Virginie Ehrlacher & Mi-Song Dupuy

Méthodes de tenseurs pour les problèmes en grande dimension

24 janvier 2024

Sorbone Université - M2 Mathématiques et Applications

Table des matières

1

Introduction

Le but de ce cours est de vous présenter les fondements théoriques et les principaux algorithmes utilisés d'une famille de méthodes, appelées *méthodes de tenseurs* , *méthodes tensorielles* ou *méthodes de faible rang*, utilisées pour traiter efficacement des problèmes en grande dimension.

On peut rencontrer ce type de problèmes dans des champs d'application extrêmement divers: physique, mécanique, économie, sciences sociales, chimie, etc. Nous allons en donner quelques exemples précis dès la prochaine section.

1.1 Malédiction de la dimensionalité

Mais commençons tout d'abord par préciser ce qui est entendu par *problèmes en grande dimension* dans le cadre de ce cours. Par problèmes en grande dimension, nous voulons désigner des problèmes, définis par un système d'équations mathématiques, et dont la solution est

- soit un tenseur (dimension finie), c'est-à-dire un élément

$$U := (U_{i_1,i_2,\dots,i_d})_{1\leq i_1\leq N_1, 1\leq i_2\leq N_2,\dots,1\leq i_d\leq N_d} \in \mathbb{R}^{N_1\times N_2\times\dots\times N_d}$$

  où $d \in \mathbb{N}^*$ est appelé *l'ordre* du tenseur et où $N_1,\dots,N_d \in \mathbb{N}^*$; identifier l'ensemble des composantes de ce tenseur par une approche naïve requiert d'identifier $N_1 \times N_2 \times \dots \times N_d$ scalaires. Dans le cas simplifié où $N_1 = N_2 = \dots = N_d = N$ pour $N \in \mathbb{N}^*$, cela implique d'identifier N^d scalaires.
- soit une fonction u dépendant de d variables $x_1,\dots,x_d$ appartenant respectivement à des ensembles $\Omega_1,\dots,\Omega_d$, sous-ensembles boréliens de $\mathbb{R}^{p_1},\dots,\mathbb{R}^{p_d}$ respectivement avec $p_1,\dots,p_d \in \mathbb{N}^*$. On a alors

$$u : \begin{cases} \Omega_1 \times \dots \times \Omega_d \to & \mathbb{K}, \\ (x_1,\dots,x_d) & \mapsto u(x_1,\dots,x_d). \end{cases}$$

  La fonction u est alors en général définie comme la solution d'une équation aux dérivées partielles définie sur le domaine $\Omega := \Omega_1 \times \dots \times \Omega_d$.

Quel est le lien entre fonctions et tenseurs? Supposons que nous cherchions ici à calculer une approximation numérique d'une fonction

$$u : \begin{cases} \Omega_1 \times \dots \Omega_d \to & \mathbb{R} \\ (x_1,\dots,x_d) \mapsto & u(x_1,\dots,x_d) \end{cases}$$

dépendant de d variables comme celle que nous avons introduite ci-dessus. Une approche numérique naïve consiste alors à introduire, pour tout $1 \leq j \leq d$, une famille libre de N_j fonctions $\phi^j_1(x_j),\dots,\phi^j_{N_j}(x_j)$ (avec $N_j \in \mathbb{N}^*$) définies sur Ω_j et ne dépendant que de la variable x_j, i.e. telles que

$$\forall 1 \leq i_j \leq N_j, \quad \phi^j_{i_j} : \begin{cases} \Omega_j \to & \mathbb{R} \\ x_j \mapsto & \phi^j_{i_j}(x_j). \end{cases}$$

Un exemple classique dans le cas où Ω_j est un intervalle de $\mathbb{R}$ (où un domaine de $\mathbb{R}^{p_j}$) est de choisir la famille $\phi_1^j,\ldots,\phi_{N_j}^j$ comme la famille des fonctions éléments finis $\mathbb{P}_1$ associée à un maillage du domaine Ω_j.

La fonction $u(x_1,\ldots,x_d)$ est alors approchée comme une combinaison linéaire de tous les produits de fonctions à une variable de la forme $\phi_{i_1}^1(x_1)\phi_{i_2}^2(x_2)\ldots\phi_{i_d}^d(x_d)$, soit

$$u(x_1,\ldots,x_d) \approx \sum_{i_1=1}^{N_1}\ldots\sum_{i_d=1}^{N_d} U_{i_1,\ldots,i_d}\phi_{i_1}^1(x_1)\phi_{i_2}^2(x_2)\ldots\phi_{i_d}^d(x_d),$$

où $U := (U_{i_1,\ldots,i_d})_{1\leq i_1\leq N_1,\ldots,1\leq i_d\leq N_d} \in \mathbb{R}^{N_1\times\ldots\times N_d}$ est un tenseur d'ordre d à identifier. Dans le cas où $N_1 = \ldots = N_d = N$, il faut donc identifier N^d scalaires pour déterminer l'approximation numérique de la fonction u.

On voit donc qu'approcher numériquement une fonction dépendant de d variables par des approches naïves ou standards nécessitent de résoudre des problèmes dont le nombre d'inconnues dépend de manière **exponentielle** du nombre de variables dont la fonction dépend. Bien sûr, lorsque d le nombre de variables est grand, de telles approches sont juste impossibles à mettre en oeuvre numériquement: ce phénomène s'appelle la *malédiction de la dimensionalité* ou encore le *fléau de la dimension* (*curse of dimensionality* en anglais), d'après le nom qui lui a été donné par Richard Bellman en 1961. Il est alors nécessaire de faire appel à des méthodes numériques dédiées pour la résolution des problèmes en grande dimension, qui permettent de réduire la complexité de tels systèmes de manière systématique, tout en conservant une bonne qualité d'approximation des modèles obtenus. Les méthodes de tenseurs qui font l'objet de ce cours sont une famille de méthodes numériques dont l'objectif est précisément de contourner cette malédiction de la dimensionalité.

1.2 Exemples de problèmes en grande dimension

Donnons ici quelques exemples d'équations aux dérivées partielles dont les solutions sont des fonctions dépendant d'un grand nombre de variables.

1.2.1 Équation de Fokker-Planck

Un premier exemple d'équations aux dérivées partielles en grande dimension est l'équation de Fokker-Planck. Celle-ci est en particulier extrêmement importante en physique statistique ou en finance. Cette équation permet de modéliser la distribution de probabilités de la solution d'une équation différentielle stochastique en dimension $d \in \mathbb{N}^*$.

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilités et soit $(X_t(\omega))_{t\geq 0}$ un processus aléatoire à valeurs dans $\mathbb{R}^d$. Pour tout $t \geq 0$ et tout $\omega \in \Omega$, on note $X_t^1(\omega),\ldots,X_t^d(\omega)$ les composantes de $X_t(\omega)$ de telle sorte que $X_t(\omega) = (X_t^1(\omega),\ldots,X_t^d(\omega))$.

Supposons que le processus $(X_t)_{\geq 0}$ soit solution d'ue équation différentielle stochastique de la forme :

$$dX_t = F(X_t)\,dt + \sigma\,dB_t \tag{1.1}$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien en dimension d, $\sigma > 0$ et $F : \mathbb{R}^d \to \mathbb{R}^d$ est une application régulière.

En physique statistique, le processus $(X_t)_{t\geq 0}$ peut représenter l'évolution de la position des atomes composant un système moléculaire. Si le système est composé de N atomes, on a alors $d = 3N$. L'application F modélise alors le champ des forces exercées sur chacun des atomes et $\sigma > 0$ représente un niveau de bruit stochastique qui est proportionnel à la valeur de la température du système.

En finance, $(X_t)_{t\geq 0}$ peut représenter l'évolution des prix d'un portefeuille d'actions au sens où X_t^i représente le prix à l'instant t de l'action i du portefeuille.

Dans ces deux situations, connaître la distribution de probabilité $u(t,x_1,\dots,x_d)$ pour $(t,x_1,\dots,x_d) \in \mathbb{R}_+ \times \mathbb{R}^d$ du processus $(X_t)_{t\geq 0}$ permet de pouvoir estimer toutes les propriétés statistiques de ce dernier. Rappelons que u est définie comme la fonction (pour être parfaitement rigoureux, il s'agit en réalité d'une mesure de probabilités) telle que pour tout ensemble mesurable $B \subset \mathbb{R}^d$,

$$\mathbb{P}\left[X_t \in B\right] = \int_B u(t,x_1,\dots,x_d)\, dx_1 \dots dx_d.$$

Si $(X_t)_{t\geq 0}$ est solution de l'équation différentielle stochastique (1.1), sa densité de probabilité u est alors solution de l'**équation de Fokker-Planck**:

$$\partial_t u - \frac{\sigma^2}{2}\Delta_{x_1,\dots,x_d} u + \mathrm{div}_{x_1,\dots,x_d}(Fu) = 0, \quad \text{dans } \mathbb{R}_+ \times \mathbb{R}^d. \tag{1.2}$$

En pratique, d peut être très grand (soit 3 fois le nombre d'atomes dans un système moléculaire, soit le nombre d'actions dans le portefeuille). En conséquence, l'utilisation de méthodes numériques classiques pour résoudre (1.2) est en général inenvisageable à cause de la malédiction de la dimensionalité.

1.2.2 Équation de Schrödinger

Mentionnons ici un autre exemple tiré d'applications en chimie quantique.

Considérons un système composé de d électrons, considérés comme des particules quantiques de masse m. Leur état, à un instant de temps $t > 0$, est représentée par une fonction $u(t,\cdot) : (\mathbb{R}^3)^d \ni (x_1,\dots,x_d) \mapsto u(t,x_1,\dots,x_d)$, appelée *fonction d'onde* qui a l'interprétation probabiliste suivante. Si $B \subset (\mathbb{R}^3)^d$ est un ensemble mesurable, la quantité $\int_B |u(t,x_1,\dots,x_d)|^2\, dx_1 \dots dx_d$ représente la probabilité de trouver, à un instant t, les électrons à des positions appartenant à l'ensemble B. Noter que cette interprétation implique que u doit vérifier la condition de normalisation

$$\int_{(\mathbb{R}^3)^d} |u(t,x_1,\dots,x_d)|^2\, dx_1 \dots dx_d = 1.$$

L'équation permettant de modéliser l'évolution de la fonction d'onde u, appelée **équation de Schrödinger**, s'écrit

$$i\hbar\partial_t u = -\frac{\hbar^2}{2m}\Delta_{x_1,\dots,x_d} u + Vu,$$

où $\hbar$ est la constante de Planck réduite, et où

$$V : \left\{ \begin{array}{ccc} (\mathbb{R}^3)^d & \to & \mathbb{R} \\ (x_1,\dots,x_d) & \mapsto & V(x_1,\dots,x_d) \end{array} \right.$$

est une fonction représentant le potentiel d'interaction du système d'électrons.

Là encore, il s'agit d'une équation aux dérivées partielles de grande dimension lorsque le nombre d'électrons dans le système est grand, qui ne peut pas être résolue par des approches classiques à cause de la malédiction de la dimensionalité.

1.3 Autres familles de méthodes numériques pour contrer la malédiction de la dimensionalité

Citons ici quelques autres approches utilisées pour contrer la malédiction de la dimensionalité. Celles-ci se décomposent en deux grandes familles de méthodes.

La première grande famille est constituée des approches dites **linéaires**, qui consistent à approcher des fonctions solutions d'équations aux dérivées partielles et dépendant d'un grand nombre de variables dans des *espaces vectoriels de petite dimension*. Tout l'enjeu dans ce type d'approches est de choisir optimalement l'espace vectoriel de petite dimension. Souvent, ce choix est fait en utilisant des notions de *sparsité*. Dans des contextes où la grande dimension du problème est liée au fait que l'on cherche à approcher la solution d'un modèle dépendant d'un grand nombre de paramètres, la *méthode des bases réduites* offre une approche intéressante, en particulier lorsqu'il s'agit d'approcher des problèmes elliptiques ou paraboliques. Le cours de M. Albert Cohen présente une introduction à cette famille de méthodes.

La deuxième grande famille de méthodes est constituée des approches dites *non-linéaires*. Dans ce type d'approches, on cherche à approcher les fonctions de grande dimension dans des ensembles **qui ne sont pas** des espaces vectoriels, mais plus généralement des variétés. Les méthodes de tenseurs qui font l'objet de ce cours sont une famille d'approches non-linéaires. Parmi celles-ci, nous pouvons également citer les approches basées sur l'utilisation de réseaux de neurones et d'apprentissage automatique. Le cours de M. Bruno Després présente une introduction à ce type d'approches.

2

Produits tensoriels d'espaces de Hilbert

Le but de ce chapitre est d'introduire les définitions fondamentales sur les notions de produits tensoriels d'espaces de Hilbert. Deux cas particuliers attireront notre attention: les produits tensoriels d'espaces de dimension finie et les produits tensoriels d'espaces de fonctions comme les espaces de Lebesgue.

2.1 Préliminaires

Nous rappelons au lecteur les notions d'espace complété d'espaces pré-hilbertiens et d'espaces vectoriels normés (sous-entendu pour le corps des réels $\mathbb{R}$).

2.1.1 Isométrie

Nous rappelons ici la définition d'une isométrie.

Définition 2.1. *Soit $(E,\|\cdot\|_E)$ et $(F,\|\cdot\|_F)$ deux espaces vectoriels normés. On dit qu'une application linéaire $J : E \to F$ est une* ***isométrie*** *si et seulement si elle préserve la norme, i.e.*

$$\forall u \in E, \quad \|J(u)\|_F = \|u\|_E.$$

Faisons ici quelques remarques.

Remarque 2.1. *Une isométrie est nécessairement une application injective. En effet, soit $u \in E$ telle que $J(u) = 0$. On a alors $\|u\|_E = \|J(u)\|_F = 0$ et nécessairement $u = 0$. Le noyau de J est donc bien réduit à $\{0\}$.*

En particulier, $J : E \to J(E)$, où $J(E)$ désigne l'image de E par J définit un isomorphisme.

Définition 2.2. *Soit $(E,\|\cdot\|_E)$ et $(F,\|\cdot\|_F)$ deux espaces vectoriels normés. On dit que ces deux espaces sont* ***isométriquement isomorphes*** *si et seulement si il exsite une isométrie $J : E \to F$ telle que $J(E) = F$. On dit que J est un* ***isomorphisme isométrique*** *et on note alors $F \approx E$.*

Remarque 2.2. *On utilise très souvent l'abus de langage suivant: lorsque deux espaces vectoriels normés $(E,\|\cdot\|_E)$ et $(F,\|\cdot\|_F)$ sont isométriquement isomorphes, on identifie souvent E à F et on dit que ces espaces sont égaux ou identiques et on écrira $(E,\langle\cdot,\cdot\rangle_E) = (F,\langle\cdot,\cdot\rangle_F)$. De plus, si $K \subset E$ est un sous-espace vectoriel de E, il pourra arriver d'écrire $K \subset F$, ce qui est un abus de langage pour signifier plus précisément $J(K) \subset F$ où J est une isométrie de E dans F telle que $J(E) = F$.*

On rappelle ici qu'un espace pré-hilbertien est un espace vectoriel muni d'un produit scalaire, mais qui n'est pas nécessairement complet pour la norme associée. Un espace vectoriel muni d'un produit scalaire qui est complet pour la norme associée est un espace de Hilbert.

Proposition 2.1. *Soit $(E,\langle\cdot,\cdot\rangle_E)$ et $(F,\langle\cdot,\cdot\rangle_F)$ deux espaces pré-hilbertiens. Alors une application linéaire $J : E \to F$ est une isométrie si et seulement si elle préserve le produit scalaire, i.e.*

$$\forall u,v \in E, \quad \langle J(u),J(v)\rangle_F = \langle u,v\rangle_E.$$

Démonstration. La preuve est une conséquence directe des identités de polarisation pour les espaces pré-hilbertiens: si $(E,\langle\cdot,\cdot\rangle_E)$ est un espace pré-hilbertien et si on note $\|\cdot\|_E$ la norme associée au produit scalaire $\langle\cdot,\cdot\rangle_E$, alors on a

$$\forall u,v \in E, \quad \langle v,u\rangle_E := \frac{\|u+v\|_E^2 - \|u-v\|_E^2}{4}.$$

2.1.2 Complété d'un espace pré-hilbertien

Définition-Théorème 2.1. *Soit H_0 un espace pré-hilbertien muni d'un produit scalaire $\langle\cdot,\cdot\rangle_{H_0}$, dont on note $\|\cdot\|_{H_0}$ la norme associée. Alors, il existe un espace de Hilbert H (muni d'un produit scalaire $\langle\cdot,\cdot\rangle_H$ et de la norme associée $\|\cdot\|_H$) et une isométrie $J : H_0 \to H$ tels que $J(H_0)$ est dense dans H.*

*L'espace de Hilbert H est unique à isomorphime isométrique près et est appelé le **complété** de H_0.*

Remarque 2.3. *En identifiant l'espace H_0 à son image par J, $J(H_0)$, on peut alors écrire $H_0 \subset H$ (en utilisant un léger abus de notation).*

Démonstration. Nous ne prouverons ici ce résultat que dans le cas où H_0 est un espace pré-hilbertien séparable.

Soit $\mathcal{K} \subset \mathbb{N}$ et $(e_k)_{k\in\mathcal{K}}$ un système orthonormé de H_0 tel que l'ensemble des combinaisons linéaires finies de $\{e_k\}_{k\in\mathcal{K}}$ soit dense dans H_0. On note H_0' l'espace dual de H_0, c'est-à-dire l'ensemble des formes linéaires $f : H_0 \to \mathbb{R}$ continues, i.e. telles que

$$\|f\|_{H_0'} := \sup_{v_0\in H_0,\ \|v_0\|_{H_0}=1} |f(v_0)| < +\infty.$$

Lemme 2.1. *L'espace H_0' est un espace de Hilbert séparable par rapport au produit scalaire*

$$\forall f,g \in H_0', \quad (f,g)_{H_0'} = \sum_{k\in\mathcal{K}} f(e_k)g(e_k),$$

qui engendre la norme $\|\cdot\|_{H_0'}$.

Soit H l'espace dual de H_0'. Alors H est aussi un espace de Hilbert. On note $J : H_0 \to H$ l'application linéaire qui à un élément $v_0 \in H_0$ renvoie la fonctionnelle $F_{v_0} \in H = (H_0')'$ telle que

$$\forall f \in H_0', \quad F_{v_0}(f) = f(v_0).$$

Montrons que J est une application bijective de H_0 dans $J(H_0)$ qui préserve le produit scalaire. En effet, il est facile de vérifier que pour tout $v_0 \in H_0$,

$$\|J(v_0)\|_H = \sup_{f\in H_0',\ \|f\|_{H_0'}=1} |F_{v_0}(f)| = \|v_0\|_{H_0}.$$

Cette relation implique en particulier que le noyau de J est réduit à $\{0\}$ et que J est donc injective. Elle est donc bijective de H_0 dans $J(H_0)$. Par ailleurs, on peut vérifier que J préserve bien le produit scalaire en utilisant l'identité de polarisation.

Il nous reste à montrer que $J(H_0)$ est dense dans H. Supposons qu'il existe $F \in H$ tel que $\langle F,J(v_0)\rangle_H = 0$ pour tout $v_0 \in H_0$. Si $\{f_j\}_{j\in\mathbb{N}}$ est une base orthonormée de H_0', alors on a

$$\begin{aligned}\langle F,J(v_0)\rangle_H &= \sum_{j\in\mathbb{N}} F(f_j)J(v_0)(f_j)\\ &= \sum_{j\in\mathbb{N}} F(f_j)f_j(v_0)\\ &= 0.\end{aligned}$$

Comme cette relation est vraie pour tout $v_0 \in H_0$, cela implique que $\sum_{j\in\mathbb{N}} F(f_j)f_j = 0$ dans H_0'. Ceci implique alors que $F(f_j) = 0$ pour tout $j \in \mathbb{N}$ et donc $F = 0$. L'espace $J(H_0)$ est donc bien dense dans H.

Montrons maintenant que H est unique à isométrie près. Soit $\widetilde{H}$ un espace de Hilbert et $\widetilde{J} : H_0 \to \widetilde{H}$ une isométrie telle que $\widetilde{J}(H_0)$ soit dense dans $\widetilde{H}$.

On considère alors l'application $S : J(H_0) \to \widetilde{J}(H_0)$ comme $S = \widetilde{J} \circ J^{-1}$ où $J^{-1} : J(H_0) \to H_0$ est l'inverse de l'application J restreinte à $J(H_0)$. Il est facile de voir que S est alors une isométrie. C'est donc une application continue et on peut l'étendre de manière unique en une application $S : \overline{J(H_0)} = H \to \overline{\widetilde{J}(H_0)} = \widetilde{H}$ dont on peut facilement vérifier qu'elle est une isométrie. D'où le résultat.

Démonstration (Preuve du Lemme 2.1). Montrons d'abord que pour tout $f \in H_0'$, $\|f\|^2_{H_0'} = (f,f)_{H_0'}$.

Montrons d'abord que pour tout $f \in H_0'$, $(f,f)_{H_0'} \leq \|f\|^2_{H_0'}$.

Pour tout $K \in \mathbb{N}$, on note $v_K := \sum_{k\in\mathcal{K},k\leq K} f(e_k)e_k$. On a alors

$$\|v_K\|^2_{H_0} = \sum_{k\in\mathcal{K},k\leq K} |f(e_k)|^2.$$

De plus,

$$f(v_K) = \sum_{k\in\mathcal{K},k\leq K} |f(e_k)|^2.$$

En conséquence, pour tout $K \in \mathbb{N}$,

$$f\left(\frac{v_K}{\|v_K\|_{H_0}}\right) = \left(\sum_{k\in\mathcal{K},k\leq K} |f(e_k)|^2\right)^{1/2} \leq \|f\|_{H_0'}$$

ce qui implique que

$$\sqrt{(f,f)_{H_0'}} = \sum_{k\in\mathcal{K}} |f(e_k)|^2 \leq \|f\|_{H_0'} < +\infty.$$

Montrons l'inégalité inverse. Soit $\epsilon > 0$ et soit $v_0 \in H_0$ tel que $\|v_0\|_{H_0} = 1$ et $|f(v_0)| \geq \|f\|_{H_0'} - \epsilon$. On a alors $v_0 = \sum_{k\in\mathcal{K}} \langle e_k,v_0\rangle_{H_0} e_k$ avec $\sum_{k\in\mathcal{K}} |\langle e_k,v_0\rangle_{H_0}|^2 = 1$.

Comme f est une forme linéaire continue, on peut montrer que

$$f(v_0) = \sum_{k\in\mathcal{K}} \langle e_k,v_0\rangle f(e_k),$$

ce qui implique que

$$|f(v_0)|^2 \leq \left(\sum_{k\in\mathcal{K}} |\langle e_k,v_0\rangle_{H_0}|^2\right)^{1/2} \left(\sum_{k\in\mathcal{K}} |f(e_k)|^2\right)^{1/2} = \sqrt{(f,f)_{H_0'}}.$$

On a donc pour tout $\epsilon > 0$ que $\|f\|_{H_0'} \leq \sqrt{(f,f)_{H_0'}} + \epsilon$. D'où le résultat.

Montrons maintenant que H_0' munit de ce produit scalaire est un espace de Hilbert. Il suffit donc de montrer qu'il est complet pour la norme associé.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans H_0'. Il est aisé de voir que pour tout $v_0 \in H_0$, la suite $(f_n(v_0))_{n\in\mathbb{N}}$ est alors une suite de Cauchy dans $\mathbb{R}$ qui est complet. Elle est donc convergente et on note sa limite $f(v_0)$. Montrons que l'application $f : H_0 \ni v_0 \mapsto f(v_0)$ appartient à H_0' et que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f dans H_0'. Il est aisé de voir que f est une application linéaire. De plus, comme $(f_n)_{n\in\mathbb{N}}$ est de Cauchy dans H_0', elle est bornée et il existe une constante $C > 0$ telle que pour tout $n \in \mathbb{N}$, $\|f_n\|_{H_0'} \leq C$. Il est aisé de vérifier que nécessairement pour tout $v_0 \in H_0$, $|f(v_0)| \leq C\|v_0\|_{H_0}$ ce qui implique bien que $f \in H_0'$.

On a par ailleurs

$$\forall \epsilon > 0, \exists N_\epsilon \in \mathbb{N}, \forall n \geq N_\epsilon, \forall p \geq 0, \|f_n - f_{n+p}\|_{H_0'} \leq \epsilon,$$

soit

$$\forall \epsilon > 0, \exists N_\epsilon \in \mathbb{N}, \forall n \geq N_\epsilon, \forall p \geq 0, \forall v_0 \in V_0, |f_n(v_0) - f_{n+p}(v_0)| \leq \epsilon\|v_0\|_{H_0}.$$

Ceci implique que

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall v_0 \in V_0, \forall p \geq 0, |f_n(v_0) - f_{n+p}(v_0)| \leq \epsilon\|v_0\|_{H_0},$$

En passant à la limite $p \to +\infty$ dans l'expression ci-dessus, on obtient alors

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall v_0 \in V_0, |f_n(v_0) - f(v_0)| \leq \epsilon\|v_0\|_{H_0},$$

soit encore

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \|f_n - f\|_{H_0'} \leq \epsilon.$$

La suite $(f_n)_{n\in\mathbb{N}}$ converge donc bien vers f dans H_0', donc H_0' est bien un espace de Hilbert.

Il reste à montrer qu'il est séparable. Pour ce faire, il suffit de montrer que l'ensemnle $\{f_k\}_{k\in\mathcal{K}}$ des formes linéaires continues sur H_0 définies par

$$f_k(e_l) = \delta_{kl}$$

pour tout $k,l \in \mathcal{K}$, forme bien une base Hilbertienne de H_0'.

Exercice 2.1. *Soit $(H,\langle\cdot,\cdot\rangle_H)$ un espace de Hilbert et soit H_0 un sous-espace vectoriel de H. Montrer que le complété de $(H_0,\langle\cdot,\cdot\rangle_H)$ est alors $(\overline{H_0}^{\|\cdot\|_H},\langle\cdot,\cdot\rangle_H)$ où $\overline{H_0}^{\|\cdot\|_H}$ désigne la fermeture (ou l'adhérence) de H_0 pour la norme $\|\cdot\|_H$.*

2.1.3 Complété d'un espace vectoriel normé

Ce résultat peut être étendu de manière plus général à des espaces vectoriels normés. Plus précisément, nous admettrons le théorème ci-dessous.

Définition-Théorème 2.2. *Soit V_0 un espace vectoriel normé muni d'une norme $\|\cdot\|$. Alors, il existe un espace de Banach V (muni d'une norme $\|\cdot\|_V$) et une isométrie $J : V_0 \to V$ tels que $J(V_0)$ soit dense dans V.*

*L'espace de Banach V est unique à isomorphisme isométrique près et est appelé le **complété** de V_0.*

2.2 Produits tensoriels d'espaces de Hilbert

Dans toute la suite de ces notes de cours, tous les espaces de Hilbert seront supposés être des espaces de Hilbert *séparables* sauf mention contraire.

2.2.1 Définition

Soit H_1, H_2 deux espaces de Hilbert sur $\mathbb{R}$, munis respectivement des produits scalaires $\langle \cdot,\cdot \rangle_{H_1}$ et $\langle \cdot,\cdot \rangle_{H_2}$. Pour tout $u_1 \in H_1$ et tout $u_2 \in H_2$, on note $u_1 \otimes u_2 : H_1 \times H_2 \to \mathbb{R}$ la forme bilinéaire continue définie par

$$\forall v_1 \in H_1, \forall v_2 \in H_2, \quad u_1 \otimes u_2(v_1,v_2) := \langle u_1,v_1 \rangle_{H_1} \langle u_2,v_2 \rangle_{H_2}.$$

Soit $H_1 \bigotimes_a H_2$ l'espace vectoriel engendré par ces formes (i.e. l'ensemble des combinaisons linéaires finies de telles formes). Cet espace vectoriel est également appelé *produit tensoriel algébrique* des espaces H_1 et H_2.

On munit l'espace $H_1 \bigotimes_a H_2$ de la forme hermitienne $\langle \cdot,\cdot \rangle_\otimes : (H_1 \bigotimes_a H_2) \times (H_1 \bigotimes_a H_2) \to \mathbb{R}$ définie par

$$\begin{aligned} \forall u_1,v_1 \in H_1, \forall u_2,v_2 \in H_2, \quad \langle u_1 \otimes u_2, v_1 \otimes v_2 \rangle_\otimes &:= \langle u_1,v_1 \rangle_{H_1} \langle u_2,v_2 \rangle_{H_2} \\ &= u_1 \otimes u_2(v_1,v_2). \end{aligned}$$

et étendue à tout l'espace $H_1 \bigotimes_a H_2$ par bilinéarité. En particulier, on a que pour tout $\eta \in H_1 \bigotimes_a H_2$, et pour tout $v_1 \in H_1$ et $v_2 \in H_2$,

$$\langle \eta, v_1 \otimes v_2 \rangle_\otimes = \eta(v_1,v_2).$$

Proposition 2.2. *La forme hermitienne $\langle \cdot,\cdot \rangle_\otimes$ est bien définie sur $H_1 \bigotimes_a H_2$ et est un produit scalaire. On notera dans la suite $\| \cdot \|_\otimes$ la norme associée.*

Démonstration. Pour montrer que la forme hermitienne $\langle \cdot,\cdot \rangle_\otimes$ est bien défini sur $H_1 \bigotimes_a H_2$, il faut montrer que pour tout $\eta,\zeta \in H_1 \bigotimes_a H_2$, la valeur de $\langle \eta,\zeta \rangle_\otimes$ ne dépend pas du choix de la représentation de η et ζ comme combinaison linéaire de produits tensoriels purs. Pour cela, il suffit de montrer que si $\eta \in H_1 \bigotimes_a H_2$ est la forme linéaire identiquement nulle, alors, pour tout $\zeta \in H_1 \bigotimes_a H_2$, on a $\langle \eta,\zeta \rangle_\otimes = 0$. Supposons que ζ s'écrive comme

$$\zeta = \sum_{k=1}^K c_k u_k^1 \otimes u_k^2,$$

avec $c_k \in \mathbb{R}$, $u_k^1 \in H_1$ et $u_k^2 \in H_2$ pour tout $1 \le k \le K$. On a alors

$$\begin{aligned} \langle \eta,\zeta \rangle_\otimes &= \left\langle \eta, \sum_{k=1}^K c_k u_k^1 \otimes u_k^2 \right\rangle_\otimes \\ &= \sum_{k=1}^K c_k \left\langle \eta, u_k^1 \otimes u_k^2 \right\rangle_\otimes \\ &= \sum_{k=1}^K c_k \eta(u_k^1,u_k^2) \\ &= 0. \end{aligned}$$

La forme bilinéaire $\langle \cdot,\cdot \rangle_\otimes$ est donc bien définie.

Montrons maintenant que $\langle \cdot,\cdot \rangle_\otimes$ est un produit scalaire. Il est facile de voir que, en introduisant une base orthonormale $(v_1^1, \dots, v_{K_1}^1)$ de $\mathrm{Vect}\{u_1^1, \dots, u_K^1\}$, et respectivement $(v_1^2, \dots, v_{K_2}^1)$ de $\mathrm{Vect}\{u_1^2, \dots, u_K^2\}$, on peut toujours écrire η sous la forme

$$\eta = \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} d_{k_1,k_2} v_{k_1}^1 \otimes v_{k_2}^2$$

pour des coefficients $d_{k_1,k_2} \in \mathbb{R}$. On obtient alors que

$$\langle \eta,\eta \rangle_\otimes = \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} |d_{k_1,k_2}|^2,$$

ce qui implique que $\langle \eta,\eta \rangle_\otimes \geq 0$ et $\langle \eta,\eta \rangle_\otimes = 0$ si et seulement si $d_{k_1,k_2} = 0$ pour tout $1 \leq k_1 \leq K_1$ et $1 \leq k_2 \leq K_2$. D'où le résultat.

Remarque 2.4. *Pour tout $u_1 \in H_1$ et $u_2 \in H_2$, on a alors*

$$\|u_1 \otimes u_2\|_\otimes = \|u_1\|_{H_1} \|u_2\|_{H_2}.$$

On appelle la norme $\| \cdot \|_\otimes$ norme produit croisé (cross product en anglais) ou *norme canonique.*

Définition 2.3. *On définit le produit tensoriel (canonique) des espaces H_1 et H_2 comme le complété de $H_1 \bigotimes_a H_2$ muni, et on note celui-ci $H_1 \bigotimes H_2$.*

Remarque 2.5. *Par définition, $H_1 \bigotimes H_2$ est donc un espace de Hilbert.*

Proposition 2.3. *Soit $\mathcal{I},\mathcal{J} \subset \mathbb{N}$ deux sous-ensembles de $\mathbb{N}$ et soit $(e_i)_{i\in\mathcal{I}}$ une base orthonormale de H_1 et $(f_j)_{j\in\mathcal{J}}$ une base orthonormale de H_2. Alors $(e_i \otimes f_j)_{(i,j)\in\mathcal{I}\times\mathcal{J}}$ forme une base orthonormale de $H_1 \bigotimes H_2$.*

Démonstration. Le fait que l'ensemble $\{e_i \otimes f_j,\, (i,j) \in \mathcal{I} \times \mathcal{J}\}$ forme une famille orthonormale de $H_1 \bigotimes H_2$ est clair par définition.

Soit S l'espace vectoriel engendré par cet ensemble (l'ensemble des combinaisons linéaires finies de ces éléments). Montrons que S est dense dans $H_1 \bigotimes H_2$.

Montrons tout d'abord que $H_1 \bigotimes_a H_2 \subset \overline{S}^{\|\cdot\|_\otimes}$ où $\overline{S}^{\|\cdot\|_\otimes}$ désigne la fermeture de S pour la norme $\| \cdot \|_\otimes$. Soit $u \in H_1$ et $v \in H_2$ et soit $(\alpha_i)_{i\in\mathcal{I}} \in \mathbb{R}^\mathcal{I}$ et $(\beta_j)_{j\in\mathcal{J}} \in \mathbb{R}^\mathcal{J}$ tels que

$$u = \sum_{i\in\mathcal{I}} \alpha_i e_i \quad \text{et } v = \sum_{j\in\mathcal{J}} \beta_j f_j.$$

On a alors en particulier

$$\|u\|_{H_1}^2 = \sum_{i\in\mathcal{I}} |\alpha_i|^2 < +\infty \quad \text{and } \|v\|_{H_2}^2 = \sum_{j\in\mathcal{J}} |\beta_j|^2 < +\infty.$$

En conséquence, on a que

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} |\alpha_i\beta_j|^2 < +\infty,$$

et

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} \alpha_i\beta_j e_i \otimes f_j$$

appartient à $\overline{S}$. Par ailleurs, on voit que pour tout $I,J \in \mathbb{N}$, on a

$$\left\| \phi \otimes \psi - \sum_{i\in\mathcal{I},i\leq I} \sum_{j\in\mathcal{J},j\leq J} \alpha_i\beta_j e_i \otimes f_j \right\|_\otimes^2 = \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} |\alpha_i|^2|\beta_j|^2 - \sum_{i\in\mathcal{I},i\leq I} \sum_{j\in\mathcal{J},j\leq J} |\alpha_i|^2|\beta_j|^2.$$

Cette quantité tend donc bien vers 0 lorsque I et J tendent vers l'infini. L'ensemble $H_1 \bigotimes_a H_2$ est donc bien inclus dans $\overline{S}$. Comme $H_1 \bigotimes H_2$ est le complété de $H_1 \bigotimes_a H_2$, on a donc nécessairement que $H_1 \bigotimes H_2 = \overline{S}$. D'où le résultat.

Nous utiliserons également dans la suite le résultat suivant:

Proposition 2.4. *Le produit tensoriel canonique $H_1 \bigotimes H_2$ est un sous-espace vectoriel de $\mathcal{L}_2(H_1,H_2)$, l'ensemble des formes bilinéaires continues définies sur $H_1 \times H_2$, et on a*

$$\forall \eta \in H_1 \bigotimes H_2, \quad \|\eta\|_{\mathcal{L}_2(H_1,H_2)} := \sup_{u_1 \in H_1, u_2 \in H_2} \frac{\eta(u_1,u_2)}{\|u_1\|_{H_1}\|u_2\|_{H_2}} \leq \|\eta\|_{\otimes}. \tag{2.1}$$

Démonstration. Commençons par prouver (2.1) pour $\eta \in H_1 \bigotimes_a H_2$. En effet, comme pour tout $u_1 \in H_1$ et $u_2 \in H_2$, on a $\|u_1 \otimes u_2\|_{\otimes} = \|u_1\|_{H_1}\|u_2\|_{H_2}$, on a

$$\begin{aligned}
\|\eta\|_{\mathcal{L}_2(H_1,H_2)} &= \sup_{u_1 \in H_1, u_2 \in H_2} \frac{\eta(u_1,u_2)}{\|u_1\|_{H_1}\|u_2\|_{H_2}} \\
&= \sup_{u_1 \in H_1, u_2 \in H_2} \frac{\langle \eta, u_1 \otimes u_2\rangle_{\otimes}}{\|u_1 \otimes u_2\|_{\otimes}} \\
&\leq \sup_{\xi \in H_1 \bigotimes H_2} \frac{\langle \eta, \xi\rangle_{\otimes}}{\|\xi\|_{\otimes}} \\
&\leq \|\eta\|_{\otimes}.
\end{aligned}$$

Comme $\mathcal{L}_2(H_1,H_2)$ est un espace de Banach, il est par conséquent aisé de montrer par densité que nécessairement, $H_1 \bigotimes H_2$ est un sous-espace vectoriel de $\mathcal{L}_2(H_1,H_2)$ et que l'inégalité (2.1) est vérifiée pour tout $\eta \in H_1 \bigotimes H_2$.

Attention cependant, il n'y a pas égalité entre ces normes ni entre ces espaces (sauf dans le cas où H_1 et H_2 sont de dimension finie).

En effet, supposons que $H_1 = H_2 = H$ et que l'espace de Hilbert H soit de dimension infinie. Soit $(e_i)_{i \in \mathbb{N}}$ une base hilbertienne de H.

Soit $a : H \times H \to \mathbb{R}$ la forme bilinéaire définie par

$$\forall u,v \in H, \quad a(u,v) = \sum_{i \in \mathbb{N}} \langle u,e_i\rangle_H \langle v,e_i\rangle_H.$$

On remarque tout d'abord que $a \in \mathcal{L}_2(H,H)$, puisqu'avec l'inégalité de Cauchy-Schwarz, on obtient que

$$|a(u,v)| \leq \left(\sum_{i \in \mathbb{N}} |\langle u,e_i\rangle_H|^2\right)^{1/2} \left(\sum_{i \in \mathbb{N}} |\langle v,e_i\rangle_H|^2\right)^{1/2} = \|u\|_H\|v\|_H.$$

En revanche $a \notin H \bigotimes H$. En effet, on peut remarquer que $a = \sum_{i \in \mathbb{N}} e_i \otimes e_i$, ce qui implique que

$$\|a\|_{\otimes} = +\infty.$$

Pour ceux qui désirent aller plus loin et qui suivent ou ont suivi le cours de théorie spectrale, on peut montrer qu'on peut identifier $\mathcal{L}_2(H,H)$ avec l'ensemble des opérateurs linéaires bornés (continus) de H dans H. Par ailleurs, l'espace $H \bigotimes H$ peut être identifié avec l'ensemble des opérateurs *de Hilbert-Schmidt*de H dans H.

Remarque 2.6. *Soit $d \in \mathbb{N}^*$ et $H_1,\ldots,H_d$ d espaces de Hilbert. On définit le produit tensoriel $H_1 \bigotimes \ldots \bigotimes H_d$ de la même manière que ce qui a été décrit précédemment pour le cas de deux espaces de Hilbert. De plus, pour toute permutation σ de l'ensemble $\{1,\ldots,d\}$, on a que $H_1 \bigotimes \ldots \bigotimes H_d = H_{\sigma(1)} \bigotimes \ldots \bigotimes H_{\sigma(d)}$ (à isométrie près).*

Soit $\mathcal{I}_1,\ldots,\mathcal{I}_d \subset \mathbb{N}$ des sous-ensembles de $\mathbb{N}$ et pour tout $1 \leq j \leq d$, soit $\{e^j_{i_j}\}_{i_j \in \mathcal{I}_j}$ une base hilbertienne de H_j. Alors l'ensemble $\{e^1_{i_1} \otimes e^2_{i_2} \times \ldots \times e^d_{i_d}\}_{i_1 \in \mathcal{I}_1, i_2 \in \mathcal{I}_2, \ldots, i_d \in \mathcal{I}_d}$ forme une base hilbertienne de $H_1 \bigotimes H_2 \bigotimes \ldots \bigotimes H_d$.

2.2.2 Espaces de dimension finie

Le but de cette section est de faire le lien entre la définition abstraite des produits tensoriels d'espaces introduite à la section précédente et aux notions de tenseurs classiques dans le cas où on considère des espaces de dimension finie.

Soit $N_1,N_2 \in \mathbb{N}^*$ et considérons $H_1 = \mathbb{R}^{N_1}$ et $H_2 = \mathbb{R}^{N_2}$ muni chacun de leur norme euclidienne canonique. Nous allons montrer la proposition suivante.

Proposition 2.5. *L'espace $\mathbb{R}^{N_1} \otimes \mathbb{R}^{N_2}$ est isométriquement isomorphe à $\mathbb{R}^{N_1\times N_2}$ muni du produit scalaire de Frobenius, défini par*

$$\forall A,B \in \mathbb{R}^{N_1\times N_2}, \quad \langle A,B\rangle_F := \mathrm{Tr}\left(A^T B\right),$$

où $A^T \in \mathbb{R}^{N_2\times N_1}$ est la matrice transposée de la matrice A.

Remarque 2.7. *Le produit scalaire de Frobenius entre deux matrices $A,B \in \mathbb{R}^{N_1\times N_2}$ est parfois noté $\langle A,B\rangle_F = A : B$. On l'appelle aussi le produit doublement contracté de A par B.*

Exercice 2.2. *Soit $A := (A_{ij})_{1\le i\le N_1, 1\le j\le N_2}, B := (B_{ij})_{1\le i\le N_1,1\le j\le N_2} \in \mathbb{R}^{N_1\times N_2}$. Montrer que*

$$\langle A,B\rangle_F = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} A_{ij}B_{ij}$$

et prouver que $\langle\cdot,\cdot\rangle_F$ définit bien un produit scalaire sur $\mathbb{R}^{N_1\times\mathbb{N}_2}$.

Soit $(e_1,\dots,e_{N_1})$ la base canonique de H_1 et $(f_1,\dots,f_{N_2})$ la base canonique de H_2. D'après la Proposition 2.3, on sait alors que $\{e_i \otimes f_j\}_{1\le i\le N_1,1\le j\le N_2}$ forme une base orthonormale de $\mathbb{R}^{N_1} \bigotimes \mathbb{R}^{N_2}$. En particulier, $\mathbb{R}^{N_1} \bigotimes \mathbb{R}^{N_2}$ muni de la norme produit est un espace vectoriel de dimension $N_1 \times N_2$.

Définissons l'application

$$J : \begin{cases} \mathbb{R}^{N_1} \bigotimes \mathbb{R}^{N_2} & \to & \mathbb{R}^{N_1\times N_2} \\ \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{ij} e_i \otimes f_j & \mapsto & (u_{ij})_{1\le i\le N_1,1\le j\le N_2}. \end{cases}$$

Exercice 2.3. *(1) Montrer que pour tout $u = (u_i)_{1\le i\le N_1} \in \mathbb{R}^{N_1}$ et tout $v = (v_j)_{1\le j\le N_2}$,*

$$\forall 1 \le i \le N_1, 1 \le j \le N_2, \quad J(u\otimes v)_{ij} = u_i v_j.$$

En déduire que

$$J(u\otimes v) = u^T v.$$

(2) Montrer que J est un isomorphisme isométrique.

On peut donc identifier $\mathbb{R}^{N_1} \bigotimes \mathbb{R}^{N_2}$ avec $\mathbb{R}^{N_1\times N_2}$ muni du produit scalaire de Frobenius.

Ce résultat se généralise très facilement au cas de produits tensoriels de d espaces euclidiens de dimension finie.

Proposition 2.6. *Soit $d \in \mathbb{N}^*$, $N_1,\dots,N_d \in \mathbb{N}^*$. On suppose qu'on munit $\mathbb{R}^{N_1},\dots,\mathbb{R}^{N_d}$ de leur produit scalaire euclidien canonique. Le produit tensoriel $\mathbb{R}^{N_1} \bigotimes \dots \bigotimes \mathbb{R}^{N_d}$ est isométriquement isomorphe à $\mathbb{R}^{N_1\times\dots\times N_d}$ muni du produit scalaire de Frobenius suivant: pour tout $A = (A_{i_1,\dots,i_d})_{1\le i_1\le N_1,\dots,1\le i_d\le N_d}, B = (B_{i_1,\dots,i_d})_{1\le i_1\le N_1,\dots,1\le i_d\le N_d} \in \mathbb{R}^{N_1\times\dots\times N_d}$,*

$$\langle A,B\rangle_F := \sum_{i_1=1}^{N_1}\dots\sum_{i_d=1}^{N_d} A_{i_1,\dots,i_d}B_{i_1,\dots,i_d}.$$

2.2.3 Espaces de fonctions

Le but de cette section est de faire le lien entre les produits tensoriels d'espaces de fonctions et des espaces fonctionnels que vous connaissez mieux, à savoir les espaces de Lebesgue et les espaces de Sobolev.

Dans toute la suite de cette section, nous noterons $\Omega_1 \subset \mathbb{R}^{p_1}$, $\Omega_2 \subset \mathbb{R}^{p_2}$, ..., $\Omega_d \subset \mathbb{R}^{p_d}$ des ensembles ouverts avec $p_1, \cdots, p_d \in \mathbb{N}^*$ et soit $\Omega := \Omega_1 \times \cdots \times \Omega_d$.

Fonctions $\mathcal{C}^\infty$ à support compact

Nous aurons besoin du résultat préliminaire suivant, qui énonce un résultat de densité dans l'espace des fonctions $\mathcal{C}^\infty$ à support compact.

Soit $\mathcal{D}(\Omega_1) \bigotimes_a \cdots \bigotimes_a \mathcal{D}(\Omega_d)$ le sous-ensemble de $\mathcal{D}(\Omega)$ des combinaisons linéaires finies de fonctions de la forme $\phi_1 \otimes \cdots \otimes \phi_d$ où $\phi_1 \in \mathcal{D}(\Omega_1)$, ..., $\phi_d \in \mathcal{D}(\Omega_d)$ et où

$$\phi_1 \otimes \cdots \otimes \phi_d : \begin{cases} \Omega = \Omega_1 \times \cdots \times \Omega_d \to & \mathbb{R} \\ \quad (x_1, \cdots, x_d) & \mapsto \phi_1(x_1) \cdots \phi_d(x_d). \end{cases}$$

On a alors le résultat de densité suivant:

Proposition 2.7. *L'ensemble $\mathcal{D}(\Omega_1) \bigotimes_a \cdots \bigotimes_a \mathcal{D}(\Omega_d)$ est dense dans $\mathcal{D}(\Omega)$ au sens suivant: pour tout $q \in \mathbb{N}$ et pour tout $\phi \in \mathcal{D}(\Omega)$, il existe une suite de fonctions $(\phi_n)_{n\in\mathbb{N}} \subset \mathcal{D}(\Omega_1) \bigotimes_a \cdots \bigotimes_a \mathcal{D}(\Omega_d)$ telle que pour tout $\alpha \in \mathbb{N}^{p_1+p_2+\cdots+p_d}$ avec $|\alpha| \leq q$, alors*

$$\|\partial^\alpha \phi - \partial^\alpha \phi_n\|_{L^\infty(\Omega)} \underset{n\to+\infty}{\longrightarrow} 0.$$

La preuve de ce résultat repose sur le résultat suivant qui est une extension du théorème de Stone-Weierstrass, dont nous ne rappellerons pas la preuve ici.

Théorème 2.1 (Stone-Weierstrass). *Soit $p \in \mathbb{N}^*$ et $K \subset \mathbb{R}^p$ un sous-ensemble compact. Soit $f \in \mathcal{C}^\infty(K)$ et $q \in \mathbb{N}$. Alors il existe une siote de fonctions polynômiales $(P_n)_{n\in\mathbb{N}}$ telle que pour tout $\alpha \in \mathbb{N}^p$ tel que $|\alpha| \leq q$, on ait*

$$\|\partial^\alpha f - \partial^\alpha P_n\|_{L^\infty(K)} \underset{n\to+\infty}{\longrightarrow} 0.$$

Démonstration. Pour simplifier, nous ne prouvons ce résultat ici que dans le cas où $d = 2$, mais la preuve s'étend aisément au cas où d est un entier quelconque.

Soit $\phi \in \mathcal{D}(\Omega_1 \times \Omega_2)$. Il existe alors $K \subset \Omega_1$ et $L \subset \Omega_2$ deux sous-ensembles compacts tels que $\mathrm{Supp}\, \phi \subset K \times L$. On définit également deux fonctions $\eta_1 \in \mathcal{D}(\Omega_1)$ et $\eta_2 \in \mathcal{D}(\Omega_2)$ telles que $\eta_1 = 1$ sur K et $\eta_2 = 1$ sur L. On note également $\widetilde{K}$ le support de η et $\widetilde{L}$ le support de η_2, qui vérifient $K \subset \widetilde{K}$ et $L \subset \widetilde{L}$. Il est aisé de vérifier qu'on a alors $\phi(x_1,x_2) = \phi(x_1,x_2)\eta_1(x_1)\eta_2(x_2)$ pour tout $(x_1,x_2) \in \Omega_1 \times \Omega_2$. Autrement dit $\phi = \phi\eta_1 \otimes \eta_2$.

D'après le théorème de Stone-Weierstrass, comme le domaine $K \times L$ est un sous-domaine compact de $\mathbb{R}^{p_1} \times \mathbb{R}^{p_2}$, il exsiste une suite de fonctions polynômiales $(P_n)_{n\in\mathbb{N}}$ telles que pour tout $\alpha \in \mathbb{N}^{p_1+p_2}$ avec $|\alpha| \leq q$,

$$\|\partial_\alpha \phi - \partial_\alpha P_n\|_{L^\infty(\widetilde{K}\times\widetilde{L})} \underset{n\to+\infty}{\longrightarrow} 0.$$

En particulier, ceci implique, puisque $\phi = 0$ sur $\widetilde{K} \times \widetilde{L} \setminus K \times L$, que

$$\|\partial_\alpha P_n\|_{L^\infty(\widetilde{K}\times\widetilde{L}\setminus K\times L)} \underset{n\to+\infty}{\longrightarrow} 0, \tag{2.2}$$

pour tout $\alpha \in \mathbb{N}^{p_1+p_2}$ avec $|\alpha| \leq q$.

On définit alors $\phi_n(x_1,x_2) = \eta_1(x_1)\eta_2(x_2)P_n(x_1,x_2)$. D'une part, comme P_n est une fonction polynômiale, on a nécessairement $\phi_n \in \mathcal{D}(\Omega_1) \bigotimes_a \mathcal{D}(\Omega_2)$ pour tout $n \in \mathbb{N}$. De plus, on a pour tout $\alpha \in \mathbb{N}^{p_1+p_2}$ avec $|\alpha| \leq q$,

$$\begin{aligned}\|\partial^\alpha\phi - \partial^\alpha\phi_n\|_{L^\infty(\Omega)} &= \|\partial^\alpha\phi - \partial^\alpha\phi_n\|_{L^\infty(\widetilde{K}\times\widetilde{L})}\\ &= \|\partial^\alpha(\phi\eta_1\otimes\eta_2) - \partial^\alpha(P_n\eta_1\otimes\eta_2)\|_{L^\infty(\widetilde{K}\times\widetilde{L})}\\ &\leq \|\partial^\alpha(\phi\eta_1\otimes\eta_2) - \partial^\alpha(P_n\eta_1\otimes\eta_2)\|_{L^\infty(\widetilde{K}\times\widetilde{L}\setminus K\times L)} + \|\partial^\alpha(\phi\eta_1\otimes\eta_2) - \partial^\alpha(P_n\eta_1\otimes\eta_2)\|_{L^\infty(K\times L)}.\end{aligned}$$

D'une part, en utilisant (2.2), on obtient

$$\|\partial^\alpha(\phi\eta_1\otimes\eta_2) - \partial^\alpha(P_n\eta_1\otimes\eta_2)\|_{L^\infty(\widetilde{K}\times\widetilde{L}\setminus K\times L)} = \|\partial^\alpha(P_n\eta_1\otimes\eta_2)\|_{L^\infty(\widetilde{K}\times\widetilde{L}\setminus K\times L)} \underset{n\to+\infty}{\longrightarrow} 0.$$

D'autre part, on a

$$\|\partial^\alpha(\phi\eta_1\otimes\eta_2) - \partial^\alpha(P_n\eta_1\otimes\eta_2)\|_{L^\infty(K\times L)} = \|\partial^\alpha\phi - \partial^\alpha P_n\|_{L^\infty(K\times L)} \underset{n\to+\infty}{\longrightarrow} 0,$$

ce qui nous donne le résultat désiré.

Nous allons voir dans la suite de cette section comment ce résultat permet d'identifier naturellement les produits tensoriels d'espaces fonctionnels classiques comme les espaces de Lebesgue et les espaces de Sobolev.

Espaces de Lebesgue L^2

On considère $H_1 = L^2(\Omega_1)$ et $H_2 = L^2(\Omega_2)$ munis de leurs produits scalaires. On rappelle que

$$\forall u_1,v_1 \in L^2(\Omega_1),\quad \langle v_1,u_1\rangle_{L^2(\Omega_1)} := \int_{\Omega_1} v_1u_1 \quad \text{et} \quad \forall u_2,v_2 \in L^2(\Omega_2),\quad \langle v_2,u_2\rangle_{L^2(\Omega_2)} := \int_{\Omega_2} v_2u_2.$$

On considère une application linéaire $J : H_1 \bigotimes_a H_2 \to L^2(\Omega_1\times\Omega_2)$ définie comme suit (puis étendue par linéarité): pour tout $u_1 \in L^2(\Omega_1)$ et pour tout $u_2 \in L^2(\Omega_2)$, $J(u_1\otimes u_2)$ est la fonction appartenant à $L^2(\Omega_1\times\Omega_2)$ telle que

$$J(u_1\otimes u_2)(x_1,x_2) = u_1(x_1)u_2(x_2) \quad \text{pour presque tout } (x_1,x_2) \in \Omega_1\times\Omega_2.$$

Exercice 2.4. *(1) Montrer que J est bien définie.*
(2) Montrer que J est une isométrie de $(H_1 \bigotimes_a H_2, \langle\cdot,\cdot\rangle_\otimes)$ dans $\left(L^2(\Omega_1\times\Omega_2), \langle\cdot,\cdot\rangle_{L^2(\Omega_1\times\Omega_2)}\right)$.
(3) En déduire que l'espace $H_1 \bigotimes H_2$ est isométriquement isomorphe à la fermeture dans $L^2(\Omega_1\times\Omega_2)$ de l'ensemble des fonctions qui s'écrivent sous la forme $\Omega_1\times\Omega_2 \ni (x_1,x_2) \mapsto \sum_{k=1}^K c_k u_k^1(x_1)u_k^2(x_2)$ pour $K\in\mathbb{N}^*$ et $c_k\in\mathbb{R}$, $u_k^1\in L^2(\Omega_1)$ et $u_k^2 \in L^2(\Omega_2)$ pour tout $1\leq k\leq K$.
(4) Montrer que $H_1 \bigotimes H_2$ est isométriquement isomorphe à $L^2(\Omega_1\times\Omega_2)$ en utilisant la Proposition 2.7.

Nous avons donc prouvé le résultat suivant:

Théorème 2.2. *Le produit tensoriel $L^2(\Omega_1)\bigotimes L^2(\Omega_2)$ est isométriquement isomorphe à $L^2(\Omega_1\times\Omega_2)$.*

Nous identifierons les deux espaces par la suite et nous utiliserons la notation $u_1\otimes u_2$ pour désigner $J(u_1\otimes u_2) \in L^2(\Omega_1\times\Omega_2)$.

Le Théorème 2.5 peut se généraliser dans le cas du produit tensoriels de d espaces de Lebesgue avec $d\in\mathbb{N}^*$.

Théorème 2.3. *Soit $d\in\mathbb{N}^*$ et pour tout $1\leq j\leq d$, soit Ω_j un sous-ensemble borélien de $\mathbb{R}^{p_j}$ avec $p_j\in\mathbb{N}^*$. Alors $L^2(\Omega_1)\bigotimes\cdots\bigotimes L^2(\Omega_d)$ est isométriquement isomorphe à $L^2(\Omega_1\times\cdots\times\Omega_d)$.*

Espaces de Bochner

Soit Ω un sous-ensemble borélien de $\mathbb{R}^p$ avec $p \in \mathbb{N}^*$ et soit μ une mesure de probabilité sur Ω. Soit V un espace de Hilbert.

On définit

$$L^2_\mu(\Omega; V) := \left\{ u : \Omega \to V, \quad \int_\Omega \|u(t)\|_V^2 \, d\mu(t) \right\}.$$

L'espace $L^2_\mu(\Omega; V)$ est appelé un **espace de Bochner**. Il s'agit d'un espace de Hilbert, muni du produit scalaire suivant:

$$\forall u,v \in L^2_\mu(\Omega; V), \quad \langle v,u \rangle_{L^2_\mu(\Omega;V)} := \int_\Omega \langle v(t),u(t) \rangle_V \, d\mu(t).$$

Soit $H_1 = L^2_\mu(\Omega)$, muni du produit scalaire

$$\forall u,v \in L^2_\mu(\Omega), \quad \langle v,u \rangle_{L^2_\mu(\Omega)} := \int_\Omega vu \, d\mu,$$

et $H_2 = V$.

On considère une application linéaire $J : H_1 \bigotimes_a H_2 \to L^2_\mu(\Omega; V)$ définie comme suit (puis étendue par linéarité): pour tout $u \in L^2_\mu(\Omega)$ et pour tout $v \in V$, $J(u \otimes v)$ est la fonction appartenant à $L^2_\mu(\Omega; V)$ telle que

$$J(u \otimes v)(t) = u(t)v \quad \text{pour presque tout } t \in \Omega.$$

Exercice 2.5. *(1) Montrer que J est bien définie.*

(2) Montrer que J est une isométrie de $(H_1 \bigotimes_a H_2, \langle\cdot,\cdot\rangle_\otimes)$ dans $\left(L^2_\mu(\Omega; V), \langle\cdot,\cdot\rangle_{L^2_\mu(\Omega;V)}\right)$.

(3) En déduire que l'espace $H_1 \bigotimes H_2$ est isométriquement isomorphe à la fermeture dans $L^2_\mu(\Omega; V)$ de l'ensemble des fonctions qui s'écrivent sous la forme $\Omega \ni t \mapsto \sum_{k=1}^K c_k u_k(t) v_k$ pour $K \in \mathbb{N}^*$ et $c_k \in \mathbb{R}$, $u_k \in L^2_\mu(\Omega)$ et $v_k \in V$ pour tout $1 \le k \le K$.

(4) En admettant que l'ensemble des combinaisons linéaires finies de fonctions de la forme $\Omega \ni t \mapsto \phi(t)v$ avec $\phi \in \mathcal{D}(\Omega)$ et $v \in V$ sont denses dans $L^2_\mu(\Omega; V)$, en déduire que $H_1 \bigotimes H_2$ est isométriquement isomorphe à $L^2_\mu(\Omega; V)$.

Nous venons donc de montrer le résultat suivant:

Théorème 2.4. *Le produit tensoriel $L^2_\mu(\Omega) \bigotimes V$ est isométriquement isomorphe à $L^2_\mu(\Omega; V)$.*

Espaces de Sobolev

L'objet de cette section est d'attirer votre attention sur le cas particulier des espaces de Sobolev, qui est différent de celui des espaces de Lebesgue vu précédemment.

On considère $H_1 = H^1_0(\Omega_1)$ et $H_2 = H^1_0(\Omega_2)$ munis de leurs produits scalaires. On rappelle que

$$\forall u_1,v_1 \in H^1_0(\Omega_1), \quad \langle v_1,u_1 \rangle_{H^1_0(\Omega_1)} := \int_{\Omega_1} v_1 u_1 + \int_{\Omega_1} \nabla v_1 \cdot \nabla u_1.$$

Le produit scalaire de $H^1_0(\Omega_2)$ est défini de manière similaire.

On considère une application linéaire $J : H_1 \bigotimes_a H_2 \to H^1_0(\Omega_1 \times \Omega_2)$ définie comme suit (puis étendue par linéarité): pour tout $u_1 \in H^1_0(\Omega_1)$ et pour tout $u_2 \in H^1_0(\Omega_2)$, $J(u_1 \otimes u_2)$ est la fonction appartenant à $H^1_0(\Omega_1 \times \Omega_2)$ telle que

$$J(u_1 \otimes u_2)(x_1,x_2) = u_1(x_1)u_2(x_2) \quad \text{pour presque tout } (x_1,x_2) \in \Omega_1 \times \Omega_2.$$

Exercice 2.6. *(1) Montrer que J est bien définie.*

*(2) Montrer que J **n'est pas** une isométrie de $(H_1 \bigotimes_a H_2, \langle \cdot,\cdot \rangle_\otimes)$ dans $\left(H^1_0(\Omega_1 \times \Omega_2), \langle \cdot,\cdot \rangle_{H^1_0(\Omega_1\times\Omega_2)}\right)$.*

L'espace $H^1_0(\Omega_1) \bigotimes H^1_0(\Omega_2)$ **n'est donc pas** isométriquement isomorphe à $H^1_0(\Omega_1 \times \Omega_2)$! On peut cependant identifier cet espace avec un autre espace de Sobolev, appelé espace de Sobolev avec dérivées mixtes.

Plus précisément, on définit

$$H^1_{\mathrm{mix}}(\Omega_1 \times \Omega_2) := \left\{ u \in H^1(\Omega_1 \times \Omega_2),\, \partial_{x_1}\partial_{x_2} u \in L^2(\Omega_1 \times \Omega_2) \right\}.$$

On munit alors l'espace $H^1_{\mathrm{mix}}(\Omega_1 \times \Omega_2)$ du produit scalaire suivant: pour tout $u,v \in H^1_{\mathrm{mix}}(\Omega_1 \times \Omega_2)$,

$$\langle v,u \rangle_{H^1_{\mathrm{mix}}(\Omega_1\times\Omega_2)} := \int_{\Omega_1\times\Omega_2} vu + \nabla v \cdot \nabla u + \partial_{x_1}\partial_{x_2} v \partial_{x_1}\partial_{x_2} u.$$

Exercice 2.7. *Montrer que $H^1_{\mathrm{mix}}(\Omega_1 \times \Omega_2)$ est un espace de Hilbert muni de ce produit scalaire.*

On définit l'espace $H^1_{0,\mathrm{mix}}(\Omega_1 \times \Omega_2)$ comme la fermeture de l'ensemble des fonctions $\mathcal{C}^\infty$ à support compact $\mathcal{D}(\Omega_1 \times \Omega_2)$ dans l'espace $H^1_{\mathrm{mix}}(\Omega_1 \times \Omega_2)$.

Exercice 2.8. *(1) Montrer que $J\,(H_1 \bigotimes_a H_2)$ est inclus dans $H^1_{0,\mathrm{mix}}(\Omega_1 \times \Omega_2)$.*
(2) Montrer que J est une isométrie de $(H_1 \bigotimes_a H_2, \langle \cdot,\cdot \rangle_\otimes)$ dans $\left(H^1_{0,\mathrm{mix}}(\Omega_1 \times \Omega_2), \langle \cdot,\cdot \rangle_{H^1_{\mathrm{mix}}(\Omega_1\times\Omega_2)}\right)$.
(3) Montrer que $J\,(H_1 \bigotimes_a H_2)$ est dense dans $H^1_{0,\mathrm{mix}}(\Omega_1 \times \Omega_2)$. En déduire que $H_1 \bigotimes H_2$ est isométriquement isomorphe à $H^1_{0,\mathrm{mix}}(\Omega_1 \times \Omega_2)$.

Nous avons donc montré le résultat suivant:

Théorème 2.5. *Le produit tensoriel $H^1_0(\Omega_1) \bigotimes H^1_0(\Omega_2)$ est isométriquement isomorphe à $H^1_{0,\mathrm{mix}}(\Omega_1 \times \Omega_2)$.*

Attention! Le produit tensoriel $H^1_0(\Omega_1) \bigotimes H^1_0(\Omega_2)$ n'est donc pas isométriquement isomorphe à $H^1_0(\Omega_1 \times \Omega_2)$! Méfiez-vous!

2.3 Complément 1: Espaces injectifs et projectifs

Soit H_1 et H_2 deux espaces de Hilbert. On peut munir l'espace vectoriel $H_1 \bigotimes_a H_2$ de normes différentes de la norme canonique $\|\cdot\|$.

Définition 2.4. *Soit $z \in H_1 \bigotimes_a H_2$. On définit*

- ***Norme projective:***

$$\|z\|_\wedge := \inf \left\{ \sum_{k=1}^K \|u^1_k \otimes u^2_k\|_\otimes, \quad u^1_k \in H_1,\, u^2_k \in H_2,\, 1 \leq k \leq K,\, z = \sum_{k=1}^K u^1_k \otimes u^2_k \right\}$$

- ***Norme injective:***

$$\|z\|_\vee := \sup_{\substack{u_1 \in H_1,\, u_2 \in H_2,\\ \|u_1 \otimes u_2\|_\otimes = 1}} \langle z, u_1 \otimes u_2 \rangle_\otimes = \sup_{u_1 \in H_1,\, u_2 \in H_2} \frac{z(u_1,u_2)}{\|u_1\|_{H_1}\|u_2\|_{H_2}}.$$

On peut aisément voir que $\|\cdot\|_\wedge$ et $\|\cdot\|_\vee$ définissent des normes sur $H_1 \bigotimes_a H_2$ (non associées à des produits scalaires) et que l'on a, pour tout $z \in H_1 \bigotimes_a H_2$,

$$\|z\|_\vee \leq \|z\|_\otimes \leq \|z\|_\wedge.$$

Définition 2.5. • *Le complété de $H_1 \bigotimes_a H_2$ muni de la norme $\|\cdot\|_\wedge$ est un espace de Banach noté $H_1 \widehat{\bigotimes} H_2$ et appelé **produit tensoriel projectif** de H_1 et H_2.*

- *Le complété de $H_1 \bigotimes_a H_2$ muni de la norme $\| \cdot \|_\vee$ est un espace de Banach noté $H_1 \widecheck{\bigotimes} H_2$ et appelé* ***produit tensoriel injectif*** *de H_1 et H_2. Il est de plus isométriquement isomorphe à $\mathcal{L}_2(H_1,H_2)$ l'ensemble des formes bilinéaires continues sur $H_1 \times H_2$.*

Proposition 2.8. *On a les inclusions suivantes (à isomorphisme isométrique près)*

$$H_1 \widehat{\bigotimes} H_2 \subset H_1 \bigotimes H_2 \subset H_1 \widecheck{\bigotimes} H_2.$$

Remarque 2.8. *Attention! Hormis dans le cas où les espaces H_1 et H_2 sont de dimension finie, il n'y a en général pas d'égalité entre ces différents espaces.*

2.4 Complément 2: Produits tensoriels d'opérateurs

2.4.1 Produits tensoriels d'opérateurs bornés

Soit A un opérateur borné sur H_1 et B un opérateur borné sur H_2. On peut alors définir l'opérateur $A \otimes B$ sur $H_1 \bigotimes_a H_2$ tel que

$$\forall u \in H_1, \forall v \in H_2, \quad A \otimes B(u \otimes v) = (Au) \otimes (Bv).$$

Exercice 2.9. *Montrer que l'opérateur $A \otimes B$ est bien défini et borné sur $H_1 \bigotimes_a H_2$. Montrer qu'il peut être étendu par continuité de manière unique à un opérateur borné sur $H_1 \bigotimes H_2$. On note cette extension également $A \otimes B$.*

Proposition 2.9. *Soit H_1 et H_2 deux espaces de Hilbert séparables. Soient A et B deux opérateurs bornés et compacts sur H_1 et H_2 respectivement. Alors $A \otimes B$ est un opérateur compact sur $H_1 \bigotimes H_2$.*

Exercice 2.10. *Prouver la Proposition 2.9.*

2.4.2 Produits tensoriels de formes sesquilinéaires continues

Soit $a : H_1 \times H_1 \to \mathbb{R}$ et $b : H_2 \times H_2 \to \mathbb{R}$ deux formes sesquilinéaires continues. On définit alors la forme sesquilinéaire $a \otimes b : (H_1 \bigotimes_a H_2) \times (H_1 \bigotimes_a H_2) \to \mathbb{R}$ telle que

$$\forall u_1,v_1 \in H_1, \forall u_2,v_2 \in H_2, \quad a \otimes b(v_1 \otimes v_2, u_1 \otimes u_2) = a(v_1,u_1)b(v_2,u_2)$$

et étendue à tout $(H_1 \bigotimes_a H_2) \times (H_1 \bigotimes_a H_2)$ par sesquilinéarité.

Proposition 2.10. *La forme sesquilinéaire $a \otimes b$ est continue sur $(H_1 \bigotimes_a H_2) \times (H_1 \bigotimes_a H_2)$ et peut être étendue par continuité de manière unique à une forme sesquilinéaire continue sur $(H_1 \bigotimes H_2) \times (H_1 \bigotimes H_2)$. On note également cette forme $a \otimes b$.*

2.4.3 Produits tensoriels d'opérateurs fermables

La suite et fin de cette section n'est pas à lire en première lecture. Il s'agit de compléments pour ceux d'entre vous qui suivent le cours de Théorie Spectrale. On ne suppose plus ici que A et B soient des opérateurs bornés, mais uniquement à domaines denses.

Soit A un opérateur sur H_1 de domaine dense $D(A)$ et B un opérateur sur H_2 de domaine dense $D(B)$. Soit $D := D(A) \bigotimes_a D(B)$. Il est alors clair que D est dense dans $H_1 \bigotimes H_2$ par définition. On peut alors définir l'opérateur $A \otimes B$ sur $H_1 \bigotimes H_2$, de domaine D tel que

$$\forall u \in D(A), \forall v \in D(B), \quad A \otimes B(u \otimes v) = (Au) \otimes (Bv).$$

Exercice 2.11. *Montrer que l'opérateur $A \otimes B$ est bien défini.*

On rappelle ci-dessous les notions d'opérateurs fermés et d'opérateurs fermables.

Définition 2.6. • *Un opérateur non borné T sur un espace vectoriel normé X à valeurs dans un espace vectoriel normé Y, de domaine $D(T)$ dense dans X est dit* ***fermé*** *si et seulement si son graphe, c'est-à-dire l'ensemble $\{(x,y) \in D(T) \times Y,\, y = Tx\}$, est un sous-ensemble fermé de $X \times Y$.*

- *Un opérateur non borné T sur un espace vectoriel normé X à valeurs dans un espace vectoriel normé Y, de domaine $D(T)$ dense dans X est dit* ***fermable*** *si et seulement si l'adhérence de son graphe dans $X \times Y$ est le graphe d'un opérateur fermé noté $\overline{T}$. L'opérateur $\overline{T}$ est alors unique et appelé la* ***fermeture*** *de l'opérateur T.*

On rappelle ici quelques caractérisations séquentielles utiles des opérateurs fermés et fermables.

Proposition 2.11. • *L'opérateur T est fermé si et seulement si pour toute suite $(x_n)_{n\in\mathbb{N}}$ dans $D(T)$ admettant dans X une limite x et telle que la suite $(T(x_n))_{n\in\mathbb{N}}$ converge dans Y vers une limite y, alors on a $x \in D(T)$ et $y = T(x)$.*

- *L'opérateur T est fermable si et seulement si pour toute suite $(x_n)_{n\in\mathbb{N}}$ dans $D(T)$ convergeant vers 0 et que la suite $(T(x_n))_{n\in\mathbb{N}}$ converge dans Y vers une limite y, alors on a $y = 0$.*

On a alors le résultat suivant:

Proposition 2.12. *Si A et B sont fermables, l'opérateur $A \otimes B$ l'est aussi.*

Définition 2.7. *Si A et B sont des opérateurs fermables, on appelle produit tensoriel de A et B la fermeture de $A \otimes B$. On le note également $A \otimes B$.*

3

Décomposition orthogonale propre

TO COMPLETE

4

Algorithmes gloutons

Nous avons vu dans un chapitre précédent comment la méthode POD pouvait être utilisée pour compresser un tenseur ou une fonction en grande dimension. On a en particulier vu qu'elle permettait d'obtenir une approximation optimale d'un tenseur $u \in H_1 \bigotimes H_2$ appartenant au produit tensoriel canonique de deux espaces de Hilbert par des tenseurs de rang plus petit ou égal à un certain entier r choisi, au moyen des décompositions POD tronquées de ce tenseur.

Cependant, la POD se heurte aux limitations intrinsèques suivantes:

(1) La décomposition POD nécessite de connaître le tenseur u pour être calculée: elle peut être utile pour compresser un tenseur connu mais pas pour résoudre une EDP dont on ne connaît pas la solution;

(2) Il n'existe pas d'équivalent de la décomposition POD dans le cas où u est un élément du produit tensoriel de plus de trois espaces de Hilbert $H_1 \bigotimes H_2 \bigotimes \cdots \bigotimes H_d$; cela peut être un problème lorsque l'on cherche à approcher des fonctions de plus de trois variables;

(3) enfin, la solution u d'EDPs classiques peut appartenir à des espaces de fonctions qu'on ne peut pas identifier comme des produits tensoriels canoniques d'espaces de fonctions dépendant uniquement d'une variable. Plus précisément, considérons par exemple $u \in H_0^1(\Omega_1 \times \cdots \times \Omega_d)$ l'unique solution du problème de Laplace:

$$\begin{cases} -\Delta_{x_1,\cdots,x_d} u(x_1,\cdots,x_d) = f(x_1,\cdots,x_d), & \text{pour } (x_1,\cdots,x_d) \in \Omega := \Omega_1 \times \cdots \times \Omega_d, \\ u(x_1,\cdots,x_d) = 0, & \text{pour } (x_1,\cdots,x_d) \in \partial\Omega. \end{cases}$$

Nous avons vu au chapitre 2 que $H_0^1(\Omega_1 \times \cdots \times \Omega_d)$ **n'est pas** égal au produit tensoriel canonique $H_0^1(\Omega_1) \bigotimes \cdots \bigotimes H_0^1(\Omega_d)$.

Les algorithmes gloutons sont une première famille de méthodes numériques qui permettent de contourner ces limitations de la POD pour calculer des approximations de rang faible de fonctions dépendant d'un grand nombre de variables. Le prix à payer est que ces approximations ne jouiront pas des mêmes propriétés d'optimalité que celles des décompositions POD tronquées que nous avons vues au chapitre précédent. Nous allons en particulier illustrer ces algorithmes gloutons sur l'exemple de la résolution du problème de Laplace mentionné ci-dessus.

4.1 Algorithme glouton

Dans le cadre de ce cours, nous allons illustrer les propriétés des algorithmes gloutons dans le cas particulier de la résolution d'un problème variationnel de type Lax-Milgram. Il est à noter que ces algorithmes sont utilisés pour résoudre d'autres types de problèmes (problèmes non-linéaires, problèmes aux valeurs propres, problèmes d'évolution...), mais nous n'aurons pas le temps de les aborder dans le cadre de ce cours.

Dans toute la suite du cours, nous considèrerons $\Omega_1 \subset \mathbb{R}^{p_1}$, $\Omega_2 \subset \mathbb{R}^{p_2}$, ..., $\Omega_d \subset \mathbb{R}^{p_d}$ des ensembles ouverts, avec $p_1,p_2,\ldots,p_d \in \mathbb{N}^*$. On notera également $\Omega := \Omega_1 \times \cdots \times \Omega_d$.

4.1.1 Problème de référence: Lax-Milgram

Soit V un espace de Hilbert, $a : V \times V \to \mathbb{R}$ une forme sesquilinéaire continue symmétrique et coercive et $l : V \to \mathbb{R}$ une forme linéaire continue. Alors, d'après le théorème de Lax-Milgram, il existe une unique solution $u \in V$ au problème variationnel:

$$\forall v \in V, \quad a(u,v) = l(v).$$

De plus, u est de manière équivalente l'unique solution du problème de minimisation

$$u = \operatorname*{argmin}_{v \in V} \mathcal{E}(v), \tag{4.1}$$

où

$$\forall v \in V, \quad \mathcal{E}(v) := \frac{1}{2} a(v,v) - l(v).$$

Exemple: Le problème de Laplace

Un premier exemple fondamental est celui de la résolution de l'équation de Laplace. Plus précisément, on supposera ici que les domaines $\Omega_1, \ldots, \Omega_d$ sont bornés et on cherche $u \in H^1_0(\Omega_1 \times \cdots \times \Omega_d)$ l'unique solution de

$$\begin{cases} -\Delta_{x_1,\cdots,x_d} u(x_1,\cdots,x_d) = f(x_1,\cdots,x_d), & \text{pour } (x_1,\cdots,x_d) \in \Omega = \Omega_1 \times \cdots \times \Omega_d, \\ u(x_1,\cdots,x_d) = 0, & \text{pour } (x_1,\cdots,x_d) \in \partial\Omega. \end{cases}$$

La fonction u est alors de manière équivalente l'unisue solution du problème variationnel suivant:

$$\forall v \in V, a(u,v) = l(v),$$

avec $V := H^1_0(\Omega_1 \times \cdots \times \Omega_d)$, et pour tout $v,w \in V$,

$$\begin{aligned} l(v) &:= \int_\Omega fv, \\ a(v,w) &:= \int_\Omega \nabla_{x_1,\cdots,x_d} v \cdot \nabla_{x_1,\cdots,x_d} w \\ &= \sum_{i=1}^d \int_\Omega \nabla_{x_i} v \cdot \nabla_{x_i} w. \end{aligned}$$

Il est aisé de montrer, en utilisant l'inégalité de Poincaré et l'inégalité de Cauchy-Schwartz, que a et l satisfont les hypothèses du théorème de Lax-Milgram que nous avons rappelées ci-dessus.

4.1.2 Dictionnaire

L'ingrédient essentiel d'un algorithme glouton pour la résolution de problèmes de type Lax-Milgram comme celui évoqué ci-dessus est la notion de *dictionnaire*. Nous en donnons la définition ci-dessous.

Définition 4.1 (Dictionnaire). *Un sous-ensemble $\Sigma \subset V$ est appelé un dictionnaire de V si et seulement si il vérifie les trois conditions suivantes:*

(i) Σ est un cône au sens suivant: pour tout $t \in \mathbb{R}$ et $z \in \Sigma$, alors $tz \in \Sigma$;

(ii) Σ est un ensemble faiblement fermé de V;

(iii) L'espace vectoriel engendré par les éléments de Σ, $\mathrm{Vect}\{\Sigma\}$, c'est-à-dire l'ensemble des combinaisons linéaires finies d'éléments de Σ, est dense dans V.

Remarque 4.1. *L'ensemble Σ est un ensemble faiblement fermé si et seulement si pour toute suite $(z_n)_{n\in\mathbb{N}}$ d'éléments de Σ convergeant faiblement dans V vers un élément z de V, alors z appartient nécessairement à Σ. Si un ensemble est faiblement fermé dans V, alors il est nécessairement fermé (au sens de la convergence forte) dans V. Cependant, la contraposée est fausse. Un ensemble (fortement) fermé dans V peut ne pas être faiblement fermé dans V. La faible fermeture est donc une notion plus exigeante que la notion de fermeture habituelle au sens de la convergence forte.*

Nous donnons ci-dessous quelques exemples simples de dictionnaires, ainsi que d'ensembles **qui ne sont pas** des dictionnaires.

Nous considèrerons deux cas particuliers:

(a) soit $V := H^1_0(\Omega_1 \times \cdots \times \Omega_d)$ et $V_1 := H^1_0(\Omega_1)$, ..., $V_d := H^1_0(\Omega_d)$. Attention, on rappelle ici que $V \neq V_1 \bigotimes \cdots \bigotimes V_d$!

(b) soit $V_1, \cdots ,V_d$ sont des espaces de Hilbert quelconques et $V = V_1 \bigotimes \cdots \bigotimes V_d$. Un exemple de telle situation est lorsque $V = L^2(\Omega_1 \times \cdots \times \Omega_d)$, $V_1 = L^2(\Omega_1)$, ..., $V_d = L^2(\Omega_d)$.

Exemple 1: Produits tensoriels purs

Considérons tout d'abord ici le cas où le dictionnaire est l'ensemble constitué des produits tensoriels purs dans le cas (a) ou (b). Plus précisément, on note

$$\Sigma_1 := \left\{ r^1 \otimes \cdots \otimes r^d,\, r^1 \in V_1,\ \cdots,\, r^d \in V_d \right\}$$

l'ensemble des tenseurs de rang au plus 1.

On a alors le résultat suivant:

Proposition 4.1. *Dans les cas (a) et (b), Σ_1 est un dictionnaire de V.*

Exemple 2: Tenseurs de rang plus petit que R dans le cas $d = 2$

Soit $R \in \mathbb{N}^*$ tel que $R \geq 2$. Considérons maintenant le cas où $d = 2$ et où le dictionnaire est l'ensemble constitué des tenseurs de rang canonique plus petit que R. Ici, la situation est différente dans le cas (a) ou dans le cas (b). Plus précisément, on note

$$\Sigma_R := \left\{ \sum_{k=1}^{R} r^1_k \otimes r^2_k,\, r^1_k \in V_1,\, r^2_k \in V_2,\, 1 \leq k \leq R \right\}$$

l'ensemble des tenseurs de rang canonique plus petit que R.

On a alors le résultat suivant:

Proposition 4.2. *Dans le cas (a), Σ_R n'est pas un dictionnaire de V. Dans le cas (b), Σ_R est un dictionnaire de V.*

Exemple 3: Tenseurs de rang plus petit que R dans le cas $d \geq 3$

Soit $R \in \mathbb{N}^*$ tel que $R \geq 2$. Considérons maintenant le cas général où $d \geq 3$ et où le dictionnaire est l'ensemble constitué des tenseurs de rang canonique plus petit que R. Ici, la situation est différente dans le cas (a) ou dans le cas (b). Plus précisément, on note

$$\Sigma_R := \left\{ \sum_{k=1}^{R} r^1_k \otimes \cdots \otimes r^d_k,\, r^1_k \in V_1,\ \cdots,\, r^d_k \in V_d,\, 1 \leq k \leq R \right\}$$

l'ensemble des tenseurs de rang canonique plus petit que R.

On a alors le résultat suivant:

Proposition 4.3. *Dans le cas (a) et (b), Σ_R n'est pas un dictionnaire de V.*

4.1.3 Algorithmes gloutons

Nous présentons à présent les deux versions les plus classiques d'algorithmes gloutons utilisés pour résoudre des problèmes de type Lax-Milgram.

Soit Σ un dictionnaire de V. Ces algorithmes sont des algorithmes itératifs dont le but est de construire, à chaque itération $n \in \mathbb{N}^*$, une approximation $u_n \in V$ de la solution u problème (4.1) sous la forme

$$u_n = \sum_{k=1}^{n} z_k^n,$$

où pour $1 \leq k \leq n$, $z_k^n \in \Sigma$. Autrement dit, cette approximation sera construite au bout de n itérations comme la somme de n éléments du dictionnaire Σ.

Un premier algorithme glouton (le plus simple) est l'**algorithme glouton pur** qui s'écrit comme suit.

Algorithme glouton pur:

- **Initialisation:** $u_0 = 0$;
- **Itération** $n \geq 1$**:** On cherche $z_n \in \Sigma$ solution de

$$z_n \in \operatorname*{argmin}_{z \in \Sigma} \mathcal{E}(u_{n-1} + z); \tag{4.2}$$

On définit $u_n := u_{n-1} + z_n$ et on passe à l'itération suivante.

Un deuxième algorithme glouton (le plus utilisé en pratique) est l'**algorithme glouton orthogonal** qui consiste à rajouter l'étape suivante à chaque itération: au lieu de calculer u_n comme simplement la somme de u_{n-1} et de z_n comme dans l'algorithme glouton pur, on construit u_n comme la meilleur combinaison linéaire de tous les éléments $z_1, \ldots, z_n$ qui ont été calculés par l'algorithme jusque-là, au sens suivant:

Algorithme glouton orthogonal:

- **Initialisation:** $u_0 = 0$;
- **Itération** $n \geq 1$**:** On cherche $z_n \in \Sigma$ solution de

$$z_n \in \operatorname*{argmin}_{z \in \Sigma} \mathcal{E}(u_{n-1} + z); \tag{4.3}$$

On cherche $\alpha^n := (\alpha_1^n, \ldots, \alpha_n^n) \in \mathbb{R}^n$ solution de

$$\alpha^n := \operatorname*{argmin}_{\beta := (\beta_1, \ldots, \beta_n) \in \mathbb{R}^n} \mathcal{E}\left(u_{n-1} + \sum_{k=1}^{n} \beta_k z_k\right). \tag{4.4}$$

On définit $u_n := \sum_{k=1}^n \alpha_k z_k$ et on passe à l'itération suivante.

Nous avons alors le théorème suivant:

Théorème 4.1. *Toutes les itérations des algorithmes gloutons pur et orthogonal sont bien définies au sens où, pour tout $n \in \mathbb{N}^*$, il existe toujours au moins une solution $z_n \in \Sigma$ aux problèmes de minimisation (4.2) et (4.3) (pas nécessairement unique) et une unique solution au problème de minimisation (4.4). De plus, la suite $(u_n)_{n \in \mathbb{N}^*}$ converge fortement dans V vers u l'unique solution de (4.1). Enfin, si V est de dimension finie, il existe un réel $\rho \in (0,1)$ et une constante $C > 0$ tels que*

$$\forall n \in \mathbb{N}^*, \quad \|u - u_n\|_V \leq C\rho^n. \tag{4.5}$$

Remarque 4.2. *Bien sûr, la valeur de ρ apparaissant dans (4.5) dépend de la dimension de V.*

A ce stade, il est naturel de se demander comment calculer un élément $z_n \in \Sigma$ solution de (4.2). Les méthodes choisies dépendent du choix du dictionnaire. Nous présentons ici un algorithme très classiquement utilisé lorsque $\Sigma = \Sigma_1$ est le dictionnaire composé des produits tensoriels purs.

4.1.4 Algorithme de directions alternées

Nous présentons ici l'algorithme de directions alternées qui est le plus classiquement utilisé pour calculer un élément $z_n \in \Sigma$ solution de (4.2) dans le cas où $\Sigma = \Sigma_1$ est le dictionnaire composé des produits tensoriels purs.

Pour simplifier la présentation, nous ne le présentons ici que dans le cas où $d = 2$ et où on considère l'algorithme glouton pur. Il se généralise cependant sans difficulté au cas où $d \geq 3$ est un entier quelconque ou pour l'algorithme glouon orthogonal. !

Dans ce cas, le problème (4.2) à résoudre à l'itération $n \geq 1$ de l'algorithme glouton (pur ou orthogonal) se réécrit comme suit:

A l'itération n de l'algorithme, on recherche donc un couple $(r_n,s_n) \in V_1 \times V_2$ solution du problème de minimisation

$$(r_n,s_n) \in \underset{(r,s)\in V_1\times V_2}{\operatorname{argmin}} \mathcal{E}(u_{n-1} + r \otimes s), \tag{4.6}$$

où $u_{n-1} = \sum_{k=1}^{n-1} r_k \otimes s_k$ est la somme des termes calculés aux itérations précédentes.

Pour tout $n \in \mathbb{N}^*$, on définit l'application $\mathcal{J}_{n-1} : V_1 \times V_2 \to \mathbb{R}$ comme suit:

$$\forall (r,s) \in V_1 \times V_2, \quad \mathcal{J}_{n-1}(r,s) := \mathcal{E}(u_{n-1} + r \otimes s).$$

On rappelle que l'espace $V_1 \times V_2$, muni du produit scalaire défini par

$$\forall (r_1,s_1),(r_2,s_2) \in H^1_0(0,1)\times H^1_0(0,1), \quad \langle (r_1,s_2),(r_2,s_2)\rangle_{H^1_0(0,1)\times H^1_0(0,1)} := \langle r_1,r_2\rangle_{H^1_0(0,1)} + \langle s_1,s_2\rangle_{H^1_0(0,1)},$$

est un espace de Hilbert.

Exercice 4.1. *(1) Montrer que l'application $\mathcal{J}_{n-1}$ est différentiable sur $V_1 \times V_2$ et calculer sa différentielle en fonction de la forme bilinéaire a et de la forme linéaire l.*

(2) En déduire que si $(r_n,s_n) \in V_1 \times V_2$ est une solution de (4.13), alors (r_n,s_n) est solution du système d'équations couplées:

$$\begin{cases} \forall s \in V_2, \quad a(r_n \otimes s_n, r_n \otimes s) = l(r_n \otimes s) - a(u_{n-1}, r_n \otimes s), \\ \forall r \in V_1, \quad a(r_n \otimes s_n, r \otimes s_n) = l(r \otimes s_n) - a(u_{n-1}, r \otimes s_n), \end{cases} \tag{4.7}$$

(3) Montrer que ces équations peuvent se réécrire sous la forme:

$$\forall s \in V_2, \quad a^{r_n}(s_n,s) = l^{u_{n-1},r_n}(s), \tag{4.8}$$

et

$$\forall r \in v_1, \quad a^{s_n}(r_n,r) = l^{u_{n-1},s_n}(r), \tag{4.9}$$

où $a^{r_n} : V_2 \times V_2 \to \mathbb{R}$ et $a^{s_n} : V_1 \times V_1 \to \mathbb{R}$ sont deux formes bilinéaires symmétriques et continues, et où $l^{u_{n-1},r_n} : V_2 \to \mathbb{R}$ et $l^{u_{n-1},s_n} : V_1 \to \mathbb{R}$ sont des formes linéaires continues, dont on donnera les expressions en fonction de a, l r_n et s_n.

(4) On supposera de plus qu'il existe $C_1 : V_1 \to \mathbb{R}_+$ et $C_2 : V_2 \to \mathbb{R}_+$ tels qeue $C_1(r) > 0$ pour tout $r \neq 0$ et $C_2(s) > 0$ pour tout $s \neq 0$ et tels que pour tout $r \in V_1$, $s \in V_2$, $\|r \otimes s\|_V \geq \max(C_1(r)\|s\|_{V_2}, C_2(s)\|r\|_{V_1})$. Montrer alors que si $r_n \neq 0$ (respectivement $s_n \neq 0$), a^{r_n} (respectivement a^{s_n}) est coercive.

(5) En déduire pour tout $r_n \in V_1$ tel que $r_n \neq 0$ (respectivement pour tout $s_n \in V_2$ tel que $s_n \neq 0$), il existe une unique solution $s_n \in V_2$ (respectivement $r_n \in V_1$) au problème variationnel (4.16) (respectivement (4.17)).

En pratique, le système d'équations (4.7) est résolu par un algorithme de point fixe, appelé directions alternées, qui s'écrit plus précisément comme suit:

Algorithme de directions alternées:

(1) **Initialisation:** On choisit $(r_n^{(0)}, s_n^{(0)}) \in V_1 \times V_2$ aléatoirement.

(2) **Itération** $m \geq 1$**:** On définit $r_n^{(m)} \in V_1$ comme l'unique solution du problème

$$\forall r \in V_1, \quad a^{s_n^{(m-1)}}(r_n^{(m)},r) = l^{u_{n-1},s_n^{(m-1)}}(r). \tag{4.10}$$

Puis, on définit $s_n^{(m)} \in V_2$ comme l'unique solution du problème

$$\forall s \in V_2, \quad a^{r_n^{(m)}}(s_n^{(m)},s) = l^{u_{n-1},r_n^{(m)}}(s). \tag{4.11}$$

Il est par ailleurs facile de voir que l'unique solution $r_n^{(m)} \in V_1$ de (4.10) est également l'unique solution de

$$r_n^{(m)} = \underset{r \in V_1}{\operatorname{argmin}}\, \mathcal{J}_{n-1}(r,s_n^{(m-1)}) = \underset{r \in V_1}{\operatorname{argmin}}\, \mathcal{E}(u_{n-1} + r \otimes s_n^{(m-1)}).$$

De même, l'unique solution $s_n^{(m)} \in V_2$ de (4.11) est également l'unique solution de

$$s_n^{(m)} = \underset{s \in V_2}{\operatorname{argmin}}\, \mathcal{J}_{n-1}(r_n^{(m)},s) = \underset{s \in V_2}{\operatorname{argmin}}\, \mathcal{E}(u_{n-1} + r_n^{(m)} \otimes s).$$

En pratique, on observe numériquement que cet algorithme de point fixe converge très souvent exponentiellement vite en fonction du nombre d'itérations m.

4.2 Cas de la résolution de l'équation de Laplace

Le but de cette section est de vous illustrer comment les algorithmes gloutons et algorithmes de directions alternées peuvent être mis en oeuvre sur la résolution de l'équation de Laplace.

Dans cette section, nous supposerons pour simplifier que $d = 2$ et que $\Omega_1 = \Omega_2 = (0,1)$ et ne considèrerons que l'algorithme glouton pur.

Le principe de l'algorithme est de construire une approximation de la solution u du problème de Laplace:

chercher $u \in H_0^1((0,1)^2)$ solution de

$$\begin{cases} -\Delta_{x_1,x_2} u(x_1,x_2) = f(x_1,x_2), & (x_1,x_2) \in (0,1)^2? \\ u(x_1,x_2) = 0, & (x_1,x_2) \in \partial(0,1)^2, \end{cases}$$

avec $f \in L^2((0,1)^2)$, sous la forme d'une somme de fonctions produits tensoriels, i.e. sous la forme

$$u(x_1,x_2) \approx \sum_{k=1}^{n} r_k \otimes s_k(x_1,x_2) = \sum_{k=1}^{n} r_k(x_1)s_k(x_2),$$

où chaque terme apparaissant dans la somme ci-dessus est calculé de manière itérative, et où pour tout $1 \leq k \leq n$, $r_k,s_k \in H_0^1(0,1)$.

Dans ce cas, en particulier, $V_1 = V_2 = H_0^1(0,1)$ et $V = H_0^1((0,1)^2)$. De plus, pour tout $v,w \in V$,

$$a(v,w) = \int_{(0,1)^2} \partial_{x_1} v \partial_{x_1} w + \partial_{x_2} v \partial_{x_2} w$$

et

$$l(v) = \int_{(0,1)^2} fv.$$

4.2.1 Au niveau continu

Dans cette section, nous allons présenter en détails comment l'algorithme glouton pur et l'algorithme des directions alternées présenté précédemment pour des problèmes de Lax-Milgram généraux peut être mis en oeuvre en pratique dans le cas de l'équation de Laplace.

Dans ce cas particulier, l'algorithme glouton s'écrit comme suit:

- **Initialisation:** Soit $u_0(x_1,x_2) := 0$.
- **Itération** $n \geq 1$**:** On choisit $r_n,s_n \in H_0^1(0,1)$ comme une solution du problème de minimisation

$$(r_n,s_n) \in \underset{r,s \in H_0^1(0,1)}{\operatorname{argmin}} \ \mathcal{E}\left(u_{n-1} + r \otimes s\right) \tag{4.12}$$

On définit ensuite $u_n = u_{n-1} + r_n \otimes s_n$.

On voit qu'à l'itération $n \geq 1$ de l'algorithme, on a par récurrence

$$u_n = u_{n-1} + r_n \otimes s_n = \sum_{k=1}^{n} r_k \otimes s_k.$$

A l'itération n de l'algorithme, on recherche donc un couple $(r_n,s_n) \in (H_0^1(0,1))^2$ solution du problème de minimisation

$$(r_n,s_n) \in \underset{(r,s) \in H_0^1(0,1)}{\operatorname{argmin}} \ \mathcal{E}(u_{n-1} + r \otimes s), \tag{4.13}$$

où $u_{n-1}(x_1,x_2) = \sum_{k=1}^{n-1} r_k(x_1)s_k(x_2)$ est la somme des termes calculés aux itérations précédentes. Pour tout $n \in \mathbb{N}^*$, on définit l'application $\mathcal{J}_{n-1} : H_0^1(0,1) \times H_0^1(0,1) \to \mathbb{R}$ comme suit:

$$\forall (r,s) \in H_0^1(0,1) \times H_0^1(0,1), \quad \mathcal{J}_{n-1}(r,s) := \mathcal{E}(u_{n-1} + r \otimes s).$$

On rappelle que l'espace $H_0^1(0,1) \times H^1(0,1)$, muni du produit scalaire défini par

$$\forall (r_1,s_1),(r_2,s_2) \in H_0^1(0,1) \times H_0^1(0,1), \quad \langle (r_1,s_2),(r_2,s_2) \rangle_{H_0^1(0,1)\times H_0^1(0,1)} := \langle r_1,r_2 \rangle_{H_0^1(0,1)} + \langle s_1,s_2 \rangle_{H_0^1(0,1)},$$

est un espace de Hilbert.

Exercice 4.2. *(1) Montrer que si* $(r_n,s_n) \in H_0^1(0,1) \times H_0^1(0,1)$ *est une solution de (4.13), alors* (r_n,s_n) *est solution du système d'équations couplées: pour tout* $r,s \in H_0^1(0,1)$,

$$\left\{\begin{array}{l} \|r_n'\|^2_{L^2(0,1)} \langle s_n,s \rangle_{L^2(0,1)} + \|r_n\|^2_{L^2(0,1)} \langle s_n',s' \rangle_{L^2(0,1)} \\ = \langle f, r_n \otimes s \rangle_{L^2((0,1)^2)} \\ -\langle \partial_{x_1} u_{n-1}, r_n' \otimes s \rangle_{L^2((0,1)^2)} + \langle \partial_{x_2} u_{n-1}, r_n \otimes s' \rangle_{L^2(0,1)^2}, \\ \\ \|s_n'\|^2_{L^2(0,1)} \langle r_n,r \rangle_{L^2(0,1)} + \|s_n\|^2_{L^2(0,1)} \langle r_n',r' \rangle_{L^2(0,1)} \\ = \langle f, r \otimes s_n \rangle_{L^2((0,1)^2)} \\ -\langle \partial_{x_1} u_{n-1}, r' \otimes s_n \rangle_{L^2((0,1)^2)} + \langle \partial_{x_2} u_{n-1}, r \otimes s_n' \rangle_{L^2(0,1)^2}. \end{array}\right. \tag{4.14}$$

(2) Montrer que $(r_n,s_n) \in H_0^1(0,1)^2$ *est de manière équivalente solution du système d'EDPs couplées suivant:*

$$\left\{\begin{array}{l} \|r_n'\|^2_{L^2(0,1)} s_n(x_2) + \|r_n\|^2_{L^2(0,1)} (-s_n''(x_2)) \\ = \int_0^1 f(x_1,x_2) r_n(x_1)\, dx_1 \\ - \int_0^1 \partial_{x_1} u_{n-1}(x_1,x_2) r_n'(x_1)\, dx_1 + \partial_{x_2} \left(\int_0^1 \partial_{x_2} u_{n-1}(x_1,x_2) r_n(x_1)\, dx_1 \right), \\ \\ \|s_n'\|^2_{L^2(0,1)} r_n(x_1) + \|s_n\|^2_{L^2(0,1)} (-r_n''(x_1)) \\ = \int_0^1 f(x_1,x_2) s_n(x_2)\, dx_2 \\ - \int_0^1 \partial_{x_2} u_{n-1}(x_1,x_2) s_n'(x_2)\, dx_2 + \partial_{x_1} \left(\int_0^1 \partial_{x_1} u_{n-1}(x_1,x_2) s_n(x_2)\, dx_2 \right). \end{array}\right. \tag{4.15}$$

(3) Montrer que ces équations peuvent se réécrire sous la forme:

$$\forall s \in H_0^1(0,1), \quad a^{r_n}(s_n,s) = l^{f,u_{n-1},r_n}(s), \tag{4.16}$$

et

$$\forall r \in H_0^1(0,1), \quad a^{s_n}(r_n,r) = l^{f,u_{n-1},s_n}(r), \tag{4.17}$$

où $a^{r_n} : H_0^1(0,1) \times H_0^1(0,1) \to \mathbb{R}$ *et* $a^{s_n} : H_0^1(0,1) \times H_0^1(0,1) \to \mathbb{R}$ *sont deux formes bilinéaires symmétriques et continues, et où* $l^{f,u_{n-1},r_n} : H_0^1(0,1) \to \mathbb{R}$ *et* $l^{f,u_{n-1},s_n} : H_0^1(0,1) \to \mathbb{R}$ *sont des formes linéaires continues, dont on donnera les expressions en fonction de* a, l r_n *et* s_n. *Montrer de plus que si* $r_n \neq 0$ *(resepctivement* $s_n \neq 0$*),* a^{r_n} *(respectivement* a^{s_n}*) est coercive.*

(4) En déduire pour tout $r_n \in H_0^1(0,1)$ *tel que* $r_n \neq 0$ *(respectivement pour tout* $s_n \in H_0^1(0,1)$ *tel que* $s_n \neq 0$*), il existe une unique solution* $s_n \in H_0^1(0,1)$ *(resepctivement* $r_n \in H_0^1(0,1)$*) au problème variationnel (4.16) (respectivement (4.17)).*

(5) On suppose qu'il existe $P \in \mathbb{N}^*$ *tel que* $f(x_1,x_2) = \sum_{p=1}^P f_p^1(x_1) f_p^2(x_2)$ *où pour tout* $1 \leq p \leq P$, $f_p^1, f_p^2 \in L^2(0,1)$. *Montrer que*

$$\begin{aligned} l^{f,u_{n-1},r_n}(s) &= \textstyle\sum_{p=1}^P \langle f_p^1, r_n\rangle_{L^2(0,1)} \langle f_p^2, s\rangle_{L^2(0,1)} \\ &\quad - \left(\textstyle\sum_{k=1}^{n-1} \langle r_k, r_n\rangle_{L^2(0,1)} \langle s_k', s'\rangle_{L^2(0,1)} + \langle r_k', r_n'\rangle_{L^2(0,1)} \langle s_k, s\rangle_{L^2(0,1)}\right) \end{aligned} \tag{4.18}$$

et

$$\begin{aligned} l^{f,u_{n-1},s_n}(r) &= \textstyle\sum_{p=1}^P \langle f_p^2, s_n\rangle_{L^2(0,1)} \langle f_p^1, r\rangle_{L^2(0,1)} \\ &\quad - \left(\textstyle\sum_{k=1}^{n-1} \langle s_k, s_n\rangle_{L^2(0,1)} \langle r_k', r'\rangle_{L^2(0,1)} + \langle s_k', s_n'\rangle_{L^2(0,1)} \langle r_k, r\rangle_{L^2(0,1)}\right) \end{aligned} \tag{4.19}$$

On rappelle que $u_{n-1}(x_1,x_2) = \sum_{k=1}^{n-1} r_k(x_1) s_k(x_2)$.

En pratique, le système d'équations (4.7) est résolu par un algorithme de point fixe, qui s'écrit plus précisément comme suit: soit $\epsilon > 0$ un critère d'erreur fixé.

Algorithme de directions alternées:

(1) **Initialisation:** On choisit $(r_n^{(0)}, s_n^{(0)}) \in H_0^1(0,1)$ aléatoirement.

(2) **Itération** $m \geq 1$**:** On définit $r_n^{(m)} \in H_0^1(0,1)$ comme l'unique solution du problème

$$\forall r \in H_0^1(0,1), \quad a^{s_n^{(m-1)}}(r_n^{(m)}, r) = l^{f,u_{n-1},s_n^{(m-1)}}(r).$$

Puis, on définit $s_n^{(m)} \in H_0^1(0,1)$ comme l'unique solution du problème

$$\forall s \in H_0^1(0,1), \quad a^{r_n^{(m)}}(s_n^{(m)}, s) = l^{f,u_{n-1},r_n^{(m)}}(s).$$

Si $\|r_n^{(m)} \otimes s_n^{(m)} - r_n^{(m-1)} \otimes s_n^{(m-1)}\|_{H_0^1((0,1)^2)} < \epsilon$, on arrête l'algorithme et on définit $r_n = r_n^{(m)}$ et $s_n = s_n^{(m)}$. Sinon, on passe à l'itération suivante: $m := m + 1$.

En pratique, on observe numériquement que cet algorithme de point fixe converge très souvent exponentiellement vite en fonction du nombre d'itérations m.

4.2.2 Discrétisation en éléments finis

Nous présentons dans cette partie comment l'algorithme PGD présenté dans la section précédente peut être implémenté en pratique à l'aide d'une discrétisation en éléments finis. Le but du TP sur la méthode PGD sera d'implémenter l'algorithme discret décrit ci-dessous. Nous nous plaçons toujours dans le cas où $d = 2$ pour simplifier.

Soit $(\phi_i)_{1\leq i\leq I}$ une famille libre de fonctions de $H^1_0(0,1)$, qui sera typiquement une base de fonctions éléments finis. On note $V_I := \mathrm{Vect}\{\phi_i(x), 1 \leq i \leq I\}$. On note $D := (D_{ij})_{1\leq i,j\leq I} \in \mathbb{R}^{I\times I}$ et $M := (M_{ij})_{1\leq i,j\leq I} \in \mathbb{R}^{I\times I}$ les matrices définies comme suit:

$$\forall 1 \leq i,j \leq I, \quad D_{ij} := \int_0^1 \phi_i'\phi_j' \quad \text{et} \quad M_{ij} := \int_0^1 \phi_i\phi_j.$$

On suppose également qu'il existe $P \in \mathbb{N}^*$ tel que $f(x_1,x_2) = \sum_{p=1}^P f^1_p(x_1)f^2_p(x_2)$ où pour tout $1 \leq p \leq P$, $f^1_p, f^2_p \in L^2(0,1)$. Pour tout $1 \leq p \leq P$, on note $F^1_p := (F^1_{p,i})_{1\leq i\leq I} \in \mathbb{R}^I$ et $F^2_p := (F^2_{p,i})_{1\leq i\leq I} \in \mathbb{R}^I$ les vecteurs définis comme suit:

$$\forall 1 \leq i \leq I, \quad F^1_{p,i} := \int_0^1 f^1_p\phi_i \quad \text{et} \quad F^2_{p,i} := \int_0^1 f^2_p\phi_i.$$

A l'itération $n \in \mathbb{N}^*$ de l'algorithme, on cherche à approcher les solutions $r_n, s_n \in H^1_0(0,1)$ des équations (4.17) et (4.16) par des fonctions $r^I_n, s^I_n \in V_I$ solutions de

$$\forall s^I \in V_I, \quad a^{r^I_n}(s^I_n, s^I) = l^{f,u^I_{n-1},r^I_n}(s^I), \tag{4.20}$$

et

$$\forall r^I \in V_I, \quad a^{s^I_n}(r^I_n, r^I) = l^{f,u^I_{n-1},s^I_n}(r^I), \tag{4.21}$$

où $u^I_{n-1}(x_1,x_2) = \sum_{k=1}^{n-1} r^I_k(x_1)s^I_k(x_2)$ est la somme des termes calculés aux itérations précédentes.

Exercice 4.3. *(1) Soit $R = (R_i)_{1\leq i\leq I}, S = (S_i)_{1\leq i\leq I} \in \mathbb{R}^I$. Montrer que si $r^I, s^I \in V_I$ sont les fonctions définies comme*

$$r^I(x) = \sum_{i=1}^I R_i\phi_i(x) \text{ et } s^I(x) = \sum_{i=1}^I S_i\phi_i(x),$$

alors

$$\int_0^1 (r^I)'(x)(s^I)'(x)\,dx = R^TDS = S^TDR \text{ et } \int_0^1 r^I(x)s^I(x)\,dx = R^TMS = S^TMR.$$

(2) Pour tout $1 \leq k \leq n$, on note $R_k := (R_{k,i})_{1\leq i\leq I} \in \mathbb{R}^I$ et $S_k := (S_{k,i})_{1\leq i\leq I} \in \mathbb{R}^I$ les vecteurs des coordonnées des fonctions r^I_k et s^I_k dans la base $(\phi_i)_{1\leq i\leq I}$ de telle sorte que:

$$r^I_k(x_1) = \sum_{i=1}^I R_{k,i}\phi_i(x_1) \quad \text{et } s^I_k(x_2) = \sum_{i=1}^I S_{k,i}\phi_i(x_2).$$

Montrer que $r^I_n, s^I_n \in V_I$ sont solutions des équations (4.20) et (4.21) si et seulement si R_n et S_n sont solutions des systèmes matriciels couplés suivants:

$$A_R(S_n)R_n = f^{n-1}_R(S_n) \text{ et } A_S(R_n)S_n = f^{n-1}_S(R_n),$$

où pour tout $R, S \in \mathbb{R}^N$,

$$\begin{aligned}
A_R(S) &= (S^T DS)M + (S^T MS)D \in \mathbb{R}^{I\times I},\\
A_S(R) &= (R^T DR)M + (R^T MR)D \in \mathbb{R}^{I\times I},\\
f_R^{n-1}(S) &= \sum_{p=1}^{P}(S^T F_p^2)F_p^1 - \sum_{k=1}^{n-1}((S_k^T DS_k)MR_k + (S_k^T MS_k)DR_k) \in \mathbb{R}^I,\\
f_S^{n-1}(R) &= \sum_{p=1}^{P}(R^T F_p^1)F_p^2 - \sum_{k=1}^{n-1}((R_k^T DR_k)MS_k + (R_k^T MR_k)DS_k) \in \mathbb{R}^I.
\end{aligned}$$

(3) Ecrire l'algorithme de point fixe présenté à la section précédente pour le calcul des fonctions r_n,s_n) sous forme discrétisée.