

# Equations aux Dérivées Partielles

Tony Lelièvre

2009-2010

Après avoir considéré dans le chapitre précédent des équations d'évolution pour des fonctions ne dépendant que du paramètre temps, nous nous intéressons dans cette partie à des équations sur des fonctions dépendant de plusieurs variables (typiquement une variable de temps et des variables de position en espace). Plus précisément, nous allons nous intéresser aux équations aux dérivées partielles.

Notre objectif est de présenter les principaux résultats concernant les propriétés qualitatives des solutions aux équations aux dérivées partielles, ainsi que les méthodes de discrétisation usuelles, en nous concentrant sur les problèmes elliptiques et paraboliques. Au passage, nous compléterons le cours d'analyse de première année [2].

Une équation aux dérivées partielles est une relation entre une fonction de plusieurs variables et ses dérivées :

$$F\left(x, u, \dots, \frac{\partial u}{\partial x_{i_1}}, \dots, \frac{\partial^2 u}{\partial x_{i_1} \partial x_{i_2}}, \dots, \frac{\partial^m u}{\partial x_{i_1} \dots \partial x_{i_m}}\right) = 0$$

où  $m$  est le degré de l'équation. Le problème est posé sur un domaine (*i.e.* un ouvert connexe)  $\Omega \subset \mathbb{R}^d$  ( $x \in \Omega$ ) et les indices  $i_k$  varient entre 1 et  $d$ . On note  $\partial\Omega$  la frontière de  $\Omega$  :

$$\partial\Omega = \overline{\Omega} \setminus \Omega.$$

La forme générale d'une équation aux dérivées partielles linéaire, scalaire, d'ordre 2 est

$$au + c \cdot \nabla u + \operatorname{div}(A\nabla u) = f \quad (1)$$

où  $a : \Omega \rightarrow \mathbb{R}$ ,  $c : \Omega \rightarrow \mathbb{R}^d$ ,  $A : \Omega \rightarrow \mathbb{R}^{d \times d}$  et  $f : \Omega \rightarrow \mathbb{R}$  sont les coefficients de l'équation aux dérivées partielles. Dans le cas où  $u$  est scalaire ( $d = 1$ ) et les coefficients sont constants, on obtient :

$$\alpha \frac{\partial^2 u}{\partial x^2} + \beta \frac{\partial^2 u}{\partial x \partial y} + \gamma \frac{\partial^2 u}{\partial y^2} + \delta \frac{\partial u}{\partial x} + \epsilon \frac{\partial u}{\partial y} + \zeta u = f, \quad (2)$$

avec  $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$  sont des scalaires. On dit que l'équation (2) est :

- *elliptique* si  $\beta^2 - 4\alpha\gamma > 0$ ,
- *parabolique* si  $\beta^2 - 4\alpha\gamma = 0$ ,
- *hyperbolique* si  $\beta^2 - 4\alpha\gamma < 0$ .

Il est plus important d'avoir en tête des exemples d'équations aux dérivées partielles appartenant à chacune de ces classes, plutôt que d'essayer de formaliser dans le détail cette classification. Donnons donc quelques exemples :

- L'équation de Laplace (ou Poisson) posée sur  $\Omega$

$$-\Delta u = f$$

est elliptique.

- L'équation de la chaleur posée sur  $Q = \mathbb{R}_+ \times \Omega$

$$\partial_t u - \Delta u = f$$

est parabolique.

- L'équation des ondes posée sur  $Q = \mathbb{R}_+ \times \Omega$

$$\partial_{t,t} u - \Delta u = f$$

est hyperbolique.

- L'équation d'advection diffusion posée sur  $Q = \mathbb{R}_+ \times \Omega$

$$\partial_t u + c \cdot \nabla u - \mu \Delta u = f$$

est parabolique si  $\mu > 0$  et hyperbolique si  $\mu = 0$ .

Dans la suite, nous nous concentrerons sur les problèmes elliptiques et paraboliques. Les problèmes hyperboliques présentent des caractéristiques différentes :

- En terme de régularité des solutions : pour des problèmes linéaires, si les données sont discontinues, la solution du problème est en général discontinue ; pour des problèmes non-linéaires, des discontinuités peuvent apparaître dans la solution, même pour des données régulières. On comparera avec le théorème de régularisation pour les solutions de l'équation de la chaleur (cf. Section 2.3.2). Mathématiquement, il faut donc travailler avec des notions de solution plus faibles.
- En terme de conditions aux limites : pour des problèmes elliptiques, pour que le problème admette une unique solution, il faut typiquement ajouter des conditions aux limites en tous les points de la frontière  $\partial\Omega$ . Pour des problèmes paraboliques, il faut ajouter des conditions aux limites en tous les points de la frontière  $\partial\Omega$ , et une condition initiale en  $t = 0$ . Pour des problèmes hyperboliques, la situation est en général plus compliquée : des conditions aux limites sont nécessaires simplement sur une partie du bord.

On renvoie au cours de calcul scientifique de première année [3] pour une introduction aux problèmes hyperboliques et à leur résolution numérique.

Avant d'attaquer l'étude des équations aux dérivées partielles, nous donnons dans la section suivante quelques rappels et compléments d'analyse nécessaires pour la suite.

## 1 Quelques rappels d'analyse fonctionnelle

Pour étudier les équations aux dérivées partielles, il faut préciser dans quel espace fonctionnel on cherche des solutions. On donne ensuite un sens différent aux opérateurs de différentiation suivant la régularité de la fonction. Nous allons passer en revue dans cette section les notions de fonctions utiles en pratique, de la notion la plus faible, à la notion la plus forte : les distributions, les mesures de Radon, les fonctions  $L^p$ , et enfin les fonctions  $H^k$ .

## 1.1 Distributions

### 1.1.1 Définition et propriétés fondamentales

Une idée fondamentale de l'analyse moderne est de considérer une fonction (disons définie sur  $\mathbb{R}$  et à valeurs dans  $\mathbb{R}$ ) non plus comme une collection de couples de réels  $(x, f(x))$  mais comme la collection des  $(\phi, \int f\phi)$  où  $\phi$  appartient à un ensemble de fonctions tests régulières. Ceci permet de généraliser la notion de fonction, et de fournir un cadre dans lequel les opérations de différentiations sont plus naturellement définies.

Rappelons la notion de distribution :

**Définition 1** Une distribution  $T$  sur  $\Omega$  est une forme linéaire sur les fonctions de classe  $\mathcal{C}^\infty$  et à support compact dans  $\Omega$  :

$$T : \begin{cases} \mathcal{C}_c^\infty(\Omega) & \rightarrow \mathbb{R} \\ \phi & \mapsto \langle T, \phi \rangle \end{cases}$$

qui vérifie la propriété de continuité suivante : pour tout compact  $K$  inclus dans  $\Omega$ , il existe un entier  $p$  et une constante  $C$  tel que :

$$\forall \phi \in \mathcal{C}_K^\infty, |\langle T, \phi \rangle| \leq C \sup_{|\alpha| \leq p, x \in K} |\partial^\alpha \phi(x)|$$

où  $\alpha = (\alpha_1, \dots, \alpha_d)$  désigne un multi-indice, de longueur  $|\alpha| = \alpha_1 + \dots + \alpha_d$ ,  $\partial^\alpha \phi = \frac{\partial^{|\alpha|} \phi}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}$  et

$$\mathcal{C}_K^\infty = \{f \in \mathcal{C}^\infty(\Omega), f \text{ à support dans le compact } K\}.$$

On note  $\mathcal{D}'(\Omega)$  l'ensemble des distributions sur  $\Omega$ . Lorsque l'entier  $p$  peut être choisi de manière indépendante de  $K$ , on dit que la distribution  $T$  est d'ordre fini, et la plus petite valeur possible de  $p$  est appelé l'ordre de  $T$ .

**Remarque 1** L'extension à des distributions à valeurs complexes est immédiate.

**Remarque 2** L'espace  $\mathcal{C}_c^\infty(\Omega)$  n'a pas de topologie très simple à définir. Par contre, pour tout compact  $K \subset \Omega$ , on peut munir l'espace  $\mathcal{C}_K^\infty$  de la topologie de la convergence uniforme des fonctions et de toutes leurs dérivées (c'est un espace de Fréchet, et donc un espace complet pour une métrique bien choisie). Noter que  $\mathcal{C}_c^\infty(\Omega)$  est la réunion des  $\mathcal{C}_K^\infty$  pour  $K$  compact inclus dans  $\Omega$ . L'ensemble des distributions est donc l'ensemble des formes linéaires sur  $\mathcal{C}_c^\infty(\Omega)$ , dont la restriction à  $\mathcal{C}_K^\infty$  est continue, pour tout compact  $K \subset \Omega$ .

Un premier exemple de distribution est la distribution de Dirac au point  $a$  définie par

$$\langle \delta_a, \phi \rangle = \phi(a).$$

C'est une distribution d'ordre 0.

Un deuxième exemple de distribution d'ordre 0 est donné par le lemme suivant :

**Lemme 1** L'application qui à toute fonction  $f \in L^1_{\text{loc}}(\Omega)$  associe la distribution

$$T_f : \begin{cases} \mathcal{C}_c^\infty(\Omega) & \rightarrow \mathbb{R} \\ \phi & \mapsto \int f\phi \end{cases}$$

est une injection.

On rappelle que  $L^1_{\text{loc}}(\Omega)$  est l'ensemble des fonctions intégrables sur tout compact de  $\Omega$ , et que deux fonctions de cet espace sont identifiées si elles sont égales presque partout. C'est la notion la plus faible de fonction au sens classique d'une application qui à une valeur en associe une autre que l'on puisse donner. Le Lemme précédent revient donc à démontrer que pour  $f \in L^1_{\text{loc}}(\Omega)$ , si  $\int f\phi = 0$  pour toute fonction  $\phi \in \mathcal{C}_c^\infty(\Omega)$ , alors  $f = 0$  presque partout (cf. [1, Lemme IV.2]). Par ce lemme, on identifie donc toute fonction  $L^1_{\text{loc}}(\Omega)$  à une distribution.

Rappelons la topologie sur l'ensemble des distributions :

**Définition 2** Une suite  $T_n$  de distributions converge vers  $T$  si et seulement si, pour toute fonction  $\phi \in \mathcal{C}_c^\infty(\Omega)$ ,  $\langle T_n, \phi \rangle$  converge vers  $\langle T, \phi \rangle$ .

**Exemple 1** Par exemple, les fonctions  $f_n(x) = n \sin(nx)$  convergent vers 0 dans  $\mathcal{D}'(\mathbb{R})$  car, par intégrations par parties, pour toute fonction  $\phi \in \mathcal{C}_c^\infty(\mathbb{R})$ ,  $\int f_n\phi = -\frac{1}{n} \int \sin(nx)\phi'$ . Remarquer qu'on a pourtant, pour tout compact  $K \subset \mathbb{R}$ ,  $\int_K |f_n| \rightarrow +\infty$  quand  $n \rightarrow \infty$  : la convergence au sens des distributions est donc moins exigeante que dans  $L^1_{\text{loc}}$ .

Un intérêt majeur de la notion de distribution est que l'opération de dérivation est toujours bien définie et continue.

**Proposition 1** Pour toute distribution  $T \in \mathcal{D}'(\Omega)$ , sa dérivée  $\partial_{x_i}T$  est une distribution définie par :

$$\langle \partial_{x_i}T, \phi \rangle = -\langle T, \partial_{x_i}\phi \rangle.$$

De plus, si  $T_n$  est une suite de distributions tendant vers  $T$ , alors  $\partial_{x_i}T_n$  converge vers  $\partial_{x_i}T$ .

Evidemment, cette notion de dérivation prolonge la notion usuelle de dérivation pour les fonctions : pour une fonction  $f$  de classe  $\mathcal{C}^1$  (et donc dérivable au sens classique : l'accroissement  $(f(x+h) - f(x))/h$  a une limite quand  $h$  tend vers 0, en tout point  $x \in \Omega$ ), la dérivée au sens distribution s'identifie avec la dérivée au sens classique.

**Exercice 1** Soit  $f(x) = x^2 \sin(1/x^2)$ . On prolonge  $f$  en  $x = 0$  par continuité. Montrer que  $f$  est partout différentiable. Que dire de la dérivée de  $T_f$  ? Indication : On vérifiera que  $f$  n'est pas  $\mathcal{C}^1$ . Pour étudier la dérivée au sens des distributions, on pourra s'appuyer sur le cas  $f(x) = \ln|x|$  vu en première année.

**Exercice 2** – Soit  $T$  une distribution de dérivée nulle. Montrer que  $T$  est une constante.

– Soit  $\chi : \mathbb{R} \rightarrow \mathbb{R}$  une fonction positive à support compact et telle que  $\int \chi = 1$ . Quelle est la limite (quand  $n \rightarrow \infty$ ) de  $n\chi(nx)$  dans  $\mathcal{D}'(\mathbb{R})$  ?

- Quelle est la limite (quand  $n \rightarrow \infty$ ) de  $\sin(nx)/(\pi x)$  dans  $\mathcal{D}'(\mathbb{R})$  ? (On pourra écrire  $\sin(nx)/(\pi x) = \frac{d}{dx} \left( \int_0^{nx} \sin t/(\pi t) dt \right)$ , et utiliser le fait que  $\int_0^\infty \sin t/t dt = \pi/2$ .)
- Question subsidiaire : Retrouver le résultat de la question précédente en utilisant la transformée de Fourier.

Un exemple important de dérivées au sens des distributions est donné par la formule de Stokes.

**Lemme 2 (Formule de Stokes)** Soit  $K$  un compact de  $\mathbb{R}^d$  de frontière régulière  $\partial K$ , et de normale sortant  $n_K$ . Alors le gradient (au sens distribution) de la fonction indicatrice de  $K$  est la distribution (d'ordre 0)  $-n_K d\sigma$  où  $\sigma$  est la mesure surfacique sur le bord  $\partial K : \forall \phi \in (C_c^\infty(\mathbb{R}^d))^d$ ,

$$\langle \nabla 1_K, \phi \rangle = - \int_K \operatorname{div}(\phi) dx = - \int_{\partial K} \phi \cdot n_K d\sigma.$$

**Remarque 3** On peut voir ce Lemme comme la définition de la mesure  $\sigma$  (encore faut-il vérifier qu'elle vérifie toutes les propriétés requises pour définir une mesure). Voici une manière de définir la mesure  $\sigma$  dans le cas où on dispose d'une paramétrisation de  $\partial K$ . Si on dispose d'une paramétrisation du bord  $\partial K$  sous la forme :

$$h : \begin{cases} U & \rightarrow \partial K \\ x & \mapsto h(x) \end{cases}$$

où  $U$  est un ouvert de  $\mathbb{R}^p$ , alors

$$\int_{\partial K} f d\sigma = \int_U f \circ h \sqrt{g_h} dx$$

où  $g_h = \det G_h$  et  $(G_h)_{i,j} = \nabla h_i \cdot \nabla h_j$  est la matrice de Gram de  $\operatorname{Jac}(h) = \nabla h$ .

**Exercice 3** Que devient la formule de Stokes en dimension 1, avec  $K = [a, b]$  ? Calculer l'aire de la sphère unité  $\mathcal{S}^2$  de  $\mathbb{R}^3$  en utilisant la paramétrisation

$$h : \begin{cases} (0, \pi) \times (-\pi, \pi) & \rightarrow \mathcal{S}^2 \\ (\theta, \phi) & \mapsto (\sin(\theta) \cos(\phi), \sin(\theta) \sin(\phi), \cos(\theta)) \end{cases}.$$

Nous terminons ces rappels sur les distributions par deux types de résultat : le lien entre théorie de la mesure et analyse fonctionnelle *via* le théorème de représentation de Riesz, et enfin la relation intégration - dérivation.

### 1.1.2 Le théorème de représentation de Riesz

Un lien important peut être fait entre mesure et distribution :

**Théorème 1 (Théorème de représentation de Riesz)** Pour toute distribution  $T$  positive (i.e. telle que  $\langle T, \phi \rangle \geq 0$  dès que  $\phi \geq 0$ ), il existe une unique mesure  $\mu$  borélienne, positive et finie sur les compacts de  $\Omega$  telle que

$$\langle T, \phi \rangle = \int \phi d\mu.$$

On peut donc naturellement identifier les mesures  $\mu$  boréliennes, positives et finies sur les compacts de  $\Omega$  (on parle de mesure de Radon<sup>1</sup> positive) aux distributions sur  $\Omega$  positives. La démonstration ressemble beaucoup au théorème de prolongement de Carathéodory (qui permet de construire la mesure de Lebesgue) et consiste à construire la mesure d'un borélien à partir de  $T$ , puis à vérifier que cette mesure vérifie les propriétés requises (cf. [7, Théorème 2.14]).

**Remarque 4** *Un corollaire de ce théorème est que les distributions positives sont nécessairement d'ordre 0. En fait, toute distribution  $T$  (à valeur réelle) d'ordre 0 s'identifie à une forme linéaire sur l'ensemble des fonctions continues à support compact, muni de la norme du sup. Si on suppose de plus cette forme linéaire bornée :*

$$\exists C, \forall \phi \in \mathcal{C}_c^\infty(\Omega), |\langle T, \phi \rangle| \leq C \|\phi\|_{L^\infty},$$

*alors elle s'écrit (cf. [7, Théorème 6.19]) sous la forme  $T_+ - T_-$ , avec  $T_+$  et  $T_-$  des distributions positives. En particulier, la distribution  $T$  peut être identifiée à la différence de deux mesures de Radon positives bornées (on parle de mesure de Radon bornée).*

**Remarque 5** *Les mesures de probabilités peuvent donc être vues comme des distributions. Une question naturelle est de relier la notion de convergence en loi (la notion la plus faible introduite dans le cours de probabilité) et la notion de convergence au sens des distributions. On peut vérifier que si  $\mu_n$  est une suite de mesures de probabilité et que  $\mu$  est également une mesure de probabilité, alors la convergence en loi de  $\mu_n$  vers  $\mu$  est équivalente à la convergence au sens des distributions de  $\mu_n$  vers  $\mu$ .*

*Pour les lecteurs connaissant les distributions tempérées et la transformée de Fourier sur les distributions tempérées, la convergence en loi de  $\mu_n$  vers  $\mu$  est en fait équivalente à la convergence de  $\mu_n$  vers  $\mu$  dans l'espace des distributions tempérées. On comprend alors pourquoi la convergence en loi est équivalente à la convergence ponctuelle des fonctions caractéristiques des lois, qui ne sont rien d'autre que les transformées de Fourier (dans l'espace des distributions tempérées) des lois.*

### 1.1.3 Intégration et dérivation

Dans l'approche élémentaire (intégrale de Riemann, et dérivation au sens classique), l'intégration et la dérivation sont des opérations inverses. Qu'en est-il avec la notion d'intégrale de Lebesgue, et la dérivation au sens des distributions ? On suppose donc ici  $d = 1$ , et que l'on travaille sur un intervalle borné de  $\mathbb{R}$ , disons  $[0, 1]$ .

Commençons tout d'abord par une des implications : si  $f \in L^1(0, 1)$  et  $F(x) = \int_0^x f(y) dy$ , alors on peut faire le lien avec la dérivée au sens classique grâce à la proposition suivante (cf. [7, Théorème 7.7]) :

**Proposition 2 (Théorème de dérivation de Lebesgue)** *Soit  $f$  une fonction de  $L^1_{\text{loc}}(\mathbb{R})$ . Alors, pour presque tout point  $x$ ,*

$$\lim_{h \rightarrow 0} \frac{1}{2h} \int_{x-h}^{x+h} |f(y) - f(x)| dy = 0.$$

---

<sup>1</sup>Une mesure de Radon sur  $\mathbb{R}^d$  est une mesure borélienne finie sur les compacts

Autrement dit, la fonction  $F$  est presque partout dérivable au sens classique, et sa dérivée est  $f$ . Il est facile de vérifier que sa dérivée au sens des distributions  $T'_F$  est la fonction  $f$ . On a donc bien dans ce cas “dérivation  $\circ$  intégration = Id”, et les notions de dérivée au sens classique et au sens des distributions coïncident.

Qu’en est-il de la réciproque ? Commençons par un contre-exemple : il existe des fonctions  $f$  continues sur  $[0, 1]$ , dérivables presque partout sur  $[0, 1]$ , avec  $f' \in L^1[0, 1]$  ( $f'$  étant ici à comprendre au sens classique et non pas au sens des distributions), et telles que la relation suivante soit **fausse** :

$$f(x) - f(0) = \int_0^x f'(y) dy. \quad (3)$$

En effet, soit  $E_0 = [0, 1]$ ,  $E_1 = [0, 1/3] \cup [2/3, 1]$ , et, par récurrence,  $E_n$  est la réunion de  $2^n$  segments fermés disjoints de longueur  $(1/3)^n$ , obtenus en enlevant un tiers au centre de chacun des intervalles de  $E_{n-1}$ . La suite des  $E_n$  est une suite décroissante d’ensemble mesurables, et donc  $E = \bigcap_{n \geq 0} E_n$  est mesurable. On vérifie que  $\lambda(E_n) = (2/3)^n$ , et donc  $\lambda(E) = 0$ . Soit maintenant  $f_n(x) = (2/3)^{-n} \int_0^x 1_{E_n}(y) dy$ . On a (cf. Figure 1)  $f_n(0) = 0$ ,  $f_n(1) = 1$  et  $f_n$  est une fonction continue, croissante, constante sur  $E_n^c$ . On vérifie de plus que, pour  $x \in [0, 1]$ ,  $|f_n(x) - f_{n+1}(x)| \leq 2^{-n}$ . La suite de fonctions  $f_n$  converge donc uniformément vers une fonction  $f$  continue croissante, telle que  $f(0) = 0$  et  $f(1) = 1$ . Par ailleurs, on a  $f'(x) = 0$  pour tout  $x \notin E$ , on donc  $f$  est dérivable et de dérivée nulle presque partout, ce qui semble contradictoire avec  $f(0) = 0$  et  $f(1) = 1$ . (L’ensemble  $E$  s’appelle l’ensemble triadique de Cantor, et  $f$  est appelée l’escalier du diable). Autrement dit, (3) n’est pas vérifié, et “la dérivée de  $f$  n’est pas  $f'$ ”...

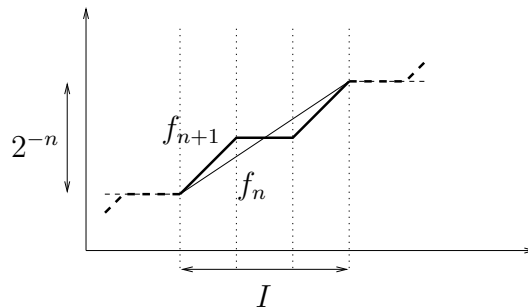


FIG. 1 – L’escalier du diable :  $I$  désigne un des  $2^n$  intervalles de  $E_n$ . En trait fin la fonction  $f_n$ , et en trait gras la fonction  $f_{n+1}$ .

Soit maintenant une fonction  $f \in L^1(0, 1)$  : on peut donc considérer la dérivée  $T'_f$  de  $f$  au sens distribution. Pour pouvoir écrire une relation telle (3), il faut tout d’abord (i) que  $T'_f$  soit une mesure borélienne (si  $T'_f = \mu$ , on peut écrire  $f(x) - f(a) = \mu[a, x]$ ), puis ensuite (ii) que cette mesure soit absolument continue par rapport à la mesure de Lebesgue, (pour écrire  $T'_f = \mu = f'(y) dy$ ).

Considérons le premier point : à quelle condition  $T'_f$  est une mesure. D’après ce qu’on a vu précédemment (cf. Remarque 4), ceci revient à écrire que  $T'_f$  est une distribution d’ordre 0, bornée comme forme linéaire sur l’ensemble des fonctions continues à support compact, muni de la norme du sup : c’est la notion de fonction à variation bornée.

**Proposition 3** On appelle fonction à variation bornée sur l'intervalle  $[0, 1]$  les fonctions de  $L^1(0, 1)$  telles que  $T'_f$  soit une distribution d'ordre 0 bornée comme forme linéaire sur l'ensemble des fonctions continues à support compact, muni de la norme du sup. Autrement dit, ce sont les fonctions  $f$  de  $L^1(0, 1)$  telles que il existe  $C$ , t.q. pour toute fonction  $\phi \in C_c^\infty((0, 1))$ ,

$$\left| \int f \phi' \right| \leq C \|\phi\|_{L^\infty}.$$

On a :

- Une fonction est à variation bornée si et seulement si elle s'écrit comme la différence de deux fonctions croissantes bornées<sup>2</sup>,
- Une fonction est à variation bornée si et seulement si

$$VT(f) = \sup_{0=x_0 < x_1 < \dots < x_N=1, N \in \mathbb{N}} \sum_{i=1}^N |f(x_i) - f(x_{i-1})| < \infty$$

où le supremum est pris sur toutes les subdivisions possibles de  $[0, 1]$ .

On généralise la notion de fonction à variation bornée aux fonctions définies sur tout  $\mathbb{R}$ , en demandant à la fonction d'être à variation bornée sur tout compact.

**Exemple 2** Les fonctions croissantes et bornées sont à variation bornée. Une fonction Lipschitzienne est à variation bornée. La fonction escalier du diable est croissante et bornée : elle est donc clairement à variation bornée. Sa dérivée au sens distribution ne comporte que la partie étrangère à la mesure de Lebesgue.

**Remarque 6 (Fonctions à variation bornée et généralisation de l'intégrale de Riemann)**

Pour une fonction  $f$  à variation bornée, on peut définir  $\int \phi df$  comme la limite de  $\sum_{i=0}^{n-1} \phi(x_i)(f(x_{i+1}) - f(x_i))$  quand  $\Delta x = 1/n$  tend vers 0 et où  $x_i = i\Delta x$ . D'après la Remarque 4, la dérivée de  $f$  au sens distributions s'écrit  $\mu_+ - \mu_-$  (où  $\mu_+$  et  $\mu_-$  sont deux mesures de Radon positives bornées) et  $\int \phi df = \int \phi d\mu_+ - \int \phi d\mu_-$ . L'intégrale  $\int \phi df$  a un sens dès que  $\phi$  est telle que  $\int |\phi| d\mu_+ < \infty$  et  $\int |\phi| d\mu_- < \infty$ .

Pour le second point, nous avons besoin du théorème de Radon-Nikodym (cf. [7, Théorème 6.10]) :

**Proposition 4 (Théorème de Radon-Nikodym)** Soit  $(X, \mathcal{A})$  un espace muni d'une  $\sigma$ -algèbre, et  $\mu$  et  $\lambda$  deux mesures positives. Alors la mesure  $\lambda$  se décompose de manière unique sous la forme

$$\lambda = \lambda_a + \lambda_s$$

où  $\lambda_a$  est absolument continu par rapport à  $\mu$  (ce qui s'écrit  $\lambda_a \ll \mu$ ) i.e.

$$\forall A \in \mathcal{A}, \mu(A) = 0 \text{ implique } \lambda_a(A) = 0,$$

---

<sup>2</sup>Cette propriété est fondamentale pour étudier les propriétés de ces fonctions : elle dit qu'on peut essentiellement se restreindre à l'étude des fonctions croissantes bornées.



et  $\lambda_s$  est singulière par rapport à  $\mu$  (ce qui s'écrit  $\lambda_s \perp \mu$ ) i.e.

il existe  $A$  et  $B$  dans  $\mathcal{A}$  disjoints tels que  $\lambda_s(A) = \lambda_s(X)$  et  $\mu(B) = \mu(X)$ .

De plus, il existe une fonction  $f$  intégrable par rapport à la mesure  $\mu$ , telle que  $\lambda_a = f\mu$ , i.e.  $\forall A \in \mathcal{A}$ ,

$$\lambda_a(A) = \int 1_A f d\mu.$$

En utilisant le théorème de Radon-Nikodym, on montre que

**Proposition 5** *Soit  $f$  une fonction à variation bornée. Alors  $f$  est dérivable (au sens classique) presque partout, et sa dérivée (au sens classique)  $f'$  est intégrable. On a*

$$T'_f = f'(x) dx + \nu$$

où  $f'(x)dx$  est la partie absolument continue par rapport à la mesure de Lebesgue, et  $\nu$  la partie étrangère.

Les fonctions à variations bornées jouent un rôle très important en analyse. On les rencontre par exemple comme solutions de problèmes hyperboliques (remarquer qu'elles peuvent être discontinues). L'ensemble des fonctions à variations bornées muni de la norme  $VT(f) + \|f\|_{L^\infty}$  est un espace de Banach, et les bornés de cet espace de Banach sont compacts dans  $L^1_{\text{loc}}$ .

Au vu de la Proposition 5, on comprend que la relation (3) ne peut être satisfaite que si  $\nu = 0$ , autrement dit si la mesure de Radon  $T'_f$  est absolument continue par rapport à la mesure de Lebesgue : c'est la notion de fonction absolument continue.

**Proposition 6** *La fonction  $f$  est dite absolument continue (sur  $(0, 1)$ ) si et seulement si elle est à variation bornée, et sa dérivée  $T'_f$  est absolument continue par rapport à la mesure de Lebesgue. Autrement dit,  $f$  est absolument continue sur  $(0, 1)$  si et seulement si elle est dans  $L^1(0, 1)$  et sa dérivée au sens distribution  $T'_f$  s'identifie à une fonction dans  $L^1(0, 1)$ . On a également la caractérisation : Une fonction est absolument continue (sur  $(0, 1)$ ) si et seulement si  $\forall \epsilon > 0, \exists \delta > 0, \forall N$ , pour toute famille  $0 \leq \alpha_1 \leq \beta_1 \leq \alpha_2 \leq \beta_2 \leq \dots \leq \alpha_N \leq \beta_N \leq 1$  tels que*

$$\sum_{i=1}^N (\beta_i - \alpha_i) < \delta$$

alors

$$\sum_{i=1}^N |f(\beta_i) - f(\alpha_i)| < \epsilon.$$

**Remarque 7** *On montre également que si  $f$  est une fonction continue, dérivable en tout point (et non pas simplement presque partout), et telle que  $f' \in L^1$ , alors  $T'_f = f'$  (cf. [7, Théorème 7.21]).*

On retiendra de cette discussion que la “véritable dérivée”, celle qui rend compte complètement de la variation de  $f$ , est la dérivée au sens des distributions. La dérivée au sens classique laisse échapper les sauts dans la variation de  $f$ .

Par exemple, la dérivée de la fonction  $f(x) = x + 1_{x > 1/2}$  (définie sur  $(0, 1)$ ) est, au sens des distributions,  $1 + \delta_{1/2}$ . Au sens classique, sa dérivée vaut presque partout 1.

## 1.2 Espaces de Sobolev

### 1.2.1 Définitions

Les espaces de Sobolev sont des espaces très naturels pour les solutions des équations aux dérivées partielles.

**Définition 3** Pour  $k \geq 1$ , l'espace de Sobolev  $H^k(\Omega)$  est l'ensemble des fonctions  $f \in L^2(\Omega)$  telles que les dérivées de  $f$  au sens distribution jusqu'à l'ordre  $k$  s'identifient à des fonctions de  $L^2(\Omega)$ . Autrement dit,

$$H^k(\Omega) = \left\{ f \in L^2(\Omega), \forall \alpha \text{ multi-indice de longueur } |\alpha| \leq k, \exists g_\alpha \in L^2(\Omega), \right. \\ \left. \partial^\alpha T_f = g_\alpha \text{ dans } \mathcal{D}'(\Omega) \right\}.$$

**Remarque 8** Par le théorème de Riesz sur l'espace de Hilbert  $L^2(\Omega)$ , une distribution  $T$  s'identifie à une fonction  $f \in L^2(\Omega)$  si et seulement si  $T$  définit une forme linéaire continue sur  $L^2(\Omega)$ , i.e.,  $\exists C, \forall \phi \in \mathcal{C}_c^\infty(\Omega), |\langle T, \phi \rangle| \leq C \|\phi\|_{L^2}$ .

On rappelle que  $H^k$  muni du produit scalaire

$$(f, g)_{H^k} = \sum_{|\alpha| \leq k} \int_{\Omega} \partial^\alpha f \partial^\alpha g$$

est un espace de Hilbert séparable.

On peut également définir l'espace de Sobolev  $H^k(\Omega)$  pour des  $k$  fractionnaires. Retenons simplement qu'on a toujours les injections continues : si  $k \leq l$ ,  $H^l(\Omega) \subset H^k(\Omega)$ .

**Remarque 9** Quand  $\Omega = \mathbb{R}^d$ , ou bien quand  $\Omega = \mathbb{T}^d$ , on peut définir les espaces de Sobolev en utilisant la transformée de Fourier et les séries de Fourier. Ainsi, sur  $\Omega = \mathbb{R}^d$ , on dit que  $f \in H^k(\mathbb{R}^d)$  si et seulement si  $f \in L^2(\mathbb{R}^d)$  et  $(1 + |\xi|^2)^{k/2} \hat{f}(\xi) \in L^2(\mathbb{R}^d)$ , où  $\hat{f}(\xi) = \int_{\mathbb{R}^d} \exp(-ix \cdot \xi) f(x) dx$  est la transformée de Fourier de  $f$ . Remarquer que cette définition a bien un sens pour  $k$  non entier.

**Définition 4** Pour  $k \geq 1$ , l'espace  $H_0^k(\Omega)$  est par définition l'adhérence des fonctions  $\mathcal{C}_c^\infty(\Omega)$  dans  $H^k(\Omega)$ .

L'espace  $H_0^k(\Omega)$  est en général strictement inclus dans  $H^k(\Omega)$ . On a cependant  $H_0^k(\mathbb{R}^d) = H^k(\mathbb{R}^d)$  : les fonctions  $\mathcal{C}_c^\infty(\Omega)$  sont denses dans  $H^k(\mathbb{R}^d)$ . On rappelle que de manière générale, si  $\Omega$  est suffisamment régulier les fonctions  $\mathcal{C}_c^\infty(\bar{\Omega})$  sont denses dans  $H^k(\Omega)$  (cf. [1, Corollaire IX.8]).

Attention à ne pas confondre  $H^2(\Omega) \cap H_0^1(\Omega)$  avec  $H_0^2(\Omega)$ .

**Exercice 4** Donner un exemple de fonction de  $H^2((0, 1)) \cap H_0^1((0, 1))$  qui n'est pas dans  $H_0^2((0, 1))$ .

Solution : Prendre une fonction régulière qui s'annule au bord, mais dont les dérivées ne s'annulent pas au bord.

Voici une caractérisation de  $H_0^1(\Omega)$  parfois utile :

**Proposition 7** Soit  $u \in L^2(\Omega)$ . Alors  $u \in H_0^1(\Omega)$  si et seulement si la fonction  $\bar{u}$  qui est le prolongement de  $u$  par 0 est une fonction de  $H^1(\mathbb{R}^d)$ . Dans ce cas, on a  $\nabla \bar{u} = \nabla u 1_\Omega$ .

**Exercice 5** Montrer que si  $u \in H_0^1(\Omega)$  est une fonction telle que  $\nabla u = 0$ , alors  $u = 0$ . Solution : on considère  $\bar{u}$  qui est le prolongement de  $u$  à  $\mathbb{R}^d$  par 0, on a, d'après la Proposition 7,  $\nabla \bar{u} = 0$ . Par conséquent  $\bar{u}$  est une fonction constante, qui ne peut être que nulle puisque  $\bar{u} \in H^1(\mathbb{R}^d)$ , et est donc en particulier de carré intégrable.

**Définition 5** L'espace  $H^{-1}(\Omega)$  est le dual de  $H_0^1(\Omega)$ , c'est-à-dire l'espace des formes linéaires continues sur  $H_0^1(\Omega)$ . C'est un Hilbert séparable inclus dans  $\mathcal{D}'(\Omega)$ .

**Exercice 6** Montrer que les fonctions de  $H^1(\mathbb{R})$  tendent vers 0 à l'infini. Peut-on en dire autant des fonctions de  $L^2(\mathbb{R})$  ?

Solution : Soit  $u \in H^1(\mathbb{R})$ .  $u$  peut être identifiée à une fonction continue, et est donc bien définie en tout point, en écrivant  $u(x) - u(0) = \int_0^x u'(y) dy$  (plus précisément  $u(x) - \int_0^x u'(y) dy$  est une distribution de dérivée nulle, donc c'est une constante, et  $\int_0^x u'(y) dy$  est clairement continue). Soit  $x_n$  une suite qui tend vers  $+\infty$ . On montre que  $|u|^2(x_n)$  est de Cauchy car  $(\int_A^\infty uu')^2 \leq \int_A^\infty u^2 \int_A^\infty (u')^2$  est aussi petit qu'on veut pour  $A$  assez grand. On en déduit que  $|u|^2(x_n)$  converge vers une certaine limite, et donc que  $|u|(x)$  a une limite quand  $x \rightarrow \infty$ . Cette limite est nécessairement 0 car  $u$  est de carré intégrable. Ce résultat est faux pour les fonctions de  $L^2(\mathbb{R})$  (considérer par exemple  $f(x) = \sum_{n \geq 2} 1_{[n-1/n^2; n+1/n^2]}(x)$ ).

## 1.2.2 Trace

**Proposition 8** Soit  $\Omega$  un ouvert borné régulier. On peut alors définir une application linéaire continue et surjective

$$\gamma_0 : \begin{cases} H^1(\Omega) & \rightarrow H^{1/2}(\partial\Omega) \\ f & \mapsto f|_{\partial\Omega} \end{cases}$$

qui prolonge l'application trace pour les fonctions continues sur  $\bar{\Omega}$ . De même, on peut définir une application linéaire continue et surjective

$$\gamma_1 : \begin{cases} H^2(\Omega) & \rightarrow H^{1/2}(\partial\Omega) \\ f & \mapsto \frac{\partial f}{\partial n}|_{\partial\Omega} \end{cases}$$

où  $n$  désigne la normale sortante à  $\Omega$  et  $\frac{\partial f}{\partial n} = \nabla f \cdot n$  est la dérivée normale.

Noter qu'on a besoin de régularité (sur la fonction, et sur l'ouvert) pour définir la trace. On sait bien par exemple que l'on ne peut pas définir la trace de fonctions de  $L^2(\Omega)$ , même pour un ouvert  $\Omega$  très régulier (par exemple  $\Omega = (0, 1)$ ).

**Remarque 10 (Régularité du domaine)** On ne définit pas plus précisément la notion d'ouvert régulier. Grosso modo, un ouvert  $\Omega \subset \mathbb{R}^d$  est de classe  $\mathcal{C}^p$  si localement, il peut être transporté par un difféomorphisme de classe  $\mathcal{C}^p$  sur  $\mathbb{R}^{d-1} \times \mathbb{R}_+$ .

**Proposition 9** Si  $\Omega$  un ouvert borné régulier, alors  $H_0^1(\Omega) = \{u \in H^1(\Omega), \gamma_0(u) = 0\}$  et  $H_0^2(\Omega) = \{u \in H^2(\Omega), \gamma_0(u) = 0, \gamma_1(u) = 0\}$ .

**Exercice 7** En utilisant la Proposition 8, montrer que  $H_0^1(\Omega) \subset \{u \in H^1(\Omega), \gamma_0(u) = 0\}$ . Solution : Comme  $\gamma_0$  est une application continue,  $\gamma_0^{-1}(\{0\}) = \{u \in H^1(\Omega), \gamma_0(u) = 0\}$  est un fermé de  $H^1(\Omega)$ . Ce fermé contient clairement l'ensemble  $C_c^\infty(\Omega)$ . Par conséquent il contient son adhérence, c'est-à-dire  $H_0^1(\Omega)$ .

**Exercice 8** Prouver la Proposition 8 dans le cas  $\Omega = (0, 1)$ . Solution : Soit une fonction  $f \in H^1((0, 1))$ . En considérant  $f - \int_0^x f'(y) dy$ , on montre que  $f$  admet un représentant continu sur  $[0, 1]$ . On peut donc écrire :  $\exists x_0 \in [0, 1], f(x) = \int_0^1 f(y) dy + \int_{x_0}^x f'(x) dx$ . Par conséquent  $|f(1)| = \left| \int_0^1 f(y) dy + \int_{x_0}^1 f'(x) dx \right| \leq \|f\|_{H^1}$  en utilisant Cauchy Schwarz (et de même pour  $f(0)$ ).

L'application trace intervient notamment dans les formules d'intégrations par parties. Par exemple, si  $\Omega$  est borné et régulier, on a :  $\forall u, v \in H^1(\Omega)$ ,

$$\int_{\Omega} \frac{\partial u}{\partial x_i} v = - \int_{\Omega} u \frac{\partial v}{\partial x_i} + \int_{\partial\Omega} u v n_i d\sigma, \quad (4)$$

où  $n$  désigne la normale sortante à  $\Omega$  ( $n_i$  est la  $i$ -ème composante de  $n$ ), et  $\sigma$  la mesure surfacique sur  $\partial\Omega$ . Dans la dernière intégrale,  $u$  et  $v$  sont à comprendre comme  $\gamma_0(u)$  et  $\gamma_0(v)$ .

**Exercice 9** En utilisant (4), montrer que pour  $\Omega$  un ouvert borné régulier,  $u \in H^2(\Omega)$  et  $v \in H^1(\Omega)$ , on a

$$- \int_{\Omega} \Delta u v = \int_{\Omega} \nabla u \cdot \nabla v - \int_{\partial\Omega} \frac{\partial u}{\partial n} v d\sigma.$$

### 1.2.3 Convergence faible

**Définition 6** Soit  $H$  un espace de Hilbert. On dit qu'une suite  $u_n$  de  $H$  converge faiblement vers  $u$  dans  $H$  si et seulement si  $u \in H$  et,  $\forall v \in H$ ,

$$\lim_{n \rightarrow \infty} (u_n, v)_H = (u, v)_H.$$

On note  $u_n \rightharpoonup u$ .

On peut vérifier que la topologie de la convergence faible est *séparée* : la limite d'une suite au sens de la convergence faible, si elle existe, est unique. Si  $H$  est de dimension finie, la convergence au sens faible est équivalente à la convergence au sens fort. Ceci est faux en dimension infinie.

Ainsi :

- $f_n(x) = \sin(2\pi n x)$  converge faiblement vers 0 dans  $L^2((0, 1))$ . (Pour le démontrer, tester contre des fonctions de classe  $C_c^\infty((0, 1))$  et utiliser une intégration par parties, puis raisonner par densité.) Pourtant  $f_n$  ne converge pas fortement vers 0 dans  $L^2((0, 1))$ .

- On considère  $f_n(x) = \sqrt{n}\chi(nx) \in L^2((0,1))$  où  $\chi : \mathbb{R} \rightarrow \mathbb{R}$  est une fonction à support dans  $(0,1)$  et telle que  $\int \chi^2 = 1$ . On vérifie que  $f_n$  converge faiblement dans  $L^2((0,1))$  vers 0. Pourtant,  $f_n$  ne converge pas fortement vers 0 dans  $L^2((0,1))$  (puisque  $\|f_n\|^2 = 1$ ).
- $f_n(x) = \chi(x-n)$  où  $\chi$  est une fonction  $\mathcal{C}^\infty$  à support dans  $(0,1)$  converge faiblement vers 0 dans  $L^2(\mathbb{R})$ . (Le prouver tout d'abord pour des fonctions  $\mathcal{C}_c^\infty(\mathbb{R})$  puis raisonner par densité.) Pourtant  $f_n$  ne converge pas fortement vers 0 dans  $L^2(\mathbb{R})$ .

**Proposition 10** Soit  $(u_n)_{n \geq 0}$  une suite de  $H$ .

1. Si  $u_n$  converge vers  $u$  fortement (i.e.  $\lim_{n \rightarrow \infty} \|u_n - u\|_H = 0$ ), alors  $u_n$  converge vers  $u$  faiblement.
2. Si  $u_n$  converge vers  $u$  faiblement, alors la suite  $u_n$  est bornée dans  $H$  et  $\|u\|_H \leq \liminf_{n \rightarrow \infty} \|u_n\|_H$ .
3. Si  $u_n$  converge vers  $u$  faiblement et  $v_n$  converge vers  $v$  fortement, alors  $\lim_{n \rightarrow \infty} (u_n, v_n)_H = (u, v)_H$ .

Comme le montre l'exercice suivant, les points 1 et 3 ne sont pas compliqués à démontrer. Le point 2 repose sur un argument plus profond : le théorème de Banach-Steinhaus (cf. [1, Théorème II.1 et Proposition III.5]).

**Exercice 10** Démontrer le premier et le dernier point. Donner un exemple d'une suite de fonctions  $u_n$  qui converge faiblement vers  $u$  dans  $L^2((0,1))$ , et telle que  $\int_0^1 (u_n)^2$  ne converge pas  $\int_0^1 u^2$ .

Solution : Pour la première question, utiliser Cauchy-Schwarz, et le fait que la suite  $u_n$  est bornée dans  $L^2$ . Pour la deuxième question, on peut considérer par exemple une base orthonormée de  $L^2(0,1)$  (on a alors  $u_n$  qui tend faiblement vers 0 puisque  $\sum_{n \geq 1} |(u, u_n)|^2 = \|u\|_{L^2}^2 < \infty$ , alors que  $\|u_n\|_{L^2}^2 = 1$ ).

**Exercice 11** Soit  $u_n$  une suite de  $H$  qui converge faiblement vers  $u$  dans  $H$ . Montrer que  $u_n$  converge fortement si et seulement si  $\|u_n\|$  converge vers  $\|u\|$ .

Solution : On a  $\|u_n - u\|^2 = \|u_n\|^2 + \|u\|^2 - 2(u_n, u)$ . Comme  $u_n$  converge faiblement vers  $u$ , on a  $\lim_{n \rightarrow \infty} (u_n, u) = \|u\|^2$ . On en déduit le résultat.

**Exercice 12** Si  $u_n$  converge faiblement vers  $u$  dans  $L^2(\Omega)$ , que dire de la convergence de  $u_n$  au sens des distributions ?

**Exercice 13** 1. Montrer que  $u_n \in H$  converge faiblement vers  $u$  si et seulement si, pour toute forme linéaire continue  $\phi : H \rightarrow \mathbb{R}$ ,  $\phi(u_n)$  converge vers  $\phi(u)$ .

2. Soit  $H_1$  et  $H_2$  deux espaces de Hilbert, et  $T : H_1 \rightarrow H_2$  une application linéaire continue (pour la topologie forte). Montrer que  $T$  est continue pour la topologie faible. Solution : Soit  $u_n \in H_1$  qui converge faiblement vers  $u$  et  $\phi$  une forme linéaire continue sur  $H_2$ . On vérifie que  $\phi \circ T$  est une forme linéaire continue sur  $H_1$  et donc  $\phi(T(u_n))$  converge vers  $\phi(T(u))$ . Ceci montre que  $T(u_n)$  converge faiblement vers  $T(u)$ .

3. Soit  $f_n$  une suite de fonctions de  $H^1(\Omega)$  qui converge faiblement vers  $f$ . Vérifier que  $f_n$  converge faiblement vers  $f$  dans  $L^2(\Omega)$  et en déduire que  $f_n$  converge vers  $f$  au sens des distributions.

**Exercice 14** Soit  $H_1$  et  $H_2$  deux espaces de Hilbert et  $T : H_1 \rightarrow H_2$  une application linéaire compacte : pour tout borné  $B \subset H_1$ ,  $T(B)$  est relativement compact dans  $H_2$  (i.e. que  $\overline{T(B)}$  est compact).

1. Montrer que  $T$  est continue. Indication : il suffit de montrer que  $\sup_{x, \|x\| \leq 1} \|T(x)\| < \infty$ .
2. Soit  $u_n$  une suite de  $H_1$  qui converge faiblement vers  $u$ . Montrer que  $T(u_n)$  converge fortement vers  $T(u)$ . Solution :  $T$  est une application continue, donc  $T(u_n)$  converge faiblement vers  $T(u)$  (cf. Exercice 13). Par ailleurs, on sait que  $u_n$  est une suite bornée, donc on peut extraire de la suite  $T(u_n)$  une sous-suite qui converge fortement vers un élément  $v \in H_2$ . Comme la convergence forte implique la convergence faible, nécessairement,  $v = T(u)$ . Ainsi la suite  $T(u_n)$  est une suite à valeur dans un compact (la fermeture de l'ensemble des  $(T(u_n))_{n \in \mathbb{N}}$ ) qui admet une seule valeur d'adhérence ( $T(u)$ ), donc  $T(u_n)$  converge vers  $T(u)$ .

L'intérêt essentiel de la convergence faible est le résultat suivant :

**Proposition 11** La boule unité de  $H$  est faiblement compacte. Autrement dit, de toute suite bornée dans  $H$ , on peut extraire une sous-suite qui converge faiblement dans  $H$ .

**Remarque 11** Dans un espace de Hilbert de dimension infinie (et en fait, dans tout espace normé de dimension infinie), la boule unité de  $H$  n'est jamais compacte pour la topologie forte. On renvoie à [1, Théorème VI.5]. On comprend alors l'intérêt de la notion de convergence faible.

**Exercice 15** Soit la fonction  $f_n(x) = \frac{1}{\sqrt{n}} 1_{[0,n]}(x)$ . Vérifier que  $\|f_n\|_{L^2} = 1$ . En déduire que  $f_n$  converge faiblement dans  $L^2$  et identifier sa limite. La suite converge-t-elle fortement dans  $L^2$  ? Solution : utiliser la compacité des fonctions  $\mathcal{C}_0^\infty(\mathbb{R})$  dans  $L^2(\mathbb{R})$ .

**Exercice 16** On se propose de donner une preuve de la Proposition 11 dans le cas où il existe une base hilbertienne  $(e_k)_{k \in \mathbb{N}}$  de  $H$ . On considère  $(x_n)$  une suite bornée de  $H$ .

1. Montrer que pour tout  $k$ ,  $((x_n, e_k))_{n \in \mathbb{N}}$  est une suite bornée de  $\mathbb{R}$ .
2. Justifier alors l'existence de fonctions  $\phi_k$  strictement croissantes de  $\mathbb{N}$  dans  $\mathbb{N}$  telles que pour tout  $k$ ,  $\lim_{n \rightarrow \infty} (x_{\phi_1 \circ \phi_2 \circ \dots \circ \phi_k(n)}, e_k) = a_k \in \mathbb{R}$ .
3. On définit pour tout  $n$ ,  $y_n = x_{\phi_1 \circ \phi_2 \circ \dots \circ \phi_n(n)}$ . Vérifier que  $y_n$  est une sous-suite de  $x_n$  et que, de plus,  $\forall k \in \mathbb{N}$ ,  $\lim_{n \rightarrow \infty} (y_n, e_k) = a_k$ .
4. Montrer que  $(y_n)$  est une suite qui converge faiblement vers une limite que l'on précisera.
5. Conclure.

La topologie faible est liée à la convexité.

**Proposition 12** Soit  $C \subset H$  un convexe<sup>3</sup> dans le Hilbert  $H$ . Alors  $C$  est faiblement fermé si et seulement si  $C$  est fortement fermé.

Tout ensemble fermé pour la topologie faible est fermé pour la topologie forte. La réciproque est fautive en général, sauf si l'ensemble est convexe. Ainsi  $H_0^1(\Omega)$  est faiblement fermé, *i.e.* est fermé pour la topologie de la convergence faible dans  $H^1(\Omega)$ .

**Proposition 13** Soit  $J : C \rightarrow ]-\infty, +\infty]$  une fonction convexe et s.c.i. (pour la topologie forte). Alors  $J$  est s.c.i. pour la topologie faible. En particulier, pour toute suite  $u_n$  qui tend faiblement vers  $u$ ,

$$J(u) \leq \liminf J(u_n).$$

On rappelle qu'une fonction  $J : H \rightarrow ]-\infty, +\infty]$  est s.c.i. (pour la topologie forte) si et seulement si pour tout  $x \in H$ , pour tout  $\epsilon > 0$ , il existe un voisinage  $V$  de  $x$  tel que, pour tout  $y \in V$ ,  $J(y) \geq J(x) - \epsilon$ .

**Exercice 17** 1. Montrer que  $J : H \rightarrow ]-\infty, +\infty]$  est une fonction s.c.i. si et seulement si pour tout  $\lambda \in \mathbb{R}$ , l'ensemble  $\{x \in H, \phi(x) \leq \lambda\}$  est fermé.

2. En utilisant la Proposition 12, en déduire que si  $J : H \rightarrow ]-\infty, +\infty]$  une fonction s.c.i. pour la topologie forte, alors  $J$  est une fonction s.c.i. pour la topologie faible.

3. Soit  $J : H \rightarrow ]-\infty, +\infty]$  une fonction s.c.i. Montrer que pour toute suite  $u_n$  qui tend vers  $u$ ,

$$J(u) \leq \liminf J(u_n).$$

En utilisant cette propriété, montrer le point 2 de la Proposition 10.

## 1.2.4 Inclusions de Sobolev

On commence par rappeler l'inégalité de Hölder :

**Lemme 3** Soit  $f \in L^p(\Omega)$  et  $g \in L^{p'}(\Omega)$  avec  $1/p + 1/p' = 1$  et  $1 \leq p, p' \leq \infty$ , alors

$$\int_{\Omega} |fg| \leq \|f\|_{L^p(\Omega)} \|g\|_{L^{p'}(\Omega)}.$$

**Exercice 18** Soit  $\Omega$  un domaine borné et  $1 \leq p < q \leq \infty$ . Montrer l'inclusion continue  $L^q(\Omega) \subset L^p(\Omega)$ .

Solution : Pour une fonction  $h \in L^q$ , utiliser l'inégalité de Hölder, avec  $g = 1$ ,  $f = |h|^p$  et l'exposant  $r = q/p$ .

**Exercice 19** Montrer que si  $f \in L^p(\Omega) \cap L^q(\Omega)$ , avec  $1 \leq p < q \leq \infty$ , alors pour tout  $p \leq r \leq q$ ,  $f \in L^r(\Omega)$  et

$$\|f\|_{L^r(\Omega)} \leq \|f\|_{L^p(\Omega)}^{\alpha} \|f\|_{L^q(\Omega)}^{1-\alpha},$$

où  $\alpha$  est tel que  $1/r = \alpha/p + (1-\alpha)/q$ .

Solution : Pour une fonction  $f \in L^p(\Omega) \cap L^q(\Omega)$ , utiliser l'inégalité de Hölder, avec  $g = |f|^{\alpha r}$ ,  $h = |f|^{(1-\alpha)r}$  et les exposants conjugués  $p/(\alpha r)$  et  $q/((1-\alpha)r)$ .

<sup>3</sup>On rappelle que  $C$  est convexe si et seulement si :  $\forall u, v \in C, \forall \theta \in [0, 1], \theta u + (1-\theta)v \in C$ .

**Exercice 20** Soit  $f$  et  $g$  deux fonctions de  $L^2(\Omega)$ . Montrer l'inégalité : pour  $\varepsilon > 0$ ,

$$\int_{\Omega} |fg| \leq \varepsilon \|f\|_{L^2(\Omega)}^2 + \frac{1}{4\varepsilon} \|g\|_{L^2(\Omega)}^2.$$

Solution : Par Cauchy Schwarz, on a  $\int fg \leq \|f\|_{L^2} \|g\|_{L^2}$ . On utilise ensuite l'inégalité  $ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2$  qui se déduit du cas bien connu  $\varepsilon = 1/2$ .

**Proposition 14 (Inclusions de Sobolev)** Soit  $\Omega$  un ouvert régulier. On a les injections continues : pour  $k \geq 1$  un entier,

- si  $d > 2k$ , alors  $H^k(\Omega) \subset L^{p^*}(\Omega)$  avec  $\frac{1}{p^*} = \frac{1}{2} - \frac{k}{d}$ .
- si  $d = 2k$ , alors  $H^k(\Omega) \subset L^q(\Omega)$ , pour tout  $q \in [2, \infty[$ .
- si  $d < 2k$ , alors  $H^k(\Omega) \subset \mathcal{C}(\overline{\Omega})$ .

Dans ce cas, si de plus  $k - d/2 > 0$  n'est pas un entier, on a  $H^k(\Omega) \subset \mathcal{C}^{n,1/2}(\overline{\Omega})$ , avec  $n = \lfloor k - d/2 \rfloor$  où  $\mathcal{C}^{n,1/2}(\overline{\Omega})$  désigne les fonctions de classe  $\mathcal{C}^n(\overline{\Omega})$  de dérivées  $n$ -ième  $1/2$ -höldérienne : pour toute dérivée  $n$ -ième  $v$  de  $u$ ,  $\exists C > 0, \forall x, y \in \overline{\Omega}, |v(x) - v(y)| \leq C|x - y|^{1/2}$ .

**Exercice 21** Dans le cas  $\Omega = \mathbb{R}^d$ , en utilisant la définition des espaces de Sobolev en terme de transformée de Fourier (cf. Remarque 9), montrer que si  $d < 2k$ ,  $H^k(\mathbb{R}^d) \subset \mathcal{C}_0(\mathbb{R}^d)$ , où  $\mathcal{C}_0(\mathbb{R}^d)$  désigne l'espace des fonctions continues sur  $\mathbb{R}^d$  qui tendent vers 0 à l'infini, muni de la norme  $L^\infty$ .

**Proposition 15 (Inclusions de Sobolev compactes. Théorème de Rellich-Kondrachov)**

Soit  $\Omega$  un ouvert régulier borné. On a les injections compactes :

- si  $d > 2k$ , alors  $H^k(\Omega) \subset L^q(\Omega)$ , pour tout  $q \in [1, p^*[, \frac{1}{p^*} = \frac{1}{2} - \frac{k}{d}$ .
- si  $d = 2k$ , alors  $H^k(\Omega) \subset L^q(\Omega)$ , pour tout  $q \in [1, \infty[$ .
- si  $d < 2k$ , alors  $H^k(\Omega) \subset \mathcal{C}(\overline{\Omega})$ .

Si  $\Omega$  est borné, on a ainsi  $H^1(\Omega) \subset L^2(\Omega)$  (quelque soit la dimension  $d$ ) avec injection compacte. On rappelle que cela signifie que l'image de tout borné de  $H^1(\Omega)$  dans  $L^2(\Omega)$  est relativement compact (cf. Exercice 14). En terme de suites, cela se traduit par les propriétés suivantes (cf. Exercices 13 et 14) :

- Si  $u_n$  est une suite de  $H^1(\Omega)$  qui converge faiblement vers  $u$  dans  $H^1(\Omega)$ , alors  $u_n$  converge fortement vers  $u$  dans  $L^2(\Omega)$ , et  $\nabla u_n$  converge faiblement vers  $\nabla u$  dans  $L^2(\Omega)$ .
- De toute suite bornée dans  $H^1(\Omega)$  on peut extraire une sous-suite qui converge faiblement dans  $H^1(\Omega)$  et fortement dans  $L^2(\Omega)$ .

On énonce maintenant un théorème très important en pratique

**Lemme 4 (Inégalité de Poincaré)** Soit  $\Omega$  un domaine borné. Alors il existe une constante  $C > 0$  (appelée constante de Poincaré du domaine  $\Omega$ ) tel que  $\forall u \in H_0^1(\Omega)$ ,

$$\int_{\Omega} u^2 \leq C \int_{\Omega} |\nabla u|^2. \quad (5)$$



*Preuve* : Raisonnons par l'absurde. Si il n'existe pas de constante  $C$  telle que (5) ait lieu, alors on peut trouver une suite  $u_n \in H_0^1(\Omega)$  telle que  $\int_{\Omega} (u_n)^2 = 1$  et

$$\int_{\Omega} |\nabla u_n|^2 < 1/n. \quad (6)$$

La suite  $u_n$  est bornée dans  $H^1(\Omega)$  et donc on peut supposer (quitte à extraire une sous-suite) que  $u_n$  converge vers une fonction  $u$  de  $H^1(\Omega)$ , la convergence étant dans  $H^1$ -faible et dans  $L^2$ -fort. De l'inégalité (6) et du fait que  $u_n$  converge dans  $L^2$ -fort, on déduit que  $u_n$  est en fait une suite de Cauchy dans  $H^1(\Omega)$  :  $\forall \epsilon > 0, \exists n_0, \forall n \geq m \geq n_0,$

$$\|u_n - u_m\|_{L^2} + \|\nabla u_n - \nabla u_m\|_{L^2} \leq \epsilon.$$

La convergence de  $u_n$  vers  $u$  est donc forte dans  $H^1(\Omega)$ . On en déduit que  $u \in H_0^1(\Omega)$ . Mais on a par ailleurs  $\int_{\Omega} u^2 = 1$  et  $\int_{\Omega} |\nabla u|^2 = 0$ .  $u$  est donc une fonction constante nulle au bord et de norme  $L^2$  égale à 1. D'où la contradiction.  $\diamond$

Le lemme 4 montre que  $u \mapsto \int_{\Omega} |\nabla u|^2$  est une norme sur  $H_0^1(\Omega)$ .

**Exercice 22** Soit une suite bornée dans  $H_0^1(\Omega)$ . D'après ce qui précède, on peut extraire une sous-suite qui converge faiblement dans  $H^1(\Omega)$ . Pourquoi la limite est dans  $H_0^1(\Omega)$  ?

Solution : Méthode 1 :  $H_0^1(\Omega)$ , muni du produit scalaire de  $H^1$  ou bien du produit scalaire  $\int \nabla u \cdot \nabla v$ , est un espace de Hilbert, sur lequel on peut donc considérer la topologie de la convergence faible. La convergence faible dans  $H^1$  et dans  $H_0^1$  impliquent la convergence au sens des distributions, et on peut donc identifier les limites. Méthode 2 : utiliser la Proposition 12.

**Exercice 23** – Montrer que l'inégalité de Poincaré (5) est aussi valable sur l'espace  $V = \{u \in H^1(\Omega), u = 0 \text{ sur } \Gamma\}$ , pour un domaine (i.e. un ouvert connexe)  $\Omega$  borné et un ouvert  $\Gamma$  non vide de  $\partial\Omega$ .

– En déduire que pour un domaine borné  $\Omega$ , il existe une constante  $C > 0$  tel  $\forall u \in H^1(\Omega)$ ,

$$\int_{\Omega} \left( u - \frac{1}{|\Omega|} \int_{\Omega} u \right)^2 \leq C \int_{\Omega} |\nabla u|^2. \quad (7)$$

Cette inégalité s'appelle l'inégalité de Poincaré-Wirtinger.

Solution : Cela revient à montrer que l'inégalité (7) est valable sur  $V = \{u \in H^1(\Omega), \int_{\Omega} u = 0\}$ .

**Remarque 12** On s'est restreint dans la Section 1.2 à des espaces de Sobolev définis par rapport à l'espace  $L^2(\Omega)$ . On peut de même définir des espaces de Sobolev par rapport à l'espace  $L^p(\Omega)$ , pour  $1 \leq p \leq \infty$ . Ainsi

$$W^{m,p}(\Omega) = \{f \in L^p(\Omega), \forall \alpha \text{ multi-indice de longueur } |\alpha| \leq k, \exists g_{\alpha} \in L^p(\Omega), \partial^{\alpha} T_f = g_{\alpha} \text{ dans } \mathcal{D}'(\Omega)\}.$$

L'avantage du cas  $p = 2$  est que les espaces sont naturellement munis d'un produit scalaire pour lequel ils sont des espaces de Hilbert. Pour  $p \neq 2$ , on a simplement des espaces de Banach. Retenons que pour  $1 < p < \infty$ , les propriétés que nous avons énoncées se généralisent à  $W^{m,p}$ . Les cas  $p = 1$  et  $p = \infty$  peuvent par contre présenter des particularités (essentiellement liées au fait que le dual de  $L^{\infty}$  est strictement plus grand que  $L^1$ ).

## 2 Analyse des problèmes elliptiques et paraboliques

On suppose dans tout ce qui suit et pour simplifier la présentation que  $\Omega$  est un domaine (*i.e.* un ouvert connexe) de  $\mathbb{R}^d$  borné (en particulier pour utiliser l'inégalité de Poincaré (5)) et régulier (pour pouvoir définir la trace, cf. Section 1.2.2). Plusieurs des résultats qui suivent se généralisent au cas où  $\Omega$  n'est pas borné ou régulier. De plus, on se limite à des équations simples (du type équation de Poisson ou équation de la chaleur) mais les résultats et les méthodes de démonstration se généralisent à des équations elliptiques ou paraboliques, moyennant une régularité suffisante sur les coefficients.

### 2.1 Motivation

Les équations paraboliques se rencontrent dans de nombreux domaines. Donnons quelques exemples.

En physique, de nombreuses équations de conservation s'écrivent sous la forme d'une équation parabolique. Prenons l'exemple de l'équation qui régit l'évolution de la température  $T(t, x)$  pour  $t \geq 0$  et  $x \in \Omega$  :

$$\begin{cases} \partial_t T - \operatorname{div}(K \nabla T) = 0, \forall t \geq 0, \forall x \in \Omega, \\ T(0, x) = T_0(x), \forall x \in \Omega, \\ T(t, x) = f(x), \forall t \geq 0, \forall x \in \partial\Omega. \end{cases} \quad (8)$$

La matrice  $K$  (qui peut éventuellement dépendre des variables  $(t, x)$ ) est la matrice de conductivité thermique et  $K \nabla T$  donne le flux de chaleur (loi de Fourier-Fick). L'équation (8) est donc une équation de conservation, qui exprime le fait que la variation de température au cours du temps en un point s'obtient en comptabilisant les flux rentrant et sortant. En effet, en intégrant l'équation sur un petit élément  $E \subset \Omega$  et en utilisant la formule de Stokes (cf. Lemme 2), on a :

$$\frac{d}{dt} \int_E T(t, x) dx = \int_{\partial E} K \nabla T \cdot n,$$

où  $n$  désigne la normale sortante à  $E$ .

En finance, l'équation de Black-Scholes qui donne le prix d'une option européenne de payoff  $\phi$  et de maturité  $T$  s'écrit :

$$\begin{cases} \partial_t P + rS \partial_S P + \frac{\sigma^2 S^2}{2} \partial_{S,S}^2 P - rP = 0, \forall t \in [0, T], \forall S \in \mathbb{R}_+, \\ P(T, S) = \phi(S), \forall S \in \mathbb{R}_+. \end{cases} \quad (9)$$

La fonction  $P$  dépend du temps  $t$  et de la valeur du sous-jacent  $S$  sur lequel porte l'option. Le taux d'intérêt  $r$  peut dépendre du temps, et la volatilité  $\sigma$  peut dépendre du temps et de la valeur du sous-jacent.

**Exercice 24** Expliquer pourquoi l'équation de Black-Scholes (9) est bien du même type que l'équation de la chaleur (8).

L'évolution de la concentration d'un polluant dans un fluide de champ de vitesse  $u(t, x)$  donné satisfait l'équation aux dérivées partielles :

$$\begin{cases} \partial_t c + u \cdot \nabla c - \operatorname{div}(K \nabla c) = 0, \forall t \geq 0, \forall x \in \Omega, \\ c(0, x) = c_0(x), \forall x \in \Omega, \\ c(t, x) = f(x), \forall t \geq 0, \forall x \in \partial\Omega. \end{cases} \quad (10)$$

On suppose que le fluide est incompressible, et donc le champ de vitesse est à divergence nulle :  $\operatorname{div}(u) = 0$ . Ici,  $K$  désigne un tenseur de diffusion. Le terme  $\operatorname{div}(K\nabla c)$  est appelé terme de diffusion. Il modélise le fait que le polluant diffuse dans le fluide, sous l'influence de l'agitation thermique. Le terme  $u \cdot \nabla c$  est appelé terme d'advection. Il modélise le fait que la concentration de polluant évolue par transport du polluant par le fluide.

**Exercice 25** *On suppose que  $K = 0$ , et donc que l'advection est le phénomène prédominant dans l'évolution de  $c$ . On suppose de plus que  $\Omega = \mathbb{R}^d$  et que la vitesse  $u$  est régulière. Ecrire une forme analytique de la solution de l'équation (10) en fonction de la condition initiale  $c_0$  et de la solution de l'équation différentielle ordinaire :*

$$\begin{cases} \frac{d}{dt} X_t^x = u(t, X_t^x), \\ X_0^x = x. \end{cases} \quad (11)$$

## 2.2 Problèmes elliptiques

### 2.2.1 Le laplacien de Dirichlet

On considère le problème de Laplace

$$\begin{cases} -\Delta u = f \text{ sur } \Omega, \\ u = 0 \text{ sur } \partial\Omega. \end{cases} \quad (12)$$

On propose d'adopter une approche variationnelle pour construire une solution. Rappelons comment écrire une formulation variationnelle du problème (12).

**Etape 1 :** *Construction de la formulation variationnelle.*

On multiplie par une fonction test, on intègre sur  $\Omega$  et on intègre par parties, sans se soucier pour l'instant de la justification. On obtient

$$\int_{\Omega} \nabla u \cdot \nabla v - \int_{\partial\Omega} \frac{\partial u}{\partial n} v = \int_{\Omega} f v. \quad (13)$$

L'objectif est de faire jouer à  $u$  et  $v$  un rôle symétrique : on veut que la formulation variationnelle ait un sens pour des  $u$  et  $v$  de même régularité. On voit ainsi que  $u, v \in H^1(\Omega)$  est approprié pour donner un sens à (13). Pour tenir compte de la condition aux limites  $u = 0$  sur  $\partial\Omega$ , on impose en fait  $u \in H_0^1(\Omega)$  ce qui a bien un sens d'après la Proposition 8. La formulation variationnelle du problème est donc :

$$\text{trouver } u \in H_0^1(\Omega), \text{ tel que } \forall v \in H_0^1(\Omega), \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v. \quad (14)$$

Pour que le second membre ait un sens, on suppose dans la suite que  $f \in L^2(\Omega)$ . Remarque que le problème est également bien défini avec  $f \in H^{-1}(\Omega)$ , puisque  $v \in H_0^1(\Omega)$ .

**Etape 2 :** *Résolution du problème sous forme variationnelle.*

On invoque ici le théorème de Lax-Milgram que l'on rappelle.

**Lemme 5 (Théorème de Lax-Milgram)** Soit  $V$  un espace de Hilbert et  $a$  une forme bilinéaire continue et coercive sur  $V$ , i.e. telle que (continuité)

$$\exists M > 0, \forall u, v \in H, |a(u, v)| \leq M \|u\| \|v\|,$$

et (coercivité)

$$\exists \alpha > 0, \forall u \in H, a(u, u) \geq \alpha \|u\|^2.$$

Soit  $L$  une forme linéaire continue sur  $H$  i.e. telle que

$$\exists C > 0, \forall v \in H, |L(v)| \leq C \|v\|.$$

Alors le problème suivant :

$$\text{trouver } u \in V, \text{ tel que } \forall v \in V, a(u, v) = L(v)$$

admet une unique solution. De plus on a

$$\|u\| \leq \frac{C}{\alpha}.$$

Ce théorème permet de montrer que le problème (14) est bien posé.

**Etape 3 : Interprétation des résultats.**

On a donc obtenu une solution au problème variationnel (14). Une question naturelle est alors de savoir en quel sens on a résolu le problème initial (12). Il est facile de voir que l'équation  $-\Delta u = f$  est vérifiée au sens des distributions. Or  $f \in L^2(\Omega)$  donc  $-\Delta u \in L^2(\Omega)$ , et l'égalité a donc lieu presque partout. On a de même  $u = 0$  presque partout sur  $\partial\Omega$  (pour la mesure surfacique sur  $\partial\Omega$ ).

On comprend que quelques tâtonnements sont possibles dans la première étape. L'important est d'obtenir une formulation variationnelle à partir de laquelle on raisonne rigoureusement : existence et unicité du résultat, puis retour au problème initial (en répondant à la question : en quel sens a-t'on construit une solution du problème initial ?).

**Remarque 13** *Le lecteur curieux se pose peut-être quelques questions sur la manière d'obtenir une "bonne formulation variationnelle". Pourquoi est-ce que la formulation variationnelle (14) du problème (12) est "la bonne" ? Pourquoi ne pas, par exemple, intégrer une nouvelle fois par parties et chercher une solution en un sens encore plus faible, par exemple : trouver  $u \in L^1(\Omega)$  tel que pour  $v \in C_c^\infty(\Omega)$ ,*

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

*Plusieurs réponses sont possibles. Tout d'abord, en vue d'utiliser le théorème de Lax-Milgram, on a envie de faire jouer à  $u$  et  $v$  un rôle symétrique, comme expliqué dans l'étape 1 ci-dessus. De plus, il faut garder un tête un principe général de l'analyse : plus on cherche des solutions en un sens faible, plus il est facile de prouver l'existence de solutions, mais plus il est difficile d'obtenir l'unicité de la solution. Ici, l'espace fonctionnel  $H^1(\Omega)$  est un bon cadre car il permet d'avoir l'existence et l'unicité de la solution. En travaillant dans  $L^1(\Omega)$  seulement, on ne pourrait pas, par exemple, donner de sens à  $u = 0$  sur  $\partial\Omega$ .*

**Exercice 26** Ecrire une formulation variationnelle pour les problèmes suivants

$$- \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left( a_{i,j} \frac{\partial u}{\partial x_j} \right) = f \text{ dans } \Omega \text{ et } u = 0 \text{ sur } \partial\Omega, \text{ où } a \text{ est une fonction bornée.}$$

$$- \Delta u + b \cdot \nabla u = f \text{ dans } \Omega \text{ et } u = 0 \text{ sur } \partial\Omega, \text{ où } b \text{ est une fonction bornée.}$$

Donner des hypothèses naturelles sur  $a$  et  $b$  pour que ces problèmes soient bien posés.

Solution :  $a$  borné inférieurement (au sens matriciel) par une constante strictement positive,  $b$  borné supérieurement par une constante suffisamment petite (par rapport à l'inverse de la constante de Poincaré du domaine), ou bien  $\operatorname{div}(b)$  suffisamment petit (par rapport à l'inverse du carré de la constante de Poincaré du domaine).

**Remarque 14** Les problèmes symétriques peuvent s'interpréter en terme de minimisation. Ainsi, on peut vérifier que la solution du problème (12) est aussi la fonction qui réalise le minimum de :

$$\min_{u \in H_0^1(\Omega)} \frac{1}{2} \int_{\Omega} |\nabla u|^2 - \int_{\Omega} f u.$$

**Exercice 27** En utilisant la surjectivité de l'application trace  $\gamma_0$ , étudier le problème suivant : pour  $g \in H^{1/2}(\Omega)$ , trouver  $u \in H^1(\Omega)$  tel que

$$\begin{cases} -\Delta u = f \text{ sur } \Omega, \\ u = g \text{ sur } \partial\Omega. \end{cases} \quad (15)$$

Montrer que  $u$  satisfait la formulation variationnelle :

$$\text{trouver } u \in \{v \in H^1(\Omega), v = g \text{ sur } \partial\Omega\}, \text{ tel que } \forall v \in H_0^1(\Omega), \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} f v. \quad (16)$$

Solution : considérer  $v = u - \bar{u}$  où  $\bar{u} \in H^1(\Omega)$  est un relèvement de la condition au bord, c'est-à-dire est tel que  $\bar{u} = g$  sur  $\partial\Omega$ .

Précisons un point de terminologie : pour une solution d'une équation aux dérivées partielles, on parle de *solution classique* du problème (12) si les dérivées sont définies au sens classique, i.e. si  $u \in C^2(\Omega)$ . On parle de *solution forte* du problème (12) si les dérivées sont des fonctions  $L_{\text{loc}}^1$ , i.e. si  $u \in H^2(\Omega)$ . Si les dérivées ne sont définies qu'au sens des distributions, on parle de *solution faible* : c'est le cas pour (12) si  $u$  est seulement une fonction de  $H^1(\Omega)$ . Par ailleurs, la formulation (12) est appelée *formulation forte* et la formulation (14) est appelée *formulation faible*.

## 2.2.2 Le laplacien de Neumann

On considère maintenant le problème

$$\begin{cases} -\Delta u = f \text{ sur } \Omega, \\ \frac{\partial u}{\partial n} = g \text{ sur } \partial\Omega. \end{cases} \quad (17)$$

**Etape 1** : Construction de la formulation variationnelle.

On obtient facilement la formulation variationnelle suivante :

$$\text{trouver } u \in H^1(\Omega), \text{ tel que } \forall v \in H^1(\Omega), \int_{\Omega} \nabla u \cdot \nabla v = \int_{\partial\Omega} gv + \int_{\Omega} fv. \quad (18)$$

Il est important de remarquer que la condition aux limites s'insère ici de manière *naturelle* dans la formulation variationnelle et n'est pas incluse dans l'espace fonctionnel pour  $u$  (d'ailleurs demander à une fonction de  $H^1(\Omega)$  d'avoir une dérivée normale égale à une fonction  $g$  au bord n'a pas de sens). La situation est donc différente du cas des conditions aux limites de Dirichlet pour lesquelles on tient compte des conditions aux limites de manière *essentielle* dans l'espace fonctionnel<sup>4</sup>.

On voit également que l'on demande typiquement  $f \in L^2(\Omega)$  et  $g \in L^2(\partial\Omega)$ , pour que (18) ait un sens.

**Etape 2 :** *Résolution du problème sous forme variationnelle.*

A ce stade, il est important de remarquer que le problème ne peut avoir de solution que si  $f$  et  $g$  vérifient une condition de compatibilité (prendre  $v = 1$  dans (18)) :

$$\int_{\Omega} f + \int_{\partial\Omega} g = 0. \quad (19)$$

Un deuxième point à noter est que si  $u$  est solution,  $u$  plus une constante est également solution. On ne peut donc pas espérer prouver qu'il existe une unique solution au problème sans restreindre l'espace fonctionnel pour  $u$ . Pour éliminer l'indétermination due à cette constante additive, on introduit (par exemple)  $V = \{u \in H^1(\Omega), \int_{\Omega} u = 0\}$  et on pose le problème sous la forme :

$$\text{trouver } u \in V, \text{ tel que } \forall v \in V, \int_{\Omega} \nabla u \cdot \nabla v = \int_{\partial\Omega} gv + \int_{\Omega} fv. \quad (20)$$

En utilisant l'inégalité de Poincaré-Wirtinger, on montre que le problème est bien coercif et donc qu'il admet une unique solution.

**Exercice 28** *Quelle est la différence entre la formulation variationnelle du problème de Dirichlet homogène ( $u = 0$  sur  $\partial\Omega$ ), et du problème de Neumann homogène ( $\partial u / \partial n = 0$  sur  $\partial\Omega$ ) ?*

**Remarque 15** *Si (19) n'est pas vérifié, on peut vérifier que (20) résout le problème initial, avec un  $f$  (ou un  $g$ ) modifié d'une constante additive, de manière à satisfaire (19).*

**Etape 3 :** *Interprétation des résultats.*

Cette étape est un peu plus compliquée que dans le cas du Laplacien de Dirichlet. En prenant  $\phi \in C_c^\infty(\Omega)$ , et on testant contre  $v = \left(\phi - \frac{1}{|\Omega|} \int_{\Omega} \phi\right) \in V$ , on vérifie facilement que le problème  $-\Delta u = f$  est vérifié au sens des distributions, et donc dans  $L^2(\Omega)$  (et en particulier presque partout). Par contre, donner un sens à la condition aux limites  $\frac{\partial u}{\partial n} = g$  sur  $\partial\Omega$  pose des difficultés car  $\frac{\partial u}{\partial n}$  n'est pas défini pour

---

<sup>4</sup>Les qualificatifs *naturelle* et *essentielle* sont des termes consacrés pour désigner la manière dont on tient compte d'une condition aux limites.

$u \in V$ . Pour pouvoir donner un sens à la condition aux limites, il faut supposer un peu plus de régularité sur  $u$ . Admettons pour le moment que  $u \in H^2(\Omega)$  (ceci découle d'un résultat énoncé ci-dessous, cf. Proposition 17, et nécessite  $g \in H^{1/2}(\partial\Omega)$ ), on peut alors intégrer par parties la formulation variationnelle pour obtenir

$$\int_{\Omega} (-\Delta u - f) v = \int_{\partial\Omega} \left( g - \frac{\partial u}{\partial n} \right) v.$$

Cette relation est valable pour toute fonction  $v \in H^1(\Omega)$  (en utilisant la relation de compatibilité (19)). Comme  $-\Delta u - f = 0$  presque partout, on a en fait  $\int_{\partial\Omega} \left( g - \frac{\partial u}{\partial n} \right) v = 0$ . On peut prendre ici n'importe quelle fonction  $v \in H^{1/2}(\partial\Omega)$  d'après la Proposition 8, et en particulier,  $v = g - \frac{\partial u}{\partial n}$ . On en déduit donc que  $\frac{\partial u}{\partial n} = g$  presque partout sur  $\partial\Omega$ .

**Exercice 29** *Ecrire une formulation variationnelle et analyser le problème suivant :*

$$\begin{cases} -\Delta u = f & \text{sur } \Omega, \\ \frac{\partial u}{\partial n} + \alpha u = g & \text{sur } \partial\Omega, \end{cases} \quad (21)$$

où  $\alpha$  est une fonction bornée, définie sur  $\partial\Omega$ .

**Exercice 30** *On pourrait penser utiliser, pour résoudre le problème de Neumann homogène, la fermeture de  $\{u \in C_c^\infty(\overline{\Omega}), \frac{\partial u}{\partial n} = 0 \text{ sur } \partial\Omega\}$  dans  $H^1(\Omega)$ . Mais cet espace est en fait  $H^1(\Omega)$  ! Se convaincre que ceci est vrai en considérant  $\Omega = (0, 1)$  et la suite de fonctions  $v_n(x) = \frac{1}{n} \mathbf{1}_{0 \leq x \leq 1/n} + x \mathbf{1}_{1/n < x < 1-1/n} + \frac{n-1}{n} \mathbf{1}_{1-1/n < x < 1}$  qui converge dans  $H^1((0, 1))$  vers  $v(x) = x$ .*

**Remarque 16 (Preuve d'existence d'une solution et estimées associées)** *Quand on regarde un problème linéaire, la preuve de l'existence d'une solution au problème implique une estimation. Ceci découle du théorème de l'application ouverte dont on rappelle un corollaire utile. Soit  $E$  et  $F$  deux espaces de Banach, et  $T : E \rightarrow F$  une application linéaire continue, et bijective, i.e. telle que  $T^{-1} : F \rightarrow E$  est bien défini. Alors  $T^{-1}$  est continue.*

*Ainsi, si on considère par exemple le problème du Laplacien de Dirichlet, on prouve que l'application  $T : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ , définie par  $T(u) = -\Delta u$  est une application linéaire continue et bijective, donc  $T^{-1}$  est continue, c'est-à-dire que qu'il existe  $C > 0$  tel que pour tout  $f \in H^{-1}(\Omega)$ , si on considère l'unique  $u \in H_0^1(\Omega)$  tel que  $-\Delta u = f$ , on a*

$$\|u\|_{H^1} \leq C \|f\|_{H^{-1}}.$$

### 2.2.3 Propriétés qualitatives

Nous commençons par le principe du maximum pour l'équation de Laplace. Nous aurons besoin pour cela du lemme suivant

**Lemme 6** *Soit  $u \in H^1(\Omega)$  et  $G$  une fonction de classe  $C^1(\mathbb{R})$  Lipschitzienne. Alors  $G(u) \in H^1(\Omega)$  et*

$$\nabla G(u) = G'(u) \nabla u.$$

De plus si  $u \in H^1(\Omega)$ , alors  $u^+ = \max(u, 0)$  est une fonction de  $H^1(\Omega)$  et

$$\nabla(u^+) = 1_{u>0} \nabla u,$$

où  $1_{u>0}$  désigne la fonction indicatrice de l'ensemble  $\{x \in \Omega, u(x) > 0\}$ .

Preuve : cf. [1, Proposition IX.5] ou [4, Lemmes 7.5 et 7.6].  $\diamond$

**Proposition 16 (Principe du maximum)** Soit  $u \in H^1(\Omega)$  la solution du problème suivant

$$\begin{cases} -\Delta u = f \text{ sur } \Omega, \\ u = g \text{ sur } \partial\Omega, \end{cases}$$

pour  $f \in H^{-1}(\Omega)$  et  $g \in H^{1/2}(\partial\Omega)$ . Si  $f \leq 0$  (au sens suivant :  $\langle f, \phi \rangle \leq 0$  pour toute fonction  $\phi \in H_0^1(\Omega)$  tel que  $\phi \geq 0$ ) et  $g \leq 0$  (au sens suivant :  $u^+ \in H_0^1(\Omega)$ ) alors  $u \leq 0$  (p.p.).

On en déduit que si  $u \in H^1(\Omega)$  est tel que  $-\Delta u \leq 0$  (resp.  $-\Delta u \geq 0$ ) alors  $\forall x \in \Omega, u(x) \leq \sup_{\partial\Omega} u$  (resp.  $u(x) \geq \inf_{\partial\Omega} u$ ).

Ce théorème a une interprétation physique claire, si on pense à  $u$  comme une température par exemple. Il permet d'obtenir des bornes  $L^\infty$  sur la solution.

**Remarque 17** Si  $\Omega$  est suffisamment régulier pour définir l'application trace, on peut montrer que pour une fonction  $u \in H^1(\Omega)$  la trace de  $u^+$  sur le bord  $\partial\Omega$  est égale à la partie positive de la trace de  $u$  sur  $\partial\Omega$ . Autrement dit, dire que  $u^+ \in H_0^1(\Omega)$  revient à dire que  $u \in H^1(\Omega)$  et  $\gamma_0(u) \leq 0$  (avec les notations de la Proposition 8). La condition énoncée dans la Proposition 16 est donc une manière de supposer que  $u$  est positive sur le bord, même si le domaine  $\Omega$  n'est pas assez régulier pour définir la trace de  $u$  sur le bord.

Preuve : En utilisant le lemme précédent, on prend  $v = u^+$  dans la formulation variationnelle (16). On obtient donc

$$\int_{\Omega} \nabla u \cdot \nabla u^+ = \int_{\Omega} f u^+.$$

Le second terme est négatif par hypothèse sur  $f$ . Le premier terme s'écrit

$$\begin{aligned} \int_{\Omega} \nabla u \cdot \nabla u^+ &= \int_{\Omega} 1_{u>0} |\nabla u|^2 \\ &= \int_{\Omega} |\nabla u^+|^2 \geq 0. \end{aligned}$$

On a donc  $\int_{\Omega} |\nabla u^+|^2 = 0$ . On en déduit que  $\nabla u^+ = 0$  (p.p.), et donc, comme  $u^+ \in H_0^1(\Omega)$ ,  $u^+ = 0$  (cf. Exercice 5). Ceci termine la preuve de  $u \leq 0$  (p.p.).  $\diamond$

**Exercice 31** Soit  $u \in H^1(\Omega)$  la solution du problème suivant

$$\begin{cases} -\Delta u + u = f \text{ sur } \Omega, \\ u = g \text{ sur } \partial\Omega, \end{cases}$$

pour  $f \in H^{-1}(\Omega)$  et  $g \in H^{1/2}(\partial\Omega)$ .



1. Montrer que si  $f \leq 0$  (au sens suivant :  $\langle f, \phi \rangle \leq 0$  pour toute fonction  $\phi \in H_0^1(\Omega)$  tel que  $\phi \geq 0$ ) et  $g \leq 0$  (au sens suivant :  $u^+ \in H_0^1(\Omega)$ ) alors  $u \leq 0$  (p.p.).
2. En déduire que (sans hypothèse de signe sur  $f$  et  $g$ )

$$\min \left\{ \inf_{\partial\Omega} g, \inf_{\Omega} f \right\} \leq u \leq \max \left\{ \sup_{\partial\Omega} g, \sup_{\Omega} f \right\}.$$

Solution : La preuve du premier résultat se calque sur la preuve de la Proposition 16. Le deuxième résultat se déduit du premier en considérant  $u - \max \{ \sup_{\partial\Omega} g, \sup_{\Omega} f \}$  et  $\min \{ \inf_{\partial\Omega} g, \inf_{\Omega} f \} - u$ .

Enonçons maintenant un résultat de régularité.

**Proposition 17 (Régularité elliptique)** Soit  $u \in H_0^1(\Omega)$  la solution du problème suivant

$$\begin{cases} -\Delta u = f \text{ sur } \Omega, \\ u = 0 \text{ sur } \partial\Omega, \end{cases}$$

pour  $f \in H^{-1}(\Omega)$ . Pour un entier  $m \geq 0$ , si  $\Omega$  est un domaine régulier (de classe  $C^{m+2}$ ) borné et que  $f \in H^m(\Omega)$ , alors  $u \in H^{m+2}(\Omega)$  et il existe une constante  $C_m$  (indépendante de  $f$ ) telle que

$$\|u\|_{H^{m+2}} \leq C_m \|f\|_{H^m}.$$

Ainsi, si  $m > d/2$ , alors  $u \in C^2(\overline{\Omega})$  et la solution est donc une solution classique.

On obtient exactement le même type de résultat pour le problème de Neumann homogène. Pour une démonstration, on renvoie par exemple à [1].

Ce résultat n'est pas du tout trivial. Il énonce que si  $\Delta u \in L^2$  (donc si une certaine combinaison linéaire des dérivées secondes de  $u$  est dans  $L^2$ ), alors toutes les dérivées secondes de  $u$  sont dans  $L^2$ . Ceci est bien sûr évident en dimension  $d = 1$ , mais la démonstration dans le cas  $d > 1$  est en général difficile.

*Attention* : l'hypothèse de régularité sur  $\Omega$  est importante. Par exemple, en dimension  $d = 2$ , si  $\Omega$  est polygonal, et contient des coins rentrants, alors  $u$  présente des singularités en  $r^{\pi/\theta}$ , où  $\theta > \pi$  est l'angle du coin rentrant, et  $r$  la distance au coin. On obtient ainsi des solutions  $u \in H^{3/2+\varepsilon}(\Omega)$ , pour un  $0 < \varepsilon < 1/2$ , mais pas  $H^2(\Omega)$ . Dans le cas où  $\Omega$  est convexe, on a cependant encore  $u \in H^2(\Omega)$ .

**Remarque 18** Il faut comprendre que les résultats de régularités sont en fait des résultats locaux (cf. [4, Theorem 8.8]). Si  $f \in L^2(\Omega')$  pour  $\Omega' \subset \Omega$ , alors pour tout ouvert  $\omega$  tel que  $\overline{\omega} \subset \Omega'$ , alors  $u \in H^2(\omega)$  et  $\|u\|_{H^2(\omega)} \leq C'(\|u\|_{H^1(\Omega')} + \|f\|_{L^2(\Omega')})$ .

**Exercice 32** Montrer que la norme  $\|u\|_{L^2} + \|\Delta u\|_{L^2}$  est équivalent à la norme  $H^2$  sur  $H^2(\Omega) \cap H_0^1(\Omega)$ .

**Exercice 33** On suppose que  $\Omega = \mathbb{R}^d$ . En utilisant la définition des espaces de Sobolev en terme de transformée de Fourier (cf. Remarque 9), démontrer la Proposition 17.

## 2.2.4 Un exemple de résultat d'homogénéisation

En simulation des matériaux, il est courant de rencontrer des difficultés de modélisation liées à la présence de microstructures dans le matériau. On peut par exemple penser à un solide composé de couches successives de deux matériaux aux propriétés différentes (cf. Figure 2). Pour calculer la réponse d'un tel solide à des sollicitations, une méthode de discrétisation classique est difficile à mettre en oeuvre car il faut utiliser un pas de discrétisation de l'ordre de la taille de la microstructure. Pour contourner ce problème numérique, on peut chercher à approcher ce solide par un solide aux propriétés équivalentes, en calculant les propriétés effectives dans la limite "taille des microstructures tend vers 0" : c'est la technique d'homogénéisation. Pour pouvoir mener à bien ce passage à la limite, il faut utiliser une hypothèse sur l'organisation des microstructures : organisation périodique, organisation aléatoire avec de bonnes propriétés d'ergodicité.

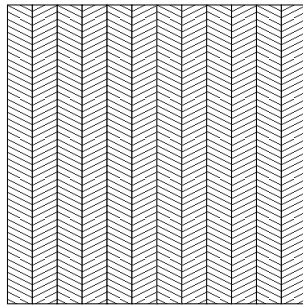


FIG. 2 – Un matériau lamellaire : le matériau est composé d'une succession de couches de deux matériaux différents.

Pour comprendre le type de résultats que l'on peut chercher à démontrer, on propose de considérer le problème modèle suivant en dimension 1 :

$$\begin{cases} -\frac{d}{dx} \left( a \left( \frac{x}{\epsilon} \right) \frac{d}{dx} u^\epsilon \right) = f \text{ sur } (0, 1), \\ u^\epsilon(0) = u^\epsilon(1) = 0, \end{cases} \quad (22)$$

où  $a : \mathbb{R} \rightarrow \mathbb{R}$  est une fonction bornée, périodique de période 1 et telle que :  $\forall x \in \mathbb{R}$ ,

$$0 < c_1 \leq a(x) \leq c_2$$

et  $f$  est une fonction bornée<sup>5</sup>. Ce problème est bien posé et admet une unique solution.

**Exercice 34** Pourquoi le problème (22) est-il bien posé ?

**Exercice 35** A quelles conditions la primitive d'une fonction périodique est périodique ?

On aimerait comprendre comment caractériser la limite de  $u^\epsilon$  quand  $\epsilon \rightarrow 0$ .

---

<sup>5</sup>Les hypothèses sur  $a$  et  $f$  ne sont pas les plus restrictives pour prouver les résultats énoncés dans cette section.

**Lemme 7** La fonction  $a(\cdot/\epsilon)$  converge faiblement dans  $L^2(0,1)$  vers la moyenne de  $a$  :  $\langle a \rangle = \int_0^1 a$ .

*Preuve* : En effet, soit  $\phi \in L^2(0,1)$  une fonction test quelconque, on veut montrer que

$$\lim_{\epsilon \rightarrow 0} \int_0^1 a(x/\epsilon) \phi(x) dx = \langle a \rangle \int_0^1 \phi(x) dx.$$

Par un argument de densité, ceci revient à montrer que pour toute fonction  $\phi \in \mathcal{C}_c^\infty((0,1))$ ,

$$\lim_{\epsilon \rightarrow 0} \int_0^1 (a(x/\epsilon) - \langle a \rangle) \phi(x) dx = 0.$$

Soit  $A(x) = \int_0^x (a(y) - \langle a \rangle) dy$ . On vérifie facilement que  $A$  est une fonction périodique, de période 1. On obtient donc :

$$\begin{aligned} \int_0^1 (a(x/\epsilon) - \langle a \rangle) \phi(x) dx &= \int_0^1 A'(x/\epsilon) \phi(x) dx, \\ &= \epsilon \int_0^1 \frac{d}{dx} (A(x/\epsilon)) \phi(x) dx, \\ &= -\epsilon \int_0^1 A(x/\epsilon) \phi'(x) dx. \end{aligned}$$

En utilisant le fait que  $A$  est une fonction  $L^\infty(\mathbb{R})$  (puisque continue et 1-périodique), on obtient alors le résultat voulu.  $\diamond$

**Exercice 36** En supposant  $\phi \in H_0^k(0,1)$ , pour un entier  $k \geq 1$ , donner un équivalent de  $\left| \int_0^1 a(x/\epsilon) \phi(x) dx - \langle a \rangle \int_0^1 \phi(x) dx \right|$  dans la limite  $\epsilon \rightarrow 0$ .

Comme problème intermédiaire, effaçons les opérateurs de dérivation dans (22), et calculons  $\lim_{\epsilon \rightarrow 0} u^\epsilon$  pour  $u^\epsilon$  qui satisfait

$$-a\left(\frac{x}{\epsilon}\right) u^\epsilon(x) = f(x). \quad (23)$$

On a alors  $u^\epsilon(x) = -\frac{1}{a}\left(\frac{x}{\epsilon}\right) f(x)$ . En utilisant le Lemme 7 pour la fonction  $1/a$ , on a alors, pour toute fonction  $v \in L^2(0,1)$

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \int_0^1 u^\epsilon(x) v(x) dx &= \lim_{\epsilon \rightarrow 0} - \int_0^1 \frac{1}{a}\left(\frac{x}{\epsilon}\right) f(x) v(x) dx \\ &= - \left\langle \frac{1}{a} \right\rangle \int_0^1 f(x) v(x) dx. \end{aligned}$$

On a donc  $u^\epsilon \rightharpoonup u^*$  dans  $L^2((0,1))$  avec  $u^* = -\left\langle \frac{1}{a} \right\rangle f$ . La limite  $u^*$  satisfait l'équation homogénéisée

$$-\frac{1}{\left\langle \frac{1}{a} \right\rangle} u^* = f.$$

Ainsi, le coefficient homogénéisé est  $1/\langle 1/a \rangle$  et non pas  $\langle a \rangle$ .

**Exercice 37** Pour une fonction  $a$  1-périodique minorée par une constante strictement positive, à quelle condition sur  $a$  a-t'on  $\langle \frac{1}{a} \rangle = \frac{1}{\langle a \rangle}$ ? Que peut-on dire si on enlève l'hypothèse de minoration sur  $a$ ?

On va voir que le problème intermédiaire (23) est assez proche du problème (22) que l'on considérait au début de cette section.

**Proposition 18** La solution  $u^\epsilon$  de (22) converge dans  $L^2((0,1))$  vers la solution  $u^* \in H_0^1((0,1))$  du problème homogénéisé

$$\begin{cases} -\frac{d}{dx} \left( \frac{1}{\langle \frac{1}{a} \rangle} \frac{d}{dx} u^* \right) = f \text{ sur } (0,1), \\ u^*(0) = u^*(1) = 0. \end{cases} \quad (24)$$

*Preuve :* On a, pour tout  $\epsilon > 0$

$$\int_0^1 a \left( \frac{x}{\epsilon} \right) \left( \frac{du^\epsilon}{dx} \right)^2 (x) dx = \int_0^1 f u^\epsilon.$$

On en déduit que :  $\exists C > 0, \forall \epsilon > 0,$

$$\|u^\epsilon\|_{H^1((0,1))} \leq C.$$

Ceci permet<sup>6</sup> de construire une suite  $u^\epsilon$  qui converge faiblement dans  $H^1((0,1))$  et fortement dans  $L^2((0,1))$  vers une fonction  $u^* \in H_0^1((0,1))$ .

On a par intégration

$$-a \left( \frac{x}{\epsilon} \right) \frac{du^\epsilon}{dx} = \int_0^x f + c_\epsilon.$$

La suite des réels  $c_\epsilon$  est bornée, et donc, on peut supposer qu'elle converge vers un réel  $c$ . Soit  $v \in L^2((0,1))$ . On a

$$\int_0^1 \frac{du^\epsilon}{dx} v = - \int_0^1 \frac{1}{a} \left( \frac{x}{\epsilon} \right) \left( \int_0^x f + c_\epsilon \right) v(x) dx.$$

On sait que  $\frac{1}{a} \left( \frac{\cdot}{\epsilon} \right)$  converge faiblement dans  $L^2$  vers  $\langle \frac{1}{a} \rangle$ . Or  $\left( \int_0^x f + c_\epsilon \right) v$  converge fortement dans  $L^2$  vers  $\left( \int_0^x f + c \right) v$ . On en déduit que  $\int_0^1 \frac{du^\epsilon}{dx} v$  converge vers  $-\int_0^1 \langle \frac{1}{a} \rangle \left( \int_0^x f + c \right) v(x) dx$ . Autrement dit  $\frac{du^\epsilon}{dx}$  converge faiblement dans  $L^2$  vers  $-\langle \frac{1}{a} \rangle \left( \int_0^x f + c \right)$ . On a donc

$$\frac{du^*}{dx} = - \left\langle \frac{1}{a} \right\rangle \left( \int_0^x f + c \right),$$

d'où  $u^*$  vérifie (24).

Il nous reste à montrer que la suite des  $u^\epsilon$  (et non pas seulement une sous-suite) converge vers la solution  $u^*$  de (24). Ceci provient simplement du fait que la limite  $u^*$  que nous avons obtenue ci-dessus ne dépend pas de la sous-suite choisie, car elle est uniquement déterminée comme solution de (24).  $\diamond$

**Exercice 38** Quel intérêt numérique peut-on trouver à ce type de résultat?

Cette section n'est qu'une très courte introduction à la problématique de l'homogénéisation des équations aux dérivées partielles pour laquelle il existe une littérature abondante. On renvoie par exemple au Chapitre 2 (et aux références données dans ce chapitre) de [5].

---

<sup>6</sup>Dans toute la suite, il faut comprendre les limites  $\epsilon \rightarrow 0$  comme des limites prises sur une suite  $\epsilon_n$  qui tend vers 0, et dont on extrait des sous-suites aussi souvent que nécessaire.

## 2.3 Problèmes paraboliques

On considère le problème de la chaleur : trouver  $u(t, x)$  tel que

$$\begin{cases} \partial_t u - \Delta u = f \text{ sur } (0, T) \times \Omega, \\ u(t, \cdot) = 0 \text{ sur } (0, T) \times \partial\Omega, \\ u(0, \cdot) = u_0 \text{ sur } \Omega. \end{cases} \quad (25)$$

### 2.3.1 Existence et unicité

On utilise la même approche que pour les problèmes elliptiques. On commence donc par chercher une formulation variationnelle de ce problème.

**Etape 1 :** *Formulation variationnelle.*

Nous faisons jouer à la variable de temps  $t$  et la variable d'espace  $x$  des rôles différents. On considère ainsi la fonction  $u(t, x)$  comme une fonction du temps, à valeur dans un espace fonctionnel en  $x$ , et on prend comme fonctions tests des fonctions qui dépendent seulement de la variable  $x$ . On obtient facilement la formulation variationnelle : trouver  $u$  tel que pour toute fonction  $v \in H_0^1(\Omega)$ ,

$$\frac{d}{dt} \int_{\Omega} uv + \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} fv. \quad (26)$$

Pour donner un sens à cette formulation variationnelle, nous demandons la régularité suivante sur  $u$  :

$$u \in \mathcal{C}^0([0, T], L^2(\Omega)) \cap L^2([0, T], H_0^1(\Omega)).$$

L'espace  $\mathcal{C}^0([0, T], L^2(\Omega))$  est l'espace des fonctions  $v(t, x)$  telles que pour tout  $t$ ,  $v(t, \cdot)$  est une fonction de  $L^2(\Omega)$  et l'application qui à  $t$  associe la fonction  $v(t, \cdot)$  est continue de  $[0, T]$  dans  $L^2(\Omega)$ . Il est muni de la norme  $\|v\|_{L^\infty((0, T), L^2(\Omega))} = \sup_{0 \leq t \leq T} \|v\|_{L^2}$  pour lequel c'est un espace de Banach.

L'espace  $L^2([0, T], H_0^1(\Omega))$  est l'espace des fonctions  $v(t, x)$  telles que pour presque tout  $t$ ,  $v(t, \cdot)$  est une fonction de  $H_0^1(\Omega)$  et  $\int_0^T \|v\|_{H_0^1}^2 < \infty$ . Il est muni de la norme

$$\|v\|_{L^2((0, T), H_0^1(\Omega))} = \left( \int_0^T \|v\|_{H_0^1}^2 \right)^{1/2} \text{ pour lequel c'est un espace de Banach.}$$

La formulation variationnelle s'écrit donc :

$$\begin{aligned} &\text{trouver } u \in \mathcal{C}^0([0, T], L^2(\Omega)) \cap L^2([0, T], H_0^1(\Omega)), \text{ tel que } u(0, \cdot) = u_0 \text{ et} \\ &\forall v \in H_0^1(\Omega), \quad \frac{d}{dt} \int_{\Omega} uv + \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} fv. \end{aligned} \quad (27)$$

La dérivée en temps dans (27) est à comprendre au sens des distributions sur  $\mathcal{D}'((0, T))$ .

**Etape 2 :** *Résolution du problème sous forme variationnelle.*

On donne ici le principe d'une méthode de résolution, sans entrer dans tous les détails techniques. Pour les détails de la démonstration, on renvoie par exemple à [6, Chapitre 11].

**Existence :** On utilise *une méthode de Galerkin* pour résoudre le problème.

Soit  $(v_n)_{n \geq 1}$  une base de l'espace de Hilbert séparable<sup>7</sup>  $H_0^1(\Omega)$  et  $V_n = \text{Vect}(v_1, \dots, v_n)$ . Soit  $u^n$  la solution du problème approché

$$\begin{aligned} & \text{trouver } u^n \in \mathcal{C}^1([0, T], V_n), \text{ tel que } u^n(0, \cdot) = u_0 \text{ et} \\ & \forall v \in V_n, \quad \frac{d}{dt} \int_{\Omega} u^n v + \int_{\Omega} \nabla u^n \cdot \nabla v = \int_{\Omega} f v. \end{aligned} \quad (28)$$

On a supposé au passage que  $v_1 = u_0$ , pour que  $u_0 \in V_n$  pour tout  $n \geq 1$  ce qui ne coûte rien si ce n'est qu'il faut que  $u_0 \in H_0^1(\Omega)$ . On peut en fait obtenir l'existence d'une solution sous l'hypothèse plus faible  $u_0 \in L^2(\Omega)$  en approchant  $u_0$  par une suite de fonction  $u_0^\epsilon \in H_0^1(\Omega)$  puis en passant à la limite.

Etant donné que  $V_n$  est un espace vectoriel de dimension finie, le problème (28) est en fait un système d'équations différentielles ordinaires *linéaire*. Il admet donc clairement une solution  $u^n$ .

On cherche maintenant à obtenir des *estimées a priori* sur  $u^n$  de telle sorte à pouvoir passer à la limite  $n \rightarrow \infty$ . Pour cela, on réécrit la formulation variationnelle sous la forme :

$$\int_{\Omega} \frac{\partial u^n}{\partial t} v + \int_{\Omega} \nabla u^n \cdot \nabla v = \int_{\Omega} f v. \quad (29)$$

La fonction  $u^n \in V_n$  peut s'écrire sous la forme  $u^n = \sum_{k=1}^n u_k^n(t) v_k(x)$ . En choisissant  $v = v_k$  dans (29), et en appliquant  $\sum_{k=1}^n u_k^n(t)$ , on a donc :

$$\int_{\Omega} \frac{\partial u^n}{\partial t} u^n + \int_{\Omega} \nabla u^n \cdot \nabla u^n = \int_{\Omega} f u^n.$$

On obtient alors (par Cauchy Schwarz et l'inégalité de Poincaré)

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int_{\Omega} |u^n|^2 + \int_{\Omega} |\nabla u^n|^2 &= \int_{\Omega} f u^n \\ &\leq \|f\|_{L^2(\Omega)} \|u^n\|_{L^2(\Omega)} \\ &\leq \sqrt{C} \|f\|_{L^2(\Omega)} \|\nabla u^n\|_{L^2(\Omega)} \\ &\leq \frac{C}{2} \|f\|_{L^2(\Omega)}^2 + \frac{1}{2} \|\nabla u^n\|_{L^2(\Omega)}^2, \end{aligned}$$

d'où,  $\forall t \in [0, T]$ ,

$$\int_{\Omega} |u^n|^2(t) + \int_0^t \int_{\Omega} |\nabla u^n|^2 \leq \int_{\Omega} |u_0|^2 + C \int_0^t \|f\|_{L^2(\Omega)}^2. \quad (30)$$

On en déduit que la suite  $u^n$  est bornée indépendamment de  $n$  en norme  $L^\infty((0, T), L^2(\Omega))$  et  $L^2((0, T), H_0^1(\Omega))$ . Ceci permet d'extraire une sous-suite de la suite  $u^n$  qui converge faiblement vers une fonction  $u \in L^\infty((0, T), L^2(\Omega)) \cap L^2((0, T), H_0^1(\Omega))$  qui satisfait la même inégalité (30).<sup>8</sup>

<sup>7</sup>Il suffit en fait que  $(v_n)_{n \geq 1}$  soit une suite dense dans  $H_0^1(\Omega)$ .

<sup>8</sup>L'espace  $L^2((0, T), H_0^1(\Omega))$  est un espace de Hilbert pour lequel on a bien défini la convergence faible. Par contre, l'espace  $L^\infty((0, T), L^2(\Omega))$  n'est pas un espace de Hilbert, et nous n'avons donc pas défini la convergence faible sur cet espace. On a en fait besoin ici du résultat suivant : de toute suite  $u^n$  bornée dans  $L^\infty(\Gamma)$  on peut extraire une sous-suite qui converge faiblement vers une fonction  $u \in L^\infty(\Gamma)$  au sens :  $\forall v \in L^1(\Gamma), \int u^n v \rightarrow \int u v$ . On parle de convergence faible \* dans ce cas. On renvoie par exemple à [1].

On peut maintenant, pour  $v = v_{n_0}$  fixé, passer à la limite  $n \rightarrow \infty$  dans (28). En faisant ensuite varier  $n_0$ , on obtient que la fonction  $u \in L^\infty((0, T), L^2(\Omega)) \cap L^2((0, T), H_0^1(\Omega))$  satisfait

$$\forall v \in H_0^1(\Omega), \frac{d}{dt} \int_{\Omega} uv + \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega} fv.$$

Pour prouver que  $u$  est bien une solution de la formulation variationnelle (27), il reste à montrer que  $u \in C^0([0, T], L^2(\Omega))$  et que  $u(0, \cdot) = u_0$ . On admet ces résultats qui découlent de lemmes d'analyse fonctionnelle qui dépassent le cadre de ce cours.

**Unicité :** Pour obtenir l'unicité d'une solution, il faut utiliser à nouveau l'estimée *a priori* ci-dessus. Puisque le problème est linéaire, prouver l'unicité revient à montrer que si  $u_0 = f = 0$ , alors la solution du problème (27) est  $u = 0$ . On réécrit pour cela la formulation variationnelle sous la forme :

$$\forall v \in H_0^1(\Omega), \left( \frac{\partial u}{\partial t}, v \right)_{L^2(\Omega)} + \int_{\Omega} \nabla u \cdot \nabla v = 0.$$

Un résultat d'analyse fonctionnel montre que l'on peut en fait choisir  $v = u$  dans cette formulation (bien que  $u$  soit une fonction dépendant du temps). On obtient alors

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} |u|^2 + \int_{\Omega} |\nabla u|^2 = 0$$

ce qui permet de montrer que  $u = 0$  (puisque  $u \in H_0^1(\Omega)$ ).

**Remarque 19** *On voit le rôle fondamental joué par l'estimation a priori*

$$\frac{d}{dt} \int_{\Omega} |u|^2 + \int_{\Omega} |\nabla u|^2 \leq C \|f\|_{L^2(\Omega)}$$

*dans cette preuve. Les estimées a priori peuvent être obtenues de manière formelle sur le problème continu<sup>9</sup>, et sont ensuite utilisées de manière rigoureuse sur le problème approché, pour démontrer des bornes sur les solutions approchées, et pouvoir passer à la limite. La démarche de la preuve ci-dessus est très générale et très fructueuse pour les problèmes en physique mathématique, car les estimations a priori ont souvent une interprétation en terme d'énergie. Avoir une compréhension du problème physique sous-jacent aide à intuiter les estimations a priori du problème.*

### **Etape 3 : Interprétation des résultats.**

Cette étape se fait comme dans le cas elliptique, et ne présente pas plus de difficultés.

On a donc démontré le résultat suivant :

---

<sup>9</sup>Le qualificatif *a priori* est justifié par cette manière de raisonner : on commence par se poser la question : si j'ai une solution du problème, quelle est *a priori* sa régularité ? Ceci guide le choix des espaces fonctionnels et de la formulation variationnelle.

**Théorème 2** Pour  $u_0 \in L^2(\Omega)$  et  $f \in L^2((0, T) \times \Omega)$ , il existe une unique solution au problème (27). De plus, cette solution satisfait l'estimée d'énergie :  $\forall t \in [0, T]$ ,

$$\int_{\Omega} |u|^2(t) + \int_0^t \int_{\Omega} |\nabla u|^2 \leq \int_{\Omega} |u_0|^2 + C \int_0^t \|f\|_{L^2(\Omega)}^2, \quad (31)$$

où  $C$  est la constante de Poincaré du domaine  $\Omega$ .

**Exercice 39** Pourquoi le résultat d'existence et unicité du théorème précédent est encore valable pour  $f \in L^2((0, T), H^{-1}(\Omega))$  ?

**Exercice 40** On cherche à résoudre le problème suivant :

$$\begin{aligned} &\text{trouver } u \in C^0([0, T], L^2(\Omega)) \cap L^2([0, T], H_0^1(\Omega)), \text{ tel que } u(0, \cdot) = u_0 \text{ et} \\ &\forall v \in H_0^1(\Omega), \frac{d}{dt} \int_{\Omega} uv + a(u, v) = \int_{\Omega} fv, \end{aligned} \quad (32)$$

où  $a$  est une forme bilinéaire continue sur  $H_0^1(\Omega)$  et qui vérifie l'inégalité de Garding :  $\forall u \in H_0^1(\Omega)$

$$a(u, u) \geq \alpha \|\nabla u\|_{L^2}^2 - \lambda \|u\|_{L^2}^2,$$

où  $\alpha > 0$  et  $\lambda \in \mathbb{R}$ . Expliquer pourquoi ce problème est bien posé en dérivant une estimée a priori.

**Exercice 41** Utiliser la méthode de Galerkin pour prouver que le problème de Laplace (14) est bien posé.

De même que dans le cas elliptique, on peut traiter le cas de l'équation de la chaleur avec des conditions de Dirichlet non-homogène et des conditions de Neumann.

### 2.3.2 Propriétés qualitatives

On a un résultat similaire au cas elliptique :

**Proposition 19** Soit  $u \in H^1((0, T), L^2(\Omega)) \cap L^2((0, T), H^1(\Omega))$  la solution du problème suivant

$$\begin{cases} \partial_t u - \Delta u = f \text{ sur } \Omega, \\ u(t, \cdot) = g \text{ sur } \partial\Omega, \\ u(0, \cdot) = u_0 \text{ sur } \Omega, \end{cases}$$

pour  $f \in L^2((0, T) \times \Omega)$ ,  $u_0 \in L^2(\Omega)$  et  $g \in H^{1/2}(\Omega)$ . Si  $f \leq 0$  (p.p.),  $u_0 \leq 0$  (p.p.) et  $g \leq 0$  (au sens  $u^+ \in H_0^1(\Omega)$  pour presque tout  $t$ ) alors  $u \leq 0$  (p.p.).

*Preuve* : On prend  $v = u^+$  dans la formulation variationnelle (27) et on obtient donc (on admet que l'on peut prendre  $v = u^+$  comme fonction test, même si  $u^+$ , qui est bien une fonction de  $H_0^1(\Omega)$  pour presque tout temps, dépend du temps)

$$\int_{\Omega} \frac{\partial u}{\partial t} u^+ + \int_{\Omega} \nabla u \cdot \nabla u^+ = \int_{\Omega} f u^+.$$



On en déduit (mêmes arguments que pour le principe du maximum dans le cas de l'équation de Poisson) :

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} |u^+|^2 + \int_{\Omega} |\nabla u^+|^2 \leq 0,$$

et donc, comme  $\int_{\Omega} |u_0^+|^2 = 0$ , pour tout  $t \geq 0$ ,

$$\frac{1}{2} \int_{\Omega} |u^+|^2 + \int_0^t \int_{\Omega} |\nabla u^+|^2 \leq 0.$$

Ceci permet de conclure que  $u^+ = 0$  et donc  $u \leq 0$  presque partout.  $\diamond$

Ainsi, si  $u$  et  $v$  désignent deux solutions de l'équation de la chaleur, avec même second membre  $f$  et mêmes conditions aux limites  $g$ , et si les conditions initiales satisfont  $u_0 \leq v_0$  alors  $u \leq v$ .

Voici un résultat de régularité pour l'équation de la chaleur :

**Proposition 20** *Soit  $u$  la solution du problème (27). On suppose  $f \in L^2((0, T) \times \Omega)$  et  $u_0 \in H_0^1(\Omega)$ . Alors la solution  $u \in L^\infty((0, T), H_0^1(\Omega)) \cap H^1((0, T), L^2(\Omega))$  et satisfait l'estimée d'énergie :  $\forall t \in [0, T]$ ,*

$$\int_{\Omega} |\nabla u|^2(t) + \int_0^t \int_{\Omega} |\partial_t u|^2 \leq \int_{\Omega} |\nabla u_0|^2 + \int_0^t \|f\|_{L^2(\Omega)}^2.$$

On en déduit que  $u \in L^2((0, T), H^2(\Omega))$ .

Ce résultat dérive de l'estimée d'énergie obtenue en prenant  $\frac{\partial u}{\partial t}$  comme fonction test. On écrit ensuite  $\Delta u = -f + \partial_t u$ , et on utilise le résultat de régularité elliptique.

**Corollaire 1** *Si  $u_0 \in L^2(\Omega)$  et  $f = 0$ , alors la solution  $u$  du problème (27) est de classe  $C^\infty$  sur  $[\varepsilon, T] \times \overline{\Omega}$  pour tout  $\varepsilon > 0$ .*

*Preuve :* D'après le Théorème 2, on sait que  $u \in C^0((0, T), L^2(\Omega)) \cap L^2((0, T), H_0^1(\Omega))$ . On en déduit en particulier que pour  $\varepsilon > 0$  fixé, il existe un  $t_0 \in (0, \varepsilon)$  tel que  $u(t_0, \cdot) \in H_0^1(\Omega)$ . En utilisant alors le résultat de régularité de la Proposition 20 sur l'intervalle de temps  $(t_0, T)$ , on en déduit que  $u \in L^2((t_0, T), H^2(\Omega))$ .

On peut alors considérer l'équation satisfaite par  $\Delta u$  pour itérer l'argument et montrer que  $u \in C^0((\varepsilon, T), H^{k-1}(\Omega)) \cap L^2((\varepsilon, T), H^k(\Omega))$  pour tout  $k$ , et, en utilisant les injections de Sobolev et Proposition 20, que  $u \in H^1((\varepsilon, T), C^\infty(\overline{\Omega}))$ . L'équation  $\partial_t u = \Delta u$  permet ensuite de conclure.  $\diamond$

## 3 Discrétisation des problèmes elliptiques et paraboliques

### 3.1 Méthode des différences finies

#### 3.1.1 Présentation de la méthode

Le principe de la méthode des différences finies est d'approcher les dérivées dans les opérateurs par des développements limités par rapport à un paramètre  $h > 0$

destiné à tendre vers 0. Ainsi :

$$\frac{du}{dx}(x) \approx \begin{cases} D_h u(x) = \frac{u(x+h)-u(x-h)}{2h}, \\ D_h^+ u(x) = \frac{u(x+h)-u(x)}{h}, \\ D_h^- u(x) = \frac{u(x)-u(x-h)}{h}, \end{cases}$$

ou encore

$$\frac{d^2u}{dx^2}(x) \approx D_h^2 u(x) = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}.$$

L'approximation  $D_h$  (resp.  $D_h^+$  et  $D_h^-$ ) de  $\frac{d}{dx}$  est qualifiée de centrée (resp. décentrée à droite, et décentrée à gauche).

On dit que l'approximation  $L_h$  d'un opérateur  $L$  est consistante et d'ordre  $r$  si, pour toute fonction  $u$  suffisamment régulière et tout point  $x$ ,  $\exists C(u, x)$  tel que, pour tout  $h > 0$ ,

$$|Lu(x) - L_h u(x)| \leq C(u, x)h^r.$$

**Exercice 42** Donner l'ordre de consistance des quatre approximations ci-dessus. Comment se comporte la constante  $C(u, x)$  en fonction de  $u$  ?

Nous prenons dans la suite l'exemple de l'équation d'advection-diffusion en dimension 1 :  $\Omega = [0, L]$ ,  $t \in [0, T]$  et

$$\begin{cases} \partial_t u + a \partial_x u - D \partial_{x,x} u = f \text{ sur } [0, L], \\ u(\cdot, 0) = g \text{ et } u(\cdot, L) = d, \\ u(0, \cdot) = u_0 \text{ sur } [0, L], \end{cases} \quad (33)$$

où les coefficients  $a$  et  $D > 0$  sont supposés constants pour simplifier la présentation. Les fonctions  $g$  et  $d$  dépendent éventuellement du temps et sont les conditions aux limites à gauche et à droite de l'intervalle  $[0, L]$ .

**Exercice 43** Rappeler sous quelles conditions sur  $a$  et  $D$  on obtient une solution à ce problème en utilisant le cadre de la Section 2.3. Montrer que l'on peut construire une solution qui satisfait l'estimée d'énergie (comparer avec (31)) :

$$\int_{\Omega} |u|^2(t) + D \int_0^t \int_{\Omega} |\partial_x u|^2(t) \leq \int_{\Omega} |u_0|^2 + C \int_0^t \|f\|_{L^2(\Omega)}^2.$$

On introduit une grille de discrétisation en temps et en espace :  $t_n = n\delta t$  avec  $0 \leq n \leq N$  et  $N = T/\delta t$ ;  $x_j = jh$  pour  $0 \leq j \leq I$  et  $I = L/h$ .

Pour la discrétisation des opérateurs en espace, on remplace  $\partial_{x,x}$  par  $D_h^2$  et  $\partial_x$  par  $D_h$ . On obtient alors une équation sur le vecteur  $U_j(t)$  ( $U_j(t)$  est censé être une approximation de  $u(t, x_j)$ , où  $u$  satisfait (33)) :

$$\begin{cases} \partial_t U_i = \sum_{j=1}^{I-1} A_{i,j} U_j + C_i + F_i \text{ pour } 1 \leq i \leq I-1, \\ U_0 = g \text{ et } U_I = d, \\ U(0) = \pi_h(u_0), \end{cases} \quad (34)$$



### 3.1.2 Convergence de la méthode

La preuve pour la convergence de la méthode est basée sur le principe général : consistance et stabilité impliquent convergence. On illustre ce principe sur le cas particulier considéré ci-dessus, mais le même schéma de preuve s'applique pour d'autres équations et schémas numériques.

La consistance s'obtient en remplaçant  $U_j^n$  dans les schémas par la vraie valeur  $u(t_n, x_j)$  et en regardant l'erreur commise à chaque pas de temps. On choisit une norme  $L^\infty$  sur les vecteurs en espace et en temps. On obtient alors facilement

- pour le schéma d'Euler Explicite,

$$\|L_{\delta t, h}^{EE}((u(t_n, x_j))_{j,n})\|_{L^\infty} \leq C(u) (\delta t + h^2),$$

- pour le schéma d'Euler Implicite,

$$\|L_{\delta t, h}^{EI}((u(t_n, x_j))_{j,n})\|_{L^\infty} \leq C(u) (\delta t + h^2),$$

- pour le schéma de Crank-Nicolson,

$$\|L_{\delta t, h}^{CN}((u(t_n, x_j))_{j,n})\|_{L^\infty} \leq C(u) (\delta t^2 + h^2).$$

La notion de stabilité est liée (comme dans le cas des équations différentielles ordinaires) à un calcul sur la manière dont les erreurs faites à chaque pas de temps vont s'accumuler. Un schéma est stable si de petites perturbations sur les données engendrent de petites perturbations sur le résultat.

**Définition 7** *On dit que le schéma numérique (pour l'Equation (33)) est stable pour une norme  $\|\cdot\|$  si pour un temps  $T > 0$  fixé, il existe une constante  $C(T)$  indépendante de  $\delta t$  et de  $h$  telle que (pour  $N = T/\delta t$ )*

$$\sup_{0 \leq n \leq N} \|U^n\| \leq C(T) \left( \sup_{0 \leq n \leq N} \|F^n\| + \sup_{0 \leq n \leq N} \|C^n\| + \|U^0\| \right).$$

Pour la norme (en espace) dans laquelle on analyse la stabilité des schémas, on considèrera dans la suite typiquement deux normes :

- la norme  $L^2$  :

$$\|U\|_2 = \left( \frac{1}{(I-1)} \sum_{j=1}^{I-1} (U_j)^2 \right)^{1/2}$$

- la norme  $L^\infty$  :

$$\|U\|_\infty = \max_{1 \leq j \leq I-1} |U_j|.$$

Remarquer que  $\|U\|_2 \leq \|U\|_\infty$ .

On a alors

**Proposition 21** *Un schéma stable et consistant est convergent. La convergence a lieu à la vitesse donnée par l'ordre de consistance, et dans la norme de stabilité.*

*Preuve* : Ici, tout est linéaire, donc les calculs sont très simples. Quand on applique le schéma aux  $u(t_n, x_j)$ , on obtient l'erreur de consistance  $\epsilon_j^n$  dont on sait par hypothèse qu'elle est majorée en norme  $L^\infty$  par  $C(u)(\delta t^{k_t} + h^{k_x})$  où  $k_t$  (resp.  $k_x$ ) désigne l'ordre de convergence en temps (resp. en espace) du schéma. Introduisons ensuite le vecteur erreur  $E_j^n = U_j^n - u(t_n, x_j)$ . Quand on applique le schéma à  $E$ , on obtient

$$L_{\delta t, h} E = -\epsilon^n.$$

Pour ce vecteur, on a par ailleurs  $C^n = 0$  et  $E^0 = 0$ . Par stabilité du schéma, on a donc

$$\sup_{0 \leq n \leq N} \|E^n\| \leq C \sup_{0 \leq n \leq N} \|\epsilon^n\|.$$

Or  $\|\epsilon^n\| \leq \|\epsilon^n\|_\infty \leq C(u)(\delta t^{k_t} + h^{k_x})$ , d'où le résultat annoncé.  $\diamond$

### 3.1.3 Analyse de stabilité

Il nous reste maintenant à analyser la stabilité des schémas que nous avons introduits. Nous allons pour cela introduire des conditions *suffisantes* de stabilité, qui s'avèrent en pratique être aussi nécessaires.

Pour le schéma d'Euler Explicite, on a

$$U^{n+1} = (\text{Id} + \delta t A)U^n + \delta t C^n + \delta t F^n,$$

et donc, par récurrence,  $\forall 0 \leq n \leq N$ ,

$$\begin{aligned} \|U^n\| &\leq \|\text{Id} + \delta t A\|^n \|U^0\| + \delta t \sum_{k=0}^{n-1} \|\text{Id} + \delta t A\|^{n-1-k} (\|C^k\| + \|F^k\|), \\ &\leq \|\text{Id} + \delta t A\|^n \|U^0\| + \delta t \sum_{k=0}^{n-1} \|\text{Id} + \delta t A\|^k \left( \sup_{0 \leq n \leq N} \|F^n\| + \sup_{0 \leq n \leq N} \|C^n\| \right), \end{aligned}$$

où  $\|\text{Id} + \delta t A\|$  désigne la norme matricielle de  $\text{Id} + \delta t A$  associée à la norme vectorielle  $\|\cdot\|$  utilisée pour analyser la stabilité. On voit que pour obtenir la stabilité du schéma numérique, il suffit d'avoir la propriété : pour tout  $\delta t > 0$  et  $h > 0$ ,

$$\|\text{Id} + \delta t A\| \leq 1. \quad (38)$$

**Remarque 20** Cette condition de stabilité est en fait une condition de stabilité absolue. Il suffirait pour avoir la stabilité du schéma numérique au sens de la Définition 7 de vérifier que  $\exists C > 0$  indépendant de  $\delta t$  et  $h$  tel que

$$\|\text{Id} + \delta t A\| \leq 1 + C\delta t. \quad (39)$$

Par des raisonnements équivalents, on voit que pour le schéma d'Euler implicite, la stabilité est vérifiée si

$$\|(\text{Id} - \delta t A)^{-1}\| \leq 1, \quad (40)$$

et que pour le schéma de Crank-Nicolson, la stabilité est vérifiée si

$$\left\| \left( \text{Id} - \frac{\delta t}{2} A \right)^{-1} \left( \text{Id} + \frac{\delta t}{2} A \right) \right\| \leq 1. \quad (41)$$

La question est maintenant : comment vérifier les inégalités (38), (40) ou (41) ?

**Exercice 44** Soit  $A$  une matrice de taille  $I \times I$ .

- Montrer que la norme matricielle associée à la norme euclidienne (norme  $L^2$ ) est  $\|A\|_2 = \sqrt{\rho(A^T A)}$ , où le rayon spectral  $\rho(A^T A)$  de  $A^T A$  est par définition la plus grande des (valeurs absolues) des valeurs propres de  $A^T A$ .
- Montrer que la norme matricielle associée à la norme  $L^\infty$  est  $\|A\|_\infty = \max_{1 \leq i \leq I} \sum_{j=1}^I |A_{i,j}|$ .

**Le cas de la norme  $L^2$**  Pour analyser la stabilité en norme  $L^2$  d'une matrice  $B$ , on peut suivre trois méthodes plus ou moins équivalentes :

- Analyser les valeurs propres de la matrice  $B^T B$ , et vérifier que ces valeurs propres sont plus petites que 1 (puisque la norme de  $B$  dans ce cas est la racine carré de la plus grande des valeurs propres de  $B^T B$ ),
- Faire une analyse de type "énergétique" en multipliant scalairement les équations par  $U^n$  et en cherchant à montrer que  $\|U^{n+1}\|_2$  est plus petit que  $\|U^n\|_2$  (penser aux estimées *a priori* que nous avons utilisées au niveau continu sur l'équation de la chaleur ci-dessus : il s'agit ici de dériver des estimées du même type, mais au niveau discret),
- Utiliser une analyse dite de Von Neumann en passant en Fourier.

Détaillons cette dernière méthode, très utile en pratique (mais qui s'applique seulement si  $a$  et  $D$  sont constants). Considérons par exemple le cas du schéma d'Euler explicite. On veut montrer que pour  $U^{n+1} = (\text{Id} + \delta t A)U^n$ ,

$$\|U^{n+1}\|_2 \leq \|U^n\|_2.$$

Pour cela, on interpole  $U^n$  sur une fonction périodique sur  $[0, L]$  (disons  $I_h(U^n)$ ) de telle sorte que

$$\int_0^L |I_h(U^n)|^2 = \|U^n\|_2^2.$$

Le schéma se réécrit sur  $v^n := I_h(U^n)$  comme

$$v^{n+1}(x) = (\text{Id} + \delta t A)v^n(x)$$

soit

$$\frac{1}{\delta t} (v^{n+1}(x) - v^n(x)) = a \frac{v^n(x+h) - v^n(x-h)}{2h} + D \frac{v^n(x+h) - 2v^n(x) + v^n(x-h)}{h^2}.$$

En passant en Fourier, on a

$$\frac{1}{\delta t} (\hat{v}^{n+1}(\xi) - \hat{v}^n(\xi)) = a \frac{\exp(ih\xi) - \exp(-ih\xi)}{2h} \hat{v}^n(\xi) + D \frac{\exp(ih\xi) - 2 + \exp(-ih\xi)}{h^2} \hat{v}^n(\xi).$$

On a donc

$$\hat{v}^{n+1}(\xi) = z(\xi) \hat{v}^n(\xi)$$

avec le facteur d'amplification

$$\begin{aligned} z(\xi) &= 1 + a\delta t \frac{\exp(ih\xi) - \exp(-ih\xi)}{2h} + \delta t D \frac{\exp(ih\xi) - 2 + \exp(-ih\xi)}{h^2}, \\ &= 1 + ia\delta t \frac{\sin(h\xi)}{h} - 4D\delta t \frac{\sin^2(h\xi/2)}{h^2}. \end{aligned}$$

On utilisant la relation de Plancherel  $\int |v^n|^2 = \int |\hat{v}^n|^2$ , on voit qu'une condition suffisante de stabilité est que

$$\forall \xi \in \mathbb{R}, |z(\xi)| \leq 1.$$

On obtient dans notre cas que cette condition est satisfaite si et seulement si :

- Si  $|a| \geq 2D/h$ , il faut  $\delta t \leq 2D/|a|^2$  (ce qui implique  $\delta t \leq h^2/(2D)$ ),
- Si  $|a| \leq 2D/h$ , il faut  $\delta t \leq h^2/(2D)$ .

On voit donc que dans tous les cas, le schéma d'Euler explicite n'est stable en norme  $L^2$  que sous la condition dite condition CFL (du nom de ses découvreurs Courant, Friedrichs et Lewy)

$$\delta t \leq h^2/(2D).$$

De plus, si l'advection domine la diffusion ( $|a| \geq 2D/h$ ), il y a une condition supplémentaire.

**Exercice 45** Vérifier les affirmations précédentes en calculant  $|z(\xi)|^2$ . (On pourra utiliser la relation  $\sin^2(h\xi) = 4 \sin^2(h\xi/2)(1 - \sin^2(h\xi/2))$  pour se ramener à l'étude du signe d'un trinôme en  $X = \sin^2(h\xi/2)$ .)

Par un raisonnement analogue, on vérifie que le schéma d'Euler implicite et le schéma de Crank Nicolson sont inconditionnellement stables en norme  $L^2$ .

**Exercice 46** Analyser la stabilité  $L^2$  des schémas d'Euler implicite et de Crank Nicolson.

**Le cas de la norme  $L^\infty$**  La stabilité en norme  $L^\infty$  est reliée au principe du maximum discret. Précisons cette notion. L'équation (33) au niveau continu vérifie un principe du maximum : en effet, si les données ( $u_0$ ,  $g$ ,  $d$  et  $f$ ) sont positives, alors  $u$  est positive. De même,

**Définition 8** On dit qu'un schéma numérique (associé à (33)) vérifie un principe du maximum discret si et seulement si, si les données ( $U_0$ ,  $C^n$ ,  $F^n$ ) sont positives, alors la solution  $U^n$  est positive.

On a ensuite

**Proposition 22** Un schéma numérique qui vérifie le principe du maximum discret est stable en norme  $L^\infty$ .

**Remarque 21** Dans le cas de l'équation de la chaleur, on sait que la solution du problème  $\partial_t u - \Delta u$  est reliée au mouvement brownien. Par exemple, si  $u(t, x)$  désigne la densité de probabilité de  $W_t$  à l'instant  $t$ ,  $u$  vérifie l'équation de la chaleur. Il est donc naturel de chercher un schéma numérique sur  $u$  qui vérifie le principe du maximum car  $u$  est une densité de probabilité donc doit rester positif.

On propose de démontrer la Proposition 22 dans le cas du schéma d'Euler explicite. La preuve est similaire pour les autres schémas.

On commence par vérifier sous quelles conditions (nécessaires et suffisantes) le schéma d'Euler explicite vérifie le principe du maximum. Il suffit de vérifier sous

quelles conditions les coefficients de la matrice  $(\text{Id} + \delta t A)$  sont positifs. On obtient les deux conditions :

$$|a|h/(2D) \leq 1 \text{ et } \delta t \leq h^2/(2D). \quad (42)$$

La première condition dit que l'advection ne doit pas dominer sur la diffusion (le nombre sans dimension  $|a|h/(2D)$  s'appelle le nombre de Péclet). La deuxième condition est simplement la condition CFL.

On suppose que le schéma d'Euler explicite vérifie le principe du maximum discret et donc que les conditions (42) sont vérifiées, et on veut maintenant vérifier que le schéma est stable en norme  $L^\infty$ . Soit  $U$  et  $V$  deux vecteurs qui satisfont  $U = (\text{Id} + \delta t A)V$ . On veut montrer que  $\|U\|_\infty \leq \|V\|_\infty$ . On introduit  $W = U - \|V\|_\infty$ . On vérifie que :

$$W = (\text{Id} + \delta t A)(V - \|V\|_\infty) + E$$

où

$$E = -\delta t \|V\|_\infty \begin{pmatrix} (-\frac{a}{2h} + \frac{D}{h^2}) \\ 0 \\ \vdots \\ 0 \\ (\frac{a}{2h} + \frac{D}{h^2}) \end{pmatrix}.$$

Or  $V - \|V\|_\infty \leq 0$  et  $E \leq 0$ , et donc  $W \leq 0$ . En considérant  $U + \|V\|_\infty$ , et en raisonnant de la même façon, on montre que  $\|U\|_\infty \leq \|V\|_\infty$ , ce qui termine la preuve de la Proposition 22 dans le cas du schéma d'Euler explicite.

Dans le cas du schéma d'Euler implicite, il faut déterminer à quelles conditions la matrice  $(\text{Id} - \delta t A)^{-1}$  a des coefficients positifs. On utilise pour cela le

**Lemme 8**  $B \in \mathbb{R}^{J \times J}$  est une M-matrice si et seulement si,  $\forall 1 \leq i, j \leq J$

$$\begin{cases} B_{i,i} > 0, \\ B_{i,j} \leq 0 \text{ si } i \neq j, \\ \sum_{j=1}^J B_{i,j} > 0. \end{cases}$$

Dans ce cas,  $B$  est inversible et  $B^{-1}$  est une matrice à coefficients positifs.

**Exercice 47** Démontrer ce lemme.

En utilisant ce lemme, et le fait que

$$\text{Id} - \delta t A = \begin{bmatrix} & & \ddots & \ddots & \ddots & & \\ & & & & & & \\ 0 \dots 0 & \delta t \left( \frac{a}{2h} - \frac{D}{h^2} \right) & (1 + \delta t \frac{2D}{h^2}) & \delta t \left( -\frac{a}{2h} - \frac{D}{h^2} \right) & 0 \dots 0 & & \\ & & \ddots & \ddots & \ddots & & \end{bmatrix},$$

on montre que le schéma d'Euler implicite est stable en norme  $L^\infty$  sous la condition

$$|a|h/(2D) \leq 1.$$

On perd donc la stabilité en norme  $L^\infty$  si l'advection domine la diffusion.

**Exercice 48** Faire la preuve de la Proposition 22 pour les schémas d'Euler implicite et de Crank Nicolson.



**Exercice 49** On considère la matrice  $A$  de taille  $I \times I$ , telle que pour tout  $i \in \{1, \dots, I\}$ ,  $A_{i,i} = 2$ , pour tout  $i \in \{2, \dots, I\}$ ,  $A_{i,i-1} = -1$  et pour tout  $i \in \{1, \dots, I-1\}$ ,  $A_{i,i+1} = -1$ .

- $A$  quel opérateur est relié la matrice  $A$  ? Solution :  $(I+1)^2 A$  est obtenue par discrétisation par différences finies de l'opposé du laplacien de Dirichlet sur  $[0, 1]$ , avec  $(I+1)$  intervalles.
- Vérifier que  $A$  n'est pas une  $M$ -matrice. Montrer que pour tout  $x \in \mathbb{R}^I$ ,  $x^T A x = \sum_{i=1}^{I-1} (x_{i+1} - x_i)^2$ , où l'on a posé  $x_{-1} = x_{I+1} = 0$ . On déduire que  $A$  est inversible.
- On introduit  $A_\varepsilon = A + \varepsilon Id$ . Montrer que  $A_\varepsilon$  est une  $M$ -matrice, pour tout  $\varepsilon > 0$ . En déduire que les coefficients de  $A^{-1}$  sont positifs. Solution : Utiliser le fait que l'application  $A \mapsto A^{-1}$  est continue sur l'ensemble des matrices inversibles. Comme les coefficients de  $A_\varepsilon^{-1}$  sont positifs, les coefficients de  $A_0^{-1} = A^{-1}$  le sont aussi par continuité.
- Soit les deux vecteurs de  $\mathbb{R}^I$  :  $e = (1, \dots, 1)^T$  et  $x = (f(\frac{1}{I+1}), \dots, f(\frac{i}{I+1}), \dots, f(\frac{I}{I+1}))^T$ , où  $f(t) = t(1-t)/2$ . Vérifier que  $Ax = \frac{e}{(I+1)^2}$ . En déduire que  $\|A^{-1}\|_\infty = \max_{1 \leq i \leq I} \sum_{j=1}^I |A_{i,j}^{-1}| \leq \frac{(I+1)^2}{8}$ . Solution : On a  $\|A^{-1}\|_\infty = \max_{1 \leq i \leq I} \sum_{j=1}^I A_{i,j}^{-1} = \|A^{-1}e\|_\infty = (I+1)^2 \|x\|_\infty \leq (I+1)^2 \|f\|_{L^\infty(0,1)}$ .

**Décentrage** Il existe une méthode pour que le schéma d'Euler implicite soit inconditionnellement stable en norme  $L^\infty$  et traiter ainsi cette difficulté dans le cas où l'advection domine la diffusion. L'idée est de modifier la discrétisation du terme d'advection  $a \partial_x u$  en remontant les caractéristiques (méthode de décentrage ou *up-winding* en anglais) : si  $a > 0$ , on approxime  $\partial_x u$  par  $D_h^+ u$  (cf. Figure 3), et si  $a < 0$ , on approxime  $\partial_x u$  par  $D_h^- u$ .

Ainsi, dans le cas  $a > 0$ , on obtient alors pour le schéma d'Euler implicite

$$\text{Id} - \delta t A = \begin{bmatrix} & & & \dots & \dots & \dots \\ 0 \dots 0 & -\delta t \frac{D}{h^2} & (1 + \delta t \frac{2D}{h^2} + \delta t \frac{a}{h}) & \delta t (-\frac{a}{h} - \frac{D}{h^2}) & 0 \dots 0 \\ & & \dots & \dots & \dots \end{bmatrix},$$

et en utilisant le lemme 8, on obtient bien que le schéma d'Euler implicite vérifie le principe du maximum et est donc inconditionnellement stable en norme  $L^\infty$ .

**Remarque 22** On peut remarquer que l'opération de décentrage correspond à rajouter de la diffusion dans le schéma numérique, puisque (pour  $a > 0$ )

$$\begin{aligned} a D_h^+ u(x) &= a \frac{u(x+h) - u(x)}{h}, \\ &= a \frac{u(x+h) - u(x-h)}{2h} + \frac{ah}{2} \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}, \\ &= a D_h u(x) + \frac{ah}{2} D_h^2 u(x). \end{aligned}$$

Tout se passe comme si on avait discrétisé l'équation initiale avec une différence finie centrée pour le terme d'advection, et une diffusion  $(D + \frac{ah}{2})$ .

Pour résumer, on a donc montré que

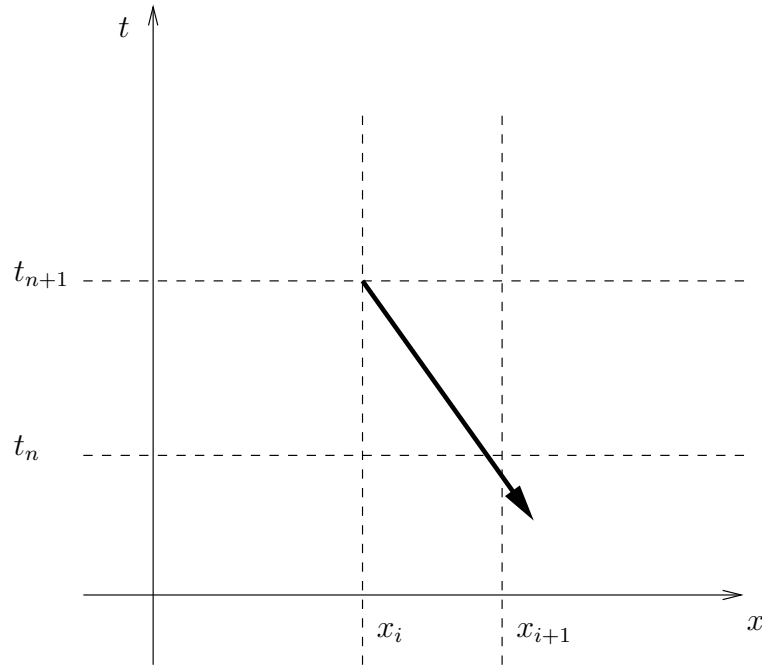


FIG. 3 – Représentation de la caractéristique (associée à l'équation  $\partial_t u = a \partial_x u$ ) issue de  $(t_{n+1}, x_i)$  dans le cas où  $a > 0$ . La flèche indique le sens des temps décroissants. Dans ce cas, on voit que décentrer le terme d'advection en “remontant le courant” (*upwinding*) consiste à approximer  $\partial_x u$  par  $D_h^+ u$ .

- le schéma d'Euler explicite n'est stable que sous une condition CFL du type  $\delta t \leq Ch^2$ ,
- quand l'advection domine la diffusion, les schémas implicites sont stables en norme  $L^\infty$  à condition de décentrer la discrétisation des termes d'advection.

Ces conclusions sont très générales et valent pour tous les schémas de discrétisation d'équations paraboliques. On observe donc que si on veut s'affranchir de la condition CFL, il faut utiliser un schéma implicite, qui sera plus coûteux car il faudra inverser un système linéaire. Si on veut obtenir une stabilité  $L^\infty$  dans le cas où l'advection domine la diffusion, il faut décentrer la discrétisation des termes d'advection, ce qui fait perdre un ordre de convergence en espace.

### 3.1.4 La résolution de systèmes linéaires

Nous avons vu que les méthodes de différences finies conduisent à résoudre des systèmes linéaires du type  $AX = B$ . Ces problèmes interviennent en fait quelque soit la méthode de discrétisation.

Un paramètre important intervenant dans la résolution d'un système linéaire est le conditionnement de la matrice. Il est relié à une notion de stabilité : pour une perturbation du second membre  $B$ , quelle est la perturbation engendrée sur  $X$  ?

**Proposition 23** *Soit  $A$  une matrice inversible,  $B$  un vecteur,  $X$  tel que  $AX = B$  et  $X + \delta X$  tel que  $A(X + \delta X) = B + \delta B$  pour une perturbation  $\delta B$  donnée. Alors,*

pour une norme vectorielle  $\|\cdot\|$  donnée, et sa norme matricielle associée, on a :

$$\frac{\|\delta X\|}{\|X\|} \leq \kappa(A) \frac{\|\delta B\|}{\|B\|},$$

où

$$\kappa(A) = \|A\| \|A^{-1}\|$$

est le conditionnement de la matrice  $A$ .

*Preuve :* Il suffit d'écrire  $\|B\| \leq \|A\| \|X\|$  et (comme  $\delta X = A^{-1} \delta B$ ),  $\|\delta X\| \leq \|A^{-1}\| \|\delta B\|$ .  $\diamond$

Si on choisit comme norme matricielle la norme  $L^2$ , le conditionnement de la matrice est

$$\kappa_2(A) = \sqrt{\frac{\max_{1 \leq i \leq I} |\lambda_i(A^T A)|}{\min_{1 \leq i \leq I} |\lambda_i(A^T A)|}},$$

où  $(\lambda_i(A^T A))_{1 \leq i \leq I}$  désigne les valeurs propres de  $A^T A$ . Si le conditionnement de la matrice est grand, il sera difficile d'obtenir une solution, en particulier avec un algorithme itératif.

On peut distinguer deux grandes classes de méthode pour résoudre un problème linéaire : les méthodes directes et les méthodes itératives.

Les méthodes directes consistent à manipuler la matrice  $A$  de telle sorte à mettre le problème sous une forme plus simple. L'exemple le plus simple est la méthode de Gauss, qui consiste à écrire  $A = LU$  avec  $L$  triangulaire inférieure et  $U$  triangulaire supérieure. L'avantage des méthodes directes est qu'elles sont stables numériquement (moyennant quelques précautions) même pour des matrices assez mal conditionnées, et qu'une fois la décomposition  $LU$  calculée, on peut résoudre le problème  $AX = B$  très rapidement pour autant de seconds membres  $B$  que l'on veut. L'inconvénient est qu'elles ne préservent pas en général la structure de la matrice  $A$  : si  $A$  a beaucoup de 0 (on parle de matrices creuses, et de telles matrices apparaissent très souvent en calcul scientifique, notamment avec les méthodes de différences finies ou d'éléments finis, cf. Exercice 50 ci-dessous), ce ne sera pas en général le cas pour  $L$  et  $U$ , et il faudra donc stocker beaucoup d'information en mémoire. De plus, ces méthodes sont en général plus lentes que les méthodes itératives.

Les méthodes itératives consistent à calculer la solution de  $AX = B$  en construisant une suite d'approximations de  $X$ . Elles ne nécessitent, pour calculer  $X_{n+1}$  en fonction de  $X_n$ , que des multiplications matrice-vecteur. Ces méthodes sont soit basées sur des décompositions de la matrice  $A$  (par exemple, la méthode de Jacobi s'écrit  $X_{n+1} = -D^{-1}(L + U)X_n + D^{-1}B$  où  $A = L + D + U$  avec  $L$  triangulaire inférieure,  $D$  diagonale et  $U$  triangulaire supérieure) soit reliées à des algorithmes d'optimisation (gradient conjugué si  $A$  est symétrique, Bi-CG-stab ou GMRES sinon). L'avantage des méthodes itératives est qu'elles peuvent utiliser avec profit le fait que la matrice  $A$  est creuse, puisqu'elles n'utilisent que des produits matrice-vecteur : on peut ainsi ne stocker que les termes non nuls de la matrice, et accélérer l'algorithme en n'effectuant dans les produits matrice-vecteur que les produits pour les termes non nuls. De plus, si on connaît une bonne approximation de la solution on peut l'utiliser comme condition initiale dans l'algorithme, pour converger plus

rapidement (penser à des problèmes paraboliques, pour lesquels on peut utiliser la solution au temps précédent comme *initial guess*). L'inconvénient des méthodes itératives est qu'elles convergent assez mal pour des matrices mal conditionnées, et que pour un même  $A$ , il faut relancer l'algorithme pour chaque nouveau second membre  $B$ .

**Remarque 23 (Préconditionnement)** *En pratique, pour s'affranchir des difficultés liées au mauvais conditionnement de  $A$  (notamment dans les méthodes itératives), on utilise la méthode du preconditionnement qui consiste à réécrire le problème  $AX = B$  sous la forme*

$$P^{-1}AX = P^{-1}B$$

*pour une matrice  $P$  bien choisie, que l'on sait facilement inverser (ou plus précisément pour laquelle on sait facilement résoudre le problème  $PY = C$ ). On veut en particulier que  $\kappa(P^{-1}A) < \kappa(A)$ . On voit qu'un choix optimal serait  $P = A$ , mais on a justement supposé que  $A$  était difficile à inverser... A l'autre extrémité des choix possibles pour  $P$ , prendre  $P = Id$  permet de calculer très facilement  $P^{-1}$ , mais ne permet pas d'améliorer le conditionnement du problème. En pratique, il faut donc "deviner" une matrice  $P$  proche de  $A$ , soit par des méthodes algébriques (considérer des morceaux de  $A$ , comme sa diagonale par exemple), soit en utilisant des informations sur  $A$  (si  $A$  provient par exemple de la discrétisation d'un problème continu compliqué, on peut construire  $P$  en simplifiant le problème au niveau continu, puis en discrétisant ce problème plus simple).*

## 3.2 Méthode des éléments finis

### 3.2.1 Principe de la méthode

On rappelle ici simplement le principe de la méthode, et on renvoie au cours de calcul scientifique de première année pour les détails. Considérons à nouveau le problème (33) que l'on rappelle :

$$\begin{cases} \partial_t u - a\partial_x u - D\partial_{x,x}u = f \text{ sur } [0, L], \\ u(\cdot, 0) = g \text{ et } u(\cdot, L) = d, \\ u(0, \cdot) = u_0 \text{ sur } [0, L]. \end{cases}$$

On considère à nouveau le cas où  $a$  et  $D$  sont constants et on suppose dans la suite et pour simplifier que  $g = d = 0$ .

La première étape consiste à dériver une formulation variationnelle :

$$\begin{aligned} &\text{Trouver } u \in \mathcal{C}^0(L^2) \cap L^2(H_0^1) \text{ tel que } u(0, \cdot) = u_0 \text{ et} \\ &\forall v \in H_0^1 \int \partial_t uv - \int a\partial_x uv + \int D\partial_x u\partial_x v = \int fv. \end{aligned} \quad (43)$$

La deuxième étape consiste ensuite à remplacer  $H_0^1$  dans (43) par un espace vectoriel de dimension fini  $V_h \subset H_0^1$ , où  $h$  est un paramètre de discrétisation en espace. On renvoie au cours [3] pour la définition des espaces d'éléments finis. La propriété importante que nous rappelons est la suivante : pour des éléments finis

$P^k$  et sous une hypothèse de régularité du maillage, on a l'estimée d'interpolation : pour tout  $l$  tel que  $0 \leq l \leq m \leq k + 1$ ,  $\exists C, \forall u \in H^m$ ,

$$\inf_{u_h \in V_h} \|u - u_h\|_{H^{l,*}} \leq Ch^{m-l} \|u\|_{H^m}, \quad (44)$$

avec

$$\|u - u_h\|_{H^{l,*}} = \sum_{K \in \mathcal{T}_h} \|u - u_h\|_{H^l(K)}.$$

Pour  $l \leq 1$  et  $k \geq 1$ ,  $\|u - u_h\|_{H^{l,*}} = \|u - u_h\|_{H^l}$ .

En particulier, si on introduit l'opérateur d'interpolation  $\pi_h : \mathcal{C}^0(\bar{\Omega}) \rightarrow V_h$ , on a :

$$\|u - \pi_h(u)\|_{H^{l,*}} \leq Ch^{m-l} \|u\|_{H^m}. \quad (45)$$

Dans le cas où la fonction  $u \in H^m(\Omega)$  n'est pas suffisamment régulière pour que ses valeurs aux noeuds soient bien définies ( $H^m(\Omega) \not\subset \mathcal{C}^0(\bar{\Omega})$ ),  $\pi_h(u)$  n'est pas défini. On peut cependant définir un opérateur d'interpolation linéaire vérifiant (45) même pour des fonctions  $u \in H^m(\Omega)$ .

On a alors le problème discret

$$\begin{aligned} &\text{Trouver } u_h \in \mathcal{C}^0(V_h) \text{ tel que } u_h(0, \cdot) = \pi_h(u_0) \text{ et} \\ &\forall v_h \in V_h \int \partial_t u_h v_h - \int a \partial_x u_h v_h + \int D \partial_x u_h \partial_x v_h = \int f v_h. \end{aligned} \quad (46)$$

Pour discrétiser ensuite le problème en temps, on utilise typiquement les mêmes schémas que dans le cas différences finies. Ainsi, le schéma d'Euler implicite s'écrit :

$$\begin{aligned} &\text{Pour } u_h^0 = \pi_h(u_0), \text{ et } n \geq 1, \text{ trouver } u_h^n \text{ tel que} \\ &\forall v_h \in V_h \int \frac{u_h^{n+1} - u_h^n}{\delta t} v_h - \int a \partial_x u_h^{n+1} v_h + \int D \partial_x u_h^{n+1} \partial_x v_h = \int f(t_{n+1}) v_h. \end{aligned} \quad (47)$$

Ce problème est bien un problème de dimension fini. En effet, si  $V_h = \text{Vect}(v_1, \dots, v_I)$ , on écrit  $u_h^n(x) = \sum_{j=1}^I U_j^n v_j(x)$  et le problème précédent se réécrit (en prenant successivement comme fonctions tests  $v_h = v_i$  pour  $i = 1 \dots I$ ) :

$$\begin{aligned} &\text{Pour } u_h^0 = \pi_h(u_0), \text{ et } n \geq 1, \text{ trouver } u_h^n \text{ tel que} \\ &M \frac{U^{n+1} - U^n}{\delta t} - AU^{n+1} = F^{n+1}, \end{aligned}$$

où la matrice de masse  $M$  est définie par

$$M_{i,j} = \int v_i v_j,$$

et la matrice de raideur  $A$  est définie par

$$A_{i,j} = - \int a \partial_x v_j v_i + \int D \partial_x v_i \partial_x v_j.$$

**Exercice 50** Expliquer pourquoi les matrices  $M$  et  $A$  sont des matrices creuses.

**Exercice 51** Dans le cas d'un maillage du segment  $[0, L]$  avec un pas d'espace constant  $h$ , montrer que  $A$  et  $M$  sont des matrices tridiagonales :

$$A = \begin{bmatrix} 0 \dots 0 & \left(-\frac{a}{2} + \frac{D}{h}\right) & \left(-\frac{2D}{h}\right) & \left(\frac{a}{2} + \frac{D}{h}\right) & 0 \dots 0 \end{bmatrix},$$

$$M = \begin{bmatrix} 0 \dots 0 & \frac{h}{6} & \frac{2h}{3} & \frac{h}{6} & 0 \dots 0 \end{bmatrix}.$$

Comparer avec la méthode des différences finies.

**Remarque 24** L'exercice précédent pourrait laisser penser que la méthode des éléments finis n'est qu'une manière compliquée d'écrire une méthode de différences finies. En fait, la méthode des éléments finis est plus générale car elle permet de bâtir une discrétisation même pour des domaines  $\Omega$  compliqués, et en utilisant des maillages non cartésiens. Le maillage peut notamment comporter plus de degrés de libertés (i.e. comporter plus de mailles) aux endroits où la solution est moins régulière. Il existe même des techniques dites de raffinement automatique pour ajouter des mailles automatiquement là où il y en a besoin, sans connaître a priori la solution. Enfin, un dernier avantage de la méthode des éléments finis est qu'elle est plus simple à analyser mathématiquement (stabilité, convergence).

### 3.2.2 Convergence et stabilité

**Stabilité** La méthode des éléments finis permet de construire des schémas numériques naturellement stables en norme d'énergie (i.e. pour la norme considérée dans les estimées a priori). Illustrons cela sur le schéma d'Euler implicite. En prenant  $v_h = u_h^{n+1}$  comme fonction test, on obtient

$$\int \frac{u_h^{n+1} - u_h^n}{\delta t} u_h^{n+1} - \int a \partial_x u_h^{n+1} u_h^{n+1} + \int D |\partial_x u_h^{n+1}|^2 = \int f(t_{n+1}) u_h^{n+1}.$$

On en déduit

$$\frac{1}{2\delta t} \left( \int |u_h^{n+1}|^2 - \int |u_h^n|^2 \right) + \frac{1}{2\delta t} \int |u_h^{n+1} - u_h^n|^2 + \int D |\partial_x u_h^{n+1}|^2 = \int f(t_{n+1}) u_h^{n+1},$$

et donc, pour un  $\epsilon > 0$  à choisir et en utilisant l'inégalité de Poincaré

$$\begin{aligned} \frac{1}{2\delta t} \left( \int |u_h^{n+1}|^2 - \int |u_h^n|^2 \right) + \int D |\partial_x u_h^{n+1}|^2 &\leq \frac{1}{4\epsilon} \int |f(t_{n+1})|^2 + \epsilon \int |u_h^{n+1}|^2, \\ &\leq \frac{1}{4\epsilon} \int |f(t_{n+1})|^2 + C\epsilon \int |\partial_x u_h^{n+1}|^2. \end{aligned} \quad (48)$$

En choisissant  $\epsilon = \frac{D}{2C}$ , il vient

$$\frac{1}{\delta t} \left( \int |u_h^{n+1}|^2 - \int |u_h^n|^2 \right) + \int D |\partial_x u_h^{n+1}|^2 \leq \frac{C}{2D} \int |f(t_{n+1})|^2$$

et donc  $\forall 0 \leq n \leq N$ ,

$$\int |u_h^n|^2 + \delta t \sum_{k=1}^n \int D|\partial_x u_h^k|^2 \leq \int |u_h^0|^2 + \frac{C}{2D} \delta t \sum_{k=1}^n \int |f(t_k)|^2.$$

En supposant par exemple  $f \in L^\infty(0, T)$  le membre de droite est majoré par  $\int |u_h^0|^2 + \frac{CT}{2D} \|f\|_{L^\infty}^2$ . Ceci donne une borne sur  $(\int |u_h^n|^2)_{0 \leq n \leq N}$ , uniforme en  $h$  et  $\delta t$ , et prouve donc la stabilité du schéma (en norme  $L^2$  si on se réfère à la Définition 7). On comparera cette estimée avec celle que l'on obtient au niveau continu (cf. Exercice 43).

**Exercice 52** *Supposons que l'on fasse la même analyse pour le problème de Neumann homogène. Dans ce cas, la formulation variationnelle est la même, mais l'espace fonctionnel dans lequel on cherche la solution est  $\mathcal{C}^0(L^2) \cap L^2(H^1)$  (et non plus  $\mathcal{C}^0(L^2) \cap L^2(H_0^1)$ ). Par conséquent, l'espace d'élément fini  $V_h$  contient des fonctions qui ne s'annulent pas sur  $\partial\Omega$  et l'on ne peut plus utiliser l'inégalité de Poincaré pour obtenir (48). Comment adapter l'argument précédent pour montrer la stabilité du schéma ?*

Solution : On obtient alors (en choisissant  $\epsilon = 1/2$ )

$$(1 - \delta t) \int |u_h^{n+1}|^2 + 2\delta t \int D|\partial_x u_h^{n+1}|^2 \leq \int |u_h^n|^2 + \delta t \int |f(t_{n+1})|^2,$$

d'où, pour  $0 < \delta t < 1$ ,

$$\int |u_h^{n+1}|^2 + 2\delta t(1 - \delta t)^{-1} \int D|\partial_x u_h^{n+1}|^2 \leq (1 - \delta t)^{-1} \int |u_h^n|^2 + \delta t(1 - \delta t)^{-1} \int |f(t_{n+1})|^2,$$

et donc, par récurrence,  $\forall 0 \leq n \leq N$ ,

$$\begin{aligned} \int |u_h^n|^2 + 2\delta t \sum_{k=1}^n (1 - \delta t)^{-n+k-1} \int D|\partial_x u_h^k|^2 &\leq (1 - \delta t)^{-n} \int |u_h^0|^2 \\ &\quad + \delta t \sum_{k=1}^n (1 - \delta t)^{-n+k-1} \int |f(t_k)|^2. \end{aligned}$$

En utilisant le fait que, pour  $0 < \delta t < 1/2$ ,  $1 < (1 - \delta t)^{-1} \leq \exp(\alpha \delta t)$  pour un  $\alpha > 0$  bien choisi, on a donc,  $\forall 0 \leq n \leq N$ ,

$$\int |u_h^n|^2 + 2\delta t \sum_{k=1}^n \int D|\partial_x u_h^k|^2 \leq \exp(\alpha T) \left( \int |u_h^0|^2 + \|f\|_{L^\infty}^2 \right).$$

**Convergence** De même que l'analyse de la stabilité se fait naturellement en norme  $L_t^\infty(L_x^2) \cap L_t^2(H_{0,x}^1)$ , la convergence est typiquement obtenue en norme  $L_t^\infty(L_x^2) \cap L_t^2(H_{0,x}^1)$ . Pour obtenir la convergence, on a besoin de régularité sur la solution du problème continu et des estimées sur le problème discret.

On peut montrer que (en prenant  $\partial_t u$ , respectivement  $\partial_t u_h$ , comme fonction test dans (43), respectivement (46)) :

$$\begin{aligned} \|\partial_t u\|_{L^2(L^2)} &\leq C(u_0, f), \\ \|\partial_t u_h\|_{L^2(L^2)} &\leq C(u_0, f), \\ \|\partial_{t,t}^2 u_h\|_{H_*^{-1}} &\leq C(u_0, f), \end{aligned} \tag{49}$$

où  $\|\partial_{t,t}^2 u_h\|_{H_*^{-1}} = \sup_{v_h \in V_h} \frac{f(\partial_{t,t}^2 u_h)v_h}{\|\partial_x v_h\|_{L^2}}$ .

Considérons tout d'abord le problème discrétisé en espace seulement. On a :  $\forall v_h \in V_h$ ,

$$\int \partial_t(u - u_h)v_h - a \int \partial_x(u - u_h)v_h + D \int \partial_x(u - u_h)\partial_x v_h = 0.$$

Ainsi, on a

$$\begin{aligned} & \int \partial_t(u - u_h)(u - u_h) - a \int \partial_x(u - u_h)(u - u_h) + D \int \partial_x(u - u_h)\partial_x(u - u_h) \\ &= \int \partial_t(u - u_h)(u - v_h) - a \int \partial_x(u - u_h)(u - v_h) + D \int \partial_x(u - u_h)\partial_x(u - v_h). \end{aligned}$$

Et donc (pour un  $\epsilon > 0$  petit),

$$\begin{aligned} & \frac{1}{2} \int |u - u_h|^2(t) + D \int_0^t \int |\partial_x(u - u_h)|^2 \\ &= \frac{1}{2} \int |u - u_h|^2(0) + \int_0^t \int \partial_t(u - u_h)(u - v_h) - a \int_0^t \int \partial_x(u - u_h)(u - v_h) \\ & \quad + D \int_0^t \int \partial_x(u - u_h)\partial_x(u - v_h), \\ &\leq \frac{1}{2} \int |u - u_h|^2(0) + \int_0^t \int \partial_t(u - u_h)(u - v_h) + \epsilon \int_0^t \int |\partial_x(u - u_h)|^2 \\ & \quad + C(\epsilon, a, D) \int_0^t \|u - v_h\|_{H^1}^2. \end{aligned}$$

En utilisant les résultats de régularités (49) et la propriété (44), on a donc

$$\begin{aligned} & \frac{1}{2} \int |u - u_h|^2(t) + (D - \epsilon) \int_0^t \int |\partial_x(u - u_h)|^2 \\ & \leq \frac{1}{2} \int |u - u_h|^2(0) + C(\epsilon, a, D) \int_0^t \inf_{v_h \in V_h} \|u - v_h\|_{H^1}^2, \\ & \leq C(\epsilon, a, D)h^2 \|u\|_{L^2(H^2)}^2. \end{aligned}$$

On peut maintenant considérer la discrétisation en temps (47) de (46). Par différence, on a :  $\forall 0 \leq n \leq N$ ,  $\forall v_h \in V_h$ ,

$$\begin{aligned} & \int \left( \frac{u_h^{n+1} - u_h^n}{\delta t} - \partial_t u_h(t_{n+1}) \right) v_h - a \int \partial_x(u_h^{n+1} - u_h(t_{n+1}))v_h \\ & \quad + D \int \partial_x(u_h^{n+1} - u_h(t_{n+1}))\partial_x v_h = 0. \end{aligned}$$

Soit  $e^n = u_h^n - u_h(t_n)$ . En prenant  $v_h = e^{n+1}$  comme fonction test, on obtient

$$\begin{aligned} & \int \left( \frac{e^{n+1} - e^n}{\delta t} \right) e^{n+1} - a \int \partial_x e^{n+1} e^{n+1} \\ & \quad + D \int \partial_x e^{n+1} \partial_x e^{n+1} = - \int \left( \frac{u_h(t_{n+1}) - u_h(t_n)}{\delta t} - \partial_t u_h(t_{n+1}) \right) e^{n+1}. \end{aligned}$$



D'où

$$\begin{aligned} & \frac{1}{2\delta t} (\|e^{n+1}\|_{L^2}^2 - \|e^n\|_{L^2}^2) + \frac{1}{2\delta t} \|e^{n+1} - e^n\|_{L^2}^2 \\ & + D \int |\partial_x e^{n+1}|^2 = - \int \left( \frac{u_h(t_{n+1}) - u_h(t_n)}{\delta t} - \partial_t u_h(t_{n+1}) \right) e^{n+1}. \end{aligned}$$

On en déduit

$$\begin{aligned} & \frac{1}{2\delta t} (\|e^{n+1}\|_{L^2}^2 - \|e^n\|_{L^2}^2) + D \int |\partial_x e^{n+1}|^2 \\ & \leq \epsilon \int |\partial_x e^{n+1}|^2 + C(\epsilon) \delta t^2 \int_{t_n}^{t_{n+1}} \|\partial_{t,t}^2 u_h\|_{H_*^{-1}}^2. \end{aligned}$$

On obtient donc finalement,  $\forall 0 \leq n \leq N$ ,

$$\|u(t_n) - u_h^n\|_{L^2} \leq C(h + \delta t).$$

**Stabilité en norme  $L^\infty$**  Nous avons vu précédemment que dans certains cas, la méthode des éléments finis est équivalente à une méthode de différences finies avec discrétisation *centrée* des termes d'advection. On peut donc s'attendre à des problèmes de stabilité en norme  $L^\infty$  quand l'advection domine la diffusion, et c'est effectivement ce qui est observé en pratique.

Nous donnons ici deux méthodes typiques pour remédier à ce problème.

- On s'inspirant de l'*upwinding* dans le cas de la méthode des différences finies (cf. remarque 22) une idée est d'ajouter de la diffusion au niveau discret, en faisant tendre ces termes supplémentaires vers 0 quand  $h \rightarrow 0$ . Dans le cadre des éléments finis, ces méthodes sont appelées méthodes de stabilisation. On peut citer par exemple la méthode SUPG (pour *Streamline Upwind Petrov Galerkin*) ou la méthode GLS (pour *Galerkin Least Square*). Voir par exemple [6].
- Une autre solution consiste à utiliser une méthode des caractéristiques pour discrétiser le terme d'advection. Le principe est d'approximer  $\partial_t u(t, x) + a(t, x) \cdot \nabla u(t, x)$  par  $\frac{1}{\delta t} (u(t, x) - u(t - \delta t, X_{t-\delta t}^{t,x}))$  ou  $X_s^{t,x}$  satisfait l'équation différentielle ordinaire :

$$\begin{cases} \frac{d}{ds} X_s^{t,x} = a(s, X_s^{t,x}), \\ X_t^{t,x} = x. \end{cases}$$

En pratique, à chaque pas de temps  $t_n$  et pour chaque noeud  $x_i$  du maillage, on calcule une approximation de la dérivée convective en "remontant la caractéristique", *i.e.* en résolvant de manière approchée l'équation différentielle ordinaire vérifiée par  $X_s^{t_n, x_i}$ , pour  $s$  allant de  $t_n$  à  $t_{n-1}$ .

## Références

- [1] H. Brézis. *Analyse fonctionnelle*. Dunod, 1999.
- [2] E. Cancès and A. Ern. *Analyse*. Cours ENPC.
- [3] A. Ern. *Calcul scientifique*. Cours ENPC.
- [4] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Springer-Verlag, 1977.

- [5] C. Le Bris. *Systèmes multiéchelles : modélisation et simulation*, volume 47 of *Mathématiques et Applications*. Springer, 2005. In French.
- [6] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, 1997.
- [7] W. Rudin. *Analyse réelle et complexe*. Dunod, 1998.