

# EDP et Approximation

Niveau M1

C. Chalons (UVSQ)

La rédaction de ces notes s'est inspirée

- du polycopié de Frédéric Lagoutière sur les équations aux dérivées partielles et leur approximation

- du livre de Brigitte Lucquin intitulé "Equations aux dérivées partielles et leurs approximations" (éditeur Ellipses Marketing, collection Mathématiques Licence-Masters)



Les équations aux dérivées partielles interviennent de manière récurrente dans la modélisation des problèmes réels. Elles traduisent le plus souvent des lois de la physique (mécanique, thermodynamique, électromagnétisme, acoustique...) écrites généralement sous forme de bilans. Elles interviennent également dans beaucoup d'autres domaines comme la chimie pour modéliser les réactions, l'économie pour étudier le comportement des marchés, la finance pour étudier les produits dérivés (options et obligations)... et sont utilisées pour modéliser et (tentar de) comprendre de nombreux systèmes complexes : notre appareil respiratoire, l'évolution météorologique et climatique, le fonctionnement d'un piano, d'un écosystème, l'évolution d'un réseau autoroutier...

Les équations aux dérivées partielles constituent encore aujourd'hui un sujet de recherche très actif. Pour la bonne et simple raison que d'une manière générale il est très difficile, voire impossible d'exhiber les solutions d'une équation aux dérivées partielles. Il s'agit donc d'essayer de montrer, lorsque cela est possible, que le problème admet une unique solution stable (on dit qu'il est "bien posé") puis de tenter d'approcher numériquement à l'aide d'un algorithme programmé sur ordinateur) là où les solutions du problème considéré. Ces problèmes ont été à l'origine de la création de nombreux concepts mathématiques nouveaux et ont permis l'effectuer des avancées concrètes dans de nombreuses applications.

Le cours propose une introduction à l'analyse mathématique et à l'approximation numérique des solutions de certaines équations aux dérivées partielles :

## Définitions

Soit  $d, m, n, s$  des entiers naturels non nuls. Soit  $\Omega$  un ouvert de  $\mathbb{R}^d$ .

On appelle système d'équations aux dérivées partielles (EDP) à coefficients réels, dans  $\mathbb{R}^d$ , de taille  $n$  et d'ordre  $m$  une relation de la forme

$$\Psi \left( x, u(x), \left( \partial_{i_1} u(x) \right)_{(i_1) \in \{1, \dots, d\}}, \left( \partial_{i_1 i_2}^2 u(x) \right)_{(i_1, i_2) \in \{1, \dots, d\}^2}, \dots, \right. \\ \left. \dots \left( \partial_{i_1, \dots, i_m}^m u(x) \right)_{(i_1, \dots, i_m) \in \{1, \dots, d\}^m} \right) = 0$$

où  $u: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$  est la solution de l'EDP (c'est l'inconnue)

et  $\Psi: \Omega \times \mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}^s$ .

$s$  représente le nombre de relations formant le système d'EDP. On aura habituellement  $s=n$ , ce qui signifie qu'on aura le même nombre d'équations que d'inconnues.

On dit que l'EDP est linéaire si et seulement si  $\Psi(x, \cdot)$  l'est pour tout  $x$  dans  $\Omega$ .

On dit que l'EDP est homogène si et seulement si  $u=0$  est solution de l'EDP.

On dit que l'équation est scalaire si  $s=n=1$ .

On étudiera uniquement dans ce cours des EDP d'ordre 1 ou 2, 1 dimension 1 ou 2 d'espace.

## Classification des EDP scalaires linéaires d'ordre 2

3/

Dans ce paragraphe on considère une EDP scalaire ( $s=m=1$ ) linéaire et d'ordre 2 ( $m=2$ ), de la forme

$$a(x)u(x) + \sum_{i=1}^d b_i(x) \partial_i u(x) + \sum_{i=1}^d \sum_{j=1}^d c_{ij}(x) \partial_{ij}^2 u(x) = f(x).$$

On note  $C(x)$  la matrice des  $(c_{ij}(x))_{(i,j) \in \{1, \dots, d\}^2}$ . A la modification près  $\tilde{c}_{ij} = \frac{1}{2}(c_{ij} + c_{ji})$  qui n'a pas d'influence sur l'EDP si la solution est de classe  $C^2$  puisque  $c_{ij} \partial_{ij}^2 u + c_{ji} \partial_{ji}^2 u = \frac{1}{2}(c_{ij} + c_{ji}) \partial_{ij}^2 u + \frac{1}{2}(c_{ij} + c_{ji}) \partial_{ji}^2 u$ , on peut toujours supposer que  $C(x)$  est symétrique. Cette matrice est donc diagonalisable dans  $\mathbb{R}$ . En particulier, ses valeurs propres sont réelles.

On note  $(\lambda_i(x))_{i=1}^d$  les valeurs propres,  $d_+(x)$  le nombre de valeurs propres strictement positives,  $d_-(x)$  le nombre de valeurs propres strictement négatives et  $d_0(x)$  la multiplicité de la valeur propre 0. Puisque le comptage s'effectue ici en tenant compte de la multiplicité des valeurs propres, on a donc  $d = d_0(x) + d_+(x) + d_-(x)$ ,  $\forall x \in \Omega$ .

### Définition

On dit que l'EDP est elliptique en  $x \in \Omega$  ssi

$$d_+(x) = d \text{ ou } d_-(x) = d$$

on dit que l'EDP est hyperbolique en  $x \in \Omega$  ssi

$$d_+(x) = d-1 \text{ et } d_-(x) = 1$$

$$\text{ou } d_+(x) = 1 \text{ et } d_-(x) = d-1$$

on dit que l'EDP est parabolique en  $x \in \Omega$  ssi

$$d_0(x) > 0$$

Déformation d'une membrane

On considère une membrane élastique horizontale soumise à un chargement vertical, et fixée à son bord. L'altitude de la membrane est solution de

$$\begin{cases} -\Delta u(x) + c(x)u(x) = f(x) & \forall x \in \Omega \\ u(x) = 0 & \forall x \in \partial\Omega \end{cases} \quad \begin{matrix} x = (x_1, x_2) \\ \Omega \subset \mathbb{R}^2 \end{matrix}$$

avec  $\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$ . Il s'agit d'un problème elliptique.

$$d=2 \quad C = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Diffusion de la chaleur

On s'intéresse à la température d'un fil sur le segment  $[0, 1]$ , que l'on note  $u(t, x)$ . Une modélisation simplifiée de ce problème est donnée par

$$\begin{cases} \partial_t u - k \partial_{xx}^2 u = f & \forall t > 0, \forall x \in ]0, 1[ \\ u(t, 0) = u(t, 1) = 0 & \forall t \geq 0 \\ u(0, x) = u_0(x) & \forall x \in ]0, 1[ \end{cases}$$

où  $f$  représente le terme de chauffage,  $k$  le coefficient de conductivité thermique et  $u_0$  l'état initial. Il s'agit d'un problème parabolique.

$$d=2 \quad C = \begin{pmatrix} 0 & 0 \\ 0 & -k \end{pmatrix}$$

Propagation d'onde

La propagation d'une onde dans un milieu infini peut être modélisée par

$$\begin{cases} \partial_{tt}^2 u - c^2 \partial_{xx}^2 u = 0 & \forall x \in \mathbb{R}, \forall t > 0 \\ u(0, x) = u_0(x) & \forall x \in \mathbb{R} \\ \partial_t u(0, x) = u_1(x) & \forall x \in \mathbb{R} \end{cases}$$

où  $c$  est la vitesse de propagation de l'onde et les fonctions  $u_0$  et  $u_1$  sont respectivement l'état et la vitesse initiaux. Il s'agit d'un problème

hyperbolique.

$$d=2 \quad C = \begin{pmatrix} 1 & 0 \\ 0 & -c^2 \end{pmatrix}$$

## Le théorème de Cauchy-Kovalevskaya

5

Il s'agit d'un théorème très général donnant des solutions pour des équations aux dérivées partielles très générales et d'ordre arbitrairement grand. Plus précisément, on considère le système (ici d'ordre 1)

$$\partial_t u_j = \sum_{i=1}^d \sum_{k=1}^n L_{ik}^j(x, u_1, \dots, u_n) \partial_i u_k + \beta^j(x, u_1, \dots, u_n), \quad \forall j=1, \dots, n$$

avec les données initiales  $u_j(0, x) = 0 \quad \forall x, \quad \forall j=1, \dots, n$ , et que l'on suppose que les coefficients  $L_{ik}^j$  et  $\beta^j$  sont analytiques au voisinage de  $(x, u_1, \dots, u_n) = (0, \dots, 0)$ , alors le problème de Cauchy admet une unique solution analytique au voisinage de  $(t, x) = (0, 0)$ .

### Remarque très importante

Ce théorème est très puissant puisqu'il fournit un résultat d'existence et d'unicité de solution pour une classe très générale d'équations aux dérivées partielles. Néanmoins, l'existence et l'unicité pour un problème soit "bien posé" n'est pas l'ingrédient principal au sens de Hadamard : la stabilité est au moins tout aussi importante. Or, il se trouve que ce théorème peut "fournir" des solutions instables, ce qui limite le champ d'application du résultat (qui n'en demeure pas moins impressionnant).

Ce, et surtout, ce théorème est la pierre angulaire de la théorie des équations paraboliques, elliptiques, hyperboliques sans la distinguer.

Afin d'illustrer ce propos, on peut considérer l'équation

$$\begin{cases} \partial_t u + i \partial_x u = 0 \\ u(0, x) = e \end{cases} \quad \text{avec } A \gg 1 \text{ donné}$$

La solution de cette équation est donnée par

$$u(t, x) = e^{A(ix+t)}$$

à nouveau alors que

$$\|u(t, x)\| = e^{2t}$$

alors que pour la fonction initiale

$$\|u(0, x)\| = 1$$

La norme en fait de la solution explose tout au long plus que celle de la fonction initiale en temps. Il y a bien existence et unicité d'une solution mais le théorème ne dit rien sur la stabilité de la solution.

Quand

Pour revenir à notre problème l'équation

$$\begin{cases} Lu = Lu_x = 0 \\ u(0, x) = e^{2x} \end{cases}$$

(est hyperbolique et) a pour solution

$$u(t, x) = e^{2x}$$

Soit alors  $\|u(t, x)\| = \|u(0, x)\| = 1$

Dans ce cas la solution est stable si le problème est bien posé au sens de Hadamard.

Mais revenons sur cette notion de stabilité dans la suite du cours. Par manque de temps, nous n'aborderons pas les problèmes paraboliques, et nous nous focaliserons en grande partie sur les problèmes elliptiques. Les systèmes hyperboliques seront abordés dans la dernière partie du cours au travers de l'équation des ondes.



# Rappels sur les distributions et les espaces de Sobolev

L'objectif de ce chapitre est de faire des rappels sur les distributions et les espaces de Sobolev qui sont indispensables à l'étude des équations aux dérivées partielles. Dans toute la suite,  $\Omega$  désigne un ouvert non vide de  $\mathbb{R}^n$ .

## Les distributions

Définition On note  $\mathcal{D}(\Omega)$  l'espace vectoriel des fonctions définies sur  $\Omega$ , à valeurs réelles, et qui sont de classe  $\mathcal{C}^\infty$  sur  $\Omega$  et à support compact inclus dans  $\Omega$ .

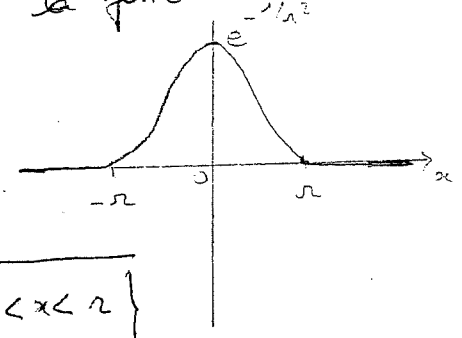
## Remarques

1) On rappelle que le support d'une fonction  $f$  est le complémentaire du plus grand ouvert sur lequel  $f$  est nulle, ou encore la fermeture de l'ensemble des points  $x \in \Omega$  pour lesquels  $f(x) \neq 0$ :

$$\text{Supp}(f) = \overline{\{x \in \Omega, f(x) \neq 0\}}$$

2) Il est important de remarquer que l'ensemble  $\mathcal{D}(\Omega)$  ainsi défini n'est pas vide. Par exemple dans le cas  $n=1$ , la fonction

$$\begin{cases} \varphi(x) = e^{-\frac{1}{x^2 - r^2}} & \text{si } -r < x < r \\ \varphi(x) = 0 & \text{sinon} \end{cases}$$



vérifie bien  $\varphi \in \mathcal{C}^\infty(\mathbb{R})$  et  $\text{Supp}(\varphi) = \overline{\{x \in \mathbb{R}, -r < x < r\}} = \{x \in \mathbb{R}, -r \leq x \leq r\} = \overline{B(0, r)}$ .

Pour  $n > 1$ , on vérifie plus généralement que la fonction

$$\begin{cases} \varphi(x) = e^{-\frac{1}{\|x\|^2 - r^2}} & \text{si } x \in B(0, r) \subset \mathbb{R}^n \\ \varphi(x) = 0 & \text{sinon} \end{cases}$$

est aussi dans  $\mathcal{D}(\Omega)$ .

3) Les fonctions de  $\mathcal{D}(\Omega)$  sont très régulières mais pas nécessairement analytiques. On remarque en effet, dans le cas  $n=1$  pour fixer les idées, que la fonction de l'exemple ci-dessus vérifie

$$\psi(x) = \psi'(x) = \psi''(x) = \dots = 0$$

et pourtant  $\psi$  n'est pas nulle dans un voisinage de  $x$ . Elle ne coïncide donc pas avec sa série de Taylor (qui est nulle) au voisinage de  $x$ .

4) On a bien sûr  $\psi \in \mathcal{D}(\Omega) \Rightarrow \psi' \in \mathcal{D}(\Omega) \Rightarrow \psi'' \in \mathcal{D}(\Omega) \dots$

### Définition (topologie sur $\mathcal{D}(\Omega)$ )

On dit qu'une suite  $(\psi_p)_{p \in \mathbb{N}}$  de  $\mathcal{D}(\Omega)$  converge vers une fonction  $\psi$  de  $\mathcal{D}(\Omega)$  si

(i)  $\exists$  un compact fixe  $K$  de  $\Omega$  contenant le support de toutes les fonctions  $\psi_p$  (au moins à partir d'un certain rang) et le support de  $\psi$ .

(ii)  $\forall \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$  un multi-indice de dérivation,

$(D^\alpha \psi_p)_{p \in \mathbb{N}}$  converge uniformément vers  $D^\alpha \psi$ , i.e.

$$\lim_{p \rightarrow +\infty} \sup_{\substack{x \in K \\ (\text{ou } x \in \Omega)}} |D^\alpha \psi_p(x) - D^\alpha \psi(x)| = 0$$

Remarque: la convergence de  $\mathcal{D}(\Omega)$  implique la convergence dans tous les  $L^p$ ,  $1 \leq p \leq +\infty$ .

Pour admettre les résultats de densité suivants:

Lemme L'espace  $\mathcal{D}(\Omega)$  est dense dans  $L^p(\Omega)$ ,  $\forall p$  t.q.  $1 \leq p < \infty$

Lemme Soit  $f \in L^1_{loc}(\Omega)$  telle que  $\forall \psi \in \mathcal{D}(\Omega)$  on ait  $\int_{\Omega} f(x) \psi(x) dx = 0$ . Alors  $f \equiv 0$  p.p sur  $\Omega$ .

Exercice sur densité de  $\mathcal{D}(\Omega)$

## Définition (distribution)

Soit  $T$  une application linéaire définie sur  $\mathcal{D}(\Omega)$  et à valeurs réelles. On dit que  $T$  est une distribution si  $T$  est continue pour la topologie définie sur  $\mathcal{D}(\Omega)$ , ie si pour toute suite  $(\varphi_p)_{p \in \mathbb{N}} \in \mathcal{D}(\Omega)$  qui converge vers  $\varphi \in \mathcal{D}(\Omega)$ , on a  $T(\varphi_p) \xrightarrow{p \rightarrow +\infty} T(\varphi)$  ds  $\mathbb{R}$ .

On note  $\mathcal{D}'(\Omega)$  l'espace vectoriel des distributions sur  $\Omega$ .

Remarque:  $T$  étant linéaire, il suffit de vérifier que  $\varphi_p \xrightarrow[p \rightarrow +\infty]{} 0 \Rightarrow T(\varphi_p) \xrightarrow[p \rightarrow +\infty]{} 0$ .

## Exemples

1) La distribution de Dirac en un point  $a$  de  $\Omega$ , notée  $\delta_a$  et définie par

$$\delta_a: \mathcal{D}(\Omega) \rightarrow \mathbb{R}$$

$$\varphi \mapsto \delta_a(\varphi) = \varphi(a)$$

est une distribution. En effet  $\varphi_p \xrightarrow[p \rightarrow +\infty]{} 0 \Rightarrow \sup_x |\varphi_p(x)| \xrightarrow[p \rightarrow +\infty]{} 0$

$$\Rightarrow |\varphi_p(a)| \xrightarrow[p \rightarrow +\infty]{} 0$$

Remarque  $\delta_0 \notin L^1(-1,1)$  = par l'absurde  $\exists \varphi \in \mathcal{D}$ ,  $\varphi(0) = \int_{-1}^1 f(x) \varphi(x) dx \neq \varphi(0)$   
 On prend  $\varphi(x) = x \psi(x) \Rightarrow 0 = \int_{-1}^1 f(x) x \psi(x) dx \neq \varphi(0) = \int_{-1}^1 f(x) \psi(x) dx \Rightarrow x f(x) = 0$  pp  $\Rightarrow f = 0$  pp  $\Rightarrow \varphi(0) = 0 \forall \varphi \in \mathcal{D}(-1,1)$   
 Absurde.

2) Soit  $f \in L^2(\Omega)$ .

On pose

$$T_f: \mathcal{D}(\Omega) \rightarrow \mathbb{R}$$

$$\varphi \mapsto T_f(\varphi) = \langle T_f, \varphi \rangle = \int_{\Omega} f \varphi$$

en remarquant que  $f \in L^2$   $\Rightarrow f \varphi$  est intégrable et donc que  $\int_{\Omega} f \varphi$  a un sens.  $T_f$  est bien une distribution. En effet

$$\varphi_p \xrightarrow[p \rightarrow +\infty]{} 0 \Rightarrow |\langle T_f, \varphi_p \rangle| \leq \int_{\Omega} |f \varphi_p| \leq \|f\|_{L^2(\Omega)} \|\varphi_p\|_{L^2(\Omega)} \xrightarrow[p \rightarrow +\infty]{} 0$$

$\Rightarrow 0$  car la convergence ds  $\mathcal{D}(\Omega)$  implique la convergence ds tous les  $L^p$ .

Remarque: Considérons l'application  $L^2(\Omega) \rightarrow \mathcal{D}'(\Omega)$   
 $\downarrow$   $T_f$

Cette application est linéaire. Montrons qu'elle est aussi injective, i.e. 4,

$$\text{que si } f \in L^2(\Omega), \int_{\Omega} f \varphi = 0 \forall \varphi \in \mathcal{D}(\Omega) \Rightarrow f = 0$$

On ne peut pas utiliser  $L^2 \subset L^1$  puis le résultat sur  $L^1$  car  $\Omega$  n'est pas nécessairement borné. On raisonne par densité.

Puisque  $\mathcal{D}(\Omega)$  est dense ds  $L^2(\Omega)$ , il existe une suite  $(\varphi_p)_p \in \mathcal{D}(\Omega)$

t.q.  $\varphi_p \xrightarrow{L^2} f$ . On a donc

$$\int_{\Omega} f \varphi = 0 \quad \forall \varphi \in \mathcal{D}(\Omega)$$

$$\Rightarrow \int_{\Omega} f \varphi_p = 0 \quad \forall p.$$

$$\text{Or } \int f^2 = \int f(f - \varphi_p) + \int f \varphi_p = \int f(f - \varphi_p) \leq \|f\|_{L^2} \|f - \varphi_p\|_{L^2} \rightarrow 0$$

On en déduit donc que  $f = 0$  p.p.

Cette remarque permet d'identifier  $L^2(\Omega)$  à un sous-ensemble de  $\mathcal{D}'(\Omega)$  en identifiant toute fonction  $f$  de  $L^2(\Omega)$  à la distribution

$T_f$ . On écrit  $L^2(\Omega) \subset \mathcal{D}'(\Omega)$ . d'injection est dense & continue.

3) Soit  $f \in L^1_{loc}(\Omega)$ .

La démarche précédente se généralise sans difficulté. En effet,

$$|\langle T_f, \varphi \rangle| \leq \int_{\Omega} |f \varphi| \leq \int_K |f| \times \sup_K |\varphi| \quad \forall \varphi \in \mathcal{D}(\Omega) \text{ de support } K$$

de sorte que l'application  $T_f$  est bien définie.

Il s'agit bien d'une distribution, toujours d'après l'inégalité précédente et en utilisant le compact  $K$  (fixe) associé à la

définition  $\varphi_p \xrightarrow[p \rightarrow +\infty]{\mathcal{D}'(\Omega)} 0$

Grâce au deuxième lemme de la page 2), l'application

$$\begin{aligned} L^1_{loc}(\Omega) &\rightarrow \mathcal{D}'(\Omega) \\ f &\rightarrow T_f \end{aligned}$$

est injective de sorte  $L^1_{loc}(\Omega)$  peut aussi être vu comme un

sous-espace de  $\mathcal{D}'(\Omega)$ .

## Définition (dérivation d'une distribution)

5/

Soit  $T$  une distribution sur  $\Omega$  et  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$  un multi-indice de dérivation. On appelle dérivée  $\alpha$ -ième de  $T$  la distribution notée  $D^\alpha T$  et définie par

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

avec  $|\alpha| = \alpha_1 + \dots + \alpha_n$ .

Vérifions que  $D^\alpha T$  est bien une distribution. Soit  $(\varphi_p)_p \in \mathcal{D}(\Omega)$  une suite telle que  $\varphi_p \xrightarrow[p \rightarrow \infty]{} 0$ . On a par définition.

$$|\langle D^\alpha T, \varphi_p \rangle| = |\langle T, D^\alpha \varphi_p \rangle|$$

Or  $\varphi_p \xrightarrow[p \rightarrow \infty]{} 0$  implique forcément que  $D^\alpha \varphi_p \xrightarrow[p \rightarrow \infty]{} 0$ , de sorte que  $T$

étant une distribution, on a  $|\langle D^\alpha T, \varphi_p \rangle| = |\langle T, D^\alpha \varphi_p \rangle| \xrightarrow[p \rightarrow \infty]{} 0$ , i.e.

$D^\alpha T$  est bien une distribution.

Le lemme suivant montre que cette nouvelle notion de dérivation n'est qu'une généralisation de la notion de dérivation usuelle pour les fonctions.

Lemme Soit  $f$  une fonction de classe  $\mathcal{C}^1(\Omega)$ . Alors, pour tout  $i=1, \dots, n$ , la dérivée partielle  $\frac{\partial f}{\partial x_i}$  de  $f$  au sens usuel coïncide avec la dérivée  $= (0, \dots, 0, \underset{i\text{ème place}}{1}, 0, \dots, 0)$ -ième de  $f$  au sens des distributions.

## Démonstration

•  $f \in \mathcal{C}^1(\Omega)$  donc  $f \in L^1_{loc}$  donc  $f$  est bien une distribution.

•  $\frac{\partial f}{\partial x_i} \in \mathcal{C}^0(\Omega)$  donc  $\frac{\partial f}{\partial x_i} \in L^1_{loc}$  donc  $\frac{\partial f}{\partial x_i}$  est bien une distribution.

Soit  $\varphi \in \mathcal{D}(\Omega)$ , et  $\alpha = (0, \dots, 0, \underset{i\text{ème place}}{1}, 0, \dots, 0)$

$$\langle D^\alpha f, \varphi \rangle = - \langle f, D^\alpha \varphi \rangle = - \langle f, \frac{\partial \varphi}{\partial x_i} \rangle = - \int_{\Omega} f \frac{\partial \varphi}{\partial x_i}$$

Soit  $K \subset \Omega$  le support de  $\frac{\partial \varphi}{\partial x_i}$ . Puisque  $f$  et  $\frac{\partial \varphi}{\partial x_i}$  sont de classe  $\mathcal{C}^1(K=K)$ , la formule de Green nous dit que  $-\int_K f \frac{\partial \varphi}{\partial x_i} = \int_K \frac{\partial f}{\partial x_i} \varphi + \int_{\partial K} f \varphi n_i = \int_{\Omega} \frac{\partial f}{\partial x_i} \varphi$ .

de sorte que  $\langle D^\alpha f, \varphi \rangle = \int_{\Omega} \frac{\partial f}{\partial x_i} \varphi \quad \forall \varphi \in \mathcal{D}(\Omega)$ .

Remarquons tout de suite néanmoins (voir exemple ci dessous) que si  $f$  n'est dérivable que presque partout alors cette dérivée ne coïncide pas nécessairement avec sa dérivée au sens des distributions.

Exemple soit  $f \in \mathcal{C}^1(\mathbb{R} - \{a\})$ , ni forcément continue, ni forcément définie en  $a$ , mais admettant une dérivée  $f'$  continue et bornée sur  $\mathbb{R} \setminus \{a\}$ .  
Aux voisinages à droite ou à gauche de  $a$ , le TAF dit que  $|f(x) - f(x')| \leq C|x - x'|$  par hyp. théor de sorte que  $f$  admet des limites  $f(a^+)$  et  $f(a^-)$  à droite et à gauche. Il est clair que  $f \in L^1_{loc}$  et est une distribution. Calculons sa dérivée au sens des distributions:

$$\begin{aligned} \text{Soit } \varphi \in \mathcal{D}(\Omega): \langle f', \varphi \rangle &= - \langle f, \varphi' \rangle = - \int_{\mathbb{R}} f \varphi' \\ &= - \int_{-\infty}^a f \varphi' - \int_a^{+\infty} f \varphi' \\ \Omega = \mathbb{R} & \\ &= \int_{\mathbb{R}} f' \varphi - \varphi(a) f(a^-) + \varphi(a) f(a^+) \\ &= \int_{\mathbb{R}} f' \varphi + [f(a^+) - f(a^-)] \varphi(a) \end{aligned}$$

On a donc  $f' = \{f'\} + [f(a^+) - f(a^-)] \delta_a$   
ou  $f'$  désigne la dérivée au sens des distributions et  $\{f'\}$  désigne la dérivée usuelle.

Les deux dérivées coïncident si  $f(a^-) = f(a^+)$  ( $f$  continue en  $a$ ).

Lorsque  $f(x) \equiv H(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$  est la fonction de Heaviside, on a

implément  $H' = \delta_0$ .

Définition (convergence des distributions)

Soit  $\{T_p\}_p$  et  $T$  des distributions sur  $\Omega$ , ie  $T_p \in \mathcal{D}'(\Omega) \forall p \in \mathbb{N}$   
 $T \in \mathcal{D}'(\Omega)$

On dit que  $T_p \xrightarrow{p \rightarrow +\infty} T$  si  $\forall \varphi \in \mathcal{D}(\Omega), \langle T_p, \varphi \rangle \rightarrow \langle T, \varphi \rangle$ .

Les deux résultats suivants nous seront utiles par la suite.

Lemme La dérivation des distributions est une application linéaire continue de  $\mathcal{D}'(\Omega)$ .

Démonstration

Soit  $\alpha$  un multi-indice et  $D^\alpha$  l'application définie par

$$D^\alpha: \mathcal{D}'(\Omega) \rightarrow \mathcal{D}'(\Omega) \\ T \rightarrow D^\alpha T$$

Montrons que cette application est continue. Soit  $T_p \xrightarrow[p \rightarrow +\infty]{\mathcal{D}'(\Omega)} T$ , ie  $\langle T_p, \varphi \rangle \rightarrow \langle T, \varphi \rangle \forall \varphi \in \mathcal{D}(\Omega)$ . Il faut mg  $D^\alpha T_p \xrightarrow[p \rightarrow +\infty]{\mathcal{D}'(\Omega)} D^\alpha T$ , ie

$$\langle D^\alpha T_p, \varphi \rangle \rightarrow \langle D^\alpha T, \varphi \rangle \forall \varphi \in \mathcal{D}(\Omega). \text{ Or}$$

$$\langle D^\alpha T_p, \varphi \rangle = (-1)^{|\alpha|} \langle T_p, D^\alpha \varphi \rangle \xrightarrow[p \rightarrow +\infty]{} (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle = \langle D^\alpha T, \varphi \rangle.$$

Lemme (Continuité de l'application  $L^2(\Omega) \rightarrow \mathcal{D}'(\Omega)$ )  
 $f \rightarrow T_f$

Soit  $f_p \xrightarrow[p \rightarrow +\infty]{L^2(\Omega)} f$ . Alors  $T_{f_p} \xrightarrow[p \rightarrow +\infty]{\mathcal{D}'(\Omega)} T_f$  et pour tout indice  $\alpha \in \mathbb{N}^n$

$$D^\alpha T_{f_p} \xrightarrow[p \rightarrow +\infty]{\mathcal{D}'(\Omega)} D^\alpha T_f.$$

Démonstration

Est  $\varphi \in \mathcal{D}(\Omega)$

$$|\langle T_{f_p} - T_f, \varphi \rangle| = \left| \int_{\Omega} (f_p - f) \varphi \right| \leq \underbrace{\|f_p - f\|_{L^2}}_{\xrightarrow[p \rightarrow +\infty]{} 0} \|\varphi\|_{L^2} \xrightarrow[p \rightarrow +\infty]{} 0$$

$$|\langle D^\alpha T_{f_p} - D^\alpha T_f, \varphi \rangle| = |(-1)^{|\alpha|} \langle T_{f_p} - T_f, D^\alpha \varphi \rangle| \leq \|f_p - f\| \|D^\alpha \varphi\|_{L^2} \xrightarrow[p \rightarrow +\infty]{} 0.$$

Pour terminer ce paragraphe de rappels sur les distributions par quelques propriétés utiles.

Lemme

Soit  $T \in \mathcal{D}'(\Omega)$  et  $f \in \mathcal{C}^\infty(\Omega)$ . Le produit  $fT$  défini par l'application  $fT: \mathcal{D}(\Omega) \rightarrow \mathbb{R}$

$$\varphi \rightarrow \langle fT, \varphi \rangle := \langle T, f\varphi \rangle$$

est une distribution

Démonstration

Tout d'abord, le produit  $f\varphi$  appartient à  $\mathcal{D}(\Omega)$  donc la définition a un sens. Montrons que  $fT$  est une distribution. Soit  $\varphi_p \xrightarrow{p \rightarrow \infty} 0$

$$\langle fT, \varphi_p \rangle = \langle T, f\varphi_p \rangle \quad \text{Il suffit donc de se convaincre, puisque}$$

$T$  est une distribution, que  $f\varphi_p \xrightarrow{p \rightarrow \infty} 0$

Soit  $K$  un compact fixe t.q.  $\text{supp } \varphi_p \subset K \forall p$ .

On a alors  $\text{supp } f\varphi_p \subset K \forall p$ . Par ailleurs, il est clair

$$\sup_{x \in K} |D^\alpha f(x) \varphi_p(x)| \xrightarrow{p \rightarrow \infty} 0 \quad \text{En effet et par exemple, si } \alpha = (1, 0, \dots, 0),$$

$$\begin{aligned} |D^\alpha f \varphi_p| &= |f D^\alpha \varphi_p + \varphi_p D^\alpha f| \\ &\leq \|f\|_{L^\infty(K)} |D^\alpha \varphi_p(x)| + \|D^\alpha f\|_{L^\infty(K)} |\varphi_p(x)| \end{aligned}$$

En prenant le sup en  $x$ , on a bien la convergence vers 0 d'où la conclusion.

Lemme

Soit  $T \in \mathcal{D}'(\Omega)$ , t.q.  $\frac{\partial T}{\partial x_i} = 0 \quad \forall i=1, \dots, n$ . Alors  $T$  est constante sur chacune des composantes connexes de  $\Omega$ .

Démonstration

On ne démontrera ce résultat que dans le cas  $n=1$  et  $\Omega = ]a, b[$ .



Montrons tout d'abord que si  $\varphi$  et  $\psi$  sont de  $\mathcal{D}(]a, b[)$  alors 9/

$$\varphi' = \psi \Leftrightarrow \int_a^b \psi(x) dx = 0$$

( $\Rightarrow$ ) évident en intégrant.

( $\Leftarrow$ ) On veut donc montrer qu'une fonction de  $\mathcal{D}(]a, b[)$  et d'intégrale nulle est la dérivée d'une fonction de  $\mathcal{D}(]a, b[)$ .

Posez  $\varphi(t) = \int_a^t \psi(x) dx$ . Il est clair que  $\varphi'(t) = \psi(t)$ .

A-t-on  $\varphi \in \mathcal{D}(]a, b[)$ ?

$\varphi \in \mathcal{D}(]a, b[)$  donc  $\exists \eta > 0$ ,  $\varphi' = 0$  sur  $]a, a + \eta[ \cup ]b - \eta, b[$

$\Rightarrow \varphi = c_1 t$  sur  $]a, a + \eta[$   
et  $\varphi = c_2 t$  sur  $]b - \eta, b[$ .

Comme  $\varphi(a) = \varphi(b) = 0$  on a  $\varphi = 0$  sur  $]a, a + \eta[ \cup ]b - \eta, b[$ .

Supposons donc maintenant que

$$\forall \varphi \in \mathcal{D}(\Omega) \quad \langle T, \varphi' \rangle = 0.$$

On veut mg  $\exists k \in \mathbb{R}$  t.g.  $\langle T, \varphi \rangle = \int_a^b k \varphi(x) dx = k \int_a^b \varphi(x) dx$ .

$$\forall \varphi \in \mathcal{D}(]a, b[)$$

Soit  $\varphi \in \mathcal{D}(]a, b[)$ . Soit  $\theta \in \mathcal{D}(]a, b[)$  t.g.  $\int_a^b \theta = 1$ .

On décompose  $\varphi$  sous la forme  $\varphi = c \theta(x) + \tilde{\varphi}(x)$

$$\text{avec } \begin{cases} \tilde{\varphi} \in \mathcal{D}(]a, b[), & \int_a^b \tilde{\varphi}(x) dx = 0 \\ c \in \mathbb{R} \end{cases}$$

(Il suffit de choisir  $c = \int_a^b \varphi(x) dx$  et  $\tilde{\varphi} = \varphi - c \theta$ )

On a alors l'existence de  $\tilde{\varphi} \in \mathcal{D}(]a, b[)$  t.g.  $\tilde{\varphi}' = \tilde{\varphi}$  et par suite

$$\begin{aligned} \langle T, \varphi \rangle &= c \langle T, \theta \rangle + \langle T, \tilde{\varphi} \rangle \\ &= c \langle T, \theta \rangle + \underbrace{\langle T, \tilde{\varphi}' \rangle}_{=0} \\ &= \langle T, \theta \rangle \int_a^b \varphi(x) dx \\ &= \int_a^b \langle T, \theta \rangle \varphi(x) dx \end{aligned}$$

On a donc  $k = \langle T, \theta \rangle$ .

Remarque Si on choisit  $\theta_1$  et  $\theta_2$  de  $\mathcal{D}(\mathbb{I}_a, b[) \text{ t.g.}$  10/

$$\int_a^b \theta_1 = \int_a^b \theta_2 = 1, \text{ alors } \langle T, \theta_1 \rangle = \langle T, \theta_2 \rangle.$$

En effet,  $\int_a^b (\theta_1 - \theta_2) = 0$  de sorte que

$\exists \varphi \text{ t.g. } \theta_1 - \theta_2 = \varphi'$  et donc

$$\begin{aligned} \langle T, \theta_1 \rangle - \langle T, \theta_2 \rangle &= \langle T, \theta_1 - \theta_2 \rangle = \langle T, \varphi' \rangle \\ &= 0 \text{ par hypothèse.} \end{aligned}$$

Il existe beaucoup d'autres résultats sur les distributions mais nous nous limitons ici à ceux qui nous seront utiles par la suite.

## Les espaces de Sobolev

Définition On note  $H^1(\Omega)$  l'ensemble défini par

$$H^1(\Omega) = \left\{ v \in L^2(\Omega), \frac{\partial v}{\partial x_i} \in L^2(\Omega) \forall i=1, \dots, n \right\}$$

où les dérivées sont comprises au sens des distributions

On munit cet espace du produit scalaire suivant :

$$(u, v)_{H^1} = (u, v)_{L^2} + \sum_{i=1}^n \left( \frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i} \right)_{L^2},$$

où  $(\cdot, \cdot)$  est le produit scalaire dans  $L^2$ , ie  $(u, v) = \int_{\Omega} uv$ , et de la norme

correspondante

$$\|v\|_{H^1} = \sqrt{(v, v)_{H^1}} = \sqrt{\|v\|_{L^2}^2 + \sum_{i=1}^n \left\| \frac{\partial v}{\partial x_i} \right\|_{L^2}^2}$$

On a alors le résultat suivant

Théorème : L'espace  $H^1$  est un espace de Hilbert, ie il est complet pour la norme issue du produit scalaire.

Démonstration Il suffit de montrer que l'espace  $H^1$  est complet, ie que

toute suite de Cauchy dans  $H^1$  converge dans  $H^1$

Soit  $(v_n)_n$  une suite de Cauchy de  $H^1$ . Par définition de la norme sur  $H^1$ , elle est de Cauchy de  $L^2$  dont on sait qu'il est un espace de Hilbert. Il existe donc  $v \in L^2$  t.q.  $v_n \xrightarrow{L^2} v$ .

En raisonnant de la même façon sur chacun des  $\frac{\partial v_n}{\partial x_i}$ , il existe des fonctions  $w_i \in L^2$  t.q.  $\frac{\partial v_n}{\partial x_i} \xrightarrow{L^2} w_i$ .

Il reste à se convaincre que  $\frac{\partial v}{\partial x_i} = w_i$ .

On a vu ds le 2<sup>e</sup> lemme page 7 que

$$\text{d'une part, } v_n \xrightarrow{L^2} v \Rightarrow \frac{\partial v_n}{\partial x_i} \xrightarrow{\mathcal{D}'} \frac{\partial v}{\partial x_i}$$

$$\text{et d'autre part, } \frac{\partial v_n}{\partial x_i} \xrightarrow{L^2} w_i \Rightarrow \frac{\partial v_n}{\partial x_i} \xrightarrow{\mathcal{D}'} w_i$$

Par unicité de la limite de  $\mathcal{D}'$ , on a donc  $\frac{\partial v}{\partial x_i} = w_i$  et donc

$$\frac{\partial v}{\partial x_i} \in L^2, \text{ de sorte que } v \in H^1.$$

Puisque  $v_n \xrightarrow{L^2} v$  et  $\frac{\partial v_n}{\partial x_i} \xrightarrow{L^2} \frac{\partial v}{\partial x_i}$ , on en déduit également que  $v_n \xrightarrow{H^1} v$  ce qui termine la démonstration.

### Remarques

- 1) Si  $\Omega$  est borné, alors  $\mathcal{C}^1(\bar{\Omega}) \subset H^1(\Omega)$
- 2)  $H^1(\Omega) \subset L^2(\Omega)$  mais  $L^2(\Omega) \not\subset H^1(\Omega)$ . Pour s'en convaincre, il suffit de prendre la fonction de Heaviside définie sur  $]-1, 1[$  et de se rappeler que  $H^1 = \delta_0 \notin L^2(]-1, 1[)$  (voir remarque page 3).
- 3)  $\mathcal{D}(\Omega) \subset H^1(\Omega)$  mais  $\mathcal{D}(\Omega)$  n'est pas dense ds  $H^1(\Omega)$ . On verra ci-après que  $\overline{\mathcal{D}(\Omega)} = H_0^1(\Omega) \subsetneq H^1(\Omega)$

( $m=0 \Rightarrow L^2$ )

Définition Soit  $m > 0$ . On note  $H^m(\Omega)$  l'ensemble défini par  $H^m(\Omega) = \{ v \in L^2(\Omega), D^\alpha v \in L^2(\Omega) \forall \alpha = (\alpha_1, \dots, \alpha_n), |\alpha| \leq m \}$

Cet espace est un espace de Hilbert pour le produit scalaire naturel

$$(u, v)_m = \sum_{|\alpha| \leq m} (D^\alpha u, D^\alpha v)_{L^2}$$

et la norme associée  $\|v\|_{H^m} = \sqrt{\sum_{|\alpha| \leq m} \|D^\alpha v\|_{L^2}^2}$  12/

On note  $|v|_{H^m} = \sqrt{\sum_{|\alpha|=m} \|D^\alpha v\|_{L^2}^2}$  Remarque: Ce n'est pas une norme mais une semi-norme propre

de sorte que  $\|v\|_{H^1} = \sqrt{\|v\|_{L^2}^2 + |v|_{H^1}^2}$

et  $\|v\|_{H^m} = \sqrt{\sum_{|\alpha| \leq m} \|D^\alpha v\|_{L^2}^2 + |v|_{H^m}^2}$

Rappelons enfin qu'il est possible de généraliser ces définitions dans un cadre autre que celui des espaces de Hilbert

Définition Soit  $m > 0$ ,  $p \geq 1$ . On note  $W^{m,p}(\Omega)$  l'ensemble défini par

$$W^{m,p}(\Omega) = \left\{ v \in L^p(\Omega), D^\alpha v \in L^p(\Omega), \forall \alpha = (\alpha_1, \dots, \alpha_n), |\alpha| \leq m \right\}$$

C'est un espace de Banach pour la norme  $1/p$ .

$$\|v\|_{W^{m,p}} = \left( \sum_{|\alpha| \leq m} \|D^\alpha v\|_{L^p}^p \right)^{1/p}$$

Définition ( $H_0^1(\Omega)$ )

On note  $H_0^1(\Omega)$  l'adhérence de  $\mathcal{D}(\Omega)$  dans  $H^1(\Omega)$ , i.e.

$$H_0^1(\Omega) := \overline{\mathcal{D}(\Omega)}^{H^1} = \left\{ v \in H^1, \exists (v_n)_n \in \mathcal{D}(\Omega), \|v_n - v\|_{H^1} \xrightarrow{n \rightarrow +\infty} 0 \right\}$$

Cet espace est par définition un sous-ensemble fermé de  $H^1(\Omega)$ , il est donc complet et  $H_0^1(\Omega)$  est donc un espace de Hilbert pour le produit scalaire  $(\cdot, \cdot)_{H^1}$  (et la norme associée).

Dans le cas d'un ouvert borné, le résultat suivant nous dit que  $H_0^1(\Omega)$  est un espace de Hilbert pour un produit scalaire (et donc une norme) plus simple que  $(\cdot, \cdot)_{H^1}$  ( $\|\cdot\|_{H^1}$ ): il s'agit du produit scalaire dont la norme associée est la semi-norme

$$|v|_{H^1} \text{ sur } H_0^1 : \|v\|_{H_0^1} = |v|_{H^1}, \text{ i.e.}$$

$$(u, v)_{H_0^1} = \sum_{|\alpha|=1} (D^\alpha u, D^\alpha v)_{L^2}$$

# Théorème (Inégalité de Poincaré)

13/

Soit  $\Omega$  un ouvert borné de  $\mathbb{R}^n$  (ou au moins borné dans une direction de l'espace). Alors il existe une constante  $C(\Omega) > 0$  telle que

$$\|v\|_{L^2} \leq C |v|_{H^1} \quad \forall v \in H_0^1(\Omega).$$

## Démonstration

Supposons tout d'abord l'inégalité vraie pour des fonctions de  $\mathcal{D}(\Omega)$ .

Alors  $\forall v \in H_0^1(\Omega)$ , posons  $(v_n)_n \in \mathcal{D}(\Omega)$  t.g.  $\|v_n - v\|_{H^1} \rightarrow 0$ .

On a par hypothèse  $\|v_n\|_{L^2} \leq C |v_n|_{H^1} \quad \forall n$ , ce qui donne par passage

à la limite  $\|v\|_{L^2} \leq C |v|_{H^1}$ . En effet, si on se focalise sur le membre

de gauche par exemple :  $\|v_n - v\|_{H^1} \rightarrow 0 \Rightarrow \|v_n - v\|_{L^2} \rightarrow 0$ . Mais

$$\|v\| - \|v_n - v\| \leq \|v_n\|_{L^2} \leq \|v_n - v\| + \|v\|$$
$$\downarrow \qquad \qquad \qquad \downarrow$$
$$\|v\| \qquad \qquad \qquad \|v\|.$$

On se focalise donc sur les fonctions  $v \in \mathcal{D}(\Omega)$ .

$$\|v\|_{L^2}^2 = \int_{\Omega} v^2(x) dx = \iiint_{x_1 \dots x_n} v^2(x_1, \dots, x_n) dx_1 \dots dx_n$$

On note  $\tilde{v}$  le prolongement par 0 de  $v$  en dehors de  $\Omega$ .  
de sorte que  $\tilde{v} \in \mathcal{D}(\mathbb{R}^n)$ . Supposons que  $\Omega$  soit borné ds la direction  $x_n$ .

$$\Rightarrow \|v\|_{L^2}^2 = \iiint_{\mathbb{R}^n} \tilde{v}^2(x_1, \dots, x_n) dx_1 \dots dx_n = \int_a^b \int_{\mathbb{R}^{n-1}} \tilde{v}^2(x_1, \dots, x_n) dx_1 \dots dx_{n-1} dx_n$$
$$= \int_a^b \left( \int_{\mathbb{R}^{n-1}} \tilde{v}^2(x', x_n) dx' \right) dx_n.$$

On a  $\tilde{v}(x', x_n) = \int_a^{x_n} \frac{\partial \tilde{v}}{\partial x_n}(x', t) dt$  (puisque  $\tilde{v}(x', a) = 0$ ),  $\forall a \leq x_n \leq b$ .

D'après l'inégalité de C.S.,  $| \tilde{v}(x', x_n) |^2 \leq (x_n - a) \int_a^{x_n} \left| \frac{\partial \tilde{v}}{\partial x_n}(x', t) \right|^2 dt \leq (b-a) \int_a^b \left| \frac{\partial \tilde{v}}{\partial x_n}(x', t) \right|^2 dt$ .

En intégrant relativement à la variable  $x'$  de  $\mathbb{R}^{n-1}$ , il vient donc

$$\int_{\mathbb{R}^{n-1}} | \tilde{v}(x', x_n) |^2 dx' \leq (b-a) \left\| \frac{\partial \tilde{v}}{\partial x_n} \right\|_{L^2(\mathbb{R}^{n-1})}^2$$

On intègre ensuite en  $x_n$ ,

$$\|\tilde{v}\|_{L^2(\mathbb{R}^n)}^2 \leq (b-a)^2 \left\| \frac{\partial \tilde{v}}{\partial x_n} \right\|_{L^2(\mathbb{R}^n)}^2$$

et donc

$$\|\tilde{v}\|_{L^2(\Omega)}^2 \leq (b-a)^2 \left\| \frac{\partial \tilde{v}}{\partial x_n} \right\|_{L^2(\Omega)}^2 \leq (b-a)^2 |v|_{H^1(\Omega)}^2$$

La constante cherchée vaut donc  $(b-a)$ .

### Remarque

1) Si  $\Omega$  est borné, en choisissant  $v = cte$ , on voit que l'inégalité de Poincaré est fautive dans  $H^1$  car  $\|v\|_{L^2} \neq 0$  et  $|v|_{H^1} = 0$  ( $v = cte \in H^1$  si  $\Omega$  borné).

2) L'exemple considéré dans 1) montre au passage que  $H_0^1 \not\subset H^1$  en général (si  $\Omega$  est borné notamment) puisque  $v = cte \in H^1$  et  $v = cte > 0 \notin H_0^1$ .

↳ votre résultat de dernière ci-dessus  $H_0^1 = H^1$  pour  $\Omega = \mathbb{R}^n$ ?

### Corollaire

Soit  $\Omega \subset \mathbb{R}^n$  un ouvert borné (au moins dans une direction). Alors la semi-norme  $|\cdot|_{H^1}$  est une norme sur  $H_0^1$  équivalente à la norme induite par celle de  $H^1$ . L'espace  $H_0^1$  est alors un espace de Hilbert pour le produit scalaire "réduit" :

$$(u, v)_{H_0^1} = (Du, Dv)_{(L^2(\Omega))^n} = \sum_i \left( \frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i} \right)_{L^2}$$

### Démonstration

Il est clair tout d'abord que  $\|v\|_{H^1} \geq |v|_{H^1}$

Ensuite, si  $v \in H_0^1$ ,

$$\|v\|_{H^1}^2 = \|v\|_{L^2}^2 + |v|_{H^1}^2 \leq (1+c) |v|_{H^1}^2$$

de sorte que  $|v|_{H^1} \leq \|v\|_{H^1} \leq \sqrt{1+c} |v|_{H^1} \quad \forall v \in H_0^1$ .

Proposition (généralisation)

On note  $H_0^m(\Omega)$  l'adhérence de  $\mathcal{D}(\Omega)$  dans  $H^m(\Omega)$ . Sur cet espace, la semi-norme  $|\cdot|_{H^m}$  est une norme équivalente à la norme induite par celle de  $H^m$  et  $H_0^m$  est un espace de Hilbert pour cette norme.

On admettra les lemmes de densité suivants.

Lemme

$$\overline{\mathcal{D}(\mathbb{R}^n)}^{H^1} = \underset{\substack{\uparrow \\ \text{définition}}}{H_0^1(\mathbb{R}^n)} = \underset{\substack{\uparrow \\ \text{résultat du lemme}}}{H^1(\mathbb{R}^n)}$$

et plus généralement

$$\overline{\mathcal{D}(\mathbb{R}^n)}^{H^m} = \underset{\substack{\uparrow \\ \text{déb}}}{H_0^m(\mathbb{R}^n)} = \underset{\substack{\uparrow \\ \text{résultat}}}{H^m(\mathbb{R}^n)}$$

Lemme On suppose que  $\Omega$  est un ouvert de  $\mathbb{R}^n$  de frontière lipschitzienne

Ainsi  $\overline{\mathcal{D}(\Omega)} = H^m(\Omega)$

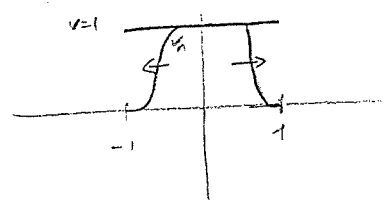
Petite explication

soit  $\Omega = ]-1, 1[$ . On a

$$H_0^1 := \overline{\mathcal{D}(\Omega)} \neq H^1$$

a) en effet, en prenant  $v=1$ , on a  $v \in H^1$  mais  $v \notin H_0^1$  car l'inégalité de Poincaré. Donc  $v$  ne peut pas être la limite de  $\mathcal{D}(\Omega)$  d'une suite de  $\mathcal{D}(\Omega)$ . Par contre,  $v$  peut être la limite de  $L^2$  d'une suite de  $\mathcal{D}(\Omega)$  par densité.

Bouger



en décalant petit à petit, on voit que ça ne va pas passer de  $L^2$  par contre à cause du passage brusque de 0 à 1, les dérivées vont devenir  $\infty$  et ne pas tendre vers 0.

b) Par contre, si on autorise des fonctions de  $\mathcal{D}(\bar{\Omega})$  ( $v$  peut être  $\neq 0$  sur le bord du supp) alors ça change tout car  $v \in \mathcal{D}(\bar{\Omega})$  signifie  $\exists \epsilon, v \neq 0$  sur une bande de  $[-1, 1]$ .  
 on peut prendre  $[-1, 1]$  comme support.  
 si  $v=1$  est déjà de  $\mathcal{D}(\bar{\Omega})$  !!!

Remarque (Généralisation à  $H^1$  de l'inégalité de Poincaré)

On peut montrer, ds le cas d'un ouvert borné convexe de  $\mathbb{R}^n$ , une inégalité de type Poincaré valable sur  $H^1$  mais le prix à payer est la présence d'un terme supplémentaire dans le membre de droite de l'inégalité. Prenons par exemple le cas  $\Omega = ]-1, 1[$ . On part de  $\int_{\Omega} |u|^2 dx$  et on veut faire apparaître l'intégrale d'une dérivée de  $u$ . On sait que  $\overline{D(\Omega)} = H^1$  donc on va raisonner par densité et prendre  $u \in \mathcal{D}(\Omega)$ .

On a donc

$$u(x) = u(y) + \int_y^x u'(t) dt$$

On intègre en  $y$ :

$$2u(x) = \int_{-1}^1 u(y) dy + \int_{-1}^1 \left( \int_y^x u'(t) dt \right) dy$$

Par Cauchy-Schwarz,

$$2|u(x)| \leq \left| \int_{-1}^1 u(y) dy \right| + 2 \int_{-1}^1 |u'(t)| dt$$

$$2|u(x)| \leq \left| \int_{-1}^1 u(y) dy \right| + 2\sqrt{x+1} \sqrt{\int_{-1}^1 |u'|^2}$$

$$\Rightarrow |u(x)| \leq \frac{1}{2} \left| \int_{-1}^1 u(y) dy \right| + \sqrt{x+1} \sqrt{\int_{-1}^1 |u'|^2}$$

$$|u(x)|^2 \leq \frac{1}{2} \left( \int_{-1}^1 u(y) dy \right)^2 + 2(x+1) \int_{-1}^1 |u'|^2$$

$$\int_{-1}^1 \sqrt{x+1} = \left[ \frac{2}{3} (x+1)^{3/2} \right]_{-1}^1$$

$$(a+b)^2 \leq 2(a^2+b^2)$$

En intégrant en  $x$ ,  $\|u\|_{L^2}^2 \leq \left( \int_{-1}^1 u(y) dy \right)^2 + 4 \|u'\|_{L^2}^2$

On a donc, en posant  $C=4$ :

$$\|u\|_{L^2}^2 \leq C \left[ \left( \int_{-1}^1 u(y) dy \right)^2 + \|u'\|_{L^2}^2 \right]$$

↑  
terme supplémentaire.

On raisonne ensuite par densité et on a

$$\forall u \in H^1, \|u\|_{L^2}^2 \leq C \left\{ \left( \int_{-1}^1 u(y) dy \right)^2 + \|u'\|_{L^2}^2 \right\}$$

Que peut-on dire de ce terme supplémentaire ?

Posons  $V = \left\{ u \in H^1, \int_{-1}^1 u(y) dy = 0 \right\}$ . On peut munir ce sous-ensemble de  $H^1$  de la  $\tilde{m}$  semi-norme que  $H^1$ .  $(V, \tilde{m})$  est fermé  $\tilde{c}$  un espace réciprocité d'un fermé  $\{0\}$  par une application linéaire continue  $\left( \begin{matrix} H^1 \rightarrow \mathbb{R} \\ v \mapsto \int_{-1}^1 v \end{matrix} \right)$ .  $V$  est fermé  $\subset$  ds le Hilbert  $H^1$  pour la norme sur  $H^1$   $\Rightarrow V$  Hilbert pour la norme sur  $H^1$ .  $V$  est aussi un Hilbert pour la semi-norme car elle est équivalente à la norme sur  $H^1$ .



# Notion de trace sur $\Gamma = \partial\Omega$

On va alors généraliser dans ce paragraphe la notion de trace sur  $\Gamma = \partial\Omega$  qui est naturellement définie pour une fonction  $f \in \mathcal{C}^0(\bar{\Omega})$ . On va d'abord considérer le cas 1D ( $\Omega \subset \mathbb{R}$ ), puis le cas du demi-espace, et enfin nous énoncerons un cas général d'un ouvert  $\Omega$  à frontière lipschitzienne.

## Le cas 1D

Théorème Si  $v \in H^1([a, b])$  avec  $a < b$ , alors  $v \in \mathcal{C}^0([a, b])$ . Par ailleurs l'application  $H^1([a, b]) \rightarrow \mathcal{C}^0([a, b])$  est continue, ie  $\exists C > 0$ ,

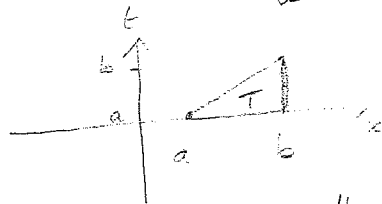
$$\forall v \in H^1([a, b]), \quad \sup_{x \in [a, b]} |v(x)| \leq C \|v\|_{H^1}.$$

(ie il existe un représentant continu sur  $[a, b]$  de  $v$ , encore noté  $v$ )

## Démonstration

Soit  $v \in H^1([a, b])$ . Posons  $\tilde{v}(x) = \int_a^x v'(t) dt$ . Cette définition a bien un sens et définit une fonction continue presque partout.  $|\tilde{v}(x) - \tilde{v}(x')| = \left| \int_{x'}^x v'(t) dt \right| \leq \|v'\|_{L^2(a, b)} \sqrt{|x-x'|} \xrightarrow{x \rightarrow x'} 0$ .  
 Il faudrait montrer que  $v$  et  $\tilde{v}$  diffèrent d'une constante près de  $L^2$  (car  $v$  admettrait bien un représentant continu, à savoir  $\tilde{v} + \text{constante}$ ), ie que  $v' = \tilde{v}'$  de  $\mathcal{D}'(\mathbb{R})$  d'après le lemme p. 6. Soit  $\varphi \in \mathcal{D}([a, b])$

$$\langle v', \varphi \rangle = -\langle \tilde{v}, \varphi' \rangle = -\int_a^b \left( \int_a^x v'(t) dt \right) \varphi'(x) dx.$$



$\mathcal{L}^1$  application  $(x, t) \mapsto v'(t) \varphi(x)$  est clairement intégrable sur  $[a, b] \times [a, b]$  car  $|v'(t) \varphi(x)| \leq C \|v'\|_{L^2} \in L^2([a, b] \times [a, b])$  donc d'après le

théorème de Fubini, on a  $\langle \tilde{v}', \varphi \rangle = -\int_a^b \left( \int_a^b \varphi(x) dx \right) v'(t) dt = \int_a^b \varphi(x) v'(x) dx$  car  $\varphi(b) = 0$ .

On a donc  $\tilde{v}' = v'$  de  $\mathcal{D}'$ .  $v$  admet donc un représentant continu, que l'on décide de toujours appeler  $v$ , qui s'écrit sous la forme

$$v = C + \int_a^x v'(t) dt.$$

On a alors  $C = v(a)$  et  $v(x) = v(a) + \int_a^x v'(t) dt \quad \forall x$ .

Donc  $v(y) = v(a) + \int_a^y v'(t) dt \quad \forall y$  et on a vite

$$v(x) = v(y) + \int_y^x v'(t) dt \quad \forall x, y$$

Ainsi, il vient  $|v(x)| \leq |v(y)| + \sqrt{b-a} \|v'\|_2$

puis par intégration en  $y$  et en utilisant C.S

$$(b-a) |v(x)| \leq \sqrt{b-a} \|v'\|_2 + \sqrt{b-a} (a,0) \|v'\|_2$$

$$\text{ie } |v(x)| \leq \frac{1}{\sqrt{b-a}} \|v'\|_2 + \sqrt{b-a} \|v'\|_2$$

En utilisant  $\epsilon = 1$  pour le produit scalaire de  $\mathbb{R}^2$ , on a donc

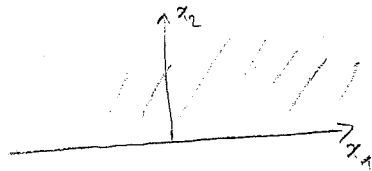
$$|v(x)| \leq \sqrt{\frac{1}{b-a} + b-a} \|v'\|_{H^1} \quad \forall x$$

ce qui conclut la démonstration du théorème

Le cas du demi-espace

On considère ici le cas  $\Omega = \mathbb{R}_+^n := \{x = (x', x_n) \in \mathbb{R}^n, x' \in \mathbb{R}^{n-1}, x_n > 0\}$

Cas n=2



Les fonctions  $v$  de  $\mathcal{D}(\overline{\mathbb{R}_+^n})$  étant régulières jusqu'au bord  $\Gamma$  de  $\Omega$ , on peut naturellement définir leur restriction  $v|_\Gamma$  sur  $\Gamma$ . Il est clair que  $v|_\Gamma := v(\cdot, 0)$  est dans  $L^2(\mathbb{R}^{n-1})$ . Nous avons plus précisément le résultat suivant.

Proposition. Soit  $v \in \mathcal{D}(\overline{\mathbb{R}_+^n})$ . On a  $v(\cdot, 0) \in L^2(\mathbb{R}^{n-1})$  et

$$\|v(\cdot, 0)\|_{L^2(\mathbb{R}^{n-1})} \leq \|v\|_{H^1(\mathbb{R}_+^n)}$$

Démonstration

Soit  $v \in \mathcal{D}(\overline{\mathbb{R}_+^n})$ .  $v$  étant très régulière et de limite nulle en  $+\infty$ , on a

$$v^2(x', 0) = - \int_0^{+\infty} 2v(x', y) \frac{\partial v}{\partial y}(x', y) dy = -2 \int_0^{+\infty} v \frac{\partial v}{\partial y}(x', y) dy \leq \int_0^{+\infty} v^2(x', y) dy + \int_0^{+\infty} \left(\frac{\partial v}{\partial y}\right)^2(x', y) dy$$

En intégrant par rapport à  $x'$ , il vient tout de suite

$$\|v(\cdot, 0)\|_{L^2(\mathbb{R}^{n-1})}^2 \leq \|v\|_{L^2(\mathbb{R}_+^n)}^2 + \left\| \frac{\partial v}{\partial y} \right\|_{L^2(\mathbb{R}_+^n)}^2 \leq \|v\|_{H^1(\mathbb{R}_+^n)}^2$$

Par densité, on a alors le résultat suivant.

Théorème Il existe une application linéaire et continue définie sur l'espace  $H^1(\mathbb{R}_+^n)$  à valeurs dans  $L^2(\mathbb{R}^{n-1})$ , notée  $\delta_0$ , telle que pour toute fonction régulière  $v \in \mathcal{D}(\mathbb{R}_+^n)$ , on ait  $\delta_0 v = v|_0$ , et  $\|\delta_0 v\|_{L^2} \leq \|v\|_{H^1}$ .  $\delta_0$  est appelée application trace. Elle généralise la notion de trace sur  $\mathcal{D}(\mathbb{R}_+^n)$ .

Démonstration

Soit  $T: \mathcal{D}(\mathbb{R}_+^n) \rightarrow L^2(\mathbb{R}^{n-1})$ . D'après le résultat précédent, cette application est linéaire et continue si on munit l'espace de départ de la norme sur  $H^1(\mathbb{R}_+^n)$ . Soit  $v \in H^1(\mathbb{R}_+^n)$ . Par densité, il existe une suite  $(v_n)_n \subset \mathcal{D}(\mathbb{R}_+^n)$  t.q.  $\|v_n - v\|_{H^1(\mathbb{R}_+^n)} \xrightarrow{n \rightarrow +\infty} 0$ . On a

$$\|Tv_n - Tv_m\|_{L^2(\mathbb{R}^{n-1})} = \|T(v_n - v_m)\|_{L^2(\mathbb{R}^{n-1})} \leq \|v_n - v_m\|_{H^1(\mathbb{R}_+^n)} \xrightarrow{n, m \rightarrow +\infty} 0$$

La suite  $(Tv_n)_n$  est donc de Cauchy dans l'espace de Hilbert  $L^2(\mathbb{R}^{n-1})$ , elle converge donc vers une limite unique notée  $\delta_0 v$ .

Il reste à démontrer deux choses : que la limite ne dépend pas de la suite considérée, et que  $\delta_0 v = v|_0$  si  $v \in \mathcal{D}(\mathbb{R}_+^n)$ .

Supposons que  $\|v_n - v\|_{H^1} \xrightarrow{n \rightarrow +\infty} 0$  et  $\|w_n - v\|_{H^1} \xrightarrow{n \rightarrow +\infty} 0$  et notons  $\delta_0^v$ , resp.  $\delta_0^w$  les limites des suites  $(Tv_n)$  et  $(Tw_n)$ . On a alors

$$\begin{aligned} \|\delta_0^v - \delta_0^w\|_{L^2} &\leq \underbrace{\|\delta_0^v - Tv_n\|}_{\xrightarrow{n \rightarrow +\infty} 0} + \underbrace{\|Tv_n - Tw_m\|}_{\xrightarrow{m \rightarrow +\infty} 0} + \underbrace{\|Tw_m - \delta_0^w\|}_{\xrightarrow{m \rightarrow +\infty} 0} \\ &\leq \|v_n - w_m\| \leq \underbrace{\|v_n - v\|}_{\xrightarrow{n \rightarrow +\infty} 0} + \underbrace{\|v - w_m\|}_{\xrightarrow{m \rightarrow +\infty} 0} \end{aligned}$$

On a donc bien égalité des limites

Si maintenant  $v \in \mathcal{D}(\mathbb{R}_+^n)$  alors  $(v_n = v)_n \xrightarrow{H^1} v$  et donc  $Tv_n = v|_0 \rightarrow \delta_0 v$  ie  $\delta_0 v = v|_0$ , de sorte que  $\delta_0$  généralise bien l'application T.

Enfin, on a  $\|Tv_n\|_{L^2} \leq \|v_n\|_{H^1}$  donc en passant à la limite

$$\|\delta_0 v\|_{L^2} \leq \|v\|_{H^1}$$

## Le cas général

20/

Plus généralement, on a le résultat suivant que l'on admettra.

Théorème On suppose que  $\Omega$  est à frontière lipschitzienne. Il existe une application linéaire et continue  $\gamma_0 : H^1(\Omega) \rightarrow L^2(\Gamma)$  telle que pour toute fonction  $v \in \mathcal{D}(\bar{\Omega})$  on ait  $\gamma_0 v = v|_{\Gamma}$ . De plus, on a

1.  $\text{Ker } \gamma_0 = H_0^1(\Omega) (= \mathcal{D}(\Omega))$

2.  $\gamma_0(H^1(\Omega)) \subsetneq L^2(\Gamma)$

3.  $\overline{\gamma_0(H^1(\Omega))} = L^2(\Gamma)$

On note  $H^{1/2}(\Gamma) := \overline{\gamma_0(H^1(\Omega))}$

### Remarques (compléments)

1) Si  $v \in H^2(\Omega)$ , on peut définir  $\gamma_0 v$  mais aussi  $\gamma_0 \left( \frac{\partial v}{\partial x_i} \right) \forall i=1, \dots, m$ .

On pose  $\gamma_1 v = \sum_{i=1}^m \nu_i \gamma_0 \left( \frac{\partial v}{\partial x_i} \right)$  où  $\nu = (\nu_1, \dots, \nu_m)$  est le vecteur normal à  $\Gamma$  orienté vers l'extérieur et de norme 1

2) Si l'ouvert  $\Omega$  est suffisamment régulier on peut montrer que l'application

$$H^2(\Omega) \rightarrow L^2(\Gamma) \times L^2(\Gamma)$$

$$v \mapsto (\gamma_0 v, \gamma_1 v)$$

est linéaire et continue et que son noyau est  $H_0^2(\Omega)$  (en particulier,

il n'est pas nécessaire que  $\gamma_0 v = 0$  et  $\gamma_0 \left( \frac{\partial v}{\partial x_i} \right) = 0 \forall i$  pour être dans  $H_0^2$ ,

il suffit que  $\gamma_0 v = 0$  et  $\gamma_1 v = 0$ , mais ceci est en fait équivalent

puisque on peut montrer que  $H_0^2 = \left\{ v \in H^2, \gamma_0 v = \gamma_0 \left( \frac{\partial v}{\partial x_i} \right) = 0 \forall i \right\}$ . Plus

généralement  $H_0^m(\Omega) = \left\{ v \in H^m(\Omega), \gamma_0(D^\alpha v) = 0 \forall \alpha, |\alpha| \leq m-1 \right\}$ .

3) En dimension 1, on a vu que  $H^1(\Omega) \subset \mathcal{C}^0(\bar{\Omega})$  et que l'application identité de  $H^1(\Omega)$  dans  $\mathcal{C}^0(\bar{\Omega})$  est continue. En dimension 2 et 3, on a  $H^2(\Omega) \subset \mathcal{C}^0(\bar{\Omega})$  et l'application identité de  $H^2(\Omega)$  dans  $\mathcal{C}^0(\bar{\Omega})$  est continue.

Plus généralement,  $H^s(\Omega) \subset C^m(\bar{\Omega})$  si  $s > m + \frac{1}{2}$  et  $\Omega \subset \mathbb{R}^n$ . 21.

Nous terminons ce chapitre par un énoncé des formules de Green.

### Proposition

1) si  $u \in H^2(\Omega)$  et  $v \in H^1(\Omega)$  alors

$$\int_{\Omega} \Delta u v = - \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Gamma} \gamma_1 u \gamma_0 v \, d\Gamma(\Omega)$$

2) si  $u \in H^1(\Omega)$  et  $v \in H^1(\Omega)$  alors

$$\int_{\Omega} \frac{\partial u}{\partial x_i} v = - \int_{\Omega} u \frac{\partial v}{\partial x_i} + \int_{\Gamma} \gamma_0 u \gamma_0 v \nu_i \, d\Gamma \quad \forall i=1, \dots, n.$$

# Formulation variationnelle de problèmes aux limites elliptiques

## Quelques problèmes modèles

- 1) Le problème modèle que nous considérerons est celui de la membrane élastique à l'équilibre et soumise à un chargement vertical dans un ouvert borné  $\Omega$  de  $\mathbb{R}^2$ . On pourra supposer que la membrane est fixe sur le bord de  $\Omega$ . L'altitude  $u(x)$  de la membrane est alors solution de l'équation elliptique suivante

$$(P_1) \quad \begin{cases} -\Delta u(x) + c(x)u(x) = f(x) & \forall x \in \Omega \\ u(x) = 0 & \forall x \in \partial\Omega, \end{cases}$$

où le laplacien  $\Delta u(x)$  est défini par

$$\begin{aligned} \Delta u(x) &= \sum_{i=1}^2 \frac{\partial^2 u}{\partial x_i^2}(x) & \forall x = (x_1, x_2) \in \mathbb{R}^2 \\ &= \frac{\partial^2 u}{\partial x_1^2}(x_1, x_2) + \frac{\partial^2 u}{\partial x_2^2}(x_1, x_2). \end{aligned}$$

et  $f$  et  $c$  sont données.

La condition aux limites  $u(x) = 0 \quad \forall x \in \partial\Omega$  s'appelle condition de Dirichlet homogène.

- 2) Il est possible de considérer d'autres conditions aux limites. Par exemple, le problème suivant

$$\begin{cases} -\Delta u(x) + c(x)u(x) = f(x) & \forall x \in \Omega \\ u(x) = g(x) & \forall x \in \partial\Omega \end{cases}$$

où  $g$  est donnée est appelé problème de Dirichlet non homogène. Au moins formellement, on voit que l'étude de ce problème se ramène à l'étude d'un problème de Dirichlet homogène

$$\begin{cases} -\Delta U(x) + c(x)U(x) = F(x) \\ U(x) = 0 & \forall x \in \partial\Omega \end{cases}$$

où on a posé  $U = u - G$ , et  $F = f + \Delta G - cG$ , où  $G$  serait une fonction dont la valeur aux bords coïncide avec  $g$ .

3) Il est enfin possible de considérer d'autres types de conditions aux limites, dont l'objectif n'est plus d'imposer une condition sur l'inconnue mais sur ses dérivées. On parle de conditions de type Neumann (par opposition aux conditions de Dirichlet ci-dessus)

le problème suivant

$$(P2) \begin{cases} -\Delta u(x) + c(x)u(x) = f(x) & \forall x \in \Omega \\ \frac{\partial u}{\partial \nu}(x) = g(x) & \forall x \in \partial\Omega \end{cases}$$

où  $\frac{\partial u}{\partial \nu} = \nabla u \cdot \vec{\nu}$  est la projection sur la normale  $\vec{\nu}$  à  $\partial\Omega$  orientée vers l'extérieur du vecteur gradient  $\nabla u$ , est appelé problème de Neumann non homogène.

Pour tous ces problèmes, nous essaierons de répondre aux questions théoriques suivantes :

- 1) Est-ce qu'il existe une solution au problème ? Si oui, dans quel espace est-elle définie ?
- 2) Le cas échéant, une telle solution est-elle unique ?
- 3) Le cas échéant, la solution dépend-elle continûment des données ?

S'il est possible de répondre positivement à ces questions, on dira que le problème est bien posé au sens de Hadamard. Nous aborderons dans un second temps l'approximation numérique de la solution exacte d'un problème elliptique bien posé au sens de Hadamard.

## Formulations variationnelles des problèmes (P1) et (P2)

23

Supposons que  $u \in H^2(\Omega)$  soit une solution du problème considéré, et soit  $v \in H^1(\Omega)$ . On suppose  $f \in L^2(\Omega)$  et que  $c \in L^\infty(\Omega)$ ,  $g \in L^2(\partial\Omega)$

### Formulation variationnelle du problème (P1)

En multipliant l'équation par  $v$  et en intégrant sur  $\Omega$ , on obtient

$$-\int_{\Omega} \Delta u v + \int_{\Omega} c u v = \int_{\Omega} f v$$

En utilisant la formule de Green, on en déduit que

$$\int_{\Omega} \nabla u \cdot \nabla v - \int_{\partial\Omega} \partial_n u \partial_n v + \int_{\Omega} c u v = \int_{\Omega} f v$$

Supposons maintenant que  $v \in H_0^1(\Omega) = \{v \in H^1(\Omega), \partial_n v = 0\}$ . Il vient

alors

$$A(u, v) = \mathcal{L}(v) \quad \forall v \in V := H_0^1(\Omega)$$

$$\text{avec } A(u, v) = \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} c u v \quad \forall v \in V.$$

$$\mathcal{L}(v) = \int_{\Omega} f v$$

Remarquons que  $u \in V$  puisque la condition aux limites du problème (P1) est  $u=0$  sur  $\partial\Omega$ .

Définition On appelle formulation variationnelle du problème (P1) le problème suivant :

$$(P1)' \left\{ \begin{array}{l} \text{Trouver } u \in V \text{ t.q. } A(u, v) = \mathcal{L}(v) \quad \forall v \in V \\ \text{avec } V = H_0^1(\Omega) \end{array} \right.$$

Dans cette formulation,  $V$  est souvent appelé espace variationnelle. On remarque qu'il s'agit d'un espace de Hilbert. On remarque également que l'application  $(u, v) \rightarrow A(u, v)$  est bilinéaire et que l'application  $v \rightarrow \mathcal{L}(v)$  est linéaire. Ces propriétés seront importantes pour l'étude de  $(P1)'$ .



Formulation variationnelle du problème (P2)

En multipliant l'équation par  $v$ , en intégrant sur  $\Omega$  et en utilisant la formule de Green, on obtient de la même manière que précédemment

$$\int_{\Omega} \nabla u \cdot \nabla v - \int_{\partial\Omega} \delta_1 u \delta_0 v + \int_{\Omega} c uv = \int_{\Omega} f v \quad \forall v \in H^1(\Omega)$$

u étant solution du problème (P2), on a ici  $\delta_1 u = g$  de sorte que

$$\int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} c uv = \int_{\Omega} f v + \int_{\partial\Omega} g \delta_0 v, \text{ où encore}$$

$$A(u, v) = \mathcal{L}(v) \quad \forall v \in V := H^1(\Omega)$$

avec  $A(u, v) = \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} c uv$

et  $\mathcal{L}(v) = \int_{\Omega} f v + \int_{\partial\Omega} g \delta_0 v$

Définition On appelle formulation variationnelle du problème (P2) le problème suivant :

$$(P2) \begin{cases} \text{Trouver } u \in V \text{ t.q. } A(u, v) = \mathcal{L}(v) \quad \forall v \in V \\ \text{avec } V := H^1(\Omega). \end{cases}$$

On remarque que  $V, A, \mathcal{L}$  vérifient les mêmes propriétés que dans le cas du problème (P1).

Interprétation des problèmes variationnels

A nous avons vu que toute solution du problème initial (P1) ou (P2) appartenant à  $H^2(\Omega)$  vérifiait le problème variationnel.

A nous étudions dans ce paragraphe la réciproque : on suppose que  $u \in V$  est solution du problème variationnel et on se demande si  $u$  vérifie le problème initial (et si oui en quel sens)

Commençons par le problème (P1).

On a  $A(u, v) = \mathcal{L}(u) \quad \forall v \in H_0^1(\Omega)$  donc en particulier

$$\int_{\Omega} \nabla u \cdot \nabla \varphi + \int_{\Omega} cu \varphi = \int_{\Omega} f \varphi \quad \forall \varphi \in \mathcal{D}(\Omega)$$

ie 
$$\sum_{i=1}^n \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial \varphi}{\partial x_i} + \int_{\Omega} cu \varphi = \int_{\Omega} f \varphi$$

Puisque  $u \in V = H_0^1$ ,  $cu \in L^2$ ,  $f \in L^2$ , on peut interpréter cette égalité au sens des distributions pour obtenir

$$\sum_{i=1}^n \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial \varphi}{\partial x_i} \right\rangle + \langle cu, \varphi \rangle = \langle f, \varphi \rangle$$

ou encore

$$- \sum_{i=1}^n \left\langle \frac{\partial^2 u}{\partial x_i^2}, \varphi \right\rangle + \langle cu, \varphi \rangle = \langle f, \varphi \rangle$$

ie 
$$- \Delta u + cu = f \quad \forall \varphi \in \mathcal{D}(\Omega)$$

On a donc obtenu, au sens des distributions, la validité de

$$- \Delta u + cu = f$$

Par ailleurs, puisque  $u \in V = H_0^1(\Omega)$ , on a tout de suite  $u = 0$  sur  $\partial\Omega$ , ie  $\gamma_0 u = 0$

On a donc montré que  $u$  était solution du problème (P1) au sens des distributions. Il est important que les égalités

$- \Delta u + cu = f$  et  $\gamma_0 u = 0$  sont donc vraies respectivement  $\forall$  ds  $L^2(\Omega)$  et  $L^2(\partial\Omega)$ , et donc presque partout sur  $\Omega$  et  $\partial\Omega$ . Il s'agit de résultat optimal que l'on pouvait espérer puisque  $f$  et  $c$  sont définies presque partout uniquement.

Remarque

1) Il n'a pas été nécessaire de supposer  $u \in H^2(\Omega)$  pour établir la réciproque. On peut néanmoins se demander si  $u$  vérifie la régularité  $H^2$ . Puisque  $\Delta u = -f + cu$ , on a immédiatement que  $\Delta u \in L^2(\Omega)$ . En une dimension d'espace, cela est

équivalent à  $u \in H^2(\Omega)$ . Dans le cas d'une dimension supérieure <sup>26</sup> à un, on peut montrer que cela reste vrai pourvu que la frontière du domaine soit suffisamment régulière.

(au sens où le résolvant du problème de  $H^2$  mais pas au sens où d'une équivalence  $u \in H^2(\Omega) \iff u \in H^1$  et  $\Delta u \in L^2$  à priori)

Passons maintenant au problème (P2).

De même que précédemment, on a

$$-\Delta u + cu = f$$

au sens des distributions, puisque  $\delta_0 \varphi = 0 \quad \forall \varphi \in \mathcal{D}(\Omega)$  (on est donc ramené aux mêmes étapes que ci-dessus), et donc aussi au sens  $L^2$ .

Il reste à se convaincre que  $\delta_1 u = g$

$$\text{On a} \quad \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} cuv = \int_{\Omega} f v + \int_{\partial\Omega} g \gamma_0 v \quad \forall v \in H^1(\Omega)$$

avec  $u \in H^1(\Omega)$ . En intégrant par partie on a

$$-\int_{\Omega} \Delta u v + \int_{\Omega} cuv + \int_{\partial\Omega} \delta_1 u \delta_0 v = \int_{\Omega} f v + \int_{\partial\Omega} g \delta_0 v \quad \forall v \in H^1$$

En combinant l'égalité  $-\Delta u + cu = f$  au sens  $L^2$ , on en déduit donc

$$\text{que} \quad \int_{\partial\Omega} (g - \delta_1 u) \delta_0 v = 0 \quad \forall v \in H^1(\Omega)$$

$$\text{et donc} \quad \forall w \in H^{1/2}(\partial\Omega), \quad \int_{\partial\Omega} (g - \delta_1 u) w = 0$$

Par densité, de  $H^{1/2}$  ds  $L^2$ , on en déduit que

$$\int_{\partial\Omega} (g - \delta_1 u) w = 0 \quad \forall w \in L^2(\partial\Omega).$$

En prenant  $w = g - \delta_1 u$  on obtient donc ce que l'on voulait, c'est-à-dire  $g = \delta_1 u$  dans  $L^2(\partial\Omega)$ . Il s'agit là encore

du résultat optimal que l'on pourrait obtenir puisque  $g \in L^2(\partial\Omega)$ .

Remarque l'égalité  $-\Delta u + cu = f$  nous dit que  $\Delta u \in L^2(\Omega)$ . De même que pour le problème (P1) on peut montrer que la solution  $u$  est dans  $H^2$  pourvu que le bord de  $\Omega$  soit suffisamment régulier, et pourvu que  $g \in H^{1/2}(\partial\Omega)$ , ainsi,  $\delta_1 u$  a bien un sens.

Remarque

Il faut avoir à l'esprit que l'on cherche la solution  $u$  dans l'espace variationnel associé à la formulation variationnelle de sorte qu'il est assez naturel, si cela est possible comme dans le cas (P1), d'introduire des hypothèses satisfaites par  $u$  de la définition de  $V$ . Inversement, il ne faut pas introduire dans l'espace  $V$  des conditions ou des hypothèses supplémentaires par rapport à celles vérifiées par la solution  $u$  du problème aux limites de départ.

Afin de montrer que les problèmes variationnels introduits plus haut admettent une unique solution, nous utiliserons le théorème de Lax-Nilgram dans des espaces de Hilbert. Avant d'énoncer et de démontrer ce théorème, nous rappelons les deux résultats suivants.

Dans tout ce paragraphe,  $H$  désigne un espace de Hilbert réel, et on note  $(\cdot, \cdot)$  le produit scalaire associé et  $\|\cdot\|$  la norme correspondante.

Théorème : projection sur un convexe fermé non vide

Soit  $C$  une partie convexe fermée non vide de  $H$ .

1) Soit  $x \in H$ . Alors il existe un unique élément noté  $P_C(x)$  de  $C$  t. q

$$\|x - P_C(x)\| \leq \|x - y\| \quad \forall y \in C.$$

De plus, on a

$$(x - P_C(x), y - P_C(x)) \leq 0 \quad \forall y \in C$$

et  $P_C(x)$  est l'unique élément de  $C$  vérifiant cette inégalité.

2) Si  $C$  est un sous-espace vectoriel de  $H$ , alors

$$(x - P_C(x), y) = 0 \quad \forall y \in C,$$

et  $P_C(x)$  est l'unique élément de  $C$  vérifiant cette égalité.

En d'autres termes, on a  $(x - P_C(x)) \in C^\perp$  et tout  $x \in H$  peut se décomposer sous la forme

$$x = \underbrace{x - P_C(x)}_{\in C^\perp} + \underbrace{P_C(x)}_{\in C}$$

de sorte que  $H = C \oplus C^\perp$ .

La décomposition est en effet unique :

$$\begin{aligned} x &= x_1 + x_2 & (x_1 - y_1) &= - (x_2 - y_2) \\ x &= y_1 + y_2 & \underbrace{\hspace{1.5cm}}_{\in C^\perp} &= \underbrace{\hspace{1.5cm}}_{\in C} \end{aligned}$$

$$\Leftrightarrow -(x_1 - y_1, x_2 - y_2) = 0$$

$$\|x_1 - y_1\| = 0 \quad \text{si } x_1 = y_1$$

$$\text{et donc } x_2 = y_2.$$

De ce résultat, on en déduit le théorème de Riesz.

Théorème de Riesz

Soit  $L: H \rightarrow \mathbb{R}$  une application linéaire et continue sur  $H$ . Alors il existe un unique élément  $u$  de  $H$  t.q

$$L(v) = (u, v) \quad \forall v \in H.$$

Démonstration

Unicité Soit  $L(v) = (u_1, v) = (u_2, v) \quad \forall v \in H$ .

alors  $(u_1 - u_2, v) = 0 \quad \forall v \in H$

En prenant  $v = u_1 - u_2$ , on obtient  $\|u_1 - u_2\| = 0$  et donc  $u_1 = u_2$ .

Existence

Si  $L(v) = 0 \quad \forall v$ , alors  $u = 0$  convient

Si non,  $\ker L \neq H$ .  $\ker L$  est un sev fermé comme image réciproque d'un fermé par une application continue ( $L$ )

D'après le théorème précédent, on a

$$H = (\ker L) \oplus (\ker L)^\perp$$

et  $(\ker L)^\perp \neq \{0\}$  puisque  $\ker L \neq H$ .

Soit  $v_0 \in (\ker L)^\perp$ ,  $\|v_0\| = 1$ .

Pour tout  $v \in H$ , on peut écrire

$$v = \frac{L(v)}{L(v_0)} v_0 + \underbrace{v - \frac{L(v)}{L(v_0)} v_0}_{=: w \in \ker L}$$

de sorte que  $(\frac{L(v)}{L(v_0)} v_0, v - \frac{L(v)}{L(v_0)} v_0) = 0$

soit  $(v_0, v - \frac{L(v)}{L(v_0)} v_0) = 0$  si  $L(v) \neq 0$ .

soit  $\frac{L(v)}{L(v_0)} (v_0, v_0) = (v_0, v)$

soit  $L(v) = (u, v)$  avec  $u = L(v_0) v_0$ .

Dans le cas  $L(v) = 0$ , est-ce que ce choix de  $u$  fonctionne, si a-t-on

$$0 = L(v_0) (v_0, v) ?$$

Oui car  $v_0 \in (\ker L)^\perp$  et  $L(v) = 0$  signifie  $v \in \ker L$ .

$y_1 \in (\ker L)^\perp$   
 $y_2 \in (\ker L)^\perp$   
 $(y_1, y) = 0 \quad \forall y \in \ker L$   
 $(y_2, y) = 0 \quad \forall y \in \ker L$   
 $\Rightarrow y_1 + \lambda y_2 \in (\ker L)^\perp$

Nous pouvons maintenant énoncer le théorème.

Théorème de Lax-Nilgram.

Soit  $V$  un espace de Hilbert réel,  $V \times V \rightarrow \mathbb{R}$  une forme  
 $(u, v) \rightarrow A(u, v)$

bilinéaire continue et coercive :

(continuité)  $\exists \eta > 0, \forall (u, v) \in V^2, |A(u, v)| \leq \eta \|u\| \|v\|,$   
 (coercivité)  $\exists \alpha > 0, \forall v \in V, A(v, v) \geq \alpha \|v\|^2,$

et  $V \rightarrow \mathbb{R}$  une forme linéaire continue :  
 $v \rightarrow L(v)$

(continuité)  $\exists C, \forall v \in V, |L(v)| \leq C \|v\|.$

Alors le problème

Trouver  $u \in V$  t.g.  $\forall v \in V, A(u, v) = L(v)$

admet une unique solution.

Démonstration

Unicité . Soit  $u_1$  et  $u_2$  t.g.

$\forall v$   $A(u_1, v) = L(v)$  de sorte que  $A(u_1 - u_2, v) = 0 \forall v$   
 $A(u_2, v) = L(v)$

En prenant  $v = u_1 - u_2$ , on obtient  
 $\alpha \|u_1 - u_2\|^2 \leq A(u_1 - u_2, u_1 - u_2) = 0$

d'où  $u_1 = u_2$ .

Existence d'après le théorème de Riesz,

$\forall u, \exists A(u)$  t.g.  $A(u, v) = (A(u), v) \forall v \in V$   
 $\exists f$  t.g.  $L(v) = (f, v) \forall v \in V.$

Ainsi,  $u \in V$  est t.g.  $A(u, v) = L(v)$   
 $\Leftrightarrow u \in V$  est t.g.  $(A(u) - f, v) \forall v \in V$   
 $\Leftrightarrow u \in V$  est t.g.  $A(u) = f.$

L'application  $L$  étant fixée, il en est de même pour  $f$ .  
 Il faut et il suffit donc de montrer que tout  $f \in V$  admet un  
 antécédent par  $A$ , c'est que  $A$  est surjective, ce que

Il est clair que  $A$  est linéaire et continue puisque  $\|A(u)\|^2 = A(u, A(u)) \leq \eta \|u\| \|A(u)\|$   
 donc  $\|A(u)\| \leq \eta \|u\|.$

Montrons tout d'abord que  $\text{Im}(A)$  est fermé (il est clair que c'est un convexe car c'est un s.e.v., et qu'il est non vide).

Soit  $A(u_p) \xrightarrow{p \rightarrow \infty} v$ . On veut mg  $v \in \text{Im} A$ .

$(A(u_p))_p$  converge donc est de Cauchy. D'après la coercivité

$$\begin{aligned} \alpha \|u_p - u_q\|^2 &\leq (A(u_p - u_q), u_p - u_q) = (A(u_p - u_q), u_p - u_q) \\ &= (Au_p - Au_q, u_p - u_q) \\ &\leq \|Au_p - Au_q\| \|u_p - u_q\| \end{aligned}$$

$$\text{de sorte que } \|u_p - u_q\| \leq \frac{1}{\alpha} \|Au_p - Au_q\|,$$

ainsi  $(u_p)$  est de Cauchy et donc converge vers un ds  $v$ .

Par continuité de  $A$ ,  $Au_p$  converge vers  $Au$  et donc  $v = Au \in \text{Im} A$ .

On a donc mg  $\text{Im}(A)$  est un convexe fermé non vide, et est un s.e.v. de  $V$ . On a donc

$$V = \text{Im}(A) \oplus (\text{Im} A)^\perp.$$

Si  $\text{Im}(A) \neq V$ , cela signifie que  $(\text{Im} A)^\perp \neq \{0\}$ .

Soit donc  $u \in V, u \neq 0$   
 $u \in (\text{Im} A)^\perp$  ie  $\forall v \in V, (u, Av) = 0$ .

ou encore  $\forall v \in V, A(v, u) = 0$ .

En prenant  $v = u$  et par coercivité on a donc

$$\alpha \|u\|^2 \leq 0$$

et donc  $u = 0$  ce qui est une contradiction.

Cela conclut la démonstration.

### Le cas particulier d'une forme symétrique

Supposons de plus que  $A$  est symétrique, ie que

$$A(u, v) = A(v, u) \quad \forall (u, v) \in V \times V.$$

Alors le problème

Trouver  $u \in V$  tq  $\forall v \in V, A(u, v) = L(v)$



est équivalent au problème de minimisation

32/

Trouver  $u \in V$  t.q.  $\forall v \in V \quad E(u) \leq E(v)$

$$\text{avec } E(v) = \frac{1}{2} A(v, v) - L(v)$$

Donnons ce résultat.

Remarquons tout d'abord que  $\forall t \in \mathbb{R}, \forall h \in V$ .

$$E(v+th) = \frac{1}{2} A(v+th, v+th) - L(v+th)$$

$$E(v+th) = E(v, v) + \frac{t}{2} A(h, v) + \frac{t}{2} A(v, h) + \frac{t^2}{2} A(h, h) - t L(h)$$

$$E(v+th) = E(v) + t (A(v, h) - L(h)) + \frac{t^2}{2} A(h, h)$$

Rappelons aussi que  $A(h, h) \geq \alpha \|h\|^2 \geq 0$ .

Il est ainsi clair que si  $u \in V$  est t.q.  $A(u, v) = L(v) \quad \forall v \in V$

alors  $E(u+h) - E(u) \geq \forall h \in V, \forall t \in \mathbb{R}$ .

$$\text{si } E(u+h) \geq E(u) \quad \forall h \in V, \forall t \in \mathbb{R}$$

et donc  $E(v) \geq E(u) \quad \forall v \in V$ .

Réciproquement, si  $u$  minimise  $E$ , alors on en déduit que

$$t (A(u, h) - L(h)) + \frac{t^2}{2} A(h, h) \geq 0 \quad \forall t, \forall h.$$

En divisant par  $t > 0$  et en faisant  $t \rightarrow 0$ , on obtient

$$A(u, h) - L(h) \geq 0 \quad \forall h \in V$$

En divisant par  $t < 0$  et en faisant  $t \rightarrow 0$ , on obtient

$$A(u, h) - L(h) \leq 0 \quad \forall h \in V$$

On a donc  $A(u, h) = L(h) \quad \forall h \in V$ ;

ce qui conclut la démonstration.

### Remarques

- 1) L'équivalence avec un problème de minimisation peut être exploitée du point de vue numérique pour calculer une approximation de  $u$  en utilisant des algorithmes classiques d'optimisation
- 2) L'inégalité suivante sera utile pour le problème de la continuité par rapport aux données :  $\alpha \|u\|^2 \leq A(u, u) = L(u) \leq c \|u\| \Rightarrow \|u\| \leq \frac{c}{\alpha}$ .

# Application du théorème de Lax-Nilgram à quelques problèmes modèles.

Nous allons appliquer dans ce paragraphe le théorème de Lax-Nilgram aux problèmes modèles (P1) et (P2).

## Application au problème (P1).

Rappelons que  $V = H_0^1$  est un espace de Hilbert pour la norme  $\|\cdot\|_{H^1}$  induite par  $H^1$ , mais aussi d'après l'inégalité de Poincaré pour la norme réduite  $\|\cdot\|_1$ . Nous utiliserons cette norme.

Observons tout d'abord que

$$|\mathcal{L}(v)| = \left| \int_{\Omega} f v \right| \underset{CS}{\leq} \|f\|_{L^2} \|v\|_{L^2} \underset{\text{Poincaré}}{\leq} C_f \|f\|_{L^2} \|v\|_1$$

de sorte que  $\mathcal{L}$  est continue avec  $C = C_f \|f\|_{L^2}$ .

Positivité, on a

$$\begin{aligned} |A(u, v)| &= \left| \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} c u v \right| \underset{CS \text{ de } L^2}{\leq} \sum_i \left\| \frac{\partial u}{\partial x_i} \right\|_{L^2} \left\| \frac{\partial v}{\partial x_i} \right\|_{L^2} + \|c\|_{L^\infty} \|u\|_{L^2} \|v\|_{L^2} \\ &\leq \|u\|_1 \|v\|_1 + \|c\|_{L^\infty} C_f^2 \|u\|_1 \|v\|_1 \\ &\underset{CS \text{ de } u \text{ et } v + \text{Poincaré}}{\leq} M \|u\|_1 \|v\|_1 \end{aligned}$$

Ainsi,  $A$  est bilinéaire et continue et  $M = 1 + C_f^2 \|c\|_{L^\infty}$ .

Étudions maintenant la coercivité de  $A$ .

On a

$$A(u, u) = \sum_{i=1}^n \int_{\Omega} \left( \frac{\partial u}{\partial x_i} \right)^2 + \int_{\Omega} c u^2$$

Ainsi, il est clair que si  $c \geq 0$  alors  $A(u, u) = |A(u, u)| \geq \alpha \|u\|_1^2$  avec  $\alpha = 1$ .

Obtons  $c^\pm$  la partie <sup>positive</sup> négative de  $c = c^- + c^+$

$$\begin{aligned} \text{On a } |A(u, u)| &\geq \|u\|_1^2 - \|c^-\|_{L^\infty} \|u\|_{L^2}^2 \\ &\geq \|u\|_1^2 - \|c^-\|_{L^\infty} C_f^2 \|u\|_1^2 \\ &\geq (1 - \|c^-\|_{L^\infty} C_f^2) \|u\|_1^2 \end{aligned}$$

Pourvu que  $c \geq 0$  ou  $\|c^-\|_{L^\infty} < \frac{1}{C_f^2}$ , on a donc la coercivité

Le théorème de Lax-Nikolski assure donc l'existence et l'unicité de  $u$ . 34,  
 Remarquons enfin que l'inégalité de la remarque 2) p 32  
 nous dit ici que

$$\|u\|_1 \leq C_0 \|f\|_{L^2}$$

$$\text{avec } C_0 = \frac{C_f}{\alpha} \text{ et } \alpha = \begin{cases} 1 & \text{si } c \geq 0 \\ 1 - \|c\|_{L^\infty} C_f^2 & \text{si } \|c\|_{L^\infty} < \frac{1}{C_f^2} \end{cases}$$

L'application qui à  $f$  associe  $u$  étant clairement linéaire  
 l'inégalité  $\|u\|_1 \leq C_0 \|f\|_{L^2}$  nous dit qu'elle est aussi  
 continue. Le problème (P1) dépend continûment de la donnée  
 $f$ . Il est donc bien posé au sens de Hadamard.

### Application au problème (P2)

On considère ici la norme usuelle sur  $H^1$  puisque  $V = H^1$ .

On a tout d'abord

$$\begin{aligned} |A(u, v)| &\leq |u_h| |v_h|_1 + \|c\|_{L^\infty} \|u\|_{L^2} \|v\|_{L^2} \leq \max(1, \|c\|_{L^\infty}) (|u_h| |v_h|_1 + \|u\|_{L^2} \|v\|_{L^2}) \\ &\leq \max(1, \|c\|_{L^\infty}) \|u\|_{H^1} \|v\|_{H^1} \\ &\quad \uparrow \\ &\quad \text{C.S. de } \mathbb{R}^2 \end{aligned}$$

$$\begin{aligned} |L(v)| &= \left| \int_{\Omega} f v + \int_{\partial\Omega} g \gamma_0 v \right| \leq \|f\|_{L^2} \|v\|_{L^2} + \|g\|_{L^2(\partial\Omega)} \|\gamma_0 v\|_{L^2(\partial\Omega)} \\ &\leq \|f\|_{L^2} \|v\|_{L^2} + \|g\|_{L^2(\partial\Omega)} C_0 \|v\|_{H^1} \end{aligned}$$

de sorte que  $|L(v)| \leq C \|v\|_{H^1}$  avec  $C = \|f\|_{L^2} + \|g\|_{L^2(\partial\Omega)} C_0$ .

Voyons maintenant la coercivité. Pour cela, on suppose qu'il existe  
 $C_0 > 0$  t.q.  $c(x) \geq C_0$  p.p sur  $\Omega$ .

Alors

$$\begin{aligned}
 |A(v,v)| &= \sum_{i=1}^r \left\| \frac{\partial v}{\partial x_i} \right\|_{L^2}^2 + \int c v^2 \\
 &\geq \sum_{i=1}^r \left\| \frac{\partial v}{\partial x_i} \right\|_{L^2}^2 + c_0 \|v\|_{L^2}^2 \\
 &\geq \min(1, c_0) \|v\|_{H^1(\Omega)}^2.
 \end{aligned}$$

d'où la coercivité.

Le théorème de Lax-Nilgram s'applique donc :

Concernant la dépendance continue de la solution en fait rapport aux données  $f$  et  $g$  du problème, on a l'existence d'une constante

$$k \text{ telle que } \|u\|_{H^1} \leq \frac{\|f\|_{L^2} + \|g\|_{L^2(\partial\Omega)} c_0}{\min(1, c_0)} \leq k (\|f\|_{L^2} + \|g\|_{L^2(\partial\Omega)}).$$

$$\text{avec } k = \max(1, c_0) / \min(1, c_0).$$

### Remarque

Remarquons que si  $c \equiv 0$ , alors le problème est mal posé. En effet on remarque tout d'abord que  $u$  n'intervenant que par le biais de son gradient, si  $u$  est solution alors  $u+k$  l'est aussi. Il n'y a donc pas unicité. Par ailleurs l'existence de la solution implique nécessairement (prendre  $v=1$ ) la relation  $\int_{\Omega} f + \int_{\partial\Omega} g = 0$ . Si cette relation n'est pas satisfaite, il ne peut pas y avoir de solution.

# Présentation de la méthode des éléments finis

36/

Notre objectif est de résoudre de manière approchée le problème variationnel général suivant :

$$(P) \quad \text{Trouver } u \in V \text{ t.q. } A(u, v) = L(v) \quad \forall v \in V$$

où  $V$  est un espace de Hilbert,  $L$  une forme linéaire continue sur  $V$  et  $A$  une forme bilinéaire, continue et  $V$ -elliptique. Le théorème de lax-Nilgram assure l'existence et l'unicité d'une solution à ce problème.

Le principe de la méthode des éléments finis va consister à résoudre de manière exacte le problème approché suivant

$$(P_h) \quad \text{Trouver } u_h \in V_h \text{ t.q. } A(u_h, v_h) = L(v_h) \quad \forall v_h \in V_h$$

où  $V_h$  est un sous-espace vectoriel fermé de  $V$ . Là encore, le théorème de lax-Nilgram assure l'existence et l'unicité d'une solution à ce problème.

Il faut alors s'assurer que la solution  $u_h$  soit simple à calculer, mais aussi qu'elle soit "proche" de la solution exacte  $u$ .

Pour le premier point, on choisira  $V_h$  un sous-espace vectoriel de dimension finie. On verra alors que la résolution du problème approché se ramène à la résolution d'un système linéaire, les inconnues étant les composantes de la solution  $u_h$  dans une base du sous-espace vectoriel  $V_h$ . Pour le second point, il faut s'assurer que  $V_h$  est "proche" de  $V$  en un sens à préciser. On sera plus précisément amené à définir une suite d'approximation  $(V_h)_{h>0}$  de  $V$  ainsi qu'une notion de "convergence" de  $V_h$  vers  $V$ .

Il ... l'abord vers 0.

Définition On dit que  $(v_h)_{h>0}$  est une suite d'approximations conforme de  $V$  si et seulement si

(i)  $\forall h > 0$ ,  $h$  assez petit,  $\dim V_h < \infty$  et  $V_h \subset V$

(ii)  $\forall v \in V, \forall \epsilon > 0, \exists \eta \in \mathbb{R}, \forall h \leq \eta,$

$\exists (v_h) \in V_h,$   
 $\|v - v_h\|_V \leq \epsilon.$

Lemme de Céa.

On se place ds les conditions du théorème de Lax-Nilgram et on considère  $V_h \subset V$  t.g.  $\dim V_h < \infty$ . Soit  $u$  la solution du problème (P) et  $u_h$  la solution du problème (P<sub>h</sub>). Alors

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V.$$

Démonstration.

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

et donc  $a(u, v_h) = L(v_h) \forall v_h \in V_h$

tandis que  $a(u_h, v_h) = L(v_h) \forall v_h \in V_h$ .

Ainsi, il est clair que

$$a(u - u_h, v_h) = 0 \forall v_h \in V_h.$$

En particulier, on a donc  $\forall v_h \in V_h$

$$\alpha \|u - u_h\|^2 \leq a(u - u_h, u - u_h) = a(u - u_h, u - v_h) \leq M \|u - u_h\| \|u - v_h\|$$

On a donc  $\|u - u_h\|_V \leq \frac{M}{\alpha} \|u - v_h\|_V \forall v_h \in V_h$  ce qui

conclut la démonstration.

Ce résultat est intéressant mais  $u$  n'étant généralement pas connue, il reste à trouver une estimation de

$$\inf_{v_h \in V_h} \|u - v_h\|.$$

Pour revenir sur ce point ultérieurement.

38,

### Calcul de la solution approchée

Obtons  $m_h$  la dimension de  $V_h$  et considérons  $\{w_h^1, \dots, w_h^{m_h}\}$  une base de  $V_h$ . Soit  $u_h \in V_h$  la solution du problème  $(P_h)$ . On note  $\{u_{h1}, \dots, u_{hm_h}\}$  les coordonnées de  $u_h$  dans la base  $\{w_h^1, \dots, w_h^{m_h}\}$ :

$$u_h = \sum_{i=1}^{m_h} u_{hi} w_h^i$$

$$\text{On pose } U_h = \begin{pmatrix} u_{h1} \\ \vdots \\ u_{hm_h} \end{pmatrix}$$

On a alors le résultat suivant.

### Proposition

$u_h \in V_h$  est solution de  $(P_h)$  si et seulement si

$$A_h U_h = B_h$$

$$\text{avec } A_{hij} = A(\underline{w}_h^j, \underline{w}_h^i) \quad \forall i, j = 1, \dots, m_h.$$

$$B_{hi} = L(w_h^i)$$

De plus, la matrice  $A_h$  est définie positive (elle est aussi symétrique si  $A$  l'est) et le système linéaire  $A_h U_h = B_h$  admet une unique solution.

### Démonstration

Il est clair que l'égalité  $A(u_h, v_h) = L(v_h) \quad \forall v_h \in V_h$  est vérifiée, par linéarité de  $A$  et  $L$ , si et seulement si elle est vraie  $\forall w_h^i$ ,  $i = 1, \dots, m_h$ . Ainsi,  $u_h$  est solution du problème si et seulement si

$$A\left(\sum_{j=1}^{m_h} u_{hj} w_h^j, w_h^i\right) = L(w_h^i) \quad \forall i = 1, \dots, m_h.$$

c'est-à-dire

$$\sum_{j=1}^{m_2} u_j A(w_2^j, w_2^i) = L(w_2^i) \quad \forall i=1, \dots, m_2$$

ce qui n'est rien d'autre que le système  $A_2 U_2 = B_2$ .

On sait déjà que ce système admet une unique solution puisque nous venons de voir qu'il est équivalent à (P<sub>2</sub>) qui lui-même admet une unique solution d'après le théorème de lax-Diagram.

Montrons que  $A_2$  est définie positive. Soit  $U = (u_i)_{i=1}^{m_2}$ ,

$$\begin{aligned} (A_2 U, U) &= \sum_{i=1}^{m_2} \sum_{j=1}^{m_2} A_{2,ij} u_j u_i = \sum_i \sum_j A(w_2^j, w_2^i) u_j u_i \\ &= A\left(\sum_j u_j w_2^j, \sum_i u_i w_2^i\right) \\ &\geq \alpha \left\| \sum_i u_i w_2^i \right\|^2 \end{aligned}$$

par  $\forall$ -ellipticité de  $A$ .

On a donc bien  $(A_2 U, U) \geq 0 \quad \forall U$

et  $(A_2 U, U) = 0 \Leftrightarrow U = 0$ ,

ce qui termine la démonstration.

Prenez en œuvre sur un exemple concret, en dimension 1

Pour simplifier, on considère le problème (P1) en dimension 1 d'espace sur le domaine  $\Omega = ]0, 1[$ :

$$\left\{ \begin{array}{l} -u''(x) + c(x)u(x) = f(x) \quad \forall x \in ]0, 1[ \\ u(0) = u(1) = 0. \end{array} \right.$$

avec  $c(x) = c \quad \forall x$ .

On rappelle que la formulation variationnelle est

$$\begin{aligned} \text{Trouver } u \in H_0^1(]0, 1[) \text{ t.q.} \\ A(u, v) = L(v) \quad \forall v \in H_0^1(]0, 1[) \end{aligned}$$

$$\text{avec } A(u, v) = \int_0^1 u'v' + \int_0^1 c u v \quad \forall u, v \in H_0^1(]0, 1[)$$

$$\text{et } L(v) = \int_0^1 f v \quad \forall v \in H_0^1(]0, 1[)$$



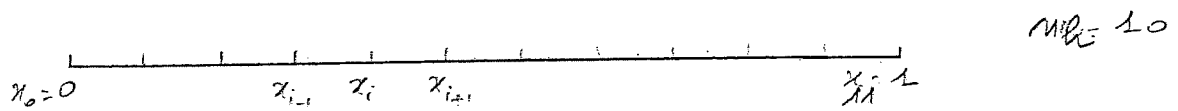
Afin de formaliser le problème approché, il nous faut construire 40,  
 la suite d'approximations conforme  $(V_h)_{h>0}$ ,  $V_h \subset H^1(\Omega) \forall h$ .  
 Cela impose en particulier que les fonctions de  $V_h$  soient continues  
 et nulles en 0 et en 1. Un critère important de la choix des  
 $V_h$  est lié au caractère creux de la matrice  $A_h$  associé, afin  
 de faciliter l'inversion du système  $A_h U_h = B_h$ . Puisque l'on a

$$A_{h,ij} = A(w_h^i, w_h^j) = \int_0^1 w_h^i{}' w_h^j{}' + \int_0^1 c w_h^i w_h^j$$

cela revient à trouver des fonctions  $w_h^i$  dont les supports soient  
 d'intersections "le plus souvent vides". Un tel exemple nous est  
 donné par les éléments P1. (A ne pas confondre avec le problème  
 (P1) !)

### Éléments P1

On partitionne le segment  $[0, 1]$  en  $M_h+1$  sous-intervalles (avec  
 $N$  un entier naturel donné) de taille  $h = \frac{1}{M_h+1}$ . On note  $x_i = ih$ ,  
 $i=0, \dots, M_h+1$  les  $(M_h+2)$  points du maillage ainsi défini :



On considère comme ensemble  $V_h$  l'ensemble des fonctions continues  
 sur  $[0, 1]$ , qui s'annulent en 0 et en 1, et qui sont affines  
 sur chaque sous-intervalle  $[\frac{i}{M_h+1}, \frac{i+1}{M_h+1}]$  avec  $i=0, \dots, M_h$ .

Sur chacun de ses sous-intervalles, toute fonction  $v \in V_h$  s'écrit donc  
 sous la forme  $v(x) = v(x_i) + \frac{x-x_i}{h} (v(x_{i+1}) - v(x_i)) \quad \forall x \in [x_i, x_{i+1}]$ .

Et en particulier  $v(x) = \frac{x}{h} v(x_1) \quad \forall x \in [0, h]$   
 $v(x) = \frac{1-x}{h} v(x_{M_h}) \quad \forall x \in [1-h, 1]$

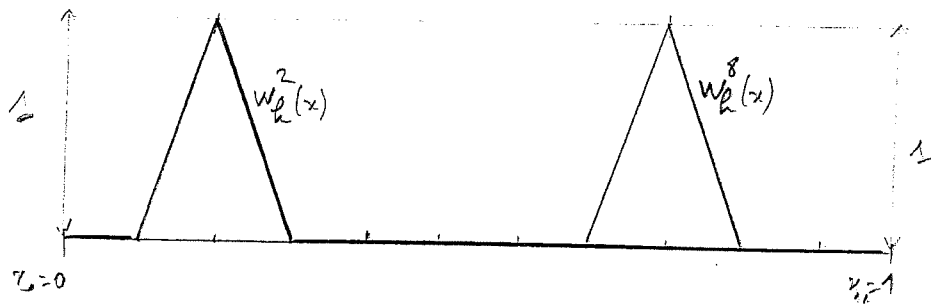
On remarque en particulier que toute fonction de  $V_h$  est déterminée  
 par la connaissance de ses valeurs aux points  $x_1, \dots, x_{M_h}$ . Ses  
 valeurs aux autres points sont donc les degrés de liberté de la fonction  $v \in V_h$ .

On est alors tenté d'introduire les fonctions de bases  $w_h^i, i=1, \dots, m_h$  411  
 affines par morceaux, sur chaque sous-intervalle  $[x_i, x_{i+1}]$ , et telles  
 que  $w_h^i(x_j) = \delta_{ij}$  i.e.  $w_h^i(x_j) = 0$  si  $j \neq i$   
 et  $w_h^i(x_i) = 1$  si  $i=j$ .

Ces fonctions de base appartiennent bien à  $V_h$  et on remarque que  
 toute fonction  $v$  de  $V_h$  peut s'écrire de manière unique comme  
 une combinaison linéaire de ces fonctions de base:

$$v(x) = \sum_{i=1}^{m_h} v(x_i) w_h^i(x) \quad \forall x \in [0,1], \forall v \in V_h.$$

La famille  $\{w_h^i, i=1, \dots, m_h\}$  forme donc une base de  $V_h$ , qui  
 est donc un espace vectoriel de dimension  $h$ .



Il reste à vérifier que  $V_h \subset V = H_0^1([0,1]) \quad \forall m_h$ .

### Lemme

Pour tout  $m_h \in \mathbb{N}^*$ ,  $V_h \subset V = H_0^1([0,1])$

### Démonstration

Il faut montrer que toute fonction  $v \in V_h$  est dans  $L^2([0,1])$ , que  
 sa dérivée au sens des distributions est dans  $L^2([0,1])$  et que sa trace  
 au bord, i.e. en 0 et en 1, est nulle. Ce dernier point est vrai par  
 construction puisque  $V_h \subset C^0([0,1])$  et  $v(0) = v(1) = 0 \quad \forall v \in V_h$ .

Le premier point est lui aussi immédiat par continuité des fonctions

Soit donc  $v \in V_h$ . Calculons la dérivée de  $v$  au sens des distributions. 42,  
 butions. Soit  $\varphi \in \mathcal{D}([0,1[$ )

$$\begin{aligned} \langle v', \varphi \rangle &= - \langle v, \varphi' \rangle = - \int_0^1 v(x) \varphi'(x) dx \\ &= - \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} v(x) \varphi'(x) dx. \end{aligned}$$

En utilisant une formule d'intégration par partie sur chacun des sous-intervalles  $[x_i, x_{i+1}]$  où  $v$  est de classe  $\mathcal{C}^1$ , on a

$$\langle v', \varphi \rangle = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} v'(x) \varphi(x) dx + \sum_{i=0}^{n-1} \{v(x_{i+1}) \varphi(x_{i+1}) - v(x_i) \varphi(x_i)\}$$

c'est à dire

$$\langle v', \varphi \rangle = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} v'(x) \varphi(x) dx + v(x_{n-1}) \varphi(x_{n-1}) - v(x_0) \varphi(x_0).$$

Comme  $\varphi(0) = \varphi(1) = 0$  et que sur chaque sous-intervalle  $[x_i, x_{i+1}]$ , la dérivée  $v'(x)$  est une constante que l'on notera  $\{v'_{i+1/2}\}$ , on obtient

$$\langle v', \varphi \rangle = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} \{v'_{i+1/2}\} \varphi(x) dx.$$

En définissant la fonction  $\{v'\}$  de la manière suivante

$$\{v'\}(x) = \{v'_{i+1/2}\} \quad \forall x \in ]x_i, x_{i+1}[,$$

on a donc montré que  $v' = \{v'\}$ . La fonction  $\{v'\}$  étant clairement de  $L^2([0,1[$ ), on a donc finalement  $v' \in L^2$  et donc  $v \in H^1$ . Ce qui termine la démonstration.

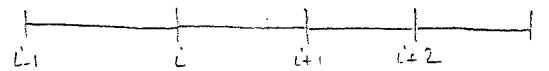
Précisons maintenant la matrice  $A_h$ . Il est clair tout d'abord que pour tout  $i=1, \dots, n_h$ , le support de la fonction de base est donné par  $[x_{i-1}, x_{i+1}]$

Ainsi, l'intersection des supports de deux fonctions  $w_h^i$  et  $w_h^j$  telles que  $|i-j| \geq 2$  est vide. est ainsi clair que

$$A(w_h^i, w_h^j) = 0 \quad \forall i, j \text{ t. q. } |i-j| > 2$$

ce qui signifie que la matrice  $A_h$  est tridiagonale. Elle est 43/ également clairement symétrique puisque  $A(v_1, v_2) = A(v_2, v_1)$ . On a enfin, puisque  $w_h^i(x) = \frac{x-x_{i-1}}{h} \mathcal{1}_{[x_{i-1}, x_i]} + \frac{x_{i+1}-x}{h} \mathcal{1}_{[x_i, x_{i+1}]}$ ,

$$\begin{aligned} A_{h,ii} &= A(w_h^i, w_h^i) = \int_0^1 w_h^{i'2} + \int_0^1 c w_h^{i2} \\ &= \int_{x_{i-1}}^{x_{i+1}} (w_h^{i'2} + c w_h^{i2}) = \int_{x_{i-1}}^{x_i} \left(\frac{1}{h}\right)^2 + c \frac{(x-x_{i-1})^2}{h^2} + \int_{x_i}^{x_{i+1}} \left(\frac{1}{h}\right)^2 + c \frac{(x_{i+1}-x)^2}{h^2} \\ &= \frac{1}{h} + \frac{c}{3h^2} [(x-x_{i-1})^3]_{x_{i-1}}^{x_i} + \frac{1}{h} + \frac{c}{3h^2} [(x_{i+1}-x)^3]_{x_i}^{x_{i+1}} \\ &= \frac{1}{h} + \frac{ch}{3} + \frac{1}{h} + \frac{ch}{3} \\ &= \frac{2}{h} + \frac{2ch}{3} \end{aligned}$$



$$\begin{aligned} A_{h,ii+1} &= A(w_h^{i+1}, w_h^i) = \int_0^1 w_h^{i+1'} w_h^{i'} + c \int_0^1 w_h^{i+1} w_h^i \\ &= \int_{x_i}^{x_{i+1}} w_h^{i+1'} w_h^{i'} + c \int_{x_i}^{x_{i+1}} w_h^{i+1} w_h^i \\ &= \int_{x_i}^{x_{i+1}} \frac{1}{h} x \left(-\frac{1}{h}\right) + c \int_{x_i}^{x_{i+1}} \frac{(x-x_i)}{h} \frac{(x_{i+1}-x)}{h} dx \\ &= -\frac{1}{h} + \frac{c}{h^2} \left\{ \int_{x_i}^{x_{i+1}} \frac{(x-x_i)^2}{2} + \left[ \frac{(x-x_i)^2 (x_{i+1}-x)}{2} \right]_{x_i}^{x_{i+1}} \right\} \\ &= -\frac{1}{h} + \frac{c}{h^2} \left\{ \frac{1}{6} [(x-x_i)^3]_{x_i}^{x_{i+1}} + 0 \right\} \\ &= -\frac{1}{h} + \frac{c}{6h^2} h^3 = -\frac{1}{h} + \frac{ch}{6} \end{aligned}$$



## Théorème

Soit  $f \in L^2(I_0, 1; \mathbb{C})$  et soit  $u \in H_0^1(I_0, 1; \mathbb{C})$  la solution du problème

$$\begin{cases} -u'' + cu = f & \text{sur } I_0, 1; \mathbb{C} \\ u(0) = u(1) = 0. \end{cases}$$

Soit  $u_h$  la solution approchée définie à l'aide de la résolution (exacte) du système linéaire

$$A_h u_h = B_h.$$

avec  $h = \frac{1}{m_h + 1}$  comme décrit ci-dessus.

Alors, il existe une constante  $D$  ne dépendant pas de  $h$  telle que

$$\|u - u_h\|_{H^1} \leq D h.$$

## Démonstration

On sait déjà d'après le lemme de Coa que

$$\|u - u_h\|_{H^1} \leq \inf_{v_h \in V_h} \|u - v_h\|_{H^1}$$

Il suffirait donc d'établir l'inégalité  $\|u - v_h\|_{H^1} \leq D h$  pour un certain élément bien choisi de  $V_h$ .

On sait que la solution exacte  $u$  est dans  $H^2(I_0, 1; \mathbb{C})$  puisque le bord de  $I_0, 1; \mathbb{C}$  est régulier, et donc que  $u$  admet un représentant dans  $\mathcal{C}^1(I_0, 1; \mathbb{C})$ .

En particulier, on peut parler de la valeur de  $u$  en un point et de sa dérivée.

Soit donc la fonction notée  $\pi_h u$  définie par

$$\pi_h u = \sum_{i=0}^{m_h} u(x_i) w_h^i.$$

Il s'agit de l'interpolé de  $u$  en chacun des nœuds du maillage. Supposons temporairement que  $u \in \mathcal{C}^2(I_0, 1; \mathbb{C})$ , et essayons d'estimer

$$\|u - \pi_h u\|_{H^1}^2 = \sum_{i=0}^{m_h} \int_{x_i}^{x_{i+1}} |u - \pi_h u|^2(x) dx + \int_{x_i}^{x_{i+1}} |u' - (\pi_h u)'|^2(x) dx.$$

$$w_h^i(x) = \begin{cases} (x_{i+1} - x)/h & \text{si } x \in (x_i, x_{i+1}] \\ (x - x_{i-1})/h & \text{si } x \in [x_{i-1}, x_i) \end{cases}$$

de sorte que sur chaque intervalle  $[x_i, x_{i+1}]$ , on a

$$(u - \Pi_2 u)(x) = u(x) - u(x_i) \frac{(x_{i+1}-x)}{h} - u(x_{i+1}) \frac{(x-x_i)}{h}$$

$$(u' - \Pi_2 u')(x) = u'(x) - \frac{u(x_{i+1}) - u(x_i)}{h}$$

Concentrons nous sur les termes impliquant les dérivées. Puisque l'on a supposé que  $u$  est de classe  $\mathcal{C}^3$ , on a d'après la formule de Taylor avec RZ

$$u(x_{i+1}) = u(x) + (x_{i+1}-x) u'(x) + \int_x^{x_{i+1}} u''(y) (x_{i+1}-y) dy$$

$$u(x_i) = u(x) + (x_i-x) u'(x) + \int_x^{x_i} u''(y) (x_i-y) dy$$

de sorte que

$$\frac{u(x_{i+1}) - u(x_i)}{h} = u'(x) + \frac{1}{h} \left\{ \int_x^{x_{i+1}} u''(y) (x_{i+1}-y) dy - \int_x^{x_i} u''(y) (x_i-y) dy \right\}$$

Or d'après l'inégalité de Cauchy-Schwarz

$$\left| \int_x^{x_{i+1}} u''(y) (x_{i+1}-y) dy \right|^2 \leq \|u''\|_{L^2(x, x_{i+1})}^2 \int_x^{x_{i+1}} (x_{i+1}-y)^2 dy = \|u''\|_{L^2(x, x_{i+1})}^2 \frac{(x_{i+1}-x)^3}{3}$$

$$\left| \int_x^{x_i} u''(y) (x_i-y) dy \right|^2 \leq \|u''\|_{L^2(x, x_i)}^2 \int_x^{x_i} (x_i-y)^2 dy = \|u''\|_{L^2(x, x_i)}^2 \frac{(x_i-x)^3}{3}$$

de sorte que

$$\|u' - \Pi_2 u'\|_2^2 = \int |u'(x) - \frac{u(x_{i+1}) - u(x_i)}{h}|^2 dx$$

$$\leq \frac{2}{3h^2} \left\{ \|u''\|_{L^2(x, x_{i+1})}^2 \frac{(x_{i+1}-x)^3}{3} + \|u''\|_{L^2(x, x_i)}^2 \frac{(x_i-x)^3}{3} \right\} \text{ sur chaque intervalle } [x_i, x_{i+1}]$$

$(a+b)^2 \leq 2(a^2+b^2)$

$$\leq \frac{2}{3h^2} h^3 \|u''\|_{L^2(x_i, x_{i+1})}^2 \times \frac{2}{3}$$

$$\leq \frac{2h}{3} \|u''\|_{L^2(x_i, x_{i+1})}^2 \times \frac{2}{3}$$

ainsi

$$\int_{x_i}^{x_{i+1}} (u' - \Pi_2 u')^2(x) dx \leq \frac{2h^2}{3} \|u''\|_{L^2(x_i, x_{i+1})}^2 \times \frac{2}{3}$$

En sommant sur  $i=0, \dots, m_h$ , on en déduit tout de suite que 47.

$$\int_0^1 |u' - (\pi_h u)'|^2(x) dx \leq \frac{2}{3} \times \frac{2h^2}{3} \|u''\|_{L^2(0,1)}^2 \quad \text{puisque } (m_{h+1}) = \frac{1}{h}$$

En utilisant l'inégalité de Binegaré, on a aussi

$$\int_0^1 |u - \pi_h u|^2(x) dx \leq \frac{4}{9} h^2 \|u''\|_{L^2(0,1)}^2 \quad \left( \begin{array}{l} \Rightarrow \|u - \pi_h u\|_{H^1}^2 \leq kh^2 \\ \Rightarrow \|u - \pi_h u\|_{H^1} \leq kh \end{array} \right)$$

On a supposé jusqu'à présent que  $u$  était  $\mathcal{C}^2([0,1])$ . Par densité de  $\mathcal{C}^2([0,1])$  dans  $H^2([0,1])$ , ces inégalités restent vraies ds  $H^2([0,1])$  donc pour la solution générale  $u \in H_0^1 \cap H^2$  du problème de Dirichlet considéré.

Ceci termine la démonstration.

## La méthode des éléments finis en dimension deux

On considère maintenant un domaine en dimension deux d'espace.

On pourra par exemple avoir à l'esprit le domaine  $\Omega = ]0,1[ \times ]0,1[$ .

L'idée générale reste bien entendu la même que précédemment et,

comme on pourrait s'y attendre, consistera à recouvrir le domaine  $\Omega$  par des éléments notés  $T_k$ ,  $k=1, \dots, N_T$  dont la "taille" sera destinée à tendre vers 0. Nous considérerons dans la suite des éléments

de forme rectangulaire et triangulaire uniquement. L'ensemble des éléments  $T_k$ ,  $k=1, \dots, N_T$  sera appelé une triangulation (même si on l'on manipule des rectangles) et noté  $\mathcal{T}_h$  où  $h$  sera appelé la taille de la triangulation et sera définie par

$$h = \max_{k \in \{1, \dots, N_T\}} h(T_k)$$

où  $h(T_k)$  désigne le diamètre de  $T_k$ , c'est-à-dire la distance maximale entre deux points quelconque de  $T_k$ . C'est ce qu'on a appelé



"taille" ci-dessus et bien sûr  $h(T_k)$  sera destinée à tendre vers 0 (48)  
 $(h \rightarrow 0 \Rightarrow h(T_k) \rightarrow 0 \forall k)$

Définition (triangulation admissible)

On dit que la triangulation  $\mathcal{T}_h =$  est admissible ssi

(i) l'intersection entre deux éléments quelconques  $T_k$  et  $T_{k'}$  est soit vide, soit réduite à un point, soit réduite à un côté

entier

(ii) tous les coins de  $\Omega$  sont des sommets d'éléments de  $\mathcal{T}_h$

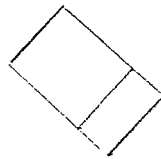
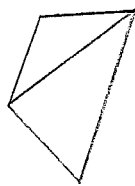
(iii) tous les coins de  $\partial(\bigcup_{k=1}^N T_k) =: \partial\Omega_h$  sont des points de  $\partial\Omega$

(iv) les éléments sont d'aire non nulle (on dit qu'ils ne sont pas dégénérés).

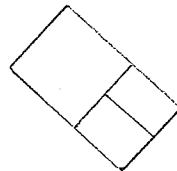
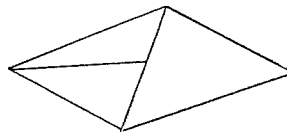
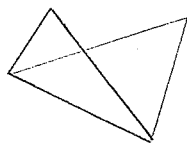
(v)  $\exists C > 0$  t.q.  $\forall h > 0 \quad \sup_{T \in \mathcal{T}_h} \frac{h(T)}{\rho(T)} \leq C$

où  $\rho(T)$  désigne le diamètre du plus grand cercle inscrit dans l'élément  $T$ .

Voici quelques exemples de triangles / rectangles  
 - admissibles

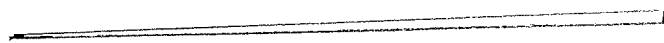


- non admissibles



La condition (v) nous dit que lorsque  $h \rightarrow 0$  (remarquez que  $C$  ne doit pas dépendre de  $h$ ) les éléments ne deviennent pas de

- plus en plus allongés :



On rappelle qu'on s'intéresse toujours au problème (P1) mais maintenant en dimension 2 d'espace sur le domaine  $\Omega = ]0,1[ \times ]0,1[$  49

$$\begin{cases} \Delta u + c(x) u(x) = f(x) \\ u(x) = 0 \text{ sur } \partial\Omega \end{cases} \quad \forall x = (x_1, x_2) \in ]0,1[ \times ]0,1[ = \Omega$$

dont la formulation variationnelle est

Trouver  $u \in H_0^1(]0,1[ \times ]0,1[)$  tel- $q$   $A(u, v) = L(v) \quad \forall v \in H_0^1(]0,1[ \times ]0,1[$

avec toujours  $A(u, v) = \iint_{\Omega} u'v' dx_1 dx_2 + \iint_{\Omega} c uv dx_1 dx_2$

et  $L(v) = \iint_{\Omega} f v dx_1 dx_2$

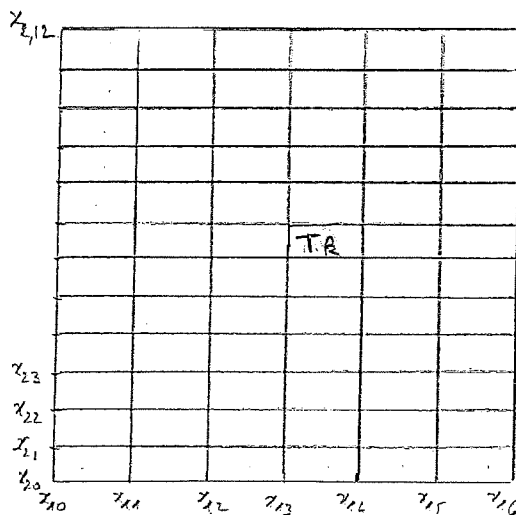
Comme dans le cas 1D, il nous faut maintenant définir les espaces  $V_h \subset H_0^1(]0,1[ \times ]0,1[) \quad \forall h$ .

### Éléments $\mathcal{D}_1$

On partitionne le carré  $[0,1] \times [0,1]$  en  $(N_1+1)$  sous-intervalles dans la direction  $x_1$  et  $(N_2+1)$  sous intervalles dans la direction  $x_2$ , où  $N_1$  et  $N_2$  sont des entiers naturels donnés. On note  $x_{1i} = i h_1$ ,  $h_1 = \frac{1}{N_1+1}$ ,  $i = 0, \dots, N_1+1$ , et  $x_{2j} = j h_2$ ,  $h_2 = \frac{1}{N_2+1}$ ,  $j = 1, \dots, N_2+1$

les points ainsi définis et on note  $T_k$  les  $(N_1+1) \times (N_2+1) =: N_T$

rectangles associés :



$$\begin{aligned} N_1 &= 5 \\ N_2 &= 11 \end{aligned}$$

Remarquons que pour cette triangulation

- le nombre de rectangles est  $N_T := (N_1+1)(N_2+1)$
- le nombre de sommets est  $N_S := (N_1+2)(N_2+2)$
- le nombre de sommets situés sur le bord est  $N_B := 2(N_1+N_2)+4$
- le nombre de sommets situés à l'intérieur est  $N_I := N_1 N_2 = N_S - N_B$

On considère comme ensemble  $V_h$  l'ensemble des fonctions continues sur  $[0,1] \times [0,1]$ , qui s'annulent sur le bord de  $\Omega$ , et dont la restriction sur chaque  $T_k$  est un polynôme de degré inférieur ou égal à 1 par rapport à chacune des deux variables  $x_1$  et  $x_2$ . On note  $\mathcal{P}_1$  cet espace vectoriel de polynômes :

$$\mathcal{P}_1 = \left\{ P(x_1, x_2) = a + b x_1 + c x_2 + d x_1 x_2, a \in \mathbb{R}, b \in \mathbb{R}, c \in \mathbb{R}, d \in \mathbb{R} \right\}$$

On considère donc

$$V_h = \left\{ v_h \in \bar{V}_h, v_h|_{\partial\Omega} = 0 \right\}$$

avec  $\bar{V}_h = \left\{ v_k \in \mathcal{C}^0(\bar{\Omega}), \forall k=1, \dots, N_T, v_k|_{T_k} \in \mathcal{P}_1 \right\}$ .

Sur chacun des rectangles  $[x_{1i}, x_{1i+1}] \times [x_{2j}, x_{2j+1}]$ ,  $i=0, \dots, N_1$ ,  $j=0, \dots, N_2$ ,

toute fonction  $v \in V_h$  s'écrit donc sous la forme

$$v(x_1, x_2) = a_{ij} + b_{ij} x_1 + c_{ij} x_2 + d_{ij} x_1 x_2$$

où les  $a_{ij}, b_{ij}, c_{ij}, d_{ij}$  dépendent du rectangle considéré.

On remarque en particulier que ces coefficients sont entièrement déterminés par la donnée des valeurs de la fonction  $v$  aux 4 coins du rectangle. En effet, si on se donne

$$v_{11i} := v(x_{1i}, x_{2i}), \quad v_{12j} := v(x_{1i}, x_{2j+1}), \quad v_{21j} := v(x_{1i+1}, x_{2j}), \quad v_{22j+1} := v(x_{1i+1}, x_{2j+1}),$$

les coefficients  $a_{ij}, b_{ij}, c_{ij}, d_{ij}$  vérifient le système suivant 511

$$\begin{cases} V_{ij} = a_{ij} + b_{ij} x_{1i} + c_{ij} x_{2j} + d_{ij} x_{1i} x_{2j} \\ V_{ij+1} = a_{ij} + b_{ij} x_{1i} + c_{ij} x_{2j+1} + d_{ij} x_{1i} x_{2j+1} \\ V_{i+1j} = a_{ij} + b_{ij} x_{1i+1} + c_{ij} x_{2j} + d_{ij} x_{1i+1} x_{2j} \\ V_{i+1j+1} = a_{ij} + b_{ij} x_{1i+1} + c_{ij} x_{2j+1} + d_{ij} x_{1i+1} x_{2j+1} \end{cases}$$

La matrice associée à ce système est la matrice  $4 \times 4$

$$A = \begin{pmatrix} 1 & x_{1i} & x_{2j} & x_{1i} x_{2j} \\ 1 & x_{1i} & x_{2j+1} & x_{1i} x_{2j+1} \\ 1 & x_{1i+1} & x_{2j} & x_{1i+1} x_{2j} \\ 1 & x_{1i+1} & x_{2j+1} & x_{1i+1} x_{2j+1} \end{pmatrix}$$

On a

$$\det A = \begin{vmatrix} 1 & x_{1i} & x_{2j} & x_{1i} x_{2j} \\ 0 & 0 & h_2 & x_{1i} h_2 \\ 0 & h_1 & 0 & x_{2j} h_1 \\ 0 & h_1 & h_2 & x_{1i+1} x_{2j+1} - x_{1i} x_{2j} \end{vmatrix} = \begin{vmatrix} 0 & h_2 & x_{1i} h_2 \\ h_1 & 0 & x_{2j} h_1 \\ h_1 & h_2 & x_{1i+1} x_{2j+1} - x_{1i} x_{2j} \end{vmatrix} = \begin{vmatrix} 0 & h_2 & x_{1i} h_2 \\ h_1 & 0 & x_{2j} h_1 \\ 0 & h_2 & x_{1i+1} x_{2j+1} - x_{1i} x_{2j} - x_{2j} h_1 \end{vmatrix}$$

$$= -h_1 h_2 [x_{1i+1} x_{2j+1} - x_{1i} x_{2j} - x_{2j} h_1 - x_{1i} h_2]$$

$$= -h_1 h_2 [x_{1i+1} \frac{h_2}{2} + x_{2j+1} \frac{h_1}{2} - x_{2j} h_1 - x_{1i} h_2]$$

$$= -h_1 h_2 [ \frac{h_2 h_1}{2} + \frac{h_1 h_2}{2} ]$$

$$= -h_1^2 h_2^2 \neq 0 \quad (h_1 \neq 0 \text{ et } h_2 \neq 0)$$

Le déterminant est non nul, les coefficients  $a_{ij}, b_{ij}, c_{ij}, d_{ij}$  sont <sup>52,</sup>  
donc déterminés de manière unique par la connaissance des valeurs  
(aux nœuds)  $v_{ij}, v_{ij+1}, v_{i+1j}, v_{i+1j+1}$ .

Toute fonction de  $\bar{V}_h$  est donc déterminée par la connaissance  
de ses valeurs aux  $N_S$  sommets, et ceci de manière unique.

Remarque Le raccord au niveau d'une arête commune  
à deux rectangles est nécessairement continu. En effet, si  
on considère par exemple l'arête  $[(x_{i1}, x_{ij}), (x_{i1}, x_{j+1})]$   
sur cette arête,  $x_1$  étant fixe à  $x_{i1}$ , les polynômes de  $\mathcal{D}_1$   
associés aux rectangles  $[x_{i1}, x_{i1}] \times [x_{ij}, x_{j+1}]$  et  $[x_{i1}, x_{i1}] \times [x_{j1}, x_{j1}]$   
sont en fait des polynômes habituels de degré 1 (dépendant  
d'une seule variable). Puisque ces polynômes ont la même  
valeur aux 2 points  $(x_{i1}, x_{ij})$  et  $(x_{i1}, x_{j+1})$ , ils sont nécessairement  
égaux sur l'arête toute entière.

Les  $N_S$  sommets sont appelés les degrés de liberté des fonctions de  
 $\bar{V}_h$ . Concernant les fonctions de  $V_h$ , on impose que la valeur  
d'une fonction de cet ensemble soit 0 sur les  $N_f$  sommets  
situés sur le bord de sorte que (en suivant la même raison-  
nement que dans la remarque précédente) les fonctions de  $V_h$   
valent bien 0 sur  $\partial\Omega$  tout entier.

Les degrés de liberté des fonctions de  $V_h$  sont donc les  $N_i = N_1 N_2$   
sommets situés à l'intérieur de  $\Omega$ . En d'autres termes toute  
fonction de  $V_h$  est uniquement déterminée par la connaissance

de ses valeurs aux  $N_i$  sommets intérieurs. Inversement, la connaissance des valeurs d'une fonction de  $V_h$  aux  $N_i$  sommets intérieurs permet de déterminer de manière unique cette fonction sur  $\bar{\Omega} = [0,1] \times [0,1]$  tout entier.

On est alors tenté, de manière similaire au cas d'une seule dimension d'espace, d'introduire les fonctions de base  $w_h^{ij}$ ,  $i=1, \dots, N_1$ ,  $j=1, \dots, N_2$  appartenant à  $V_h$  et telles que

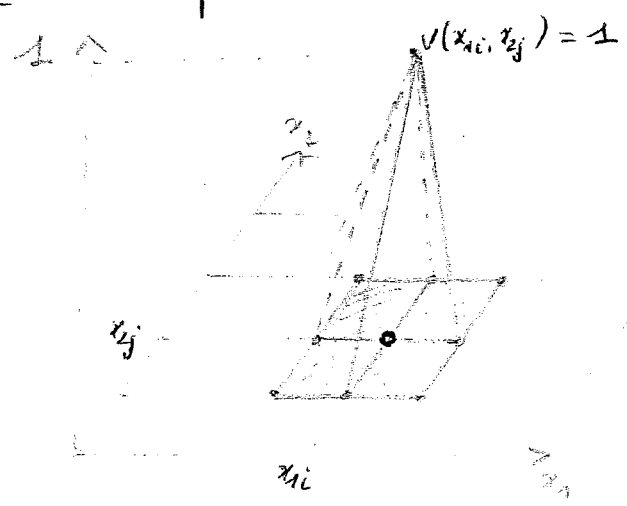
$$w_h^{ij}(x_{1k}, x_{2l}) = \begin{cases} 1 & \text{si } k=i \text{ et } l=j \\ 0 & \text{sinon} \end{cases}$$

Ces fonctions de base appartiennent bien à  $V_h$  et on remarque que tout élément de  $V_h$  peut s'écrire de manière unique comme une combinaison linéaire de ces fonctions de base

$$v(x_1, x_2) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} v(x_{1i}, x_{2j}) w_h^{ij}(x_1, x_2)$$

$\forall (x_1, x_2) \in [0,1] \times [0,1]$   
 $\forall v \in V_h$

La famille  $\{ w_h^{ij}, i=1, \dots, N_1, j=1, \dots, N_2 \}$  forme donc une base de  $V_h$  qui est donc un espace vectoriel de dimension  $N_i = N_1 N_2$ .



Il reste à vérifier que  $V_h \subset H_0^1(\mathcal{J}_{0,1} \mathbb{C} \times \mathcal{J}_{0,1} \mathbb{C})$

Lemme

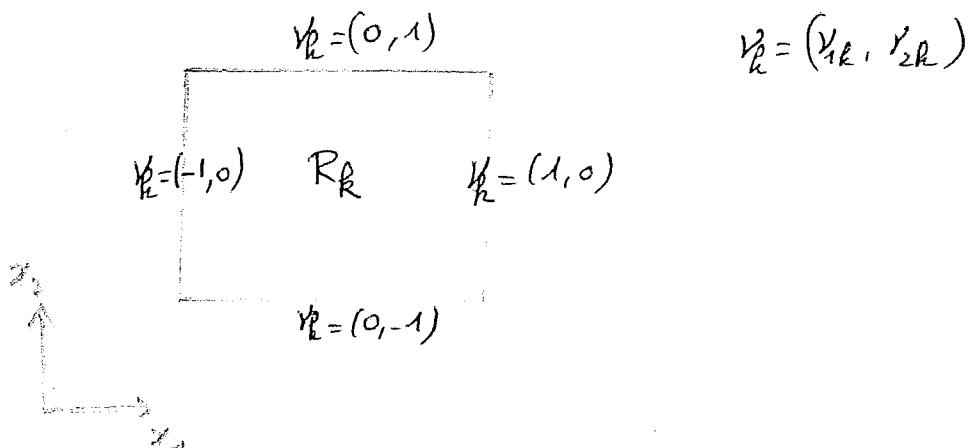
Pour tout  $N_1$  et tout  $N_2$ , dans  $\mathcal{W}^*$ ,  $V_h \subset H_0^1(\mathcal{J}_{0,1} \mathbb{C} \times \mathcal{J}_{0,1} \mathbb{C})$

Démonstration

Il faut montrer que  $\forall v \in V_h$ ,  $v \in L^2$ ,  $v' \in L^2$  et  $v|_{\partial\Omega} = 0$  avec  $\Omega = \mathcal{J}_{0,1} \mathbb{C} \times \mathcal{J}_{0,1} \mathbb{C}$ . Ce dernier point est vrai par construction puisque  $V_h \subset \mathcal{C}^\infty(\mathcal{J}_{0,1} \mathbb{C} \times \mathcal{J}_{0,1} \mathbb{C})$  et puisqu'on impose aux fonctions d'être nulles sur les  $N_T$  sommets situés sur  $\partial\Omega$  (ou à un plus haut que cela impliquerait  $v|_{\partial\Omega} = 0 \forall v \in V_h$ ). Ce dernier point aussi est vrai par continuité des fonctions de  $V_h$  sur  $\mathcal{J}_{0,1} \mathbb{C} \times \mathcal{J}_{0,1} \mathbb{C}$ . Il reste donc le second point comme dans le cas 1D soit  $v \in V_h$  et  $\varphi \in \mathcal{D}(\Omega)$

$$\begin{aligned} \left\langle \frac{\partial v}{\partial x_i}, \varphi \right\rangle &= - \left\langle v, \varphi'_i \right\rangle = - \iint_{\Omega} v(x_1, x_2) \varphi'_i(x_1, x_2) dx_1 dx_2 \\ &= - \sum_{k=1}^{N_T} \int_{T_k} v(x_1, x_2) \varphi'_i(x_1, x_2) dx_1 dx_2 \quad \text{avec } \varphi'_i := \frac{\partial \varphi}{\partial x_i} \\ &= - \sum_{k=1}^{N_T} \int_{T_k} v(x_1, x_2) \frac{\partial \varphi}{\partial x_i}(x_1, x_2) dx_1 dx_2 \\ &= \sum_{k=1}^{N_T} \int_{T_k} \frac{\partial v}{\partial x_i}(x_1, x_2) \varphi(x_1, x_2) dx_1 dx_2 - \sum_{k=1}^{N_T} \int_{\partial T_k} (v \varphi \nu_{i,k})(x_1, x_2) dx_1 dx_2 \end{aligned}$$

où  $\nu_{i,k}$  représente la composante d'indice  $i$  de la normale  $\nu_k$  à  $\partial T_k$  orientée vers l'extérieur de  $T_k$



Regardons de plus près la quantité

55/

$$S := \sum_{k=1}^{N_1} \int_{\partial T_k} (v \Psi v_{i_k}) (x_1, x_2) dx_1 dx_2$$

Il est clair que  $S$  se décompose en somme d'intégrales sur des arêtes de rectangles. Plus précisément, soit l'arête considérée est sur le bord de  $\Omega$  et dans ce cas  $\Psi = 0$  sur cette arête et donc la contribution intégrale associée est nulle, soit l'arête n'est pas située sur le bord et dans ce cas elle est commune à deux éléments  $T_k$  et  $T_{k'}$ . Puisque sur cette arête les produits  $v \Psi$  provenant de  $T_k$  et  $T_{k'}$  sont identiques par continuité, et que  $v_{i_{k'}} = -v_{i_k}$ , les deux contributions intégrales associées à une telle arête se compensent entre elles pour donner 0. On a finalement  $S = 0$ .

On a donc montré que

$$\langle \frac{\partial v}{\partial x_i}, \Psi \rangle = \sum_{k=1}^{N_1} \int_{T_k} \frac{\partial v}{\partial x_i} (x_1, x_2) \Psi (x_1, x_2) dx_1 dx_2$$

où rappelons-le,  $\frac{\partial v}{\partial x_i}$  à l'intérieur de l'intégrale coïncide avec la dérivée classique de la restriction de  $v$  à  $T_k$ . La distribution  $\frac{\partial v}{\partial x_i}$

coïncide donc avec la fonction  $\left\{ \frac{\partial v}{\partial x_i} \right\}$  définie par

$$\left\{ \frac{\partial v}{\partial x_i} \right\} (x_1, x_2) = \frac{\partial v}{\partial x_i} (x_1, x_2) \quad \forall (x_1, x_2) \in T_k$$

qui est clairement dans  $L^2([0,1] \times [0,1] \subset \mathbb{C})$ , ce qui termine la démonstration

### Calcul de la solution discrète

La solution approchée se trouve dans  $V_h$  et la déterminer signifie donc déterminer ses coordonnées dans la base  $w^{ij}$ ,  $i=1, \dots, N_1$ ,  $j=1, \dots, N_2$  chacun des éléments de cette base étant associé à un sommet interne. On notera dans ce paragraphe  $w^{k(i,j)} \equiv w^k$  ces éléments et  $u_h$  les coordonnées de  $u_h$  (la solution approchée) dans cette base :



$$u_h = \sum_{k=1}^{N_i} u_{hk} w^k$$

56/

Pour connaître  $u_h$ , il suffit donc de calculer la matrice  $A_h$  et le second membre  $B_h$  du système linéaire  $A_h U_h = B_h$  avec

$$U_h = \begin{pmatrix} u_{h1} \\ \vdots \\ u_{hN_i} \end{pmatrix}$$

On a

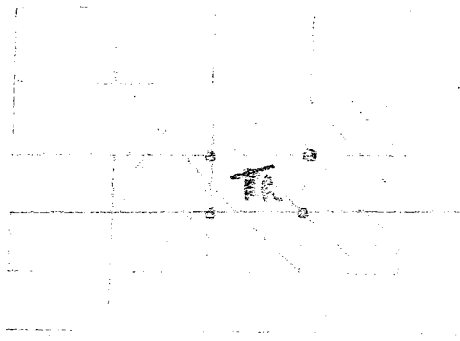
$$A_{hij} = \int_{\Omega} \nabla w^i \cdot \nabla w^j + c w^i w^j = \sum_{k=1}^{N_T} A_{hij}(T_k)$$

$$\text{avec } A_{hij}(T_k) = \int_{T_k} \nabla w^i \cdot \nabla w^j + c w^i w^j$$

$$\text{et } B_{hi} = \int_{\Omega} f w^i = \sum_{k=1}^{N_T} B_{hi}(T_k)$$

$$\text{avec } B_{hi}(T_k) = \int_{T_k} f w^i$$

Cette écriture montre que le calcul des coefficients de la matrice et du second membre se ramène à une somme de contributions élémentaire sur chacun des rectangles de la triangulation.

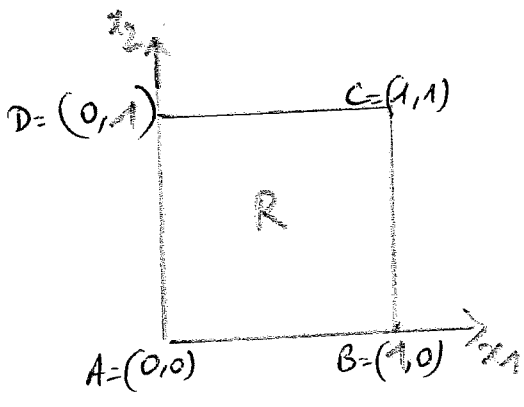


On remarque également que pour un rectangle donné  $T_k$ , n'interviennent dans le calcul effectif des  $A_{hij}(T_k)$  que les indices  $i$  et  $j$  associés à

des sommets de  $T_k$  : la contribution  $A_{k,ij}(T_k)$  est tout de suite nulle si l'un au moins des indices  $i$  et  $j$  ne correspond pas à un sommet de  $T_k$ . Cette remarque est valable aussi pour le calcul de  $B_{h,i}(T_k)$  :  $B_{h,i}(T_k) = 0$  si l'indice ne correspond pas à un sommet de  $T_k$ .

La somme des contributions élémentaires  $A_{k,ij}(T_k)$  et  $B_{h,i}(T_k)$  permettant de calculer  $A_{k,ij}$  et  $B_{h,i}$  s'appelle l'assemblage.

Il est souvent plus simple, puisque l'ensemble des calculs à effectuer se ramène à des intégrales sur un rectangle  $T_k$ , de faire le calcul une fois sur un rectangle de référence, par exemple  $[0,1] \times [0,1]$ , puis de transposer les calculs par homothétie et translation à tous les rectangles de la triangulation. C'est ce que nous allons faire.



Soit  $R = [0,1] \times [0,1]$  le rectangle de référence. Calculons les fonctions de base  $w^A, w^B, w^C, w^D$  associées aux points  $A, B, C$  et  $D$  sur cet élément. Par définition, les  $w^\alpha$  sont des polynômes de la forme  $a + bx_1 + cx_2 + dx_1x_2$  tels que  $w^\alpha(\alpha) = 1$  et  $w^\alpha(\beta) = 0$  pour  $\beta \neq \alpha, \beta = A, B, C, D, \alpha = A, B, C, D$ .  
Faisons le calcul pour  $w^A$  par exemple.

On cherche  $w^A(x_1, x_2) = a + bx_1 + cx_2 + dx_1x_2$   
 avec  $a, b, c, d$  tels que

$$\begin{aligned} w^A(0,0) &= a = 1 \\ w^A(1,0) &= a+b = 0 \\ w^A(1,1) &= a+b+c+d = 0 \\ w^A(0,1) &= a+c = 0 \end{aligned}$$

de sorte que  $a=1, b=-1, c=-1, d=1$ . On a donc

$$w^A(x_1, x_2) = 1 - x_1 - x_2 + x_1x_2$$

ce qui donne sous forme factorisée

$$w^A(x_1, x_2) = (1-x_1)(1-x_2)$$

De la même manière on trouve

$$w^B(x_1, x_2) = x_1(1-x_2)$$

$$w^C(x_1, x_2) = x_1x_2$$

$$w^D(x_1, x_2) = (1-x_1)x_2$$

La connaissance des valeurs de  $w \in \mathcal{D}_1$  aux points  $A, B, C, D$  a permis de déterminer  $w$  uniquement.  
 On vient sans le savoir de définir ce que l'on appelle un élément fini unisolvant, dont on donne maintenant une définition rigoureuse.

Définition (élément fini unisolvant)

Soit  $T$  une partie compacte connexe et d'intérieur non vide du plan. Soit  $\mathcal{P}(T)$  un espace vectoriel de dimension finie formé de polynômes définis sur  $T$  et à valeurs réelles.  
 Soit  $\Sigma(T)$  un espace vectoriel de dimension finie  $d$  formé de formes linéaires définies sur  $\mathcal{P}(T)$ , et dont une base est notée  $\varphi_1, \dots, \varphi_d$  :  $\Sigma(T) = \text{Vect} \{ \varphi_1, \dots, \varphi_d \}$   
 $\varphi_i : \mathcal{P}(T) \rightarrow \mathbb{R}, \quad \forall i = 1, \dots, d$   
 $p \rightarrow \varphi_i(p)$

On dit que le triplet  $(T, P, \Sigma)$  est unisolvant si et seulement si  $\forall (d_1, \dots, d_d) \in \mathbb{R}$ , il existe un unique polynôme de  $P(T)$  tel que  $\varphi_i(p) = d_i \quad \forall i = 1, \dots, d$ . 59/

Dans l'exemple ci-dessus, on a

$$T = [0, 1] \times [0, 1]$$

$$P(T) = \left\{ p: T \rightarrow \mathbb{R} \right. \\ \left. (x_1, x_2) \mapsto a + bx_1 + cx_2 + dx_1x_2 \right. \\ \left. (a, b, c, d) \in \mathbb{R}^4 \right\} = \mathcal{P}$$

$$\Sigma = \text{Vect} \{ \varphi_A, \varphi_B, \varphi_C, \varphi_D \}$$

avec  $\varphi_A: P(T) \rightarrow \mathbb{R}$   
 $p \rightarrow p(A)$

$$\varphi_B: P(T) \rightarrow \mathbb{R}$$

$$p \rightarrow p(B)$$

$$\varphi_C: P(T) \rightarrow \mathbb{R}$$

$$p \rightarrow p(C)$$

$$\varphi_D: P(T) \rightarrow \mathbb{R}$$

$$p \rightarrow p(D)$$

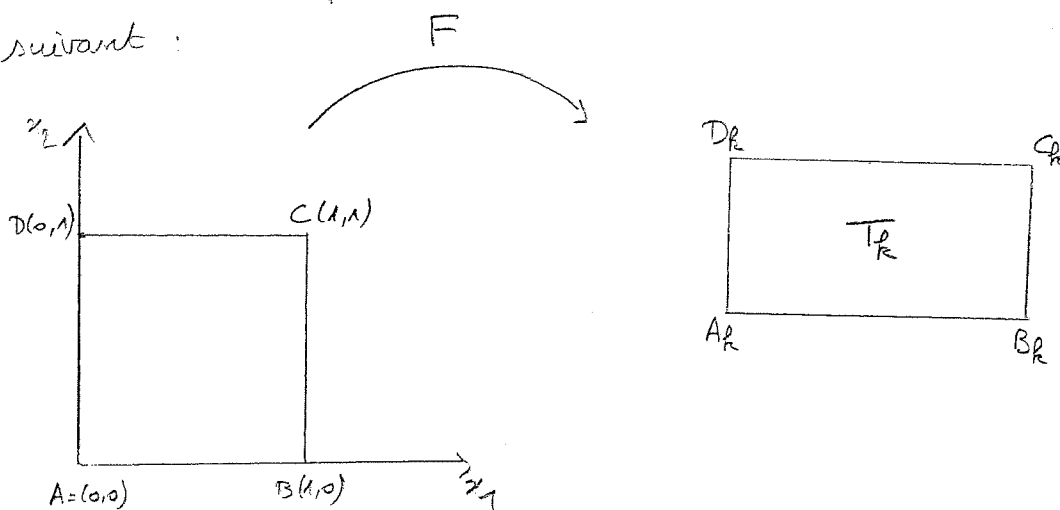
La connaissance des valeurs d'un polynôme de  $P(T)$  aux points  $A, B, C, D$  suffit en effet à le déterminer de manière unique

Remarque La famille  $\{ \varphi_A, \varphi_B, \varphi_C, \varphi_D \}$  est bien libre. En effet si on a  $x_A \varphi_A + x_B \varphi_B + x_C \varphi_C + x_D \varphi_D = 0$  alors en prenant successivement  $p = w^A, p = w^B, p = w^C$  et  $p = w^D$  on trouve  $x_A = 0, x_B = 0, x_C = 0, x_D = 0$ .

# Passage de l'élément de référence à un élément courant

60/

Il existe une unique transformation affine  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  permettant de passer de l'élément de référence à un élément courant  $T_R$  selon le schéma suivant :



Plus précisément, on a clairement

$$F_R(x_1, x_2) = A_R + x_1 \vec{A_R B_R} + x_2 \vec{A_R D_R}$$

Concernant les fonctions de base, on a (si on note  $w^{A_R}$  la fonction de base associée au point  $A_R$ ) :

$$\begin{cases} w^{A_R}(F_R(x_1, x_2)) = w^A(x_1, x_2) \\ w^{B_R}(F_R(x_1, x_2)) = w^B(x_1, x_2) \\ w^{C_R}(F_R(x_1, x_2)) = w^C(x_1, x_2) \\ w^{D_R}(F_R(x_1, x_2)) = w^D(x_1, x_2) \end{cases}$$

où encore

$$\begin{aligned} w^{A_R}(x_1, x_2) &= w^A(F_R^{-1}(x_1, x_2)) \\ w^{B_R}(x_1, x_2) &= w^B(F_R^{-1}(x_1, x_2)) \\ w^{C_R}(x_1, x_2) &= w^C(F_R^{-1}(x_1, x_2)) \\ w^{D_R}(x_1, x_2) &= w^D(F_R^{-1}(x_1, x_2)) \end{aligned}$$

$$\text{avec } F_R^{-1}(x_1, x_2) = \left( \frac{x_1 - x_1^{A_R}}{x_1^{B_R} - x_1^{A_R}}, \frac{x_2 - x_2^{A_R}}{x_2^{D_R} - x_2^{A_R}} \right)$$

## Exemple de calcul à partir de l'élément de référence

61,

Calculons par exemple  $A_{hii}(T_R)$ , en supposant que  $c(x) \equiv c \forall x$ .

$$A_{hii}(T_R) = \int_{T_R} \nabla w^i \cdot \nabla w^i + c w^i w^i$$

On suppose que l'indice  $i$  correspond par exemple au sommet  $C_h$ . En reprenant les notations ci-dessus, on a

$$A_{hii}(T_R) = \int_{T_R} \nabla w^{C_h} \cdot \nabla w^{C_h} + c w^{C_h} w^{C_h} dx$$

$$\text{avec } w^{C_h}(x_1, x_2) = W^c(F_R^{-1}(x_1, x_2))$$

$$\text{de sorte que } \nabla w^{C_h}(x_1, x_2) = \text{Jac}(F_R^{-1}(x_1, x_2))^t \nabla W^c(F_R^{-1}(x_1, x_2))$$

$$\text{On a ici } \text{Jac}(F_R^{-1}(x_1, x_2))^t = \begin{pmatrix} 1/h_1 & 0 \\ 0 & 1/h_2 \end{pmatrix} \quad \det \text{Jac} = 1/h_1 h_2$$

En faisant le changement de variable

$$(y_1, y_2) = F_R^{-1}(x_1, x_2) \quad \text{on a donc}$$

$$A_{hii}(T_R) = \int_0^1 \int_0^1 \left[ \left\| \begin{pmatrix} 1/h_1 & 0 \\ 0 & 1/h_2 \end{pmatrix} \nabla W^c(y_1, y_2) \right\|^2 + c (W^c)^2(y_1, y_2) \right] h_1 h_2 dy_1 dy_2$$

$$= \int_0^1 \int_0^1 \left[ \left\| \begin{pmatrix} 1/h_1 & 0 \\ 0 & 1/h_2 \end{pmatrix} \begin{pmatrix} y_2 \\ y_1 \end{pmatrix} \right\|^2 + c y_1^2 y_2^2 \right] h_1 h_2 dy_1 dy_2$$

$$= \int_0^1 \int_0^1 \left[ \frac{y_2^2}{h_1^2} + \frac{y_1^2}{h_2^2} + c y_1^2 y_2^2 \right] h_1 h_2 dy_1 dy_2$$

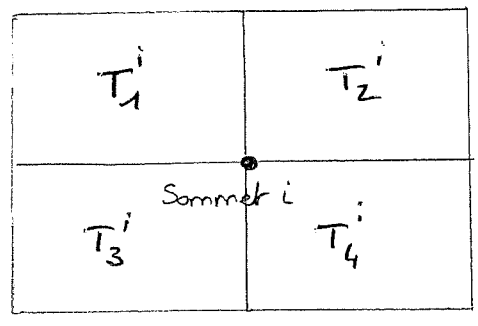
$$= \frac{h_1 h_2}{3} \left[ \frac{1}{h_1^2} + \frac{1}{h_2^2} \right] + h_1 h_2 \int_0^1 \int_0^1 y_1^2 y_2^2 dy_1 dy_2$$

$$= \frac{h_1 h_2}{3} \left[ \frac{1}{h_1^2} + \frac{1}{h_2^2} \right] + c \frac{h_1 h_2}{9}$$

Abons que ce résultat est indépendant de  $T_k$ . Des calculs similaires montrent que ce résultat est en réalité indépendant du sommet de  $T_k$  considéré. En utilisant les notations de la figure ci-dessus, on a donc

$$A_{h_{ii}}(T_1^i) = A_{h_{ii}}(T_2^i) = A_{h_{ii}}(T_3^i) = A_{h_{ii}}(T_4^i) = \frac{h_1 h_2}{3} \left[ \frac{1}{h_1^2} + \frac{1}{h_2^2} \right] + \frac{c h_1 h_2}{9}$$

Puisque les éléments  $T_1^i, T_2^i, T_3^i, T_4^i$  constituent le support de  $w^i$ , on a donc finalement  $A_{h_{ii}} = 4 \frac{h_1 h_2}{3} \left[ \frac{1}{h_1^2} + \frac{1}{h_2^2} \right] + 4 \frac{c h_1 h_2}{9}$ .

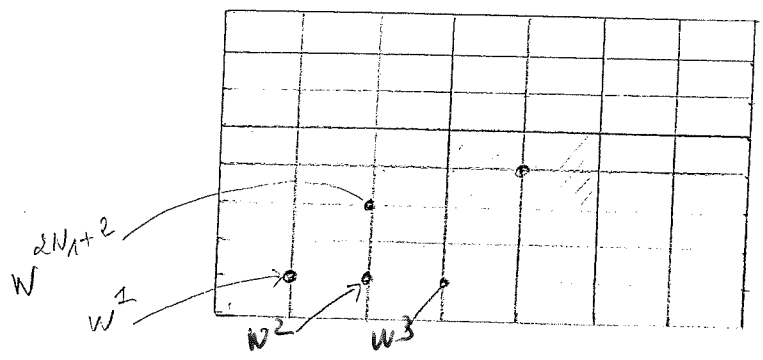


Support de  $w^i$

Structure de la matrice

Afin de préciser la structure de la matrice  $A_h$ , il convient au préalable de se fixer une numérotation des fonctions de base  $w^k, k=1, \dots, N_1 N_2$ . On choisit de numérotter ces fonctions de base en parcourant les sommets du maillage ligne par ligne, en commençant naturellement par la première :

$$\begin{aligned} k(1,1) &= 1 & k(1,2) &= N_1 + 1 \\ k(2,1) &= 2 & k(2,2) &= N_1 + 2 \\ & \vdots & & \vdots \\ k(N_1,1) &= N_1 & k(N_1,2) &= 2N_1 \end{aligned}$$

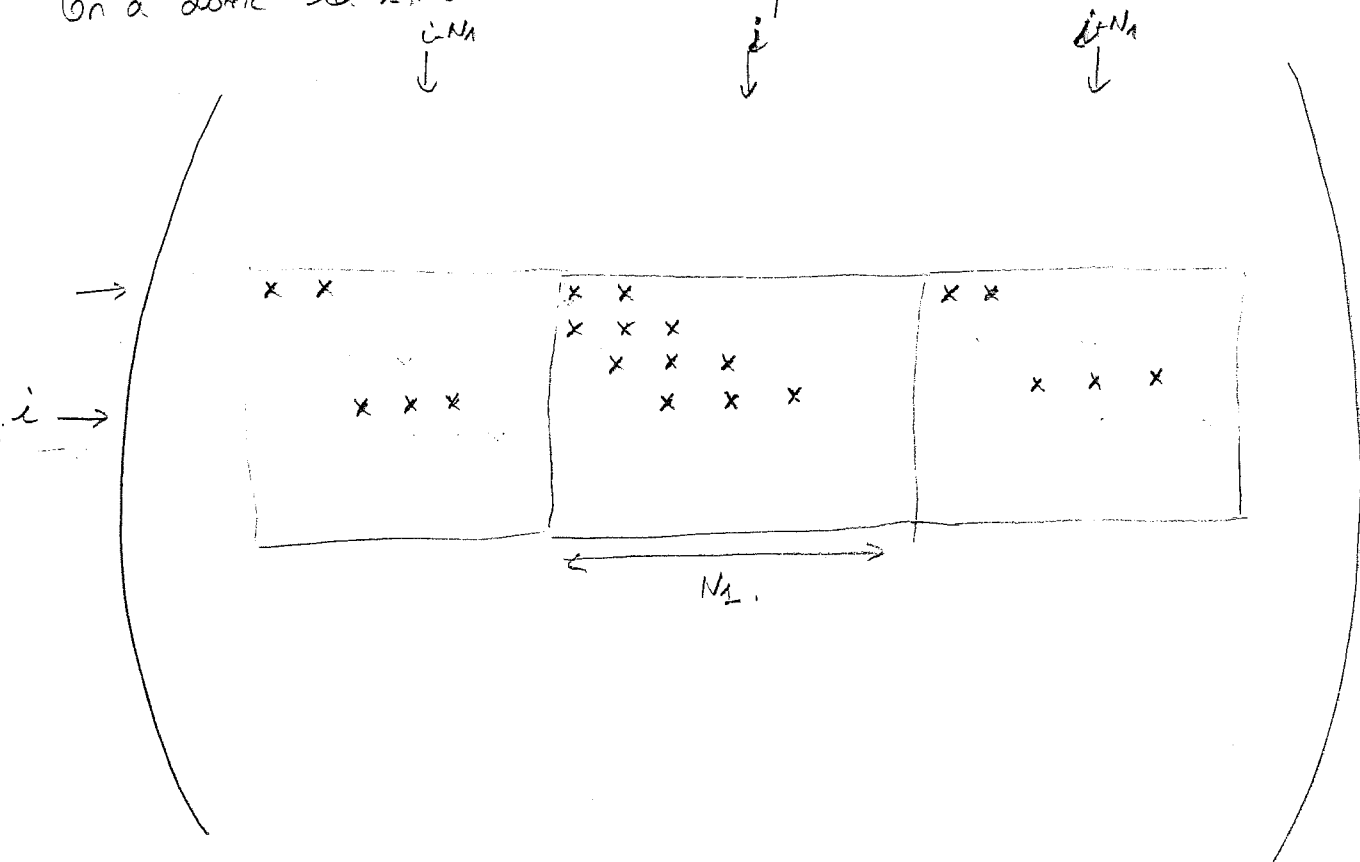


Avec cette convention de numérotation, on voit que les supports de 63,  
deux fonctions de base  $w^i$  et  $w^j$  sont non disjointes si et seulement

si

$j = i - 1$ ou $j = i$ ou $j = i + 1$	ou	$j = i + N_1$ $j = i - N_1$ $j = i + 1 + N_1$ $j = i + 1 - N_1$ $j = i - 1 - N_1$ $j = i - 1 + N_1$
---------------------------------------------	----	--------------------------------------------------------------------------------------------------------------------

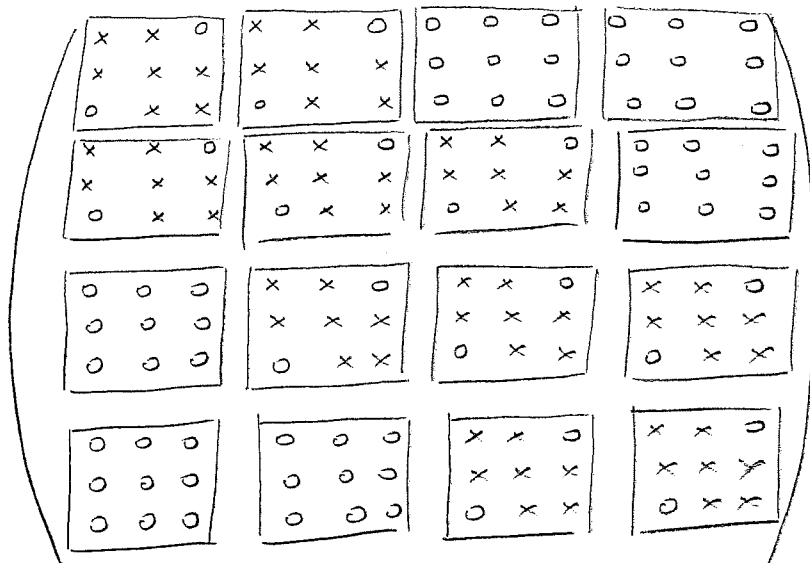
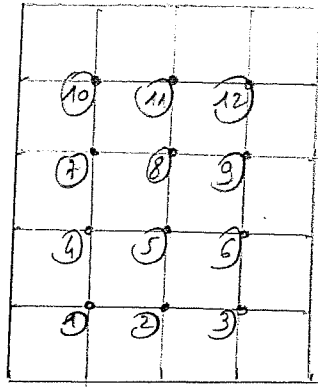
On a donc la structure suivante pour la matrice  $A_h$  :



On observe alors que la matrice est tridiagonale par blocs de tailles  $N_1 \times N_1$ .

On peut se convaincre plus facilement de cette structure en considérant des valeurs concrètes (et pas trop grandes !) de  $N_1$  et  $N_2$ . Par exemple  $N_1 = 3$  et  $N_2 = 4$ , ce qui génère déjà une matrice  $12 \times 12$  et qui correspond au maillage suivant :





Matrice tridiagonale  
par blocs de taille  $N \times N$

### Autres exemples d'éléments finis rectangulaires

Il est possible d'utiliser d'autres éléments plus précis, au sens où les polynômes utilisés sont de degré plus élevés. Il est tout à fait possible de mener sur ces nouveaux éléments des calculs similaires à ce qui vient d'être fait pour les éléments  $\mathcal{P}_1$ . Nous ne les détaillerons pas.

Il est important, pour les éléments que nous allons proposer, de se convaincre de deux points :

- 1) l'unissabilité de l'élément fini
- 2) la propriété  $V_h \subset V$  où  $V$  est l'espace variationnel, ici  $H_0^1(\mathcal{I}_0, \mathcal{I} \times \mathcal{I}_0, \mathcal{I})$

Élément fini  $\mathcal{Q}^2$

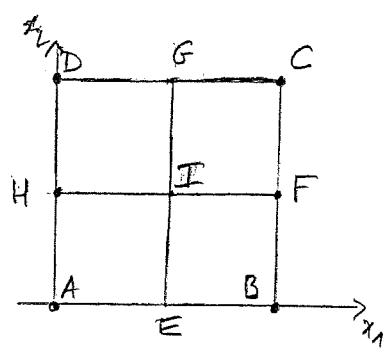
On se place pour simplifier sur l'élément de référence.

Soit  $T = [0, 1] \times [0, 1]$

Soit  $P(T) = \text{Vect} \{ 1, x_1, x_1^2, x_2, x_2^2, x_1 x_2, x_1^2 x_2, x_1 x_2^2, x_1^2 x_2^2 \} = \mathcal{Q}^2$

$\equiv$  l'ensemble des polynômes de degré  $\leq 2$  par rapport à chaque variable  $x_1$  et  $x_2$ , et définis sur  $T$ .

Pour définir  $\Sigma(T)$ , on utilise les notations suivantes



On pose alors  $\Sigma = \text{Vect} \{ \varphi_A, \varphi_B, \varphi_C, \varphi_D, \varphi_E, \varphi_F, \varphi_G, \varphi_H, \varphi_I \}$

avec  $\varphi_A: P(T) \rightarrow \mathbb{R}, P \rightarrow P(A)$        $\varphi_B: P(T) \rightarrow \mathbb{R}, P \rightarrow P(B)$        $\varphi_I: P(T) \rightarrow \mathbb{R}, P \rightarrow P(I)$ .

On montre que l'élément fini  $(T, P(T), \Sigma(T))$  est unisolvant.

On considère alors  $V_k = \{ v_k \in \bar{V}_k, v_k|_{\Omega} = 0 \}$

avec  $\bar{V}_k = \{ v_k \in \mathcal{C}^0(\bar{\Omega}), \forall k=1, \dots, N_T, v_k|_{T_k} \in \mathcal{Q}_2 \}$

$\mathcal{Q}_2 = \{ p(x_1, x_2) = a + b x_1 + c x_1^2 + d x_2 + e x_2^2 + f x_1 x_2 + g x_1^2 x_2 + h x_1 x_2^2 + i x_1^2 x_2^2 \}$

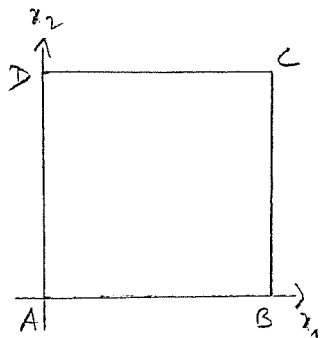
Élément fini  $\mathcal{Q}^3$

$T = [0, 1] \times [0, 1]$

$P(T) = \text{Vect} \{ 1, x_1, x_1^2, x_1^3, x_2, x_2^2, x_2^3, x_1^2 x_2, x_1 x_2^2, x_1^3 x_2, x_1 x_2^3, x_1^2 x_2^2, x_1^2 x_2^3, x_1^3 x_2^2, x_1^3 x_2^3 \}$

$= \mathcal{Q}^3$

dim  $\mathcal{Q}^3 = 16$ .



On pose ensuite

$$\Sigma = \text{Vect} \left\{ \begin{array}{cccc} \varphi_A^{00}, & \varphi_A^{01}, & \varphi_A^{10}, & \varphi_A^{11} \\ \varphi_B^{00}, & \varphi_B^{01}, & \varphi_B^{10}, & \varphi_B^{11} \\ \varphi_C^{00}, & \varphi_C^{01}, & \varphi_C^{10}, & \varphi_C^{11} \\ \varphi_D^{00}, & \varphi_D^{01}, & \varphi_D^{10}, & \varphi_D^{11} \end{array} \right\}$$

avec  $\varphi_x^{ij} = P(T) \rightarrow \mathbb{R}$   
 $P \rightarrow \frac{\partial^{i+j}}{\partial x_1^i \partial x_2^j} P(x)$

avec  $x = A, B, C, D$   
 $i = 0, 1$   
 $j = 0, 1$

Il y a donc 16 éléments de l'espace qui engendre  $\Sigma$ .

On montre que l'élément fini  $(T, P(T), \Sigma(T))$  est unisolvant.

A partir de cet élément on construit les espaces discrets  $V_h$  et  $\bar{V}_h$  en posant

$$V_h = \{ v_h \in \bar{V}_h, v_h|_{\partial\Omega} = 0 \}$$

avec  $\bar{V}_h = \{ v_h \in \mathcal{G}^1(\bar{\Omega}), \forall k=1, \dots, N_T, v_h|_{T_k} \in \mathcal{P}_3 \}$

De manière intéressante, on montre que  $\bar{V}_h \subset H_2(\mathbb{I}_0, 1\mathbb{C} \times \mathbb{I}_0, 1\mathbb{C})$   
 et donc  $V_h \subset H_0^1(\mathbb{I}_0, 1\mathbb{C} \times \mathbb{I}_0, 1\mathbb{C})$

## Éléments finis triangulaires

Nous proposons de considérer dans ce paragraphe des éléments finis triangulaires. Nous aurons toujours en tête l'application au même problème modèle, avec donc en particulier  $\Omega = \mathbb{I}_0, 1\mathbb{C} \times \mathbb{I}_0, 1\mathbb{C}$ , mais il est important d'avoir à l'esprit que l'intérêt des éléments finis triangulaires réside dans le fait qu'ils permettent de traiter naturellement des domaines  $\Omega$  plus généraux, notamment de frontière polygonale par morceaux.

## Éléments P1

On utilise les mêmes notations que dans le cas rectangulaire. Plus précisément, on partitionne le carré  $\Omega = ]0,1[ \times ]0,1[$  à l'aide de  $N_T$  triangles.

On note  $q^{(i)}$  les points du maillage ainsi formé, pour  $i = 1, \dots, N_S$ , c'est-à-dire les  $N_S$  sommets de tous les triangles de la triangulation. Parmi ces  $N_S$  sommets,  $N_p$  sont situés sur le bord de  $\Omega$ , et  $N_i = N_S - N_p$  sont situés à l'intérieur de  $\Omega$ .

On note  $P_1$  l'espace vectoriel de polynômes de degré inférieur ou égal à un, i.e.

$$P_1 = \left\{ P(x_1, x_2) = a + bx_1 + cx_2, a \in \mathbb{R}, b \in \mathbb{R}, c \in \mathbb{R} \right\}$$

et on considère les espaces  $V_h$  et  $\bar{V}_h$  suivants

$$V_h = \left\{ v_h \in \bar{V}_h, v_h|_{\partial\Omega} = 0 \right\}$$

$$\bar{V}_h = \left\{ v_h \in C_0(\Omega), \forall k = 1, \dots, N_T, v_h|_{T_k} \in P_1 \right\}$$

Proposition Soit  $T = \left\{ (x_1, x_2), 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, x_1 + x_2 \leq 1 \right\}$

le triangle de référence ABC avec  $A = (0,0), B = (1,0), C = (0,1)$ . Soit  $P(T)$

l'ensemble  $P(T) = \left\{ p: T \rightarrow \mathbb{R}, (x_1, x_2) \rightarrow a + bx_1 + cx_2, a \in \mathbb{R}, b \in \mathbb{R}, c \in \mathbb{R} \right\}$

et  $\Sigma = \text{Vect} \{ \varphi_A, \varphi_B, \varphi_C \}$  avec  $\varphi_A: P(T) \rightarrow \mathbb{R}, p \rightarrow p(A)$      $\varphi_B: P(T) \rightarrow \mathbb{R}, p \rightarrow p(B)$      $\varphi_C: P(T) \rightarrow \mathbb{R}, p \rightarrow p(C)$

Alors le triplet  $(T, P, \Sigma)$  est unisolvant.

Démonstration Soit  $\alpha_A, \alpha_B$  et  $\alpha_C$  trois réels quelconques et soit  $p \in P(T)$ , i.e.

$p(x_1, x_2) = a + bx_1 + cx_2$ . Il s'agit de déterminer de manière unique les coefficients  $a, b$  et  $c$  tels que  $\begin{cases} p(0,0) = a = \alpha_A \\ p(1,0) = a+b = \alpha_B \\ p(0,1) = a+c = \alpha_C \end{cases}$

On trouve immédiatement  $\begin{cases} a = \alpha_A \\ b = \alpha_B - \alpha_A \\ c = \alpha_C - \alpha_A \end{cases}$  i.e.  $p(x_1, x_2) = \alpha_A + (\alpha_B - \alpha_A)x_1 + (\alpha_C - \alpha_A)x_2$

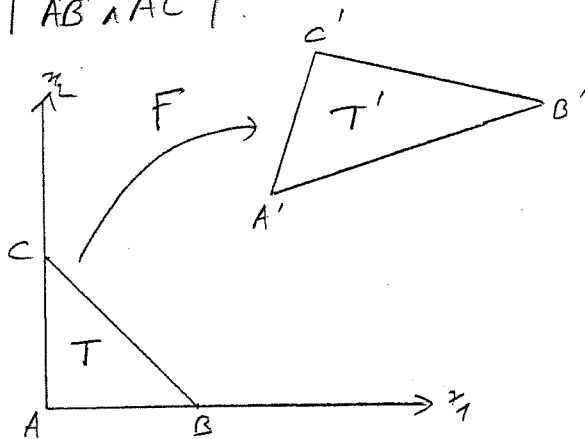
L'élément fini est donc unisolvant.

Remarque: L'univocance reste bien entendu vraie pour tout triangle 63/

$T = A'B'C'$  quelconque, i.e. avec  $A' = (x'_A, y'_A)$ ,  $B' = (x'_B, y'_B)$ ,  $C' = (x'_C, y'_C)$ , pourvu qu'il ne soit pas dégénéré, i.e. pourvu que les points  $A, B$  et  $C$  ne soient pas alignés, ou encore pourvu que  $\vec{A'B'} \wedge \vec{A'C'} = \pm 2A(T)$  soit non nulle. Rappelons en effet que

$$\vec{A'B'} = \begin{pmatrix} x'_B - x'_A \\ y'_B - y'_A \end{pmatrix}, \quad \vec{A'C'} = \begin{pmatrix} x'_C - x'_A \\ y'_C - y'_A \end{pmatrix}, \quad \vec{A'B'} \wedge \vec{A'C'} = (x'_B - x'_A)(y'_C - y'_A) - (y'_B - y'_A)(x'_C - x'_A)$$

$$A(A'B'C') = \frac{1}{2} |\vec{A'B'} \wedge \vec{A'C'}|$$



Ainsi, sous ces conditions, l'application

$$F(x_1, x_2) = A' + x_1 \vec{A'B'} + x_2 \vec{A'C'}$$

qui envoie  $T$  sur  $T'$  est inversible et on a

$$F^{-1}(y_1, y_2) = \left( \frac{y_1 - x'_A}{x'_B - x'_A}, \frac{y_2 - y'_A}{y'_C - y'_A} \right) M^{-1} \quad \text{avec } M = \begin{pmatrix} x'_B - x'_A & y'_B - y'_A \\ x'_C - x'_A & y'_C - y'_A \end{pmatrix}$$

Notons en effet que  $|\det M| = 2A(A'B'C')$ .

Il s'agit donc de montrer, pour se convaincre de l'univocance, qu'il existe un unique polynôme  $p' \in \mathcal{P}_1$  t.q.

$$\begin{cases} p'(A') = \alpha_{A'} \\ p'(B') = \alpha_{B'} \\ p'(C') = \alpha_{C'} \end{cases} \quad \text{et ceci } \forall \alpha_{A'}, \alpha_{B'}, \alpha_{C'}$$

$$\text{i.e. } \begin{cases} p'(F(A)) = \alpha_{A'} \\ p'(F(B)) = \alpha_{B'} \\ p'(F(C)) = \alpha_{C'} \end{cases}$$

Posons  $p'(x_1, x_2) = p(F^{-1}(x_1, x_2))$

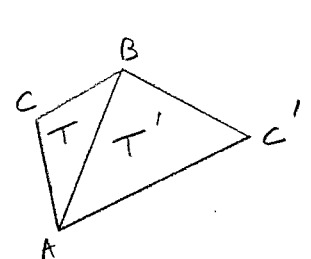
ou  $p(x_1, x_2) = \alpha_{A'} + (\alpha_{B'} - \alpha_{A'})x_1 + (\alpha_{C'} - \alpha_{A'})x_2$  est l'unique polynôme de  $\mathbb{P}_1$  t. q

$p(A) = \alpha_{A'}$ ,  $p(B) = \alpha_{B'}$ ,  $p(C) = \alpha_{C'}$  (voir preuve ci-dessus)

Il est clair que  $p'$  ainsi défini appartient à  $\mathbb{P}_1$  et qu'il vérifie les contraintes imposées. Il est unique par unicité de  $p$ .

L'unicovariance nous permet de dire que toute fonction de  $\mathbb{V}_n$  est entièrement déterminée par la connaissance de ses valeurs aux  $N_T$  sommets, et ceci de manière unique.

Définissons que la fonction ainsi définie est bien continue au niveau d'une arête commune à deux triangles  $T$  et  $T'$ .



- $A(x_{1A}, x_{2A})$
- $B(x_{1B}, x_{2B})$
- $C(x_{1C}, x_{2C})$
- $C'(x_{1C'}, x_{2C'})$

soit  $v$  t. q  $\forall k=1, \dots, N_T, \forall V_{Tk} \in \mathbb{P}_1$ .

On note  $p(x_1, x_2) = a x_1 + b x_2$  les polynômes de  $\mathbb{P}_1$  tels que  
et  $p'(x_1, x_2) = a' + b' x_1 + c' x_2$

$$\begin{aligned}
 p(A) = p'(A) &= v(A) & \text{et} & & p(C) &= v(C) \\
 p(B) = p'(B) &= v(B) & & & p'(C') &= v(C')
 \end{aligned}$$

⚠ le ' ne représente pas la dérivée !

sur le segment  $[AB]$ , on a  $x_2 = \frac{x_{2B} - x_{2A}}{x_{1B} - x_{1A}} (x - x_{1A}) + x_{2A}$

de sorte que sur ce segment, les polynômes  $p$  et  $p'$  peuvent être vus comme des polynômes de degré 1 en  $x_1$ , et ne dépendant plus de  $x_2$ . Puisqu'ils coïncident aux deux points  $A$  et  $B$ , ils coïncident donc sur tout le segment  $[AB]$  et on a donc la continuité du raccord entre  $p$  et  $p'$  sur l'arête  $[AB]$ .

De même que précédemment, les  $N_b$  sommets sont appelés les degrés de liberté des fonctions de  $V_h$ . Concernant les fonctions de  $V_h$  on impose que la valeur d'une fonction de cet ensemble soit 0 sur les  $N_f$  sommets situés sur le bord de  $\Omega$ , de sorte que en suivant le même raisonnement que celui qu'on vient de faire concernant la continuité au niveau d'une arête, les fonctions de  $V_h$  valent bien 0 sur  $\partial\Omega$  tout entier.

Les degrés de liberté des fonctions de  $V_h$  sont donc les  $N_i$  sommets situés à l'intérieur de  $\Omega$ . En d'autres termes, toute fonction de  $V_h$  est uniquement déterminée par la connaissance de ses valeurs aux  $N_i$  sommets intérieurs.

De même que précédemment, on introduit donc les fonctions de base  $w_h^i$ , définies sur  $\bar{\Omega}$  et appartenant à  $V_h$  et telles que

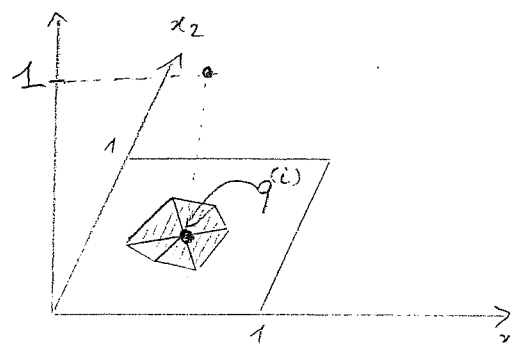
$$w_h^i(q^j) = \begin{cases} 1 & \text{si } i=j \\ 0 & \text{sinon} \end{cases} \quad \forall i=1, \dots, N_i$$

ou les  $q^i$  représentent les  $N_i$  sommets intérieurs.

Tout élément  $v \in V_h$  peut s'écrire de manière unique comme combinaison linéaire de ces  $N_i$  fonctions de base

$$v(x_1, x_2) = \sum_{i=1}^{N_i} v(q^i) w_h^i(x_1, x_2) \quad \forall (x_1, x_2) \in \bar{\Omega}$$

La famille  $\{w_h^i, i=1, \dots, N_i\}$  forme donc une base de  $V_h$  qui est donc un espace vectoriel de dimension  $N_i$ .



La partie hachurée correspond au support de la fonction de base  $q^i$ .

Il reste à vérifier que  $V_h \subset H_0^1(\Gamma_0, \Gamma \times \Gamma_0, \Gamma)$

7/1

Lemme

On a  $V_h \subset H_0^1(\Gamma_0, \Gamma \times \Gamma_0, \Gamma)$

La démonstration est laissée en exercice.

Éléments de calcul de la solution discrète

De même que précédemment, on a  $A_h U_h = B_h$  avec

$$A_{hij} = \int_{\Omega} \nabla w^i \cdot \nabla w^j + c w^i w^j = \sum_{k=1}^{N_T} A_{hij}(T_k)$$

$$\text{avec } A_{hij}(T_k) = \int_{T_k} \nabla w^i \cdot \nabla w^j + c w^i w^j$$

$$\text{et } B_{hi} = \int_{\Omega} f w^i = \sum_{k=1}^{N_T} B_{hi}(T_k)$$

$$\text{avec } B_{hi}(T_k) = \int_{T_k} f w^i$$

et on note que pour un triangle  $T_k$  donné,  $A_{hij}(T_k)$  et  $B_{hi}(T_k)$  sont automatiquement nulles si  $i$  et  $j$  ne sont pas associés à un sommet du triangle  $T_k$ .

De même que précédemment, il est souvent plus simple de considérer

comme triangle  $T_k$  le triangle de référence  $T = ABC$  avec  $A(0,0)$ ,  $B(1,0)$ ,  $C(0,1)$ ,

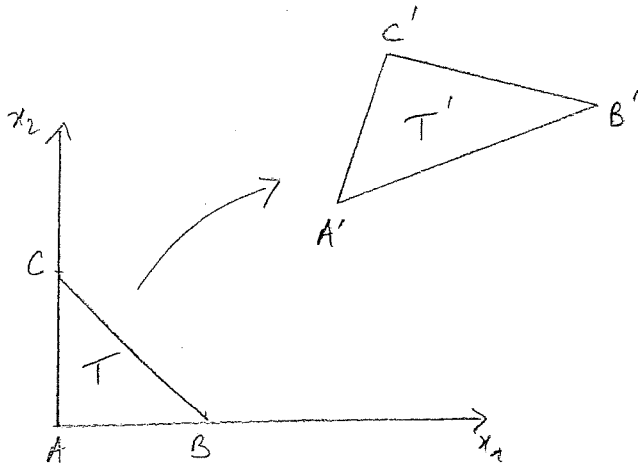
de calculer sur ce triangle les contributions  $A_{hij}(T)$  et  $B_{hi}(T)$  associées,

soit avec des notations claires  $A_{hAB}(T)$ ,  $A_{hAC}(T)$ ,  $A_{hCB}(T)$ ,  $A_{hAA}(T)$ ,

$A_{hBB}(T)$ ,  $A_{hCC}(T)$ ,  $B_{hA}(T)$ ,  $B_{hB}(T)$ ,  $B_{hC}(T)$ , et procéder ensuite au passage de l'élément de référence à un élément courant.



Exemple



Tout d'abord, on a clairement

$$W^A(x_1, x_2) = 1 - x_1 - x_2$$

$$W^B(x_1, x_2) = x_1$$

$$W^C(x_1, x_2) = x_2$$

$$\nabla W^A(x_1, x_2) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\nabla W^B(x_1, x_2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\nabla W^C(x_1, x_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Ainsi, en choisissant  $c = c_0$ , on a par exemple

$$A_{RAA}(T) = \iint_T [2 + c(1 - x_1 - x_2)^2] dx_1 dx_2$$

$$A_{RAA}(T) = 2A(T) + c_0 \iint (1 - x_1 - x_2)^2 dx_1 dx_2$$

$$A_{RAA}(T) = 2A(T) + c_0 \int_0^1 \left( \int_0^{1-x_2} (1 - x_1 - x_2)^2 dx_1 \right) dx_2$$

$$A_{RAA}(T) = 2A(T) + c_0 \int_0^1 \left[ -\frac{1}{3}(1 - x_1 - x_2)^3 \right]_0^{1-x_2} dx_2$$

$$A_{RAA}(T) = 2A(T) + c_0 \int_0^1 +\frac{1}{3}(1 - x_2)^3 dx_2$$

$$A_{RAA}(T) = 2A(T) + \frac{c_0}{3} \left[ -\frac{1}{4}(1 - x_2)^4 \right]_0^1$$

$$A_{RAA}(T) = 2A(T) + \frac{c_0}{12}$$

On procède de la même manière pour calculer les autres termes, ce qui donne par exemple pour  $A_{RBC}(T)$ :

$$\begin{aligned}
 A_{hBC}(T) &= \iint_T (0 + c_0 x_1 x_2) dx_1 dx_2 \\
 &= c_0 \int_0^1 \left( \int_0^{1-x_2} x_1 x_2 dx_1 \right) dx_2 \\
 &= c_0 \int_0^1 \left[ x_2 \frac{x_1^2}{2} \right]_0^{1-x_2} dx_2 \\
 &= c_0 \int_0^1 \frac{x_2}{2} (1-x_2)^2 dx_2 \\
 &= \frac{c_0}{2} \left[ \frac{x_2^4}{4} - 2 \frac{x_2^3}{3} + \frac{x_2^2}{2} \right]_0^1 \\
 &= \frac{c_0}{2} \left[ \frac{1}{4} - \frac{2}{3} + \frac{1}{2} \right] \\
 &= \frac{c_0}{24} [3 - 8 + 6] = \frac{c_0}{24}
 \end{aligned}$$

$$x_2(1-x_2)^2 = x_2^3 - 2x_2^2 + x_2$$

Et ainsi de suite...

Si l'on passe maintenant de l'élément de référence à un élément courant  $T' = A'B'C'$ , on a par exemple

$$A_{hB'C'}(T') = \iint_{T'} \nabla w^{B'} \nabla w^{C'} + c w^{B'} w^{C'} dx_1 dx_2$$

avec  $w^{B'}(x_1, x_2) = w^B(F^{-1}(x_1, x_2))$   
 $w^{C'}(x_1, x_2) = w^C(F^{-1}(x_1, x_2))$

$$\begin{aligned}
 \nabla w^{B'}(x_1, x_2) &= \text{Jac}(F^{-1}(x_1, x_2))^k \nabla w^B(F^{-1}(x_1, x_2)) \\
 \nabla w^{C'}(x_1, x_2) &= \text{Jac}(F^{-1}(x_1, x_2))^k \nabla w^C(F^{-1}(x_1, x_2))
 \end{aligned}$$

avec rappelons-le

$$F(x_1, x_2) = A' + x_1 \vec{A'B'} + x_2 \vec{A'C'}$$

$$F^{-1}(x_1, x_2) = (x_1 - x_A) \vec{e}_1 + (x_2 - y_A) \vec{e}_2 \Pi^{-1}$$

$$\text{Jac}(F^{-1}(x_1, x_2))^k = M^{-1}$$

$$|\det \Pi^{-1}|^k = \frac{1}{|\det M|} = \frac{1}{2A(A'B'C')}$$

et  $\Pi = \begin{pmatrix} x_{B'} - x_{A'} & y_{B'} - y_{A'} \\ x_{C'} - x_{A'} & y_{C'} - y_{A'} \end{pmatrix} \quad |\det \Pi| = 2A(A'B'C')$

Ainsi, en faisant le changement de variables  $(y_1, y_2) = F^{-1}(x_1, x_2)$ , on a 741

$$A_{B'C'}(T') = \iint_{T'} \left[ M^{-1} \nabla W^B(F^{-1}(x_1, x_2)) \cdot \Pi^{-1} \nabla W^C(F^{-1}(x_1, x_2)) + \epsilon_3 W^B(F^{-1}(x_1, x_2)) W^C(F^{-1}(x_1, x_2)) \right] dx_1 dx_2$$

$$= 2 A(A'B'C') \iint_T \left[ M^{-1} \nabla W^B(y_1, y_2) \cdot \Pi^{-1} \nabla W^C(y_1, y_2) + \epsilon_3 W^B(y_1, y_2) W^C(y_1, y_2) \right] dy_1 dy_2$$

Or  $\Pi^{-1} = \frac{1}{\det \Pi} \begin{pmatrix} y_{C'} - y_{A'} & -(x_{C'} - x_{A'}) \\ -(y_{B'} - y_{A'}) & x_{B'} - x_{A'} \end{pmatrix}$   $\nabla W^B(y_1, y_2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$   $\nabla W^C(y_1, y_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

de sorte que  $M^{-1} \nabla W^B(y_1, y_2) \cdot \Pi^{-1} \nabla W^C(y_1, y_2) = \frac{1}{(\det \Pi)^2} \begin{pmatrix} y_{C'} - y_{A'} \\ -(y_{B'} - y_{A'}) \end{pmatrix} \cdot \begin{pmatrix} -(x_{C'} - x_{A'}) \\ x_{B'} - x_{A'} \end{pmatrix}$

$$= \frac{-[(x_{C'} - x_{A'})(y_{C'} - y_{A'}) + (x_{B'} - x_{A'})(y_{B'} - y_{A'})]}{(\det \Pi)^2} =: 1$$

Cette quantité ne dépend pas de  $(y_1, y_2)$ .

Finalement, on obtient donc

$$A_{B'C'}(T') = 2 A(A'B'C') \left[ FA(T) + \frac{\epsilon_3}{24} \right]$$

Cette quantité dépend de la géométrie de l'élément courant  $A'B'C'$ .

# Éléments de programmation

En théorie, le calcul du coefficient  $A_{hij}$  de la matrice  $A_h$  se fait en sommant chacune des contributions élémentaires  $A_{hij}(T_k)$  sur chacun des triangles  $T_k$  de la triangulation, ce qui est long et coûteux (c'est  $\forall i, j$ )

En pratique, on procède autrement. On rappelle tout d'abord que les seules valeurs  $A_{hij}(T_k)$  éventuellement non nulles correspondent à des indices  $i$  et  $j$  associés aux sommets du triangle  $T_k$ .

Ainsi, pour assembler la matrice, on procède plutôt en faisant une simple boucle sur les triangles  $T_k$  de la triangulation (cette boucle est "unique", au sens où elle ne sera pas répétée pour tous les indices  $i$  et  $j$ ) et à l'intérieur de cette boucle, si pour chaque triangle  $T_k$  de la triangulation, on introduit deux boucles imbriquées sur  $l=1,2,3$  et  $m=1,2,3$  permettant de calculer les contributions  $A_{hlm}(T_k)$ . L'algorithme s'écrit schématiquement de la manière suivante :

$$A_h = 0$$

boucle sur  $k=1, \dots, N_T$

    boucle sur  $l=1,2,3$

        calcul de l'indice global  $i$  correspondant à l'indice  $l$  sur  $T_k$

        boucle sur  $m=1,2,3$

            calcul de l'indice global  $j$  correspondant à l'indice  $m$  sur  $T_k$

            calcul de  $A_{hlm}(T_k)$

$$A_{hij} \neq A_{hij} + A_{hlm}(T_k)$$

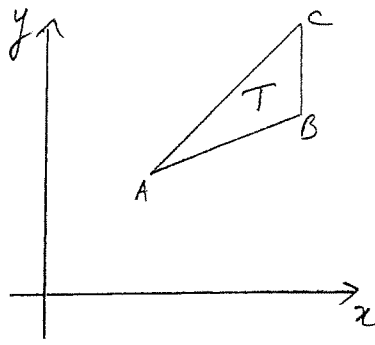
        fin boucle

    fin boucle

fin boucle

## Coordonnées barycentriques

On se donne un triangle  $ABC$  quelconque (mais non dégénéré)



$$A(x_A, y_A)$$

$$B(x_B, y_B)$$

$$C(x_C, y_C)$$

### Proposition

Soit  $\Pi(x, y)$  un point quelconque du plan. Alors il existe un unique triplet  $(\lambda_A, \lambda_B, \lambda_C) \in \mathbb{R}^3$  t. q

$$M = \lambda_A A + \lambda_B B + \lambda_C C$$

$$\text{et } \lambda_A + \lambda_B + \lambda_C = 1.$$

Les coordonnées sont appelées coordonnées barycentriques de  $\Pi$ .

Les propriétés suivantes sont également vérifiées :

(i) les coordonnées barycentriques sont des fonctions affines des coordonnées cartésiennes  $(x, y)$

(ii) les coordonnées cartésiennes sont des fonctions linéaires des coordonnées barycentriques  $(\lambda_A, \lambda_B, \lambda_C)$ .

iii)  $\Pi \in (AB) \Leftrightarrow \lambda_C = 0$

$$\Pi \in (AC) \Leftrightarrow \lambda_B = 0$$

$$\Pi \in (BC) \Leftrightarrow \lambda_A = 0$$

iv)  $\Pi \in ABC \Leftrightarrow 0 \leq \lambda_x \leq 1$  avec  $x = A, B, C$ .

### Démonstration

Montrer l'existence et l'unicité des coordonnées barycentriques revient à montrer l'existence et l'unicité d'une solution  $(\lambda_A, \lambda_B, \lambda_C)$  au système suivant :

$$(*) \begin{cases} \lambda_A + \lambda_B + \lambda_C = 1 \\ x_A \lambda_A + x_B \lambda_B + x_C \lambda_C = x \\ y_A \lambda_A + y_B \lambda_B + y_C \lambda_C = y \end{cases}$$

La matrice associée à ce système est la matrice

$$\begin{pmatrix} 1 & 1 & 1 \\ x_A & x_B & x_C \\ y_A & y_B & y_C \end{pmatrix}$$

dont le déterminant vaut clairement

$$D = \begin{vmatrix} 1 & 0 & 0 \\ x_A & x_B - x_A & x_C - x_A \\ y_A & y_B - y_A & y_C - y_A \end{vmatrix} = \pm 2 A(ABC)$$

Il est donc non nul puisqu'on a supposé le triangle ABC non dégénéré.

Les points (i), (ii) sont immédiats d'après  $\textcircled{*}$ .

Concernant (iii), montrons par exemple que

$$\Pi \in (AB) \Leftrightarrow \lambda_C = 0$$

$$\lambda_C = 0 \Leftrightarrow \begin{cases} \lambda_A + \lambda_B = 1 \\ x \lambda_A + x_C \lambda_B = x \\ y_A \lambda_A + y_B \lambda_B = y \end{cases} \Leftrightarrow \begin{cases} \lambda_B = 1 - \lambda_A \\ x - x_B = \lambda_A (x_A - x_B) \\ y - y_B = \lambda_A (y_A - y_B) \end{cases} \Leftrightarrow \begin{cases} \lambda_B = 1 - \lambda_A \\ \vec{B\Pi} = \lambda_A \vec{AB} \end{cases}$$

La relation  $\vec{B\Pi} = \lambda_A \vec{AB}$  signifie bien que  $\Pi \in (AB)$ .

Le point (iv) est certainement le plus technique à vérifier.

On a

$$\textcircled{*} \Leftrightarrow \begin{cases} \lambda_A = 1 - \lambda_B - \lambda_C \\ x - x_A = \lambda_B (x_B - x_A) + \lambda_C (x_C - x_A) \\ y - y_A = \lambda_B (y_B - y_A) + \lambda_C (y_C - y_A) \end{cases}$$

$$\Leftrightarrow \begin{cases} \lambda_A = 1 - \lambda_B - \lambda_C \\ \vec{A\Pi} = \lambda_B \vec{AB} + \lambda_C \vec{AC} \end{cases}$$

Il est ainsi clair que  $\Pi \in ABC \Rightarrow \begin{cases} \lambda_B \geq 0 \\ \lambda_C \geq 0 \\ \lambda_A \leq 1. \end{cases}$

En inversant les rôles de A, B et C, on a donc

$$\Pi \in ABC \Rightarrow \begin{cases} 0 \leq \lambda_A \leq 1 \\ 0 \leq \lambda_B \leq 1 \\ 0 \leq \lambda_C \leq 1 \end{cases}$$

Il reste à vérifier la réciproque.

Si  $0 \leq \lambda_A \leq 1$   
 $0 \leq \lambda_B \leq 1$   
 $0 \leq \lambda_C \leq 1$

alors  $\Pi \in \text{Conv}(\vec{AB}, \vec{AC})$  car  $\lambda_B \geq 0, \lambda_C \geq 0$   
 $\Pi \in \text{Conv}(\vec{CA}, \vec{CB})$  car  $\lambda_A \geq 0, \lambda_B \geq 0$   
 $\Pi \in \text{Conv}(\vec{BA}, \vec{BC})$  car  $\lambda_A \geq 0, \lambda_C \geq 0$ .

$\Pi$  appartient donc au triangle ABC.

# Equations de transport et equations des ondes

## Equation de transport

On considère l'équation scalaire suivante

$$\partial_t u + c \partial_x u = 0, \quad u = u(x, t), \quad \forall t > 0 \quad \forall x \in \mathbb{R}$$

où  $u$  est la fonction inconnue, et  $c$  la vitesse de transport. On se donne la valeur à l'instant  $t=0$  de la fonction  $u$ , à savoir

$$u(x, 0) = u_0(x) \quad \forall x \in \mathbb{R}$$

où  $u_0$  est une fonction supposée de classe  $\mathcal{C}^1$ .

Aussi allons montrer que ce problème admet une unique solution. Référons nous pour cela ce que l'on appellera les courbes caractéristiques comme les courbes de  $\mathbb{R}^2$  définies par  $(x(t), t)$  où la fonction  $X(t)$  est solution de l'équation différentielle

$$X'(t) = c$$

Soi, puisque  $c$  est une constante, on a simplement

$$X(t) = c(t - t^*) + X(t^*) \quad \forall t^*$$

Le long de ces courbes caractéristiques, on remarque que la fonction  $u$  est constante puisque par définition

$$\begin{aligned} \frac{d}{dt}(u(X(t), t)) &= (\partial_t u + X'(t) \partial_x u)(X(t), t) \\ &= (\partial_t u + c \partial_x u)(X(t), t) = 0. \end{aligned}$$

Ainsi, pour tout point  $(x^*, t^*)$  du plan avec  $t^* > 0$ , en considérant la courbe caractéristique définie par

$$\begin{cases} X'(t) = c \\ X(t^*) = x^* \end{cases}$$

qui est donc donnée par  $X(t) = c(t - t^*) + x^*$ , on a donc

$$u(x^*, t^*) = u(X(t^*), t^*) = u(X(0), 0) = u_0(X(0)) = u_0(x^* - ct^*)$$



La solution de l'équation de transport

$$\begin{cases} \partial_t u + c \partial_x u = 0 & \forall t > 0, \forall x \in \mathbb{R} \\ u(x, 0) = u_0(x) \end{cases}$$

avec  $c \in \mathbb{R}$  et  $u_0 \in \mathcal{C}^1(\mathbb{R})$  est donc donnée par

$$u(x, t) = u_0(x - ct)$$

On voit bien qu'il s'agit d'une équation de transport puisque la valeur de la solution en un point  $(x, t)$  coïncide avec la valeur de la donnée initiale au point  $x = ct$  et  $ct$  correspond bien à la distance que l'on parcourt pendant un temps  $t$  si l'on se déplace à la vitesse  $c$ .

Remarque

La méthode des caractéristiques utilisée ci-dessus peut être généralisée au cas d'une équation de transport à vitesse variable  $v(x, t)$  qui s'écrit maintenant  $\partial_t u + v(x, t) \partial_x u = 0 \quad \forall t > 0, \forall x \in \mathbb{R}$ , puisque

la fonction  $u$  est toujours constante le long des caractéristiques définies par

$$\begin{cases} X'(t) = v(X(t), t) \\ X(t^*) = x^* \end{cases}$$

En effet

$$\begin{aligned} \frac{d}{dt} u(X(t), t) &= (\partial_t u + X'(t) \partial_x u)(X(t), t) \\ &= (\partial_t u + v \partial_x u)(X(t), t) \\ &= 0 \end{aligned}$$

De la même manière que précédemment, on a donc

$$u(x^*, t^*) = u_0(X(0))$$

où  $X$  est la courbe caractéristique définie par

$$\begin{cases} X'(t) = v(X(t), t) \\ X(t^*) = x^* \end{cases}$$

Remarque

Il est aussi possible de résoudre avec cette méthode des équations de transport non homogène de la forme  $\partial_t u + v(x, t) \partial_x u = f$ , à ce prix que la fonction  $f$  est constante le long des caractéristiques puisque

$$\frac{d}{dt} u(x(t), t) = f(x(t), t)$$

En a alors, si  $x(t) = x$ ,

$$u(x, t) = u_0(x(0)) + \int_0^t f(x(s), s) ds$$

Ce terme est nul par une équation homogène.

### L'équation de transport conservative

On considère maintenant l'équation, dite conservative, à vitesse variable.

$$\begin{cases} \partial_t u + \partial_x (v(x, t) u) = 0 & \forall t > 0, \forall x \in \mathbb{R} \\ u(x, 0) = u_0(x) \end{cases}$$

où la vitesse  $v$  est telle que  $v(x, t) = 0$   $\forall x \leq a$   
 $\forall x \geq b$ .

Cette équation est dite conservative car en intégrant l'équation en  $x$  on obtient

$$\int_a^b u(x, t) dx = \int_a^b u_0(x) dx \quad \forall t > 0$$

L'intégrale en  $x$  de la solution sur l'intervalle  $[a, b]$  est donc conservée à chaque instant.

Pour déterminer la solution de cette équation conservative on procède comme précédemment en remarquant que  $u$  vérifie aussi

$$\partial_t u + v(x, t) \partial_x u = -u(x, t) \partial_x v(x, t)$$

de sorte que  $u$  n'est plus constante le long des caractéristiques mais vérifie (voir remarque précédente)

$$\frac{d}{dt} u(x(t), t) = -u(x(t), t) \partial_x v(x(t), t)$$

ou encore  $\frac{d}{dt} \ln |u(x(t), t)| = -\partial_x v(x(t), t)$  si  $u(x(t), t) \neq 0$ .

Ce qui donne  $\ln |u(x(t), t)| = \ln |u_0(x(0))| - \int_0^t \partial_x v(x(s), s) ds$

et donc  $u(x(t), t) = u_0(x(0)) e^{-\int_0^t \partial_x v(x(s), s) ds}$

# Equation des ondes

Le système de l'équation des ondes sur  $\mathbb{R}$  s'écrit

$$\begin{cases} \partial_t^2 u - c^2 \partial_x^2 u = 0 & \forall x \in \mathbb{R}, \forall t > 0 \\ u(x, 0) = u_0(x) & \forall x \in \mathbb{R} \\ \partial_t u(x, 0) = u_1(x) & \forall x \in \mathbb{R} \end{cases}$$

où  $c$  est la vitesse de propagation et où les fonctions  $u_0$  et  $u_1$  supposées de classe  $\mathcal{C}^2(\mathbb{R})$  et  $\mathcal{C}^1(\mathbb{R})$  désignent respectivement l'état et la "vitesse" initiale.

Il est utile et intéressant de faire tout de suite un lien avec l'équation de transport considérée précédemment en introduisant les fonctions

$$\begin{cases} v_1 = \partial_t u \\ v_2 = \partial_x u \end{cases}$$

et en remarquant que d'une part l'équation  $\partial_t^2 u - c^2 \partial_x^2 u = 0$  s'écrit

$\partial_t v_1 - c \partial_x v_2 = 0$  tandis que l'égalité des dérivées croisées de  $u$  s'écrit  $\partial_t v_2 - c \partial_x v_1 = 0$ . Avec ce changement de variables, le système de l'équation

des ondes s'écrit donc

$$\begin{cases} \partial_t v_1 - c \partial_x v_2 = 0 & \forall x \in \mathbb{R}, \forall t > 0 \\ \partial_t v_2 - c \partial_x v_1 = 0 & \forall x \in \mathbb{R}, \forall t > 0 \\ v_1(x, 0) = u_1(x) & \forall x \in \mathbb{R} \\ v_2(x, 0) = c u_0'(x) & \forall x \in \mathbb{R} \end{cases}$$

En introduisant maintenant les variables

$$\begin{cases} \vec{w} = \frac{1}{2} (v_2 + v_1) \\ \overleftarrow{w} = \frac{1}{2} (v_2 - v_1) \end{cases} \quad (\Leftrightarrow) \quad \begin{cases} v_2 = \vec{w} + \overleftarrow{w} \\ v_1 = \vec{w} - \overleftarrow{w} \end{cases}$$

on obtient

$$\begin{cases} \partial_t \vec{w} + c \partial_x \vec{w} = 0 & \forall x \in \mathbb{R}, \forall t > 0 \\ \partial_t \overleftarrow{w} - c \partial_x \overleftarrow{w} = 0 & \forall x \in \mathbb{R}, \forall t > 0 \\ \vec{w}(x, 0) = (c u_0'(x) - u_1(x)) / 2 \\ \overleftarrow{w}(x, 0) = (c u_0'(x) + u_1(x)) / 2 \end{cases}$$

de sorte que  $\vec{w}$  et  $\vec{w}$  satisfont des équations de transport à vitesse  $\pm c$ .  
 On notera que les équations sur  $\vec{w}$  et  $\vec{w}$  sont totalement découplées, c'est-à-dire qu'elles peuvent être résolues indépendamment l'une de l'autre. Plus précisément, on a d'après l'étude de l'équation de transport

$$\vec{w}(x,t) = \frac{1}{2} [c u_0'(x-ct) - u_1(x-ct)]$$

$$\vec{w}(x,t) = \frac{1}{2} [c u_0'(x+ct) + u_1(x+ct)]$$

et donc

$$\begin{cases} v_1(x,t) = \frac{1}{2} [u_1(x-ct) + u_1(x+ct)] - \frac{c}{2} [u_0'(x-ct) - u_0'(x+ct)] \\ v_2(x,t) = -\frac{1}{2} [u_1(x-ct) - u_1(x+ct)] + \frac{c}{2} [u_0'(x-ct) + u_0'(x+ct)] \end{cases}$$

On remarque que ces formules ne dépendent du couple  $(x,t)$  que par le biais des quantités  $x-ct$  et  $x+ct$ .

Pour revenir à la fonction  $u$ , l'égalité  $v_1(x,t) = \int_t^x u(x,t)$  entraîne

$$u(x,t) = u_0(x) + \int_0^t v_1(x,s) ds$$

c'est-à-dire

$$u(x,t) = u_0(x) + \frac{1}{2} \int_0^t [u_1(x-cs) + u_1(x+cs)] ds - \frac{c}{2} \int_0^t [u_0'(x-cs) - u_0'(x+cs)] ds$$

$$u(x,t) = u_0(x) + \frac{1}{2} \int_0^t [u_1(x-cs) + u_1(x+cs)] ds + \frac{1}{2} [u_0(x-ct) - u_0(x) + u_0(x+ct) - u_0(x)]$$

ce

$$u(x,t) = \frac{1}{2} [u_0(x-ct) + u_0(x+ct)] + \frac{1}{2} \int_0^t [u_1(x-cs) + u_1(x+cs)] ds$$

où encore, puisque  $\int_0^t u_1(x-cs) ds = \int_{y=x-ct}^{y=x} u_1(y) \frac{dy}{c}$

et  $\int_0^t u_1(x+cs) ds = \int_{y=x}^{y=x+ct} u_1(y) \frac{dy}{c}$ ,

Cette formule est connue sous le nom de formule de d'Alembert

$$u(x,t) = \frac{1}{2} [u_0(x-ct) + u_0(x+ct)] + \frac{1}{2c} \int_{x-ct}^{x+ct} u_1(y) dy$$

## Le cas du demi-espace $x > 0$

61

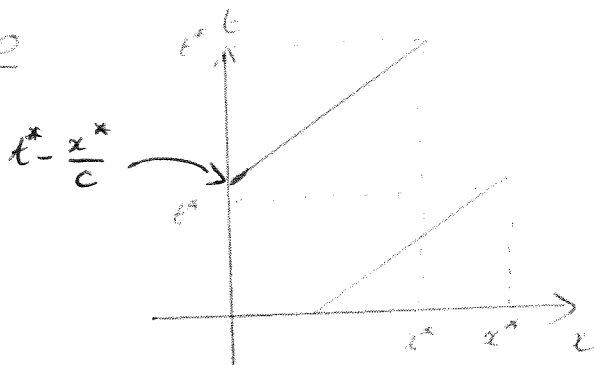
On considère dans un premier temps le cas de l'équation de transport à vitesse constante  $c$  posée sur l'espace  $x > 0$ , avec une condition de bord supplémentaire pour  $x = 0$ . Plus précisément, on considère le système

$$\begin{cases} 2u + c \partial_x u = 0 & \forall (x, t) \in \mathbb{R}^+ \times \mathbb{R}^+ \\ u(x, 0) = u_0(x) & \forall x \geq 0 \\ u(0, t) = g(t) & \forall t \geq 0 \end{cases}$$

Dans un souci de régularité, on voit donc qu'il est nécessaire de supposer la relation de compatibilité  $g(0) = u_0(0)$ .

Pour déterminer la solution de ce système, on écrit que  $u$  est constante le long des caractéristiques  $(x(t), t)$  avec  $x(t) = x^* + c(t - t^*)$ .

### Le cas $c > 0$



Il y a 3 possibilités :

→ si  $x^* - ct^* > 0$  alors

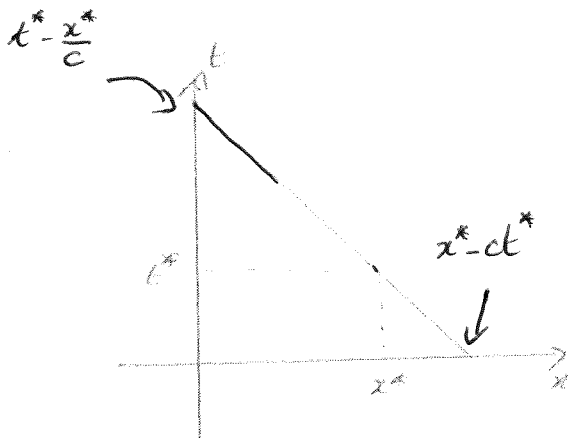
$$u(x^*, t^*) = u_0(x^* - ct^*)$$

→ si  $x^* - ct^* < 0$  alors la caractéristique que touche l'axe  $x = 0$  avant de toucher l'axe  $t = 0$  et dans ce cas  $u(x^*, t^*) = g\left(t^* - \frac{x^*}{c}\right)$

→ si  $x^* - ct^* = 0$  alors

$$u(x^*, t^*) = u_0(0) = g(0)$$

### Le cas $c < 0$



Dans ce cas  $x^* - ct^*$  est toujours strictement positif  $\forall (x^*, t^*) \neq (0, 0)$  et donc

$$u(x^*, t^*) = u_0(x^* - ct^*)$$

En note que dans ce cas on doit avoir la relation de compatibilité  $u_0(y) = g\left(-\frac{y}{c}\right)$  puisque  $u(x^*, t^*) = u_0(x^* - ct^*) = g\left(t^* - \frac{x^*}{c}\right)$ ,  
( $u$  est constante le long de la caractéristique)

Pour résumer ces deux cas, la solution est donc donnée par

$$u(x,t) = \begin{cases} u_0(x-ct) & \text{si } x-ct \geq 0 \\ g(t, \frac{x}{c}) & \text{si } x-ct < 0 \end{cases}$$

(avec en plus la relation de compatibilité  $u_0(y) = g(-\frac{y}{c})$  si  $c < 0$  et  $u_0(0) = g(0)$  tout le temps)

En considérant maintenant le cas de l'équation des ondes posée sur l'axe  $x > 0$  avec une condition de frontière de Neumann homogène, autrement dit le système

$$\begin{cases} \partial_{tt}^2 u - c^2 \partial_{xx}^2 u = 0 & \forall t > 0, \forall x > 0 \\ u(x,0) = u_0(x) & \forall x > 0 \\ \partial_x u(x,0) = u_1(x) & \forall x > 0 \\ \partial_x u(0,t) = 0 & \forall t > 0 \end{cases}$$

On suppose pour fixer les idées que  $c > 0$ .

En utilisant les calculs déjà effectués sur l'équation des ondes posée sur  $\mathbb{R}$  on aboutit donc aux équations de transport suivantes posées sur  $x > 0$  pour  $\vec{w}$  et  $\vec{w}$

$$\begin{cases} \partial_t \vec{w} + c \partial_x \vec{w} = 0 & \forall x > 0, \forall t > 0 \\ \vec{w}(x,0) = (c u_0'(x) - u_1(x))/2 \\ \vec{w}(x=0,t) = \frac{1}{2} (c \partial_x u - \partial_t u)(x=0,t) = -\frac{1}{2} \partial_t u(0,t) \end{cases}$$

$$\begin{cases} \partial_t \vec{w} - c \partial_x \vec{w} = 0 & \forall x > 0, \forall t > 0 \\ \vec{w}(x,0) = (c u_0'(x) + u_1(x))/2 \\ \vec{w}(x=0,t) = \frac{1}{2} (c \partial_x u + \partial_t u)(x=0,t) = \frac{1}{2} \partial_t u(0,t) \end{cases}$$

Les équations sont couplées par  $\partial_t u$ . En utilisant les résultats sur l'équation de transport posée sur  $x > 0$ , on a (rappelons que  $c > 0$  pour fixer les idées):

$$\vec{w}(x,t) = \begin{cases} \frac{1}{2} [c u_0'(x-ct) - u_1(x-ct)] & \text{si } x-ct \geq 0 \\ -\frac{1}{2} \partial_t u(0, t - \frac{x}{c}) & \text{si } x-ct < 0 \end{cases}$$

et, puisque  $-c < 0$

$$\overleftarrow{w}(x,t) = \frac{1}{2} [c u_0'(x+ct) + u_1(x+ct)]$$

avec la relation de compatibilité

$$\frac{1}{2} [c u_0'(x+ct) + u_1(x+ct)] = \frac{1}{2} 2u(0,t)$$

On en déduit donc que

$$\overrightarrow{w}(x,t) = \begin{cases} \frac{1}{2} [c u_0'(x-ct) - u_1(x-ct)] & \text{si } x-ct \geq 0 \\ -\frac{1}{2} [c u_0'(x+ct) + u_1(x+ct)] & \text{si } x-ct < 0 \end{cases}$$

$$\text{et } \overleftarrow{w}(x,t) = \frac{1}{2} [c u_0'(x+ct) + u_1(x+ct)]$$

Dans le cas  $x-ct \geq 0$ , la solution est donc donnée par la formule de d'Alembert obtenue précédemment puisque  $\overrightarrow{w}$  et  $\overleftarrow{w}$  coïncident avec les valeurs obtenues dans le cas  $x \geq 0$ .

Dans le cas  $x-ct < 0$ , on a

$$\overrightarrow{w}(x,t) = -\frac{1}{2} [c u_0'(ct-x) + u_1(ct-x)]$$

$$\overleftarrow{w}(x,t) = \frac{1}{2} [c u_0'(x+ct) + u_1(x+ct)]$$

d'où

$$\begin{cases} v_1(x,t) = \frac{1}{2} [u_1(ct-x) + u_1(x+ct)] + \frac{c}{2} [u_0'(ct-x) + u_0'(x+ct)] \\ v_2(x,t) = -\frac{1}{2} [u_1(ct-x) - u_1(x+ct)] - \frac{c}{2} [u_0'(ct-x) - u_0'(x+ct)] \end{cases}$$

Une fois encore, on remarque que ces formules ne dépendent que du couple  $(x,t)$  par le biais des quantités  $x-ct$  et  $x+ct$

Pour revenir à la fonction  $u$ , l'égalité  $v_t(x,t) = \partial_x u(x,t)$  entraîne

$$u(x,t) = u_0(x) + \int_{x/c}^t v_t(x,s) ds$$

c'est-à-dire, en remarquant que  $t=0 \Rightarrow x=0$  dans le cas  $x-ct \leq 0$ ,

$$u(x,t) = u(x, \frac{x}{c}) + \frac{1}{2} (u_0(ct-x) - u_0(0) + u_0(x+ct) - u_0(2x)) + \frac{1}{2} \int_{x/c}^t [u_1(cs-x) + u_1(cs+x)] ds = \frac{1}{2} [u_0(ct-x) + u_0(ct+x)] + \frac{1}{2c} \int_0^{ct-x} u_1(y) dy + \frac{1}{2c} \int_{2x}^{ct+x} u_1(y) dy$$

(il reste donc  $u(x, \frac{x}{c}) - \frac{1}{2} (u_0(0) + u_0(2x))$  qui vaut  $\frac{1}{2c} \int_0^{2x} u_1(y) dy$  d'après la formule de D'Alembert)

Enfin, on obtient la formule de d'Alembert généralisée suivante pour la solution de l'équation des ondes sur le demi-espace  $x > 0$ :

$$u(x,t) = \begin{cases} \frac{1}{2} [u_0(x-ct) + u_0(x+ct)] + \frac{1}{2c} \int_{x-ct}^{x+ct} u_1(y) dy & \text{si } x-ct \geq 0 \\ \frac{1}{2} [u_0(ct-x) + u_0(ct+x)] + \frac{1}{2c} \int_0^{ct-x} u_1(y) dy + \frac{1}{2c} \int_0^{ct+x} u_1(y) dy & \text{si } x-ct \leq 0 \end{cases}$$

### Conservation de l'énergie

L'objectif de ce paragraphe est de montrer que l'énergie du système  $E$ , définie à chaque instant  $t$  par

$$E(t) = \frac{1}{2} \int_0^l [(\partial_t u)^2 + c^2 (\partial_x u)^2] (x,t) dx$$

est conservée au cours du temps, c'est-à-dire qu'elle coïncide avec sa valeur à l'instant initiale donnée par

$$E(0) = \frac{1}{2} \int_0^l [u_1^2(x) + c^2 (u_0')^2(x)] dx,$$



pour le système de l'équation des ondes suivant posé sur  $[0,1]$  et/ou en espace et avec des conditions de Dirichlet homogène sur les bords.

$$\begin{cases} \partial_{tt}^2 u - c^2 \partial_{xx}^2 u = 0 & \forall x \in ]0,1[, \forall t > 0 \\ u(x,0) = u_0(x) & \forall x \in ]0,1[ \\ \partial_x u(x,0) = u_1(x) & \forall x \in ]0,1[ \\ u(0,t) = u(1,t) = 0 & \forall t \geq 0 \end{cases}$$

Il s'agit d'un résultat important permettant de démontrer l'unicité d'une solution du système mais aussi la stabilité de cette solution par rapport aux données  $u_0$  et  $u_1$  et nous reviendrons sur ces points, et montrons tout d'abord que  $E'(t) = 0 \forall t$ . En dérivant simplement l'énergie en fonction du temps, il vient

$$E'(t) = \int_0^1 (\partial_t u) (\partial_{tt}^2 u)(x,t) dx + c^2 \int_0^1 (\partial_x u) (\partial_{tx}^2 u)(x,t) dx.$$

En intégrant par parties la deuxième contribution, il vient

$$\int_0^1 (\partial_x u) (\partial_{tx}^2 u)(x,t) dx = \left[ (\partial_x u) (\partial_t u)(\cdot, t) \right]_{x=0}^{x=1} - \int_0^1 (\partial_{xx}^2 u) (\partial_t u)(x,t) dx$$

Puisque  $u(x=0,t) = u(x=1,t) = 0$ , on a aussi en dérivant par rapport à  $t$   $\partial_x u(0,t) = \partial_x u(1,t) = 0$  de sorte que  $\left[ (\partial_x u) (\partial_t u)(\cdot, t) \right]_{x=0}^{x=1} = 0 - 0 = 0$ .

Pour obtenir finalement

$$E'(t) = \int_0^1 (\partial_t u) \left[ \partial_{tt}^2 u - c^2 \partial_{xx}^2 u \right](x,t) dx = 0.$$

$\forall t \geq 0$ , on a donc  $E(t) = E(0)$ .

Ce résultat permet de montrer l'unicité d'une solution du problème.

En effet, par linéarité la différence  $v = u - \bar{u}$  entre deux solutions  $u$  et  $\bar{u}$  du problème vérifie le système

$$\begin{cases} \partial_{tt}^2 v - c^2 \partial_{xx}^2 v = 0 & \forall x \in ]0,1[, \forall t > 0, \\ v(x,0) = 0 & \forall x \in ]0,1[ \\ \partial_x v(x,0) = 0 & \forall x \in ]0,1[ \\ v(0,t) - v(1,t) = 0 & \forall t \geq 0 \end{cases}$$

et seule que son énergie vaut 0 à tout instant (puisque'elle est nulle

initialement)

111)

$$0 = \frac{1}{2} \int_0^1 [(\partial_t v)^2 + c^2 (\partial_x v)^2](x, t) dx$$

On obtient donc immédiatement que  $\partial_t v = \partial_x v = 0 \quad \forall (x, t)$  de sorte que  $v$  est constante en temps et en espace. En tenant compte des conditions aux limites nulles, on en déduit que cette constante est nulle, ce qui donne  $v = 0$  et donc  $u = \bar{u}$ . Ce qui prouve l'unicité de la solution.

De même, ce résultat permet de montrer la dépendance continue de la solution du problème vis à vis des données  $u_0$  et  $u_1$ . En effet, si on note  $u$  la solution associée aux données  $u_0$  et  $u_1$ , et  $\bar{u}$  la solution associée aux données  $\bar{u}_0$  et  $\bar{u}_1$ , la linéarité des équations nous donne

$$\frac{1}{2} \int_0^1 [(\partial_t (u - \bar{u}))^2 + c^2 (\partial_x (u - \bar{u}))^2](x, t) dx = \frac{1}{2} \int_0^1 [(\bar{u}_1 - u_1)^2(x) + c^2 (\bar{u}_0' - u_0')^2(x)] dx$$

Et ainsi, il est clair que si  $u_1 \rightarrow \bar{u}_1$  et  $u_0 \rightarrow \bar{u}_0$ , on obtient que  $u - \bar{u} \rightarrow 0$  en utilisant les mêmes arguments qu'auparavant.

## Approximation numérique des solutions - Brève introduction à la méthode

### des volumes finis

On va se focaliser dans cette brève introduction sur l'équation de transport sur  $\mathbb{R}$  et à vitesse constante

$$\begin{cases} \partial_t u + c \partial_x u = 0 & \forall t > 0, \forall x \in \mathbb{R} \\ u(x, 0) = u_0(x) \end{cases}$$

Le principe de la méthode des volumes finis consiste à introduire un pas  $\Delta x$ , un maillage de  $\mathbb{R}$  formé de cellules (ou volumes)

$\mathcal{C}_j = [x_{j-1/2}, x_{j+1/2}[$  de taille  $\Delta x$ ,  $x_{j+1/2} = (j+1/2)\Delta x$ , tel que

$$\mathbb{R} = \bigcup_{j \in \mathbb{Z}} \mathcal{C}_j,$$



un pas de temps  $\Delta t$ , et de calculer à chaque instant  $t^n = n\Delta t$  une approximation notée  $u_j^n$  de la valeur moyenne de la solution exacte à l'instant  $t^n$  sur la cellule  $G_j$ .

$$u_j^n \approx \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^n) dx.$$

Partant de la donnée initiale  $u_0(x)$  supposée connue, on peut définir la suite  $(u_j^0)_{j \in \mathbb{Z}}$  en posant simplement

$$u_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x) dx.$$

Pour définir la méthode des volumes finis considérée, il s'agit donc de proposer une relation de récurrence permettant de définir la suite  $(u_j^{n+1})_{j \in \mathbb{Z}}$  à partir de la connaissance de la suite  $(u_j^n)_{j \in \mathbb{Z}}$ .

Pour cela, intégrons l'équation  $\partial_t u + c \partial_x u = 0$  sur un volume de contrôle

$$[x_{j-1/2}, x_{j+1/2}] \times [t^n, t^{n+1}] :$$

$$\int_{x_{j-1/2}}^{x_{j+1/2}} \int_{t^n}^{t^{n+1}} (\partial_t u + c \partial_x u)(x, t) dx dt = 0$$

c'est-à-dire

$$\int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^{n+1}) dx - \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^n) dx + \int_{t^n}^{t^{n+1}} cu(x_{j+1/2}, t) dt - \int_{t^n}^{t^{n+1}} cu(x_{j-1/2}, t) dt = 0$$

où encore

$$\frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^{n+1}) dx = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^n) dx - \frac{\Delta t}{\Delta x} \left( \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} cu(x_{j+1/2}, t) dt - \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} cu(x_{j-1/2}, t) dt \right)$$

Cette relation, qui est exacte, suggère donc de proposer une relation (3) de récurrence de la forme

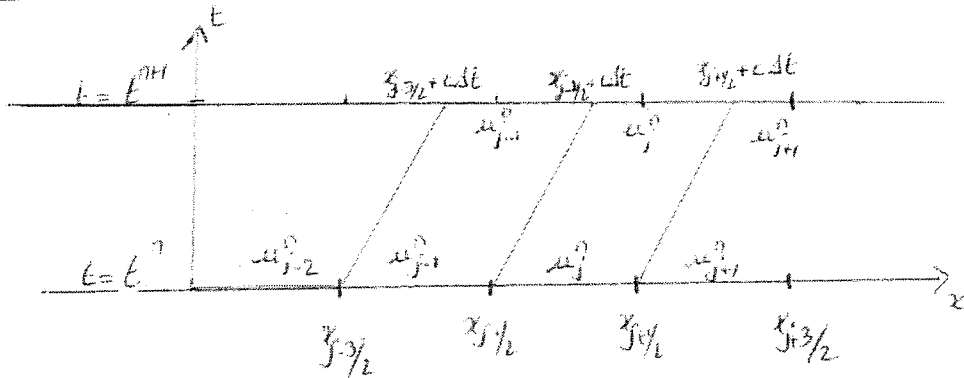
$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} \left( f_{j+1/2}^n - f_{j-1/2}^n \right) \quad \forall j \in \mathbb{Z}, \forall n \geq 0,$$

où  $f_{j+1/2}^n$  représenterait une approximation de la valeur du flux  $cu$  au point  $x_{j+1/2}$  entre les instants  $t^n$  et  $t^{n+1}$ :

$$f_{j+1/2}^n \approx \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} cu(x_{j+1/2}, t) dt.$$

Pour définir la méthode, il suffit donc de proposer une approximation de cette valeur moyenne des flux :

Supposons que  $c$  soit positif dans un premier temps.



Le dessin ci-dessus représente dans le plan  $(x, t)$  le transport des quantités  $u_j^n$  à la vitesse  $c > 0$ . On observe donc qu'au niveau de l'interface  $x_{j+1/2}$ ,  $u$  prend la valeur  $u_j^n$  (et donc  $cu$  prend la valeur  $c u_j^n$ ) entre les instants  $t^n$  et  $t^{n+1}$ , pourvu que  $x_{j-1/2} + c\Delta t \leq x_{j+1/2}$ , c'est-à-dire pourvu que

$$c\Delta t \leq \Delta x.$$

Cette condition, qui porte le nom de condition CFL (pour Courant-Friedrichs-Lewy), exprime le fait qu'une fois le choix de  $\Delta x$  effectué, celui de  $\Delta t$  doit être fait de sorte à vérifier l'inégalité  $\Delta t \leq \frac{\Delta x}{c}$ . Le schéma obtenu porte le nom de schéma décentré s'il s'agit donc

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} \left( f_{j+1/2}^n - f_{j-1/2}^n \right) \quad \forall n \geq 0 \quad \forall j \in \mathbb{Z} \quad \text{avec } f_{j+1/2}^n = c u_j^n.$$

## Le cas $c < 0$

16/

Dans le cas  $c < 0$ , un raisonnement similaire montre que le schéma décentré s'écrit naturellement

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} (f_{j+1/2}^n - f_{j-1/2}^n) \quad \forall n \geq 0 \quad \forall j \in \mathbb{Z} \quad \text{avec } f_{j+1/2}^n = c u_{j+1}^n$$

sous la condition CFL

$$\Delta t \leq \frac{\Delta x}{|c|}$$

## Application à l'équation des ondes

En utilisant des notations naturelles, le schéma décentré pour l'équation des ondes s'écrit en variables  $(w, \bar{w})$  sous la forme suivante

$$\begin{aligned} w_j^{n+1} &= w_j^n - \frac{c \Delta t}{\Delta x} (w_{j+1}^n - w_{j-1}^n) \\ \bar{w}_j^{n+1} &= \bar{w}_j^n + \frac{c \Delta t}{\Delta x} (w_{j+1}^n - w_j^n) \end{aligned}$$

sous la condition CFL  $\Delta t \leq \frac{\Delta x}{c}$  ( $c > 0$ ).

## Remarque

Il est possible de montrer la stabilité et la convergence du schéma décentré, et que celui-ci est d'ordre 1. Nous n'aborderons pas ces questions d'analyse numérique par manque de temps.