GRAVEJAT Philippe

Agrégation

3/5/16

Chapitre VII : Résolution exacte et approchée des équations aux dérivées partielles

Introduction

Nous nous intéressons à la résolution exacte et approchée des équations aux dérivées partielles. Etant donné un sous-ensemble ouvert Ω de $\mathbb{R}^N$, une équation aux dérivées partielles d'ordre $p \in \mathbb{N}^*$ s'écrit de façon générale sous la forme :

$\forall x \in \Omega, \ F(x, u(x), du(x), d^2u(x), \ldots, d^pu(x)) = 0,$

où l'inconnue u est une fonction définie sur Ω et à valeurs dans $\mathbb{R}^M$ et la fonction F est définie sur $\Omega \times \mathbb{R}^M \times \mathbb{R}^{MN} \times \ldots \mathbb{R}^{MN^p}$ et à valeurs dans $\mathbb{R}^Q$. Les notations d^ku désignent ici les différentielles d'ordre k de l'inconnue u. En général, l'équation aux dérivées partielles précédente est munie de conditions aux limites de la forme :

$\forall x \in \partial\Omega, \ G(x, u(x), du(x), \ldots, d^pu(x)) = 0,$

où la fonction G est définie sur $\partial\Omega \times \mathbb{R}^M \times \ldots \times \mathbb{R}^{MN^p}$ et à valeurs dans $\mathbb{R}^R$. Résoudre l'équation aux dérivées partielles consiste alors à déterminer l'existence d'une solution qui satisfait ces conditions aux limites.

Dans la suite, notre objectif n'est pas de proposer un catalogue exhaustif de méthodes de résolution des équations aux dérivées partielles. Nous nous contentons d'illustrer quelques-unes des méthodes les plus classiques en les appliquant aux exemples les plus simples d'équations aux dérivées partielles. Nous nous limitons ainsi aux équations aux dérivées partielles linéaires pour lesquelles la fonction F dépend de manière linéaire des variables u, du, ... et d^pu, et nous nous restreignons aux équations en dimension un. Ceci ne signifie pas que l'entier N est égal à 1, mais que nous différencions une variable d'espace x, qui appartient à la droite réelle $\mathbb{R}$, à laquelle vient éventuellement s'ajouter une variable de temps t elle-aussi dans $\mathbb{R}$. Le nombre N est donc égal à 2 lorsque nous considérons ces deux variables simultanément, ou à 1 lorsque la variable temporelle est omise.

Notre premier exemple est fourni par les équations de Laplace et de Poisson. En dimension quelconque, l'équation de Poisson se présente sous la forme :

$$-\Delta u = f,$$

où les fonctions u et f sont définies sur Ω et à valeurs réelles. L'opérateur laplacien Δ est donné par l'identité : $\Delta = \sum_{j=1}^{N} \partial_j^2 = \sum_{j=1}^{N} \partial_{jj} = \sum_{j=1}^{N} \frac{\partial^2}{\partial x_j^2}$.

Cette équation se réduit à l'équation de Laplace :

$$-\Delta u = 0,$$

lorsque la fonction f est identiquement nulle. Les solutions u de l'équation de Laplace sont appelées fonctions harmoniques.

En dimension un, soit lorsque $N = 1$, les équations de Poisson et de Laplace se réduisent aux équations différentielles ordinaires :

$$-u'' = f.$$

Leur analyse diffère néanmoins de l'étude classique des équations différentielles ordinaires puisqu'il est naturel de munir ces équations de conditions aux limites, ou ici aux bords, de type Dirichlet :

$$\forall x \in \partial\Omega, \quad u(x) = g(x),$$

où g est une fonction définie sur $\partial\Omega$ et à valeurs réelles. La résolution du problème de Dirichlet

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

ne rentre en effet pas dans le cadre théorique du théorème de Cauchy-Lipschitz.

Afin de résoudre ce problème, nous introduisons dans la suite la notion de dérivée faible dans le cadre de la théorie des espaces de Sobolev.

Nous appliquons alors le théorème de Lax-Milgram pour construire des solutions faibles de l'équation avant de recourir à des résultats de régularité elliptique pour obtenir des solutions classiques de l'équation. Nous mettons enfin en oeuvre la méthode des différences finies pour calculer de manière approchée ces solutions.

Notre second exemple est donné par les équations de transport linéaires. En dimension quelconque, ces équations s'écrivent sous la forme :

$$\partial_t u + \sum_{j=1}^{N} c_j \, \partial_j u = f,$$

où les fonctions u, $c_1, \dots, c_N$ et f sont définies sur $\Omega \times I$ et à valeurs réelles. L'ensemble Ω demeure ici un ouvert de $\mathbb{R}^N$ et I désigne un intervalle de $\mathbb{R}$ qui contient 0. Il est naturel de munir ces équations de conditions aux limites, ou ici de conditions initiales, du type :

$$\forall x \in \Omega, \quad u(x,0) = u_0(x),$$

où la donnée initiale u_0 est définie sur Ω et à valeurs réelles. En dimension un, soit lorsque $N=1$, ce problème de Cauchy se réduit au système :

$$\left\{ \begin{array}{l} \partial_t u + c \, \partial_x u = f \quad \text{sur } \Omega \times I, \\ \text{et } u(\cdot, 0) = u_0 \quad \text{sur } \Omega. \end{array} \right.$$

Nous abordons la résolution de ce problème par la méthode des caractéristiques, que nous étendons ensuite à l'équation des ondes

$$\partial_{tt} u - \partial_{xx} u = f,$$

où les fonctions u et f sont définies sur $\mathbb{R}^2$ et à valeurs réelles. Ceci nous permet d'établir la formule de D'Alembert pour le problème de Cauchy :

$$\left\{ \begin{array}{l} \partial_{tt} u - \partial_{xx} u = 0 \quad \text{sur } \mathbb{R}^2, \\ \text{et } u(\cdot, 0) = u_0 \text{ et } \partial_t u(\cdot, 0) = u_1 \quad \text{sur } \mathbb{R}, \end{array} \right.$$

où les conditions initiales u_0 et u_1 sont définies sur $\mathbb{R}$ et à valeurs réelles. Nous mettons ensuite en oeuvre le schéma de type différences finies de Lax-Friedrichs afin de calculer de manière approchée les solutions du problème de Cauchy associé à une équation de transport linéaire.

Nous concluons enfin par l'exemple des équations de la chaleur, qui se présentent en dimension quelconque sous la forme:

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

où les fonctions u et f sont définies sur $\Omega \times I$ et à valeurs réelles. L'ensemble Ω désigne ici un ouvert de $\mathbb{R}^M$ et I renvoie à un intervalle de $\mathbb{R}$ qui contient 0. Ces équations sont en général munies des conditions aux limites

$$\begin{cases} \forall t \in I, \ \forall x \in \partial\Omega, \ u(x,t) = g(x,t), \\ \forall x \in \Omega, \ u(x,0) = u_0(x), \end{cases}$$

où la condition au bord g est définie sur $\partial\Omega \times I$ et à valeurs réelles, tandis que la condition initiale u_0 est définie sur Ω et à valeurs réelles.

Nous résolvons le problème de Cauchy en dimension un :

$$\begin{cases} \partial_t u - \partial_{xx} u = f \quad \text{sur } \Omega \times I, \\ u = g \quad \text{sur } \partial\Omega \times I, \\ u(\cdot, 0) = u_0 \quad \text{sur } \Omega, \end{cases}$$

par le calcul de la solution fondamentale lorsque Ω est la droite réelle $\mathbb{R}$, puis par la méthode de séparation des variables lorsque Ω se réduit à un intervalle borné de $\mathbb{R}$. Nous approchons ensuite les solutions de ce second problème par les méthodes de type différences finies d'Euler explicites et implicites.

Les applications des méthodes classiques de résolution des équations aux dérivées partielles que nous analysons dans la suite ne se réduisent bien sûr pas aux seuls exemples basiques que nous avons mentionnés. Nous renvoyons à l'ouvrage de L. C. Evans intitulé "Partial Differential Equations" pour une approche théorique bien plus exhaustive des équations aux dérivées partielles. Nous mentionnons par ailleurs dans la suite plusieurs ouvrages ayant trait aux diverses méthodes numériques que nous étudions, et à leurs possibles généralisations.

I Équations elliptiques du second ordre

Dans ce paragraphe, nous restreignons l'analyse des équations elliptiques du second ordre à un intervalle ouvert et borné de $\mathbb{R}$. Quitte à translater et à dilater, nous supposons dans la suite que cet intervalle est égal à : $I =]0, 1[$.

1. Définitions et exemples

Définitions : (i) Un opérateur différentiel L sur I est un opérateur du second ordre si et seulement s'il existe des fonctions a, b et c (bornées) sur I telles que :

$$\forall u \in \mathcal{C}^2(I), \quad L(u) := -a u'' + b u' + c u.$$

(ii) Un opérateur différentiel L sur I est un opérateur du second ordre sous forme divergence si et seulement s'il existe des fonctions a, b et c (bornées) sur I telles que :

$$\forall u \in \mathcal{C}^2(I), \quad L(u) := -(a u')' + b u' + c u.$$

Exemples : (i) L'opérateur laplacien $L(u) = -u''$ est un opérateur différentiel du second ordre (sous forme divergence).

(ii) (Lorsque la fonction potentielle V est bornée sur I,) l'opérateur de Schrödinger $L(u) = -u'' + V u$ est un opérateur différentiel du second ordre (sous forme divergence).

Un opérateur $L(u) = -a u'' + b u' + c u$ du second ordre s'écrit sous la forme divergence :

$$L(u) = -(a u')' + (b + a') u' + c u,$$

dès que la fonction a est dérivable sur I (, à dérivée bornée). L'opération inverse est possible sous les mêmes hypothèses. Afin de simplifier la présentation, nous nous restreignons dans la suite à des opérateurs différentiels du second ordre sous forme divergence.

Définition : Soit a, b et c des fonctions bornées sur I. L'opérateur L du second ordre sous forme divergence défini par :
$$\forall u \in \mathcal{E}^2(I), \quad L(u) = -(au')' + bu' + cu,$$
est elliptique si et seulement s'il existe un nombre strictement positif λ tel que :
$$\forall x \in I, \quad a(x) \geq \lambda.$$

Exemple : L'opérateur laplacien est elliptique sur I avec une constante d'ellipticité $\lambda = 1$.

En dimension un comme ici, la terminologie elliptique n'est pas des plus claire. En dimension N quelconque, un opérateur L du second ordre sous forme divergence s'écrit sous la forme :
$$\forall u \in \mathcal{E}^2(\Omega), \quad L(u) = -\sum_{1 \leq i,j \leq N} \partial_i (a_{ij} \partial_j u) + \sum_{k=1}^{N} b_k \partial_k u + cu,$$
où les fonctions $(a_{ij})_{1 \leq i,j \leq N}$, $(b_k)_{1 \leq k \leq N}$ et c sont bornées sur un ouvert Ω de $\mathbb{R}^N$. L'opérateur est alors elliptique si et seulement s'il existe un nombre strictement positif λ tel que :
$$\forall x \in \Omega, \ \forall \xi \in \mathbb{R}^N, \quad \sum_{1 \leq i,j \leq N} a_{ij} \xi_i \xi_j \geq \lambda \sum_{i=1}^{N} \xi_i^2.$$
En dimension $N = 2$, cette inégalité assure que les lignes de niveau de la forme quadratique $\xi \mapsto \sum_{1 \leq i,j \leq 2} a_{ij} \xi_i \xi_j$ sont des ellipses, ce qui explique la terminologie d'opérateurs elliptiques.

Définitions : (i) Une équation différentielle linéaire du second ordre sur I sous forme divergence est une équation de la forme :
$$-(au')' + bu' + cu = f,$$
où a, b et c sont des fonctions bornées sur I, et f est une fonction mesurable sur I.

(ii) L'équation différentielle linéaire du second ordre sur I sous forme divergence :
$$-(au')' + bu' + cu = f,$$

est elliptique si et seulement si l'opérateur L du second ordre sous forme divergence qui lui est associé :

$$L(u) = -(a\,u')' + b\,u' + c\,u,$$

est elliptique.

<u>Exemples</u> : Les équations de Poisson, et de Laplace, $-u'' = f$, sont des équations elliptiques linéaires du second ordre sous forme divergence.

<u>Définitions</u> : Soit a, b et c, des fonctions bornées sur I, f, une fonction mesurable sur I, et (α, β), deux nombres réels.

(i) Le problème de Dirichlet du second ordre associé aux coefficients a, b et c, au second membre f, et aux données aux bords α et β est le système différentiel :

$$(\mathcal{E}) \begin{cases} \forall x \in I, -(a(x)u'(x))' + b(x)\,u'(x) + c\,u(x) = f(x), \\ u(0) = \alpha \quad, u(1) = \beta. \end{cases}$$

(ii) Le problème de Dirichlet du second ordre est elliptique si et seulement si l'opérateur L du second ordre sous forme divergence qui lui est associé :

$$L(u) = -(a\,u')' + b\,u' + c\,u,$$

est elliptique.

<u>Exemple</u> : Le problème de Dirichlet :

$$(\mathcal{E}) \begin{cases} \forall x \in I, \quad -u''(x) = f(x), \\ u(0) = \alpha, \; u(1) = \beta, \end{cases}$$

est elliptique.

Dans la suite, nous cherchons à résoudre les problèmes de Dirichlet elliptiques du second ordre tant sur le plan théorique que sur le plan pratique. Lorsque les fonctions a, b, c et f ne sont pas assez régulières, il n'est déjà pas immédiat de donner un sens rigoureux aux équations du second ordre sous-jacentes. Afin de résoudre cette difficulté, nous introduisons maintenant la notion de dérivation faible, et les espaces de Sobolev associés.

2. Dérivation faible et espaces de Sobolev

Désignons par la notation $\mathcal{C}_c^\infty(I)$ l'ensemble des fonctions de classe $\mathcal{C}^\infty$ sur I, à support compact inclus dans I.

a. Notion de dérivée faible

Cette notion est définie à partir de la formule d'intégration par parties.

Définition: L'espace des fonctions localement intégrables sur I est défini par:

$$L^1_{loc}(I) := \left\{ f : I \to \mathbb{R} \text{ t.q. } f \text{ est mesurable sur } I \text{ et } \forall\, 0 < a \leq b < 1,\ \int_a^b |f(x)|\, dx < +\infty \right\}.$$

Exemple: La fonction $x \mapsto \frac{1}{x}$ est localement intégrable sur $]0,1[$, mais n'est pas intégrable sur $]0,1[$.

Définition: Soit $f \in L^1_{loc}(I)$. La fonction f possède une dérivée faible sur I si et seulement s'il existe une fonction $g \in L^1_{loc}(I)$ telle que
$$\forall \varphi \in \mathcal{C}_c^\infty(I),\ \int_I f\,\varphi' = -\int_I g\,\varphi.$$

Lemme: Soit $f \in L^1_{loc}(I)$. Si la fonction f possède une dérivée faible sur I, alors, cette dérivée faible est l'unique fonction $g \in L^1_{loc}(I)$ telle que:
$$\forall \varphi \in \mathcal{C}_c^\infty(I),\ \int_I f\,\varphi' = -\int_I g\,\varphi.$$
La dérivée faible de la fonction f est alors notée: $f' = g$.

Preuve:

Supposons qu'il existe deux fonctions g_1 et g_2 de $L^1_{loc}(I)$ telles que:

$\forall \varphi \in \mathcal{C}_c^\infty(I)$, $-\int_I g_1 \varphi = \int_I f \varphi' = -\int_I g_2 \varphi$.

Dans ce cas, la fonction $h = g_2 - g_1$ satisfait:

$\forall \varphi \in \mathcal{C}_c^\infty(I)$, $\int_I h \varphi = 0$.

Pour $m \in \mathbb{N}$, nous introduisons alors les intervalles:

$I_m = \left] \frac{1}{m+3}, 1 - \frac{1}{m+3} \right[$,

ainsi que des fonctions plateaux $\chi_m \in \mathcal{C}_c^\infty(I)$ telles que:

$\chi_m = 0$ sur $I \setminus I_{2m}$, $0 \leq \chi_m \leq 1$ sur I, et $\chi_m = 1$ sur I_m.

Il vient:

$\forall \varphi \in \mathcal{C}^\infty(I)$, $\int_I h \chi_m \varphi = 0$.

De plus, la fonction $h \chi_m$ est intégrable sur I, et elle se prolonge en une fonction intégrable sur $\mathbb{R}$ en posant:

$\forall x \in \mathbb{R} \setminus I$, $[h \chi_m](x) = 0$.

Nous introduisons maintenant une approximation de l'identité $(\rho_n)_{n \in \mathbb{N}}$, soit une suite de fonctions $\rho_n \in \mathcal{C}_c^\infty(\mathbb{R})$ telles que:

$\forall n \in \mathbb{N}$, $\rho_n \geq 0$, $\int_{\mathbb{R}} \rho_n = 1$ et $\operatorname{supp}(\rho_n) \subset \left] -\frac{1}{n+2}, \frac{1}{n+2} \right[$.

Nous obtenons:

$\forall n \in \mathbb{N}, \forall x \in \mathbb{R}$, $[h\chi_m] * \rho_n(x) = \int_{\mathbb{R}} [h\chi_m](x-y)\, \rho_n(y)\, dy = 0$.

Nous vérifions par ailleurs que:

$$\int_{\mathbb{R}} \left| [h\chi_m] * \rho_n(x) - [h\chi_m](x) \right| dx \leq \int_{\mathbb{R}} dx \int_{\mathbb{R}} dy \left| [h\chi_m](x-y) - [h\chi_m](x) \right| \rho_n(y)$$
$$\leq \int_{\mathbb{R}} \rho_n(y) \sup_{|y| \leq \frac{1}{n+2}} \left\| [h\chi_m](\cdot - y) - h\chi_m \right\|_{L^1(\mathbb{R})} dy$$

Lorsque $g \in L^1(\mathbb{R})$, nous savons qu'il existe une fonction $g_\varepsilon \in \mathcal{C}_c^0(\mathbb{R})$ telle que:

$\| g - g_\varepsilon \|_{L^1(\mathbb{R})} \leq \varepsilon$,

pour tout nombre strictement positif ε. En particulier,

$$\sup_{|y| \leq \frac{1}{n+2}} \| g(\cdot - y) - g \|_{L^1(\mathbb{R})} \leq 2\varepsilon + \sup_{|y| \leq \frac{1}{n+2}} \| g_\varepsilon(\cdot - y) - g_\varepsilon \|_{L^1(\mathbb{R})}.$$

Nous déduisons alors de l'uniforme continuité de la fonction g_ε et de son support compact que:

$$\sup_{|y| \leq \frac{1}{n+2}} \| g_\varepsilon(\cdot - y) - g_\varepsilon \|_{L^1(\mathbb{R})} \xrightarrow[n \to +\infty]{} 0.$$

Ceci nous permet de conclure que:

$$\sup_{|y| \leq \frac{1}{n+1}} \| g(\cdot - y) - g \|_{L^2(\mathbb{R})} \underset{n \to +\infty}{\longrightarrow} 0.$$

En particulier, quel que soit le nombre strictement positif δ, il existe un entier N tel que :

$$\forall n \geq N, \quad \sup_{|y| \leq \frac{1}{n+1}} \| g(\cdot - y) - g \|_{L^2(\mathbb{R})} \leq \delta.$$

Il s'ensuit que :

$$\forall n \geq N, \quad \| (h \chi_m) * \rho_n - h \chi_m \|_{L^2} \leq \delta \int_{\mathbb{R}} \rho_n(y)\, dy = \delta.$$

Nous aboutissons donc à :

$$\forall n \geq N, \quad \| h \chi_m \|_{L^2} \leq \delta,$$

De sorte que la fonction $h \chi_m$ est nulle presque partout. Par recouvrement dénombrable, il en va de même pour la fonction h. D'où $g_2 = g_1$ p.p., ce qui conclut la preuve.

Exemples : (i) Si f est de classe $\mathcal{C}^1$ sur I, alors, f a une dérivée faible sur I, et sa dérivée faible est égale à sa dérivée classique.

(ii) La fonction g définie par :

$$\forall x \in]0,1[, \; g(x) = 0 \text{ si } x \leq \tfrac{1}{2}, \; x - \tfrac{1}{2} \text{ sinon}$$

n'est pas de classe $\mathcal{C}^1$ sur I, mais a pour dérivée faible la fonction :

$$\text{pp } x \in]0,1[, \; g'(x) = 0 \text{ si } x \leq \tfrac{1}{2}, \; 1 \text{ sinon.}$$

b. Espaces de Sobolev $H^1(I)$ et $H^1_0(I)$

Définition : L'espace de Sobolev $H^1(I)$ est défini par :

$$H^1(I) := \{ f \in L^2(I) \text{ t.q. } f \text{ a une dérivée faible } f' \text{ sur } I \text{ et } f' \in L^2(I) \}.$$

Exemple : Si f est de classe $\mathcal{C}^1$ sur $[0,1]$, alors, la restriction de la fonction f à l'intervalle I appartient à l'espace de Sobolev $H^1(I)$.

De manière équivalente à la définition précédente par les dérivées faibles, l'espace de Sobolev $H^1(I)$ est l'ensemble des fonctions de carré intégrable sur I, qui sont égales à une primitive de fonction de carré intégrable.

<u>Théorème</u> : Soit $f \in H^1(I)$. Il existe un nombre réel (ou complexe) A tel que :

$$\text{pp } t\, x \in I, \quad f(x) = A + \int_{\frac{1}{2}}^{x} f'(y)\, dy.$$

<u>Preuve</u> :

Soit $\forall x \in I$, $g(x) = \int_{\frac{1}{2}}^{x} f'(y)\, dy$. Comme $f' \in L^2(I) \subset L^1(I)$, la fonction g est bien définie sur I. De plus, par l'inégalité de Cauchy - Schwarz, elle vérifie :

$$\forall (x, x') \in I^2, \quad |g(x') - g(x)| \leq |x - x'|^{\frac{1}{2}} \left(\int_I (f')^2 \right)^{\frac{1}{2}}.$$

De sorte que g est continue sur I. Montrons maintenant que g a une dérivée faible sur I qui est égale à f'. En effet,

$$\begin{aligned}
\forall \varphi \in \mathcal{C}_c^\infty(I), \quad \int_I g\, \varphi' &= \int_0^1 \int_{\frac{1}{2}}^{x} f'(y)\, \varphi'(x)\, dy\, dx \\
&= \int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^{x} f'(y)\, \varphi'(x)\, dy\, dx + \int_{\frac{1}{2}}^{1} \int_{\frac{1}{2}}^{x} f'(y)\, \varphi'(x)\, dx \\
&= - \int_0^{\frac{1}{2}} dy \int_0^{y} dx\, f'(y)\, \varphi'(x) + \int_{\frac{1}{2}}^{1} dy \int_y^1 dx\, f'(y)\, \varphi'(x) \\
&= - \int_0^1 f' \varphi.
\end{aligned}$$

Il vient donc :

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \quad \int_I (g - f)\, \varphi' = 0.$$

Soit alors $\chi \in \mathcal{C}_c^\infty(I)$ telle que $\int_I \chi = 1$. Posons :

$$\forall x \in I, \quad \Phi(x) = \int_0^x \left(\varphi(y) - \left(\int_I \varphi \right) \chi(y) \right) dy.$$

La fonction Φ appartient à $\mathcal{C}_c^\infty(I)$ et :

$$\forall x \in I, \quad \Phi'(x) = \varphi(x) - \left(\int_I \varphi \right) \chi(x).$$

D'après l'identité précédente, nous avons donc :

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \quad \int_I (g - f) \left(\varphi - \left(\int_I \varphi \right) \chi \right) = 0,$$

soit encore :

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \quad \int_I \left[g - f - \int_I (g - f) \chi \right] \varphi = 0.$$

D'après la démonstration de l'unicité d'une dérivée faible, nous concluons que

$$\text{pp } t\, x \in I, \quad g(x) = f(x) + \int_I (g - f) \chi.$$

La conclusion s'ensuit pour $A = - \int_I (g - f) \chi$.

Une première conséquence de cette caractérisation est la description fonctionnelle suivante de l'espace $H^1(I)$.

<u>Théorème</u>: L'espace de Sobolev $H^1(I)$ est un espace de Hilbert pour le produit scalaire:

$$\forall (f,g) \in H^1(I)^2, \langle f, g\rangle_{H^1(I)} := \int_I \left[f'(x)\,\overline{g'(x)} + f(x)\,\overline{g(x)} \right] dx.$$

Le sous-espace $\mathcal{C}^\infty(\mathbb{R}) \cap H^1(I)$ est dense dans l'espace $H^1(I)$.

<u>Preuve</u>:

L'espace de Sobolev $H^1(I)$ est un espace vectoriel, et l'application $\langle\, ,\, \rangle_{H^1(I)}$ définit un produit scalaire sur cet espace.

De plus, si $(f_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour la norme associée, alors, $(f_n)_{n\in\mathbb{N}}$ et $(f'_n)_{n\in\mathbb{N}}$ sont des suites de Cauchy de $L^2(I)$. Il existe donc des fonctions f et g de $L^2(I)$ telles que:

$$f_n \xrightarrow[n\to+\infty]{} f \text{ dans } L^2(I) \text{ et } f'_n \xrightarrow[n\to+\infty]{} g \text{ dans } L^2(I)$$

Comme

$$\forall n \in \mathbb{N}, \ \forall \varphi \in \mathcal{C}_c^\infty(I), \int_I f_n \varphi' = -\int_I f'_n \varphi,$$

il vient à la limite $n \to +\infty$:

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \int_I f\varphi' = -\int_I g\varphi.$$

La fonction g est donc la dérivée faible f' de f. En particulier, f appartient à $H^1(I)$, et:

$$f_n \xrightarrow[n\to+\infty]{} f \text{ dans } H^1(I),$$

D'où la complétude de l'espace de Sobolev $H^1(I)$.

Soit alors $f \in H^1(I)$. Par le théorème précédent, il existe un nombre réel A tel que:

$$\text{pp } t\, x \in I, \ f(x) = A + \int_{\frac{1}{2}}^x f'(y)\, dy.$$

Soit alors $(g_n)_{n\in\mathbb{N}} \in \mathcal{C}_c^\infty(I)^{\mathbb{N}}$ telle que:

$$g_n \xrightarrow[n\to+\infty]{} f' \text{ dans } L^2(I).$$

Notons:

$$\forall x \in I, \ \forall n \in \mathbb{N}, \ f_n(x) = A + \int_{\frac{1}{2}}^x g_n(y)\, dy.$$

Les fonctions f_n se prolongent en fonction de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, qui appartiennent donc à $H^1(I)$. De plus,

$$f'_n = g_n \underset{n\to+\infty}{\longrightarrow} f' \text{ dans } L^2(I),$$

Et, par l'inégalité de Cauchy-Schwarz,

$$\forall n \in \mathbb{N}, \quad \int_I |f_n - f|^2 = \int_I \left| \int_{\frac{1}{2}}^{x} (g_n(y) - f'(y))\, dy \right|^2 dx$$
$$\leq \int_I \left| x - \frac{1}{2} \right| \int_I |g_n - f'|^2\, dx$$
$$\underset{n\to+\infty}{\longrightarrow} 0,$$

D'où la densité de $\mathcal{C}^\infty([0,1]) \cap H^1(I)$ dans $H^1(I)$.

Une seconde conséquence de la caractérisation par les primitives réside dans le théorème d'injection de Sobolev, qui assure que chaque fonction de $H^1(I)$ possède un unique représentant continu sur le segment $[0,1]$. L'injection de $H^1(I)$ dans l'espace $\mathcal{C}^0([0,1])$ est de plus continue.

<u>Théorème d'injection de Sobolev</u>: (i) Soit $f \in H^1(I)$. La fonction f admet un unique représentant qui se prolonge en une fonction continue sur $[0,1]$. Nous identifions f à ce représentant dans la suite.

(ii) Soit $f \in H^1(I)$. Alors:

$$\forall (x,y) \in [0,1]^2, \quad f(y) = f(x) + \int_x^y f'(z)\, dz$$

(iii) L'application $\begin{array}{ccc} H^1(I) & \to & \mathcal{C}^0([0,1]) \\ f & \mapsto & f \end{array}$ est bien définie, linéaire et continue.

<u>Preuve:</u>

Par le théorème précédent, il existe un nombre réel A tel que:

$$\text{pp}^t\ x \in I, \quad f(x) = A + \int_{\frac{1}{2}}^{x} f'(y)\, dy$$

Notons:

$$\forall x \in I, \quad \tilde{f}(x) = A + \int_{\frac{1}{2}}^{x} f'(y)\, dy.$$

Par l'inégalité de Cauchy-Schwarz, la fonction $\tilde{f}$ se prolonge en une application continue sur $[0,1]$, qui est l'unique représentant de f dans cette classe de fonctions. De plus,

$$\forall (x,y) \in [0,1]^2, \quad \tilde{f}(y) - \tilde{f}(x) = \int_{\frac{1}{2}}^{y} f'(z)\, dz - \int_{\frac{1}{2}}^{x} f'(z)\, dz = \int_x^y f'(z)\, dz.$$

Il vient alors:

$$|f(y)| = \int_0^1 |f(y)|\,dx \leq \int_0^1 \left[|f(x)| + \int_x^y |f'(z)|\,dz \right] dx,$$

De sorte que par l'inégalité de Cauchy-Schwarz,

$$|f(y)| \leq \|f\|_{L^2(I)} + \|f'\|_{L^2(I)} \leq \|f\|_{H^1(I)}.$$

L'application $\begin{array}{l} H^1(I) \to \mathcal{C}^0([0,1]) \\ \quad f \mapsto f \end{array}$ est donc bien définie, linéaire et continue

Grâce au théorème d'injection de Sobolev, il est possible de définir la notion de données au bord pour une fonction $f \in H^1(I)$. En particulier, il est possible d'introduire le sous-ensemble $H^1_0(I)$ des fonctions nulles sur $\{0,1\}$.

Définition: L'espace de Sobolev $H^1_0(I)$ est défini par:

$$H^1_0(I) = \{ f \in H^1(I) \text{ t.q. } f(0) = f(1) = 0 \}.$$

Les fonctions de $H^1_0(I)$ satisfont l'inégalité de Poincaré.

Théorème (Inégalité de Poincaré): Soit $f \in H^1_0(I)$. Alors:

$$\|f\|_{L^2(I)} \leq \frac{1}{\sqrt{2}} \|f'\|_{L^2(I)}$$

Preuve:

En effet,

$$\forall x \in I, \quad f(x) = \int_0^x f'(y)\,dy$$

De sorte que, par l'inégalité de Cauchy-Schwarz,

$$\int_0^1 f^2 \leq \int_0^1 \left(\int_0^x |f'(y)|\,dy \right)^2 dx \leq \int_0^1 x \int_I (f')^2\,dx = \frac{1}{2} \int_I (f')^2$$

Cette inégalité conduit au résultat suivant.

Théorème: L'espace de Sobolev $H^1_0(I)$ est un espace de Hilbert pour le produit scalaire:

$$\forall (f,g) \in H^1_0(I)^2, \quad \langle f, g \rangle_{H^1_0(I)} = \int_I f'(x)\,\overline{g'(x)}\,dx.$$

De plus, cet espace est l'adhérence du sous-espace $\mathcal{C}_c^\infty(I)$

dans $H^1(I)$.

Preuve:

Il résulte du théorème d'injection de Sobolev que $H^1_0(I)$ est un sous-espace fermé de $H^1(I)$ qui contient le sous-espace $\mathcal{C}^\infty_c(I)$. De plus, si $f \in H^1_0(I)$, alors, il existe une suite $(g_n)_{n\in\mathbb{N}} \in \mathcal{C}^\infty_c(I)^{\mathbb{N}}$ telle que:

$$g_n \underset{n\to+\infty}{\longrightarrow} f' \text{ dans } L^2(I).$$

Par l'inégalité de Cauchy-Schwarz, il vient:

$$\left|\int_0^1 (g_n - f')\right| \leq \|g_n - f'\|_{L^2(I)} \underset{n\to+\infty}{\longrightarrow} 0,$$

De sorte que:

$$\int_0^1 g_n \underset{n\to+\infty}{\longrightarrow} \int_0^1 f' = 0.$$

Soit alors $\chi \in \mathcal{C}^\infty_c(I)$ telle que $\int_I \chi = 1$. Posons:

$$\forall x \in I,\ f_n(x) = \int_0^x \left[g_n(y) - \left(\int_0^1 g_n\right)\chi(y)\right] dy.$$

Les fonctions f_n appartiennent à $\mathcal{C}^\infty_c(I)$ et nous pouvons vérifier que:

$$f_n \underset{n\to+\infty}{\longrightarrow} f \text{ dans } H^1(I).$$

Ceci assure que $H^1_0(I)$ est l'adhérence de $\mathcal{C}^\infty_c(I)$ dans $H^1(I)$.

Enfin, par l'inégalité de Poincaré,

$$\forall f \in H^1_0(I),\ \|f\|_{H^1(I)} \leq \sqrt{\tfrac{3}{2}}\, \|f\|_{H^1_0(I)},$$

De sorte que toute suite de Cauchy de $H^1_0(I)$ pour la norme $\|\ \|_{H^1_0(I)}$ est une suite de Cauchy pour la norme $\|\ \|_{H^1(I)}$. Le fait que $H^1_0(I)$ soit un espace de Hilbert pour la norme $\|\ \|_{H^1_0(I)}$ résulte alors de son caractère fermé dans $H^1(I)$.

c. Espaces de Sobolev $H^k(I)$

Soit $k \in \mathbb{N}$, avec $k \geq 2$.

Définition: L'espace de Sobolev $H^k(I)$ est défini par:

$$H^k(I) := \{ f \in L^2(I) \text{ t.q. } \forall\, 0 \leq j \leq k-1,\ f^{(j)} \text{ a une dérivée faible } f^{(j+1)} \text{ et } f^{(j+1)} \in L^2(I) \}$$

Exemple: Si f est de classe $\mathcal{C}^k$ sur $[0,1]$, alors la restriction de la fonction f a l'intervalle I appartient à l'espace de Sobolev $H^k(I)$.

L'espace de Sobolev $H^k(I)$ peut aussi se définir suivant le principe de récurrence qui suit.

<u>Lemme</u>: L'espace de Sobolev $H^k(I)$ est égal à:

$$H^k(I) := \{ f \in L^2(I) \text{ t.q. } f \text{ a une dérivée faible } f' \text{ sur } I \text{ et } f' \in H^{k-1}(I) \}.$$

<u>Preuve</u>:

Immédiat.

En particulier, les espaces de Sobolev $H^k(I)$ sont emboîtés: $H^k(I) \subset H^{k-1}(I) \subset \dots \subset H^1$

Grâce à ce lemme, il est possible d'établir les propriétés fonctionnelles suivantes de l'espace de Sobolev $H^k(I)$.

<u>Théorème</u>: L'espace de Sobolev $H^k(I)$ est un espace de Hilbert pour le produit scalaire:

$$\forall (f,g) \in H^k(I)^2, \ \langle f, g \rangle_{H^k(I)} = \sum_{j=0}^{k} \int_I f^{(j)} \overline{g^{(j)}} = \langle f, g \rangle_{L^2(I)} + \langle f', g' \rangle_{H^{k-1}(I)}.$$

Le sous-espace $\mathcal{C}^\infty(\mathbb{R}) \cap H^k(I)$ est dense dans l'espace $H^k(I)$.

<u>Preuve</u>:

L'espace de Sobolev $H^k(I)$ est un espace vectoriel, et l'application $\langle , \rangle_{H^k}$ est un produit scalaire sur cet espace. Son caractère complet s'établit par récurrence à l'aide du lemme précédent comme dans le cas de l'espace $H^1(I)$. Il en va de même pour la densité du sous-espace $\mathcal{C}^\infty(\mathbb{R}) \cap H^k(I)$ qui s'établit aussi par récurrence à l'aide du lemme précédent et de la formule:

$$\forall f \in H^k(I), \ \forall x \in [0,1], \ f(x) = f(0) + \int_0^x f'(y)\,dy.$$

Le théorème d'injection de Sobolev s'étend aussi à l'espace $H^k(I)$.

Théorème d'injection de Sobolev : (i) Soit $f \in H^k(I)$. La fonction f admet un unique représentant qui se prolonge en une fonction de classe $\mathcal{C}^{k-1}$ sur $[0,1]$.

(ii) Nous identifions f à ce représentant dans la suite, et nous avons alors :

$$\forall (x,y) \in [0,1]^2, \quad f(y) = f(x) + \int_x^y f'(z)\,dz$$

(iii) L'application $\begin{array}{ccc} H^k(I) & \to & \mathcal{C}^{k-1}(I) \\ f & \mapsto & f \end{array}$ est bien définie, linéaire et continue.

Preuve :

Comme $H^k(I) \subset H^1(I)$, la fonction f peut s'identifier à son unique représentant continu, lequel satisfait la formule de (ii). Par récurrence sur $k \geq 2$, la fonction f' est de classe $\mathcal{C}^{k-2}$ sur $[0,1]$ par la formule de (ii). Le caractère continu de l'injection de (iii) résulte aussi d'une récurrence sur $k \geq 1$ et du lemme précédent.

3. Solutions faibles et théorème de Lax-Milgram

Nous nous donnons deux nombres réels α et β, des fonctions a, b et c mesurables et bornées sur I, et une fonction f mesurable sur I, et nous considérons le problème de Dirichlet du second ordre sous forme divergence :

$$(\mathcal{E}) \left\{ \begin{array}{l} \forall x \in I, \quad -(a(x)\,u'(x))' + b(x)\,u'(x) + c(x)\,u(x) = f(x), \\ u(0) = \alpha \quad \text{et} \quad u(1) = \beta. \end{array} \right.$$

Sous des hypothèses à préciser sur les fonctions a, b et c, nous allons établir l'existence de solutions de ce problème à l'aide du théorème de Lax-Milgram.

a. Solutions classiques, fortes et faibles

Nous commençons par donner un sens à la notion de solutions du problème de Dirichlet précédent. En principe, la résolution d'une équation différentielle ordinaire du second ordre requiert une régularité de type $\mathcal{C}^2$. Une solution classique possède une régularité de ce type, et satisfait l'équation en tout point.

<u>Définition</u>: Le problème de Dirichlet précédent admet une solution classique u si et seulement s'il existe une fonction $u \in \mathcal{C}^1(I) \cap \mathcal{C}^0([0,1])$ telle que la fonction au' de classe $\mathcal{C}^1$ sur I, et :

$$\left\{\begin{array}{l} \forall x \in I, \ -(au')'(x) + b(x)\, u'(x) + c(x)\, u(x) = f(x), \\ \text{et } u(0) = \alpha \text{ et } u(1) = \beta. \end{array}\right.$$

Il n'existe pas toujours de solution classique à un problème de Dirichlet.

<u>Exemple</u>: Soit $\forall x \in I$, $f(x) = 1$ si $0 < x < \frac{1}{2}$, 0 si $\frac{1}{2} < x < 1$. Le problème de Dirichlet :

$$\left\{\begin{array}{l} \forall x \in I, \ -u''(x) = f(x), \\ \text{et } u(0) = \alpha \ \text{ et } \ u(1) = \beta, \end{array}\right.$$

ne possède pas de solution u de classe $\mathcal{C}^2$ sur I.

L'introduction des dérivées faibles permet de remédier à ce type de problèmes. Elle conduit à la notion de solution forte du problème de Dirichlet précédent pour laquelle les dérivées qui apparaissent dans l'équation sont des dérivées faibles, et l'équation n'est satisfaite que presque partout.

<u>Définition</u>: Le problème de Dirichlet précédent admet une solution forte u si et seulement s'il existe une fonction $u \in \mathcal{C}^0([0,1])$ telle que les fonctions u et au' admettent des dérivées faibles sur I et :

$$\left\{\begin{array}{l} \text{p.p.t. } x \in I, \ -(au')'(x) + b(x)\, u'(x) + c(x)\, u(x) = f(x), \\ \text{et } u(0) = \alpha \ \text{ et } u(1) = \beta. \end{array}\right.$$

<u>Exemple</u>: Soit $\forall x \in I$, $f(x) = 1$ si $0 < x < \frac{1}{2}$, 0 si $\frac{1}{2} < x < 1$. La fonction u définie par : $\forall x \in I$, $u(x) = -\frac{1}{2}\left(x^2 - \frac{3}{4}x\right)$ si $0 \leq x \leq \frac{1}{2}$, $\frac{1}{8}(1-x)$ si $\frac{1}{2} \leq x \leq 1$, est une solution forte du problème de Dirichlet :

$$\left\{\begin{array}{l} \forall x \in I, \ -u''(x) = f(x), \\ u(0) = 0 \ \text{ et } \ u(1) = 0. \end{array}\right.$$

Dans la définition précédente des solutions fortes, la condition de continuité de la fonction u permet de donner un sens aux conditions aux bords. Cette condition n'est pas restrictive au sens où toute fonction dans un espace de Sobolev $H^k(I)$ est continue par le théorème d'injection de Sobolev.

Il est possible d'affaiblir encore la notion de solution. Il suffit de revenir à la définition des dérivées faibles à partir de la formule d'intégration par parties. L'équation n'est plus satisfaite ponctuellement, mais uniquement à travers cette formule.

<u>Définition</u>: Le problème de Cauchy précédent admet une solution faible u si et seulement s'il existe une fonction $u \in \mathcal{C}^0([0,1])$ telle que la fonction u admette une dérivée faible sur I et :

$$\text{et} \begin{cases} \forall \varphi \in \mathcal{C}_c^\infty(I), \int_I \left(a u' \varphi' + b u' \varphi + c u \varphi\right) = \int_I f \varphi, \\ u(0) = \alpha \text{ et } u(1) = \beta. \end{cases}$$

Dans cette définition, la condition de continuité de la fonction u est à nouveau introduite afin de donner un sens aux conditions aux bords.

Remarquons qu'une solution classique du problème de Dirichlet précédent est une solution forte de ce problème, et qu'une solution forte de ce problème est une solution faible de ce problème. En général, soit pour des équations aux dérivées partielles arbitraires en dimension quelconque, les réciproques de ces deux affirmations sont fausses. L'avantage de la notion de solution faible réside dans le fait qu'elle permet une construction a priori plus facile des solutions puisque celles-ci satisfont des conditions moins restrictives. Dans le cadre des problèmes de Dirichlet elliptiques qui nous intéressent ici, ces solutions faibles seront (sous des hypothèses ad hoc) des solutions fortes, voire classiques, grâce aux théorèmes de régularité elliptiques que nous présenterons dans la suite.

<u>b. Théorème de Lax-Milgram</u>

La construction des solutions faibles repose sur le théorème de Lax - Milgram

Théorème de Lax - Milgram: Soit H, un espace de Hilbert muni d'un produit scalaire $\langle .,. \rangle_H$. Considérons une forme bilinéaire $a : H \times H \to \mathbb{R}$, et supposons que cette forme bilinéaire est:

(i) continue, c'est-à-dire qu'il existe un nombre positif A tel que:

$\forall (x,y) \in H^2, \ |a(x,y)| \leq A \, \|x\|_H \, \|y\|_H$.

(ii) coercive, c'est-à-dire qu'il existe un nombre strictement positif Λ tel que

$\forall x \in H, \ a(x,x) \geq \Lambda \|x\|_H^2$.

Si φ est une forme linéaire continue sur H, alors, il existe un unique vecteur $x_\varphi \in H$ tel que:

$\forall y \in H, \quad a(x_\varphi, y) = \varphi(y)$.

Preuve:

(i) Supposons qu'il existe deux vecteurs x_φ et $\tilde{x}_\varphi$ de H tels que:

$\forall y \in H, \ a(x_\varphi, y) = \varphi(y) = a(\tilde{x}_\varphi, y)$.

Par bilinéarité, nous avons alors:

$a(x_\varphi - \tilde{x}_\varphi, y) = 0$,

De sorte que:

$\Lambda \|x_\varphi - \tilde{x}_\varphi\|_H^2 \leq a(x_\varphi - \tilde{x}_\varphi, x_\varphi - \tilde{x}_\varphi) = 0$,

Soit $x_\varphi = \tilde{x}_\varphi$. D'où l'unicité du vecteur x_φ si existence.

(ii) Par bilinéarité, pour tout $x \in H$, l'application $y \mapsto a(x,y)$ est une forme linéaire sur H, qui est de plus continue par l'hypothèse (i).
Par le théorème de représentation de Riesz, il existe donc un unique vecteur $A(x) \in H$ tel que:

$\forall y \in H, \ a(x,y) = \langle A(x), y \rangle_H$.

Par bilinéarité de a, l'application A est bien définie et linéaire sur H.
De plus, nous avons:

$\forall x \in H, \ \langle A(x), A(x) \rangle_H = a(x, A(x)) \leq A \|x\|_H \|A(x)\|_H$,

De sorte que A est une application linéaire continue sur H.

(iii) Par coercivité de a, nous avons aussi:

$$\forall x \in H, \quad \lambda \|x\|_H^2 \leq a(x,x) = \langle A(x), x\rangle_H \leq \|A(x)\|_H \, \|x\|_H .$$

En particulier, il vient:

$$\forall x \in H \setminus \{0\}, \quad \|A(x)\|_H \geq \lambda \|x\|_H .$$

L'application A est donc injective, et son image est de plus fermée. En effet, si une suite $(y_n = A(x_n))_{n\in\mathbb{N}}$ satisfait:

$$y_n \underset{n\to+\infty}{\longrightarrow} y \text{ dans } H,$$

alors, par l'inégalité précédente, la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy dans H, de sorte qu'il existe $x \in H$ tel que:

$$x_n \underset{n\to+\infty}{\longrightarrow} x \text{ dans } H.$$

Il s'ensuit que:

$$y_n = A(x_n) \underset{n\to+\infty}{\longrightarrow} A(x) \text{ dans } H,$$

Soit que $y = A(x) \in \mathrm{Im}(A)$. Ce sous-espace est donc fermé dans H.

(iv) Dans ce cas, nous savons que:

$$H = \mathrm{Im}(A) \oplus \mathrm{Im}(A)^{\perp}.$$

Si $\mathrm{Im}(A)^{\perp} \neq \{0\}$, il existe alors un vecteur $x \in H\setminus\{0\}$ tel que:

$$\forall y \in H, \quad \langle A(y), x\rangle_H = 0$$

En particulier, nous obtenons:

$$0 = \langle A(x), x\rangle_H = a(x,x) \geq \lambda \|x\|_H^2 ,$$

De sorte que $x = 0$, ce qui est absurde! Aussi savons-nous que:

$$\mathrm{Im}(A) = H.$$

(v) En conclusion, l'application A est un isomorphisme continu de H sur H. Si φ est une forme linéaire continue sur H, il découle du théorème de représentation de Riesz qu'il existe un vecteur $z \in H$ tel que:

$$\forall y \in H, \quad \varphi(y) = \langle z, y\rangle_H .$$

Soit alors $x_\varphi \in H$ tel que : $z = A(x_\varphi)$. Nous concluons que:

$$\forall y \in H, \quad a(x_\varphi, y) = \langle A(x_\varphi), y\rangle_H = \langle z, y\rangle_H = \varphi(y),$$

D'où l'existence du vecteur x_φ.

Remarque : Dans le cas où la forme bilinéaire a est symétrique c'est-à-dire lorsque :

$$\forall (x,y) \in H^2,\ a(x,y) = a(y,x),$$

cette forme bilinéaire est un produit scalaire sur H, et sa norme associée est équivalente à la norme $\|\cdot\|_H$ grâce aux hypothèses (i) et (ii). En particulier, si ℓ est une forme linéaire continue sur H, elle reste continue pour la norme associée au produit scalaire a. Le théorème de Lax-Milgram est alors une conséquence directe du théorème de représentation de Riesz pour le produit scalaire a.

c. Construction des solutions faibles

Nous considérons dans un premier temps que les conditions aux bords sont égales à :

$\alpha = \beta = 0$,

De sorte que nous étudions le problème de Dirichlet :

$$\text{Et}\begin{cases} \forall x \in I,\ -(a(x)\ u'(x))' + b(x)\ u'(x) + c(x)\ u(x) = f(x), \\ u(0) = u(1) = 0. \end{cases}$$

Définition : La forme bilinéaire B associée à ce problème de Dirichlet est définie par :

$$\forall (u,v) \in H_0^1(I)^2,\ B(u,v) = \int_I (a u' v' + b u' v + c u v).$$

Sous les hypothèses que nous avons annoncées au début de ce paragraphe, cette forme bilinéaire est bien définie et continue sur $H_0^1(I)^2$.

Lemme : Supposons que les fonctions a, b et c sont mesurables et bornées sur I.
La forme bilinéaire B est alors bien définie et continue sur $H_0^1(I)^2$.

Preuve :

Il vient en effet :

$$\int_I |a u' v'| + |b u' v| + |c u v| \leq \max\{\|a\|_{L^\infty(I)}, \|b\|_{L^\infty(I)}, \|c\|_{L^\infty(I)}\}$$

$$\times \int_I \left[(u')^2 + \frac{(v')^2}{2} + \frac{u^2}{2} + v^2 \right],$$

D'où le caractère bien défini et continu de B sur $H_0^1(I)^2$.

Lorsque le problème de Dirichlet est elliptique, la forme bilinéaire B satisfait la minoration suivante.

__Lemme:__ Supposons de plus qu'il existe un nombre strictement positif λ tel que: $pp\, t\, x \in I,\ a(x) \geq \lambda$. Il existe alors un nombre strictement positif γ et un nombre positif β tel que:

$$\forall u \in H_0^1(\Omega),\ B(u,u) \geq \gamma \|u\|_{H^1}^2 - \beta \|u\|_{L^2}^2.$$

Si, de plus, **nous avons**:

$$pp\, t\, x \in I, \begin{cases} b(x) = 0, \\ \text{et } c(x) \geq 0, \end{cases}$$

alors, il est possible de choisir: $\beta = 0$.

__Preuve:__

(i) Rappelons l'inégalité remarquable $2ab \leq a^2 + b^2$ pour $(a,b) \in \mathbb{R}^2$. Il découle de cette inégalité et de l'hypothèse d'ellipticité que:

$$\forall u \in H_0^1(\Omega),\ B(u,u) \geq \lambda \int_I (u')^2 - \frac{\lambda}{2} \int_I (u')^2 - \frac{\|b\|_{L^\infty(I)}^2}{2\lambda} \int_I u^2 - \|c\|_{L^\infty(I)}$$
$$\times \int_I u^2,$$

De sorte que les nombres $\gamma = \frac{\lambda}{2}$ et $\beta = \|c\|_{L^\infty(I)} + \frac{\|b\|_{L^\infty(I)}^2}{2\lambda}$ conviennent.

(ii) Si, de plus,

$$pp\, t\, x \in I, \begin{cases} b(x) = 0, \\ \text{et } c(x) \geq 0, \end{cases}$$

alors, il vient:

$$\forall u \in H_0^2(\Omega),\ B(u,u) \geq \lambda \int_I (u')^2$$

Cependant, l'inégalité de Poincaré assure que:

$$\int_I u^2 \leq \frac{1}{2} \int_I (u')^2,$$

De sorte que:

$$B(u,u) \geq \frac{\lambda}{2} \int_I (u')^2 + \lambda \int_I u^2 \geq \gamma \|u\|_{H^1}^2.$$

Grâce au lemme précédent, nous sommes dans les conditions d'application du théorème de Lax-Milgram dès lors que $\beta = 0$. Ceci conduit au résultat d'existence et d'unicité suivant des solutions faibles.

Théorème : Supposons que les fonctions a, b et c sont mesurables et bornées sur I, et qu'il existe un nombre strictement positif α tel que :
$\text{pp } t\ x \in I,\ a(x) \geq \alpha.$
Il existe un nombre positif β tel que, quelle que soit la fonction $f \in L^2(I)$, il existe une unique solution faible $u \in H^1_0(I)$ du problème de Dirichlet :

$$\begin{cases} \forall x \in I, & -(au')'(x) + b(x)u'(x) + (c(x)+\beta)u(x) = f(x), \\ \text{et} & u(0) = u(1) = 0. \end{cases}$$

Si, de plus, nous avons :

$$\text{pp } t\ x \in I, \begin{cases} b(x) = 0, \\ \text{et } c(x) \geq 0, \end{cases}$$

alors, il est possible de choisir : $\beta = 0$.

Preuve :
Considérons la forme bilinéaire B_β définie par :

$$\forall (u,v) \in H^1_0(I)^2,\quad B_\beta(u,v) = B(u,v) + \beta \int_I uv,$$

où β est le nombre positif du lemme précédent. Par les lemmes précédents, B_β est une forme bilinéaire, continue et coercive sur $H^1_0(I)$.
Soit alors :

$$\forall v \in H^1_0(I),\quad \varphi(v) = \int_I fv.$$

Il résulte de l'inégalité de Cauchy-Schwarz que φ est une forme linéaire continue sur $H^1_0(I)$. Par le théorème de Lax-Milgram, il existe donc une unique fonction $u \in H^1_0(I)$ telle que :

$$\forall v \in H^1_0(I),\quad B_\beta(u,v) = \varphi(v) \iff \int_I (au'v' + bu'v + (c+\beta)uv) = \int_I fv.$$

Par le théorème d'injection de Sobolev, u est continue sur $[0,1]$ et vérifie $u(0) = u(1) = 0$.
En conclusion, comme $C_c^\infty(I)$ est un sous-espace de $H^1_0(I)$, u est une solution

faible du problème de Dirichlet considéré. Sous les conditions :

$$\text{pp } t\ x \in I, \begin{cases} b(x) = 0, \\ \text{et } \ c(x) \geq 0, \end{cases}$$

nous savons par le lemme précédent qu'il est possible de choisir $\beta = 0$.
Supposons alors qu'une fonction $\tilde{u} \in H_0^1(I)$ soit une solution faible du problème de Dirichlet. Par définition, nous avons :

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \quad \int_I \left[a\, \tilde{u}'\, \varphi' + b\, \tilde{u}'\, \varphi + (c+\beta)\, \tilde{u}\, \varphi \right] = \int_I f \varphi.$$

Comme $\tilde{u} \in H_0^1(I)$, nous pouvons combiner la densité du sous-espace $H_0^2(I)$ dans $H_0^1(I)$, et la continuité des applications bilinéaire B_β et linéaire φ afin d'obtenir la formule :

$$\forall v \in H_0^1(I), \quad B_\beta(\tilde{u}, v) = \varphi(v).$$

Il s'ensuit que $\tilde{u} = u$, ce qui établit l'unicité de u.

<u>Exemple</u> : Soit $f \in L^2(I)$. Il existe une unique solution faible $u \in H_0^1(I)$ du problème de Dirichlet :

$$\begin{cases} \forall x \in I, \ -u''(x) = f(x), \\ \text{et } \ u(0) = u(1) = 0. \end{cases}$$

Cet énoncé ne résout pas entièrement la question de l'existence des solutions faibles du fait de la nécessité d'ajouter le paramètre β lorsque les fonctions b et c ne satisfont pas les conditions ad hoc. Nous renvoyons à l'ouvrage de L. C. Evans intitulé "Partial Differential Equations" pour des compléments à ce sujet. Dans la suite, nous nous placerons dans des situations où le théorème précédent fournit des solutions faibles pour $\beta = 0$, par exemple, en supposant que :

$$\text{pp } t\ x \in I, \begin{cases} b(x) = 0, \\ \text{et } \ c(x) \geq 0. \end{cases}$$

Revenons auparavant au problème de Dirichlet général pour lequel les conditions aux bords α et β sont des nombres réels quelconques. Dans ce cas, nous introduisons la fonction affine w définie par :

$\forall x \in [0,1],\ w(x) = \alpha + (\beta - \alpha)\, x.$

Si $u \in \mathcal{C}^0([0,1])$ est une solution faible du problème de Dirichlet :

$$\begin{cases} \forall x \in I, & -(a u')'(x) + b(x)\, u'(x) + c(x)\, u(x) = f(x), \\ \text{et} & u(0) = \alpha \quad \text{et} \quad u(1) = \beta, \end{cases}$$

alors, la fonction $\tilde{u} = u - w$ est continue sur $[0,1]$, avec $\tilde{u}(0) = \tilde{u}(1) = 0$, admet une dérivée faible $\tilde{u}' = u' - \beta + \alpha$ sur I, et vérifie :

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \quad \int_I \left[a\, \tilde{u}' \varphi' + b\, \tilde{u}' \varphi + c\, \tilde{u}\, \varphi\right] = \int_I \left[f + c\, w + b(\beta - \alpha)\right]\varphi + \int_I a(\beta - \alpha)\, \varphi'.$$

Lorsque la fonction a admet une dérivée faible $a' \in L^2(I)$, nous aboutissons à la formule :

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \int_I \left[a\, \tilde{u}' \varphi' + b\, \tilde{u}' \varphi + c\, \tilde{u}\, \varphi\right] = \int_I \tilde{f}\, \varphi,$$

avec $\tilde{f} = f + c\,w + (\beta - \alpha)(b - a') \in L^2(I)$. Nous pouvons donc nous ramener au théorème précédent. En fait, nous pouvons même établir le résultat suivant.

__Corollaire :__ Supposons que les fonctions a, b et c sont mesurables et bornées sur I, et qu'il existe un nombre strictement positif λ tel que :

$$\text{ppt } x \in I, \ a(x) \geq \lambda.$$

Il existe un nombre positif β tel que, quelle que soit la fonction $f \in L^2(I)$, il existe une unique solution faible $u \in H^1(I)$ du problème de Dirichlet :

$$\begin{cases} \forall x \in I, & -(a u')'(x) + b(x)\, u'(x) + (c(x) + \beta)\, u(x) = f(x), \\ \text{et} & u(0) = \alpha \quad \text{et} \quad u(1) = \beta. \end{cases}$$

Si, de plus, nous avons :

$$\text{ppt } x \in I, \begin{cases} b(x) = 0, \\ \text{et } c(x) \geq 0, \end{cases}$$

alors, il est possible de choisir $\beta = 0$.

__Preuve :__

Gardons les notations précédentes et posons :

$$\forall v \in H_0^1(I), \ \varphi(v) = \int_I \left(f + c\,w + (\beta - \alpha)\, b\right) v + \int_I (\beta - \alpha)\, a\, v'.$$

Il résulte de l'inégalité de Cauchy-Schwarz que φ est une forme linéaire continue sur $H_0^1(I)$. Comme B_β est une forme bilinéaire, continue et coercive sur $H_0^1(I)$, par le théorème de Lax - Milgram, il existe une unique fonction $\tilde{u} \in H_0^1(I)$ telle que :

$$\forall v \in H_0^1(I), \quad B_\beta(\tilde{u}, v) = \ell(v) \iff \int_I \left[a\tilde{u}'v' + b\tilde{u}'v + (c+\beta)\tilde{u}v\right]$$
$$= \int_I \left(f + cw + (\beta-\alpha)b\right)v + \int_I (\beta-\alpha)av'$$

Notons $u = \tilde{u} + w$. La fonction u appartient à $H^1(I)$, donc est continue sur $[0,1]$, avec : $u(0) = \alpha$ et $u(1) = \beta$. Comme $\mathcal{C}_c^\infty(I)$ est un sous-espace de $H_0^1(I)$, u est bien une solution faible du problème de Cauchy considéré. L'unicité de cette solution résulte comme dans le théorème précédent de l'unicité de $\tilde{u}$ par le théorème de Lax - Milgram. Le fait de pouvoir choisir $\beta = 0$ provient à nouveau de la coercivité de la forme bilinéaire B lorsque :

$$\text{pp } t\, x \in I, \begin{cases} b(x) = 0, \\ \text{et } c(x) \geq 0. \end{cases}$$

<u>Exemple</u> : Soit $f \in L^2(I)$. Il existe une unique solution faible $u \in H^1(I)$ du problème de Dirichlet :

$$\begin{cases} \forall x \in I, \; -u''(x) = f(x), \\ \text{et } u(0) = \alpha \text{ et } u(1) = \beta. \end{cases}$$

4. Régularité elliptique et solutions classiques

Nous nous restreignons désormais au cas où la fonction b est identiquement nulle sur I, la fonction c est positive sur I, et la fonction a satisfait l'hypothèse d'ellipticité :

$$\text{pp } t\, x \in I, \; a(x) \geq \Lambda,$$

pour un nombre strictement positif Λ. Nous savons alors qu'il existe une unique solution faible $u \in H^1(I)$ du problème de Dirichlet :

$$\begin{cases} \forall x \in I, \; -(au')'(x) + c(x)u(x) = f(x), \\ \text{et } u(0) = \alpha \text{ et } u(1) = \beta, \end{cases}$$

Dès lors que le second membre f appartient à l'espace $L^2(I)$. Nous souhaitons désormais savoir sous quelles conditions ce problème de Dirichlet possède des solutions fortes, ou classiques. Si tel est le cas, nous savons que ces solutions sont des solutions faibles, donc sont égales à u (dès lors que les solutions fortes ou classiques appartiennent à l'espace $H^2(I)$). Le problème de l'existence de solutions fortes ou classiques se ramène donc à établir que la solution faible u possède la régularité suffisante pour être une solution forte ou classique. Les résultats de régularité elliptique qui suivent nous permettent de résoudre cette seconde question.

a. Construction des solutions fortes

Sous les hypothèses précédentes, la solution faible u est en fait une solution forte.

Théorème: Soit $(\alpha,\beta)\in\mathbb{R}^2$. Supposons que les fonctions a et c sont mesurables et bornées sur I, et qu'il existe un nombre strictement positif $\wedge$ tel que:

$$\text{ppt } x\in I, \quad \text{et}\begin{cases} a(x)\geq \wedge, \\ c(x)\geq 0. \end{cases}$$

Quelle que soit la fonction $f\in L^2(I)$, il existe une unique solution forte $u\in H^2(I)$ du problème de Dirichlet:

$$\text{et}\begin{cases} \forall x\in I, \; -(au')'(x) + c(x)u(x) = f(x), \\ u(0)=\alpha \quad \text{et} \quad u(1)=\beta. \end{cases}$$

Preuve:

En cas d'existence, cette solution forte est la solution faible $u\in H^1(I)$ obtenue au paragraphe précédent, d'où son unicité. Montrons maintenant que u est bien une solution forte.

Nous savons déjà que u est continue sur $[0,1]$, avec $u(0)=\alpha$ et $u(1)=\beta$, et que u a une dérivée faible $u'\in L^2(I)$. Comme u est une solution faible, nous avons de plus:

$$\forall \varphi\in \mathcal{C}_c^\infty(I), \quad \int_I a u'\varphi' = \int_I (f-cu)\varphi.$$

La fonction au' a donc une dérivée faible égale à $-f+cu \in L^2(I)$, et nous avons de plus l'identité :

$pp\, t\, x \in I,\ -(a u')'(x) + c(x)\, u(x) = f(x),$

De sorte que u est bien une solution forte de ce problème de Dirichlet.

Lorsque la fonction a possède une dérivée faible a' mesurable et bornée sur I, nous pouvons établir que la solution forte u appartient à l'espace $H^2(I)$.

<u>Corollaire:</u> Supposons de plus que la fonction a possède une dérivée faible a' mesurable et bornée sur I. Alors la solution forte u du théorème précédent appartient à l'espace $H^2(I)$ et elle vérifie:

$pp\, t\, x \in I,\ -a(x)\, u''(x) - a'(x)\, u'(x) + c(x)\, u(x) = f(x).$

<u>Preuve</u>:

Comme les fonctions a et a' sont bornées sur I, la fonction a appartient à l'espace $H^1(I)$. En particulier, il existe une suite de fonctions $(a_n)_{n\in\mathbb{N}}$ de $\mathcal{C}^\infty(\mathbb{R}) \cap H^1(I)$ telle que

$a_n \underset{n\to+\infty}{\longrightarrow} a$ dans $H^1(I)$.

Par le théorème d'injection de Sobolev, nous en déduisons que:

$a_n \underset{n\to+\infty}{\longrightarrow} a$ unif. sur $[0,1]$,

De sorte que, par l'hypothèse d'ellipticité, nous pouvons supposer que:

$\forall n \in \mathbb{N},\ \forall x \in [0,1],\ a_n(x) \geqslant \frac{\Lambda}{2} > 0.$

Considérons alors une fonction $\varphi \in \mathcal{C}_c^\infty(I)$. Les fonctions $\frac{\varphi}{a_n}$ sont bien définies sur I et appartiennent elles-aussi à $\mathcal{C}_c^\infty(I)$, donc à $H_0^1(I)$ avec:

$$\left(\frac{\varphi}{a_n}\right)' = \frac{\varphi'}{a_n} - \frac{a'_n \varphi}{a_n^2}.$$

Les propriétés de convergence précédentes assurent que:

$$\text{Et}\begin{cases} \dfrac{\varphi}{a_n} \underset{n\to+\infty}{\longrightarrow} \dfrac{\varphi}{a} \text{ dans } L^2(I), \\ \dfrac{\varphi'}{a_n} \underset{n\to+\infty}{\longrightarrow} \dfrac{\varphi'}{a} \text{ dans } L^2(I), \\ -\dfrac{a'_n \varphi}{a_n^2} \underset{n\to+\infty}{\longrightarrow} -\dfrac{a'\varphi}{a^2} \text{ dans } L^2(I), \end{cases}$$

De sorte que la suite $\left(\frac{\varphi}{a_n}\right)_{n\in\mathbb{N}}$ est convergente dans $H^1(I)$ de limite égale à la fonction $\frac{\varphi}{a}$. Cette fonction appartient donc à $H_0^1(I)$ et elle vérifie :

$$\left(\frac{\varphi}{a}\right)' = \frac{\varphi'}{a} - \frac{a'}{a^2}\varphi.$$

Comme u est une solution forte du problème de Dirichlet, c'est aussi une solution faible qui satisfait donc l'identité :

$$\forall \theta \in \mathcal{C}_c^\infty(I), \quad \int_I a u' \theta' = \int_I (f - cu)\,\theta.$$

Par densité du sous-espace $\mathcal{C}_c^\infty(I)$ dans $H_0^1(I)$, cette identité reste vraie si $\theta \in H_0^1(I)$. Il vient donc :

$$\int_I a\, u' \left(\frac{\varphi}{a}\right)' = \int_I (f - cu)\frac{\varphi}{a},$$

Soit :

$$\int_I u'\varphi' = \int_I \left[\frac{a'u'}{a} + \frac{f - cu}{a}\right]\varphi.$$

En conclusion, la fonction u' possède une dérivée faible $u'' = -\frac{1}{a}[a' u' - cu + f]$, laquelle appartient à l'espace $L^2(I)$. La fonction u est donc dans $H^2(I)$, et elle satisfait de plus :

$$\text{pp } x \in I, \quad -a(x)u''(x) - a(x)\,u'(x) + c(x)u(x) = f(x).$$

<u>Exemple :</u> Soit $f \in L^2(I)$. Il existe une unique solution forte $u \in H^2(I)$ du problème de Dirichlet :

$$\begin{cases} \forall x \in I, \quad -u''(x) = f(x), \\ \text{Et} \quad u(0) = \alpha \ \text{ et } \ u(1) = \beta. \end{cases}$$

<u>Remarques :</u> (i) La notion de solution forte correspond parfois au cadre fonctionnel du corollaire précédent, soit au fait que les dérivées u, u' et u'' ont un sens presque partout, et que l'équation du second ordre est satisfaite presque partout par ces dérivées.

(ii) Par le théorème d'injection de Sobolev, sous les hypothèses du corollaire précédent, la solution forte u est de classe $\mathcal{C}^1$ sur le segment $[0,1]$. Ce n'est pas néanmoins une solution classique, puisque la dérivée u'' appartient certes à $L^2(I)$, mais n'est pas continue sur I.

b. Construction des solutions classiques

Sous les hypothèses du corollaire précédent, la dérivée u'' de la solution forte satisfait:

$$\text{pp } t\ x \in I,\quad u''(x) = \frac{1}{a(x)}\left[-a'(x)\,u'(x) + c(x)\,u(x) - f(x)\right],$$

puisque la fonction a n'est jamais nulle par l'hypothèse d'ellipticité. Afin que cette fonction puisse être continue, et que la solution u soit classique, il est nécessaire d'ajouter des hypothèses de continuité pour les fonctions a, c et f.

__Théorème:__ Soit $(\alpha, \beta) \in \mathbb{R}^2$. Supposons que les fonctions a et c sont continues et bornées sur I, et qu'il existe un nombre strictement positif λ tel que:

$$\forall x \in I, \begin{cases} a(x) \geq \lambda, \\ \text{et } c(x) \geq 0. \end{cases}$$

Quelle que soit la fonction f continue sur I, avec $f \in L^2(I)$, il existe une unique solution classique $u \in \mathcal{C}^1(I)$, avec u' bornée sur I, du problème de Dirichlet:

$$\begin{cases} \forall x \in I, \quad -(a u')'(x) + c(x)\,u(x) = f(x), \\ \text{et } u(0) = \alpha \text{ et } u(1) = \beta. \end{cases}$$

__Preuve:__

En cas d'existence, cette solution classique appartient à l'espace $H^1(I)$, de sorte qu'il s'agit de la solution forte u du paragraphe précédent, d'où son unicité. Montrons maintenant que u est bien une solution classique.

Nous savons déjà que u est continue sur $[0,1]$, avec $u(0) = \alpha$ et $u(1) = \beta$, que u a une dérivée faible $u' \in L^2(I)$, et que la fonction $a u'$ a une dérivée faible donnée par:

$$(a u')' = -f + c u.$$

Comme a est bornée sur I et que la fonction $cu - f$ appartient à $L^2(I)$, nous en déduisons que $a u' \in H^1(I)$. Par le théorème d'injection de Sobolev, la fonction $a u'$ se prolonge donc en une fonction continue sur $[0,1]$. Par l'hypothèse d'ellipticité, la fonction u est de classe $\mathcal{C}^1$ sur I, avec u' bornée sur I. La formule $(au')' = cu - f$ assure alors que la fonction au' est de classe $\mathcal{C}^1$ sur I, et que la fonction u est une solution classique du problème de Dirichlet.

Lorsque la fonction a est de classe $\mathcal{C}^1$ sur I, nous pouvons établir que la solution classique u est de classe $\mathcal{C}^2$ sur I.

<u>Corollaire</u> : Supposons de plus que la fonction a est de classe $\mathcal{C}^1$ sur I. Alors la solution classique u du théorème précédent est de classe $\mathcal{C}^2$ sur I, et elle vérifie :

$$\forall x \in I, \ -a(x)\, u''(x) - a'(x)\, u'(x) + c(x)\, u(x) = f(x).$$

<u>Preuve</u> :

Comme $u' = \frac{au'}{a}$, par l'hypothèse d'ellipticité, la fonction u' est de classe $\mathcal{C}^1$ sur I, et :

$$\forall x \in I, \quad u''(x) = -\frac{a'(x)\, u'(x)}{a(x)} + \frac{1}{a(x)} (au')'(x).$$

La fonction u est donc de classe $\mathcal{C}^2$ sur I, et elle satisfait l'équation :

$$\forall x \in I, \quad u''(x) = \frac{1}{a(x)} \left[-a'(x)\, u'(x) + c(x)\, u(x) - f(x) \right].$$

<u>Exemple</u> : Soit f une fonction continue sur I, telle que $f \in L^2(I)$. Il existe une unique solution classique $u \in \mathcal{C}^2(I)$, avec u' bornée sur I, du problème de Dirichlet :

$$\begin{cases} \forall x \in I, \ -u''(x) = f(x) \\ \text{et} \quad u(0) = \alpha \ \text{ et } \ u(1) = \beta. \end{cases}$$

<u>Remarque</u> : La notion de solution classique correspond parfois au cadre fonctionnelle du corollaire précédent, soit au fait que u est de classe $\mathcal{C}^2$ sur I, et que l'équation du second ordre est satisfaite partout.

c. <u>Propriétés qualitatives des solutions classiques</u>

Commençons par les propriétés de régularité. Sous les hypothèses du corollaire précédent, nous avons :

$\forall x \in I,\ u''(x) = \frac{1}{a(x)}\left[-a'(x)\,u'(x) + c(x)\,u(x) - f(x)\right],$

et nous savons que les fonctions u et u' sont de classe $\mathcal{C}^1$ sur I. Lorsque a est de classe $\mathcal{C}^2$ sur I, et que c et f sont de classe $\mathcal{C}^1$ sur I, nous en déduisons que u'' est de classe $\mathcal{C}^1$ sur I, soit que u est de classe $\mathcal{C}^3$ sur I. Cet argument est un exemple de régularisation elliptique, qui conduit de même à l'énoncé suivant.

Théorème : Soit $(\alpha, \beta) \in \mathbb{R}^2$ et $k \in \mathbb{N}$. Supposons que la fonction a est bornée et de classe $\mathcal{C}^{k+1}$ sur I, que la fonction c est bornée et de classe $\mathcal{C}^k$ sur I, et qu'il existe un nombre strictement positif λ tel que :

$$\forall x \in I, \begin{cases} a(x) \geqslant \lambda, \\ \text{et } c(x) \geqslant 0. \end{cases}$$

Quelle que soit la fonction f de classe $\mathcal{C}^k$ sur I, avec $f \in L^2(I)$, la solution classique u du problème de Dirichlet :

$$\begin{cases} \forall x \in I,\ -(au')'(x) + c(x)\,u(x) = f(x), \\ \text{et } u(0) = \alpha \quad \text{et } u(1) = \beta, \end{cases}$$

donnée par le paragraphe précédent, est de classe $\mathcal{C}^{k+2}$ sur I.

De plus, ce résultat demeure valable lorsque k est remplacé par ∞.

Preuve :

La preuve est par récurrence sur $k \in \mathbb{N}$. Au rang 0, c'est exactement le corollaire précédent, et nous avons de plus :

$\forall x \in I,\ u''(x) = \frac{1}{a(x)}\left(-a'(x)\,u'(x) + c(x)\,u(x) - f(x)\right).$

Supposons le résultat vrai jusqu'au rang k. Au rang $k+1$, nous savons donc que u est de classe $\mathcal{C}^{k+2}$ sur I, de sorte que par l'équation précédente, u'' est de classe $\mathcal{C}^{k+1}$ sur I, et donc, que u est de classe $\mathcal{C}^{k+3}$ sur I, ce qui achève la récurrence.

Dans le cas où k est remplacé par ∞, la solution u est de classe $\mathcal{C}^{l+2}$ pour tout entier l, donc de classe $\mathcal{C}^\infty$

<u>Exemple:</u> Soit $f \in \mathcal{C}^\infty(I)$, telle que $f \in L^2(I)$. L'unique solution classique u, avec u' bornée sur I, du problème de Dirichlet :

$$\text{Et}\begin{cases} \forall x \in I, \ -u''(x) = f(x), \\ u(0)=\alpha \text{ et } u(1) = \beta, \end{cases}$$

est de classe $\mathcal{C}^\infty$ sur I.

Les solutions classiques satisfont de plus le principe du maximum. Lorsque la fonction c est identiquement nulle sur I, la version faible de ce principe s'exprime sous la forme suivante.

<u>Théorème:</u> Soit $(\alpha, \beta) \in \mathbb{R}^2$. Supposons que la fonction a est de classe $\mathcal{C}^1$ sur $[0, 1]$, et qu'il existe un nombre strictement positif Λ tel que:

$\forall x \in I, \ a(x) \geq \Lambda.$

Étant donnée une fonction f continue sur I, considérons une solution $u \in \mathcal{C}^2(I) \cap \mathcal{C}^0([0,1])$ du problème de Dirichlet:

$$\text{Et}\begin{cases} \forall x \in I, \ -\left(a(x)\, u'(x)\right)' = f(x), \\ u(0) = \alpha \text{ et } u(1) = \beta. \end{cases}$$

Si la fonction f est négative sur I, alors :

$$\max_{x \in [0,1]} u(x) = \max \{\alpha, \beta\}.$$

<u>Preuve:</u>

Soit $(\varepsilon, \lambda) \in]0, +\infty[^2$. Posons :

$\forall x \in [0,1], \ v_\varepsilon(x) = u(x) + \varepsilon e^{\lambda x}.$

La fonction v_ε est continue sur $[0,1]$, de classe $\mathcal{C}^2$ sur I, et elle satisfait :

$\forall x \in I, \ -\left(a(x)\, v'_\varepsilon(x)\right)' = f(x) - \varepsilon\left[a(x)\lambda^2 + a'(x)\lambda\right] e^{\lambda x}.$

Comme a' est continue sur $[0,1]$, elle est bornée par un nombre positif Π sur cet ensemble, de sorte que, par la condition d'ellipticité,

$\forall x \in I, \ a(x)\lambda^2 + a'(x)\lambda \geq \Lambda\lambda^2 - \Pi\lambda = \lambda(\Lambda\lambda - \Pi).$

Choisissons alors : $\lambda = \frac{\Pi + 1}{\Lambda}$. Nous obtenons :

$\forall x \in I, \ a(x)\lambda^2 + a'(x)\lambda \geq \frac{\Pi+1}{\Lambda} \Rightarrow -\left(a(x)\, v'_\varepsilon(x)\right)' \leq -\frac{\Pi+1}{\Lambda}\varepsilon < 0.$

Considérons alors un point $x_\varepsilon \in [0,1]$ tel que la fonction v_ε, qui est continue sur $[0,1]$, admette son maximum en x_ε. Si x_ε appartient à I, alors :

$$\text{Et}\begin{cases} v'_\varepsilon(x_\varepsilon) = 0, \\ v''_\varepsilon(x_\varepsilon) \leq 0, \end{cases}$$

De sorte que:

$$-(a(x_\varepsilon) v'_\varepsilon(x_\varepsilon))' = -a(x_\varepsilon) v''_\varepsilon(x_\varepsilon) - a'(x_\varepsilon) v'(x_\varepsilon) \geq 0,$$

Ce qui est absurde ! Nous en déduisons que $x_\varepsilon \in \{0, 1\}$, ce qui induit que :

$$\underset{x \in [0,1]}{\text{Max}}\ v_\varepsilon(x) = \text{Max}\{\alpha + \varepsilon, \beta + \varepsilon e^\Lambda\}.$$

En particulier, nous avons :

$$\forall x \in [0,1],\ u(x) \leq \underbrace{\text{Max}\{\alpha+\varepsilon, \beta+\varepsilon e^\Lambda\} - \varepsilon e^{\lambda x}}_{\underset{\varepsilon \to 0}{\longrightarrow} \text{Max}\{\alpha, \beta\}}.$$

Il s'ensuit que :

$$\forall x \in [0,1],\ u(x) \leq \text{Max}\{\alpha, \beta\},$$

Et la conclusion en découle.

Nous pouvons étendre ce résultat au cas où la fonction c est positive, et obtenir le résultat suivant.

__Théorème__ : Soit $(\alpha, \beta) \in \mathbb{R}^2$. Supposons que la fonction a est de classe $\mathcal{C}^1$ sur $[0,1]$, la fonction c est continue sur $[0,1]$, et qu'il existe un nombre strictement positif Λ tel que:

$$\forall x \in I, \text{Et}\begin{cases} a(x) \geq \Lambda, \\ c(x) \geq 0. \end{cases}$$

Etant donnés une fonction f continue sur I, considérons une solution $u \in \mathcal{C}^2(I) \cap \mathcal{C}^0([0,1])$ du problème de Dirichlet :

$$\text{Et}\begin{cases} \forall x \in I, -(a(x) u'(x))' + c(x) u(x) = f(x), \\ u(0) = \alpha \text{ et } u(1) = \beta. \end{cases}$$

Si la fonction f est négative sur I, alors :

$$\underset{x \in [0,1]}{\text{Max}}\ u(x) \leq \text{Max}\{0, \alpha, \beta\}.$$

__Preuve__:

Nous reprenons les notations de la preuve du théorème précédent. Pour tout nombre strictement positif ε, la fonction v_ε définie par:

$\forall x \in [0,1], \; v_\varepsilon(x) = u(x) + \varepsilon e^{\lambda x}$,

est continue sur $[0,1]$, de classe C^2 sur I, et satisfait:

$\forall x \in I, \; -(a(x)\, v'_\varepsilon(x))' + c(x)\, v_\varepsilon(x) = f(x) + \varepsilon\left[-a(x)\lambda^2 - a'(x)\lambda + c(x)\right] e^{\lambda x}$

Comme les fonctions a' et c sont continues sur $[0,1]$, elles sont bornées par un nombre positif M sur I, de sorte que, par la condition d'ellipticité,

$\forall x \in I, \; a(x)\lambda^2 + a'(x)\lambda - c(x) \geq \alpha\lambda^2 - M(\lambda+1)$.

Choisissons alors λ suffisamment grand pour que: $\alpha\lambda^2 - M(\lambda+1) \geq 1$.

Nous obtenons:

$\forall x \in I, \; -(a(x)\, v'_\varepsilon(x))' + c(x)\, v_\varepsilon(x) \leq -\varepsilon < 0$.

Considérons alors un point $x_\varepsilon \in [0,1]$ tel que:

$\max\limits_{x \in [0,1]} v_\varepsilon(x) = v_\varepsilon(x_\varepsilon)$.

Si $v_\varepsilon(x_\varepsilon) > 0$ et si $x_\varepsilon \in I$, alors:

$\begin{cases} v'_\varepsilon(x_\varepsilon) = 0, \\ \text{et } v''_\varepsilon(x_\varepsilon) \leq 0, \end{cases}$

De sorte que:

$-(a(x_\varepsilon)\, v'_\varepsilon(x_\varepsilon))' + c(x_\varepsilon)\, v_\varepsilon(x_\varepsilon) \geq 0$,

Ce qui est absurde ! Il s'ensuit que $v_\varepsilon(x_\varepsilon) \leq 0$ ou $x_\varepsilon \in \{0,1\}$.

Dans ces deux cas, nous avons:

$\max\limits_{x \in [0,1]} v_\varepsilon(x) \leq \max\{0, \alpha+\varepsilon, \beta+\varepsilon e^{\lambda}\}$,

et la conclusion s'ensuit comme dans la preuve du théorème précédent.

Une conséquence intéressante du principe du maximum faible précédent est donnée par l'énoncé suivant.

Corollaire: Sous les hypothèses du théorème précédent, si la fonction f est positive sur I, et si $(\alpha, \beta) \in \mathbb{R}_+^2$, alors:

$\forall x \in I, \; u(x) \geq 0$.

Preuve:

Nous appliquons le théorème précédent à la fonction $-u$, qui satisfait l'équation :

$$\forall x \in I, \quad -\left(a(x)(-u(x))\right)' + c(x)(-u(x)) = -f(x) \leq 0.$$

Nous obtenons :

$$\underset{x \in [0,1]}{\mathrm{Max}} \left[-u(x)\right] \leq \mathrm{Max}\,\{0, -\alpha, -\beta\} = 0,$$

De sorte que :

$$\forall x \in [0,1], \; u(x) \geq 0.$$

<u>Exemple</u> : Soit $(\alpha, \beta) \in \mathbb{R}_+^2$ et f, une fonction continue et positive sur $[0,1]$. L'unique solution classique u, avec u' bornée sur I, du problème de Dirichlet :

$$(E)\begin{cases} \forall x \in I, \; -u''(x) = f(x), \\ u(0) = \alpha \text{ et } u(1) = \beta, \end{cases}$$

est positive sur I.

Nous ne détaillerons pas plus les différentes versions du principe du maximum pour les problèmes de Dirichlet du type précédent. Nous renvoyons à l'ouvrage de L. C. Evans intitulé "Partial Differential Equations" pour de plus amples détails à ce sujet (en particulier, quant au principe du maximum fort).

5. <u>Méthode des différences finies pour les équations elliptiques du second ordre</u> :

Nous nous intéressons désormais à la résolution approchée des équations elliptiques du second ordre par la méthode des différences finies. Nous nous restreignons au cas de l'équation de Schrödinger :

$$\forall x \in I, \quad -u''(x) + c(x)u(x) = f(x),$$

pour des fonctions c et f définies et continues sur $[0,1]$, et dans le cas où la fonction c est positive. Nous avons établi au paragraphe précédent que le problème de Dirichlet associé possède une unique solution u de classe $\mathcal{C}^2$ sur I (avec u' bornée sur I) pour tout couple de données aux bords $(u(0), u(1)) = (\alpha, \beta) \in \mathbb{R}^2$, et nous cherchons à la calculer de manière approchée.

1. Principe des méthodes de différences finies

Les méthodes de différences finies reposent sur un processus de discrétisation. Étant donné un entier N, nous introduisons la subdivision du segment $[0,1]$

$$x_0 = 0 < x_1 < \dots < x_N < x_{N+1} = 1,$$

de pas constant $h = \frac{1}{N+1}$. Les nombres $(x_i)_{0 \leq i \leq N+1}$ sont donc donnés par les formules:

$$\forall\, 0 \leq i \leq N+1, \quad x_i = \frac{i}{N+1} = hi.$$

Notre objectif est de calculer des valeurs approchées $(u_i)_{0 \leq i \leq N+1}$ de la solution $(u(x_i))_{0 \leq i \leq N+1}$ aux points $(x_i)_{0 \leq i \leq N+1}$. Dans ce but, nous devons calculer une valeur approchée des dérivées secondes $(u''(x_i))_{1 \leq i \leq N}$.

Comme la fonction u est de classe $\mathcal{C}^2$ sur I, nous savons que :

$$\forall\, 1 \leq i \leq N, \quad \begin{cases} u(x_{i+1}) = u(x_i) + h\, u'(x_i) + \frac{h^2}{2} u''(x_i) + \underset{h \to 0}{o}(h^2), \\ \text{Et } u(x_{i-1}) = u(x_i) - h\, u'(x_i) + \frac{h^2}{2} u''(x_i) + \underset{h \to 0}{o}(h^2), \end{cases}$$

lorsque le pas h est assez petit. Il s'ensuit que:

$$\forall\, 1 \leq i \leq N, \quad u''(x_i) = \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} + \underset{h \to 0}{o}(1),$$

Nous aboutissons ainsi à un calcul approché des dérivées secondes $u''(x_i)$ par les valeurs:

$$u''(x_i) \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}.$$

Nous en déduisons un schéma numérique pour la résolution de ce problème :

$$\text{Et} \begin{cases} u_0 = \alpha, \\ \forall\, 1 \leq i \leq N, \quad -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c(x_i)\, u_i = f(x_i), \\ u_{N+1} = \beta. \end{cases}$$

Il s'agit du schéma aux différences finies pour la résolution du problème de Dirichlet

$$\text{Et} \begin{cases} \forall x \in I, \quad -u''(x) + c(x)\, u(x) = f(x), \\ u(0) = \alpha \text{ et } u(1) = \beta. \end{cases}$$

Cette terminologie provient du fait que nous avons remplacé une dérivée (seconde en l'occurence) par un quotient aux différences finies.

En effet, lorsque le pas h est assez petit, nous pouvons approcher la dérivée $u'(x_i)$ par l'un des quotients aux différences finies suivants :

(i) le quotient décentré à gauche :

$$u'(x_i) \approx \frac{u(x_i) - u(x_{i-1})}{h},$$

(ii) le quotient centré :

$$u'(x_i) \approx \frac{u(x_{i+1}) - u(x_{i-1})}{2h},$$

(iii) le quotient décentré à droite :

$$u'(x_i) \approx \frac{u(x_{i+1}) - u(x_i)}{h}.$$

Le calcul approché précédent de la dérivée seconde $u''(x_i)$ s'interprète comme une itération des formules précédentes :

$$u''(x_i) \approx \frac{u'(x_{i+1}) - u'(x_i)}{h} \approx \frac{1}{h}\left(\frac{u(x_{i+1}) - u(x_i)}{h} - \frac{u(x_i) - u(x_{i-1})}{h}\right).$$

Ce choix de combiner quotients décentrés à gauche et à droite est le seul qui conduise à une expression approchée de $u''(x_i)$, qui ne dépende que des valeurs de u aux points x_{i-1}, x_i et x_{i+1}. Cette propriété rend l'usage de cette expression plus simple que pour les autres choix possibles.

b. Définition du schéma aux différences finies

Le schéma aux différences finies pour la résolution du problème de Dirichlet

$$\left\{\begin{array}{l} \forall x \in I,\ -u''(x) + c(x)\,u(x) = f(x), \\ u(0) = \alpha \quad \text{et} \quad u(1) = \beta, \end{array}\right.$$

conduit au calcul des valeurs approchées $(u_i)_{0 \leq i \leq N+1}$ de la solution u aux points $(x_i = hi)_{0 \leq i \leq N+1}$ par les formules :

$$\left\{\begin{array}{l} u_0 = \alpha, \\ \forall\, 1 \leq i \leq N,\ -\dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c(x_i)\,u_i = f(x_i), \\ u_{N+1} = \beta. \end{array}\right.$$

Ce schéma s'écrit sous la forme matricielle :

$$A_c\,(U) = b,$$

où la matrice $A_c \in \mathcal{M}_N(\mathbb{R})$ est définie par :

$$A_c = \frac{1}{h^2}\begin{pmatrix} 2+c(x_1)h^2 & -1 & & 0 \\ -1 & 2+c(x_2)h^2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & \ddots & -1 \\ 0 & & -1 & 2+c(x_N)h^2 \end{pmatrix},$$

et le vecteur $b \in \mathbb{R}^N$ est donné par :

$$b = \begin{pmatrix} f(x_1) + \frac{\alpha}{h^2} \\ f(x_2) \\ \vdots \\ f(x_{N-1}) \\ f(x_N) + \frac{\beta}{h^2} \end{pmatrix}.$$

La mise en place pratique de ce schéma requiert de résoudre le système $A_c(U)=b$ d'inconnue $U = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} \in \mathbb{R}^N$, ce qui nécessite d'établir l'inversibilité de la matrice A_c.

__Lemme :__ Soit $N \in \mathbb{N}$ et $h = \frac{1}{N+1}$. Supposons que la fonction c est positive sur I. Alors, la matrice A_c est symétrique, définie, positive, donc inversible. En particulier, le schéma aux différences finies est bien défini.

__Preuve :__

La matrice A_c est symétrique, et elle satisfait :

$$\forall\, V = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \in \mathbb{R}^N, \quad {}^t v\, A_c\, v = \sum_{i=1}^{N} c(x_i)\, v_i^2 + \frac{1}{h^2}\left(v_1^2 + \sum_{i=2}^{N} (v_i - v_{i-1})^2 + v_N^2\right).$$

Cette quantité est positive, et nulle si et seulement si $V=0$. La matrice A_c est donc définie positive, ce qui suffit à établir son caractère inversible.

__Exemple :__ La matrice $A_0 = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix}$ est symétrique définie positive, donc inversible.

En pratique, la résolution du système $A_c(U)=b$ passe par la mise en oeuvre d'une méthode de résolution approchée des systèmes linéaires. Comme la matrice A_c est tridiagonale, le calcul d'une factorisation de type LU (ou plutôt de type Cholesky, puisque la matrice A_c est aussi symétrique définie positive) se révèle des plus rapides, et conduit à la résolution du système $A_c(U)=b$ en de l'ordre de N opérations. Nous renvoyons au chapitre sur la résolution des systèmes linéaires pour de plus amples détails à ce sujet.

Nous supposons désormais que nous savons résoudre exactement le système $A_c(V)=b$ et nous nous intéressons à la convergence des valeurs approchées ainsi calculées vers la solution du problème de Dirichlet. Comme dans le cas des méthodes à un pas pour les équations différentielles ordinaires, la preuve de cette convergence repose sur les notions de consistance et de stabilité.

c. Consistance et ordre du schéma aux différences finies

<u>Définition</u>: Soit $N \in \mathbb{N}$ et $1 \leq i \leq N$. Considérons une solution $u \in \mathcal{C}^2(I) \cap \mathcal{C}^0([0,1])$ du problème de Dirichlet:

$$\text{(E)}\begin{cases} \forall x \in I, \ -u''(x) + c(x)\, u(x) = f(x), \\ u(0) = \alpha \ \text{ et } \ u(1) = \beta. \end{cases}$$

L'erreur de consistance e_i au point $x_i = \frac{i}{N+1}$ relative à la solution u est définie par:

$$e_i = -\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} + c(x_i)\, u(x_i) - f(x_i).$$

Comme dans le cas des équations différentielles ordinaires, l'erreur de consistance mesure l'erreur qui est faite lorsque l'équation différentielle du second ordre **est** remplacée par le schéma numérique. Cette erreur n'est pas égale à l'erreur finale entre la valeur exacte $u(x_i)$ au point x_i, et la valeur approchée u_i, laquelle est donnée par la différence $u(x_i) - u_i$. Cependant, il semble clair que l'erreur finale sera au moins de l'ordre de $\max\limits_{1 \leq i \leq N} |e_i|$, quantité qu'il est donc important de pouvoir contrôler.

<u>Définition</u>: Le schéma aux différences finies est consistant si et seulement si, quelle que soit la solution $u \in \mathcal{C}^2(I) \cap \mathcal{C}^0([0,1])$ du problème de Dirichlet:

$$\text{(E)}\begin{cases} \forall x \in I, \ -u''(x) + c(x)\, u(x) = f(x), \\ u(0) = \alpha \ \text{ et } \ u(1) = \beta, \end{cases}$$

les erreurs de consistance $(e_i)_{1 \leq i \leq N}$ aux points $(x_i)_{1 \leq i \leq N}$ relatives à la solution u satisfont:

$$\max_{1 \leq i \leq N} |e_i| \xrightarrow[N \to +\infty]{} 0.$$

Lorsque les fonctions c et f sont continues sur $[0,1]$, le schéma aux différences finies est consistant.

<u>Lemme</u>: Supposons que les fonctions c et f sont continues sur $[0,1]$. Alors le schéma aux différences finies est consistant.

<u>Preuve</u>:

Si $u \in \mathcal{C}^2(I) \cap \mathcal{C}^0([0,1])$ est une solution du problème de Dirichlet, alors, u se prolonge en une fonction de classe $\mathcal{C}^2$ sur $[0,1]$, de sorte que la dérivée seconde est uniformément continue sur $[0,1]$.

La formule de Taylor-Lagrange assure de plus l'existence de nombres $\alpha_i \in [x_i, x_{i+1}]$ et $\beta_i \in [x_{i-1}, x_i]$ tels que :

$$\begin{cases} u(x_{i+1}) = u(x_i) + h\,u'(x_i) + \frac{h^2}{2} u''(\alpha_i), \\ u(x_{i-1}) = u(x_i) - h\,u'(x_i) + \frac{h^2}{2} u''(\beta_i). \end{cases}$$

Il vient donc :

$$\begin{aligned} e_i &= -\frac{1}{2}\left[u''(\alpha_i) + u''(\beta_i)\right] + c(x_i)\,u(x_i) - f(x_i) \\ &= \frac{1}{2}\left[u''(x_i) - u''(\alpha_i) - u''(\beta_i) + u''(x_i)\right], \end{aligned}$$

De sorte que, par uniforme continuité de u'' sur $[0,1]$,

$$\max_{1 \leq i \leq N} |e_i| \xrightarrow[N \to +\infty]{} 0.$$

La notion de consistance ne donne aucune information sur la taille des erreurs de consistance. Afin de préciser cette taille, il est nécessaire de recourir à la notion d'ordre du schéma numérique.

<u>Définition</u>: Soit $p \in \mathbb{N}^*$. Le schéma aux différences finies est d'ordre (au moins égal à) p si et seulement si, quelle que soit la solution $u \in \mathcal{C}^{p+2}(I) \cap \mathcal{C}^p([0,1])$ du problème de Dirichlet,

$$\begin{cases} \forall x \in I, \ -u''(x) + c(x)\,u(x) = f(x), \\ u(0) = \alpha \text{ et } u(1) = \beta, \end{cases}$$

il existe un nombre positif C tel que les erreurs de consistance

$(e_i)_{1\leq i\leq N}$ aux points $(x_i)_{1\leq i\leq N}$ relatives à la solution u satisfont:
$$\max_{1\leq i\leq N} |e_i| \leq C h^p.$$

Remarque: Le schéma est (parfois) dit d'ordre exactement égal à p s'il est d'ordre (au moins égal à) p, mais pas d'ordre (au moins égal à) $p+1$. Les deux définitions précédentes cohabitent dans la littérature.

Lemme: Supposons que les fonctions c et f sont de classe $\mathcal{C}^2$ sur $[0,1]$. Alors le schéma aux différences finies est d'ordre (au moins égal à) 2.

Preuve:
Si $u \in \mathcal{C}^4(I) \cap \mathcal{C}^2([0,1])$ est une solution du problème de Dirichlet, alors, u se prolonge en une fonction de classe $\mathcal{C}^4$ sur $[0,1]$. Par la formule de Taylor-Lagrange, il existe alors des nombres $\alpha_i \in (x_i, x_{i+1})$ et $\beta_i \in (x_{i-1}, x_i)$ tels que:
$$\begin{cases} u(x_{i+1}) = u(x_i) + h\, u'(x_i) + \frac{h^2}{2} u''(x_i) + \frac{h^3}{6} u'''(x_i) + \frac{h^4}{24} u^{(4)}(\alpha_i), \\ \text{et } u(x_{i-1}) = u(x_i) - h\, u'(x_i) + \frac{h^2}{2} u''(x_i) - \frac{h^3}{6} u'''(x_i) + \frac{h^4}{24} u^{(4)}(\beta_i), \end{cases}$$
De sorte que:
$$e_i = -\frac{h^2}{24}\left(u^{(4)}(\alpha_i) + u^{(4)}(\beta_i)\right).$$
Comme la fonction $u^{(4)}$ est bornée sur $[0,1]$, nous concluons que:
$$\max_{1\leq i\leq N} |e_i| \leq \frac{1}{12} \max_{x\in[0,1]} |u^{(4)}(x)|\, h^2,$$
De sorte que le schéma est d'ordre (au moins égal à) 2.

Il est ici utile de remarquer que la précision du schéma aux différences finies dépend non seulement de son ordre, mais également de la régularité de la solution u que l'on cherche à approcher. Par exemple, si u n'est que de classe $\mathcal{C}^3$ sur $[0,1]$, les calculs précédents établissent que l'erreur de consistance est contrôlée par:
$$\max_{1\leq i\leq N} |e_i| \leq \frac{1}{3} \max_{x\in[0,1]} |u^{(3)}(x)|\, h.$$
La précision obtenue lorsque h tend vers 0 est ainsi bien moins bonne que pour une solution de classe $\mathcal{C}^4$ sur $[0,1]$.

2. Stabilité du schéma aux différences finies

En pratique, le calcul des valeurs approchées $(u_i)_{0 \le i \le N}$ est entâchée d'erreurs d'arrondis. Afin que ces valeurs approchés soient significatives, il est crucial de maîtriser ces erreurs d'arrondis.

<u>Définition:</u> Le schéma aux différences finies est stable si et seulement s'il existe un nombre positif S tel que, quelles que soient les nombres $(u_i)_{0 \le i \le N+1}$, $(u'_i)_{0 \le i \le N+1}$ et $(\delta_i)_{1 \le i \le N}$ tels que:

$$\text{Et}\begin{cases} u_0 = u'_0 = \alpha, \\ \forall 1 \le i \le N, \ \text{Et}\begin{cases} -\dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c(x_i)\, u_i = f(x_i), \\ -\dfrac{u'_{i+1} - 2u'_i + u'_{i-1}}{h^2} + c(x_i)\, u'_i = f(x_i) + \delta_i, \end{cases} \\ u_{N+1} = u'_{N+1} = \beta, \end{cases}$$

nous avons l'inégalité:

$$\underset{0 \le i \le N+1}{\text{Max}} |u'_i - u_i| \le S \underset{1 \le i \le N}{\text{Max}} |\delta_i|.$$

Le nombre S est alors une constante de stabilité du schéma.

La notion de stabilité garantit que de petites erreurs d'arrondis à chaque point $(x_i)_{1 \le i \le N}$ ne conduisent pas à une erreur finale incontrolable.

Afin d'établir la stabilité du schéma aux différences finies, nous revenons à sa forme matricielle. D'après la définition ci-dessus, les vecteurs

$$\delta U = \begin{pmatrix} u'_1 - u_1 \\ \vdots \\ u'_N - u_N \end{pmatrix} \quad \text{et} \quad \Delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_N \end{pmatrix}$$ satisfont l'équation:

$$A_c\,(\delta U) = \Delta,$$

De sorte que:

$$\delta U = A_c^{-1}(\Delta).$$

Etablir la stabilité du schéma revient donc à estimer la norme $|||\cdot|||_\infty$ de la matrice A_c^{-1}:

$$|||A_c^{-1}|||_\infty = \text{Max}\left\{ \|A_c(x)\|_\infty = \underset{1 \le i \le N}{\text{Max}} \left|[A_c(x)]_i\right|,\ x \in \mathbb{R}^N \text{ t.q. } \|x\|_\infty = 1 \right\}.$$

Rappelons que la valeur de cette norme est donnée par la formule :

$$|||A_c^{-1}|||_\infty = \max_{1 \leq i \leq N} \sum_{j=1}^{N} \left|(A_c^{-1})_{i,j}\right|.$$

Pour évaluer cette quantité, nous utilisons la version discrète suivante du principe du maximum faible.

Lemme : Supposons que la fonction c est positive sur I, et considérons un vecteur $b \in \mathbb{R}^N$ tel que : $\forall\, 1 \leq i \leq N,\ b_i \geq 0$. Si le vecteur $u \in \mathbb{R}^N$ est solution de l'équation :

$$A_c(u) = b,$$

alors : $\forall\, 1 \leq i \leq N,\ u_i \geq 0$.

Preuve :

Considérons le plus petit entier $1 \leq k \leq N$ tel que :

$$u_k = \min_{1 \leq i \leq N} u_i,$$

et supposons que $u_k < 0$. Si $k = 1$, alors, nous savons que :

$$b_1 = \frac{2u_1 - u_2}{h^2} + c(x_1)\, u_1 \geq 0.$$

Il vient donc :

$$u_2 \leq (2 + c(x_1) h^2)\, u_1 \leq u_1 + (1 + c(x_1) h^2)\, u_1 < u_1,$$

Ce qui est absurde ! Si $k = N$, alors, nous avons de même :

$$b_N = \frac{2u_N - u_{N-1}}{h^2} + c(x_N)\, u_N \geq 0,$$

De sorte que :

$$u_{N-1} \leq (2 + c(x_N) h^2)\, u_N \leq u_{N-1} + (1 + c(x_N) h^2)\, u_N < u_{N-1},$$

Ce qui est aussi absurde ! Il s'ensuit que $1 < k < N$, ce qui implique que :

$$\frac{2u_k - u_{k+1} - u_{k-1}}{h^2} + c(x_k)\, u_k = b_k \geq 0.$$

Il vient alors :

$$u_{k+1} - u_k + u_{k-1} - u_k = c(x_k)\, u_k\, h^2 \leq 0.$$

Comme $u_{k+1} - u_k \geq 0$ et $u_{k-1} - u_k \geq 0$, nous en déduisons que :

$$u_{k+1} = u_{k-1} = u_k,$$

Ce qui contredit la définition de l'entier k.

Nous déduisons de cette version discrète du principe du maximum faible l'estimation suivante de la matrice A_0^{-1}.

__Lemme:__ Supposons que la fonction c est positive sur I. L'inverse de la matrice A_c vérifie alors :

$$|||A_c^{-1}|||_\infty \leq \frac{1}{8}.$$

__Preuve:__

(i) Considérons la matrice $A_0 := \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix}$ associée au cas où la fonction c est identiquement nulle.

Pour chaque entier $1 \leq i \leq N$, nous désignons par $e_i = (\delta_{i,j})_{1 \leq j \leq N}$ les vecteurs de la base canonique de $\mathbb{R}^N$. Par le principe du maximum faible, nous savons que :

$$\forall\, 1 \leq j \leq N, \ \left[A_0^{-1}(e_i)\right]_j \geq 0.$$

Comme $\left[A_0^{-1}(e_i)\right]_j = \left[A_0^{-1}\right]_{j,i}$, nous en déduisons que la matrice A_0^{-1} est à coefficients positifs. Notons ici que cette propriété s'étend à la matrice A_c^{-1} dans le cas où la fonction c est positive sur I. En particulier, nous avons :

$$|||A_0^{-1}|||_\infty = \underset{1 \leq i \leq N}{\operatorname{Max}} \sum_{j=1}^{N} \left(A_0^{-1}\right)_{i,j}.$$

Afin d'évaluer cette quantité, nous introduisons le problème de Dirichlet :

$$\begin{cases} \forall x \in I, \ -u''(x) = 1, \\ u(0) = u(1) = 0, \end{cases}$$

dont l'unique solution u est donnée par la formule :

$$\forall x \in [0,1], \ u(x) = \frac{1}{2}x(1-x).$$

Pour chaque entier $1 \leq i \leq N$, la formule de Taylor-Lagrange pour cette fonction s'écrit :

$$\begin{cases} u(x_{i+1}) = u(x_i) + h\, u'(x_i) + \frac{h^2}{2}, \\ u(x_{i-1}) = u(x_i) - h\, u'(x_i) + \frac{h^2}{2}, \end{cases}$$

De sorte que :

$$-\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} = 1.$$

Nous concluons que les vecteurs $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ et $u = \begin{pmatrix} u(x_1) \\ \vdots \\ u(x_N) \end{pmatrix}$ satisfont l'identité :

$$A_0(u) = e,$$

Ce qui équivaut à :

$$u = A_0^{-1}(e).$$

En particulier, nous avons :

$$\forall 1 \leq i \leq N, \sum_{j=1}^{N} \left[A_0^{-1}\right]_{i,j} = \left[A_0^{-1}(e)\right]_j = u(x_j),$$

De sorte que :

$$|||A_0^{-1}|||_\infty = \underset{1\leq j\leq N}{\mathrm{Max}} \; u(x_j) \leq \frac{1}{8}.$$

(ii) Dans le cas général où la fonction c est seulement positive, nous écrivons :

$$A_0^{-1} - A_c^{-1} = A_0^{-1}(A_c - A_0)A_c^{-1}.$$

Comme $A_c - A_0 = \begin{pmatrix} c(x_1) & & 0 \\ & \ddots & \\ 0 & & c(x_N) \end{pmatrix}$ est à coefficients positifs, et que les matrices

A_0^{-1} et A_c^{-1} sont aussi à coefficients positifs, nous en déduisons que la matrice $A_0^{-1} - A_c^{-1}$ est à coefficients positifs. Comme la matrice A_c^{-1} est à coefficients positifs, nous concluons que :

$$|||A_c^{-1}|||_\infty \leq |||A_0^{-1}|||_\infty \leq \frac{1}{8}.$$

Nous pouvons enfin démontrer la stabilité du schéma aux différences finies.

Théorème : Supposons que la fonction c est positive sur I. Le schéma aux différences finies est stable pour une constante de stabilité égale à $\frac{1}{8}$.

Preuve :

Considérons des nombres $(u_i)_{0\leq i\leq N+1}$, $(u'_i)_{0\leq i\leq N+1}$ et $(\delta_i)_{1\leq i\leq N}$ tels que :

$$\text{Et}\begin{cases} u_0 = u'_0 = \alpha, \\ \forall 1\leq i\leq N, \begin{cases} -\dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c(x_i)\,u_i = f(x_i), \\ \text{Et} \; -\dfrac{u'_{i+1} - 2u'_i + u'_{i-1}}{h^2} + c(x_i)\,u'_i = f(x_i) + \delta_i, \end{cases} \\ u_{N+1} = u'_{N+1} = \beta. \end{cases}$$

Les vecteurs $\delta U = \begin{pmatrix} u'_1 - u_1 \\ \vdots \\ u'_N - u_N \end{pmatrix}$ et $\Delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_N \end{pmatrix}$ satisfont l'équation :

$$A_c(\delta U) = \Delta,$$

De sorte que :

$$\underset{0\leq i\leq N+1}{\operatorname{Max}} |u'_i - u_i| = \|\delta U\|_\infty = \|A_c^{-1}(\Delta)\|_\infty \leq \frac{1}{8}\|\Delta\|_\infty = \frac{1}{8}\underset{1\leq i\leq N}{\operatorname{Max}}|\delta_i|,$$

D'où la stabilité du schéma aux différences finies.

Remarquons que la stabilité du schéma aux différences finies est inconditionnelle : il n'est pas nécessaire d'imposer une quelconque condition sur le pas h pour que le schéma soit stable.

c. Convergence du schéma aux différences finies

__Définition__ : Le schéma aux différences finies est convergent si et seulement si, quelle que soit la solution $u \in \mathcal{C}^2(I) \cap \mathcal{C}^0([0,1])$ du problème de Dirichlet :

$$\text{Et}\begin{cases} \forall x \in I, \ -u''(x) + c(x)\,u(x) = f(x), \\ u(0) = \alpha \ \text{ et } \ u(1) = \beta, \end{cases}$$

les nombres $(u_i)_{0\leq i\leq N+1}$ définis par les identités :

$$\text{Et}\begin{cases} u_0 = \alpha, \\ \forall\, 1\leq i\leq N, \ -\dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c(x_i)\,u_i = f(x_i), \\ u_{N+1} = \beta, \end{cases}$$

satisfont à la convergence :

$$\underset{0\leq i\leq N+1}{\operatorname{Max}} |u_i - u(x_i)| \to 0,$$

lorsque le pas $h = \dfrac{1}{N+1}$ tend vers 0.

Cette notion garantit que le schéma fournit une approximation valable de la solution $u \in \mathcal{C}^2(I) \cap \mathcal{C}^0([0,1])$ du problème de Dirichlet considéré puisque l'erreur globale

$$E_h = \underset{0\leq i\leq N+1}{\operatorname{Max}} |u_i - u(x_i)|,$$

est aussi petite que souhaité lorsque le pas h tend vers 0.

__Théorème__ : Supposons que les fonctions c et f sont continues sur $[0,1]$, et que c est positive sur I. Alors le schéma aux différences finies est convergent.

Preuve :

La convergence du schéma résulte de sa consistance et de sa stabilité. Considérons en effet les erreurs de consistance $(e_i)_{1 \leq i \leq N}$ définies par :

$\forall\ 1 \leq i \leq N,\quad e_i = -\dfrac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} + c(x_i)\,u(x_i) - f(x_i)$.

Les suites $(u_i)_{0 \leq i \leq N+1}$ et $(u(x_i))_{0 \leq i \leq N+1}$ satisfont alors :

$$\text{Et}\begin{cases} u_0 = u(x_0) = \alpha, \\ \forall 1 \leq i \leq N, \begin{cases} -\dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c(x_i)\,u_i = f(x_i), \\ \text{Et } -\dfrac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} + c(x_i)\,u(x_i) = f(x_i) + e_i, \end{cases} \\ u_{N+1} = u(x_{N+1}) = \beta, \end{cases}$$

De sorte que la stabilité du schéma induit que :

$$\max_{0 \leq i \leq N+1} |u_i - u(x_i)| \leq \frac{1}{8} \max_{1 \leq i \leq N} |e_i|.$$

La conclusion découle alors de la consistance du schéma.

La notion d'ordre permet de calculer la précision du schéma.

Théorème de Gerschgorin : Supposons que les fonctions c et f sont de classe $\mathcal{C}^2$ sur $[0,1]$, et que c est positive sur I. Considérons une solution $u \in \mathcal{C}^4(I) \cap \mathcal{C}^2([0,1])$ du problème de Dirichlet :

$$\text{Et}\begin{cases} \forall x \in I,\ -u''(x) + c(x)\,u(x) = f(x), \\ u(0) = \alpha \text{ et } u(1) = \beta. \end{cases}$$

Les nombres $(u_i)_{0 \leq i \leq N+1}$ définis par les identités :

$$\text{Et}\begin{cases} u_0 = \alpha, \\ \forall\ 1 \leq i \leq N,\ -\dfrac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + c(x_i)\,u_i = f(x_i), \\ u_{N+1} = \beta, \end{cases}$$

satisfont :

$$\max_{0 \leq i \leq N+1} |u_i - u(x_i)| \leq \frac{h^2}{96} \max_{x \in [0,1]} |u^{(4)}(x)|.$$

Preuve :

Comme pour la preuve du théorème précédent, la stabilité du schéma induit que les erreurs de consistance $(e_i)_{1 \leq i \leq N}$ vérifient :

$$\max_{0 \leq i \leq N+1} |u_i - u(x_i)| \leq \frac{1}{8} \max_{1 \leq i \leq N} |e_i|.$$

Les calculs effectués pour établir l'ordre du schéma aux différences finies montrent alors que :

$$\max_{1 \leq i \leq N} |e_i| \leq \frac{1}{12} \max_{x \in [0,1]} |u^{(4)}(x)|,$$

D'où la conclusion ci-dessus.

Notons, en guise de conclusion, que les schémas aux différences finies s'étendent aux problèmes de Dirichlet plus généraux du type :

$$\begin{cases} \forall x \in I, & -\big(a(x)\, u'(x)\big)' + b(x)\, u'(x) + c(x)\, u(x) = f(x), \\ \text{Et} & u(0) = \alpha \text{ et } u(1) = \beta, \end{cases}$$

et qu'ils restent alors convergents sous des hypothèses d'ellipticité, de régularité, de positivité, et de petitesse, ad hoc pour les fonctions a, b, c et f.

II Équations de transport linéaires