0(\Omega))',\boldsymbol{V}_0(\Omega)} = 0.$$

Donc, d'après le théorème de de Rham,

$$\exists! p \in L^2_{vmn}(\Omega), \quad \boldsymbol{f}+\nu\,\Delta\boldsymbol{u} = \nabla p \text{ dans } (\boldsymbol{H}^1_0(\Omega))'.$$

Si on rassemble le tout, on dispose donc de $(\boldsymbol{u},p) \in \boldsymbol{V}_0(\Omega)\times L^2_{vmn}(\Omega)$ telles que

$$-\nu\,\Delta\boldsymbol{u}+\nabla p = \boldsymbol{f} \text{ dans } \Omega.$$

Ainsi, $(\boldsymbol{u},p)$ est bien une solution du modèle de Stokes (4.5), avec dépendance continue. En effet, pour $\boldsymbol{u}$, on a vu que $|\boldsymbol{u}|_{1,\Omega} \le \nu^{-1}\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}$. Et, pour p, on a

$$\begin{aligned}
\|p\| &\le C_\nabla \|\nabla p\|_{(\boldsymbol{H}^1_0(\Omega))'} \text{ (cf. (4.9))}\\
&\le C_\nabla\big(\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \nu\,\|\Delta\boldsymbol{u}\|_{(\boldsymbol{H}^1_0(\Omega))'}\big)\\
&\le C_\nabla\big(\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \nu\,|\boldsymbol{u}|_{1,\Omega}\big) \text{ (proposition 4.1)}\\
&\le 2C_\nabla\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}.
\end{aligned}$$

L'unicité de la solution est une conséquence immédiate : si on a deux solutions du modèle de Stokes, leur différence $(\delta\boldsymbol{u},\delta p)$ vérifie ce modèle avec donnée nulle et, en reprenant le raisonnement précédent, on trouve que $\delta\boldsymbol{u}=0$, puis que $\delta p = 0$. On a donc démontré le

Théorème 4.5 *Soit $\nu>0$. Le modèle de Stokes (4.5)-(4.6) est bien posé :*

$$\forall \boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))', \begin{cases} \exists!(\boldsymbol{u},p) \in \boldsymbol{H}^1_0(\Omega)\times L^2_{vmn}(\Omega) \text{ sol. de (4.5)-(4.6)},\\ \text{avec } |\boldsymbol{u}|_{1,\Omega} \le \nu^{-1}\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} \text{ et } \|p\| \le 2C_\nabla\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}. \end{cases} \tag{4.12}$$

Pour le modèle de Stokes "généralisé" (4.7)-(4.8), la démarche est de s'appuyer sur le théorème 4.3 pour décomposer la solution $\boldsymbol{u}$ selon $\boldsymbol{H}^1_0(\Omega) = \boldsymbol{V}_0(\Omega) \overset{\perp}{\oplus} \boldsymbol{V}_0(\Omega)^\perp$, et de reprendre les résultats obtenus pour le modèle de Stokes. Soient donc $(\boldsymbol{f},g) \in (\boldsymbol{H}^1_0(\Omega))'\times L^2_{vmn}(\Omega)$.
On décompose orthogonalement la composante $\boldsymbol{u}$ de la solution en $\boldsymbol{u}=\boldsymbol{u}_0+\boldsymbol{u}_\perp$, avec $(\boldsymbol{u}_0,\boldsymbol{u}_\perp) \in \boldsymbol{V}_0(\Omega)\times\boldsymbol{V}_0(\Omega)^\perp$. On reformule (4.7) en tenant compte de cette décomposition :

$$\begin{cases} \text{Trouver } (\boldsymbol{u}_0,\boldsymbol{u}_\perp,p) \in \boldsymbol{V}_0(\Omega)\times\boldsymbol{V}_0(\Omega)^\perp\times L^2_{vmn}(\Omega) \text{ telles que}\\ -\nu\,\Delta\boldsymbol{u}_0 - \nu\,\Delta\boldsymbol{u}_\perp + \nabla p = \boldsymbol{f} \text{ dans } \Omega\\ \operatorname{div}\boldsymbol{u}_\perp = g \text{ dans } \Omega. \end{cases} \tag{4.13}$$

Or, d'après le théorème 4.3, on sait que

$$\exists!\boldsymbol{v}_g \in \boldsymbol{V}_0(\Omega)^\perp \text{ tel que } \operatorname{div}\boldsymbol{v}_g = g, \text{ avec } |\boldsymbol{v}_g|_{1,\Omega} \le C_{\rm div}\,\|g\|.$$

Pour résoudre (4.13), on doit donc choisir $\boldsymbol{u}_\perp = \boldsymbol{v}_g$ (qui dépend continûment de la donnée g). A partir de là, on remarque que $(\boldsymbol{u}_0,p)$ résout

$$\begin{cases} \text{Trouver } (\boldsymbol{u}_0,p) \in \boldsymbol{V}_0(\Omega)\times L^2_{vmn}(\Omega) \text{ telles que}\\ -\nu\,\Delta\boldsymbol{u}_0 + \nabla p = \boldsymbol{f} + \nu\,\Delta\boldsymbol{u}_\perp \text{ dans } \Omega. \end{cases}$$

C'est donc une solution du modèle de Stokes (4.5), avec la donnée $\boldsymbol{f}' = \boldsymbol{f} + \nu\,\Delta\boldsymbol{u}_\perp \in (\boldsymbol{H}^1_0(\Omega))'$. On peut alors utiliser le théorème 4.5 pour caractériser $(\boldsymbol{u}_0, p)$ et conclure à l'existence d'une solution de (4.7)-(4.8). On a bien dépendance continue, car la donnée $\boldsymbol{f}'$ est bornée par :

$$\begin{aligned}
\|\boldsymbol{f}'\|_{(\boldsymbol{H}^1_0(\Omega))'} &\leq \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \nu\,\|\Delta\boldsymbol{u}_\perp\|_{(\boldsymbol{H}^1_0(\Omega))'} \\
&\leq \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \nu\,|\boldsymbol{u}_\perp|_{1,\Omega} \text{ (proposition 4.1)} \\
&\leq \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \nu\,C_{\rm div}\,\|g\|.
\end{aligned}$$

On a donc les dépendances continues de la solution, par composante,

$$\begin{aligned}
|\boldsymbol{u}_\perp|_{1,\Omega} &\leq C_{\rm div}\,\|g\|\,; \\
|\boldsymbol{u}_0|_{1,\Omega} &\leq \nu^{-1}\|\boldsymbol{f}'\|_{(\boldsymbol{H}^1_0(\Omega))'} \leq \nu^{-1}\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + C_{\rm div}\,\|g\|\,; \\
\|p\| &\leq 2C_\nabla\|\boldsymbol{f}'\|_{(\boldsymbol{H}^1_0(\Omega))'} \leq 2C_\nabla\big(\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \nu\,C_{\rm div}\,\|g\|\big).
\end{aligned}$$

L'unicité de la solution suit... On a donc démontré le

Théorème 4.6 *Soit $\nu > 0$. Le modèle de Stokes "généralisé" (4.7)-(4.8) est bien posé :*

$$\forall(\boldsymbol{f}, g) \in (\boldsymbol{H}^1_0(\Omega))' \times L^2_{vmn}(\Omega),\ \left\{\begin{array}{l} \exists!(\boldsymbol{u}, p) \in \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega) \text{ sol. de (4.7)-(4.8)}, \\ \text{avec} \quad |\boldsymbol{u}|_{1,\Omega} \leq \nu^{-1}\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \sqrt{3}C_{\rm div}\,\|g\| \text{ et} \\ \qquad\quad \|p\| \leq 2C_\nabla\big(\|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'} + \nu\,C_{\rm div}\,\|g\|\big). \end{array}\right. \tag{4.14}$$

Remarque 4.7 *Lorsque $g = 0$, l'estimation (4.14) est identique à (4.12).*

La difficulté, si on veut approcher les formulations variationnelles, est la prise en compte de la contrainte sur la divergence de la vitesse. En d'autres termes, comment construire une approximation conforme dans $\boldsymbol{V}_0(\Omega)$, c'est-à-dire avec des champs discrets à valeurs vectorielles appartenant à $\boldsymbol{H}^1_0(\Omega)$, et *à divergence nulle* ? Dans la suite, on va décrire une technique de pénalisation, qui permet de prendre en compte cette contrainte en pénalisant les champs à divergence non-nulle. On développera d'autres approches au §5.

4.3 Approximation par pénalisation

On s'intéresse ici à l'approximation du modèle de Stokes (4.5)-(4.6), avec une viscosité cinématique $\nu > 0$ donnée [2]. On va décrire une première méthode d'approximation, qui prend en compte la contrainte sur la divergence de la vitesse par pénalisation.

2. On peut reprendre tous les calculs de la section pour établir une dépendance explicite des erreurs en fonction de ν.

4.3.1 Pénalisation abstraite

Pour $\varepsilon > 0$, on considère la formulation variationnelle "pénalisée"

$$\left\{\begin{array}{l} \text{Trouver } \boldsymbol{u}_\varepsilon \in \boldsymbol{H}^1_0(\Omega) \text{ telle que} \\ \forall \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega), \quad \nu \displaystyle\int_\Omega \boldsymbol{\nabla u}_\varepsilon : \boldsymbol{\nabla v}\, d\Omega + \frac{1}{\varepsilon}\int_\Omega \operatorname{div} \boldsymbol{u}_\varepsilon \operatorname{div} \boldsymbol{v}\, d\Omega = \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{(\boldsymbol{H}^1_0(\Omega))',\boldsymbol{H}^1_0(\Omega)}. \end{array}\right. \tag{4.15}$$

L'idée étant que $\frac{1}{\varepsilon}$ soit "grand", on choisit par exemple $\varepsilon \in]0,1]$. A partir de là, on observe que la forme bilinéaire définie sur $\boldsymbol{H}^1_0(\Omega) \times \boldsymbol{H}^1_0(\Omega)$:

$$a_\varepsilon \ : \ (\boldsymbol{v},\boldsymbol{w}) \mapsto \nu \int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w}\, d\Omega + \frac{1}{\varepsilon}\int_\Omega \operatorname{div} \boldsymbol{v} \operatorname{div} \boldsymbol{w}\, d\Omega \tag{4.16}$$

est continue (voir (1.9)), coercive et symétrique. D'après le théorème de Lax-Milgram,

$$\begin{array}{l} \forall \varepsilon > 0,\ \exists C_\varepsilon > 0,\ \forall \boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))', \\ \exists ! \boldsymbol{u}_\varepsilon \text{ solution de (4.15) telle que } |\boldsymbol{u}_\varepsilon|_{1,\Omega} \le C_\varepsilon \, \|\boldsymbol{f}\|_{\boldsymbol{H}^1_0(\Omega))'}. \end{array}$$

On peut "comparer" la solution $\boldsymbol{u}_\varepsilon$ à la solution $(\boldsymbol{u}, p)$ du modèle de Stokes (4.5)-(4.6).

Théorème 4.8 *La solution $\boldsymbol{u}_\varepsilon$ de (4.15) est telle que*

$$\exists C_\nu > 0,\ \forall \varepsilon \in]0,1],\ \forall \boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))', \quad |\boldsymbol{u} - \boldsymbol{u}_\varepsilon|_{1,\Omega} + \|p + \varepsilon^{-1}\operatorname{div} \boldsymbol{u}_\varepsilon\| \le C_\nu\, \varepsilon^{1/2} \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}.$$

Démonstration : 1ère étape : évaluer le comportement de $\|\operatorname{div} \boldsymbol{u}_\varepsilon\|$. On choisit $\boldsymbol{v} = \boldsymbol{u}_\varepsilon$ dans (4.15) :

$$\nu\, |\boldsymbol{u}_\varepsilon|^2_{1,\Omega} + \varepsilon^{-1} \|\operatorname{div} \boldsymbol{u}_\varepsilon\|^2 = \langle \boldsymbol{f}, \boldsymbol{u}_\varepsilon\rangle_{(\boldsymbol{H}^1_0(\Omega))',\boldsymbol{H}^1_0(\Omega)}. \tag{4.17}$$

Dans (4.17), on observe que chaque terme de gauche est positif. En particulier, on peut les majorer séparément par celui de droite.
Tout d'abord,

$$\nu\, |\boldsymbol{u}_\varepsilon|^2_{1,\Omega} \le \langle \boldsymbol{f}, \boldsymbol{u}_\varepsilon\rangle_{(\boldsymbol{H}^1_0(\Omega))',\boldsymbol{H}^1_0(\Omega)} \le \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}\, |\boldsymbol{u}_\varepsilon|_{1,\Omega},$$

et ainsi $|\boldsymbol{u}_\varepsilon|_{1,\Omega} \le \nu^{-1} \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}$. Puis, si on reprend (4.17), il suit

$$\|\operatorname{div} \boldsymbol{u}_\varepsilon\|^2 \le \varepsilon \langle \boldsymbol{f}, \boldsymbol{u}_\varepsilon\rangle_{(\boldsymbol{H}^1_0(\Omega))',\boldsymbol{H}^1_0(\Omega)} \le \varepsilon \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}\, |\boldsymbol{u}_\varepsilon|_{1,\Omega} \le \varepsilon \nu^{-1} \|\boldsymbol{f}\|^2_{(\boldsymbol{H}^1_0(\Omega))'},$$

et donc $\|\operatorname{div} \boldsymbol{u}_\varepsilon\| \le \varepsilon^{1/2} \nu^{-1/2} \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}$.
2ème étape : on pose $\boldsymbol{u}'_\varepsilon = \boldsymbol{u} - \boldsymbol{u}_\varepsilon \in \boldsymbol{H}^1_0(\Omega)$, $p'_\varepsilon = p + \varepsilon^{-1} \operatorname{div} \boldsymbol{u}_\varepsilon \in L^2_{vmn}(\Omega)$ et $g_\varepsilon = -\operatorname{div} \boldsymbol{u}_\varepsilon \in L^2_{vmn}(\Omega)$. A l'aide de (4.6) et (4.15), on en déduit que $(\boldsymbol{u}'_\varepsilon, p'_\varepsilon)$ est solution de

$$\left\{\begin{array}{l} \text{Trouver } (\boldsymbol{u}'_\varepsilon, p'_\varepsilon) \in \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega) \text{ telles que} \\ \forall (\boldsymbol{v}, q) \in \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega), \\ \qquad \nu \displaystyle\int_\Omega \boldsymbol{\nabla u}'_\varepsilon : \boldsymbol{\nabla v}\, d\Omega - \int_\Omega p'_\varepsilon \operatorname{div} \boldsymbol{v}\, d\Omega - \int_\Omega q \operatorname{div} \boldsymbol{u}'_\varepsilon\, d\Omega = -\int_\Omega q\, g'_\varepsilon\, d\Omega. \end{array}\right. \tag{4.18}$$

C'est exactement la formulation variationnelle (4.8) du modèle de Stokes "généralisé", avec la donnée $(0, g_\varepsilon)$. Puisque $\|g_\varepsilon\| \le \varepsilon^{1/2} \nu^{-1/2} \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'}$, on déduit finalement du théorème 4.6 qu'il existe $C_\nu > 0$ indépendante de $\varepsilon > 0$ et telle que

$$|\boldsymbol{u}'_\varepsilon|_{1,\Omega} + \|p'_\varepsilon\| \le C_\nu\, \varepsilon^{1/2} \|\boldsymbol{f}\|_{(\boldsymbol{H}^1_0(\Omega))'},$$

ce qui est la conclusion annoncée. ∎

4.3.2 Discrétisation de la formulation variationnelle "pénalisée"

Le but est de calculer une approximation de la solution $(\boldsymbol{u}, p)$ du modèle de Stokes avec $\boldsymbol{f} \in (\boldsymbol{H}^1_0(\Omega))'$ donnée : on passe par la discrétisation de la formulation variationnelle "pénalisée". Au §3.4.1, nous avons déjà introduit $(V^0_h)_h$, une suite de sous-espaces discrets de $H^1_0(\Omega)$. Pour l'adapter à notre cas, il suffit de considèrer $(\boldsymbol{V}^0_h)_h$ une suite de sous-espaces discrets de $\boldsymbol{H}^1_0(\Omega)$, en posant $\boldsymbol{V}^0_h = (V^0_h)^{\mathtt{d}}$.

Pour ε et h donnés, la formulation variationnelle discrète associée à (4.15) est :

$$\left\{ \begin{array}{l} \text{Trouver } \boldsymbol{u}_{\varepsilon,h} \in \boldsymbol{V}^0_h \text{ telle que} \\ \forall \boldsymbol{v}_h \in \boldsymbol{V}^0_h, \quad a_\varepsilon(\boldsymbol{u}_{\varepsilon,h}, \boldsymbol{v}_h) = \langle \boldsymbol{f}, \boldsymbol{v}_h \rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)} \end{array} \right. . \tag{4.19}$$

où la forme bilinéaire a_ε est donnée par (4.16).

On utilise la méthode des éléments finis, dans un domaine Ω supposé polygonal si $\mathtt{d} = 2$, ou polyédrique si $\mathtt{d} = 3$. Le paramètre de discrétisation est égal au pas du maillage, noté h. Pour construire les espaces discrets $(\boldsymbol{V}^0_h)_h$, on s'appuie sur l'élément fini de Lagrange d'ordre $k \geq 1$ pour construire $(V^0_h)_h$, et sur une famille de maillages régulière : on parle d'*éléments finis de Lagrange vectoriels d'ordre k*.
Ci-dessous, on propose l'analyse numérique de la méthode sous forme d'exercice.

Exercice : Tout d'abord, *on laisse ε et h varier.*

1. Montrer que, pour tout $\varepsilon \in]0,1]$ et tout $\boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{H}^1_0(\Omega)$,

$$|a_\varepsilon(\boldsymbol{v}, \boldsymbol{w})| \leq (\nu + \mathtt{d})\varepsilon^{-1} |\boldsymbol{v}|_{1,\Omega} |\boldsymbol{w}|_{1,\Omega}.$$

2. A l'aide du lemme de Céa, en déduire que

$$\exists C_1 > 0,\ \forall \varepsilon \in]0,1],\ \forall h, \quad |\boldsymbol{u}_\varepsilon - \boldsymbol{u}_{\varepsilon,h}|_{1,\Omega} \leq C_1 \varepsilon^{-1/2} \inf_{\boldsymbol{v}_h \in \boldsymbol{V}^0_h} |\boldsymbol{u}_\varepsilon - \boldsymbol{v}_h|_{1,\Omega}.$$

3. En utilisant l'inégalité triangulaire, en conclure que

$$\exists C_2 > 0,\ \forall \varepsilon \in]0,1],\ \forall h, \quad |\boldsymbol{u} - \boldsymbol{u}_{\varepsilon,h}|_{1,\Omega} \leq C_2 \left(\varepsilon^{1/2} + \varepsilon^{-1/2} \inf_{\boldsymbol{v}_h \in \boldsymbol{V}^0_h} |\boldsymbol{u}_\varepsilon - \boldsymbol{v}_h|_{1,\Omega} \right).$$

Supposons que :
- pour tout $\varepsilon > 0$ la solution $\boldsymbol{u}_\varepsilon$ de la formulation variationnelle "pénalisée" (4.15) est de régularité $\boldsymbol{H}^2$ "par morceaux" sur une partition de Ω ;
- la norme $\boldsymbol{H}^2$ "par morceaux" de la famille de solutions $(\boldsymbol{u}_\varepsilon)_\varepsilon$ est uniformément bornée.

4. Sous ces conditions, rappeler pourquoi, si on utilise les éléments finis de Lagrange d'ordre k sur une famille régulière de maillages qui respectent la partition,

$$\exists C_3 > 0,\ \forall \varepsilon \in]0,1],\ \forall h, \quad \inf_{\boldsymbol{v}_h \in \boldsymbol{V}_h^0} |\boldsymbol{u}_\varepsilon - \boldsymbol{v}_h|_{1,\Omega} \leq C_3\, h^{\min(1,k)}.$$

   Puisque $\min(1,k) = 1$, on choisit l'élément fini de Lagrange d'ordre $k = 1$ dans l'estimation de la vitesse de convergence. L'idée est maintenant de *choisir ε en fonction de h, afin d'assurer la "meilleure" convergence possible.*

5. Concernant la vitesse... Pour $h > 0$, on note $\widetilde{\boldsymbol{u}_h} = \boldsymbol{u}_{\varepsilon,h}$ la solution de la formulation variationnelle discrète (4.19) pour $\varepsilon = h$. Montrer qu'on a l'erreur entre $\boldsymbol{u}$ et $\widetilde{\boldsymbol{u}_h}$

$$\exists C_4 > 0,\ \forall h, \quad |\boldsymbol{u} - \widetilde{\boldsymbol{u}_h}|_{1,\Omega} \leq C_4\, h^{1/2}.$$

6. Concernant la pression... Soit l'inconnue discrète auxiliaire $p_{\varepsilon,h} = -\varepsilon^{-1}\mathrm{div}\,\boldsymbol{u}_{\varepsilon,h}$. Montrer que

$$\exists C_5 > 0,\ \forall \varepsilon \in]0,1],\ \forall h, \quad \|p - p_{\varepsilon,h}\| \leq C_5 \left(\varepsilon^{1/2} + \varepsilon^{-1}\, |\boldsymbol{u}_\varepsilon - \boldsymbol{u}_{\varepsilon,h}|_{1,\Omega}\right).$$

   Pour $h > 0$, on note $\widetilde{p_h} = -\varepsilon^{-1}\mathrm{div}\,\boldsymbol{u}_{\varepsilon,h}$, où $\boldsymbol{u}_{\varepsilon,h}$ est la solution de la formulation variationnelle discrète (4.19) pour $\varepsilon = h^{1/2}$. Montrer qu'on a l'erreur entre p et $\widetilde{p_h}$

$$\exists C_6 > 0,\ \forall h, \quad \|p - \widetilde{p_h}\| \leq C_6\, h^{1/4}.$$

7. Que pensez-vous des erreurs obtenues ? La méthode numérique est-elle intéressante ?

Notes de cours 5

Problèmes avec contraintes - Compléments

On s'intéresse encore à la résolution des modèles de Stokes (4.5)-(4.6), et de Stokes "généralisé" (4.7)-(4.8), à l'aide de techniques différentes, mais complémentaires, de celles proposées au §4. En particulier, la condition sur la divergence reste explicite dans la formulation variationnelle étudiée. On étudie donc la formulation variationnelle "abstraite"

$$\left\{ \begin{array}{l} \text{Trouver } u \in \mathcal{V} \text{ tel que} \\ \forall w \in \mathcal{V}, \quad a(u,w) = \langle \ell, w \rangle_{\mathcal{V}}, \end{array} \right. \tag{5.1}$$

avec :

— $\mathcal{V} = \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega)$, muni de la norme $\|(\boldsymbol{v},q)\|_{\mathcal{V}} = (|\boldsymbol{v}|^2_{1,\Omega} + \|q\|^2)^{1/2}$;

— $a((\boldsymbol{v},q),(\boldsymbol{w},r)) = \nu \displaystyle\int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w}\, d\Omega - \int_\Omega q \operatorname{div} \boldsymbol{w}\, d\Omega - \int_\Omega r \operatorname{div} \boldsymbol{v}\, d\Omega$;

— $\langle \ell, (\boldsymbol{w},r) \rangle_{\mathcal{V}} = \langle \boldsymbol{f}, \boldsymbol{w} \rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)} - \displaystyle\int_\Omega r\, g\, d\Omega$ ($g = 0$ pour Stokes).

On conserve les notations et le formalisme du §4.

5.1 Approche tout-en-un

On raisonne dans l'espace $\mathcal{V}$, c'est-à-dire que la contrainte est placée au même niveau que l'équation. Voir le §5.2 pour une approche alternative. En préambule, on rappelle que la forme a définie sur $\mathcal{V} \times \mathcal{V}$ *n'est pas coercive*. En effet,

$$\forall q \in L^2_{vmn}(\Omega), \quad a((0,q),(0,q)) = 0.$$

On passe donc par une condition inf-sup, ou de stabilité.

5.1.1 Condition inf-sup abstraite

Le but est ici d'établir une *condition inf-sup* (cf. (2.13) avec $W = V = \mathcal{V}$) pour la forme a. On sait déjà qu'il en existe une, d'après le caractère bien posé de la formulation variationnelle, établi

au théorème 4.6. Ci-dessous, on souhaite en construire une explicitement, via une construction que l'on pourra "adapter" à la formulation variationnelle discrétisée. Pour $(\boldsymbol{v}, q) \in \mathcal{V}$ non-nul, nous allons chercher $(\boldsymbol{w}^\star, r^\star) \in \mathcal{V}$ non-nul, dépendant linéairement de $(\boldsymbol{v}, q)$, et tel que

$$|a((\boldsymbol{v}, q), (\boldsymbol{w}^\star, r^\star))| \geq \alpha' \|(\boldsymbol{v}, q)\|_{\mathcal{V}} \, \|(\boldsymbol{w}^\star, r^\star)\|_{\mathcal{V}},$$

avec une constante $\alpha' > 0$ indépendante de $(\boldsymbol{v}, q)$. On raisonne par étapes. L'enchainement est "naturel", mais technique...

1. **Cas particulier $q = 0$.** On a $\|(\boldsymbol{v}, 0)\|_{\mathcal{V}} = |\boldsymbol{v}|_{1,\Omega}$, et

$$a((\boldsymbol{v}, 0), (\boldsymbol{w}, r)) = \nu \int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w} \, d\Omega - \int_\Omega r \, \mathrm{div}\, \boldsymbol{v} \, d\Omega.$$

   Si on choisit $\boldsymbol{w}^\star = \boldsymbol{v}$ et $r^\star = 0$, on a $a((\boldsymbol{v}, 0), (\boldsymbol{w}^\star, r^\star)) = \nu \, |\boldsymbol{v}|^2_{1,\Omega}$, et $\|(\boldsymbol{w}^\star, r^\star)\|_{\mathcal{V}} = |\boldsymbol{v}|_{1,\Omega}$. Il suit

$$|a((\boldsymbol{v}, 0), (\boldsymbol{w}^\star, r^\star))| = \nu \, \|(\boldsymbol{v}, 0)\|_{\mathcal{V}} \|(\boldsymbol{w}^\star, r^\star)\|_{\mathcal{V}}.$$

2. **Cas particulier $\boldsymbol{v} = 0$.** On a $\|(0, q)\|_{\mathcal{V}} = \|q\|$, et

$$a((0, q), (\boldsymbol{w}, r)) = -\int_\Omega q \, \mathrm{div}\, \boldsymbol{w} \, d\Omega.$$

   D'après le théorème 4.3, on sait qu'il existe $\boldsymbol{w}_q \in \boldsymbol{H}^1_0(\Omega)$ tel que $\mathrm{div}\, \boldsymbol{w}_q = q$, et $|\boldsymbol{w}_q|_{1,\Omega} \leq C_{\mathrm{div}} \, \|q\|$, avec $C_{\mathrm{div}} > 0$ indépendante de q. Si on choisit $\boldsymbol{w}^\star = -\boldsymbol{w}_q$ et $r^\star = 0$, on a $a((0, q), (\boldsymbol{w}^\star, r^\star)) = \|q\|^2$, et $\|(\boldsymbol{w}^\star, r^\star)\|_{\mathcal{V}} = |\boldsymbol{w}_q|_{1,\Omega}$. Il suit

$$|a((0, q), (\boldsymbol{w}^\star, r^\star))| \geq \|q\| \, \frac{|\boldsymbol{w}_q|_{1,\Omega}}{C_{\mathrm{div}}} = \frac{1}{C_{\mathrm{div}}} \|(0, q)\|_{\mathcal{V}} \, \|(\boldsymbol{w}^\star, r^\star)\|_{\mathcal{V}}.$$

3. **Cas général**. On cherche $\boldsymbol{w}^\star$ sous la forme $\boldsymbol{w}^\star = \alpha \boldsymbol{v} - \beta \boldsymbol{w}_q$, avec $\alpha, \beta \in \mathbb{R}$. On a

$$\begin{aligned} a((\boldsymbol{v}, q), (\boldsymbol{w}^\star, r)) \;\; &= \;\; \alpha\nu \, |\boldsymbol{v}|^2_{1,\Omega} - \beta\nu \int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w}_q \, d\Omega \\ &\qquad -\alpha \int_\Omega q \, \mathrm{div}\, \boldsymbol{v} \, d\Omega + \beta \|q\|^2 - \int_\Omega r \, \mathrm{div}\, \boldsymbol{v} \, d\Omega. \end{aligned}$$

   D'après les premier et quatrième termes, on impose d'abord $\alpha > 0$ et $\beta > 0$. Ensuite, on choisit $r^\star = -\alpha q$ pour annuler la somme des troisième et cinquième termes. Avec ce choix pour $(\boldsymbol{w}^\star, r^\star)$, on obtient

$$a((\boldsymbol{v}, q), (\boldsymbol{w}^\star, r^\star)) = \alpha\nu \, |\boldsymbol{v}|^2_{1,\Omega} + \beta \|q\|^2 - \beta\nu \int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w}_q \, d\Omega.$$

   Il reste maintenant à contrôler le dernier terme à l'aide des autres. Pour cela, on utilise l'inégalité de Young [1], avec $\eta > 0$ quelconque :

$$\begin{aligned} -\beta\nu \int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w}_q \, d\Omega \;\; &\geq \;\; -\frac{\beta\nu\eta}{2} |\boldsymbol{v}|^2_{1,\Omega} - \frac{\beta\nu}{2\eta} |\boldsymbol{w}_q|^2_{1,\Omega} \\ &\geq \;\; -\frac{\beta\nu\eta}{2} |\boldsymbol{v}|^2_{1,\Omega} - \frac{\beta\nu (C_{\mathrm{div}})^2}{2\eta} \|q\|^2. \end{aligned}$$

1. L'inégalité de Young s'écrit : $\forall x, y \in \mathbb{R}$, $\forall \eta > 0$, $2xy \leq \eta \, x^2 + \eta^{-1} y^2$, en développant $(\eta^{1/2} x - \eta^{-1/2} y)^2 \geq 0$.

D'où les minorations

$$\begin{aligned} a((\boldsymbol{v},q),(\boldsymbol{w}^\star,r^\star)) &\geq (\alpha\nu - \frac{\beta\nu\eta}{2})|\boldsymbol{v}|^2_{1,\Omega} + \beta(1 - \frac{\nu(C_{\rm div})^2}{2\eta})\|q\|^2 \\ (\text{choix : } \eta = \frac{\alpha}{\beta} > 0) &\geq \frac{\alpha\nu}{2}|\boldsymbol{v}|^2_{1,\Omega} + \beta(1 - \frac{\beta\nu(C_{\rm div})^2}{2\alpha})\|q\|^2 \\ (\text{choix : } \alpha = \beta\nu(C_{\rm div})^2) &\geq \frac{\beta\nu^2(C_{\rm div})^2}{2}|\boldsymbol{v}|^2_{1,\Omega} + \frac{\beta}{2}\|q\|^2 \\ (\text{choix : } \beta = 1) &\geq \frac{\nu^2(C_{\rm div})^2}{2}|\boldsymbol{v}|^2_{1,\Omega} + \frac{1}{2}\|q\|^2. \end{aligned}$$

D'où finalement, si on pose $\gamma = \frac{1}{2}\min(\nu^2(C_{\rm div})^2, 1) > 0$:

$$a((\boldsymbol{v},q),(\boldsymbol{w}^\star,r^\star)) \geq \gamma\|(\boldsymbol{v},q)\|^2_{\mathcal{V}}, \tag{5.2}$$

où $\boldsymbol{w}^\star = \nu(C_{\rm div})^2\boldsymbol{v} - \boldsymbol{w}_q$ et $r^\star = -\nu(C_{\rm div})^2 q$. On note que [2]

$$|\boldsymbol{w}^\star|_{1,\Omega} \leq \nu(C_{\rm div})^2|\boldsymbol{v}|_{1,\Omega} + C_{\rm div}\ \|q\|, \quad \|r^\star\| = \nu(C_{\rm div})^2\|q\|.$$

D'où

$$\begin{aligned} \|(\boldsymbol{w}^\star,r^\star)\|^2_{\mathcal{V}} &\leq 2\nu^2(C_{\rm div})^4|\boldsymbol{v}|^2_{1,\Omega} + 2(C_{\rm div})^2\,\|q\|^2 + \nu^2(C_{\rm div})^4\|q\|^2 \\ &\leq (\gamma')^2\|(\boldsymbol{v},q)\|^2_{\mathcal{V}}, \end{aligned}$$

où on a posé $\gamma' = \max\left(\sqrt{2}\nu(C_{\rm div})^2, C_{\rm div}\,[2 + \nu^2(C_{\rm div})^2]^{1/2}\right) > 0$.
Dans le cas général, on a donc prouvé que

$$|a((\boldsymbol{v},q),(\boldsymbol{w}^\star,r^\star))| \geq \frac{\gamma}{\gamma'}\|(v,\boldsymbol{q})\|_{\mathcal{V}}\,\|(w^\star,\boldsymbol{r}^\star)\|_{\mathcal{V}}.$$

On en déduit une condition inf-sup (cf. (2.13) avec $W = V = \mathcal{V}$) pour la forme a.

Proposition 5.1 *La forme a apparaissant dans les formulations variationnelles de Stokes (4.6) et de Stokes "généralisé" (4.8) vérifie une condition inf-sup :*

$$\forall(\boldsymbol{v},q) \in \mathcal{V}, \qquad \sup_{(\boldsymbol{w},r)\in\mathcal{V}\setminus\{0\}} \frac{|a((\boldsymbol{v},q),(\boldsymbol{w},r))|}{\|(\boldsymbol{w},r)\|_{\mathcal{V}}} \geq \alpha'_{St}\,\|(\boldsymbol{v},q)\|_{\mathcal{V}},$$

où on a posé

$$\alpha'_{St} = \frac{\frac{1}{2}\min(\nu^2(C_{\rm div})^2, 1)}{\max\left(\sqrt{2}\nu(C_{\rm div})^2, C_{\rm div}\,[2 + \nu^2(C_{\rm div})^2]^{1/2}\right)} > 0.$$

Remarque 5.2 *Comme indiqué en préambule, cette condition inf-sup permet de recouvrer le caractère bien posé de la formulation variationnelle en appliquant le corollaire BNB (corollaire 2.16), puisque la forme a est symétrique.*

2. Si on n'est pas intéressé par une majoration explicite, on note que l'application linéaire $(\boldsymbol{v},q) \mapsto (\boldsymbol{w}^\star,r^\star)$ ainsi construite, est continue de $\mathcal{V} = \boldsymbol{H}^1_0(\Omega) \times L^2_{vmn}(\Omega)$ dans lui-même. En d'autres termes

$$\exists C_T > 0,\ \forall(\boldsymbol{v},q) \in \mathcal{V}, \quad \|(\boldsymbol{w}^\star,r^\star)\|_{\mathcal{V}} \leq C_T\|(\boldsymbol{v},q)\|_{\mathcal{V}}.$$

5.1.2 Utilisation de la T-coercivité

Comme on l'a vu au §2.3.4 pour la diffusion, on peut également raisonner à l'aide de la T-coercivité. En effet, l'application linéaire $\mathtt{T}$ définie par

$$\mathtt{T}((\boldsymbol{v},q)) = (\boldsymbol{w}^\star, r^\star) = (\nu(C_{\rm div})^2\boldsymbol{v} - \boldsymbol{w}_q, -\nu(C_{\rm div})^2 q),$$

appartient à $\mathcal{L}(\mathcal{V})$, et vérifie $|||\mathtt{T}||| \leq \gamma'$. De plus, elle est bijective :

— $\mathtt{T}((\boldsymbol{v},q)) = (0,0)$ signifie que $\nu(C_{\rm div})^2\boldsymbol{v} - \boldsymbol{w}_q = 0$ et $q = 0$, c'est-à-dire que $(\boldsymbol{v},q) = (0,0)$, d'où l'injectivité ;

— pour $(\boldsymbol{w},r) \in \mathcal{V}$, si on choisit $q = -\nu^{-1}(C_{\rm div})^{-2} r$ et $\boldsymbol{v} = \nu^{-1}(C_{\rm div})^{-2}(\boldsymbol{w} - \nu^{-1}(C_{\rm div})^{-2}\boldsymbol{w}_r)$, alors $(\boldsymbol{v},q) \in \mathcal{V}$ est tel que $\mathtt{T}((\boldsymbol{v},q)) = (\boldsymbol{w},r)$, d'où la surjectivité.

D'après ce qui précède, on déduit de (5.2) que la forme $a(\cdot,\cdot)$ est T-coercive, cf. définition 2.17 avec $\underline{\alpha} = \gamma = \frac{1}{2}\min(\nu^2(C_{\rm div})^2, 1)$. On peut construire une formulation variationnelle équivalente à (5.1), toujours posée dans $\mathcal{V}$, pour laquelle la forme bilinéaire est continue et coercive : le théorème de Lax-Milgram permet de conclure que celle-ci est bien posée. Enfin, on vérifie que la constante de stabilité apparaissant de la définition 2.6 est inférieure ou égale à $1/\alpha'_{St}$.

5.1.3 Condition inf-sup discrète

On utilise la méthode des éléments finis, dans un domaine Ω supposé polygonal si $\mathtt{d} = 2$, ou polyédrique si $\mathtt{d} = 3$. Le paramètre de discrétisation est égal au pas du maillage, noté h. Les éléments finis seront choisis sous certaines conditions. Précisément, des conditions suffisantes pour obtenir une condition inf-sup discrète uniforme.

On note cette fois $(\boldsymbol{V}_h)_h$ la suite des sous-espaces discrets de $\boldsymbol{H}^1_0(\Omega)$, et $(Q_h)_h$ la suite des sous-espaces discrets de $L^2_{vmn}(\Omega)$, et $\mathcal{V}_h = \boldsymbol{V}_h \times Q_h$ les sous-espaces discrets de $\mathcal{V}$ correspondants.

Remarque 5.3 *Pour les espaces discrets $(Q_h)_h$, on passe habituellement par une approximation conforme dans $L^2(\Omega)$. Mais il n'est pas aisé de construire des sous-ensembles de fonctions à moyenne nulle... En pratique, la pression étant définie* à une constante près*, on fixe cette constante en* annulant *un des degrés de liberté (pour h donné).*

Pour $h > 0$ donné, la formulation variationnelle discrète du modèle de Stokes "généralisé" s'écrit :

$$\left\{\begin{array}{l} \text{Trouver } (\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h \text{ telles que} \\ \forall (\boldsymbol{v}_h, q_h) \in \boldsymbol{V}_h \times Q_h, \\ \qquad \nu \displaystyle\int_\Omega \boldsymbol{\nabla}\boldsymbol{u}_h : \boldsymbol{\nabla}\boldsymbol{v}_h \, d\Omega - \int_\Omega p_h \,\mathrm{div}\, \boldsymbol{v}_h \, d\Omega - \int_\Omega q_h \,\mathrm{div}\, \boldsymbol{u}_h \, d\Omega \\ \qquad\qquad = \langle \boldsymbol{f}, \boldsymbol{v}_h \rangle_{(\boldsymbol{H}^1_0(\Omega))', \boldsymbol{H}^1_0(\Omega)} - \displaystyle\int_\Omega q_h \, g \, d\Omega. \end{array}\right. \tag{5.3}$$

Pour établir son caractère bien posé, ainsi que la convergence de l'erreur vers 0, on a vu qu'il fallait disposer d'une condition inf-sup discrète uniforme (3.17), cf. le théorème 3.15. *L'idée est,*

encore une fois (se souvenir de §3.4.2), de transposer au cas discret : soit la démonstration de la condition inf-sup conclue par la proposition 5.1 ; soit l'utilisation de la T*-coercivité du §5.1.2.* Dans la suite, nous allons donc réaliser cette transposition, en déterminant *sous quelles conditions* on peut la mener à bien. En particulier, on verra (voir (5.4)-(5.5) ci-dessous) qu'*on ne peut pas définir l'approximation de* $\boldsymbol{H}^1_0(\Omega)$ *par* $(\boldsymbol{V}_h)_h$ *indépendamment de l'approximation de* $L^2_{vmn}(\Omega)$ *par* $(Q_h)_h$.

Pour $(\boldsymbol{v}_h, q_h) \in \mathcal{V}_h$ non-nul, nous allons chercher $(\boldsymbol{w}^\star_h, r^\star_h) \in \mathcal{V}_h$ non-nul, dépendant linéairement de $(\boldsymbol{v}_h, q_h)$, et tel que

$$|a((\boldsymbol{v}_h, q_h), (\boldsymbol{w}^\star_h, r^\star_h))| \geq \alpha_\dagger \|(v_h, \boldsymbol{q}_h)\|_{\mathcal{V}} \, \|(w^\star_h, \boldsymbol{r}^\star_h)\|_{\mathcal{V}},$$

avec une constante $\alpha_\dagger > 0$ indépendante de $(\boldsymbol{v}_h, q_h)$ et de h. Si on transpose directement le choix de l'étape 3 pour l'obtention de la condition inf-sup abstraite, on choisit :

$$\boldsymbol{w}^\star = \nu(C_{\rm div}\,)^2 \boldsymbol{v}_h - \boldsymbol{w}_{q_h} \text{ et } r^\star = -\nu(C_{\rm div}\,)^2 q_h,$$

avec $\boldsymbol{w}_{q_h} \in \boldsymbol{H}^1_0(\Omega)$ tel que $\operatorname{div} \boldsymbol{w}_{q_h} = q_h$, et $|\boldsymbol{w}_{q_h}|_{1,\Omega} \leq C_{\rm div}\, \|q_h\|$ ($C_{\rm div}\, > 0$ indépendante de q_h). La *difficulté* dans le choix ci-dessus est, qu'en général, $\boldsymbol{w}_{q_h} \notin \boldsymbol{V}_h$. Pour la contourner, on veut trouver $\boldsymbol{w}^+_h \in \boldsymbol{V}_h$ tel que "$\operatorname{div} \boldsymbol{w}^+_h = q_h$" *en un sens à préciser*, et $|\boldsymbol{w}^+_h|_{1,\Omega} \leq C^+ \|q_h\|$, avec $C^+ > 0$ indépendante de q_h et de h. Si cela est possible, on choisira à partir de là

$$\boldsymbol{w}^\star_h = \nu(C^+)^2 \boldsymbol{v}_h - \boldsymbol{w}^+_h \in \boldsymbol{V}_h \text{ et } r^\star_h = -\nu(C^+)^2 q_h \in Q_h,$$

et on pourra reproduire les calculs de la condition inf-sup abstraite, C^+ remplaçant $C_{\rm div}\,$.
La démarche est la suivante. On doit trouver des espaces discrets $(\boldsymbol{V}_h, Q_h)_h$ tels que

$$\exists C_\pi > 0,\ \forall h,\ \exists \pi_h \in \mathcal{L}(\boldsymbol{H}^1_0(\Omega), \boldsymbol{V}_h) \text{ possédant les propriétés :}$$

$$\forall \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega), \quad |\pi_h \boldsymbol{v}|_{1,\Omega} \leq C_\pi |\boldsymbol{v}|_{1,\Omega}\,; \tag{5.4}$$

$$\forall \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega),\ \forall q'_h \in Q_h, \quad \int_\Omega q'_h \operatorname{div}(\pi_h \boldsymbol{v})\, d\Omega = \int_\Omega q'_h \operatorname{div} \boldsymbol{v}\, d\Omega. \tag{5.5}$$

En effet, si on dispose d'approximations $(\boldsymbol{V}_h, Q_h)_h$ et d'opérateurs $(\pi_h)_h$ satisfaisant ces propriétés, on va voir qu'on peut choisir $\boldsymbol{w}^+_h = \pi_h \boldsymbol{w}_{q_h}$. Pourquoi ? Les réponses sont simples !
D'après (5.4) et la définition de $\boldsymbol{w}_{q_h}$, on a

$$|\boldsymbol{w}^+_h|_{1,\Omega} = |\pi_h \boldsymbol{w}_{q_h}|_{1,\Omega} \leq C_\pi |\boldsymbol{w}_{q_h}|_{1,\Omega} \leq C^+ \|q_h\|, \text{ avec } C^+ = C_\pi C_{\rm div}\, > 0. \tag{5.6}$$

Et, d'après (5.5) et la définition de $\boldsymbol{w}_{q_h}$, on a

$$\forall q'_h \in Q_h, \quad \int_\Omega q'_h \operatorname{div} \boldsymbol{w}^+_h\, d\Omega = \int_\Omega q'_h \operatorname{div}(\pi_h \boldsymbol{w}_{q_h})\, d\Omega = \int_\Omega q'_h \operatorname{div} \boldsymbol{w}_{q_h}\, d\Omega = \int_\Omega q'_h q_h\, d\Omega. \tag{5.7}$$

Remarque 5.4 *Ci-dessus, on remarque que la propriété "*$\operatorname{div} \boldsymbol{w}_h^+ = q_h$*" est réalisée en un sens faible, puisqu'on impose uniquement que*

$$(\operatorname{div} \boldsymbol{w}_h^+ - q_h) \in Q_h^\perp, \ \text{où } L^2(\Omega) = Q_h \overset{\perp}{\oplus} Q_h^\perp.$$

Mais ceci est suffisant, puisque la condition inf-sup discrète doit être imposée variationnellement dans $\boldsymbol{V}_h \times Q_h$*, voir ci-après.*

Pour résumer ce qui précède, on choisit finalement

$$\boldsymbol{w}_h^\star = \nu (C^+)^2 \boldsymbol{v}_h - \pi_h \boldsymbol{w}_{q_h} \in \boldsymbol{V}_h \text{ et } r_h^\star = -\nu (C^+)^2 q_h \in Q_h,$$

avec $(\pi_h)_h$ qui vérifie (5.4)-(5.5), et $C^+ = C_\pi C_{\mathrm{div}}$. Si on reprend les calculs, on trouve

$$\begin{aligned}
a((\boldsymbol{v}_h, q_h), (\boldsymbol{w}_h^\star, r_h^\star)) &= \nu \int_\Omega \boldsymbol{\nabla} \boldsymbol{v}_h : \boldsymbol{\nabla} \boldsymbol{w}_h^\star \, d\Omega - \int_\Omega q_h \operatorname{div} \boldsymbol{w}_h^\star \, d\Omega - \int_\Omega r_h^\star \operatorname{div} \boldsymbol{v}_h \, d\Omega \\
&= \nu^2 (C^+)^2 \, |\boldsymbol{v}_h|_{1,\Omega}^2 - \nu \int_\Omega \boldsymbol{\nabla} \boldsymbol{v}_h : \boldsymbol{\nabla} (\pi_h \boldsymbol{w}_{q_h}) \, d\Omega \\
&\quad \cancel{-\nu (C^+)^2 \int_\Omega q_h \operatorname{div} \boldsymbol{v}_h \, d\Omega} + \int_\Omega q_h \operatorname{div} (\pi_h \boldsymbol{w}_{q_h}) \, d\Omega + \cancel{\nu (C^+)^2 \int_\Omega q_h \operatorname{div} \boldsymbol{v}_h \, d\Omega}. \\
\text{(propriété (5.7))} &= \nu^2 (C^+)^2 \, |\boldsymbol{v}_h|_{1,\Omega}^2 - \nu \int_\Omega \boldsymbol{\nabla} \boldsymbol{v}_h : \boldsymbol{\nabla} (\pi_h \boldsymbol{w}_{q_h}) \, d\Omega + \|q_h\|^2.
\end{aligned}$$

On conclut le calcul comme précédemment, grâce à l'inégalité de Young ($\eta > 0$) :

$$\begin{aligned}
-\nu \int_\Omega \boldsymbol{\nabla} \boldsymbol{v}_h : \boldsymbol{\nabla} (\pi_h \boldsymbol{w}_{q_h}) \, d\Omega &\geq -\frac{\eta \nu}{2} |\boldsymbol{v}_h|_{1,\Omega}^2 - \frac{\nu}{2\eta} |\pi_h \boldsymbol{w}_{q_h}|_{1,\Omega}^2 \\
\text{(propriété (5.6))} &\geq -\frac{\eta \nu}{2} |\boldsymbol{v}_h|_{1,\Omega}^2 - \frac{\nu (C^+)^2}{2\eta} \|q_h\|^2.
\end{aligned}$$

On choisit, toujours à la suite du calcul abstrait, $\eta = \nu (C^+)^2 > 0$, pour trouver finalement :

$$\begin{aligned}
a((\boldsymbol{v}_h, q_h), (\boldsymbol{w}_h^\star, r_h^\star)) &\geq \frac{1}{2} \nu^2 (C^+)^2 \, |\boldsymbol{v}_h|_{1,\Omega}^2 + \frac{1}{2} \|q_h\|^2 \\
&\geq \frac{1}{2} \min(\nu^2 (C^+)^2, 1) \big(|\boldsymbol{v}_h|_{1,\Omega}^2 + \|q_h\|^2 \big) \\
&= \frac{1}{2} \min(\nu^2 (C^+)^2, 1) \|(\boldsymbol{v}_h, q_h)\|_{\mathcal{V}}^2.
\end{aligned}$$

Par ailleurs, similairement au cas abstrait :

$$\begin{aligned}
\|(\boldsymbol{w}_h^\star, r_h^\star)\|_{\mathcal{V}} &\leq \Big(2\nu^2 (C^+)^4 |\boldsymbol{v}_h|_{1,\Omega}^2 + 2 (C^+)^2 \, \|q_h\|^2 + \nu^2 (C^+)^4 \|q_h\|^2 \Big)^{1/2} \\
&\leq \max \big(\sqrt{2} \nu (C^+)^2, C^+ [2 + \nu^2 (C^+)^2]^{1/2} \big) \big\| (\boldsymbol{v}_h, q_h) \|_{\mathcal{V}}.
\end{aligned}$$

La condition inf-sup discrète uniforme suit...

On a à nouveau utilisé une preuve "adaptant" la T-coercivité du §2.3.4 : en effet, $\mathtt{T}_h \in \mathcal{L}(\boldsymbol{V}_h \times Q_h)$ définie par

$$\mathtt{T}_h((\boldsymbol{v}_h, q_h)) = (\nu (C^+)^2 \boldsymbol{v}_h - \pi_h \boldsymbol{w}_{q_h}, -\nu (C^+)^2 q_h),$$

est telle que, lorsque les conditions (5.4)-(5.5) sont vérifiées, pour tout h, la forme a restreinte à $(\boldsymbol{V}_h \times Q_h) \times (\boldsymbol{V}_h \times Q_h)$ est $\mathtt{T}$-coercive au sens de la définition 2.17, avec une constante $\underline{\alpha}$ indépendante de h. De plus, la norme de $\mathtt{T}_h$ est indépendante de h. On renvoie au chapitre 9 (voir la définition 9.15 pour la "$\mathtt{T}$-coercivité discrète (uniforme en h)").

Proposition 5.5 *Si les espaces discrets $(\boldsymbol{V}_h, Q_h)_h$ et les opérateurs $(\pi_h)_h$ satisfont les propriétés (5.4)-(5.5), alors la forme a apparaissant dans les formulations variationnelles de Stokes (4.6) et de Stokes "généralisé" (4.8) vérifie une condition inf-sup discrète uniforme :*

$$\forall h,\ \forall (\boldsymbol{v}_h, q_h) \in \mathcal{V}_h, \qquad \sup_{(\boldsymbol{w}_h, r_h) \in \mathcal{V}_h \setminus \{0\}} \frac{|a((\boldsymbol{v}_h, q_h), (\boldsymbol{w}_h, r_h))|}{\|(\boldsymbol{w}_h, r_h)\|_{\mathcal{V}}} \geq \alpha_{\dagger,St}\, \|(\boldsymbol{v}_h, q_h)\|_{\mathcal{V}},$$

où on a posé

$$\alpha_{\dagger,St} = \frac{\frac{1}{2}\min(\nu^2 (C^+)^2, 1)}{\max\left(\sqrt{2}\nu (C^+)^2, C^+ [2 + \nu^2 (C^+)^2]^{1/2}\right)} > 0.$$

On peut donc appliquer le lemme de Céa (théorème 3.13), et en déduire la convergence de l'erreur vers 0 si la propriété d'approximabilité (3.11) est vérifiée (voir le théorème 3.15).

5.1.4 Exemples de discrétisation

Pour construire les sous-espaces discrets $(\boldsymbol{V}_h, Q_h)_h$ vérifiant (5.4)-(5.5), on s'appuie sur une famille régulière de maillages de Ω, formés de simplexes. Nous donnons deux exemples de discrétisation par éléments finis, en suivant [3].

Premièrement, l'élément fini MINI de degré $k \geq 1$. On donne la définition pour l'élément fini MINI de degré $k = 1$:

$$\begin{aligned} \boldsymbol{V}_h &= \{\boldsymbol{v}_h \in \boldsymbol{C}^0(\overline{\Omega}) \text{ tel que } \boldsymbol{v}_{h|\partial\Omega} = 0 \text{ et } \boldsymbol{v}_{h|T_\ell} \in \{\boldsymbol{P}^1(T_\ell) + \mathbf{Bulle}(T_\ell)\},\ \forall \ell = 1, L\}, \\ Q_h &= \{q_h \in \mathcal{C}^0(\overline{\Omega}) \text{ tel que } q_{h|T_\ell} \in P^1(T_\ell),\ \forall \ell = 1, L\}, \end{aligned} \tag{5.8}$$

où, pour tout simplexe T, on a défini l'ensemble des "bulles" [3] selon

$$\mathrm{Bulle}(T) = \mathrm{Vect}\Big(\prod_{i=1,\mathtt{d}+1} \lambda_i^T\Big),$$

avec $(\lambda_i^T)_{i=1,\mathtt{d}+1}$ les coordonnées barycentriques dans T (on a donc $\dim(\mathrm{Bulle}(T)) = 1$) ; et comme d'habitude $\mathbf{Bulle}(T) = (\mathrm{Bulle}(T))^{\mathtt{d}}$. Il est prouvé, pp. 470-471 pour $\mathtt{d} = 2$ et pp. 491-492 pour $\mathtt{d} = 3$ dans [3], que l'élément fini MINI vérifie (5.4)-(5.5).
Pour la propriété d'approximabilité, on note que ces espaces discrets contiennent les espaces composés des éléments finis de Lagrange P^1 (vectoriels pour $\boldsymbol{V}_h$, scalaires pour Q_h) ; la propriété d'approximabilité étant vraie en norme $\|\cdot\|_{\boldsymbol{H}^1(\Omega)}$, voir le §3.4.1, et elle est a fortiori vraie

3. Une propriété notable d'une "bulle" est qu'elle s'annule sur ∂T, c'est-à-dire que son support est égal à T.

en norme $|\cdot|_{1,\Omega}$ pour $(\boldsymbol{V}_h)_h$ dans $\boldsymbol{H}^1_0(\Omega)$, et en norme $\|\cdot\|$ pour $(Q_h)_h$ dans $L^2_{vmn}(\Omega)$.

Comme annoncé plus haut, d'après le lemme de Céa et la propriété d'approximabilité, on a donc convergence de l'erreur vers 0 pour l'élément fini MINI (théorème 3.15) :

$$\lim_{h\to 0}\left(|\boldsymbol{u}-\boldsymbol{u}_h|_{1,\Omega}+\|p-p_h\|\right)=0.$$

Concernant enfin la vitesse de convergence, si de plus la solution $(\boldsymbol{u},p)$ de Stokes, ou de Stokes "généralisé", est de régularité $\boldsymbol{H}^2\times H^1$ "par morceaux", on peut établir que : [4]

$$\exists C_{St}>0,\ \forall h,\quad |\boldsymbol{u}-\boldsymbol{u}_h|_{1,\Omega}+\|p-p_h\|\le C_{St}\,h.$$

Deuxièmement, l'élément fini de Taylor-Hood $\boldsymbol{P}^{k+1}-P^k$ pour $k\ge 1$. [5] La définition pour $k=1$ est :

$$\begin{array}{l}\boldsymbol{V}_h=\{\boldsymbol{v}_h\in\boldsymbol{C}^0(\overline{\Omega})\text{ tel que }\boldsymbol{v}_{h|\partial\Omega}=0\text{ et }\boldsymbol{v}_{h|T_\ell}\in\boldsymbol{P}^2(T_\ell),\ \forall\ell=1,L\},\\ Q_h=\{q_h\in\mathcal{C}^0(\overline{\Omega})\text{ tel que }q_{h|T_\ell}\in P^1(T_\ell),\ \forall\ell=1,L\}.\end{array}\tag{5.9}$$

Pour cet élément fini, les conditions (5.4)-(5.5) sont établies dans [3], pp. 496-504, pour $\mathtt{d}=2,3$. La propriété d'approximabilité est encore une fois une conséquence de celle déjà énoncée pour les éléments finis de Lagrange, et on en déduit la convergence de l'erreur vers 0 pour l'élément fini de Taylor-Hood :

$$\lim_{h\to 0}\left(|\boldsymbol{u}-\boldsymbol{u}_h|_{1,\Omega}+\|p-p_h\|\right)=0.$$

Et pour la vitesse de convergence, si la solution $(\boldsymbol{u},p)$ de Stokes, ou de Stokes "généralisé", est de régularité $\boldsymbol{H}^2\times H^1$ "par morceaux", on a :[4]

$$\exists C_{St}>0,\ \forall h,\quad |\boldsymbol{u}-\boldsymbol{u}_h|_{1,\Omega}+\|p-p_h\|\le C_{St}\,h.$$

5.2 Approche un-plus-un : formulation variationnelle mixte

On choisit maintenant d'"isoler" la contrainte sur la divergence. Précisément, on écrit

$$\left\{\begin{array}{l}\text{Trouver }(\boldsymbol{u},p)\in\boldsymbol{H}^1_0(\Omega)\times L^2_{vmn}(\Omega)\text{ telles que}\\ \forall\boldsymbol{v}\in\boldsymbol{H}^1_0(\Omega),\quad \nu\displaystyle\int_\Omega\boldsymbol{\nabla u}:\boldsymbol{\nabla v}\,d\Omega-\int_\Omega p\,\mathrm{div}\,\boldsymbol{v}\,d\Omega=\langle\boldsymbol{f},\boldsymbol{v}\rangle_{(\boldsymbol{H}^1_0(\Omega))',\boldsymbol{H}^1_0(\Omega)}\\ \forall q\in L^2_{vmn}(\Omega),\quad \displaystyle\int_\Omega q\,\mathrm{div}\,\boldsymbol{u}\,d\Omega=\int_\Omega q\,g\,d\Omega.\end{array}\right.\tag{5.10}$$

4. Si la solution $(\boldsymbol{u},p)$ est plus régulière, la vitesse de convergence sera meilleure sous réserve qu'on augmente l'ordre de l'élément fini de façon adéquate.

5. Pour $k=0$, c'est-à-dire pour l'élément fini $\boldsymbol{P}^1-P^0$, on n'a pas de condition inf-sup discrète uniforme !

5.2.1 Formulation variationnelle mixte abstraite

De façon abstraite, une *formulation variationnelle mixte* s'écrit

$$\left\{\begin{array}{l} \text{Trouver } (u,p)\in V\times Q \text{ tel que} \\ \forall v\in V \quad c(u,v)+\overline{b(v,p)}=\langle f,v\rangle_V, \\ \forall q\in Q, \quad b(u,q)=\langle g,q\rangle_Q, \end{array}\right. \tag{5.11}$$

avec :

— V et Q deux espaces de Hilbert définis sur $\mathbb{C}$;
— $c(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V\times V$;
— $b(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V\times Q$;
— $f\in V'$ et $g\in Q'$ donnés.

On parle de formulation mixte, car si on compare à la formulation initiale (5.1), on a isolé la "contrainte" en découpant la forme a en deux parties, la forme b exprimant la contrainte. [6]

Pour le modèle de Stokes "généralisé", [7] on a, cf. (5.10) :

— $V=\boldsymbol{H}^1_0(\Omega)$ muni de $\|\boldsymbol{v}\|_V=|\boldsymbol{v}|_{1,\Omega}$, $Q=L^2_{vmn}(\Omega)$ muni de $\|q\|_Q=\|q\|$;
— $c(\boldsymbol{v},\boldsymbol{w})=\nu\displaystyle\int_\Omega \boldsymbol{\nabla v}:\boldsymbol{\nabla w}\,d\Omega$, $b(\boldsymbol{v},q)=-\displaystyle\int_\Omega q\,\mathrm{div}\,\boldsymbol{v}\,d\Omega$;
— $\langle f,\boldsymbol{v}\rangle_V=\langle \boldsymbol{f},\boldsymbol{v}\rangle_{(\boldsymbol{H}^1_0(\Omega))',\boldsymbol{H}^1_0(\Omega)}$; $\langle g,q\rangle_Q=-\displaystyle\int_\Omega q\,g\,d\Omega$.

Cas $g=0$

On va commencer par étudier le cas où $g=0$. On introduit le "noyau de la forme b" : [8]

$$V_0=\{v\in V \text{ tel que } b(v,q)=0,\ \forall q\in Q\}.$$

Comme la forme b est continue, V_0 est un sous-espace vectoriel fermé de V. On a donc la décomposition orthogonale $V=V_0\overset{\perp}{\oplus}V_0^\perp$, et résoudre la formulation variationnelle mixte (5.11) avec $g=0$ *est équivalent à résoudre* :

$$\left\{\begin{array}{l} \text{Trouver } (u,p)\in V_0\times Q \text{ tel que} \\ \forall v\in V_0, \quad c(u,v)=\langle f,v\rangle_V, \\ \forall v'\in V_0^\perp, \quad c(u,v')+\overline{b(v',p)}=\langle f,v'\rangle_V. \end{array}\right.$$

6. En termes de "système linéaire", la formulation variationnelle mixte (5.11) est équivalente à :

$$\left\{\begin{array}{ll} \text{Trouver } (u,p)\in V\times Q \text{ tel que} & \\ \mathtt{C}u+\mathtt{B}^*p=F & \text{dans } V \\ \mathtt{B}u=G & \text{dans } Q \end{array}\right.$$

où d'une part $\mathtt{C}\in\mathcal{L}(V,V)$ est l'application linéaire associée à c, $\mathtt{B}\in\mathcal{L}(V,Q)$ est l'application linéaire associée à b (voir (2.11)), et $\mathtt{B}^*\in\mathcal{L}(Q,V)$ est définie par : $\forall v\in V$, $\forall q\in Q$, $b(v,q)=(v,\mathtt{B}^*q)_V$; et, d'autre part, $F\in V$ est associé à f, et $G\in Q$ est associé à g, par le théorème de Riesz.

7. Pour le modèle de Stokes, on choisit $g=0$.

8. Précisément, $V_0=\ker\mathtt{B}$.

On raisonne *en deux temps.*

Etape 1. On commence par la résolution de

$$\left\{\begin{array}{l}\text{Trouver } u \in V_0 \text{ tel que}\\ \forall v \in V_0, \quad c(u,v) = \langle f_0, v\rangle_{V_0},\end{array}\right. \tag{5.12}$$

avec $f_0 \in (V_0)'$ définie par $v \mapsto \langle f, v\rangle_V$. D'après le théorème BNB (théorème 2.12), cette formulation variationnelle à un espace est bien posée (au sens de Hadamard) si, et seulement si

$$\text{la forme } c_{|V_0\times V_0} \text{ vérifie une condition inf-sup et une condition de solvabilité.} \tag{5.13}$$

On suppose à partir de maintenant que ces conditions sont vérifiées : alors, il existe $C_0 > 0$ telle que, pour tout $f_0 \in (V_0)'$, il existe une solution unique u, et

$$\|u\|_V \leq C_0 \|f_0\|_{(V_0)'}.$$

Etape 2. Soit $f_\perp \in (V_0^\perp)'$ définie par $v' \mapsto \langle f, v'\rangle_V$. On s'intéresse à la résolution de

$$\left\{\begin{array}{l}\text{Trouver } p \in Q \text{ tel que}\\ \forall v' \in V_0^\perp, \quad \underline{b}(p,v') = \langle f_\perp, v'\rangle_V - c(u,v'),\end{array}\right. \tag{5.14}$$

où on a noté $\underline{b}$ la forme sesquilinéaire et continue sur $Q \times V_0^\perp$ définie par

$$\forall q \in Q,\ \forall v' \in V_0^\perp, \quad \underline{b}(q,v') = \overline{b(v',q)}. \tag{5.15}$$

On reconnait cette fois une formulation variationnelle à deux espaces. D'après le théorème BNB (théorème 2.12), cette formulation variationnelle est bien posée (au sens de Hadamard) si, et seulement si,

$$\text{la forme } \underline{b}_{|Q\times V_0^\perp} \text{ vérifie une condition inf-sup et une condition de solvabilité.}$$

On note que la définition de $\underline{b}$ et celle de $V_0^\perp$ dépendent de b... Ceci a des conséquences sur la condition de solvabilité. En effet, soit $v' \in V_0^\perp$ telle que $\underline{b}(q,v') = 0$ pour tout $q \in Q$: ceci signifie que v appartient au "noyau de b", égal à V_0, et ainsi $v' \in V_0^\perp \cap V_0 = \{0\}$! La condition de solvabilité sur $\underline{b}$ est toujours satisfaite. Ainsi, la formulation variationnelle (5.14) est bien posée (au sens de Hadamard) si, et seulement si,

$$\text{la forme } \underline{b}_{|Q\times V_0^\perp} \text{ vérifie une condition inf-sup.}$$

Si on revient à la forme b, cette condition s'écrit de façon équivalente :

$$\exists \beta > 0,\ \forall q \in Q, \quad \sup_{v' \in V_0^\perp \setminus \{0\}} \frac{|b(v',q)|}{\|v'\|_V} \geq \beta \, \|q\|_Q. \tag{5.16}$$

La différence avec la condition inf-sup usuelle est que les deux arguments sont inversés : on parlera de *condition inf-sup inversée.*

Remarque 5.6 *Dans (5.16), on peut également considérer le supremum pour tout $v' \in V \setminus \{0\}$, puisque l'orthogonal de $V_0^\perp$ est égal au "noyau de b".*

Si on suppose que les conditions (5.13) et (5.16) sont vérifiées, on en conclut que, si on choisit $v'_p \in V_0^\perp \setminus \{0\}$ réalisant (5.16) pour la solution p de la formulation variationnelle (5.14), on a

$$\begin{aligned} \|p\|_Q &\leq \frac{1}{\beta}\frac{|b(v'_p,p)|}{\|v'_p\|_V} \\ (5.14) \quad &\leq \frac{1}{\beta}\left(\frac{|\langle f, v'_p\rangle_V|}{\|v'_p\|_V} + \frac{|c(u,v'_p)|}{\|v'_p\|_V}\right) \\ &\leq \frac{1}{\beta}\left(\|f_\perp\|_{(V_0^\perp)'} + |||c|||\,\|u\|_V\right) \\ &\leq \frac{1}{\beta}\left(\|f_\perp\|_{(V_0^\perp)'} + C_0\,|||c|||\,\|f_0\|_{(V_0)'}\right). \end{aligned}$$

Puisque par définition $\max\left(\|f_0\|_{(V_0)'}, \|f_\perp\|_{(V_0^\perp)'}\right) \leq \|f\|_{V'}$, on a établi la

Proposition 5.7 *Soit à résoudre la formulation variationnelle (5.11) avec $f \in V'$ et $g = 0$. On suppose que la forme c vérifie les conditions (5.13), et que la forme b vérifie la condition inf-sup inversée (5.16). Alors, il existe une solution $(u,p) \in V \times Q$ et une seule à (5.11), avec dépendance continue par rapport à $\|f\|_{V'}$.*

Pour le modèle de Stokes, on note que, la forme $c : (\boldsymbol{v},\boldsymbol{w}) \mapsto \nu \displaystyle\int_\Omega \boldsymbol{\nabla v} : \boldsymbol{\nabla w}\, d\Omega$ étant coercive sur $\boldsymbol{H}_0^1(\Omega)$, elle l'est a fortiori sur $\boldsymbol{V}_0(\Omega) \subset \boldsymbol{H}_0^1(\Omega)$: les conditions (5.13) sont satisfaites. En outre, la condition (5.16) sur la forme b est également vraie : en effet, d'après le théorème 4.3, on a, pour tout $q \in L^2_{vmn}(\Omega) \setminus \{0\}$,

$$\sup_{\boldsymbol{v}\in \boldsymbol{V}_0(\Omega)^\perp\setminus\{0\}} \frac{\left|\displaystyle\int_\Omega q \operatorname{div} \boldsymbol{v}\, d\Omega\right|}{|\boldsymbol{v}|_{1,\Omega}} \geq \frac{\left|\displaystyle\int_\Omega q \operatorname{div} \boldsymbol{v}_q\, d\Omega\right|}{|\boldsymbol{v}_q|_{1,\Omega}} = \frac{\|q\|^2}{|\boldsymbol{v}_q|_{1,\Omega}} \geq \frac{1}{C_{\rm div}}\|q\|.$$

Cas g quelconque

On suppose que les hypothèses de la proposition 5.7 sont satisfaites. Si on revient à (5.11), la nouveauté est qu'on cherche $u \in V$ tel que : $\forall q \in Q$, $b(u,q) = \langle g, q\rangle_Q$, avec a priori $g \neq 0$. Si on décompose u selon $V = V_0 \overset{\perp}{\oplus} V_0^\perp$, on cherche $(u_0,u') \in V_0 \times V_0^\perp$ avec $u = u_0 + u'$. Ainsi, résoudre la formulation variationnelle mixte (5.11) avec g quelconque *est équivalent à résoudre* :

$$\left\{\begin{array}{l} \text{Trouver } (u_0,u',p) \in V_0 \times V_0^\perp \times Q \text{ tel que} \\ \quad \forall q \in Q, \quad b(u',q) = \langle g, q\rangle_Q, \\ \quad \forall v \in V_0, \quad c(u_0,v) + c(u',v) = \langle f, v\rangle_V, \\ \quad \forall v' \in V_0^\perp, \quad c(u_0,v') + c(u',v') + \overline{b(v',p)} = \langle f, v'\rangle_V. \end{array}\right.$$

On raisonne à nouveau *en deux temps.*

Etape 1. On considère la (nouvelle) formulation variationnelle à deux espaces

$$\left\{\begin{array}{l} \text{Trouver } u' \in V_0^\perp \text{ tel que} \\ \forall q \in Q, \quad b(u', q) = \langle g, q\rangle_Q. \end{array}\right. \tag{5.17}$$

Etudions son caractère bien posé sous les hypothèses de la proposition 5.7.

On introduit l'application linéaire $\underline{\mathtt{B}} \in \mathcal{L}(Q, V_0^\perp)$ associée à la forme $\underline{b}$ définie en (5.15) :

$$\forall q \in Q,\ \forall v' \in V_0^\perp, \quad (\underline{\mathtt{B}} q, v')_V = \underline{b}(q, v') = \overline{b(v', q)}. \tag{5.18}$$

Par hypothèse la formulation variationnelle (5.14) est bien posée : $\underline{\mathtt{B}}^{-1}$ existe, et $\underline{\mathtt{B}}^{-1} \in \mathcal{L}(V_0^\perp, Q)$.

Revenons à (5.17).

Concernant la condition inf-sup, pour $v' \in V_0^\perp \setminus \{0\}$, on observe que :

$$\begin{aligned} \sup_{q\in Q\setminus\{0\}} \frac{|b(v', q)|}{\|q\|_Q} &\geq \frac{|b(v', \underline{\mathtt{B}}^{-1} v')|}{\|\underline{\mathtt{B}}^{-1} v'\|_Q} \\ (\text{définition (5.18)}) &= \frac{\|v'\|_V^2}{\|\underline{\mathtt{B}}^{-1} v'\|_Q} \\ (\|\underline{\mathtt{B}}^{-1} v'\|_Q \leq |||\underline{\mathtt{B}}^{-1}|||\, \|v'\|_V) &\geq \frac{1}{|||\underline{\mathtt{B}}^{-1}|||}\, \|v'\|_V. \end{aligned}$$

Concernant la condition de solvabilité, soit $q \in Q$ tel que $b(v', q) = 0$ pour tout $v' \in V_0^\perp$: d'après (5.16), on a immédiatement $q = 0$.

Ainsi, (5.17) est bien posée sous les hypothèses de la proposition 5.7, et on a caractérisé la partie de la soution $u' \in V_0^\perp$: il existe $C' > 0$ telle que, pour tout $g \in Q'$, il existe une solution unique u', et

$$\|u'\|_V \leq C' \|g\|_{Q'}.$$

Etape 2. On définit $f' \in V'$ tel que [9]

$$\forall v \in V, \quad \langle f', v\rangle_V = \langle f, v\rangle_V - c(u', v),$$

et on résout maintenant

$$\left\{\begin{array}{l} \text{Trouver } (u_0, p) \in V_0 \times Q \text{ tel que} \\ \quad \forall v \in V_0, \quad c(u_0, v) = \langle f', v\rangle_V, \\ \quad \forall v' \in V_0^\perp, \quad c(u_0, v') + \overline{b(v', p)} = \langle f', v'\rangle_V. \end{array}\right.$$

C'est exactement la formulation variationnelle résolue pour le cas $g = 0$, avec la donnée f' : les mêmes conclusions s'appliquent.

Pour résumer, dans le cas général, on a établi le

9. On a la dépendance continue :

$$\|f'\|_{V'} = \sup_{v\in V\setminus\{0\}} \frac{|\langle f', v\rangle_V|}{\|v\|_V} \leq \|f\|_{V'} + \sup_{v\in V\setminus\{0\}} \frac{|c(u', v)|}{\|v\|_V} \leq \|f\|_{V'} + C' |||c|||\, \|g\|_{Q'}.$$

Théorème 5.8 (Babuska-Brezzi) *Soit à résoudre la formulation variationnelle (5.11) avec $f \in V'$ et $g \in Q'$. On suppose que la forme c vérifie les conditions (5.13), et que la forme b vérifie la condition inf-sup inversée (5.16). Alors, il existe une solution $(u,p) \in V \times Q$ et une seule à (5.11), avec dépendance continue par rapport à $\|f\|_{V'}$ et $\|g\|_{Q'}$.*

Pour le modèle de Stokes "généralisé", la constante inf-sup correspondant à (5.16) est égale à

$$\inf_{q \in L^2_{vmn}(\Omega)\setminus\{0\}} \sup_{\boldsymbol{w} \in \boldsymbol{H}^1_0(\Omega)\setminus\{0\}} \frac{\left| \int_\Omega q \operatorname{div} \boldsymbol{w} \, d\Omega \right|}{\|q\| \, |\boldsymbol{w}|_{1,\Omega}}.$$

On l'appelle *constante inf-sup de Ladyzhenskaya-Babuska-Brezzi*. L'étude de sa valeur en tant que fonction du domaine Ω a fait l'objet de nombreux travaux, voir [16].

5.2.2 Formulation variationnelle mixte discrète

On introduit $(V_\delta)_\delta$ et $(Q_\delta)_\delta$ deux suites d'espaces vectoriels de dimensions finies, avec $V_\delta \subset V$ et $Q_\delta \subset Q$ pour tout $\delta > 0$. Pour δ donné, la *formulation variationnelle mixte discrète* s'écrit

$$\left\{ \begin{array}{l} \text{Trouver } (u_\delta, p_\delta) \in V_\delta \times Q_\delta \text{ tel que} \\ \forall v_\delta \in V_\delta, \quad c(u_\delta, v_\delta) + \overline{b(v_\delta, p_\delta)} = \langle f, v_\delta \rangle_V, \\ \forall q_\delta \in Q_\delta, \quad b(u_\delta, q_\delta) = \langle g, q_\delta \rangle_Q. \end{array} \right. \tag{5.19}$$

Pour l'approximation des formulations variationnelles mixtes, la difficulté est, qu'en général, *on n'approche pas V_0, le "noyau de b", de façon conforme*. En outre, comme on l'a déjà constaté au §3.3, *la condition inf-sup (5.16) sur $\underline{b}$ ne se transmet pas au cadre discret*. On introduit donc un "noyau discret de la forme b" : [10]

$$V_{0\delta} = \{ v_\delta \in V_\delta \text{ tel que } b(v_\delta, q_\delta) = 0, \ \forall q_\delta \in Q_\delta \}.$$

On a le résultat suivant, tiré du chapitre II de [18], qui permet d'estimer l'erreur.

Théorème 5.9 *On suppose que les deux conditions ci-dessous sont vérifiées.*

— *Coercivité* uniforme *sur les noyaux discrets :*

$$\exists \alpha^\dagger > 0, \ \forall \delta, \ \forall v_\delta \in V_{0\delta}, \quad |c(v_\delta, v_\delta)| \geq \alpha^\dagger \|v_\delta\|_V^2. \tag{5.20}$$

— *Condition inf-sup inversée discrète* uniforme *sur b :*

$$\exists \beta^\dagger > 0, \ \forall \delta, \ \forall q_\delta \in Q_\delta, \quad \sup_{v_\delta \in V_\delta \setminus \{0\}} \frac{|b(v_\delta, q_\delta)|}{\|v_\delta\|_V} \geq \beta^\dagger \, \|q_\delta\|_Q. \tag{5.21}$$

Alors, pour tout δ, la formulation variationnelle mixte discrète (5.19) est bien posée.
De plus, on a l'estimation d'erreur

$$\exists C^\dagger > 0, \ \forall (f,g) \in V' \times Q', \ \forall \delta, \quad \|u - u_\delta\|_V + \|p - p_\delta\|_Q \leq C^\dagger \Big(\inf_{v_\delta \in V_\delta} \|u - v_\delta\|_V + \inf_{q_\delta \in Q_\delta} \|p - q_\delta\|_Q \Big).$$

10. Précisément, $V_{0\delta} = \ker \mathtt{B}_\delta$ où $\mathtt{B}_\delta \in \mathcal{L}(V_\delta, Q_\delta)$ est l'application linéaire associée à $b_{|V_\delta \times Q_\delta}$.

On peut en déduire la convergence de l'erreur vers 0 si la propriété d'approximabilité (3.11) est vérifiée pour la famille $(V_\delta)_\delta$ dans V, ainsi que pour la famille $(Q_\delta)_\delta$ dans Q.

Concernant la condition inf-sup inversée discrète uniforme (5.21), on a le résultat ci-dessous, également tiré du chapitre II de [18], qui énonce une condition abstraite équivalente.

Théorème 5.10 `(Lemme de Fortin)` *On suppose que $b(\cdot,\cdot)$ vérifie la condition inf-sup inversée (5.16). Alors, la condition inf-sup inversée discrète uniforme (5.21) est équivalente à*

$$\begin{aligned}\exists \beta^\dagger > 0,\ \forall \delta,\ \exists \pi_\delta \in \mathcal{L}(V, V_\delta)\ &\textit{possédant les propriétés :}\\ &\forall v \in V, \quad \|\pi_\delta v\|_V \le \beta^\dagger \|v\|_V\,;\\ &\forall v \in V,\ \forall q_\delta \in Q_\delta, \quad b(\pi_\delta v, q_\delta) = b(v, q_\delta).\end{aligned}$$

Si on revient aux modèles de Stokes, on considère $(\boldsymbol{V}_h)_h$ une suite de sous-espaces discrets de $\boldsymbol{H}^1_0(\Omega)$, et $(Q_h)_h$ une suite de sous-espaces discrets de $L^2_{vmn}(\Omega)$, et on discrétise la formulation variationnelle, exprimée sous forme mixte, dans $\boldsymbol{V}_h \times Q_h$. La coercivité uniforme sur les noyaux discrets (5.20) est une conséquence directe de la coercivité de la forme c sur $\boldsymbol{H}^1_0(\Omega) \times \boldsymbol{H}^1_0(\Omega)$. De façon remarquable (mais est-ce bien surprenant ?), on note que les conditions du lemme de Fortin sont identiques aux conditions (5.4)-(5.5). Ainsi, on peut reprendre les éléments finis MINI ou de Taylor-Hood pour définir les espaces discrets.

Exercice : On reprend l'étude de la condition de stabilité pour le modèle de Stokes réalisée dans le §5.1, lorsque la viscosité ν est "petite". On a obtenu à la proposition 5.1 la constante

$$\alpha'_{St}(\nu) = \frac{\frac{1}{2}\min(\nu^2 (C_{\rm div})^2, 1)}{\max\left(\sqrt{2}\nu (C_{\rm div})^2, C_{\rm div}\,[2 + \nu^2 (C_{\rm div})^2]^{1/2}\right)} > 0.$$

En particulier, $\alpha'_{St}(\nu)$ se comporte comme $\frac{C_{\rm div}}{2\sqrt{2}}\nu^2 = O(\nu^2)$ lorsque $\nu \to 0^+$.
Montrer que, par un choix approprié des paramètres α, β, η dans le cas général (item 3 de la démonstration de la proposition 5.1), on peut trouver une constante $\alpha'_{St}(\nu)$ (strictement positive !) qui se comporte comme $O(\nu)$ lorsque $\nu \to 0^+$.

Notes de cours 6

Introduction à la physique des matériaux électromagnétiques négatifs

Deux domaines de la physique sont en pleine effervescence depuis une dizaine d'années : la plasmonique qui s'intéresse aux ondes électromagnétiques à la surface de particules métalliques aux fréquences optiques, et les métamatériaux, qui sont des matériaux artificiels donnant lieu à des phénomènes électromagnétiques extraordinaires. Dans ces deux domaines, les phénomènes intéressants résultent du changement de signe de la permittivité diélectrique entre le milieu extérieur d'une part (l'air ou le vide) et le métal ou le métamatériau d'autre part. Ce changement de signe est responsable d'une perte de coercivité des équations à résoudre, et nous verrons dans la suite comment y remédier à l'aide des outils qui ont été introduits dans les chapitres précédents.

Dans ce chapitre, nous rappelons tout d'abord les modèles physiques de Drude et de Drude-Lorentz qui permettent d'exprimer la permittivité diélectrique ε d'un métal en fonction de la fréquence. Ces modèles dispersifs permettent d'identifier une plage de fréquences où la partie réelle de ε est négative. Les modèles de métamatériaux sont ensuite rapidement discutés.

Dans une seconde partie, nous expliquons dans des configurations simples pourquoi un changement de signe de la permittivité permet d'une part l'existence de plasmons de surface, et d'autre part le phénomène de réfraction négative.

6.1 Les modèles de Drude et Drude-Lorentz

On rappelle que les ondes électromagnétiques sont modélisées par les équations de Maxwell qui s'écrivent [12]

$$\begin{cases} \mathbf{rot}\,\mathcal{E} + \mu\dfrac{\partial \mathcal{H}}{\partial t} = 0, \\ \mathbf{rot}\,\mathcal{H} - \varepsilon\dfrac{\partial \mathcal{E}}{\partial t} = \mathcal{J} \end{cases}$$

où $\mathcal{E}$ et $\mathcal{H}$ désignent les champs électrique et magnétique et $\mathcal{J}$ la densité de courant. Ce sont des fonctions de la variable d'espace $\boldsymbol{x}$ et du temps t. Les coefficients ε et μ sont la permittivité diélectrique et la perméabilité magnétique. Ils caractérisent le matériau.

On s'intéresse dans la suite au régime périodique établi de fréquence $\omega > 0$, ce qui signifie que toutes les quantités dépendent du temps de la façon suivante :

$$\mathcal{E}(\boldsymbol{x},t) = \mathrm{Re}(\boldsymbol{E}(\boldsymbol{x})e^{-i\omega t}),\ \ \mathcal{H}(\boldsymbol{x},t) = \mathrm{Re}(\boldsymbol{H}(\boldsymbol{x})e^{-i\omega t}) \text{ et } \mathcal{J}(\boldsymbol{x},t) = \mathrm{Re}(\boldsymbol{J}(\boldsymbol{x})e^{-i\omega t}).$$

On obtient ainsi les équations de Maxwell dites *harmoniques* :

$$\left\{ \begin{array}{l} \mathbf{rot}\,\boldsymbol{E} = i\omega\mu\boldsymbol{H}, \\ \mathbf{rot}\,\boldsymbol{H} = -i\omega\varepsilon\boldsymbol{E} + \boldsymbol{J}. \end{array} \right.$$

Par abus de langage, on dira encore que $\boldsymbol{E}$ et $\boldsymbol{H}$ sont les champs électrique et magnétique, et $\boldsymbol{J}$ la densité de courant. Notons que les composantes des champs $\boldsymbol{E}$, $\boldsymbol{H}$ et $\boldsymbol{J}$ sont a priori à valeurs complexes. Souvent, le problème mathématique consiste à trouver $\boldsymbol{E}$ et $\boldsymbol{H}$ qui sont les inconnues alors que $\boldsymbol{J}$ est une donnée.

Dans les études mathématiques usuelles, on suppose le plus souvent que les coefficients ε et μ sont des constantes strictement positives, ou des fonctions de l'espace, bornées et minorées par des constantes strictement positives. Les résultats d'analyse mathématique ou d'analyse numérique exploitent cette hypothèse de positivité. Notre objectif dans ce qui suit est de montrer que sous certaines hypothèses physiques, il peut être pertinent de considérer des coefficients négatifs.

Construction du modèle de Drude-Lorentz Les explications ci-dessous sont largement inspirées du cours de Raphaël Grandin de l'Institut de Physique du Globe de Paris.

Dans le vide, ε et μ sont des constantes strictement positives que l'on note généralement ε_0 et μ_0. Dans un milieu différent du vide, une onde électromagnétique modifie elle-même le milieu dans lequel elle se propage, en mettant en mouvement les charges présentes. Cela produit ce qu'on appelle la polarisation et l'aimantation. Pour décrire correctement ces phénomènes, il faudrait une théorie quantique, mais on peut raisonnablement se contenter d'une théorie qualitative, obtenue à l'aide de principes classiques.

Les modèles que nous allons décrire sont dûs à Paul Drude, physicien allemand (1863-1906), et Hendrik Antoon Lorentz, physicien hollandais (1853-1926).

On considère que les noyaux des atomes sont fixes alors que les électrons de masse m et de charge $(-e)$ sont mobiles. Ces électrons subissent :

— la force de Lorentz due à la présence du champ électromagnétique,

— une force de frottement causée par les collisions avec l'environnement,

— une force de rappel, provoquée par l'interaction avec les noyaux.

Si $\mathcal{U}$ désigne la position d'un électron, son mouvement est donc décrit par une équation du type suivant :

$$m\frac{d^2\mathcal{U}}{dt^2} = -e\left(\mathcal{E} + \frac{d\mathcal{U}}{dt} \wedge \mu\mathcal{H}\right) - m\gamma\frac{d\mathcal{U}}{dt} - K\mathcal{U}$$

où γ et K sont des constantes positives. Si on néglige la partie de la force de Lorentz due au champ magnétique, alors $\mathcal{U}$ est parallèle au champ électrique $\mathcal{E}$. En régime périodique établi de fréquence ω, le mouvement des électrons est lui-même harmonique en temps

$$\mathcal{U}(t) = \mathrm{Re}(\boldsymbol{u}e^{-i\omega t})$$

et l'on obtient finalement la relation suivante entre le mouvement d'un électron autour d'une position $\boldsymbol{x}$ et le champ électrique en $\boldsymbol{x}$:

$$\boldsymbol{u} = \frac{-\frac{e}{m}}{\omega_0^2 - \omega^2 - i\omega\gamma}\boldsymbol{E}$$

où l'on a posé

$$\omega_0 = \sqrt{\frac{K}{m}}.$$

On va tenir compte maintenant du fait que le milieu contient une densité volumique N d'électrons. Ces électrons par leur mouvement créent un courant de densité

$$\mathcal{J} = -Ne\frac{d\mathcal{U}}{dt}, \text{ soit } \boldsymbol{J} = i\omega Ne\boldsymbol{u}.$$

On en déduit que le courant créé par le champ électrique faisant se déplacer les électrons est lui-même proportionnel au champ électrique puisque l'on a :

$$\boldsymbol{J} = \sigma\boldsymbol{E} \text{ avec } \sigma = \frac{-i\omega N\frac{e^2}{m}}{\omega_0^2 - \omega^2 - i\omega\gamma}.$$

Ce courant devient alors une source dans les équations de Maxwell puisqu'on rappelle que (dans le vide)

$$\mathbf{rot}\,\boldsymbol{H} = -i\omega\varepsilon_0\boldsymbol{E} + \boldsymbol{J}.$$

En injectant l'expression de $\boldsymbol{J}$ dans cette équation, on obtient finalement une équation du type

$$\mathbf{rot}\,\boldsymbol{H} = -i\omega\varepsilon\boldsymbol{E}$$

où

$$\varepsilon = \varepsilon_0 - \frac{\sigma}{i\omega}$$

que l'on récrit sous la forme suivante :

$$\varepsilon = \varepsilon_0\left(1 - \frac{\omega_P^2}{\omega^2 - \omega_0^2 + i\omega\gamma}\right)$$

avec

$$\omega_P^2 = \frac{Ne^2}{m\varepsilon_0}.$$

Cette pulsation ω_P est appelée la pulsation plasma (d'où le P).

Commentons le modèle obtenu.

— On remarque tout d'abord que la permittivité diélectrique du matériau, contrairement à celle du vide, dépend de la pulsation ω (et donc de la fréquence $\omega/2\pi$). On dit que le matériau est *dispersif*.
— Lorsque $\gamma > 0$, et donc $\mathrm{Im}(\varepsilon) > 0$, les collisions engendrent des pertes énergétiques dans le matériau (on parle d'effet Joule dans un métal). On dit alors que le matériau est *dissipatif*.
— Le choix des paramètres ω_0 et γ permet de différencier les matériaux. Ainsi par exemple, on considère que dans un métal conducteur, les électrons se déplacent librement de sorte que l'on peut supposer que $K = \omega_0 = 0$. On obtient ainsi le modèle dit de Drude. Au contraire, dans un diélectrique, les électrons bougent peu de sorte que les effets de friction peuvent être négligés et on suppose donc que $\gamma = 0$. On obtient ainsi le modèle dit de Lorentz. Autrement dit, un diélectrique n'est pas dissipatif.
— Dans le cas général où ω_0 et γ ne sont pas nuls, on parle de modèle de Drude-Lorentz.

Nous allons revenir plus en détails sur le cas des métaux.

Le cas des métaux Comme on l'a dit plus haut, la permittivité d'un métal peut être modélisée en tant que fonction de ω par le modèle de Drude suivant

$$\varepsilon = \varepsilon_0 \left(1 - \frac{\omega_P^2}{\omega^2 + i\omega\gamma}\right)$$

avec $\gamma > 0$. On a donc :

$$\mathrm{Re}(\varepsilon) = \varepsilon_0 \left(1 - \frac{\omega_P^2}{\omega^2 + \gamma^2}\right), \quad \mathrm{Im}(\varepsilon) = \varepsilon_0 \frac{\omega_P^2 \gamma}{\omega\,(\omega^2 + \gamma^2)}.$$

Les deux courbes correspondantes sont représentées sur la figure 6.1. Classiquement, en électronique, les métaux sont exploités à basse fréquence, là où la partie imaginaire de ε domine. C'est le classique effet Joule. Si ω est petit devant γ, on utilise l'approximation suivante de ε sous la forme :

$$\varepsilon = \varepsilon_0 \left(1 - \frac{\omega_P^2}{i\omega\gamma}\right) = \varepsilon_0 + i\frac{\sigma}{\omega}$$

où la conductivité σ est réelle. Ce qui est plus original et a été permis par des avancées expérimentales récentes, c'est d'exploiter le métal à plus faute fréquences, dans le domaine des fréquences optiques. Plus précisément, la plasmonique correspond à une plage de fréquences telles que ω est inférieur à ω_P mais grand devant γ, de sorte que ε a une partie réelle négative et une partie imaginaire négligeable. Ceci est possible pour les métaux rares tels que l'or et l'argent. Par exemple pour l'argent, on a $\gamma \approx 0,113 PHz$ et $\omega_P = 13,3 PHz$ (PHz signifie pétahertz, soit 10^{15} Hz). La longueur d'onde associé à la pulsation ω_P est $\lambda_P \approx 142$nm. Cette longueur d'onde se trouve dans l'ultraviolet, juste en-dessous des longueurs d'onde du domaine visible.

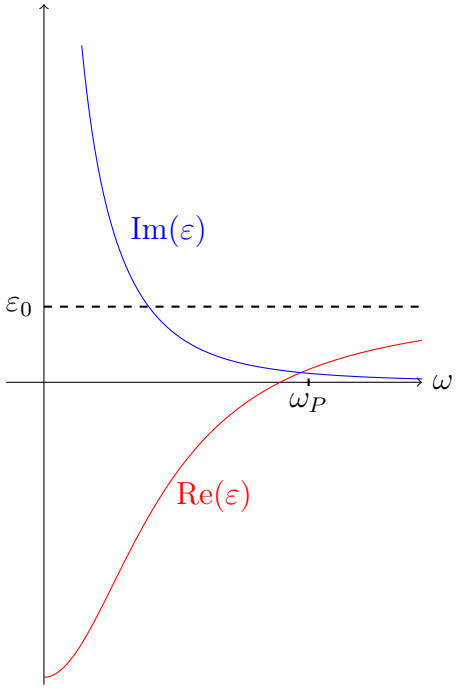

FIGURE 6.1 – Parties réelles et imaginaires de la permittivité d'un métal en fonction de la fréquence

On peut dans ce cas utiliser l'approximation suivante, que l'on appelle modèle de Drude non dissipatif :

$$\varepsilon = \varepsilon_0 \left(1 - \frac{\omega_P^2}{\omega^2}\right). \tag{6.1}$$

Les métamatériaux Nous avons vu qu'il était raisonnable dans certaines conditions de considérer qu'un métal a une permittivité réelle négative. Une question naturelle, étant donnée la symétrie apparente des équations de Maxwell, est la suivante : existe-t-il des matériaux dont la perméabilité magnétique μ est réelle négative ? A première vue, cela semble impossible car il n'existe pas de charges magnétiques ponctuelles comparables aux charges électriques que sont les électrons. Idéalement, on souhaiterait même trouver des matériaux où à la fois ε et μ sont réels et négatifs. En effet, il a été montré théoriquement [22] que cela donnerait lieu à des phénomènes extraordinaires comme la réfraction négative, et permettrait par exemple de réaliser de façon simple des lentilles parfaites [21]. L'idée a été de concevoir des matériaux micro-structurés, qui puissent être modélisés lorsque la longueur d'onde est grande devant la taille des micro-structures, comme des matériaux avec constantes *effectives* ε et μ réelles négatives. Ceci a été en effet réalisé dans le domaine des micro-ondes en assemblant de petits résonateurs. Aujourd'hui, de nombreux physiciens proposent de nouvelles idées pour réaliser des matériaux étranges. Les mathématiciens quant à eux s'intéressent à la justification du processus d'*homogénéisation* [7]. Signalons finalement que pour respecter les grands principes physiques tels que la causalité, ε et μ doivent être de la forme suivante (ou d'une forme similaire avec

plusieurs pôles) :

$$\varepsilon = \varepsilon_0 \left(1 - \frac{\Omega_E^2}{\omega^2 - \omega_E^2 + i\omega\gamma_E}\right) \text{ et } \mu = \mu_0 \left(1 - \frac{\Omega_H^2}{\omega^2 - \omega_H^2 + i\omega\gamma_H}\right)$$

avec $\gamma_{E/H} > 0$.

6.2 Deux effets du changement de signe des constantes électromagnétiques

6.2.1 Les plasmons de surface

Considérons un métal dont la permittivité est notée ε_m (m pour métal) et qui occupe le demi-espace $x_2 < 0$. On suppose que le demi-espace $x_2 > 0$ est vide. Un plasmon de surface est une onde électromagnétique qui se propage parallèlement à la surface, soit par exemple dans la direction x_1, et qui est confinée au voisinage de la surface $x_2 = 0$. Cela nous conduit en particulier à rechercher des solutions des équations de Maxwell (sans terme source $\boldsymbol{J}$) indépendantes de x_3. Or il est bien connu (et facile à vérifier) que les équations de Maxwell se découplent alors en deux sous-systèmes correspondant respectivement aux ondes Transverse Électriques (telles que $H_1 = H_2 = E_3 = 0$) et aux ondes Transverses Magnétiques (telles que $E_1 = E_2 = H_3 = 0$). Le plasmon est une onde Transverse Magnétique, donc nous allons nous concentrer sur ce cas.

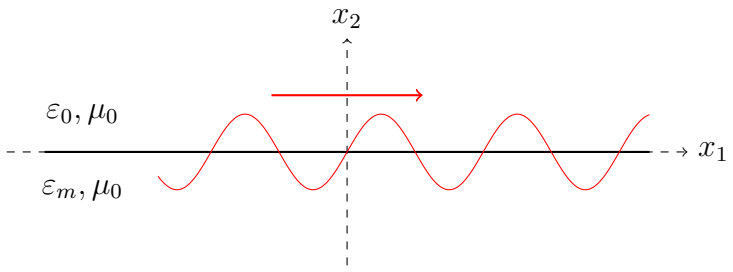

FIGURE 6.2 – Configuration géométrique pour l'étude du plasmon de surface

On peut montrer que le couple $(\boldsymbol{E}, \boldsymbol{H})$ donné par

$$\boldsymbol{H} = \begin{pmatrix} 0 \\ 0 \\ u \end{pmatrix} \text{ et } \boldsymbol{E} = \frac{-1}{i\omega\varepsilon} \begin{pmatrix} \partial_2 u \\ -\partial_1 u \\ 0 \end{pmatrix} \tag{6.2}$$

est solution des équations de Maxwell (avec $\boldsymbol{J} = 0$) si et seulement si la fonction scalaire u vérifie l'équation

$$\text{div}\left(\frac{1}{\varepsilon}\nabla u\right) + \omega^2 \mu u = 0. \tag{6.3}$$

Ici μ est constante, égale à μ_0, et ε est une fonction de x_2 (et de ω) telle que

$$\varepsilon(x_2) = \begin{cases} \varepsilon_m(\omega) & \text{si } x_2 < 0, \\ \varepsilon_0 & \text{si } x_2 > 0. \end{cases}$$

Dire que l'on cherche un onde guidée par la surface revient alors à chercher u de la forme suivante

$$u(x_1, x_2) = \hat{u}(x_2)e^{i\beta x_1}, \text{ avec } \hat{u} \in L^2(\mathbb{R}) \text{ et } \beta \in \mathbb{R}. \tag{6.4}$$

La condition $\hat{u} \in L^2(\mathbb{R})$ va imposer le confinement. Si on réintroduit la dépendance temporelle qui avait été éliminée dans les équations, on fait apparaître le terme

$$e^{i(\beta x_1 - \omega t)}$$

qui correspond à une propagation dans la direction x_1 (si $\omega, \beta > 0$) à la vitesse ω/β. En injectant la forme (6.4) dans l'équation (6.3), on obtient l'équation différentielle suivante portant sur $\hat{u}$:

$$\frac{-\beta^2}{\varepsilon(x_2)}\hat{u} + \frac{d}{dx_2}\left(\frac{1}{\varepsilon(x_2)}\frac{d\hat{u}}{dx_2}\right) + \omega^2\mu_0\hat{u} = 0$$

qui doit être vérifiée, au sens des distributions, sur $\mathbb{R}$. Cela signifie que $\hat{u}$ doit vérifier d'une part ces équations différentielles à coefficients constants de part de d'autre de $x_2 = 0$:

$$\begin{aligned} &\frac{d^2\hat{u}}{dx_2^2} + (\omega^2\varepsilon_0\mu_0 - \beta^2)\hat{u} = 0 \quad \text{si } x_2 > 0, \\ &\frac{d^2\hat{u}}{dx_2^2} + (\omega^2\varepsilon_m\mu_0 - \beta^2)\hat{u} = 0 \quad \text{si } x_2 < 0, \end{aligned}$$

et d'autre part, les relations de continuité suivantes en $x_2 = 0$:

$$\hat{u}(0^-) = \hat{u}(0^+) \text{ et } \frac{1}{\varepsilon_m}\frac{d\hat{u}}{dx_2}(0^-) = \frac{1}{\varepsilon_0}\frac{d\hat{u}}{dx_2}(0^+). \tag{6.5}$$

On remarque que si ε_m est négatif, la dérivée de $\hat{u}$ change de signe de part et d'autre de 0, et c'est ce qui va permettre à u de décroitre de part et d'autre de l'interface. Plus précisément, pour que la condition $\hat{u} \in L^2(\mathbb{R})$ soit satisfaite (avec $\hat{u}$) non identiquement nul), il est nécessaire que les deux conditions suivantes soient vérifiées :

$$\begin{aligned} &\omega^2\varepsilon_0\mu_0 - \beta^2 =: -\gamma_0^2 < 0, \\ &\omega^2\varepsilon_m\mu_0 - \beta^2 =: -\gamma_m^2 < 0. \end{aligned} \tag{6.6}$$

Si c'est le cas, en choisissant γ_0 et γ_m strictement positifs, on trouve que $\hat{u}$ est nécessairement de la forme suivante :

$$\hat{u}(x_2) = \begin{cases} A_0 e^{-\gamma_0 x_2} & \text{si } x_2 > 0, \\ A_m e^{\gamma_m x_2} & \text{si } x_2 < 0, \end{cases}$$

où A_0 et A_m sont des constantes complexes à déterminer. En injectant la forme ci-dessus dans les conditions de continuité (6.5), on obtient un système linéaire homogène de deux équations à deux inconnues en (A_0, A_m) :

$$\left| \begin{array}{l} A_0 = A_m \\ -\dfrac{\gamma_0}{\varepsilon_0} A_0 = \dfrac{\gamma_m}{\varepsilon_m} A_m \end{array} \right. \tag{6.7}$$

qui n'a des solutions non triviales que si

$$-\frac{\gamma_0}{\varepsilon_0} = \frac{\gamma_m}{\varepsilon_m}.$$

On voit immédiatement que ceci n'est possible que si ε_m est négatif. On comprend donc pourquoi la plasmonique ne s'est développée que lorsqu'on a pu observer la surface du métal aux fréquences optiques, et donc dans une gamme de fréquences où la permittivité diélectrique du métal est (approximativement) réelle et négative.

Si $\varepsilon_m < 0$, en injectant l'expression (6.6) de γ_0 et γ_m dans la relation ci-dessus, on obtient finalement la relation de dispersion suivante, qui nous donne le carré de la vitesse de l'onde en fonction des coefficients du milieu :

$$\frac{\omega^2}{\beta^2} = \frac{1}{\mu_0} \left(\frac{1}{\varepsilon_0} + \frac{1}{\varepsilon_m} \right), \tag{6.8}$$

et qui fournit une seconde condition d'existence des plasmons de surface. En effet, on voit que le terme de droite de l'égalité (6.8) doit être positif, et donc que l'on doit avoir $\varepsilon_m + \varepsilon_0 < 0$.

Remarque 6.1 *Pour aller plus loin, on doit préciser la relation qui donne ε_m en fonction de ω. Si l'on adopte le modèle de Drude non dissipatif* (6.1)*, on trouve que la condition $\varepsilon_m + \varepsilon_0 < 0$ se traduit en fréquence par la condition $\omega < \omega_{SP}$ où*

$$\omega_{SP} = \frac{\omega_P}{\sqrt{2}}$$

(SP pour Surface Plasmon). On peut ensuite représenter β en fonction de ω comme sur la figure 6.3. On observe une asymptote en ω_{SP}, qui se traduit par le fait que la longueur d'onde du plasmon de surface tend vers 0 lorsque ω tend vers ω_{SP}. C'est aussi un phénomène très inhabituel de pouvoir générer de toutes petites longueurs d'onde sans augmenter la fréquence. On peut alors envisager d'imager de très petits objets.

6.2.2 La réfraction négative

Nous allons maintenant expliquer pourquoi on peut observer un phénomène de réfraction négative à l'interface entre un diélectrique (ou du vide) et un métamatériau. On va à nouveau considérer une expérience où la géométrie et le champ électromagnétique sont indépendants de la variable x_3. Plus précisément, on considère cette fois que le demi-espace $x_2 < 0$ est vide et que le demi-espace $x_2 > 0$ est occupé par un métamatériau tel que (à la fréquence considérée)

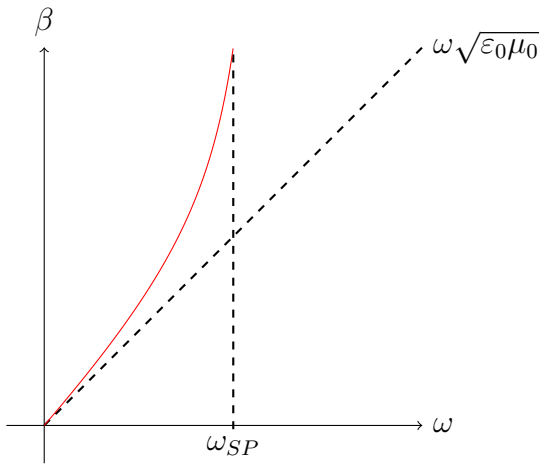

FIGURE 6.3 – Courbe de dispersion du plasmon de surface pour le modèle de Drude non dissipatif

sa permittivité ε_m (m pour métamatériau) et sa perméabilité magnétique μ_m sont toutes deux des réels négatifs. On se souviendra le moment venu que l'on a négligé les effets de dissipation, qui se manifesteraient par une partie imaginaire de ε_m et/ou de μ_m strictement positive.

Tout comme dans le paragraphe précédent, on peut alors se restreindre au calcul d'ondes Transverses Magnétiques de la forme (6.2) où u est une fonction scalaire à valeurs complexes satisfaisant (6.3). Contrairement au paragraphe précédent, nous nous intéressons ici à l'effet de l'interface sur des ondes planes qui se propagent dans une direction non parallèle à celle-ci. Ainsi par exemple, dans le vide, une onde plane correspond à une solution u de la forme

$$u(x_1, x_2) = A_0 e^{i(k_1^0 x_1 + k_2^0 x_2)}$$

avec A_0 une amplitude complexe et (k_1^0, k_2^0) un vecteur d'onde de composantes réelles, qui donne la direction de propagation de l'onde. Pour qu'une telle fonction satisfasse (6.3) dans le vide, il faut et il suffit que la relation de dispersion suivante soit vérifiée :

$$(k_1^0)^2 + (k_2^0)^2 = \omega^2 \varepsilon_0 \mu_0. \tag{6.9}$$

Une onde qui se propage vers l'interface doit avoir une composante k_2^0 positive.

Dans le métamatériau, on peut également calculer des ondes planes. On a cette fois

$$u(x_1, x_2) = A_m e^{i(k_1^m x_1 + k_2^m x_2)}$$

avec A_m une amplitude complexe et (k_1^m, k_2^m) un vecteur d'onde de composantes réelles vérifiant la relation de dispersion

$$(k_1^m)^2 + (k_2^m)^2 = \omega^2 \varepsilon_m \mu_m. \tag{6.10}$$

A ce niveau, le fait que ε_m et μ_m soient négatifs ne semblent pas avoir d'importance puisque c'est leur produit, qui est lui positif, qui apparait dans la relation de dispersion.

Classiquement, l'étude de la réfraction par une interface plane consiste à calculer les solutions u du problème global, dans tout le plan, de la forme

$$u(x_1,x_2)=\begin{cases} e^{i(k_1^0x_1+k_2^0x_2)}+Re^{i(k_1^0x_1-k_2^0x_2)} & \text{pour } x_2<0,\\ Te^{i(k_1^mx_1+k_2^mx_2)} & \text{pour } x_2>0,\end{cases}$$

avec $k_2^0>0$. Dans le vide, la solution est la somme d'une onde qui se propage vers l'interface (appelée onde incidente) et d'une onde réfléchie. Dans le métamatériau, la solution correspond à l'onde transmise. Les coefficients R et T sont des nombres complexes appelés coefficients de réflexion et de transmission. La fonction u ci-dessus vérifie les équations souhaitées de part et d'autre de l'interface dès lors que les relations de dispersion (6.9) et (6.10) sont satisfaites. Pour que l'équation (6.3) soit globalement vérifiée, il faut en plus assurer les conditions de raccord à l'interface qui s'écrivent :

$$u(x_1,0^-)=u(x_1,0^+) \text{ et } \frac{1}{\varepsilon_0}\frac{\partial u}{\partial x_2}(x_1,0^-)=\frac{1}{\varepsilon_m}\frac{\partial u}{\partial x_2}(x_1,0^+),\quad \forall x_1\in\mathbb{R}.$$

Ceci n'est possible que si $k_1^0=k_1^m$, que nous pouvons noter k_1 dans la suite et choisir, sans perte de généralité tel que $k_1\geq 0$. Les conditions de transmission conduisent alors au système linéaire suivant, pour les coefficients R et T :

$$\left|\begin{array}{l} 1+R=T\\ 1-R=\nu T\end{array}\right. \quad \text{où } \nu=\frac{k_2^m\varepsilon_0}{k_2^0\varepsilon_m}. \tag{6.11}$$

On vérifie que ce système est solvable si et seulement si $\nu\neq -1$.

Tous ces calculs sont très classiques lorsque l'on s'intéresse à une interface entre deux matériaux positifs (c'est-à-dire à constantes ε et μ positives). Le phénomène de réfraction correspond au changement de direction de l'onde, entre l'onde incidente et l'onde transmise. L'onde incidente a pour direction le vecteur (k_1,k_2^0) avec $k_1>0$ et $k_2^0>0$. Dans un matériau positif, l'onde transmise quant à elle a pour direction le vecteur (k_1,k_2^m), où l'on a choisi $k_2^m>0$ puisque l'onde transmise s'éloigne de l'interface. Ces deux vecteurs (k_1,k_2^0) et (k_1,k_2^m) ont leur deuxième composante k_2 positive : on parle de réfraction positive.

Dans le cas d'un matériau négatif, on observe une réfraction négative. Comment cela est-il possible ? On peut remarquer que dans la relation (6.10), étant donné $k_1^m=k_1$, il existe deux valeurs possibles de k_2^m opposées l'une de l'autre. Et nous allons expliquer pourquoi dans le cas d'un matériau négatif, il faut choisir k_2^m négatif. Une façon de s'en convaincre est, comme on l'a annoncé au début du paragraphe, de réintroduire une faible dissipation dans le métamatériau. Autrement dit, on suppose par exemple que la permittivité est égale à $\varepsilon_m+i\delta_m$ avec $\delta_m>0$ un paramètre que l'on va faire tendre vers 0. La relation de dispersion s'écrit alors

$$k_1^2+(k_2^m)^2=\omega^2(\varepsilon_m+i\delta_m)\mu_m.$$

Etant donné k_1, cette équation en k_2^m admet deux solutions complexes opposées l'une de l'autre :

$$\pm\sqrt{\omega^2(\varepsilon_m+i\delta_m)\mu_m-k_1^2}.$$

La solution physiquement acceptable est celle qui conduit à un champ exponentiellement décroissant loin de l'interface (et non exponentiellement croissant). Autrement dit, on veut $\mathrm{Im}(k_2^m) > 0$. On vérifie aisément, en remarquant que le nombre sous la racine a une partie réelle positive et une partie imaginaire négative, que ceci entraîne $\mathrm{Re}(k_2^m) < 0$, et donne à la limite $\delta_m = 0$ une valeur négative de k_2^m comme on l'avait annoncé.

Ceci est surprenant : comment une onde qui s'éloigne de l'interface peut-elle avoir un vecteur d'onde qui pointe vers l'interface ? L'explication est que dans le métamatériau, une onde plane a une vitesse de phase opposée à sa vitesse de groupe. Autrement dit, elle se déplace comme Michael Jackson avec sa moon-walk, et avance en reculant. Donc l'onde transmise se déplace en réalité dans la direction $(-k_1, -k_2^m)$. Elle s'éloigne bien de l'interface mais a changé de direction parallèlement à l'interface, c'est la réfraction négative.

Pour conclure, on peut noter que dans le cas idéal où $\varepsilon_m = -\varepsilon_0$ et $\mu_m = -\mu_0$, $k_2^m = -k_2^0$ et le coefficient ν qui apparaît dans le système (6.11) est égal à 1. La transmission est parfaite dans ce cas et la réflexion nulle.

Notes de cours 7

Etude des problèmes de transmission avec changement de signe : premiers résultats

Nous avons vu au chapitre précédent qu'il pouvait être pertinent et intéressant de considérer des problèmes d'électromagnétisme où les coefficients ε et/ou μ changent de signe dans le milieu. Dans ce chapitre et le suivant, nous allons nous intéresser à un problème modèle très simple de ce type, faisant intervenir un coefficient variable noté σ qui change de signe (mais sans s'annuler) dans le domaine. Ce problème écrit sous forme faible fait intervenir une forme bilinéaire qui n'est pas coercive, à cause du changement de signe de σ. L'objet de ce chapitre est de montrer à l'aide de la T-coercivité que, sous certaines conditions sur les coefficients et/ou sur la géométrie, le problème est bien posé. On verra aussi brièvement comment assurer la convergence d'une méthode de discrétisation.

7.1 Description du problème de transmission

On note Ω un ouvert borné de $\mathbb{R}^{\mathtt{d}}$, où $\mathtt{d} > 1$ (en pratique on s'intéressera aux cas $\mathtt{d} = 2$ et $\mathtt{d} = 3$). On suppose qu'il existe une interface Σ (une courbe si $\mathtt{d} = 2$ et une surface si $\mathtt{d} = 3$) telle que $\Omega = \Sigma \cup \Omega_1 \cup \Omega_2$, où Ω_1 et Ω_2 sont deux ouverts disjoints tels que $\partial\Omega_1 \cap \partial\Omega_2 = \Sigma$. On suppose enfin que Ω, Ω_1 et Ω_2 ont chacun une frontière lipschitzienne, et on note $\Gamma_j = \partial\Omega_j \cap \partial\Omega$, $j = 1, 2$ (on peut avoir $\Gamma_1 = \emptyset$ ou $\Gamma_2 = \emptyset$, voir l'exemple de droite de la figure 7.1 pour lequel $\Gamma_2 = \emptyset$). On introduit alors σ une fonction telle que

$$\sigma \in L^\infty(\Omega) \text{ et } \frac{1}{\sigma} \in L^\infty(\Omega), \text{ avec } \sigma(\boldsymbol{x}) > 0 \text{ p.p. } \boldsymbol{x} \in \Omega_1 \text{ et } \sigma(\boldsymbol{x}) < 0 \text{ p.p. } \boldsymbol{x} \in \Omega_2, \tag{7.1}$$

$$\tag{7.2}$$

et on pose

$$a(v, w) = \int_\Omega \sigma \nabla v \cdot \nabla w \, d\Omega \quad \text{et} \quad \ell(w) = \int_\Omega f w \, d\Omega. \tag{7.3}$$

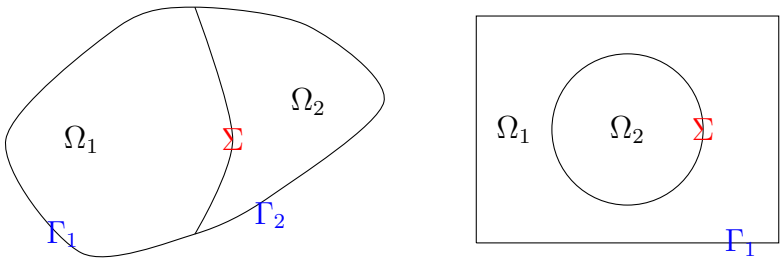

FIGURE 7.1 – Exemples de domaines considérés

On cherche alors à résoudre, pour $f \in L^2(\Omega)$, le problème suivant :

$$(P) \quad \begin{cases} \text{Trouver } u \in H_0^1(\Omega) \text{ tel que} \\ \forall w \in H_0^1(\Omega), \quad a(u,w) = \ell(w). \end{cases}$$

Le problème (P) peut être reformulé de façon équivalente :

$$\begin{cases} \text{Trouver } u \in H_0^1(\Omega) \text{ tel que} \\ -\mathrm{div}\,(\sigma \nabla u) = f \text{ dans } \Omega \end{cases}$$

où l'équation aux dérivées partielles peut s'écrire

$$-\sum_{j=1}^{\mathtt{d}} \frac{\partial}{\partial x_j}\left(\sigma \frac{\partial u}{\partial x_j}\right) = f$$

et est à prendre au sens des distributions.

Compte tenu des hypothèses sur la fonction σ, la forme bilinéaire $a(\cdot,\cdot)$ n'est pas coercive. En effet, on peut trouver une fonction $u^+ \in H_0^1(\Omega)$ et une fonction $u^- \in H_0^1(\Omega)$ telles que $a(u^+,u^+) > 0$ et $a(u^-,u^-) < 0$.[1] Il suffit alors de remarquer que la fonction $t \mapsto a(tu^+ + (1-t)u^-, tu^+ + (1-t)u^-)$ qui est continue s'annule nécessairement pour une valeur de t entre 0 et 1. Autrement dit, il existe $u \in H_0^1(\Omega)$, $u \neq 0$, tel que $a(u,u) = 0$.

Dans la suite, pour $v \in L^2(\Omega)$, on note $v_j = v_{|\Omega_j}$ pour $j = 1, 2$.

On s'intéressera plus particulièrement au cas où σ est une fonction constante par morceaux, telle que

$$\sigma(\boldsymbol{x}) = \sigma_j \quad \text{p.p. } \boldsymbol{x} \in \Omega_j, j = 1,2. \tag{7.4}$$

D'après (7.1), on a dans ce cas $\sigma_1 > 0$ et $\sigma_2 < 0$, et on appelle $\kappa_\sigma = \sigma_2/\sigma_1 < 0$ le contraste.

1. C'est évident si σ est une fonction localement continue au voisinage d'un point où elle prend une valeur strictement positive et au voisinage d'un point où elle prend une valeur strictement négative, mais c'est également vrai dans le cas général (en considérant des points de Lebesgue de la fonction).

Exercice : Sous l'hypothèse (7.4), montrer que le couple (u_1, u_2) vérifie

$$\left| \begin{array}{l} -\Delta u_j = \dfrac{f}{\sigma_j} \text{ dans } \Omega_j, j = 1, 2, \\ u_j = 0 \text{ sur } \Gamma_j, j = 1, 2, \\ u_1 = u_2 \text{ sur } \Sigma, \\ \sigma_1 \dfrac{\partial u_1}{\partial n_\Sigma} = \sigma_2 \dfrac{\partial u_2}{\partial n_\Sigma} \text{ sur } \Sigma, \end{array} \right. \tag{7.5}$$

où $\partial/\partial n_\Sigma$ désigne la dérivée normale sur Σ, la normale $\boldsymbol{n}_\Sigma$ étant orientée par exemple de Ω_1 vers Ω_2.

Observons que la dernière équation de (7.5) peut se réécrire comme suit

$$\frac{\partial u_1}{\partial n_\Sigma} = \kappa_\sigma \frac{\partial u_2}{\partial n_\Sigma} \text{ sur } \Sigma.$$

Toute la difficulté vient du fait que κ_σ est négatif.

7.2 La T-coercivité

La forme bilinéaire $a(\cdot, \cdot)$ n'étant pas coercive, nous ne pouvons pas espérer montrer le caractère bien posé du problème (P) à l'aide du théorème de Lax-Migram. Nous allons utiliser le concept de T-coercivité, déjà introduite dans le §2.3.4. On rappelle les résultats principaux ci-dessous.

Définition 7.1 *Soit V un espace de Hilbert et $a(\cdot, \cdot)$ une forme bilinéaire continue de $V \times V$ (dans $\mathbb{R}$ ou $\mathbb{C}$). On dit que la forme bilinéaire $a(\cdot, \cdot)$ est* T*-coercive sur V s'il existe $\mathtt{T} \in \mathcal{L}(V)$ bijective telle que la forme bilinéaire $(v, w) \mapsto a(v, \mathtt{T}w)$ est coercive sur V.*

On a alors le résultat suivant :

Théorème 7.2 *Soit V un espace de Hilbert, $a(\cdot, \cdot)$ une forme bilinéaire continue sur $V \times V$ et ℓ une forme linéaire continue sur V. On suppose de plus que la forme bilinéaire $a(\cdot, \cdot)$ est* T*-coercive sur V. Alors le problème*

$$\left\{ \begin{array}{l} \text{Trouver } u \in V \text{ tel que} \\ \forall w \in V, \quad a(u, w) = \ell(w), \end{array} \right. \tag{7.6}$$

est bien posé.

Démonstration : La preuve est immédiate. En effet, le problème (7.6) est équivalent au problème suivant

$$\left\{ \begin{array}{l} \text{Trouver } u \in V \text{ tel que} \\ \forall v \in V, \quad a(u, \mathtt{T}v) = \ell(\mathtt{T}v), \end{array} \right.$$

qui est bien posé d'après le théorème de Lax-Milgram. ■

Supposons qu'on soit dans le cadre du théorème précédent. Que peut-on en déduire concernant l'approximation du problème ? La situation favorable est décrite dans le théorème suivant :

Théorème 7.3 *Sous les hypothèses du théorème 7.2, on note V_h un sous-espace de dimension finie de V tel que*

$$\mathtt{T}(V_h) \subset V_h. \tag{7.7}$$

Alors le problème approché

$$\begin{cases} \text{Trouver } u_h \in V_h \text{ tel que} \\ \forall w_h \in V_h, \quad a(u_h, w_h) = \ell(w_h), \end{cases} \tag{7.8}$$

est bien posé, et il existe une constante $C > 0$, indépendante du choix de V_h, telle que

$$\|u - u_h\|_V \leq C \inf_{v_h \in V_h} \|u - v_h\|_V \tag{7.9}$$

où u désigne la solution du problème (7.6) *et u_h celle du problème* (7.8)*.*

Démonstration : On peut tout d'abord remarquer que nécessairement $\mathtt{T}(V_h) = V_h$, puisque $\mathtt{T}$ est bijective. Par conséquent, le problème suivant est équivalent à (7.8) :

$$\begin{cases} \text{Trouver } u_h \in V_h \text{ tel que} \\ \forall v_h \in V_h, \quad a(u_h, \mathtt{T}v_h) = \ell(\mathtt{T}v_h). \end{cases} \tag{7.10}$$

Il suffit alors de raisonner comme dans la démonstration du lemme de Céa (théorème 3.4) : on a

$$a(u_h, \mathtt{T}(u_h - v_h)) = 0 \quad \forall v_h \in V_h,$$

d'où

$$a(u - u_h, \mathtt{T}(u - u_h)) = a(u - u_h, \mathtt{T}(u - v_h)).$$

On en déduit que pour tout $v_h \in V_h$

$$\underline{\alpha}\|u - u_h\|_V^2 \leq |a(u - u_h, \mathtt{T}(u - u_h))| = |a(u - u_h, \mathtt{T}(u - v_h))| \leq |||a||| \, |||\mathtt{T}||| \, \|u - u_h\|_V \|u - v_h\|_V$$

où $\underline{\alpha} > 0$ est la constante de $\mathtt{T}$-coercivité. Le résultat s'en déduit avec $C = |||a||| \, |||\mathtt{T}|||/\underline{\alpha}$. ■

L'objectif dans la suite du chapitre est de construire des applications linéaires continues $\mathtt{T}$ telles que la forme bilinéaire $a(\cdot, \cdot)$ définie en (7.3) soit $\mathtt{T}$-coercive. Une remarque fondamentale est que $\mathtt{T}$ n'est pas unique. D'une part, il est clair que si $\mathtt{T}$ convient, alors $\lambda\mathtt{T}$ convient également pour tout $\lambda \neq 0$ (avec $|\lambda|\underline{\alpha}$ la constante de $\mathtt{T}$-coercivité). D'autre part, on va proposer dans la suite plusieurs catégories possibles d'applications $\mathtt{T}$.

7.3 Le cas des contrastes grands ou petits

On cherche donc $\mathtt{T} \in \mathcal{L}(H^1_0(\Omega))$ bijective pour laquelle

$$\exists \underline{\alpha} > 0 \quad \text{tel que} \quad \forall u \in H^1_0(\Omega) \quad \left| \int_\Omega \sigma \nabla u \cdot \nabla(\mathtt{T}u)\, d\Omega \right| \geq \underline{\alpha} \int_\Omega |\nabla u|^2\, d\Omega. \tag{7.11}$$

On remarque qu'on a utilisé ici le fait que la semi-norme $|\cdot|_{1,\Omega}$ est une norme sur $H^1_0(\Omega)$ grâce à l'inégalité de Poincaré.

On se place sous l'hypothèse (7.4), c'est-à-dire que σ prend deux valeurs constantes de signes opposés de part et d'autre de Σ. La difficulté venant du changement de signe de σ, on souhaiterait que l'application $\mathtt{T}$ combatte en quelque sorte ce changement de signe. Un choix tentant serait de poser $\mathtt{T}u = u$ là où σ est positif, et $\mathtt{T}u = -u$ là où σ est négatif. Mais malheureusement, ceci n'est pas possible puisque $\mathtt{T}u$ ainsi défini n'appartient pas en général à $H^1(\Omega)$. En effet, si u ne s'annule pas sur l'interface Σ, alors $\mathtt{T}u$ construit ainsi aurait un saut de trace à travers Σ, égal à $2u$. L'idée est alors de corriger cette définition pour rétablir l'appartenance de $\mathtt{T}u$ à $H^1(\Omega)$ (et à $H^1_0(\Omega)$).

On notera dans la suite pour $j = 1, 2$

$$H^1_{0,\Gamma_j}(\Omega_j) = \left\{ v_j \in H^1(\Omega_j) \text{ tel que } v_{j|\Gamma_j} = 0 \right\}.$$

Notons que, si $\Gamma_j = \emptyset$, alors $H^1_{0,\Gamma_j}(\Omega_j) = H^1(\Omega_j)$. Si $\Gamma_j \neq \emptyset$, les fonctions de $H^1_{0,\Gamma_j}(\Omega_j)$, qui s'annulent par définition sur Γ_j, ne s'annulent pas, en général, sur Σ. Enfin, on peut vérifier que l'on a l'égalité $H^1_{0,\Gamma_j}(\Omega_j) = \left\{ v \in L^2(\Omega_j) \text{ tel que } \exists \tilde{v} \in H^1_0(\Omega),\ v = \tilde{v}_{|\Omega_j} \right\}$.

Lemme 7.4 *Soit* $\mathtt{R}$ *une application linéaire continue de* $H^1_{0,\Gamma_1}(\Omega_1)$ *dans* $H^1_{0,\Gamma_2}(\Omega_2)$ *telle que*

$$\forall u_1 \in H^1_{0,\Gamma_1}(\Omega_1) \quad (\mathtt{R}u_1)|_\Sigma = (u_1)|_\Sigma \tag{7.12}$$

Alors $\mathtt{T}$ *défini pour tout* $u \in H^1_0(\Omega)$ *par*

$$\mathtt{T}u = \begin{cases} u_1 & \text{dans } \Omega_1 \\ -u_2 + 2\mathtt{R}u_1 & \text{dans } \Omega_2 \end{cases} \tag{7.13}$$

appartient à $\mathcal{L}(H^1_0(\Omega))$*, et vérifie*

$$(\mathtt{T} \circ \mathtt{T})u = u \quad \forall u \in H^1_0(\Omega).$$

C'est donc en particulier une application bijective.

Démonstration : On peut vérifier que les traces de $\mathtt{T}u$ de part et d'autre de Σ coïncident bien puisqu'on a sur Σ :

$$u_1 = -u_2 + 2\mathtt{R}u_1 = -u_1 + 2u_1,$$

où l'on a utilisé (7.12) et la continuité de u à travers Σ (voir le théorème 1.36). Il est alors facile de vérifier que $\mathtt{T}u \in H^1_0(\Omega)$. La continuité de $\mathtt{T}$ résulte ensuite de la continuité de $\mathtt{R}$. Enfin, un calcul simple montre que

$$(\mathtt{T}\circ\mathtt{T})u = \begin{cases} u_1 & \text{dans } \Omega_1 \\ -(\mathtt{T}u)_2 + 2\mathtt{R}u_1 = -(-u_2 + 2\mathtt{R}u_1) + 2\mathtt{R}u_1 = u_2 & \text{dans } \Omega_2 \end{cases}$$

et on a bien $(\mathtt{T}\circ\mathtt{T})u = u$ comme annoncé. ■

Remarque 7.5 *Il existe de nombreuses applications* $\mathtt{R}$ *remplissant les conditions du lemme. On peut penser par exemple au relèvement harmonique, qui consiste à poser* $\mathtt{R}u = \phi$ *où* ϕ *est l'unique solution du problème suivant :*

$$\Delta\phi = 0 \text{ dans } \Omega_2, \quad \phi = 0 \text{ sur } \Gamma_2 \quad \text{et} \quad \phi = u_1 \text{ sur } \Sigma.$$

Nous allons maintenant chercher à voir sous quelles conditions $a(\cdot,\cdot)$ est $\mathtt{T}$-coercive, pour une application $\mathtt{T}$ de la forme (7.13).

Théorème 7.6 *On suppose qu'il existe* $\mathtt{R} \in \mathcal{L}(H^1_{0,\Gamma_1}(\Omega_1), H^1_{0,\Gamma_2}(\Omega_2))$*, vérifiant* (7.12)*, et telle que pour une constante* $C > 0$*, on a :*

$$\int_{\Omega_2} |\nabla(\mathtt{R}u_1)|^2 \, d\Omega \le C \int_{\Omega_1} |\nabla u_1|^2 \, d\Omega, \quad \forall u_1 \in H^1_{0,\Gamma_1}(\Omega_1). \tag{7.14}$$

Alors la forme bilinéaire

$$a \ : \ (v,w) \mapsto \int_\Omega \sigma \nabla v \cdot \nabla w \, d\Omega$$

est $\mathtt{T}$*-coercive sur* $H^1_0(\Omega)$ *dès que le contraste* $\kappa_\sigma = \sigma_2/\sigma_1$ *vérifie*

$$\kappa_\sigma > -\frac{1}{\|\mathtt{R}\|^2} \tag{7.15}$$

où l'on a posé

$$\|\mathtt{R}\| = \sup_{u_1 \in H^1_{0,\Gamma_1}(\Omega_1), u_1 \neq 0} \frac{\sqrt{\displaystyle\int_{\Omega_2} |\nabla(\mathtt{R}u_1)|^2 \, d\Omega}}{\sqrt{\displaystyle\int_{\Omega_1} |\nabla u_1|^2 \, d\Omega}}.$$

Remarque 7.7 *Lorsque* $\Gamma_1 \neq \emptyset$ *et* $\Gamma_2 \neq \emptyset$ *(voir la figure 7.1 à gauche), alors* (7.14) *équivaut à la continuité de* $\mathtt{R}$*. En effet, on a à la fois une inégalité de Poincaré sur* Ω_1 *et sur* Ω_2 *; par conséquent la semi-norme* $|\cdot|_{1,\Omega_1}$ *est équivalente à la norme* $\|\cdot\|_{H^1(\Omega_1)}$ *dans* $H^1_{0,\Gamma_1}(\Omega_1)$*, et il en est de même dans* $H^1_{0,\Gamma_2}(\Omega_2)$*. Dans ce cas et avec ce choix de norme,* $\|\mathtt{R}\|$ *est donc égal au module de continuité* $|||\mathtt{R}|||$*. Dans la situation où soit* Γ_1 *est vide, soit* Γ_2 *est vide (voir la figure 7.1 à droite), on renvoie à l'exercice page 91.*

Démonstration : Prenons $u \in H^1_0(\Omega)$ et calculons $a(u, \mathtt{T}u)$. Par définition, on a

$$a(u, \mathtt{T}u) = \sigma_1 \int_{\Omega_1} |\nabla u_1|^2 \, d\Omega + \sigma_2 \int_{\Omega_2} \nabla u_2 \cdot (-\nabla u_2 + 2\nabla(\mathtt{R}u_1)) \, d\Omega.$$

On en déduit en utilisant l'inégalité de Cauchy-Schwarz et les propriétés de σ que

$$a(u, \mathtt{T}u) \geq \int_{\Omega} |\sigma| |\nabla u|^2 \, d\Omega - 2|\sigma_2| \sqrt{\int_{\Omega_2} |\nabla u_2|^2 \, d\Omega} \sqrt{\int_{\Omega_2} |\nabla(\mathtt{R}u_1)|^2 \, d\Omega}$$

qui implique, par définition de $\|\mathtt{R}\|$:

$$a(u, \mathtt{T}u) \geq \int_{\Omega} |\sigma| |\nabla u|^2 \, d\Omega - 2|\sigma_2| \|\mathtt{R}\| \sqrt{\int_{\Omega_2} |\nabla u_2|^2 \, d\Omega} \sqrt{\int_{\Omega_1} |\nabla u_1|^2 \, d\Omega}.$$

L'inégalité de Young nous dit alors que pour tout $\delta > 0$, on a

$$2\sqrt{\int_{\Omega_2} |\nabla u_2|^2 \, d\Omega} \sqrt{\int_{\Omega_1} |\nabla u_1|^2 \, d\Omega} \leq \delta \int_{\Omega_2} |\nabla u_2|^2 \, d\Omega + \frac{1}{\delta} \int_{\Omega_1} |\nabla u_1|^2 \, d\Omega.$$

En combinant les deux dernières inégalités, on trouve que pour tout $\delta > 0$, on a

$$a(u, \mathtt{T}u) \geq \left(\sigma_1 - \frac{1}{\delta} |\sigma_2| \|\mathtt{R}\| \right) \int_{\Omega_1} |\nabla u_1|^2 \, d\Omega + |\sigma_2| (1 - \delta \|\mathtt{R}\|) \int_{\Omega_2} |\nabla u_2|^2 \, d\Omega,$$

ce qui prouve la $\mathtt{T}$-coercivité à condition que

$$\frac{|\sigma_2| \|\mathtt{R}\|}{\sigma_1} < \delta < \frac{1}{\|\mathtt{R}\|},$$

d'où le résultat. ∎

Nous venons de montrer que si le contraste est assez petit en valeur absolue, bien qu'il soit négatif, la forme bilinéaire $a(\cdot, \cdot)$ est $\mathtt{T}$-coercive, de sorte que le problème de transmission (P) est bien posé. Dans notre approche, nous avons choisi d'introduire le terme correctif $\mathtt{R}u_1$ dans la définition de $\mathtt{T}u$ sur Ω_2. Mais on peut sans rien changer au problème multiplier σ et f par -1, ce qui revient à inverser les rôles de Ω_1 et Ω_2. Cela nous conduit à considérer $\mathtt{R} \in \mathcal{L}(H^1_{0,\Gamma_2}(\Omega_2), H^1_{0,\Gamma_1}(\Omega_1))$, vérifiant

$$\forall u_2 \in H^1_{0,\Gamma_2}(\Omega_2) \quad (\mathtt{R}u_2)|_{\Sigma} = (u_2)|_{\Sigma}, \tag{7.16}$$

et $\mathtt{T} \in \mathcal{L}(H^1_0(\Omega))$ définie pour tout $u \in H^1_0(\Omega)$ par

$$\mathtt{T}u = \begin{cases} u_1 - 2\mathtt{R}u_2 & \text{dans } \Omega_1 \\ -u_2 & \text{dans } \Omega_2 \end{cases} \tag{7.17}$$

On montre ainsi exactement de la même façon que le problème (P) est bien posé pour des contrastes assez grands en valeur absolue, à condition d'avoir également une condition de Dirichlet sur une partie du bord de Ω_2 (ce qui est vrai pour le domaine de gauche mais ne l'est pas pour le domaine de droite de la figure 7.1). Tout ceci est résumé dans le théorème suivant.

Théorème 7.8 *Soient* $\mathtt{R}_{1\to2} \in \mathcal{L}(H^1_{0,\Gamma_1}(\Omega_1), H^1_{0,\Gamma_2}(\Omega_2))$, $\mathtt{R}_{2\to1} \in \mathcal{L}(H^1_{0,\Gamma_2}(\Omega_2), H^1_{0,\Gamma_1}(\Omega_1))$ *telles que*

$$\forall u_1 \in H^1_{0,\Gamma_1}(\Omega_1), \forall u_2 \in H^1_{0,\Gamma_2}(\Omega_2) \quad (\mathtt{R}_{1\to2}u_1)|_\Sigma = (u_1)|_\Sigma \text{ et } (\mathtt{R}_{2\to1}u_2)|_\Sigma = (u_2)|_\Sigma, \tag{7.18}$$

et telles que pour une constante $C > 0$ *:*

$$\left| \begin{array}{l} \displaystyle\int_{\Omega_2} |\nabla(\mathtt{R}_{1\to2}u_1)|^2 \, d\Omega \le C \int_{\Omega_1} |\nabla u_1|^2 \, d\Omega, \quad \forall u_1 \in H^1_{0,\Gamma_1}(\Omega_1), \\ \displaystyle\int_{\Omega_1} |\nabla(\mathtt{R}_{2\to1}u_2)|^2 \, d\Omega \le C \int_{\Omega_2} |\nabla u_2|^2 \, d\Omega, \quad \forall u_2 \in H^1_{0,\Gamma_2}(\Omega_2). \end{array} \right. \tag{7.19}$$

Alors la forme bilinéaire

$$a \ : \ (v,w) \mapsto \int_\Omega \sigma \nabla v \cdot \nabla w \, d\Omega$$

est $\mathtt{T}$*-coercive sur* $H^1_0(\Omega)$ *dès que le contraste* κ_σ *vérifie*

$$\kappa_\sigma \notin \left[-\|\mathtt{R}_{2\to1}\|^2, -\frac{1}{\|\mathtt{R}_{1\to2}\|^2} \right], \tag{7.20}$$

où l'on a posé

$$\begin{aligned} \|\mathtt{R}_{1\to2}\| &= \sup_{u_1 \in H^1_{0,\Gamma_1}(\Omega_1), u_1 \neq 0} \frac{\sqrt{\displaystyle\int_{\Omega_2} |\nabla(\mathtt{R}_{1\to2}u_1)|^2 \, d\Omega}}{\sqrt{\displaystyle\int_{\Omega_1} |\nabla u_1|^2 \, d\Omega}}, \text{ et} \\ \|\mathtt{R}_{2\to1}\| &= \sup_{u_2 \in H^1_{0,\Gamma_2}(\Omega_2), u_2 \neq 0} \frac{\sqrt{\displaystyle\int_{\Omega_1} |\nabla(\mathtt{R}_{2\to1}u_2)|^2 \, d\Omega}}{\sqrt{\displaystyle\int_{\Omega_2} |\nabla u_2|^2 \, d\Omega}}. \end{aligned}$$

Sous cette condition, le problème de transmission (P) *est bien posé.*

Exercice : On se place dans le domaine de droite de la figure 7.1. Montrer qu'il n'existe pas d'application linéaire $\mathtt{R}_{2\to1}$ de $H^1_{0,\Gamma_2}(\Omega_2)$ dans $H^1_{0,\Gamma_1}(\Omega_1)$ vérifiant (7.16), et telle que

$$\exists C > 0, \ \forall u_2 \in H^1_{0,\Gamma_2}(\Omega_2), \quad \int_{\Omega_1} |\nabla(\mathtt{R}_{2\to1}u_2)|^2 \, d\Omega \le C \int_{\Omega_2} |\nabla u_2|^2 \, d\Omega.$$

Pour cela, on considérera u_2 constant dans Ω_2.

Pour le domaine situé à droite sur la figure 7.1, on ne peut donc pas appliquer la démarche proposée au théorème 7.8 pour démontrer que le problème est bien posé pour un contraste assez proche de 0. Néanmoins le résultat demeure vrai et peut se démontrer par une construction légèrement différente de l'application $\mathtt{T}$.

Notons finalement que dans le cas où σ est une fonction variable dans chaque sous-domaine, on peut montrer que le problème (P) est bien posé dès lors que

$$\frac{\sup\limits_{x\in\Omega_2}|\sigma_2(\boldsymbol{x})|}{\inf\limits_{x\in\Omega_1}\sigma_1(\boldsymbol{x})}<\frac{1}{\|\mathtt{R}_{1\to 2}\|^2}\quad\text{ou}\quad\frac{\inf\limits_{x\in\Omega_2}|\sigma_2(\boldsymbol{x})|}{\sup\limits_{x\in\Omega_1}\sigma_1(\boldsymbol{x})}>\|\mathtt{R}_{2\to 1}\|^2.$$

Notes de cours 8

Etude des problèmes de transmission avec changement de signe : compléments

Le premier objectif de ce chapitre est de préciser le résultat obtenu au chapitre 7 pour certaines configurations géométriques simples. On cherche en particulier à déterminer toutes les valeurs possibles du contraste conduisant à un problème bien posé. On montre également comment généraliser la notion de T-coercivité, pour traiter les problèmes bien posés au sens de l'alternative de Fredholm. On repart pour cela du problème (P), sous les hypothèses (7.1)-(7.3), avec un coefficient σ constant par morceaux (7.4). On note $\kappa_\sigma = \sigma_2/\sigma_1 < 0$ le contraste.

8.1 Le cas des domaines symétriques

On va considérer dans ce paragraphe le cas particulier où le domaine Ω est symétrique par rapport à l'interface Σ. Pour fixer les idées, on note $\boldsymbol{x} = (x, y, z)$ si on est dans $\mathbb{R}^3$ et $\boldsymbol{x} = (x, y)$ si on est dans $\mathbb{R}^2$, et on suppose que Σ est contenu dans l'hyperplan $x = 0$. On note alors S

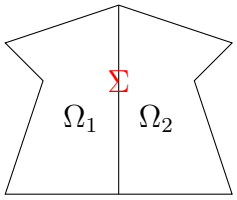

FIGURE 8.1 – Exemple de domaine symétrique

l'application qui à $\boldsymbol{x}$ associe son symétrique par rapport à l'axe $x = 0$. Autrement dit,

$$S(x, y, z) = (-x, y, z) \text{ en 3D, et } S(x, y) = (-x, y) \text{ en 2D.}$$

On démontre alors le résultat suivant.

Théorème 8.1 *Supposons que $S(\Omega_1) = \Omega_2$. Alors le problème (P) est bien posé si et seulement si le contraste κ_σ est différent de -1.*

Démonstration :

Etape 1 : supposons que $\kappa_\sigma \neq -1$ et montrons que le problème (P) est bien posé. D'après le théorème 7.8, il suffit de construire des applications $\mathtt{R}_{1\to 2}$ et $\mathtt{R}_{2\to 1}$ de normes égales à 1. Or ceci est réalisé de façon très simple en posant

$$\mathtt{R}_{1\to 2} u_1 = u_1 \circ S \text{ et } \mathtt{R}_{2\to 1} u_2 = u_2 \circ S.$$

En effet, d'une part la condition (7.18) est bien vérifiée puisque pour $\boldsymbol{x} \in \Sigma$, $S\boldsymbol{x} = \boldsymbol{x}$. D'autre part, par exemple en 2D :

$$\begin{aligned}
\int_{\Omega_2} |\nabla(u_1 \circ S)|^2 \, d\Omega &= \int_{\Omega_2} \left[\left(-\frac{\partial u_1}{\partial x}(-x,y) \right)^2 + \left(\frac{\partial u_1}{\partial y}(-x,y) \right)^2 \right] dxdy \\
&= \int_{\Omega_1} \left[\left(\frac{\partial u_1}{\partial x}(x,y) \right)^2 + \left(\frac{\partial u_1}{\partial y}(x,y) \right)^2 \right] dxdy \\
&= \int_{\Omega_1} |\nabla u_1|^2 \, d\Omega.
\end{aligned}$$

Etape 2 : il reste à montrer que le problème (P) est mal posé quand $\kappa_\sigma = -1$. Pour cela, il suffit de remarquer que dans ce cas, le problème admet un noyau de dimension infinie. Autrement dit, lorsque $f = 0$, (P) admet un espace de solutions de dimension infinie. D'après (7.5), un élément du noyau est une fonction u telle que (u_1, u_2) est solution du problème suivant

$$\left| \begin{array}{l}
-\Delta u_j = 0 \text{ dans } \Omega_j, j = 1,2, \\
u_j = 0 \text{ sur } \Gamma_j, j = 1,2, \\
u_1 = u_2 \text{ sur } \Sigma, \\
\dfrac{\partial u_1}{\partial x} = -\dfrac{\partial u_2}{\partial x} \text{ sur } \Sigma,
\end{array} \right. \tag{8.1}$$

puisque la normale à Σ est orientée selon x.

Notons V_Σ l'espace des traces des fonctions de $H^1_0(\Omega)$ sur Σ (qui est de dimension infinie) :

$$V_\Sigma = \left\{ \lambda \in L^2(\Sigma) \text{ tel que } \exists u \in H^1_0(\Omega),\ \lambda = u_{|\Sigma} \right\}.$$

Alors pour tout $\varphi \in V_\Sigma$, il existe un unique $u \in H^1_0(\Omega)$ tel que $u|_\Sigma = \varphi$ et tel que pour $j = 1, 2$, $\Delta u = 0$ dans Ω_j. Il suffit en effet pour s'en convaincre de considérer le problème de Dirichlet suivant (dont on sait qu'il est bien posé [15]) dans chaque sous-domaine Ω_j :

$$\left| \begin{array}{l}
-\Delta u_j = 0 \text{ dans } \Omega_j, \\
u_j = 0 \text{ sur } \Gamma_j, \\
u_j = \varphi \text{ sur } \Sigma.
\end{array} \right. \tag{8.2}$$

Mais par ailleurs, posons $\tilde{u}_2 = u_1 \circ S$, soit par exemple en 2D,

$$\tilde{u}_2(x,y) = u_1(-x,y) \quad \forall (x,y) \in \Omega_2.$$

On vérifie alors aisément que $\tilde{u}_2$ est solution du même problème de Dirichlet (8.2) dans Ω_2 que u_2. Par unicité du problème de Dirichlet, on en déduit que $u_2 = \tilde{u}_2 = u_1 \circ S$. Autrement dit, la fonction u est symétrique par rapport à Σ. En 2D par exemple, on a donc

$$u(-x,y) = u(x,y) \quad \forall (x,y) \in \Omega,$$

d'où l'on déduit que

$$\frac{\partial u_1}{\partial x}(0,y) = -\frac{\partial u_2}{\partial x}(0,y) \quad \forall (0,y) \in \Sigma.$$

Donc (u_1, u_2) est une solution de (8.1). On a ainsi construit pour chaque φ un élément du noyau, et il est facile d'en conclure que la dimension du noyau est, tout comme celle de V_Σ, infinie. ■

Supposons maintenant que l'on ait $S(\Omega_1) = \Omega_2$ et que le contraste κ_σ soit différent de -1, de sorte que le problème (P) soit bien posé. On veut alors le résoudre par éléments finis. D'après le théorème 7.3, on peut énoncer le résultat suivant :

Théorème 8.2 *Supposons que $S(\Omega_1) = \Omega_2$ et que le contraste κ_σ est différent de -1. Soit V_h un sous-espace de dimension finie de $H_0^1(\Omega)$ tel que*

$$\forall u_h \in V_h \quad u_h \circ S \in V_h. \tag{8.3}$$

Alors le problème approché (où $a(\cdot,\cdot)$ et $\ell(\cdot)$ sont définies par (7.3))

$$(P_h) \quad \left\{ \begin{array}{l} \text{Trouver } u_h \in V_h \text{ tel que} \\ \forall w_h \in V_h, \quad a(u_h, w_h) = \ell(w_h). \end{array} \right.$$

est bien posé, et on a l'estimation d'erreur (7.9) où u désigne la solution du problème exact (P) et u_h celle du problème approché (P_h).

Démonstration : D'après le théorème 7.3, il suffit de vérifier que la condition (8.3) implique que $\mathtt{T}(V_h) \subset V_h$, pour les applications $\mathtt{T}$ suivantes :

$$\mathtt{T}u_h = \left\{ \begin{array}{ll} (u_h)_1 & \text{dans } \Omega_1 \\ -(u_h)_2 + 2(u_h)_1 \circ S & \text{dans } \Omega_2 \end{array} \right. \text{, et } \mathtt{T}u_h = \left\{ \begin{array}{ll} (u_h)_1 - 2(u_h)_2 \circ S & \text{dans } \Omega_1 \\ -(u_h)_2 & \text{dans } \Omega_2 \end{array} \right. , \tag{8.4}$$

ce qui est effectivement le cas. ■

Concrètement, cela signifie que pour résoudre (P) par éléments finis et assurer la convergence de la méthode, il suffit de construire un maillage symétrique (et d'utiliser le même ordre pour les éléments finis de part et d'autre de l'interface). En pratique, on observe en effet de bien meilleurs résultats lorsque l'on utilise un maillage symétrique, même si la méthode semble converger avec un maillage quelconque (ce qui est prouvé sous certaine conditions moins fortes que la symétrie du maillage). Pour fixer les idées, on introduit en 2D le domaine (voir la figure 8.2)

$$\Omega =]-1,1[\times]0,1[\text{ et } \Omega_2 =]0,1[\times]0,1[.$$

Les résultats suivants sont tirés de [11]. On reproduit trois exemples de maillages à la figure 8.3 : tous ces maillages sont conformes à l'interface, c'est-à-dire que tous les triangles sont soit inclus dans $\overline{\Omega_1}$, soit dans $\overline{\Omega_2}$. Pour réaliser les tests et calculer précisément l'erreur, on se donne la solution exacte u, qu'on choisit régulière par morceaux, et on calcule la donnée f correspondante. On construit la formulation variationnelle (7.8) à l'aide d'éléments finis de Lagrange P_1 : les espaces discrets $(V_h)_h$ sont construits à partir d'une famille de maillages

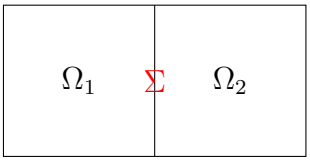

FIGURE 8.2 – Domaine symétrique "type"

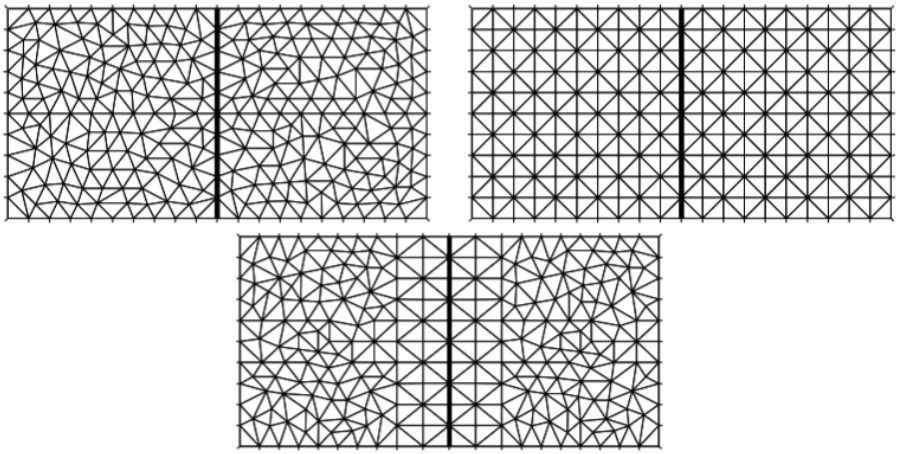

FIGURE 8.3 – Trois exemples de maillages, non-symétrique en haut à gauche, symétrique en haut à droite, et localement symétrique au voisinage de l'interface en bas.

régulière. Comme la solution est régulière par morceaux, l'ordre théorique de convergence de l'erreur est de 1 en norme $|\cdot|_{1,\Omega}$ et de 2 en norme $\|\cdot\|_{L^2(\Omega)}$ si l'on utilise un maillage symétrique. Les tests ont été réalisés à l'aide du logiciel `Freefem++`. On choisit deux valeurs du contraste $\kappa_\sigma \in \{-2, -1.001\}$, et on étudie l'influence des maillages sur l'erreur, mesurée en norme $\|\cdot\|_{L^2(\Omega)}$. Les échelles sont en log-log, voir les figures 8.4-8.5, de sorte que l'on s'attend à trouver une droite de pente -2. Faisons les commentaires suivants :

- Que les maillages soient (complètement) symétriques, localement symétriques dans un voisinage de Σ, ou quelconques, on observe que l'erreur converge vers 0. Par contre, les ordres des erreurs sont différents d'une classe de maillage à l'autre. Ceci est valable pour les deux valeurs de κ_σ, et les différences sont plus marquées lorsqu'on se rapproche de la valeur critique -1.
- De plus, les vitesses de convergence sont différentes. On ne retrouve l'ordre théorique (2) que pour les maillages symétriques. Pour les deux autres classes de maillage, on semble être dans le régime "pré-asymptotique" de convergence (on converge, mais pas à l'ordre prévu par la théorie).
- Enfin, la convergence n'est pas monotone, sauf si les maillages sont complètement symétriques.

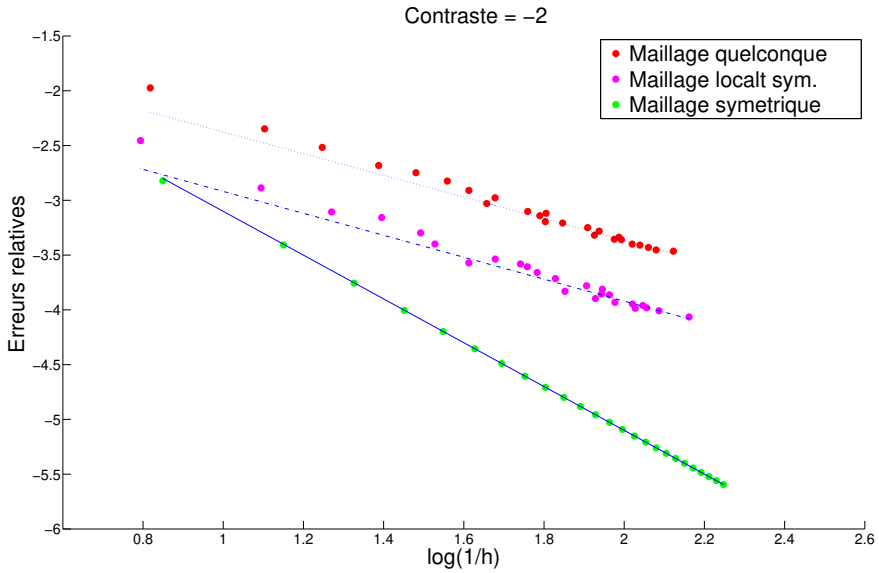

FIGURE 8.4 – Domaine symétrique : contraste $\kappa_\sigma = -2$.

8.2 Le cas des interfaces à coin

8.2.1 Le carré dans un carré

Considérons maintenant le domaine 2D de la figure 8.6 tel que Ω_2 correspond au quart du carré Ω. Alors on a le résultat suivant :

Théorème 8.3 *Soient* $\Omega =]-1,1[\times]-1,1[$ *et* $\Omega_2 =]0,1[\times]0,1[$. *Alors le problème* (P) *est bien posé si le contraste* κ_σ *n'appartient pas à l'intervalle* $[-3,-1/3]$.

Démonstration : Il suffit de construire des applications $\mathtt{R}_{1\to 2}$ et $\mathtt{R}_{2\to 1}$ de normes au carré égales à 3. Pour la construction de $\mathtt{R}_{2\to 1}$, les symétries du domaine conduisent naturellement à proposer une construction à valeur dans les trois sous-domaines de Ω_1 isométriques à Ω_2 :

$$\mathtt{R}_{2\to 1}u_2(x,y) = \begin{cases} u_2(-x,y) & \text{si} \quad (x,y) \in]-1,0[\times]0,1[, \\ u_2(x,-y) & \text{si} \quad (x,y) \in]0,1[\times]-1,0[, \\ u_2(-x,-y) & \text{si} \quad (x,y) \in]-1,0[\times]-1,0[. \end{cases}$$

Il est clair que cette application est admissible au sens où :

- $\mathtt{R}_{2\to 1}u_2$ appartient à H^1 dans chaque sous-domaine, est continu aux interfaces entre sous-domaines, et ainsi $\mathtt{R}_{2\to 1}u_2 \in H^1(\Omega_1)$ (d'après le théorème 1.36) ;
- en outre, $\mathtt{R}_{2\to 1}u_2$ s'annule sur la frontière extérieure et vérifie la condition (7.18) ;
- enfin, il est facile de voir que

$$\int_{\Omega_1} |\nabla(\mathtt{R}_{2\to 1}u_2)|^2 \, d\Omega = 3 \int_{\Omega_2} |\nabla u_2|^2 \, d\Omega.$$

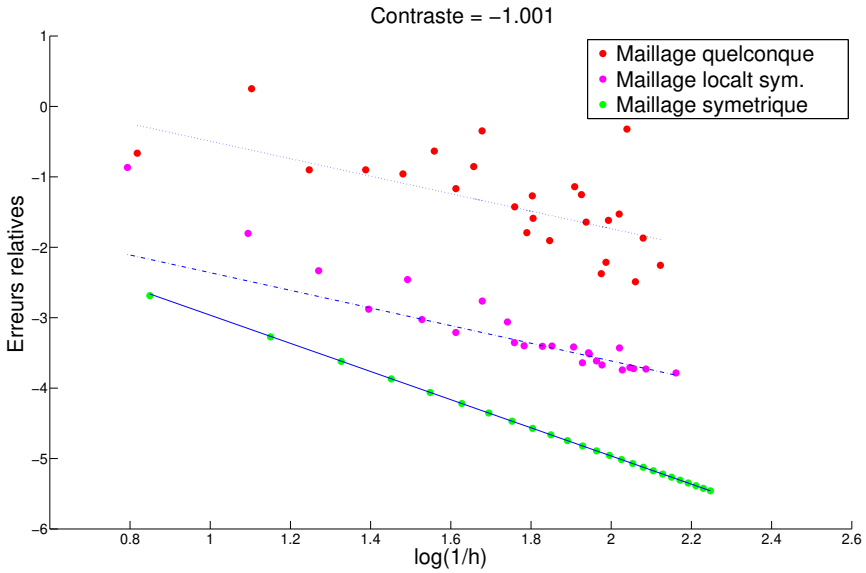

FIGURE 8.5 – Domaine symétrique : contraste $\kappa_\sigma = -1.001$.

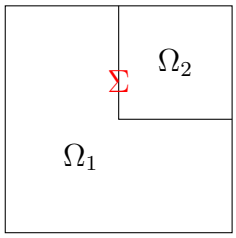

FIGURE 8.6 – Exemple où l'interface a un angle droit

La construction de $\mathtt{R}_{1\to 2}$ se fait par une sorte de repliement. On pose pour $(x, y) \in \Omega_2$:

$$\mathtt{R}_{1\to 2}u_1(x, y) = u_1(-x, y) - u_1(-x, -y) + u_1(x, -y),$$

et à nouveau, on constate que cette application vérifie toutes les conditions requises. Enfin, pour calculer sa norme, on utilise les relations suivantes, valables pour tout M-uplet de réels $a_1, a_2, ...a_M$ (c'est l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^M$) :

$$\left(\sum_{i=1}^{M} a_i\right)^2 \leq M \sum_{i=1}^{M} a_i^2. \tag{8.5}$$

On trouve donc en utilisant à nouveau les symétries que

$$\int_{\Omega_2} |\nabla(\mathtt{R}_{1\to 2}u_1)|^2 \, d\Omega \leq 3 \int_{\Omega_1} |\nabla u_1|^2 \, d\Omega,$$

l'égalité étant réalisée si pour tout (x, y), on a

$$u_1(-x, y) = u_1(x, -y) = -u_1(-x, -y).$$

■

Remarque 8.4 *On peut noter que l'intervalle* $[-3,-1/3]$ *contient la valeur* -1 *qui était déjà exclue dans le cas des domaines symétriques.*

A nouveau, la construction des applications $\mathtt{R}_{1\to 2}$ et $\mathtt{R}_{2\to 1}$ par symétrie permet d'énoncer le résultat suivant pour le problème discret :

Théorème 8.5 *Supposons que le contraste* κ_σ *n'appartient pas à l'intervalle* $[-3,-1/3]$. *Soit* V_h *un sous-espace éléments finis de* $H_0^1(\Omega)$, *construit à l'aide d'un maillage possédant la double symétrie en* x *et en* y, *alors le problème* (P_h) *est bien posé, et on a l'estimation d'erreur* (7.9) *où* u *désigne la solution du problème* (P) *et* u_h *celle du problème* (P_h).

Remarque 8.6 — *On peut remarquer que les résultats précédents s'appliquent directement à toute géométrie possédant cette double symétrie en* x *et en* y, *comme par exemple celles de la figure 8.7.*

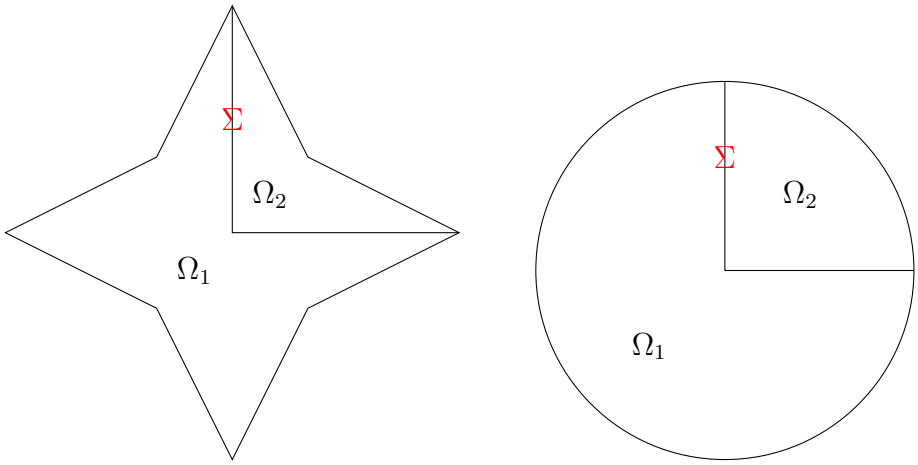

FIGURE 8.7 – Domaines symétriques en x et en y.

— *Nous avons démontré que, pour ces géométries à double symétrie, le problème* (P) *est bien posé dès que le contraste* κ_σ *n'appartient pas à l'intervalle* $[-3,-1/3]$. *En fait, on peut même démontrer que ce résultat est optimal, et que dès que l'interface présente un angle droit, le problème devient mal posé pour un contraste appartenant à cet intervalle [5]. C'est pourquoi on dit que* $[-3,-1/3]$ *est l'intervalle critique. Il est intéressant de noter que la valeur* 3 *correspond exactement au rapport entre les deux angles de part et d'autre du coin,* $3\pi/2$ *et* $\pi/2$.

— *On peut établir des résultats similaires en 3D, par exemple pour un cube dans un cube :* $\Omega_2 = \{(x,y,z); 0 < x < 1, 0 < y < 1, 0 < y < 1\}$ *et* $\Omega = \{(x,y,z); |x| < 1, |y| < 1, |z| < 1\}$. *En généralisant (en 3D) les applications* $\mathtt{R}_{1\to 2}$ *et* $\mathtt{R}_{2\to 1}$ *introduites au cours de la démonstration du théorème 8.3, on montre alors que le problème est bien posé si* $\kappa_\sigma \notin [-7,-1/7]$. *Mais contrairement au cas 2D, ce résultat n'est pas optimal.*

8.2.2 Le camembert

Que se passe-t-il si l'interface présente un coin d'angle α quelconque ? On va considérer le cas où Ω est un disque dont un secteur constitue Ω_2 :

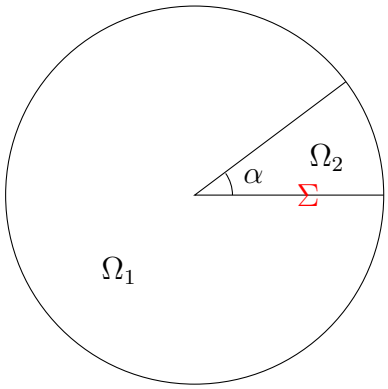

FIGURE 8.8 – Le camembert.

De façon naturelle, on repèrera ici les points par leurs coordonnées polaires (r, θ). On démontre alors le résultat suivant :

Théorème 8.7 *Soient* $\Omega = \{(r,\theta); r < 1\}$ *et* $\Omega_2 = \{(r,\theta); 0 < \theta < \alpha \text{ et } r < 1\}$. *Alors le problème* (P) *est bien posé si le contraste* κ_σ *n'appartient pas à l'intervalle* $]-\mathcal{I}_\alpha, -1/\mathcal{I}_\alpha[$ *où*

$$\mathcal{I}_\alpha = \max\left(\frac{2\pi-\alpha}{\alpha}, \frac{\alpha}{2\pi-\alpha}\right).$$

Remarque 8.8
- *On remarque que ce résultat est identique à celui du paragraphe précédent lorsque* $\alpha = \pi/2$, *puisque* $\mathcal{I}_\alpha$ *est alors égal à 3.*
- *On peut aussi noter que si* $\alpha = \pi$, $\mathcal{I}_\alpha = 1$ *et on retrouve le résultat du cas symétrique.*
- *Enfin, on note que plus l'angle est petit, plus l'intervalle critique est grand, tendant à la limite à remplir tout l'axe réel négatif.*

Démonstration : Il suffit de construire des applications $\mathtt{R}_{1\to2}$ et $\mathtt{R}_{2\to1}$ de normes au carré égales à $\mathcal{I}_\alpha$. On propose pour cela de choisir par exemple $\mathtt{R}_{2\to1}$ de la forme suivante, où γ est un paramètre réel positif à fixer :

$$\mathtt{R}_{2\to1}u_2(r,\theta) = u(r,-\gamma\theta).$$

Il est clair qu'on a la première condition de raccord (en $\theta = 0$)

$$\mathtt{R}_{2\to1}u_2(r,0) = u(r,0),$$

mais on veut également la deuxième condition de raccord (en $\theta = \alpha$)

$$\mathtt{R}_{2\to1}u_2(r,-2\pi+\alpha) = u(r,\alpha),$$

ce qui est réalisé en prenant

$$\gamma = \frac{\alpha}{2\pi-\alpha}.$$

On calcule alors la norme de l'application :

$$\begin{aligned}
\int_{\Omega_1} |\nabla(\mathtt{R}_{2\to1}u_2)|^2\, d\Omega &= \int_{-2\pi+\alpha}^{0} d\theta \int_0^1 r\, dr \left(\left| \frac{\partial}{\partial r}(u(r,-\gamma\theta) \right|^2 + \left| \frac{\partial}{r\partial\theta}(u(r,-\gamma\theta)) \right|^2 \right) \\
&= \int_{-2\pi+\alpha}^{0} d\theta \int_0^1 r\, dr \left(\left| \frac{\partial u}{\partial r}(r,-\gamma\theta) \right|^2 + \gamma^2 \left| \frac{\partial u}{r\partial\theta}(r,-\gamma\theta) \right|^2 \right) \\
(\theta' = -\gamma\theta) \qquad &= \int_0^{\alpha} \frac{1}{\gamma} d\theta' \int_0^1 r\, dr \left(\left| \frac{\partial u}{\partial r}(r,\theta') \right|^2 + \gamma^2 \left| \frac{\partial u}{r\partial\theta}(r,\theta') \right|^2 \right) \\
&\leq \max(\gamma, 1/\gamma) \int_{\Omega_2} |\nabla u_2|^2\, d\Omega,
\end{aligned}$$

qui est le résultat attendu. On peut de faire de même pour l'application $\mathtt{R}_{1\to2}$. ■

L'inconvénient de la construction proposée dans cette preuve est qu'elle ne "passe pas" au discret, au sens où ne peut pas réaliser la condition $\mathtt{T}(V_h) \subset V_h$ avec des éléments finis. C'est pourquoi une construction alternative a été proposée, qui fonctionne lorsque l'angle α est un multiple rationnel de π. On renvoie à [4] pour plus de détails.

8.3 Une idée de la démarche générale

8.3.1 T-coercivité et localisation

Pour une géométrie arbitraire comme sur la figure 7.1, on ne peut pas construire de façon aussi simple une application $\mathtt{T}$ assurant la $\mathtt{T}$-coercivité pour certaines valeurs du contraste. On décide alors d'être moins exigeant. Au lieu de chercher à restaurer la coercivité de la forme bilinéaire $a(\cdot,\cdot)$, on va se contenter d'obtenir la coercivité "à une perturbation compacte près". On s'appuie sur le théorème suivant :

Théorème 8.9 *Soit V un espace de Hilbert, $a(\cdot,\cdot)$ et $b(\cdot,\cdot)$ deux formes bilinéaires continues sur $V\times V$ et ℓ une forme linéaire continue sur V. On suppose que la forme bilinéaire $a(\cdot,\cdot)+b(\cdot,\cdot)$ est $\mathtt{T}$-coercive sur V et que $\mathtt{B} \in \mathcal{L}(V)$ définie par*

$$\forall v, w \in V, \quad (\mathtt{B}v, w)_V = b(v, w)$$

est compacte. Alors le problème

$$\left\{ \begin{array}{l} \text{Trouver } u \in V \text{ tel que} \\ \forall w \in V, \quad a(u,w) = \ell(w). \end{array} \right. \tag{8.6}$$

relève de l'alternative de Fredholm : il est bien posé si et seulement si le problème homogène l'est.

Démonstration : En effet, si on note $\mathtt{A} \in \mathcal{L}(V)$ définie par

$$\forall v, w \in V, \quad (\mathtt{A}v, w)_V = a(v, w),$$

les hypothèses du théorème permettent de montrer que $\mathtt{A}+\mathtt{B}$ est un isomorphisme. ■

Nous allons appliquer ce théorème à l'exemple suivant. On suppose que

$$\Omega = \{(x,y); -d < x < 1, 0 < y < 1\},$$

et

$$\Omega_1 = \{(x,y); -d < x < 0, 0 < y < 1\} \text{ et } \Omega_2 = \{(x,y); 0 < x < 1, 0 < y < 1\}$$

où d est un réel positif. Dans le cas où $d = 1$, on a vu que le problème (P) est bien posé si et seulement si $\kappa_\sigma \neq -1$. On voudrait maintenant savoir ce qu'on peut dire si $d \neq 1$. Supposons pour fixer les idées que $d > 1$. En procédant comme dans la preuve du théorème 8.1, on peut

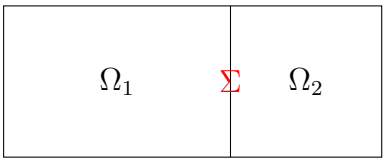

FIGURE 8.9 – Domaine rectangulaire

montrer que (P) est bien posé si $\kappa_\sigma < -1$. En effet, si S désigne la symétrie par rapport à l'interface $\Sigma = \{(x,y) \in \Omega; x = 0\}$, on peut définir une application $\mathtt{R}_{2\to1}$ comme suit :

$$\mathtt{R}_{2\to1}u_2(x,y) = \begin{cases} u_2(-x,y) & \text{si} \quad (x,y) \in]-1,0[\times]0,1[, \\ 0 & \text{si} \quad (x,y) \in]-d,-1[\times]0,1[. \end{cases}$$

En revanche, on ne peut pas aussi facilement construire une application optimale $\mathtt{R}_{1\to2}$. Comme on sait que la difficulté vient du changement de signe de σ à travers l'interface Σ, on introduit une fonction de troncature pour localiser l'action de l'application $\mathtt{R}_{1\to2}$ au voisinage de Σ. Soit $\chi \in \mathcal{D}(\mathbb{R})$, à valeurs dans $[0,1]$, telle que $\chi(x) = 0$ si $|x| > 1$ et $\chi(0) = 1$. Pour simplifier les calculs, on peut choisir une fonction χ qui est paire : $\chi(-x) = \chi(x)$. On définit alors l'application $\mathtt{R}_{1\to2}$ comme suit :

$$\mathtt{R}_{1\to2}u_1(x,y) = \chi(x)u_1(-x,y), \ \forall (x,y) \in]0,1[\times]0,1[,$$

et l'application $\mathtt{T}$ associée :

$$\mathtt{T}u(x,y) = \begin{cases} u(x,y) & \text{pour } (x,y) \in \Omega_1 \\ -u(x,y) + 2\chi(x)u(-x,y) & \text{pour } (x,y) \in \Omega_2. \end{cases}$$

Comme dans tous les exemples précédents, $(\mathtt{T} \circ \mathtt{T})u = u$ donc $\mathtt{T}$ est bijective.

On a alors pour $u \in H^1_0(\Omega)$,

$$a(u, \mathtt{T}u) = \int_\Omega |\sigma||\nabla u|^2 \, d\Omega - 2|\sigma_2| \int_{\Omega_2} \nabla u(x,y).\nabla(\chi(x)u(-x,y)) \, d\Omega.$$

Cela nous conduit à poser :

$$b(u, \mathtt{T}v) = 2|\sigma_2| \int_{\Omega_2} \frac{\partial u}{\partial x}(x,y)\chi'(x)v(-x,y)\, d\Omega, \tag{8.7}$$

de sorte que

$$a(u, \mathtt{T}u) + b(u, \mathtt{T}u) = \int_\Omega |\sigma||\nabla u|^2\, d\Omega - 2|\sigma_2| \int_{\Omega_2} \chi(x)\nabla u(x,y).\nabla(u(-x,y))\, d\Omega.$$

Il est très facile en reproduisant la preuve du théorème 7.6 de montrer que $a(\cdot,\cdot) + b(\cdot,\cdot)$ est $\mathtt{T}$-coercive sur $H_0^1(\Omega)$ dès que $\kappa_\sigma > -1$. Il ne reste plus qu'à montrer la compacité de l'application $\mathtt{B}$ associée à la forme $b(\cdot,\cdot)$.

Soit donc $(u_n)_n$ une suite bornée de $H_0^1(\Omega)$. On veut montrer que la suite $(v_n)_n = (\mathtt{B}u_n)_n$ admet une sous-suite qui converge dans $H_0^1(\Omega)$. Par définition de $\mathtt{B}$:

$$\|v_n - v_m\|^2_{H_0^1(\Omega)} = ((\mathtt{B}(u_n - u_m), (v_n - v_m))_{H_0^1(\Omega)} = b((u_n - u_m), (v_n - v_m)).$$

Mais comme $\mathtt{T}$ est égal à son inverse, on a

$$b(u, v) = b(u, \mathtt{T}(\mathtt{T}v)).$$

D'après (8.7), on en déduit finalement que

$$\|v_n - v_m\|^2_{H_0^1(\Omega)} = 2|\sigma_2| \int_{\Omega_2} \frac{\partial(u_n - u_m)}{\partial x}(x,y)\chi'(x)(\mathtt{T}v_n - \mathtt{T}v_m)(-x,y)\, d\Omega,$$

d'où, en utilisant l'inégalité de Cauchy-Schwarz

$$\|v_n - v_m\|^2_{H_0^1(\Omega)} \leq 2|\sigma_2|\, \|\chi'\|_{L^\infty(]-1,1[)}\, \|u_n - u_m\|_{H_0^1(\Omega)} \|\mathtt{T}v_n - \mathtt{T}v_m\|_{L^2(\Omega_2)}.$$

Or la suite $(\mathtt{T}v_n)_n$ est bornée dans $H_0^1(\Omega)$, donc d'après le lemme de Rellich (puisque Ω est borné), on peut en extraire une sous-suite $(\mathtt{T}v_{n'})_{n'}$ qui converge dans $L^2(\Omega)$. On en déduit que la suite $(v_{n'})_{n'}$ est de Cauchy dans $H_0^1(\Omega)$, ce qui permet de conclure.

8.3.2 Modes plasmoniques

En poursuivant sur l'exemple précédent, si $\kappa_\sigma \neq -1$, on sait que le problème (P) est bien posé si et seulement si le problème homogène l'est. Il suffit donc de chercher des solutions du problème homogène suivant :

$$\left|\begin{array}{l} -\Delta u_j = 0 \text{ dans } \Omega_j, j = 1,2, \\ u_j = 0 \text{ sur } \Gamma_j, j = 1,2, \\ u_1 = u_2 \text{ sur } \Sigma, \\ \sigma_1 \dfrac{\partial u_1}{\partial n_\Sigma} = \sigma_2 \dfrac{\partial u_2}{\partial n_\Sigma} \text{ sur } \Sigma. \end{array}\right. \tag{8.8}$$

La séparation de variables conduit à rechercher, pour chaque entier $n \geq 1$, des solutions de la forme :

$$u_j(x,y) = \varphi_j(x)\sin(n\pi y),\ j = 1,2.$$

Les deux premières équations de (8.8) imposent alors l'existence de constantes A_1 et A_2 telles que

$$\varphi_1(x) = A_1\sinh(n\pi(x+d)) \text{ et } \varphi_2(x) = A_2\sinh(n\pi(x-1)).$$

Les deux dernières équations de (8.8) conduisent ensuite au système linéaire suivant

$$\left|\begin{array}{l} A_1\sinh(n\pi d) = A_2\sinh(n\pi) \\ \sigma_1 A_1\cosh(n\pi d) = -\sigma_2 A_2\cosh(n\pi) \end{array}\right.$$

qui n'admet de solution non triviale que si

$$\frac{\sigma_2}{\sigma_1} = -\frac{\tanh(n\pi)}{\tanh(n\pi d)}.$$

On trouve donc une suite de contrastes $(\kappa_\sigma^n)_{n\geq 0} = -\tanh(n\pi)/\tanh(n\pi d)$, $n \geq 1$ pour lesquels le problème (P) est mal posé, et possède un noyau de dimension 1. Cette suite de contrastes tend vers -1. Lorsque

$$\kappa_\sigma \notin \{-1\} \cup \{(\kappa_\sigma^n)_n,\ n \geq 0\},$$

le problème homogène n'a que la solution triviale, et le problème de départ est donc bien posé.

Les modes que nous avons calculés sont appelés les modes plasmoniques.

8.3.3 Géométries quelconques

Comment résoudre (7.1)-(7.3), pour un coefficient σ constant par morceaux (7.4), si l'interface est "quelconque", typiquement régulière (pour $\mathtt{d} = 2,3$), ou polygonale ($\mathtt{d} = 2$) ? En utilisant une modification via localisation des applications $\mathtt{T}$ respectivement définies en (7.13) et (7.17), on peut prouver l'optimalité de l'approche développée précédemment pour établir la $\mathtt{T}$-coercivité à une perturbation compacte près, cf. [5, 4]. En effet, grâce à cette approche on peut démontrer qu'il existe un *intervalle critique* $I_\Sigma \subset \mathbb{R}^*_-$ tel que

— si $\kappa_\sigma \in I_\Sigma$: le problème avec coefficient changeant de signe est mal posé dans $H^1_0(\Omega)$;

— si $\kappa_\sigma \notin I_\Sigma$: on peut établir une inégalité de Gärding

$$\exists \underline{\alpha}_\sigma, \underline{\beta}_\sigma > 0,\ \forall v \in H^1_0(\Omega),\ \left|\int_\Omega \sigma\,\nabla v \cdot \overline{\nabla(\mathtt{T}\,v)}\,d\Omega\right| \geq \underline{\alpha}_\sigma\,|v|^2_{1,\Omega} - \underline{\beta}_\sigma\|v\|^2_{L^2(\Omega)}. \tag{8.9}$$

En d'autres termes, le problème avec coefficient constant par morceaux changeant de signe est bien posé, au sens de l'alternative de Fredholm (voir le théorème 8.9).

Les bornes de l'intervalle critique I_Σ dépendent de la valeur des angles aux sommets de l'interface (si elle en possède...). Dans tous les cas, -1 appartient à I_Σ . Et, si l'interface est de classe $\mathcal{C}^2$ et sans extrémités, on a l'égalité $I_\Sigma = \{-1\}$ (ce résultat avait déjà été établi dans [17]).

Exercice : On se place dans un domaine Ω symétrique par rapport à l'interface Σ. Vérifier que les applications $\mathtt{R}_{1\to 2}$ et $\mathtt{R}_{2\to 1}$ construites à l'aide du relèvement harmonique (cf. remarque 7.5) sont de normes égales à 1.

Notes de cours 9

Synthèse et compléments

On revoit et complète les notes précédentes en revisitant la notion de problèmes bien posés, leur étude et leur discrétisation, que l'on considère ici dans le cadre d'espaces fonctionnels définis sur $\mathbb{K} = \mathbb{C}$. Ci dessous, $|\cdot|$ désigne donc le module. On y inclut la T-coercivité. On applique ces résultats aux problèmes avec coefficient changeant de signe.

9.1 Problèmes bien posés et T-coercivité

Nous rappelons les outils d'analyse fonctionnelle traitant de la notion de problèmes bien posés (voir §2.3), que nous reformulons en utilisant la théorie de la T-coercivité [6, 11] (voir aussi les §2.3.4 et §7.2).

9.1.1 Définitions

Soient :

— V et W deux espaces de Hilbert définis sur $\mathbb{C}$, de produits scalaires hermitiens respectifs $(\cdot,\cdot)_V$ et $(\cdot,\cdot)_W$. On note $\|\cdot\|_V$ et $\|\cdot\|_W$ les normes associées ;

— $a(\cdot,\cdot)$ une forme sesquilinéaire (anti-linéaire par rapport au second argument) et continue sur $V \times W$. On utilise la norme $|||\cdot|||$ pour mesurer a, voir (1.3) :

$$|||a||| = \sup_{v\in V\setminus\{0\},\, w\in W\setminus\{0\}} \frac{|a(v,w)|}{\|v\|_V \|w\|_W}.$$

On rappelle que W' est l'espace des formes anti-linéaires et continues sur W, et on note $\langle\cdot,\cdot\rangle_W$ les crochets de dualité pour l'action d'un élément de W' sur un élément de W.

Pour $\ell \in W'$ donné, on étudie le problème écrit sous forme variationnelle,

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ \forall w \in W, \quad a(u,w) = \langle \ell, w\rangle_W. \end{cases} \tag{9.1}$$

Le caractère bien posé (au sens de Hadamard) est énoncé à la définition 2.6, et reformulé à la proposition 2.8. On y a introduit $\mathtt{A} \in \mathcal{L}(V,W)$ associée à $a(\cdot,\cdot)$, définie par :

$$\forall (v,w) \in V \times W, \quad (\mathtt{A}v, w)_W = a(v,w), \tag{9.2}$$

et on a démontré le résultat suivant

Proposition 9.1 *Soient V, W deux espaces de Hilbert, et $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V \times W$. Alors la formulation variationnelle (9.1) est bien posée si, et seulement si, l'inverse de* $\mathtt{A}$ *existe et est continu :* $\mathtt{A}^{-1} \in \mathcal{L}(W,V)$.

9.1.2 La T-coercivité pour établir le caractère bien posé

Pour assurer le caractère bien posé, les conditions inf-sup et de solvabilité sont des *conditions nécessaires et suffisantes*, d'après le théorème BNB (théorème 2.12). Nous les rappelons ci-dessous.

Définition 9.2 *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire sur $V \times W$. Alors $a(\cdot,\cdot)$ vérifie une* condition inf-sup *si, et seulement si,*

$$\exists \alpha' > 0, \ \forall v \in V, \ \sup_{w \in W \setminus \{0\}} \frac{|a(v,w)|}{\|w\|_W} \geq \alpha' \|v\|_V. \tag{9.3}$$

Dans ce cas, on appelle constante inf-sup *la plus grande valeur de α' dans (9.3), égale à*

$$\inf_{v \in V \setminus \{0\}} \sup_{w \in W \setminus \{0\}} \frac{|a(v,w)|}{\|v\|_V \, \|w\|_W}.$$

Définition 9.3 *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire sur $V \times W$. Alors $a(\cdot,\cdot)$ vérifie une* condition de solvabilité *si, et seulement si,*

$$\{w \in W \text{ tel que } a(v,w) = 0, \ \forall v \in V\} = \{0\}. \tag{9.4}$$

D'autre part, on a vu que la T-coercivité introduite à la définition 7.1 est une *condition suffisante*, voir le §2.3.4 (cas $\mathbb{K} = \mathbb{R}$) et le théorème 7.2.

Définition 9.4 *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire sur $V \times W$. On dit que $a(\cdot,\cdot)$ est* T-coercive *si, et seulement si,*

$$\exists \mathtt{T} \in \mathcal{L}(V,W), \ \text{bijective}, \ \exists \underline{\alpha} > 0, \ \forall v \in V, \ |a(v, \mathtt{T}v)| \geq \underline{\alpha} \|v\|_V^2. \tag{9.5}$$

Remarque 9.5 *Lorsque $V = W$, si la forme $a(\cdot,\cdot)$ est coercive, alors elle est également* T*-coercive. En effet, $\mathtt{T} = \mathtt{I}$ et $\underline{\alpha}$ égale à la constante de coercivité conviennent dans (9.5).*

Dans la définition ci-dessus, il suffit de trouver une application linéaire, continue et bijective $\mathtt{T}$ de V dans W pour assurer la $\mathtt{T}$-coercivité. Cette application $\mathtt{T}$ *n'est pas unique !*
A priori, la $\mathtt{T}$-coercivité semble plus forte que les conditions inf-sup et de solvabilité, puisque c'est une condition suffisante. Et en effet, l'existence d'une application linéaire et continue $\mathtt{T}$ vérifiant (9.5) permet de retrouver (9.3) et (9.4).
Démonstration : Soit $v \in V \setminus \{0\}$, on choisit $w_v = \mathtt{T}v \in W \setminus \{0\}$. Par construction, on a $\|w_v\|_W \leq |||\mathtt{T}||| \, \|v\|_V$, et donc

$$\frac{1}{|||\mathtt{T}||| \, \|v\|_V} \leq \frac{1}{\|w_v\|_W}.$$

On en déduit que

$$\sup_{w \in W \setminus \{0\}} \frac{|a(v,w)|}{\|w\|_W} \geq \frac{|a(v,w_v)|}{\|w_v\|_W} = \frac{|a(v,\mathtt{T}v)|}{\|w_v\|_W} \geq \frac{\underline{\alpha}}{|||\mathtt{T}|||} \|v\|_V,$$

c'est-à-dire que (9.3) est vraie avec $\alpha' = \underline{\alpha} \, |||\mathtt{T}|||^{-1}$.
Ensuite, soit $w \in W$ tel que $a(v,w) = 0$ pour tout $v \in V$. Puisque $\mathtt{T}^{-1}$ existe, on peut choisir $v_w = \mathtt{T}^{-1}w$. Par définition de w, on a en particulier

$$0 = a(v_w, w) = a(v_w, \mathtt{T}v_w).$$

D'après (9.5), on en conclut que $v_w = 0$, et donc que $w = 0$, c'est-à-dire que (9.4) est vraie. ∎

En fait, ces propriétés sont équivalentes entre elles, ce qui signifie au final que la $\mathtt{T}$-coercivité est une *condition nécessaire et suffisante* pour assurer le caractère bien posé !

Théorème 9.6 (`Caractère bien posé`) *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V \times W$. Les trois assertions ci-dessous sont équivalentes :*

(i) la formulation variationnelle (9.1) est bien posée ;

(ii) la forme $a(\cdot,\cdot)$ vérifie une condition inf-sup et une condition de solvabilité ;

(iii) la forme $a(\cdot,\cdot)$ est $\mathtt{T}$-coercive.

Démonstration : D'après le théorème BNB, on sait que (ii) $\Longrightarrow$ (i).
D'après ce qu'on a vu ci-dessus, (iii) $\Longrightarrow$ (ii).
On montre pour finir que (i) $\Longrightarrow$ (iii). D'après (i), $\mathtt{T} = \mathtt{A}$ est une bijection de $\mathcal{L}(V,W)$, et on a $\mathtt{T}^{-1} = \mathtt{A}^{-1} \in \mathcal{L}(W,V)$ d'après le théorème de l'isomorphisme de Banach. Alors, pour tout $v \in V$, on a la majoration $\|v\|_V \leq \left|\left|\left|\mathtt{T}^{-1}\right|\right|\right| \, \|\mathtt{T}v\|_W$, et on trouve ensuite que

$$a(v,\mathtt{T}v) = (\mathtt{A}v, \mathtt{T}v)_W = \|\mathtt{T}v\|_W^2 \geq \|v\|_V^2 / \left|\left|\left|\mathtt{T}^{-1}\right|\right|\right|^2,$$

et la forme $a(\cdot,\cdot)$ est $\mathtt{T}$-coercive. ∎

Corollaire 9.7 (`Stabilité`) *Sous les hypothèses du théorème 9.6, lorsque la formulation variationnelle (9.1) est bien posée, pour chaque $\mathtt{T} \in \mathcal{L}(V,W)$ et $\underline{\alpha} > 0$ réalisant (9.5), on a l'estimation de stabilité :*

$$\forall \ell \in W', \ \|u\|_V \leq \frac{|||\mathtt{T}|||}{\underline{\alpha}} \|\ell\|_{W'}.$$

Remarque 9.8 *Dans l'esprit de la remarque 2.7, pour "optimiser" la stabilité, on cherchera donc* $\mathtt{T}$ *et* $\underline{\alpha}$ *tels que* $|||\mathtt{T}|||/\underline{\alpha}$ *soit le plus petit possible.*

Démonstration : On note que

$$\underline{\alpha}\|u\|_V^2 \leq |a(u,\mathtt{T}u)| = |\langle \ell, \mathtt{T}u\rangle_W| \leq \|\ell\|_{W'}\|\mathtt{T}u\|_W \leq |||\mathtt{T}|||\|\ell\|_{W'}\|u\|_V,$$

et le résultat suit. ∎

Plaçons-nous dans le cas $W = V$. On a vu au corollaire 2.16, ou corollaire BNB, que si la forme $a(\cdot,\cdot)$ est hermitienne, alors la condition inf-sup (9.3) est une condition nécessaire et suffisante pour assurer le caractère bien posé. De même, si la forme $a(\cdot,\cdot)$ est *hermitienne*, la définition 9.4 peut être modifiée en

Définition 9.9 *Soit* $a(\cdot,\cdot)$ *une forme sesquilinéaire et hermitienne sur* $V\times V$. *On dit que* $a(\cdot,\cdot)$ *est* $\mathtt{T}$-coercive *si, et seulement si,*

$$\exists \mathtt{T}\in\mathcal{L}(V),\ \exists\underline{\alpha}>0,\ \forall v\in V,\ |a(v,\mathtt{T}v)|\geq\underline{\alpha}\|v\|_V^2. \tag{9.6}$$

En d'autres termes, lorsque la forme $a(\cdot,\cdot)$ est hermitenne, la bijectivité de $\mathtt{T}$ n'est plus nécessaire. Un second corollaire du théorème 9.6 résume les résultats dans ce cas.

Corollaire 9.10 `(Cas hermitien)` *Soit* $a(\cdot,\cdot)$ *une forme sesquilinéaire, continue et hermitienne sur* $V\times V$. *Les trois assertions ci-dessous sont équivalentes :*

(i) la formulation variationnelle (9.1) est bien posée ;

(ii)$_{herm}$ *la forme* $a(\cdot,\cdot)$ *vérifie une condition inf-sup ;*

(iii)$_{herm}$ *la forme* $a(\cdot,\cdot)$ *est* $\mathtt{T}$*-coercive au sens de la définition 9.9.*

Démonstration : Tout d'abord, (ii)$_{herm}$ $\Longrightarrow$ (i) d'après le corollaire 2.16.
Puis, (i) $\Longrightarrow$ (iii)$_{herm}$ d'après le théorème 9.6.
Enfin, (iii)$_{herm}$ $\Longrightarrow$ (ii)$_{herm}$: en effet, (9.6) implique que $\mathtt{T}$ est injective. Donc, pour tout $v\in V\setminus\{0\}$, on a $\mathtt{T}v\neq 0$ et, puisque $\|\mathtt{T}v\|_W\leq|||\mathtt{T}|||\,\|v\|_V$, il suit comme précédemment que

$$\frac{|a(v,\mathtt{T}v)|}{\|\mathtt{T}v\|_W}\geq\underline{\alpha}\frac{\|v\|_V^2}{\|\mathtt{T}v\|_W}\geq\frac{\underline{\alpha}}{|||\mathtt{T}|||}\|v\|_V,$$

ce qui permet de retrouver la condition inf-sup (9.3). ∎

9.2 Approximation des formulations variationnelles

L'idée derrière l'introduction de la $\mathtt{T}$-coercivité pour résoudre le problème (9.1) est que l'on peut construire une application linéaire $\mathtt{T}$ *explicite* pour établir (9.5) dans le cas général ($V\neq W$), ou (9.6) dans le cas $V=W$ avec une forme hermitienne. On a déjà réalisé la construction d'applications $\mathtt{T}$ pour les problèmes abstraits à résoudre (diffusion neutronique à une ou deux inconnues, modèle de Stokes, problème avec coefficient changeant de signe).

En outre, ceci nous a permis d'obtenir la convergence de méthodes d'approximation "bien construites". Pour cela, on a fait en sorte que les propriétés nécessaires à l'obtention de la condition inf-sup abstraite soient *transposables après discrétisation*, afin d'obtenir la condition inf-sup discrète uniforme qui assure la convergence via le lemme de Céa (notes de cours 3). On y revient ici, en y ajoutant la T-coercivité discrète uniforme.

9.2.1 Approximation du problème exact

On veut approcher la solution du problème (9.1), que nous supposons bien posé : on introduit $(V_\delta)_\delta$ et $(W_\delta)_\delta$ deux suites d'espaces vectoriels de dimensions finies, avec $V_\delta \subset V$ et $W_\delta \subset W$ pour tout δ (approximation conforme) ; on munit V_δ de la norme $\|\cdot\|_V$, respectivement W_δ de la norme $\|\cdot\|_W$. Par convention, le paramètre δ prend des valeurs strictement positives, et on a $\lim_{\delta\to 0}(\dim(V_\delta)) = \lim_{\delta\to 0}(\dim(W_\delta)) = +\infty$, avec la propriété d'approximabilité rappelée ci-dessous.

Définition 9.11 *La famille $(V_\delta)_\delta$ vérifie la propriété d'approximabilité de V si, et seulement si*

$$\begin{array}{l} \exists V^\star \text{ s.e.v. dense de } V, \quad \forall \delta,\ \exists r_\delta\ :\ V^\star \to V_\delta \text{ tels que} \\ \forall v^\star \in V^\star, \quad \lim_{\delta\to 0} \|v^\star - r_\delta v^\star\|_V = 0. \end{array} \tag{9.7}$$

Remarque 9.12 *Comme on le verra plus tard, au corollaire 9.19, la propriété d'approximabilité est nécessaire pour les éléments de V, mais pas pour ceux de W...*

En particulier, on peut avoir $V = W$, et choisir des espaces $(V_\delta)_\delta$ d'approximation des solutions *différents* des espaces $(W_\delta)_\delta$ d'approximation des fonctions-test. Plus généralement, lorsqu'on choisit $(V_\delta)_\delta \neq (W_\delta)_\delta$, on parle d'approximation de Petrov-Galerkin.
L'approximation du problème (9.1) s'écrit

$$\left\{ \begin{array}{l} \text{Trouver } u_\delta \in V_\delta \text{ tel que} \\ \forall w_\delta \in W_\delta,\ a_\delta(u_\delta, w_\delta) = \langle \ell_\delta, w_\delta \rangle_{W_\delta}, \end{array} \right. \tag{9.8}$$

avec des formes sesquilinéaires $(a_\delta)_\delta$, où a_δ est définie sur $V_\delta \times W_\delta$, et antilinéaires $(\ell_\delta)_\delta$, où ℓ_δ est définie sur W_δ, qui peuvent être différentes des formes exactes $a_{|V_\delta \times W_\delta}$ et $\ell_{|W_\delta}$ (on "généralise" quelque peu le formalisme du §3...). Pour des discrétisations par éléments finis, ceci est le cas si on utilise des formules approchées pour calculer les diverses quantités (formules de quadrature), ou bien si le maillage ne recouvre pas exactement le domaine dans lequel le problème exact est posé. Pour prendre en compte les écarts entre formes approchées et exactes, on introduit, pour tout $\delta > 0$ et tout $v_\delta \in V_\delta$, les *termes de consistance*

$$Cons_{\ell,\delta} = \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|\langle \ell - \ell_\delta, w_\delta \rangle_{W_\delta}|}{\|w_\delta\|_W}, \tag{9.9}$$

$$Cons_{a,\delta}(v_\delta) = \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|(a - a_\delta)(v_\delta, w_\delta)|}{\|w_\delta\|_W}. \tag{9.10}$$

En termes de normes, on a respectivement

$$Cons_{\ell,\delta} = \|\ell_{|W_\delta} - \ell_\delta\|_{W'_\delta}, \text{ et}$$
$$\sup_{v_\delta \in V_\delta \setminus \{0\}} \frac{Cons_{a,\delta}(v_\delta)}{\|v_\delta\|_V} = \sup_{v_\delta \in V_\delta \setminus \{0\}, w_\delta \in W_\delta \setminus \{0\}} \frac{|(a_{|V_\delta \times W_\delta} - a_\delta)(v_\delta, w_\delta)|}{\|v_\delta\|_V \|w_\delta\|_W} = |||a_{|V_\delta \times W_\delta} - a_\delta|||.$$

En règle générale, on écrira ℓ au lieu de $\ell_{|W_\delta}$ et de même a au lieu de $a_{|V_\delta \times W_\delta}$.

On peut aussi reformuler le problème approché (9.8) sous une forme équivalente

$$\begin{cases} \text{Trouver } u_\delta \in V_\delta \text{ tel que} \\ \mathtt{A}_\delta u_\delta = \mathtt{f}_\delta \text{ dans } W_\delta, \end{cases} \tag{9.11}$$

après introduction de $\mathtt{A}_\delta \in \mathcal{L}(V_\delta, W_\delta)$ et $\mathtt{f}_\delta \in W_\delta$ tels que

$$\forall (v_\delta, w_\delta) \in V_\delta \times W_\delta,\ (\mathtt{A}_\delta v_\delta, w_\delta)_W = a_\delta(v_\delta, w_\delta),\ \text{et } \forall w_\delta \in W_\delta,\ (\mathtt{f}_\delta, w_\delta)_W = \langle \ell_\delta, w_\delta \rangle_{W_\delta}.$$

On rappelle que $|||\mathtt{A}_\delta||| = |||a_\delta|||$ pour tout δ.

9.2.2 Caractère bien posé et stabilité des problèmes approchés

A partir de là on peut étudier le caractère bien posé des problèmes approchés (9.8) ou (9.11), ainsi que l'erreur $u - u_\delta$, en prenant notamment en compte les termes de consistance.
Pour δ donné, dire que le problème approché (9.8) ou (9.11) est bien posé équivaut à avoir $\mathtt{A}_\delta$ bijective. Tout d'abord, pour que $\mathtt{A}_\delta \in \mathcal{L}(V_\delta, W_\delta)$ soit bijective, il est nécessaire que :

$$\boxed{\dim(V_\delta) = \dim(W_\delta)},$$

et on fait cette hypothèse (pour tout δ) à partir de maintenant. Voir le §3.3 pour une discussion approfondie. En particulier, on y a noté que lorsque les dimensions de V_δ et W_δ sont égales, alors le problème discret (9.8) ou (9.11) est bien posé si, et seulement si, la condition inf-sup discrète est vérifiée ($\mathtt{A}_\delta$ injective), ou si, et seulement si, la condition de solvabilité discrète est vérifiée ($\mathtt{A}_\delta$ surjective). La condition inf-sup discrète s'écrit :

$$\forall \delta > 0,\ \exists \alpha_\delta > 0,\ \forall v_\delta \in V_\delta,\ \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a_\delta(v_\delta, w_\delta)|}{\|w_\delta\|_W} \geq \alpha_\delta \|v_\delta\|_V. \tag{9.12}$$

Ainsi, la condition (9.12) est vérifiée si, et seulement si,

$$\forall \delta > 0,\ \exists \alpha_\delta > 0,\ \forall v_\delta \in V_\delta,\ \|\mathtt{A}_\delta v_\delta\|_W \geq \alpha_\delta \|v_\delta\|_V,$$

puisque, pour tout $z_\delta \in W_\delta$, on a

$$\|z_\delta\|_W = \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|(z_\delta, w_\delta)|}{\|w_\delta\|_W} = \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|(w_\delta, z_\delta)|}{\|w_\delta\|_W}.$$

D'après le théorème 9.6, si (9.12) est vraie, alors les problèmes approchés sont bien posés dans $V_\delta \times W_\delta$ pour tout δ.
On introduit la transconjuguée $\mathtt{A}_\delta^* \in \mathcal{L}(W_\delta, V_\delta)$ définie par :

$$\forall (v_\delta, w_\delta) \in V_\delta \times W_\delta, \quad (v_\delta, \mathtt{A}_\delta^* w_\delta)_V = a_\delta(v_\delta, w_\delta).$$

On a le même type d'estimation sur $\mathtt{A}_\delta^*$.

Corollaire 9.13 *On suppose la condition inf-sup discrète (9.12) vérifiée : $\mathtt{A}_\delta^*$ est bijective, et*

$$\forall \delta > 0, \forall w_\delta \in W_\delta, \ \|\mathtt{A}_\delta^* w_\delta\|_V \geq \alpha_\delta \|w_\delta\|_W.$$

Démonstration : Soit $\delta > 0$. On sait que $\mathtt{A}_\delta$ est bijective.
Vérifions que $\mathtt{A}_\delta^*$ est également bijective. Puisque $\dim(V_\delta) = \dim(W_\delta)$, il suffit de démontrer l'injectivité : soit $w_\delta \in W_\delta$ tel que $\mathtt{A}_\delta^* w_\delta = 0$. Par définition, on a donc

$$\forall v'_\delta \in V_\delta, \quad 0 = (v'_\delta, \mathtt{A}_\delta^* w_\delta)_V = a_\delta(v'_\delta, w_\delta) = (\mathtt{A}_\delta v'_\delta, w_\delta)_W.$$

Or $\mathtt{A}_\delta$ est surjective dans W_δ, on en déduit donc que $w_\delta = 0$ et $\mathtt{A}_\delta^*$ est bijective comme annoncé.
Soit $y_\delta \in V_\delta$, on évalue maintenant $\|(\mathtt{A}_\delta^*)^{-1} y_\delta\|_W$. On note que, pour tout $v'_\delta \in V_\delta$, on a

$$(\mathtt{A}_\delta v'_\delta, (\mathtt{A}_\delta^*)^{-1} y_\delta)_W \overset{\text{def. } \mathtt{A}_\delta}{=} a_\delta(v'_\delta, (\mathtt{A}_\delta^*)^{-1} y_\delta) \overset{\text{def. } \mathtt{A}_\delta^*}{=} (v'_\delta, \mathtt{A}_\delta^* (\mathtt{A}_\delta^*)^{-1} y_\delta)_V = (v'_\delta, y_\delta)_V.$$

A partir de là, on trouve que

$$\begin{aligned}
\|(\mathtt{A}_\delta^*)^{-1} y_\delta\|_W &= \sup_{w'_\delta \in W_\delta \setminus \{0\}} \frac{|(w'_\delta, (\mathtt{A}_\delta^*)^{-1} y_\delta)_W|}{\|w'_\delta\|_W} \\
(\mathtt{A}_\delta \text{ bijective}) &= \sup_{v'_\delta \in V_\delta \setminus \{0\}} \frac{|(\mathtt{A}_\delta v'_\delta, (\mathtt{A}_\delta^*)^{-1} y_\delta)_W|}{\|\mathtt{A}_\delta v'_\delta\|_W} \\
&= \sup_{v'_\delta \in V_\delta \setminus \{0\}} \frac{|(v'_\delta, y_\delta)_V|}{\|\mathtt{A}_\delta v'_\delta\|_W} \\
(\|\mathtt{A}_\delta v'_\delta\|_W \geq \alpha_\delta \|v'_\delta\|_V) &\leq \frac{1}{\alpha_\delta} \sup_{v'_\delta \in V_\delta \setminus \{0\}} \frac{|(v'_\delta, y_\delta)_V|}{\|v'_\delta\|_V} \\
&= \frac{1}{\alpha_\delta} \|y_\delta\|_V.
\end{aligned}$$

Soit maintenant $w_\delta \in W_\delta$, on déduit de ce qui précède qu'il existe $y_\delta \in V_\delta$ tel que $w_\delta = (\mathtt{A}_\delta^*)^{-1} y_\delta$, et

$$\alpha_\delta \|w_\delta\|_W = \alpha_\delta \|(\mathtt{A}_\delta^*)^{-1} y_\delta\|_W \leq \|y_\delta\|_V = \|\mathtt{A}_\delta^* w_\delta\|_V,$$

ce qui est le résultat attendu. ∎

Dans la suite, on se placera dans le cas où la condition inf-sup discrète est *uniforme*.

Définition 9.14 *La famille des formes $(a_\delta)_\delta$ vérifie une* condition inf-sup discrète uniforme *si, et seulement si*

$$\exists \alpha_\dagger > 0, \ \forall \delta > 0, \ \forall v_\delta \in V_\delta, \quad \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a_\delta(v_\delta, w_\delta)|}{\|w_\delta\|_W} \geq \alpha_\dagger \|v_\delta\|_V. \tag{9.13}$$

On énonce une condition de $\mathtt{T}$-*coercivité discrète* (uniforme).

Définition 9.15 *La famille des formes* $(a_\delta)_\delta$ *est* uniformément $\mathtt{T}_\delta$-coercive *si, et seulement si*

$$\begin{aligned}\exists \alpha^\star, \beta^\star > 0,\ \forall \delta > 0,\ \exists \mathtt{T}_\delta \in \mathcal{L}(V_\delta, W_\delta),\\ |||\mathtt{T}_\delta||| \leq \beta^\star \text{ et } \forall v_\delta \in V_\delta,\ |a_\delta(v_\delta, \mathtt{T}_\delta v_\delta)| \geq \alpha^\star \|v_\delta\|_V^2.\end{aligned} \tag{9.14}$$

On établit maintenant l'équivalence entre la condition inf-sup discrète uniforme, et l'uniforme $\mathtt{T}_\delta$-coercivité des formes $(a_\delta)_\delta$.

Théorème 9.16 (Caractère bien posé) *Soit une famille* $(a_\delta)_\delta$ *de formes sesquilinéaires, continues et uniformément bornées. Les trois assertions ci-dessous sont équivalentes :*

(i) les problèmes discrets (9.8) sont bien posés pour tout δ*, et les applications linéaires* $(\mathtt{A}_\delta^{-1})_\delta$ *sont uniformément bornées ;*

(ii) la famille des formes $(a_\delta)_\delta$ *vérifie une condition inf-sup discrète uniforme ;*

(iii) la famille des formes $(a_\delta)_\delta$ *est uniformément* $\mathtt{T}_\delta$*-coercive.*

Démonstration : (i) $\Longrightarrow$ (iii). On choisit $\mathtt{T}_\delta = \mathtt{A}_\delta$. D'après (i), $\mathtt{T}_\delta$ est bijective.
Puisque les formes $(a_\delta)_\delta$ sont uniformément bornées, on a $|||\mathtt{T}_\delta||| \leq \sup_\delta |||\mathtt{A}_\delta||| < \infty$.
Par hypothèse, $\mathtt{T}_\delta^{-1} = \mathtt{A}_\delta^{-1}$ est telle que $\left|||\mathtt{T}_\delta^{-1}|||\right| \leq C^\star$, avec $C^\star = \sup_\delta \left|||\mathtt{A}_\delta^{-1}|||\right| < \infty$. Soit $v_\delta \in V_\delta$, on a tout d'abord $\|v_\delta\|_V \leq \left|||\mathtt{T}_\delta^{-1}|||\right| \|\mathtt{T}_\delta v_\delta\|_W \leq C^\star \|\mathtt{T}_\delta v_\delta\|_W$, et on trouve

$$a_\delta(v_\delta, \mathtt{T}_\delta v_\delta) = (\mathtt{A}_\delta v_\delta, \mathtt{T}_\delta v_\delta)_W = \|\mathtt{T}_\delta v_\delta\|_W^2 \geq \|v_\delta\|_V^2/(C^\star)^2.$$

Ainsi, $(a_\delta)_\delta$ est uniformément $\mathtt{T}_\delta$-coercive.
(iii) $\Longrightarrow$ (ii). Soit $v_\delta \in V_\delta \setminus \{0\}$, on a $0 < \|\mathtt{T}_\delta v_\delta\|_W \leq \beta^\star \|v_\delta\|_V$, et

$$\sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a_\delta(v_\delta, w_\delta)|}{\|w_\delta\|_W} \geq \frac{|a_\delta(v_\delta, \mathtt{T}_\delta v_\delta)|}{\|\mathtt{T}_\delta v_\delta\|_W} \geq \alpha^\star \frac{\|v_\delta\|_V^2}{\|\mathtt{T}_\delta v_\delta\|_W} \geq \frac{\alpha^\star}{\beta^\star} \|v_\delta\|_V.$$

Il suit que $(a_\delta)_\delta$ vérifie une condition inf-sup discrète uniforme.
(ii) $\Longrightarrow$ (i). D'après le théorème 9.6 et l'égalité $\dim(V_\delta) = \dim(W_\delta)$, si $a_\delta(\cdot,\cdot)$ vérifie une condition inf-sup, le problème discret (9.8) est bien posé. Ainsi $\mathtt{A}_\delta^{-1}$ est bien défini. Pour tout $z_\delta \in W_\delta$, on définit $v_\delta = \mathtt{A}_\delta^{-1} z_\delta \in V_\delta$. On déduit de (9.13) que

$$\|\mathtt{A}_\delta^{-1} z_\delta\|_V = \|v_\delta\|_V \leq \frac{1}{\alpha_\dagger} \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a_\delta(v_\delta, w_\delta)|}{\|w_\delta\|_W} = \frac{1}{\alpha_\dagger} \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|(z_\delta, w_\delta)_W|}{\|w_\delta\|_W} = \frac{1}{\alpha_\dagger} \|z_\delta\|_W,$$

c'est-à-dire que les $(\mathtt{A}_\delta^{-1})_\delta$ sont uniformément bornées. ∎

9.2.3 Etude d'erreur

A partir de ce qui précède, on obtient en outre une estimation de l'erreur si (9.13) ou (9.14) est vraie, incorporant les termes de consistance. On généralise ici les lemmes de Céa déjà obtenus au §3.

Théorème 9.17 (Lemme de Céa) *Sous les hypothèses du théorème 9.16, on a l'estimation d'erreur*

$$\exists C>0,\ \forall \ell\in W',\ \forall \delta,\quad \|u-u_\delta\|_V \le C\left\{Cons_{\ell,\delta}+\inf_{v_\delta\in V_\delta}\Big(\|u-v_\delta\|_V+Cons_{a,\delta}(v_\delta)\Big)\right\}. \quad (9.15)$$

Démonstration : Par hypothèse, (9.13) est vraie. Pour tout $v_\delta\in V_\delta$, il existe donc $w_\delta\in W_\delta$ tel que

$$\alpha_\dagger\|u_\delta-v_\delta\|_V\|w_\delta\|_V\le|a_\delta(u_\delta-v_\delta,w_\delta)|.$$

Par ailleurs, on a les égalités successives :

$$\begin{aligned} a_\delta(u_\delta-v_\delta,w_\delta) &= \langle\ell_\delta,w_\delta\rangle_{W_\delta}-a_\delta(v_\delta,w_\delta)\\ &= \langle\ell_\delta-\ell,w_\delta\rangle_{W_\delta}+\langle\ell,w_\delta\rangle_W-a_\delta(v_\delta,w_\delta)\\ &= \langle\ell_\delta-\ell,w_\delta\rangle_{W_\delta}+a(u,w_\delta)-a_\delta(v_\delta,w_\delta)\\ &= \langle\ell_\delta-\ell,w_\delta\rangle_{W_\delta}+a(u-v_\delta,w_\delta)+(a-a_\delta)(v_\delta,w_\delta). \end{aligned}$$

Et, par définition des termes de consistance, on trouve

$$\|u_\delta-v_\delta\|_V\le\frac{1}{\alpha_\dagger}(Cons_{\ell,\delta}+|||a|||\,\|u-v_\delta\|_V+Cons_{a,\delta}(v_\delta)).$$

D'après l'inégalité triangulaire, on a $\|u-u_\delta\|_V\le\|u-v_\delta\|_V+\|u_\delta-v_\delta\|_V$, et on en déduit finalement (9.15) avec $C=\max\left(\frac{1}{\alpha_\dagger},\frac{|||a|||}{\alpha_\dagger}+1\right)>0$. ∎

Remarque 9.18 *Si les termes de consistance sont nuls, on aboutit à (9.15) avec $C=\frac{|||a|||}{\alpha_\dagger}+1$: on a retrouvé le résultat du lemme de Céa (le théorème 3.13).*

Corollaire 9.19 *Sous les hypothèses du théorème 9.16, on a convergence de l'approximation pour toute donnée $\ell\in W'$, c'est-à-dire que $\lim_{\delta\to0}\|u-u_\delta\|_V=0$, si :*

— *d'une part les termes de consistance tendent vers 0 :*

$$\lim_{\delta\to0}Cons_{\ell,\delta}=0\quad et\quad \lim_{\delta\to0}|||a-a_\delta|||=0\,;$$

— *d'autre part, la famille $(V_\delta)_\delta$ vérifie la propriété d'approximabilité (9.7).*

Démonstration : On utilise (9.15). Par hypothèse, il est clair que $Cons_{\ell,\delta}$ et $\inf_{v_\delta\in V_\delta}\|u-v_\delta\|_V$ tendent vers 0 lorsque δ tend vers 0, mais le terme de consistance $Cons_{a,\delta}(v_\delta)$ doit être traité avec précaution, puisqu'il dépend de v_δ qui intervient aussi dans $\|u-v_\delta\|_V$. On écrit donc que

$$\forall\delta,\quad \|u-u_\delta\|_V\le C\left\{Cons_{\ell,\delta}+\inf_{v_\delta\in V_\delta\ tq\ \|v_\delta\|_V\le2\|u\|_V}\Big(\|u-v_\delta\|_V+Cons_{a,\delta}(v_\delta)\Big)\right\}.$$

En effet, on note que, pour toute suite $(v_\delta^\star)_\delta$ approchant u, c'est-à-dire $\lim_{\delta\to0^+}\|u-v_\delta^\star\|_V=0$, il existe $\delta_0>0$ tel que, pour tout $\delta\le\delta_0$, on a nécessairement $\|v_\delta^\star\|_V\le2\|u\|_V$. A partir de là, on a les majorations

$$\begin{aligned} \forall\delta\le\delta_0,\quad \|u-u_\delta\|_V &\le C\Big\{Cons_{\ell,\delta}+\|u-v_\delta^\star\|_V+Cons_{a,\delta}(v_\delta^\star)\Big\}\\ &\le C\Big\{Cons_{\ell,\delta}+\|u-v_\delta^\star\|_V+|||a-a_\delta|||\|v_\delta^\star\|_V\Big\}\\ &\le C\Big\{Cons_{\ell,\delta}+\|u-v_\delta^\star\|_V+2|||a-a_\delta|||\|u\|_V\Big\} \end{aligned}$$

et on en déduit le résultat annoncé. ■

Si $a_\delta = a$ pour tout $\delta > 0$, l'approche historique est d'utiliser le résultat (voir [18])

Théorème 9.20 (Lemme de Fortin) *On suppose la condition inf-sup (9.3) vérifiée. Alors la forme a vérifie une condition inf-sup discrète uniforme si, et seulement si*

$$\begin{array}{l} \exists \beta' > 0 \quad et \quad \forall \delta,\ \exists \pi_\delta \in \mathcal{L}(W, W_\delta) \text{ tels que} \\ \forall \delta,\ |||\pi_\delta||| \leq \beta' \quad et \quad \forall \delta,\ \forall w \in W,\ \forall v_\delta \in V_\delta, \quad a(v_\delta, \pi_\delta w - w) = 0. \end{array} \tag{9.16}$$

Démonstration : On suppose (9.16) vérifiée. Soit $v_\delta \in V_\delta \setminus \{0\}$:

$$\begin{aligned}
\sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a(v_\delta, w_\delta)|}{\|w_\delta\|_W} &\geq \sup_{w \in W \ tq \ \pi_\delta w \neq 0} \frac{|a(v_\delta, \pi_\delta w)|}{\|\pi_\delta w\|_W} \\
\text{(le sup est positif)} &= \sup_{w \in W \ tq \ \pi_\delta w \neq 0 \ et \ a(v_\delta, \pi_\delta w) \neq 0} \frac{|a(v_\delta, \pi_\delta w)|}{\|\pi_\delta w\|_W} \\
(\forall w \in W,\ a(v_\delta, \pi_\delta w) = a(v_\delta, w)) &= \sup_{w \in W \ tq \ \pi_\delta w \neq 0 \ et \ a(v_\delta, w) \neq 0} \frac{|a(v_\delta, w)|}{\|\pi_\delta w\|_W} \\
(\forall w \in W,\ \|\pi_\delta w\|_W \leq \beta' \|w\|_W) &\geq \frac{1}{\beta'} \sup_{w \in W \ tq \ \pi_\delta w \neq 0 \ et \ a(v_\delta, w) \neq 0} \frac{|a(v_\delta, w)|}{\|w\|_W} \\
(a(v_\delta, w) \neq 0 \Longrightarrow \pi_\delta w \neq 0) &= \frac{1}{\beta'} \sup_{w \in W \ tq \ a(v_\delta, w) \neq 0} \frac{|a(v_\delta, w)|}{\|w\|_W} \\
\text{(le sup est positif)} &= \frac{1}{\beta'} \sup_{w \in W \setminus \{0\}} \frac{|a(v_\delta, w)|}{\|w\|_W} \\
\text{(d'après (9.3))} &\geq \frac{\alpha'}{\beta'} \|v_\delta\|_V.
\end{aligned}$$

Réciproquement, on suppose que (9.13) est vraie :

$$\exists \alpha_\dagger > 0,\ \forall \delta > 0,\ \forall v_\delta \in V_\delta, \quad \sup_{w_\delta \in W_\delta \setminus \{0\}} \frac{|a(v_\delta, w_\delta)|}{\|w_\delta\|_W} \geq \alpha_\dagger \|v_\delta\|_V.$$

Pour construire les applications linéaires $(\pi_\delta)_\delta$ vérifiant (9.16), on utilise les propriétés de la transconjuguée $\mathtt{A}_\delta^*$ de $\mathtt{A}_\delta$, cf. le corollaire 9.13. On fixe $\delta > 0$. Pour $w \in W$, la forme anti-linéaire $v_\delta \mapsto \overline{a(v_\delta, w)}$ appartient à V_δ' : il existe $y_\delta \in V_\delta$ unique tel que

$$\forall v_\delta \in V_\delta, \quad (y_\delta, v_\delta)_V = \overline{a(v_\delta, w)}.$$

Puisque $\mathtt{A}_\delta^*$ est une bijection de W_δ dans V_δ, il existe $w_\delta \in W_\delta$ unique tel que $y_\delta = \mathtt{A}_\delta^* w_\delta$. Et par construction, on a pour tout $v_\delta \in V_\delta$,

$$a(v_\delta, w) = \overline{(y_\delta, v_\delta)_V} = \overline{(\mathtt{A}_\delta^* w_\delta, v_\delta)_V} = (v_\delta, \mathtt{A}_\delta^* w_\delta)_V \overset{\text{def. } \mathtt{A}_\delta^*}{=} a(v_\delta, w_\delta).$$

A partir de là, on introduit l'application $\pi_\delta : w \mapsto w_\delta$, qui est linéaire de W dans W_δ. Et, pour finir, on a, pour tout $w \in W$,

$$\begin{aligned}
\alpha_\dagger \|\pi_\delta w\|_W &\leq \|\mathtt{A}_\delta^*(\pi_\delta w)\|_V \\
&= \sup_{v_\delta \in V_\delta \setminus \{0\}} \frac{|(v_\delta, \mathtt{A}_\delta^*(\pi_\delta w))_V|}{\|v_\delta\|_V} \\
\text{(définition de } \mathtt{A}_\delta^*) &= \sup_{v_\delta \in V_\delta \setminus \{0\}} \frac{|a(v_\delta, \pi_\delta w)|}{\|v_\delta\|_V} \\
\text{(définition de } \pi_\delta w) &= \sup_{v_\delta \in V_\delta \setminus \{0\}} \frac{|a(v_\delta, w)|}{\|v_\delta\|_V} \\
&\leq |||a||| \, \|w\|_W,
\end{aligned}$$

ce qui prouve que le module de continuité de π_δ est (uniformément) borné par $\beta' = \alpha_\dagger^{-1}|||a|||$. ■

On revient maintenant à l'approche utilisant la $\mathtt{T}$-coercivité. Par hypothèse, le problème (9.1) est bien posé. Selon le théorème 9.6, il existe une application linéaire $\mathtt{T} \in \mathcal{L}(V, W)$ telle que la forme $a(\cdot, \cdot)$ est $\mathtt{T}$-coercive. On utilise la connaissance d'une telle application linéaire $\mathtt{T}$ dans les deux résultats ci-dessous.

Corollaire 9.21 *Soit $\mathtt{T} \in \mathcal{L}(V, W)$ réalisant la $\mathtt{T}$-coercivité. On suppose que $\mathtt{T}V_\delta \subset W_\delta$ pour tout δ, et que $\lim_{\delta \to 0} |||a - a_\delta||| = 0$. Alors il existe $\delta_0 > 0$ tel que la famille $(a_\delta)_{\delta \leq \delta_0}$ est uniformément $\mathtt{T}_\delta$-coercive, et on a l'estimation d'erreur*

$$\exists C > 0, \delta_0 > 0, \forall \delta \leq \delta_0, \ \|u - u_\delta\|_V \leq C \Big\{ Cons_{\ell,\delta} + \inf_{v_\delta \in V_\delta} \Big(\|u - v_\delta\|_V + Cons_{a,\delta}(v_\delta) \Big) \Big\}. \quad (9.17)$$

Démonstration : Si on choisit $\mathtt{T}_\delta = \mathtt{T}_{|V_\delta}$ pour tout δ, on a pour tout $v_\delta \in V_\delta$:

$$\begin{aligned}
|a_\delta(v_\delta, \mathtt{T}_\delta v_\delta)| = |a_\delta(v_\delta, \mathtt{T}v_\delta)| &= |a(v_\delta, \mathtt{T}v_\delta) + (a_\delta - a)(v_\delta, \mathtt{T}v_\delta)| \\
&\geq (\underline{\alpha} - |||a_\delta - a||| \, |||\mathtt{T}|||) \|v_\delta\|_V^2.
\end{aligned}$$

Par hypothèse, il existe $\delta_0 > 0$ tel que $\underline{\alpha} - |||a_\delta - a||| \, |||\mathtt{T}||| \geq \frac{1}{2}\underline{\alpha}$ pour tout $\delta \in]0; \delta_0]$. Si on pose $\alpha^\star = \frac{1}{2}\underline{\alpha}$, on en déduit que $(a_\delta)_{\delta \leq \delta_0}$ est uniformément $\mathtt{T}_\delta$-coercive selon la définition 9.15 et, en reprenant la démonstration du lemme de Céa, on aboutit ensuite à (9.17). ■

On note que, puisque $\ker(\mathtt{T}) = \{0\}$ et $\dim(V_\delta) = \dim(W_\delta)$, la condition $\mathtt{T}V_\delta \subset W_\delta$ est équivalente à la condition $\mathtt{T}V_\delta = W_\delta$ d'après le théorème du rang.
Enfin, si on ne peut pas utiliser directement l'application $\mathtt{T}$ pour obtenir la $\mathtt{T}$-coercivité discrète uniforme, on peut utiliser le résultat plus général ci-dessous.

Corollaire 9.22 *Soit $\mathtt{T} \in \mathcal{L}(V, W)$ réalisant la $\mathtt{T}$-coercivité. On suppose qu'il existe une famille d'applications linéaires et continues $(\mathtt{T}_\delta)_\delta$ avec $\mathtt{T}_\delta \in \mathcal{L}(V_\delta, W_\delta)$, telle que $|||\mathtt{T}_{|V_\delta} - \mathtt{T}_\delta||| \to 0$, et que $\lim_{\delta \to 0} |||a - a_\delta||| = 0$. Alors il existe $\delta_0 > 0$ tel que la famille $(a_\delta)_{\delta \leq \delta_0}$ vérifie la $\mathtt{T}$-coercivité discrète, et on a l'estimation d'erreur (9.17).*

Démonstration : Pour tout δ et pour tout $v_\delta \in V_\delta$, on a :

$$\begin{aligned}
|a_\delta(v_\delta, \mathtt{T}_\delta v_\delta)| &= |a(v_\delta, \mathtt{T}_\delta v_\delta) + (a_\delta - a)(v_\delta, \mathtt{T}_\delta v_\delta)| \\
&= |a(v_\delta, \mathtt{T} v_\delta) - a(v_\delta, (\mathtt{T}_{|V_\delta} - \mathtt{T}_\delta)v_\delta) + (a_\delta - a)(v_\delta, \mathtt{T}_\delta v_\delta)| \\
&\geq |a(v_\delta, \mathtt{T} v_\delta)| - |a(v_\delta, (\mathtt{T}_{|V_\delta} - \mathtt{T}_\delta)v_\delta)| - |(a_\delta - a)(v_\delta, \mathtt{T}_\delta v_\delta)| \\
&\geq (\underline{\alpha} - |||a||| \, |||\mathtt{T}_{|V_\delta} - \mathtt{T}_\delta||| - |||a - a_\delta||| \, |||\mathtt{T}_\delta|||) \, \|v_\delta\|_V^2.
\end{aligned}$$

Par hypothèse, il existe $\delta_1 > 0$ tel que $|||\mathtt{T}_\delta||| \leq 2|||\mathtt{T}|||$ pour tout $\delta \in]0; \delta_1]$. Puis il existe $\delta_0 \in]0, \delta_1]$ tel que $\alpha^\star = \underline{\alpha} - |||a||| \, |||\mathtt{T}_{|V_\delta} - \mathtt{T}_\delta||| - 2 \, |||a - a_\delta||| \, |||\mathtt{T}||| \geq \frac{1}{2}\underline{\alpha}$ pour tout $\delta \in]0; \delta_0]$. Le résultat suit. ∎

9.3 Bref retour sur le problème avec coefficient changeant de signe

On revient au problème (P) sous les hypothèses (7.1)-(7.3), et pour un coefficient σ constant par morceaux (7.4).

9.3.1 Géométries simples

L'exemple le plus simple est celui d'une géométrie (forme du domaine Ω) symétrique par rapport à l'interface Σ, pour un contraste différent de -1, voir le §8.1. Lorsqu'on discrétise par éléments finis (espaces discrets $(V_h)_h$), on rappelle que si on a $\mathtt{T}V_h \subset V_h$ pour une application $\mathtt{T}$ réalisant la $\mathtt{T}$-coercivité, alors d'après le corollaire 9.21 on a automatiquement convergence... Or on a vu que pour le cas de la géométrie symétrique, on utilisait une transformation basée sur la symétrie par rapport à l'hyperplan contenant Σ (voir les théorèmes 8.1 et 8.2). Ceci nous indique que si on utilise des maillages symétriques par rapport à l'hyperplan (les *maillages* $\mathtt{T}$*-conformes*), alors on doit avoir convergence !

9.3.2 Géométries quelconques

Dans le cas d'une interface "quelconque" (voir le §8.3.3), lorsque le contraste n'appartient pas à l'intervalle critique, on peut établir une inégalité de Gärding (8.9).
Si on en revient au problème abstrait (9.1) avec $V = W$, on énonce

Définition 9.23 *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire et continue sur $V \times V$. $a(\cdot,\cdot)$ est* $\mathtt{T}$-coercive+compacte *si, et seulement si,*

$$\begin{array}{l} \exists \mathtt{C} \in \mathcal{L}(V) \text{ compacte} \\ \exists \mathtt{T} \in \mathcal{L}(V) \text{ bijective} \end{array}, \ \exists \underline{\alpha} > 0, \ \underline{\beta} \in \mathbb{R}, \ \forall v \in V, \ |a(v, \mathtt{T}v)| \geq \underline{\alpha}\|v\|_V^2 - \underline{\beta}\|\mathtt{C}v\|_V^2. \tag{9.18}$$

On peut établir le résultat (voir par exemple [10, Chapitre 2])

Lemme 9.24 *Soit $a(\cdot,\cdot)$ une forme sesquilinéaire, continue et hermitienne sur $V \times V$. Les deux assertions ci-dessous sont équivalentes :*

(i) la formulation variationnelle (9.1) est bien posée au sens de Fredholm ;
(ii) la forme $a(\cdot,\cdot)$ est `T`*-coercive+compacte.*

Dans ce contexte, si le problème exact (9.1) est bien posé (au sens d'Hadamard), on peut généraliser les corollaires 9.21 ou 9.22 pour assurer la convergence des solutions discrètes vers la solution exacte [11, 4].

Bibliographie

[1] **R.A. Adams, J.J.F. Fournier**, *Sobolev spaces. Second edition*, Academic Press, New York (2003).

[2] **C. Amrouche, P.G. Ciarlet, C. Mardare**, On a lemma of Jacques-Louis Lions and its relation to other fundamental results, *J. Math. Pures Appl.*, **104**, 207–226 (2015).

[3] **D. Boffi, F. Brezzi, M. Fortin**, *Mixed and hybrid finite element methods and applications*, Springer-Verlag (2013).

[4] **A.-S. Bonnet-Ben Dhia, C. Carvalho, P. Ciarlet Jr.**, Mesh requirements for the finite element approximation of problems with sign-changing coefficients, *Numer. Math.*, **138**, 801–838 (2018).

[5] **A.-S. Bonnet-Ben Dhia, L. Chesnel, P. Ciarlet Jr.**, T-coercivity for scalar interface problems between dielectrics and metamaterials, *Math. Mod. Num. Anal.*, **46**, 1363–1387 (2012).

[6] **A.-S. Bonnet-Ben Dhia, P. Ciarlet Jr., C. M. Zwölf**, Time harmonic wave diffraction problems in materials with sign-shifting coefficients, *J. Comput. Appl. Math.*, **234**, 1912–1919 (2010). (Corrigendum *J. Comput. Appl. Math.*, **234**, 2616 (2010))

[7] **G. Bouchitté, C. Bourel, D. Felbacq**, Homogenization near resonances and artificial magnetism in 3D dielectric metamaterials, *Archive for Rational Mechanics and Analysis*, **225**, 1233–1277, (2017).

[8] **H. Brezis**, *Analyse fonctionnelle. Théorie et applications*, Masson, Paris (1983). *Functional analysis, Sobolev spaces and Partial Differential Equations*, Universitext, Springer (2011).

[9] **J. Bussac, P. Reuss**, *Traité de neutronique*, Hermann, Paris (1985).

[10] **C. Carvalho**, *Etude mathématique et numérique de structures plasmoniques avec des coins*, Thèse de l'Ecole Polytechnique (2015).

[11] **L. Chesnel, P. Ciarlet Jr.**, T-coercivity and continuous Galerkin methods : application to transmission problems with sign changing coefficients, *Numer. Math.*, **124**, 1–29 (2013).

[12] **P. Ciarlet Jr.**, *Notes de cours sur les équations de Maxwell et leur approximation*, Notes de cours (2022).

[13] **P. Ciarlet Jr., E. Jamelot**, *Outils mathématiques et algorithmiques pour le calcul scientifique*, Notes de cours (2019).

[14] **P. Ciarlet Jr., E. Jamelot, F. Kpadonou**, Domain decomposition methods for the diffusion equation with low-regularity solution, *Computers Math. Applic.*, **74**, 2369–2384 (2017).

[15] **P. Ciarlet Jr., E. Lunéville**, *La méthode des éléments finis : de la théorie à la pratique*, Editions ISTE, Coll. Méthodes Numériques (2022).

[16] **M. Costabel, M. Dauge**, On the inequalities of Babuska-Aziz, Friedrichs and Horgan-Payne, *Arch. Rational Mech. Anal.*, **217**, 873–898 (2015).

[17] **M. Costabel, E. Stephan**, A direct boundary integral method for transmission problems, *J. of Math. Anal. and Appl.*, **106**, 367–413 (1985).

[18] **V. Girault, P.-A. Raviart**, *Finite element methods for Navier-Stokes equations*, Springer Series in Computational Mathematics, **5**, Springer Verlag, Berlin (1986).

[19] **P. Grisvard**, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics, 24, Pitman, London (1985).

[20] **S. Nicaise, J. Venel**, A posteriori error estimates for a finite element approximation of transmission problems with sign changing coefficients, *J. Comput. Appl. Math.*, **235**, 4272–4282 (2011).

[21] **J.B. Pendry**, Negative refraction makes a perfect lens, *Phys. Rev. Letters*, **85**, 3966–3969 (2000).

[22] **V.G. Veselago**, The electrodynamics of substances with simultaneously negative values of ε and μ, *Soviet Physics-Uspekhi*, **10**, 509–514 (1968).

Index

application linéaire
  compacte, 3
  continue, 3
approximation
  conforme, 33
  convergence, 37, 39
  vitesse de convergence, 37

carte locale, 10
condition
  de solvabilité, 25
  de stabilité, 25
  de stabilité discrète, 37
  de stabilité discrète uniforme, 39
  inf-sup, 25
  inf-sup inversée, 68
  inf-sup inversée discrète uniforme, 71
  inf-sup discrète, 37
  inf-sup discrète uniforme, 39, 112
consistance, 110
constante inf-sup, 25, 107
  de Ladyzhenskaya-Babuska-Brezzi, 71
corollaire
  Banach-Necas-Babuska, 28
  Banach-Necas-Babuska discret, 38
crochets de dualité, 3, 6

dispersion
  métamatériau, 81, 82
  plasmon, 80
  vide, 81
distribution, 6
  convergence, 6
  dérivée, 7
distribution vectorielle, 6
domaine, 10

éléments finis
  Lagrange, 41
  Lagrange - calculs élémentaires, 41
  Lagrange - degrés de liberté, 41
  Lagrange vectoriel, 57
  MINI, 65
  Raviart-Thomas, 44
  Raviart-Thomas - calculs élémentaires, 47
  Raviart-Thomas - degrés de liberté, 45
  Taylor-Hood, 66
erreur, 35
espace
  Banach, 3
  Hilbert, 3
  $H^1(\cdot)$, 7
  $H^m(\cdot)$, $m \in \mathbb{N}$, 7
  $H_0^m(\cdot)$, $m \in \mathbb{N}$, 8
  $\boldsymbol{H}(\mathrm{div}\,;\cdot)$, 14
  $L^p(\cdot)$, 4
  réflexif, 3
  séparable, 2
exposant
  de décalage, 42

flux, 11
fonction
  lipschitzienne, 2
  support, 2
forme
  coercive, 21, 22
  continue, 3

hermitienne, 21
symétrique, 22
T-coercive, 32, 86, 107, 109, 113
formulation variationnelle
bien posée, 24
diffusion, 20, 29
mixte, 67
mixte discrète, 71
Stokes, 51, 52
formule de Green, 10
frontière
lipschitzienne, 10
régulière, 10
vecteur normal, 10
frontière
bornée, 11

inégalité
Young, 60
intégration par parties, 10, 15

lemme
Céa, 35, 38, 114
Fortin, 72, 115

maillage
famille régulière, 41
pas, 40
T-conforme, 117
matériau
métal, 76
métamatériau, 77
mode plasmonique, 104
modèle
diffusion neutronique, 19
Drude-Lorentz, 73
problème de transmission, 84
Stokes, 49
modèle mathématique, 1
bien posé (au sens de Fredholm), 101
bien posé (au sens de Hadamard), 24
bien posé, 1

opérateur d'interpolation, 42
opérateur de dérivation
divergence, 1, 9
gradient, 1, 8
Laplacien, 2
rotationnel, 2
ouvert
borné, 5

partition, 42
pénalisation, 56
plasmon, 78
prolongement par continuité, 4
propriété
d'approximabilité, 36
de décalage, 42

réfraction négative, 80
régularité par morceaux, 42, 57

saut, 17

théorème
Babuska-Brezzi, 71
Banach-Necas-Babuska, 25
de Rham, 53
isomorphisme de Banach, 4
Lax-Milgram, 21
Riesz, 3
trace, 11
trace normale, 15
trace, 11
trace normale, 15
transformation de Piola, 47