



**TECHNIQUES
DE L'INGÉNIEUR**

Réf. : AF500 V1

Méthode des différences finies pour les EDP stationnaires

Cet article est issu de : **Sciences fondamentales | Mathématiques**

par **Pierre SPITERI**

Pour toute question :
Service Relation clientèle
Techniques de l'Ingénieur
Immeuble Pleyad 1
39, boulevard Ornano
93288 Saint-Denis Cedex

Par mail :
infos.clients@teching.com
Par téléphone :
00 33 (0)1 53 35 20 20

Document téléchargé le : **02/05/2017**
Pour le compte : **7200043660 - centralesupelec // 138.195.79.110**

Méthode des différences finies pour les EDP stationnaires

par **Pierre SPITERI**

Docteur ès sciences mathématiques

Professeur à l'École nationale supérieure d'électronique, d'électrotechnique, d'informatique, d'hydraulique et de télécommunication de Toulouse

1. Position du problème.....	AF 500 - 3
2. Discrétisation du problème. Cas du problème monodimensionnel.....	— 3
3. Discrétisation du problème. Cas de domaines carré et cubique.....	— 6
4. Quelques propriétés de la matrice du système linéaire.....	— 8
4.1 Inversibilité du système	— 8
4.2 Critères pratiques d'inversibilité	— 11
5. Majoration d'erreur	— 13
6. Discrétisation du problème dans le cas d'un domaine quelconque.....	— 14
Références bibliographiques	— 16

L'observation d'un phénomène conduit toujours le scientifique à une modélisation qui s'accompagne elle-même d'une mise en équation du problème étudié ; très souvent, les modèles obtenus sont constitués par des équations différentielles ou des équations aux dérivées partielles (EDP) ; malheureusement, les méthodes analytiques de résolution de ce type de problèmes mathématiques ne s'appliquent qu'à une classe très limitée d'équations. À l'aide d'hypothèses simplificatrices, plus ou moins justifiées suivant la valeur des paramètres intervenant dans le modèle, le scientifique se ramène à un type de problèmes qu'il sait résoudre, de manière formelle ; ainsi utilise-t-il des modèles très simplifiés pour représenter les phénomènes observés qui sont souvent complexes.

La plupart du temps, les solutions des équations simplifiées ne représentent le phénomène que dans le domaine où les hypothèses simplificatrices ont un sens ; par contre, lorsque les valeurs des paramètres ne rentrent pas dans ce cadre, la solution obtenue n'a pas toujours un grand rapport avec l'observation. Pour avoir une approche plus fine du phénomène étudié, il faut donc prendre en compte dans les équations les termes qui rendent impossible la résolution analytique du problème. On se trouve donc dans une impasse et il faut trouver un compromis permettant à la fois de représenter les observations le plus exactement possible et de résoudre les équations décrivant le régime de fonctionnement du phénomène.

Cependant, avant d'envisager la résolution du problème d'équations aux dérivées partielles, il convient d'effectuer une étude analytique des équations intervenant dans le modèle ; à ce stade, le scientifique doit se poser des questions sur l'existence, l'unicité de la(des) solution(s), la sensibilité de la(des) solution(s) aux perturbations, la croissance ou la décroissance des solutions en fonction du temps, l'existence de points de bifurcation, etc., ce qui conduit à la résolution de problèmes mathématiques extrêmement complexes,

qui cependant sont des éléments de validation de modèles mathématiques élaborés par le scientifique.

Par ailleurs les récents progrès du calcul automatique ont permis la mise en œuvre de méthodes de calcul qu'il n'était pas concevable d'envisager auparavant. Ces méthodes numériques ont permis notamment :

- la possibilité d'effectuer une grande quantité de calculs dans des temps très brefs ;*
- la prise en compte de non-linéarité dans toutes sortes d'équations ainsi que la résolution de ces dernières.*

Cependant, il faut bien être conscient que la résolution numérique d'une équation aux dérivées partielles conduit à l'obtention d'une solution approchée. En effet, comme on le verra ultérieurement, le principe des méthodes numériques repose sur l'approximation du problème continu initial par un système discret d'équations algébriques linéaires ou non linéaires (selon la nature du problème initial) de grande dimension. La résolution de ce système défini dans un espace de dimension finie nécessite d'effectuer :

— une analyse d'erreur systématique, compte tenu de la construction des schémas d'approximation et permettant de situer la proximité de la solution approchée par rapport à la solution exacte du problème initial ;

— une analyse de la stabilité des schémas numériques, compte tenu de l'influence possible de l'accumulation d'une part des erreurs systématiques inhérentes aux procédés d'approximation et d'autre part des erreurs d'arrondi (ou de chute) ou de troncature dues à la mauvaise représentation des nombres réels en machines, et dont l'effet peut être de provoquer une perte importante de précision et de dénaturer complètement les résultats obtenus ; cette phase est l'analogue discret de l'étude de la sensibilité des solutions aux perturbations, évoquée ci-dessus ;

— des recherches d'algorithmes numériques performants, en liaison avec la grande taille des systèmes à inverser, algorithmes s'exécutant très rapidement ;

— des procédés de parallélisation efficaces sur machines multiprocesseurs des algorithmes numériques de résolution, parallélisation destinée à réduire le temps de restitution d'un travail informatique.

Dans cet article, nous abordons l'étude et la discréétisation par la méthode des différences finies d'équations aux dérivées partielles stationnaires constituées par des opérateurs elliptiques [1]. Nous considérerons essentiellement le problème de Poisson avec conditions aux limites de Dirichlet homogènes ; nous montrerons comment la technique de discréétisation peut être étendue à des situations distinctes concernant soit des opérateurs aux dérivées partielles, comme le problème de convection-diffusion par exemple, soit des conditions aux limites différentes comme les conditions aux limites de Neumann, les conditions aux limites de Fourier (dénommées encore conditions de Robin ou conditions aux limites mixtes), les conditions aux limites mêlées de Dirichlet-Neumann et les conditions aux limites périodiques. Par ailleurs, on verra que la discréétisation d'équations aux dérivées partielles stationnaires conduit à la résolution d'un système algébrique linéaire de grande dimension ; la question essentielle est d'établir rapidement et en faisant le minimum de calculs, la régularité (c'est-à-dire l'inversibilité) de la matrice associée à ce système linéaire. À ce sujet, nous dégagerons les propriétés générales vérifiées par ces matrices, ce qui nous permettra d'exhiber des critères systématiques de régularité des matrices issues de la discréétisation d'équations aux dérivées partielles stationnaires. On insistera également sur les difficultés numériques de résolution des systèmes algébriques issus de la discréétisation par la méthode des différences finies d'équations aux dérivées partielles stationnaires. Les propriétés des matrices de discréétisation nous permettront de donner des résultats généraux de majoration de la norme de l'erreur obtenue en comparant la solution exacte et la solution approchée issue du schéma numérique.

Il convient de noter que cette étude préalable concernant la discréétisation par la méthode des différences finies d'équations aux dérivées partielles stationnaires sera utilisée dans l'article suivant [AF 501] lorsque nous aborderons la discréétisation par la méthode des différences finies d'équations aux dérivées

partielles d'évolution. Par ailleurs, la résolution numérique d'équations aux dérivées partielles conduisant à la résolution de systèmes linéaires creux de grande dimension, on donne dans l'article [AF 502] quelques algorithmes classiques et bien adaptés à la résolution de ce type de système linéaire.

En résumé, cette étude sur la méthode des différences finies pour résoudre des équations aux dérivées partielles se décompose en trois articles :

- [AF 500] Méthode des différences finies pour les EDP stationnaires ;
- [AF 501] Méthode des différences finies pour les EDP d'évolution ;
- [AF 502] Algorithmes numériques pour la résolution des grands systèmes.

1. Position du problème

Soit Ω un domaine borné quelconque de \mathbb{R}^2 et $\partial\Omega$ sa frontière. On se propose de déterminer la solution $u(x, y)$ du classique problème de Poisson avec les conditions aux limites de Dirichlet homogènes :

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

avec Δ opérateur laplacien défini dans \mathbb{R}^2 par $\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$,

$f(x, y)$ une fonction quelconque donnée.

Il convient de noter de façon préalable que la régularité de la fonction f induit celle de la solution du problème précédent. Nous admettrons le résultat suivant [1] :

Proposition 1. Si la frontière $\partial\Omega$ est régulière et si f est une fonction m fois continûment différentiable, alors la solution $u(x, y)$ du problème de Poisson est $m+2$ fois continûment différentiable.

Remarque

La régularité de la frontière est assurée si son graphe est continu et si $\partial\Omega$ ne comporte pas de point de rebroussement.

Ainsi, si f est une fonction continue, u est deux fois continûment différentiable. Ce résultat admet implicitement que le problème précédent admet une solution unique ; or ce type de résultat est délicat à établir et nous l'admettrons dans le cas du problème modèle. Cependant, si l'on modifie légèrement les conditions aux limites du problème, l'unicité de la solution n'est plus garantie. Pour illustrer cette remarque, considérons le cas du problème de Poisson avec conditions aux limites de Neumann non homogènes :

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

avec $\frac{\partial u}{\partial n}$ dérivée normale définie par :

$$\frac{\partial u}{\partial n} = \vec{\nabla} u \cdot \vec{n} = \frac{\partial u}{\partial x} \cos(\vec{Ox}, \vec{n}) + \frac{\partial u}{\partial y} \sin(\vec{Ox}, \vec{n})$$

$\vec{\nabla} u$ gradient de la fonction u ,

\vec{n} normale au domaine Ω orientée vers l'extérieur.

Il est alors facile de vérifier que toute fonction $v = u + k$, où k est une constante quelconque, est aussi solution du problème précédent.

dent. Ainsi il n'y a pas unicité de la solution pour le problème de Poisson avec conditions aux limites de Neumann non homogènes. Cependant si l'on considère le problème suivant :

$$\begin{cases} -\Delta u + qu = f \text{ dans } \Omega \quad (q > 0) \\ \frac{\partial u}{\partial n} = g \text{ sur } \partial\Omega \end{cases}$$

on montre que la présence du terme réel positif q a pour influence d'assurer l'unicité de la solution.

Lorsqu'il modélise un phénomène, le physicien sait bien que la solution physique de son problème existe et est unique ; par contre, il ne connaît pas la pertinence de sa modélisation. À ce stade, les considérations sur l'existence et l'unicité de la solution fournissent un premier élément de pertinence du modèle. Un autre élément de pertinence est fourni par des données quantitatives constituées par l'expression de la solution de l'équation aux dérivées partielles. Pour une donnée quelconque des seconds membres de ces dernières, il est parfois difficile de donner une expression analytique de la solution. Les progrès de l'informatique, en particulier des méthodes numériques, ont permis de résoudre ce type de problème sur ordinateur. Dans la suite de cet article, nous exposerons la méthode des différences finies dans le cas où le domaine Ω est le segment unité, le carré unité ou le cube unité ; nous verrons à la fin de cet article les modifications à apporter dans le cas où Ω est un domaine borné quelconque de \mathbb{R}^2 ou \mathbb{R}^3 .

2. Discrétisation du problème. Cas du problème monodimensionnel

Pour introduire les techniques de différences finies, considérons le cas du problème de Poisson monodimensionnel avec conditions aux limites de Dirichlet homogènes. Ce problème peut être illustré sur le plan physique comme suit : soit une corde élastique de longueur unité, attachée en chacune de ses extrémités ; on agit sur cette corde avec une force $f(x)$, perpendiculaire à la corde, dans le plan contenant la corde. On se propose de déterminer, pour tout $x \in \Omega = [0, 1]$, le déplacement $u(x)$ de la corde soumise à la force $f(x)$, déplacement compatible avec les conditions aux limites $u(0) = u(1) = 0$. Sous l'hypothèse des petits déplacements, on montre que le problème d'élasticité précédent est modélisé par l'équation aux dérivées partielles suivante :

$$\begin{cases} -\frac{d^2 u(x)}{dx^2} = f(x), \quad x \in]0, 1[\\ u(0) = u(1) = 0 \end{cases} \quad (1)$$

Remarquons que, si les conditions aux limites ne sont pas homogènes, c'est-à-dire si $u = g$ (donnée) sur $\partial\Omega$, on peut se ramener au cas homogène par un simple changement de variable du type $v = u - g$, compte tenu des propriétés de linéarité de l'opérateur de dérivation.

La discréttisation du problème précédent consiste à remplacer par une technique appropriée le problème continu par un système linéaire algébrique. L'approximation s'effectue en deux étapes successives. La première consiste à discréttiser le domaine Ω ; pour cela, n étant un nombre entier strictement positif donné, on découpe l'intervalle $[0, 1]$ en $(n+1)$ parties de longueur h , telles que $h = \frac{1}{n+1}$; **h est le pas de discréttisation.** On considère alors le maillage constitué par les $n+2$ points $x_j = jh$, $j = 0, 1, \dots, n, n+1$; pour $j = 1, \dots, n$, on cherche une approximation de $u(x_j)$ aux points du maillage en considérant les n égalités suivantes :

$$-\frac{d^2u(x_j)}{dx^2} = f(x_j), \quad j = 1, \dots, n$$

La seconde étape consiste à discréttiser les opérateurs; à ce stade, on remplace les dérivées par des quotients différentiels faisant intervenir les valeurs de la fonction inconnue u aux points du maillage; pour cela, on suppose que la fonction u est quatre fois continûment différentiable, ce qui est une hypothèse abusive car non satisfaite dans tous les cas, compte tenu du résultat de la proposition 1; on demande donc plus de régularité à la fonction $u(x)$, qu'elle n'en a en réalité. Cependant, on verra que cette hypothèse est nécessaire pour obtenir des majorations d'erreur. Compte tenu de cette hypothèse abusive, on considère les développements en série de Taylor limités à l'ordre 4 de $u(x_{j+1})$ et $u(x_{j-1})$, autour du point x_j ; on a alors :

$$\begin{aligned} u(x_{j+1}) &= u(x_j) + h \frac{du(x_j)}{dx} + \frac{h^2}{2!} \frac{d^2u(x_j)}{dx^2} \\ &\quad + \frac{h^3}{3!} \frac{d^3u(x_j)}{dx^3} + \frac{h^4}{4!} \frac{d^4u(\tilde{x}_j)}{dx^4}, \quad x_j < \tilde{x}_j < x_{j+1} \\ u(x_{j-1}) &= u(x_j) - h \frac{du(x_j)}{dx} + \frac{h^2}{2!} \frac{d^2u(x_j)}{dx^2} \\ &\quad - \frac{h^3}{3!} \frac{d^3u(x_j)}{dx^3} + \frac{h^4}{4!} \frac{d^4u(\hat{x}_j)}{dx^4}, \quad x_{j-1} < \hat{x}_j < x_j \end{aligned}$$

En additionnant les deux expressions précédentes, on obtient finalement :

$$\begin{aligned} \frac{d^2u(x_j)}{dx^2} &= \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2} \\ &\quad - \frac{h^2}{4!} \left(\frac{d^4u(\hat{x}_j)}{dx^4} + \frac{d^4u(\tilde{x}_j)}{dx^4} \right) \end{aligned}$$

On a alors la majoration d'erreur pour l'approximation du laplacien :

Lemme 1. Soit $u \in \mathbf{C}^4([0, 1])$, une fonction quelconque et on pose :

$$M_4 = \text{Sup} \left(\left| \frac{d^4u(x)}{dx^4} \right|, 0 \leq x \leq 1 \right)$$

alors, on a :

$$\left| \frac{d^2u(x_j)}{dx^2} - \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2} \right| \leq \frac{h^2}{12} M_4$$

et le schéma numérique est précis à l'ordre 2.

Remarque

Au lieu de considérer la discréttisation du problème modèle, on peut considérer le problème de convection-diffusion monodimensionnel, avec conditions aux limites de Dirichlet homogènes :

$$\begin{cases} -\frac{d^2u(x)}{dx^2} + a \frac{du(x)}{dx} + bu(x) = f(x), & x \in]0, 1[, b \geq 0 \\ u(0) = u(1) = 0 \end{cases}$$

Pour ce type de problème, on a aussi besoin d'avoir une approximation de la dérivée première de u au point x_j ; celle-ci s'obtient de façon analogue en considérant les développements limités de $u(x_{j+1})$ et $u(x_{j-1})$; par exemple, en soustrayant les développements limités précédents, on obtient le schéma centré suivant :

$$\frac{du(x_j)}{dx} = \frac{u(x_{j+1}) - u(x_{j-1})}{2h} - \frac{h^2}{6} \frac{d^3u(x_j)}{dx^3} + O(h^3)$$

On peut également considérer d'autres schémas numériques pour approcher la dérivée première de la fonction $u(x)$ au point x_j ; ces derniers sont toujours construits à partir des développements limités de $u(x_{j+1})$ et $u(x_{j-1})$.

Exemple : le premier développement conduit à définir le schéma aux différences finies en avant (ou différences finies progressives) suivant :

$$\frac{du(x_j)}{dx} = \frac{u(x_{j+1}) - u(x_j)}{h} - \frac{h}{2} \frac{d^2u(x_j)}{dx^2} + O(h)$$

alors que le second développement conduit à définir le schéma aux différences finies en arrière (ou différences finies rétrogrades) suivant :

$$\frac{du(x_j)}{dx} = \frac{u(x_j) - u(x_{j-1})}{h} + \frac{h}{2} \frac{d^2u(x_j)}{dx^2} + O(h)$$

Ces deux derniers schémas sont également utiles lorsque l'on doit discréttiser des conditions aux limites de type Neumann, qui se réduisent, dans le cas de problème monodimensionnel, à une condition sur la dérivée première en les points $x=0$ et $x=1$. On utilisera également ces deux derniers schémas dans le cas de la discréttisation de la partie temporelle de problèmes d'évolution (cf. article [AF 501]).

Remarque

Il existe d'autres types de schémas de discréttisation pour approcher les valeurs des dérivées successives de la fonction $u(x)$ aux points du maillage. Le principe de construction des schémas est le même et revient à considérer des combinaisons linéaires des développements limités de l'expression de la fonction u en les points du maillage; si l'on suppose que la solution $u(x)$ est six fois continûment différentiable, on peut obtenir, par exemple, le schéma suivant, précis à l'ordre 4 :

$$\begin{aligned} \frac{d^2u(x_j)}{dx^2} &= \frac{u(x_{j-2}) + 16u(x_{j-1}) - 30u(x_j) + 16u(x_{j+1}) - u(x_{j+2})}{12h^2} \\ &+ O(h^4) \end{aligned}$$

pour l'approximation de la dérivée seconde au point x_j . Le lecteur est renvoyé à [2] pour des informations supplémentaires.

Revenons à la discréttisation du laplacien monodimensionnel qui, au stade où nous nous trouvons, n'est pas complètement terminée; en effet, si l'on remplace la dérivée seconde par le quotient différentiel, on obtient :

$$\begin{aligned} \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2} &- \frac{h^2}{4!} \left(\frac{d^4u(\hat{x}_j)}{dx^4} + \frac{d^4u(\tilde{x}_j)}{dx^4} \right) \\ &= f(x_j), \quad 1 \leq j \leq n \end{aligned}$$

ce qui fait intervenir la dérivée quatrième dans le schéma de discréttisation ; or on ne connaît pas les points \hat{x}_j et \tilde{x}_j , ce qui rend l'expression précédente inopérante. Cependant, si le pas de discréttisation h est petit, compte tenu du résultat du lemme 1, on peut écrire, en négligeant le terme en h^2 :

$$\left\{ \begin{array}{l} \frac{-u(x_{j+1}) + 2u(x_j) - u(x_{j-1})}{h^2} \approx f(x_j), \quad 1 \leq j \leq n \\ u(x_0) = u(x_{n+1}) = 0 \end{array} \right.$$

Par analogie avec l'écriture précédente, on définit le schéma numériquue suivant, pour l'approximation du problème (1) :

$$\left\{ \begin{array}{l} \frac{-u_{j+1} + 2u_j - u_{j-1}}{h^2} = f_j = f(x_j), \quad 1 \leq j \leq n \\ u_0 = u_{n+1} = 0 \end{array} \right.$$

Ces relations définissent les n équations d'un système algébrique linéaire qui s'écrit matriciellement :

$$AU = F$$

dans lequel A est une matrice de dimension n , U et F sont des vecteurs également de dimension n ; l'écriture des n équations définissant le schéma conduit à expliciter A , U et F comme suit :

$$A = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}$$

$$U = (u_1, u_2, \dots, u_n)^t$$

$$F = (h^2 f_1 + u_0, h^2 f_2, \dots, h^2 f_{n-1}, h^2 f_n + u_{n+1})^t$$

Il est à noter que la matrice A est symétrique ; cette propriété importante est à rapprocher du fait que l'opérateur de diffusion, représenté par la dérivée seconde, est lui-même autoadjoint. À ce stade, on remarque que cette dernière propriété de l'opérateur continu est conservée lorsque l'on discrétise ce même opérateur. De plus, la matrice A a une structure creuse tridiagonale ; on a obtenu une telle structure à cause de la numérotation lexicographique des points du maillage. Si l'on modifie la numérotation des points du maillage, la structure de la matrice A sera également modifiée.

Remarque

Notons que, dans l'écriture du système algébrique linéaire associé au problème de diffusion, les conditions aux limites apparaissent dans la première et la dernière composante du vecteur F . Compte tenu du fait que, dans le problème modèle, les conditions aux limites de Dirichlet sont homogènes, on aurait pu omettre de faire figurer les termes u_0 et u_{n+1} intervenant dans ces deux dernières composantes. Cependant, dans le cas de conditions aux limites de Dirichlet non homogènes, il convient de tenir compte des valeurs u_0 et u_{n+1} .

Remarque

On verra au lemme 4 que la matrice A est symétrique définie positive dans la mesure où la relation suivante est vérifiée :

$$U^t A U = u_1^2 + u_n^2 + \sum_{i=1}^{n-1} (u_{i+1} - u_i)^2, \quad \forall U \in \mathbb{R}^n$$

■ Nous avons évoqué précédemment d'autres types de conditions aux limites que les conditions aux limites de Dirichlet qui traduisent chaque fois des réalités physiques différentes ; parmi les plus classiques citons :

— les **conditions aux limites de type Neumann**, qui traduisent le fait qu'un flux est imposé à la frontière $\partial\Omega$; dans le cas du problème de Poisson monodimensionnel avec conditions de Neumann homogènes, on aura donc à résoudre le problème suivant :

$$\left\{ \begin{array}{l} -\frac{d^2 u(x)}{dx^2} = f(x), \quad x \in]0, 1[\\ \frac{du(0)}{dx} = 0 \\ \frac{du(1)}{dx} = 0 \end{array} \right.$$

Remarquons que, comme indiqué précédemment, ce problème n'a pas de solution unique puisque toute solution translatée par une constante est encore solution ; pour cette raison, on considère plutôt le problème suivant :

$$\left\{ \begin{array}{l} -\frac{d^2 u(x)}{dx^2} + qu(x) = f(x), \quad x \in]0, 1[, \quad (q > 0) \\ \frac{du(0)}{dx} = 0 \\ \frac{du(1)}{dx} = 0 \end{array} \right. \quad (2)$$

— les **conditions aux limites de type Fourier (ou de Robin)**, correspondant à une combinaison linéaire entre les conditions aux limites de Dirichlet et celles de Neumann ; on considère donc le problème suivant :

$$\left\{ \begin{array}{l} -\frac{d^2 u(x)}{dx^2} = f(x), \quad x \in]0, 1[\\ \frac{du(0)}{dx} + ru(0) = g, \quad \frac{du(1)}{dx} + ru(1) = h, \quad (r > 0) \end{array} \right. \quad (3)$$

avec h et g des données du problème ;

— les **conditions aux limites mélées de type Dirichlet-Neumann** correspondant à la situation où des phénomènes physiques différents agissent sur des portions distinctes de la frontière ; par exemple, on peut avoir des conditions de Dirichlet homogènes sur une partie de $\partial\Omega$ et des conditions de Neumann sur une autre partie ; dans le cas du problème de Poisson monodimensionnel, on aura donc à résoudre le problème suivant :

$$\left\{ \begin{array}{l} -\frac{d^2 u(x)}{dx^2} = f(x), \quad x \in]0, 1[\\ \frac{du(0)}{dx} = g, \quad u(1) = 0 \end{array} \right. \quad (4)$$

— les **conditions aux limites de type périodique**, qui se traduisent dans le cas du problème de Poisson monodimensionnel par :

$$\left\{ \begin{array}{l} -\frac{d^2 u(x)}{dx^2} + qu(x) = f(x), \quad x \in]0, 1[, \quad (q > 0) \\ \frac{du(0)}{dx} = \frac{du(1)}{dx} \\ u(0) = u(1) \end{array} \right. \quad (5)$$

Il est clair que, lorsque l'on sait traiter numériquement les conditions aux limites de type Dirichlet et de type Neumann, on est en mesure de traiter tous les cas évoqués précédemment.

- Si l'on considère le cas du **problème de Poisson monodimensionnel** (3) (avec $q \neq 0$) **muni des conditions aux limites de Neumann homogènes**, on obtient facilement le schéma suivant :

$$\begin{cases} \frac{-u_{j+1} + 2u_j - u_{j-1}}{h^2} + qu_j = f_j = f(x_j), & 1 \leq j \leq n \\ u_0 - u_1 = 0, \quad u_{n+1} - u_n = 0 \end{cases}$$

qui définit un système linéaire de dimension $(n+2)$, la matrice associée A ayant l'allure suivante :

$$A = \begin{pmatrix} 1 & -1 & & & & & \\ -1 & 2+\delta & -1 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 2+\delta & -1 \\ & & & & & -1 & 1 \end{pmatrix}$$

avec $\delta = q h^2$.

Cette matrice est symétrique définie positive ; en effet, on obtient facilement la relation (voir lemme 4 ci-après) :

$$U^t A U = \sum_{i=0}^n (u_{i+1} - u_i)^2 + \delta \|U\|_2^2, \quad \forall U \in \mathbb{R}^{n+2}$$

et l'on constate que lorsque $\delta = q = 0$, cette matrice n'est plus définie positive et, par conséquent, elle est singulière ; cette difficulté est l'analogue discret de la non-unicité de la solution du problème continu lorsque $q = 0$.

Remarque

La discréttisation du problème de Poisson monodimensionnel muni des conditions aux limites de type Fourier ne présente pas de difficultés majeures.

- Considérons à présent la **discréttisation du problème de Poisson monodimensionnel muni des conditions aux limites de type mêlées** (4) ; dans ce cas, le schéma de discréttisation par différences finies s'écrit :

$$\begin{cases} \frac{-u_{j+1} + 2u_j - u_{j-1}}{h^2} = f_j = f(x_j), & 1 \leq j \leq n \\ u_0 - u_1 = hg, \quad u_{n+1} = 0 \end{cases}$$

Ces relations définissent un système linéaire de dimension $(n+1)$, la matrice A associée étant définie par :

$$A = \begin{pmatrix} 1 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}$$

Cette matrice est symétrique définie positive dans la mesure où la relation suivante est vérifiée :

$$U^t A U = u_n^2 + \sum_{i=0}^{n-1} (u_{i+1} - u_i)^2, \quad \forall U \in \mathbb{R}^{n+1}$$

- Pour terminer, considérons le **problème de Poisson monodimensionnel muni des conditions aux limites périodiques** ; dans ce cas, le schéma de discréttisation par différences finies s'écrit :

$$\begin{cases} \frac{-u_{j+1} + 2u_j - u_{j-1}}{h^2} + qu_j = f_j = f(x_j), & 1 \leq j \leq n \\ u_0 = u_{n+1} \end{cases}$$

La matrice associée étant définie par :

$$A = \begin{pmatrix} 2+\delta & -1 & & & & & & -1 \\ -1 & 2+\delta & -1 & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & -1 & 2+\delta & -1 & & \\ & & & & -1 & 2+\delta & -1 & \\ & & & & & -1 & 2+\delta & \end{pmatrix}$$

Cette matrice est symétrique définie positive.

Remarque

Dans la suite, on pourra, de manière analogue, discréttiser des problèmes aux limites munis de toutes les conditions aux limites considérées précédemment ; en fait, pour alléger l'exposé, nous ne considérerons que les **conditions aux limites de Dirichlet homogènes**.

3. Discréttisation du problème. Cas de domaines carré et cubique

- Soit à présent $\Omega = [0, 1] \times [0, 1]$, correspondant au carré unité. On souhaite discréttiser le problème de Poisson avec conditions aux limites de Dirichlet homogènes dans ce domaine Ω , c'est-à-dire :

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

Ce problème modélise, par exemple, la situation d'une membrane élastique qui, au repos, occupe un domaine Ω du plan xOy . On suppose de plus que cette membrane est attachée le long de son contour $\partial\Omega$ et qu'elle est soumise à une force de densité f par unité de surface, perpendiculaire au plan xOy . L'inconnue du problème est alors le déplacement $u(x, y)$.

La démarche de discréttisation sera exactement la même que celle vue au paragraphe 2, c'est-à-dire qu'elle se décomposera en deux étapes :

- discréttisation du domaine Ω ;
- discréttisation des opérateurs définissant les équations du problème de Poisson.

Pour discréttiser le domaine Ω , on discréttise chacun des intervalles fermés $[0, 1]$, de façon analogue à la technique étudiée au

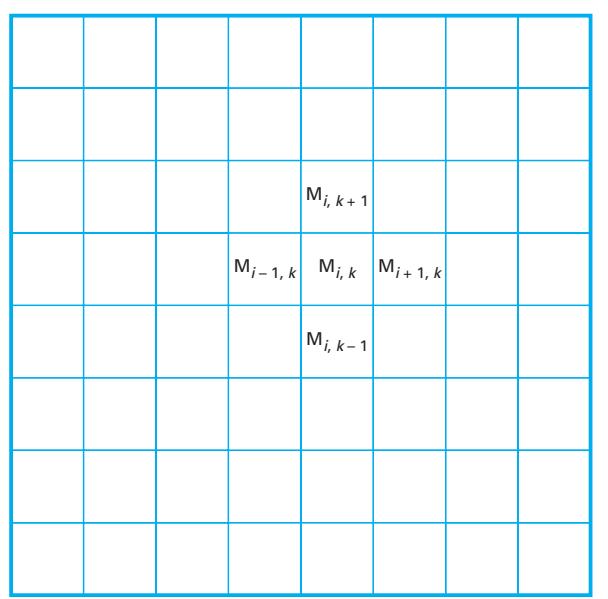


Figure 1 – Maillage différences finies uniforme

paragraphe 2. Soit n un entier naturel et h le pas de discréttisation défini par $h = \frac{1}{n+1}$; on définit un réseau de points sur l'axe des abscisses ainsi que sur l'axe des ordonnées comme indiqué sur la figure 1; autrement dit, on définit les nombres réels $x_i = ih$ et $y_k = kh$ pour $(i, k) \in \{0, \dots, n+1\} \times \{0, \dots, n+1\}$. À partir de ces données, on définit la grille de points $M_{i,k}$, $(i, k) \in \{0, \dots, n+1\} \times \{0, \dots, n+1\}$, cette grille étant constituée par les droites d'équation $x = x_i$ et $y = y_k$, $0 \leq i, k \leq n+1$. Alors, pour $1 \leq i, k \leq n$, on peut écrire :

$$-\Delta u(M_{i,k}) = f(M_{i,k})$$

les valeurs de u , pour les points d'indices i et k nuls ou égaux à $n+1$, étant nulles car correspondant aux conditions aux limites. Donc, comme au paragraphe 2, au lieu de chercher $u(x,y)$ dans tout le domaine Ω , on va chercher une approximation de cette valeur aux points de grille contenus dans Ω .

La discréttisation des opérateurs s'effectue en adaptant ce qui a été établi au paragraphe précédent. On suppose donc que l'hypothèse abusive est vérifiée, c'est-à-dire que la solution $u(x,y)$ est quatre fois continûmement différentiable par rapport à ses arguments. Comme précédemment, on obtient les schémas de discréttisation en combinant les développements limités d'une part de $u(M_{i+1,k})$ et $u(M_{i-1,k})$ et d'autre part de $u(M_{i,k+1})$ et $u(M_{i,k-1})$. Tous calculs effectués, on obtient les relations suivantes pour la discréttisation des dérivées secondes en x et en y :

$$\begin{aligned} \frac{\partial^2 u(M_{i,k})}{\partial x^2} &= \frac{u(M_{i+1,k}) - 2u(M_{i,k}) + u(M_{i-1,k})}{h^2} \\ &\quad - \frac{h^2}{4!} \left(\frac{\partial^4 u(\hat{M}_{i,k})}{\partial x^4} + \frac{\partial^4 u(\tilde{M}_{i,k})}{\partial x^4} \right) \end{aligned}$$

avec $\hat{M}_{i,k} \in [M_{i,k}, M_{i+1,k}]$, $\tilde{M}_{i,k} \in [M_{i-1,k}, M_{i,k}]$,

$$\text{et } \frac{\partial^2 u(M_{i,k})}{\partial y^2} = \frac{u(M_{i,k+1}) - 2u(M_{i,k}) + u(M_{i,k-1})}{h^2} - \frac{h^2}{4!} \left(\frac{\partial^4 u(\ddot{M}_{i,k})}{\partial y^4} + \frac{\partial^4 u(\ddot{\tilde{M}}_{i,k})}{\partial y^4} \right)$$

avec $\ddot{M}_{i,k} \in [M_{i,k}, M_{i,k+1}]$, $\ddot{\tilde{M}}_{i,k} \in [M_{i,k-1}, M_{i,k}]$.

On a alors la majoration d'erreur pour l'approximation du laplacien :

Lemme 2. Soit $u \in C^4(\Omega)$, une fonction quelconque ; on pose :

$$M_4 = \text{Sup} \left(\left| \frac{\partial^4 u(M)}{\partial x^4} \right|, M \in \bar{\Omega} \right) + \text{Sup} \left(\left| \frac{\partial^4 u(M)}{\partial y^4} \right|, M \in \bar{\Omega} \right)$$

On a alors l'estimation suivante, pour tout $(i, k) \in \{1, \dots, n\} \times \{1, \dots, n\}$:

$$\left| \frac{u(M_{i+1,k}) + u(M_{i,k+1}) - 4u(M_{i,k}) + u(M_{i,k-1}) + u(M_{i-1,k})}{h^2} - \Delta u(M_{i,k}) \right| \leq \frac{h^2}{12} M_4$$

Par conséquent, de la même façon qu'au paragraphe 2, le paramètre h étant supposé petit, on peut négliger les termes faisant intervenir les dérivées quatrièmes, ce qui permet de déduire le schéma d'approximation suivant pour résoudre le problème de Poisson dans le domaine Ω considéré :

$$\begin{cases} \frac{-u_{i+1,k} - u_{i,k+1} + 4u_{i,k} - u_{i,k-1} - u_{i-1,k}}{h^2} = f_{i,k}, & 1 \leq i, k \leq n \\ u_{0,k} = u_{i,0} = u_{n+1,k} = u_{i,n+1} = 0, & 0 \leq i, k \leq n+1 \end{cases}$$

avec $f_{i,k} = f(M_{i,k})$.

Comme précédemment, cet ensemble de relations définit un système linéaire :

$$AU = F$$

de dimension n^2 ; la matrice A est symétrique, pentadiagonale et possède la structure tridiagonale par blocs suivante :

$$A = \begin{pmatrix} B - I & & & & \\ -I & B - I & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \\ & & & & -I & B - I \\ & & & & & -I & B \end{pmatrix}$$

avec I matrice identité de dimension n ,

B matrice tridiagonale de dimension n suivante :

$$B = \begin{pmatrix} 4 & -1 & & & & \\ -1 & 4 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \\ & & & & -1 & 4 & -1 \\ & & & & & -1 & 4 \end{pmatrix}$$

De plus, les composantes du vecteur U sont $u_{i,k}$ pour $(i, k) \in \{1, \dots, n\} \times \{1, \dots, n\}$ et sont ordonnées de la façon suivante :

$$U = (U_1; U_2; \dots; U_n)^t$$

chaque vecteur U_i étant un vecteur de dimension n et représentant les valeurs de la solution approchée du problème de Poisson sur une ligne (ou une colonne en fonction de la numérotation adoptée pour les points du maillage).

Les composantes du vecteur F sont $h^2 f_{i,k}$, pour $(i, k) \in \{1, \dots, n\} \times \{1, \dots, n\}$, auxquelles il convient d'ajouter les valeurs frontières dans le cas où celles-ci seraient différentes de zéro. Il convient de noter que le vecteur F possède une structure par blocs compatible avec celle du vecteur U et représentée ci-dessous :

$$F = (F_1; F_2; \dots; F_n)^t$$

Si le domaine Ω est le cube unité défini par $\Omega = [0, 1] \times [0, 1] \times [0, 1]$, la technique de discréétisation du problème de Poisson avec conditions aux limites de Dirichlet homogènes est identique à celle que nous venons d'exposer dans le cas où le domaine Ω est le carré unité. Afin de ne pas alourdir l'exposé, nous donnons simplement le résultat de la discréétisation par différences finies, en signalant simplement les modifications. Le pas de discréétisation étant toujours noté h et défini de la même façon que précédemment à partir de la donnée d'un entier naturel n , le schéma de discréétisation associé au point $M_{i,k,\ell}$ de coordonnées x_i, y_k, z_ℓ est donné par les relations suivantes :

$$\left\{ \begin{array}{l} \frac{-u_{i+1,k,\ell} - u_{i,k+1,\ell} - u_{i,k,\ell+1} + 6u_{i,k,\ell} - u_{i,k,\ell-1} - u_{i,k-1,\ell} - u_{i-1,k,\ell}}{h^2} = f_{i,k,\ell}, 1 \leq i, k, \ell \leq n \\ u_{0,k,\ell} = u_{n+1,k,\ell} = u_{i,0,\ell} = u_{i,n+1,\ell} = u_{i,k,0} = u_{i,k,n+1} = 0, \\ 0 \leq i, k, \ell \leq n+1 \end{array} \right.$$

Cet ensemble de relations définit un système algébrique linéaire :

$$AU = F$$

avec A matrice de dimension n^3 , qui possède la structure par blocs suivante :

$$A = \begin{pmatrix} B_2 - I & & & \\ -I & B_2 - I & & \\ & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots & \ddots \\ & & & & -I & B_2 - I \\ & & & & & -I & B_2 \end{pmatrix}$$

B_2 matrice de dimension n^2 , représentant la matrice de discréétisation sur un plan, B_2 ayant la structure tridiagonale par blocs suivante :

$$B_2 = \begin{pmatrix} C - I & & & & \\ -I & C - I & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & \ddots \\ & & & & -I & C - I \\ & & & & & -I & C \end{pmatrix}$$

C matrice tridiagonale de dimension n ayant la structure suivante :

$$C = \begin{pmatrix} 6 & -1 & & & & & \\ -1 & 6 & -1 & & & & \\ & -1 & 6 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 6 & -1 \\ & & & & & -1 & 6 \end{pmatrix}$$

Remarque

Comme pour le problème monodimensionnel, on peut également définir d'autres schémas numériques. Le lecteur est renvoyé aux références [2] et [3] pour la définition de schémas numériques distincts de celui considéré précédemment.

À ce stade de l'exposé, et pour l'ensemble des problèmes que nous avons discréétisés, on est en droit de se poser les questions suivantes :

question n° 1 : le système linéaire $AU = F$ admet-il une solution unique et comment la déterminer ?

question n° 2 : les composantes du vecteur U sont-elles de « bonnes » approximations de la solution exacte du problème considéré ? En d'autres termes, comment donner une majoration de la norme de l'erreur entre la solution exacte et la solution approchée, calculée en les points du maillage.

Ces questions sont, en général, difficiles et seront abordées respectivement aux paragraphes 4 et 5.

4. Quelques propriétés de la matrice du système linéaire

Dans ce paragraphe, lorsqu'il n'y aura aucune confusion, on notera A les matrices de discréétisation par différences finies obtenues aux deux paragraphes précédents ; lorsque cela sera nécessaire, on précisera ces matrices en les notant respectivement A_{1D} , A_{2D} et A_{3D} selon que le domaine Ω est le segment unité, le carré unité ou le cube unité.

4.1 Inversibilité du système

Le but du présent sous-paragraphe est essentiellement de répondre partiellement à la première question ; en effet au niveau de ce sous-paragraphe, nous nous intéresserons à l'inversibilité de la matrice A , et nous nous efforcerons de dégager des critères garantissant la régularité de la matrice tout en minimisant l'effort de calcul. La partie algorithmique concernant le calcul effectif du vecteur U , question délicate compte tenu de la taille des matrices à inverser, sera étudiée dans l'article [AF 502] de ce traité.

Dans un premier temps, nous allons vérifier l'**inversibilité de la matrice A_{1D}** , en vérifiant que son déterminant est différent de zéro ; on a le résultat suivant :

Lemme 3. Soit n la dimension de la matrice A_{1D} . On a alors $\det(A_{1D}) = n + 1$.

En effet, notons $A^{(n)} = A_{1D}$, cette matrice de dimension n ; le calcul du déterminant de la matrice $A^{(n)}$ conduit à la relation de récurrence :

$$\det(A^{(n)}) = 2 \det(A^{(n-1)}) - \det(A^{(n-2)})$$

ce qui établit le résultat, compte tenu du fait que, d'une part, on obtient à partir de la relation précédente que $\det(A^{(n)}) = \det(A^{(n-1)}) + 1$ et d'autre part que $\det(A^{(1)}) = 2$ et $\det(A^{(2)}) = 3$.

Remarque

Pour les matrices A_{2D} et A_{3D} , le calcul du déterminant n'est pas trivial ; de plus, ce calcul ne présente pas un intérêt majeur. Cependant nous vérifierons que le déterminant de ces matrices est différent de zéro par des techniques plus générales et plus faciles à mettre en œuvre, ce qui entraîne la régularité des matrices considérées.

Corollaire 1. La matrice A_{1D} est inversible.

Remarque

Ce résultat s'étend au cas des matrices A_{2D} et A_{3D} .

■ Les matrices A_{1D} , A_{2D} et A_{3D} possèdent une autre propriété importante qui garantit leurs invariabilités ; en effet, ces matrices sont **définies positives**. Cette propriété est intéressante et découle des propriétés de positivité de l'opérateur laplacien ; elle est encore valable, dans certaines conditions, pour l'opérateur de convection-diffusion. Cependant, cette propriété joue un rôle important pour établir les majorations d'erreur que nous aborderons au paragraphe 5. Nous indiquons ci-dessous un résultat préliminaire, valable pour les matrices A_{1D} , A_{2D} et A_{3D} , que nous étendrons à des situations plus générales à la fin du présent paragraphe.

Lemme 4. Les matrices A_{1D} et A_{2D} sont symétriques définies positives et on a, de plus, les relations suivantes :

$$\begin{aligned} U^t A_{1D} U &= u_1^2 + u_n^2 + \sum_{i=1}^{n-1} (u_{i+1} - u_i)^2, \quad \forall U \in \mathbb{R}^n \\ U^t A_{2D} U &= \sum_{j=1}^n \left(u_{1,j}^2 + u_{n,j}^2 + \sum_{i=1}^{n-1} (u_{i+1,j} - u_{i,j})^2 \right) \\ &\quad + \sum_{i=1}^n \left(u_{i,1}^2 + u_{i,n}^2 + \sum_{j=1}^{n-1} (u_{i,j+1} - u_{i,j})^2 \right), \quad \forall U \in \mathbb{R}^{n^2} \end{aligned}$$

Les calculs de ces relations sont laissés à titre d'exercice.

Remarque

On peut obtenir le même type d'estimation pour la matrice A_{3D} .

■ Par ailleurs, on verra que les **valeurs propres** des matrices considérées jouent un rôle important pour l'**analyse de la stabilité des schémas numériques associés** à la discréttisation des équations aux dérivées partielles considérées ; cette dernière propriété de stabilité traduit en fait la sensibilité du schéma à ne pas amplifier d'une part la propagation des erreurs d'approximation effectuées lors de la construction des schémas numériques et d'autre part les erreurs de chute ou de troncation dues à la mauvaise représentation des nombres réels en machine. Ce phénomène numérique intervient aussi bien lors de la résolution de problèmes stationnaires que de problèmes d'évolution, les critères d'analyse étant cependant distincts dans les deux cas. Concernant la résolution de l'équation de Poisson avec conditions aux limites de Dirichlet homogènes dans le segment unité, le carré unité ou le cube unité, on a le résultat suivant (voir la preuve de ce résultat dans les références [4] [5]) :

Lemme 5. Les valeurs propres des matrices A_{1D} , A_{2D} et A_{3D} ont pour expressions :

$$\begin{aligned} \lambda_k(A_{1D}) &= 4 \sin^2\left(\frac{k\pi}{2(n+1)}\right) = 4 \sin^2\left(\frac{k\pi h}{2}\right), \quad k = 1, 2, \dots, n \\ \lambda_{k,\ell}(A_{2D}) &= 4 \left[\sin^2\left(\frac{k\pi}{2(n+1)}\right) + \sin^2\left(\frac{\ell\pi}{2(n+1)}\right) \right] \\ &= 4 \left[\sin^2\left(\frac{k\pi h}{2}\right) + \sin^2\left(\frac{\ell\pi h}{2}\right) \right], \quad k, \ell = 1, 2, \dots, n \\ \lambda_{k,\ell,m}(A_{3D}) &= 4 \left[\sin^2\left(\frac{k\pi}{2(n+1)}\right) + \sin^2\left(\frac{\ell\pi}{2(n+1)}\right) \right. \\ &\quad \left. + \sin^2\left(\frac{m\pi}{2(n+1)}\right) \right] = 4 \left[\sin^2\left(\frac{k\pi h}{2}\right) + \sin^2\left(\frac{\ell\pi h}{2}\right) \right. \\ &\quad \left. + \sin^2\left(\frac{m\pi h}{2}\right) \right], \quad k, \ell, m = 1, 2, \dots, n \end{aligned}$$

Les composantes des vecteurs propres V associés ont pour expressions :

$$v_i^k(A_{1D}) = \sin\left(\frac{\pi i k}{n+1}\right), \quad i, k = 1, \dots, n$$

$$v_{i,j}^{k,\ell}(A_{2D}) = \sin\left(\frac{\pi i k}{n+1}\right) \sin\left(\frac{\pi j \ell}{n+1}\right), \quad i, j, k, \ell = 1, \dots, n$$

$$v_{i,j,p}^{k,\ell,m}(A_{3D}) = \sin\left(\frac{\pi i k}{n+1}\right) \sin\left(\frac{\pi j \ell}{n+1}\right) \sin\left(\frac{\pi p m}{n+1}\right), \quad i, j, p, k, \ell, m = 1, \dots, n$$

Pour terminer sur les questions de valeurs propres associées aux matrices issues de la discréttisation des équations aux dérivées partielles, considérons les matrices suivantes :

$$B = \begin{pmatrix} a & b & & & & \\ c & a & b & & & \\ c & a & b & & & \\ \cdot & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ c & a & b & & & \\ c & a & & & & \end{pmatrix}$$

et :

$$A = \begin{pmatrix} B & dI & & & & \\ eI & B & dI & & & \\ eI & B & dI & & & \\ \cdot & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ eI & B & dI & & & \\ eI & B & & & & \end{pmatrix}$$

Ces matrices sont issues de la discréttisation du problème de convection-diffusion avec conditions aux limites de Dirichlet homogènes défini soit dans le segment unité, soit dans le carré unité : pour ces problèmes, on a l'expression suivante des valeurs propres de ces matrices :

$$\lambda_k(B) = a - 2\sqrt{bc} \cos\left(\frac{k\pi}{n+1}\right), \quad k = 1, \dots, n$$

$$\lambda_{k,\ell}(A) = a - 2\sqrt{bc} \cos\left(\frac{k\pi}{n+1}\right) - 2\sqrt{de} \cos\left(\frac{\ell\pi}{n+1}\right), \quad k, \ell = 1, \dots, n$$

Les composantes des vecteurs propres associés sont identiques à celles indiquées au lemme 5. Il convient de noter que la forme simple du domaine permet d'effectuer cette détermination par un calcul direct. Dans le cas de domaine quelconque, comme ceux que nous évoquerons au paragraphe 6, une détermination numérique des valeurs propres s'avère nécessaire.

Avant de donner des résultats plus généraux, indiquons une application importante du résultat indiqué dans le lemme 5 précédent ; en effet, comme cela a été indiqué précédemment, les nombres réels sont mal représentés en machine dans la mesure où ils sont codés sur un nombre fini de « digits » (nombres binaires). Cette représentation approximative des nombres réels a pour conséquence que les calculs en machine sont entachés d'erreurs de chute ou de troncature. Lorsque le nombre d'opérations arithmétiques devient très grand, on peut s'interroger sur la précision des résultats obtenus et se demander si l'accumulation des erreurs de chute ou de troncature ne va pas dénaturer la solution obtenue numériquement. Pour essayer d'analyser ce type de phénomène, très courant en calcul scientifique, les numériciens ont introduit un indicateur d'inversion numérique, constitué par le **nombre de conditionnement** $C(A)$ associé à une matrice carrée A .

Définition 1. Soit A une matrice carrée de dimension n . On appelle norme matricielle induite le nombre :

$$\|A\| = \max_{x \neq 0} \left(\frac{\|Ax\|}{\|x\|} \right)$$

où $\|\cdot\|$ désigne une norme vectorielle quelconque de \mathbb{R}^n .

Remarque

On vérifie immédiatement que :

$$\|A\| = \max_{\|x\|=1} (\|Ax\|)$$

De plus on a également :

$$\|Ax\| \leq \|A\| \|x\|, \forall x \in \mathbb{R}^n$$

Pour une matrice carrée quelconque, on note $\|A\|_1$, $\|A\|_2$, $\|A\|_\infty$ les normes matricielles induites par les normes vectorielles $\|x\|_1 = \sum_{i=1}^n |x_i|$, $\|x\|_2 = \sqrt{\left(\sum_{i=1}^n |x_i|^2\right)}$ (norme euclidienne) et $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ (norme uniforme). On a alors la caractérisation suivante (cf. références [6] [7]) :

Lemme 6. Pour toute matrice carrée A , d'ordre n , on a :

$$\begin{aligned} \|A\|_1 &= \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{i,j}| \right) \\ \|A\|_2 &= \sqrt{\rho(A^t A)} \end{aligned}$$

et

$$\|A\|_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{i,j}| \right)$$

avec $\rho(B) = \max_{1 \leq i \leq n} (|\lambda_i(B)|)$ rayon spectral de la matrice B , c'est-à-dire la valeur propre de plus grand module de la matrice considérée.

Corollaire 2. Si A est une matrice carrée symétrique ou, plus généralement, hermitienne, on a :

$$\|A\|_2 = \rho(A)$$

Dans la mesure où l'on connaît les valeurs propres des matrices A_{1D} , A_{2D} et A_{3D} , le résultat du corollaire 2 est le plus intéressant pour déterminer les normes matricielles de ces dernières. On peut à présent définir le nombre de conditionnement $C(A)$ de la matrice A .

Définition 2. Soit $\|\cdot\|$ une norme matricielle ; le conditionnement d'une matrice régulière A , associée à cette norme est le nombre :

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

Nous noterons $C_2(A) = \|A\|_2 \|A^{-1}\|_2$, le conditionnement associé à la norme matricielle induite par la norme euclidienne.

On peut énoncer le résultat suivant [6] [7] :

Théorème 1.

1) Si le conditionnement est calculé avec une norme induite, alors $C(A) \geq 1$.

2) Si μ_{\max} et μ_{\min} sont respectivement la plus grande et la plus petite valeur singulière de la matrice A , alors :

$$C_2(A) = \frac{\mu_{\max}}{\mu_{\min}}$$

Il résulte du corollaire 2 et du point 2 du théorème précédent le résultat suivant :

Corollaire 3. Si A est une matrice symétrique ou, plus généralement, hermitienne, on a :

$$C_2(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$$

avec $\lambda_i(A)$, $i = 1, 2, \dots, \dim(A)$ les valeurs propres de la matrice A .

Remarque

Sur les questions de conditionnement, les définitions et résultats précédents restent axiomatiques. Cependant, cette notion de conditionnement a un grand intérêt pratique. Nous dirons que la matrice A est bien conditionnée si son nombre de conditionnement n'est pas beaucoup plus grand que l'unité ; dans le cas contraire, si $C(A)$ est grand, la matrice A est mal conditionnée, auquel cas les calculs effectués pour la résolution du système linéaire $AU = F$ sont sensibles à la propagation des erreurs de chute ou de troncature.

Considérons à présent la résolution du système $A_{1D} U = F$, pour la résolution numérique du problème de Poisson monodimensionnel avec conditions aux limites de Dirichlet homogènes ; remarquons que $\lambda_{\max}(A_{1D}) = \lambda_n(A_{1D})$ et que $\lambda_{\min}(A_{1D}) = \lambda_1(A_{1D})$; donc :

$$C_2(A_{1D}) = \frac{\lambda_n(A_{1D})}{\lambda_1(A_{1D})} \rightarrow +\infty, \text{ si } n \rightarrow \infty$$

Donc si n tend vers l'infini (c'est-à-dire si h tend vers zéro), la matrice A_{1D} devient mal conditionnée, et la propagation des erreurs de chute ou de troncature peut dénaturer complètement le

résultat calculé numériquement. Ce résultat est à rapprocher de celui du lemme 1, qui précisait que l'approximation par différences finies de l'opérateur laplacien était d'autant meilleure que h était petit, c'est-à-dire que n était grand. Notons que, pour les matrices A_{2D} et A_{3D} , le nombre de conditionnement associé à la norme matricielle induite par la norme euclidienne a le même comportement lorsque le pas de discréttisation h tend vers zéro.

4.2 Critères pratiques d'inversibilité

Nous avons établi, au début de ce paragraphe, l'inversibilité des matrices A_{1D} , A_{2D} et A_{3D} ; comme indiqué précédemment, ces matrices sont les matrices de discréttisation du problème de Poisson, avec conditions aux limites de Dirichlet homogènes, dans des domaines Ω réguliers constitués par le segment unité, le carré unité ou le cube unité. Or lorsque les problèmes sont définis en dimension 2 ou 3, les domaines Ω correspondants peuvent être de forme quelconque et les critères d'inversibilité, basés sur le calcul du déterminant de la matrice, sur la vérification directe de la propriété de définit positivité de cette dernière ou encore sur le calcul des valeurs propres, peuvent être difficiles à mettre en œuvre, compte tenu de la complexité des calculs à effectuer.

- Dans ce paragraphe, nous allons donner des **propriétés supplémentaires des matrices** A_{1D} , A_{2D} et A_{3D} , vérifiables avec le minimum de calculs et qui assureront l'**inversibilité des matrices de discréttisation**.

Définition 3. Les matrices A_{1D} , A_{2D} et A_{3D} ont une structure bande, c'est-à-dire qu'il existe un nombre entier positif ℓ tel que $a_{ij} = 0$ pour $|i - j| > \ell$. L'entier ℓ s'appelle la **demi-largeur de bande**.

Remarque

On verra que la demi-largeur de bande dépend de la numérotation des points du maillage. Pour les problèmes modèles considérés précédemment, on obtient une telle structure de matrice à cause de la forme particulièrement simple du domaine Ω ainsi qu'également à cause de la numérotation lexicographique des points du maillage. Dans le cas de domaines quelconques, la matrice de discréttisation a une structure creuse, essentiellement à cause de la construction du schéma de discréttisation.

Définition 4. Une matrice A est diagonale dominante stricte si les conditions suivantes sont vérifiées :

$$|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|, \quad \forall i \in \{1, \dots, \dim(A)\}$$

À partir de cette définition, on obtient un critère d'inversibilité de la matrice, particulièrement simple [7] :

Théorème 2. Une matrice à diagonale dominante stricte est inversible.

Le résultat précédent ne s'applique pas strictement dans le cas du problème modèle ; en effet, la matrice de discréétisation n'est pas diagonale dominante stricte. Par contre, ce résultat s'applique au cas du problème tridimensionnel suivant :

$$\begin{cases} -\Delta u + qu = f, \text{ dans } \Omega, (q > 0) \\ u = 0, \text{ sur } \partial\Omega \end{cases}$$

car, dans ce cas, $a_{i,i} = 6 + qh^2$, et pour $i \neq j$, soit $a_{i,j} = -1$, soit $a_{i,j} = 0$, la matrice considérée ayant de plus la même structure que la matrice du problème de Poisson avec conditions aux limites de Dirichlet homogènes. On vérifie aisément dans ce cas la propriété de diagonale dominance stricte de la matrice de discréttisation qui est donc inversible. On obtiendrait un résultat similaire pour les matrices de discréttisation de problème analogue défini dans un domaine monodimensionnel et bidimensionnel. Dans le cas du problème modèle, on peut retrouver un résultat analogue d'inversibilité des matrices de discréttisation en mettant en évidence d'autres propriétés de ces matrices, systématiquement vérifiables avec le minimum de calculs.

Définition 5. Une matrice A est diagonale dominante si la condition suivante est vérifiée :

$$|a_{i,i}| \geq \sum_{j \neq i} |a_{i,j}|, \quad \forall i \in \{1, \dots, \dim(A)\}$$

Définition 6. Une matrice A est diagonale fortement dominante si A est diagonale dominante et s'il existe au moins un indice k pour lequel l'inégalité :

$$|a_{k,k}| > \sum_{j \neq k} |a_{k,j}|$$

est vérifiée.

On vérifie alors aisément que les matrices de discréétisation A_{1D} , A_{2D} et A_{3D} sont diagonales dominantes et même fortement diagonales dominantes. Cependant, ces propriétés de diagonale dominance ou de diagonale dominance forte n'assurent pas l'inversibilité de ces matrices ; en effet, ces dernières possèdent de plus la propriété d'irréductibilité définie ci-dessous :

Définition 7. Une matrice réelle ou complexe A est réductible, s'il existe une matrice de permutation P , de même dimension, telle que :

$$PAP^t = \begin{pmatrix} B_{1,1} & B_{1,2} \\ 0 & B_{2,2} \end{pmatrix}$$

où les matrices $B_{i,i}$, pour $i = 1, 2$, sont des matrices carrées. Si une matrice n'est pas réductible, elle est irréductible.

La difficulté réside à présent dans la vérification de ce type de propriété. On dispose pour cela d'un résultat de caractérisation des matrices irréductibles [8] :

Lemme 7. Une condition nécessaire et suffisante pour qu'une matrice A soit irréductible est que, pour tout couple d'indices (i, j) , $i \neq j$, il existe au moins un ensemble d'indices i_1, i_2, \dots, i_k ($i_k = j$), avec $k \geq 1$, tels que les éléments $a_{i_1, i_1}, a_{i_1, i_2}, \dots, a_{i_{k-1}, j}$ soient tous différents de zéro.

Remarque

Il faut bien noter que, dans la suite d'éléments $a_{i_1, i_1}, a_{i_1, i_2}, \dots, a_{i_{k-1}, j}$ le numéro de ligne d'un élément est égal au numéro de colonne de son prédecesseur.

Exemple 1 : la matrice

$$A = \begin{pmatrix} a_1 & b_1 \\ c_1 & a_2 & b_2 \\ & \ddots & & \\ & \ddots & \ddots & \\ & & \ddots & b_{n-1} \\ & & & c_n & a_n \end{pmatrix}$$

telle que les coefficients a_i, b_i, c_i sont tous différents de zéro est irréductible. Si l'un, au moins, des coefficients b_i ou c_i est nul, alors la matrice A est réductible. Pour vérifier que la matrice est irréductible, on applique le lemme précédent ; en effet de deux choses l'une :

— si $1 \leq i < j \leq n$, alors les éléments de la partie triangulaire supérieure de la matrice A , à savoir $A_{i, i+1} = b_i, A_{i+1, i+2} = b_{i+1}, \dots, A_{j-1, j} = b_{j-1}$ satisfont les conditions du lemme 7 ;

— si $1 \leq j < i \leq n$, alors les éléments de la partie triangulaire inférieure de la matrice A , à savoir $A_{i, i-1} = c_i, A_{i-1, i-2} = c_{i-1}, \dots, A_{j+1, j} = c_{j+1}$ satisfont les conditions du lemme 7.

Donc si les éléments codiagonaux de la matrice A sont tous différents de zéro, la matrice est bien irréductible.

En fait, si on reprend la grille constituée par les points de discrémination du problème monodimensionnel, on constate que, pour la numérotation lexicographique des points du maillage, les seuls coefficients de la matrice différents de zéro représentent la liaison qui existe entre n'importe quel point du maillage et ses deux voisins respectivement gauche et droit. L'irréductibilité traduit donc le fait que, dans le cas d'un problème aux limites monodimensionnel, il est toujours possible de trouver un chemin discret reliant n'importe quel point du maillage à n'importe quel autre point du maillage, ce qui, pour les équations aux dérivées partielles discrétisées, définit un critère pratique d'irréductibilité.

Exemple 2 : on considère soit la matrice A_{2D} soit la matrice A_{3D} issues respectivement de la discrémination par différences finies du problème de Poisson défini soit dans le carré unité soit dans le cube unité. Comme dans l'étude précédente, il est encore possible de mettre en évidence une suite de coefficients non nuls satisfaisant les hypothèses du lemme 7, ce qui assure l'irréductibilité de la matrice. Cette vérification ne présente pas de difficultés majeures ; cependant, ce procédé est moins commode à mettre en œuvre. Il est peut-être plus commode de retourner aux schémas de discrémination associés aux problèmes considérés, ainsi qu'aux grilles de points correspondantes définies dans les domaines Ω considérés. Pour une numérotation donnée des points du maillage, la propriété d'irréductibilité des matrices considérées découle immédiatement en considérant un chemin discret reliant n'importe quel point du maillage à n'importe quel autre point de la grille et dont les indices correspondent justement aux indices de lignes et de colonnes figurant dans la suite de coefficients non nuls satisfaisant les hypothèses du lemme 7.

On complète les définitions précédentes par la notion suivante :

Définition 8. Une matrice A est diagonale dominante irréductible si A est irréductible et diagonale fortement dominante.

Ces rappels étant effectués, on peut énoncer le résultat suivant permettant de conclure à la régularité des matrices de discrémination diagonale dominante irréductibles [8] :

Théorème 3. Une matrice diagonale dominante irréductible est inversible.

On a donc vérifié la régularité des matrices A_{1D} , A_{2D} et A_{3D} en effectuant le minimum de calculs.

À partir des résultats précédents, on peut indiquer également des critères simples à mettre en œuvre, permettant de vérifier que les matrices A_{1D} , A_{2D} et A_{3D} associées au problème de Poisson sont définies positives.

Théorème 4. Soit A une matrice à diagonale dominante stricte ou à diagonale dominante irréductible ; si, de plus, ses éléments diagonaux sont strictement positifs, alors :

$$\operatorname{Re}(\lambda_i(A)) > 0, \forall i \in \{1, 2, \dots, \dim(A)\}$$

où les nombres $\lambda_i(A)$ sont les valeurs propres de la matrice A et $\operatorname{Re}(\lambda_i(A))$ désigne la partie réelle des valeurs propres $\lambda_i(A)$.

Corollaire 4. Soit A une matrice symétrique réelle, de coefficients diagonaux strictement positifs et à diagonale dominante stricte ou à diagonale dominante irréductible ; alors A est une matrice définie positive.

Pour terminer ce sous-paragraphe, indiquons quelques propriétés des matrices symétriques définies positives ainsi que des matrices définies positives, qui nous serviront à établir simplement les majorations d'erreur.

Proposition 2. Soit A une matrice symétrique définie positive ; on a alors l'inégalité suivante :

$$\alpha \|U\|_2^2 \leq \langle AU, U \rangle \leq \beta \|U\|_2^2, \forall U \in \mathbb{R}^n$$

avec α plus petite valeur propre de la matrice A ,
 β plus grande valeur propre (ou encore norme matricielle) de A .

En effet, on sait que d'une part les valeurs propres de la matrice A sont réelles et strictement positives ; d'autre part on sait aussi que ses vecteurs propres sont orthogonaux. Si on considère la fonctionnelle $U \in \mathbb{R}^n \rightarrow J(U) = \langle AU, U \rangle$, où $\langle \cdot, \cdot \rangle$ dénote le produit scalaire standard, on obtient alors facilement les inégalités cherchées en décomposant le vecteur U dans la base des vecteurs propres.

Proposition 3. Soit A une matrice définie positive, on a alors l'inégalité suivante :

$$\alpha \|U\|_2^2 \leq \langle AU, U \rangle \leq \beta \|U\|_2^2, \forall U \in \mathbb{R}^n$$

avec β norme matricielle de la matrice A ,
 α nombre réel strictement positif.

Cette propriété découle d'un raisonnement simple de compacité. En effet, soit :

$$S = \{U \in \mathbb{R}^n, \|U\|_2 = 1\}$$

la boule unité de \mathbb{R}^n et Q l'ensemble des sous-espaces vectoriels de \mathbb{R}^n . Soit $S_Q = S \cap Q$, l'intersection des ensembles S et Q . Le produit scalaire étant une fonction continue, l'application $U \mapsto J(U)$ de \mathbb{R}^n dans \mathbb{R} est continue ; l'ensemble S_Q étant fermé et borné, la restriction de J à S_Q est également continue ; donc il existe un vecteur $U_0 \in S_Q$ tel que $J(U_0) = \alpha$ réalise le minimum de $J(U)$ sur l'ensemble S_Q ; la matrice A étant définie positive, on a alors $J(U_0) = \alpha > 0$. Soit alors un vecteur $V \in Q$, $V \neq 0$; posons :

$$U = \frac{V}{\|V\|_2} \in S_Q$$

donc :

$$\langle AU, U \rangle = \langle A \frac{V}{\|V\|_2}, \frac{V}{\|V\|_2} \rangle \geq J(U_0) = \alpha$$

ce qui entraîne $\langle AV, V \rangle \geq \alpha \|V\|_2^2$. L'autre inégalité découle d'une part de l'inégalité de Schwarz et d'autre part de la définition de la norme matricielle.

Remarque

L'inégalité précédente traduit une propriété de coercivité discrète.

Définition 11. On dit que le schéma numérique est d'ordre p , s'il existe une constante C , indépendante de h , telle que $E \leq C h^p$.

Remarque

Plus la valeur de p sera élevée, plus l'approximation par différences finies sera meilleure.

On peut établir simplement le résultat suivant :

Lemme 8. Si la solution exacte du problème de Poisson, avec conditions aux limites de Dirichlet homogènes, est quatre fois continûment différentiable, on a $E \leq C h^2$, avec :

$$C = \frac{M_4}{12}, \text{ où } M_4 = \text{Sup} \left(\left| \frac{d^4 u(x)}{dx^4} \right|, 0 \leq x \leq 1 \right)$$

et le schéma numérique :

$$\begin{cases} \frac{-u_{j+1} + 2u_j - u_{j-1}}{h^2} = f_j = f(x_j), & 1 \leq j \leq n \\ u_0 = u_{n+1} = 0 \end{cases}$$

est d'ordre 2.

Définition 12. On appelle erreur au point x_j la quantité $e_j = u(x_j) - u_j$, $1 \leq j \leq n$.

La définition de l'erreur de troncature permet d'écrire :

$$\frac{-u(x_{j+1}) + 2u(x_j) - u(x_{j-1})}{h^2} = f(x_j) + E_j, \quad 1 \leq j \leq n$$

alors que celle du schéma numérique conduit à la relation :

$$\frac{-u_{j+1} + 2u_j - u_{j-1}}{h^2} = f(x_j), \quad 1 \leq j \leq n$$

En soustrayant membre à membre ces deux relations, on en déduit que :

$$\begin{cases} \frac{-e_{j+1} + 2e_j - e_{j-1}}{h^2} = E_j, & 1 \leq j \leq n \\ e_0 = e_{n+1} = 0 \end{cases}$$

donc les composantes de l'erreur sont solutions du schéma numérique avec comme second membre les composantes de l'erreur de troncature. On déduit alors, du lemme 8, le résultat de majoration d'erreur suivant :

Théorème 5. On suppose que la solution du problème de Poisson, avec conditions aux limites de Dirichlet homogènes, est quatre fois continûment différentiable. On a alors l'estimation suivante :

$$\|e\|_2 \leq \frac{M_4}{12 \alpha} h^3$$

avec $\|e\|_2$ norme euclidienne du vecteur erreur e dont les composantes sont égales à e_j .

En effet, soit \mathbf{E} le vecteur de composantes E_j ; le vecteur e étant solution du système :

$$A_{1D} e = h^2 \mathbf{E}$$

En multipliant scalairement la relation précédente par e , compte tenu d'une part du résultat de la proposition 2 et d'autre part de l'inégalité de Cauchy-Schwarz, on obtient :

$$\|e\|_2 \leq \frac{h^2}{\alpha} \|\mathbf{E}\|_2 \leq \frac{\sqrt{n}}{\alpha} Ch^4 \leq \frac{C}{\alpha} h^3$$

car $\sqrt{n} \leq n+1$, ce qui établit le résultat.

Remarque

On constate que $\|e\|_2 \rightarrow 0$, lorsque $h \rightarrow 0$; ainsi la norme de l'erreur tend vers zéro avec h et, théoriquement, la précision augmente; par contre, dans cette même situation, le système linéaire $A_{1D} U = F$ est mal conditionné, ce qui peut nuire à la précision.

Ces résultats sont transposables au cas de l'analyse d'erreur pour l'approximation du problème de Poisson avec conditions aux limites de Dirichlet homogènes, **dans le cas où Ω est le carré ou le cube unité**. Nous traitons ci-dessous le premier cas et laissons au lecteur le soin d'étudier le second; en fait l'étude consiste en une adaptation des concepts précédents à cette situation. Nous donnons rapidement la définition de l'erreur de troncature dans le cas bidimensionnel :

Définition 13. On considère le problème de Poisson défini dans le carré unité, avec conditions aux limites de Dirichlet homogènes; on appelle erreur de troncature du schéma numérique au point de discréttisation $M_{j,k} = (x_j, y_k)$ la quantité $E_{j,k}$ définie par :

$$E_{j,k} = \frac{-u(M_{j+1,k}) - u(M_{j,k+1}) + 4u(M_{j,k}) - u(M_{j,k-1}) - u(M_{j-1,k})}{h^2} - f(M_{j,k}), \quad 1 \leq j, k \leq n$$

Définition 14. Le schéma numérique est consistant si la quantité $E = \max_{1 \leq j, k \leq n} (E_{j,k})$ tend vers zéro avec h .

Lemme 9. Si la solution exacte du problème de Poisson, avec conditions aux limites de Dirichlet homogènes, est quatre fois continûment différentiable par rapport à ses arguments, on a $E \leq C h^2$, avec :

$$C = \frac{M_4}{12}, \text{ où } M_4 = \operatorname{Sup} \left(\left| \frac{\partial^4 u(M)}{\partial x^4} \right|, M \in \bar{\Omega} \right) + \operatorname{Sup} \left(\left| \frac{\partial^4 u(M)}{\partial y^4} \right|, M \in \bar{\Omega} \right)$$

et le schéma numérique :

$$\begin{cases} \frac{-u_{j+1,k} - u_{j,k+1} + 4u_{j,k} - u_{j,k-1} - u_{j-1,k}}{h^2} = f_{j,k}, & 1 \leq j, k \leq n \\ u_{0,k} = u_{i,0} = u_{n+1,k} = u_{i,n+1} = 0, & 1 \leq i, k \leq n \end{cases}$$

est d'ordre 2.

Définition 15. On appelle erreur au point $M_{j,k} = (x_j, y_k)$ la quantité $e_{j,k}$ définie par :

$$e_{j,k} = u(x_j, y_k) - u_{j,k}, \quad 1 \leq j, k \leq n$$

Pour obtenir une majoration de la norme euclidienne du vecteur erreur e de composantes $e_{j,k}$, on procède comme dans le cas du problème monodimensionnel; en effet, en retranchant aux relations définissant l'erreur de troncature celles définissant le schéma numérique, on obtient aisément :

$$\begin{cases} \frac{-e_{j+1,k} - e_{j,k+1} + 4e_{j,k} - e_{j,k-1} - e_{j-1,k}}{h^2} = E_{j,k}, & 1 \leq j, k \leq n \\ e_{0,k} = e_{i,0} = e_{n+1,k} = e_{i,n+1} = 0, & 1 \leq i, k \leq n \end{cases}$$

et, comme précédemment, le vecteur e est solution du système $A_{2D} e = h^2 \mathbf{E}$, où \mathbf{E} est le vecteur de composantes $E_{j,k}$, $1 \leq j, k \leq n$. En effectuant le même type de calculs que ceux exposés précédemment, on obtient le résultat suivant :

Théorème 6. On suppose que la solution du problème de Poisson, avec conditions aux limites de Dirichlet homogènes, est quatre fois continûment différentiable. On a alors l'estimation suivante :

$$\|e\|_2 \leq \frac{M_4}{12 \alpha} h^3$$

avec $\|e\|_2$ norme euclidienne du vecteur erreur e dont les composantes sont égales à $e_{j,k}$,

α plus petite valeur propre de la matrice A_{2D} .

Remarque

Pour l'analyse d'erreur concernant le problème de Poisson défini dans le cube unité, on obtiendrait, de la même façon, le même type d'estimation. Si l'on considère à présent, non plus le problème de diffusion, mais le problème de convection-diffusion défini dans le segment unité, le carré unité ou le cube unité, on obtiendrait le même type d'estimation, à condition toutefois que le schéma de discréttisation considéré permette d'obtenir une matrice A de discréttisation définie positive; en effet, cette matrice n'est plus symétrique. Cette propriété de définie-positivité peut se vérifier dans certains cas (par exemple, schéma centré) par un calcul direct; la majoration d'erreur se déduit aisément en utilisant le résultat de la proposition 3.

6. Discréttisation du problème dans le cas d'un domaine quelconque

Soit Ω un ouvert quelconque que, pour simplifier, nous considérerons inclus dans \mathbb{R}^2 . On considère toujours la résolution numérique du problème de Poisson avec conditions aux limites de Dirichlet homogènes sur la frontière $\partial\Omega$ de Ω , dont on rappelle la formulation :

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

Pour approcher le problème précédent par différences finies, la démarche est identique à celle exposée au paragraphe 3. Cependant, compte tenu du fait que le domaine Ω est quelconque, la

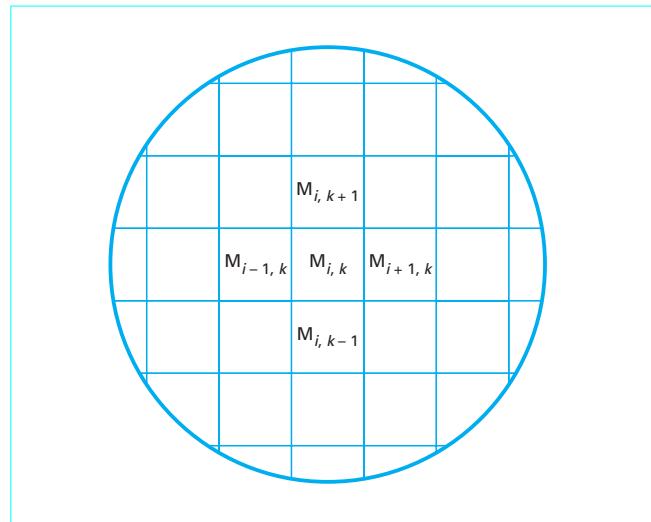


Figure 2 – Maillage différences finies non uniforme

grille de points sera formée de droites parallèles aux axes, équidistantes ou non, comme indiqué sur la figure 2. On va donc distinguer deux types de points de grille :

- les points M dont les quatre voisins dénotés N (Nord), S (Sud), E (Est), O (Ouest), situés à une distance h (pas de discrémination) du point considéré, appartiennent à $\bar{\Omega} = \Omega \cup \partial\Omega$;
- les points M dont au moins un des quatre voisins dénotés N (Nord), S (Sud), E (Est), O (Ouest), situés à une distance h du point considéré, se trouve en dehors de $\bar{\Omega}$.

Pour le premier type de points, on considérera le schéma numérique établi au paragraphe 3, que nous écrirons comme suit :

$$-u_E - u_N + 4u_M - u_S - u_O = h^2 f_M$$

Pour le second type de points, il faut définir un schéma numérique adapté à la situation ; pour cela, on se place dans le cas le plus général où les distances du point M à ses quatre voisins cardinaux, à présent notés M_i , $i=1$ à 4, sont toutes distinctes et notées par le pas h_i , $i=1$ à 4 (cf. figure 3). On suppose que la solution du problème est à présent trois fois continûment différentiable. Afin de définir une approximation de la dérivée seconde par rapport à la variable x , nous considérons les développements limités à l'ordre 3 de la solution u , aux points M_1 et M_3 ; on a alors :

$$u(M_1) = \left(u + h_1 \frac{\partial u}{\partial x} + \frac{h_1^2}{2} \frac{\partial^2 u}{\partial x^2} \right)(M) + \frac{h_1^3}{6} \frac{\partial^3 u(\bar{M}_1)}{\partial x^3}, \quad \bar{M}_1 \in]M, M_1[$$

$$u(M_3) = \left(u - h_3 \frac{\partial u}{\partial x} + \frac{h_3^2}{2} \frac{\partial^2 u}{\partial x^2} \right)(M) - \frac{h_3^3}{6} \frac{\partial^3 u(\bar{M}_3)}{\partial x^3}, \quad \bar{M}_3 \in]M_3, M[$$

Combinons ces deux relations afin d'éliminer la dérivée première par rapport à x au point M et divisons le résultat obtenu par le facteur $\frac{h_1 h_3}{2} (h_1 + h_3)$; il vient finalement :

$$\begin{aligned} -\frac{\partial^2 u(M)}{\partial x^2} &= \frac{-2}{h_1(h_1+h_3)} u(M_1) + \frac{2}{h_1 h_3} u(M) + \frac{-2}{h_3(h_1+h_3)} u(M_3) \\ &\quad + \frac{1}{3(h_1+h_3)} \left(h_1^2 \frac{\partial^3 u(\bar{M}_1)}{\partial x^3} - h_3^2 \frac{\partial^3 u(\bar{M}_3)}{\partial x^3} \right) \end{aligned}$$

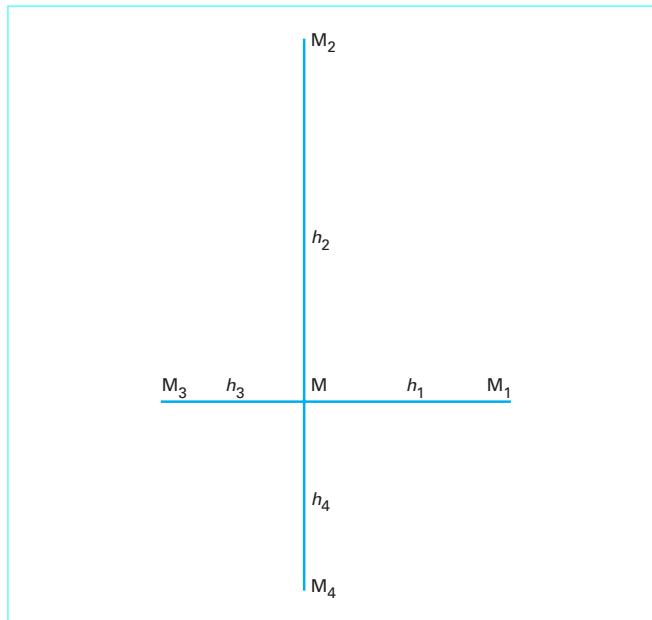


Figure 3 – Position des points dans un maillage non uniforme

De la même façon, pour approcher la dérivée seconde par rapport à la variable y , on obtient la relation suivante :

$$\begin{aligned} -\frac{\partial^2 u(M)}{\partial y^2} &= \frac{-2}{h_2(h_2+h_4)} u(M_2) + \frac{2}{h_2 h_4} u(M) + \frac{-2}{h_4(h_2+h_4)} u(M_4) \\ &\quad + \frac{1}{3(h_2+h_4)} \left(h_2^2 \frac{\partial^3 u(\bar{M}_2)}{\partial y^3} - h_4^2 \frac{\partial^3 u(\bar{M}_4)}{\partial y^3} \right) \end{aligned}$$

On en déduit le lemme suivant, permettant de donner une estimation du laplacien :

Lemme 10. Soit $\bar{h} = \max_{1 \leq i \leq 4} (h_i)$; supposons de plus que la solution du problème de Poisson avec conditions aux limites homogènes soit trois fois continûment différentiable. Alors, on a l'estimation suivante :

$$\left| \frac{2u(M_1)}{h_1(h_1+h_3)} + \frac{2u(M_2)}{h_2(h_2+h_4)} - \left(\frac{2}{h_1 h_3} + \frac{2}{h_2 h_4} \right) u(M) + \frac{2u(M_3)}{h_3(h_1+h_3)} + \frac{2u(M_4)}{h_4(h_2+h_4)} - \Delta u(M) \right| \leq \frac{2\bar{h}}{3} M_3$$

avec :

$$M_3 = \text{Sup} \left(\left| \frac{\partial^3 u(M)}{\partial x^3} \right|, M \in \bar{\Omega} \right) + \text{Sup} \left(\left| \frac{\partial^3 u(M)}{\partial y^3} \right|, M \in \bar{\Omega} \right)$$

Soit U le vecteur de composantes u_M , où $\{M\}$ est l'ensemble des points de grille intérieurs à Ω ; on définit par analogie avec les relations précédentes le schéma numérique de discrémination au point M de la grille :

$$\begin{aligned} \frac{-2u_{M_1}}{h_1(h_1+h_3)} + \frac{-2u_{M_2}}{h_2(h_2+h_4)} + \left(\frac{2}{h_1 h_3} + \frac{2}{h_2 h_4} \right) u_M + \frac{-2u_{M_3}}{h_3(h_1+h_3)} \\ + \frac{-2u_{M_4}}{h_4(h_2+h_4)} = f(M), \quad \forall M \end{aligned}$$

Il convient de noter qu'il y a autant de relations du type précédent qu'il y a de points M intérieurs à la grille ; comme précédemment on détermine la solution approchée U en résolvant un système linéaire $AU = F$.

Remarque

Si $h_1 = h_2 = h_3 = h_4 = h$, on retrouve le schéma classique.

On déduit de l'écriture du schéma de discréétisation le résultat suivant :

Lemme 11. La matrice A est une matrice tridiagonale par blocs, chaque bloc diagonal étant lui même tridiagonal. De plus la matrice A est symétrique, diagonale dominante irréductible ; les coefficients diagonaux de la matrice sont de plus strictement positifs.

Corollaire 5. La matrice A est définie positive et par conséquent régulière.

Remarque

La matrice A a une structure bande. Cependant les blocs hors diagonaux n'ont pas une structure diagonale simple analogue à celle rencontrée lors de la discréétisation du même problème lorsque le domaine Ω était le carré unité. Cela provient du fait que, le domaine Ω n'étant pas régulier, lorsqu'on le discréétise, le nombre de points lorsque l'on passe d'une ligne (respectivement colonne) à une ligne (respectivement colonne) voisine n'est pas le même.

Remarque

La discréétisation du même problème lorsque Ω est un domaine ouvert borné, inclus dans \mathbb{R}^3 , ne présente pas de difficultés majeures ; on obtient une matrice à structure par blocs comportant une diagonale positive et six codiagonales ; de plus, les blocs diagonaux ont une structure triple diagonale, les blocs hors diagonaux n'étant pas forcément diagonaux pour les mêmes raisons que celles exposées à la remarque précédente.

Remarque

Lorsque nous avons discréétisé le problème de Poisson dans le domaine Ω quelconque, nous avons supposé que la solution du problème était trois fois continûment différentiable, alors qu'auparavant on avait supposé qu'elle était quatre fois continûment différentiable. En fait, il n'est pas nécessaire de demander plus de régularité à cette solution car, dans tous les cas, l'erreur de troncature associée au schéma est en $o(\bar{h})$; en fait, la matrice de discréétisation étant symétrique définie positive, cette estimation de l'erreur de troncature nous permet d'avoir, comme précédemment, une estimation de la norme euclidienne d'erreur en $o(\bar{h}^2)$.

Remarque

On a constaté que la dimension de la matrice était égale au nombre de points de discréétisation intérieurs au maillage. On a donc à inverser des matrices bandes de grande dimension. Cela constitue une difficulté de résolution de tels systèmes algébriques linéaires que nous aborderons dans l'article [AF 502]. Cependant, si l'on considère une méthode directe de résolution, type méthode de Gauss, on verra au paragraphe 1 de cet article que le nombre d'opérations arithmétiques est proportionnel à $\dim(A)\ell^2$, où ℓ est la demi-largeur de bande de la matrice. On a donc intérêt à adopter une numérotation des points du maillage qui minimise cette demi-largeur de bande, afin de diminuer le nombre d'opérations arithmétiques, compte tenu, également, du mauvais conditionnement de la matrice A . Dans le cas d'un domaine Ω allongé, cette diminution de la demi-largeur de bande s'obtiendra aisément en numérotant les points du maillage dans le sens du côté le plus court ; on aura plus de blocs mais de tailles plus petites.

Exemple : dans le cas où Ω est un rectangle, une minimisation de la demi-largeur de bande s'obtient en numérotant les points du maillage suivant la largeur.

Pour de plus amples renseignements concernant cet article sur la méthode des différences finies pour les EDP stationnaires le lecteur pourra consulter les références [9] [10] [11] [12] [13] et [14].

Références bibliographiques

- [1] DAUTRAY (R.) et LIONS (J.-L.). – *Analyse mathématique et calcul numérique pour les sciences et les techniques*. Tome 1 à tome 9, Masson (1988).
- [2] FORSYTHE (G.) et WASOV (W.). – *Finite difference methods for partial differential equations*. Wiley (1960).
- [3] RICHTMYER (R.D.) et MORTON (K.W.). – *Difference methods for initial value problems*. John Wiley and Sons (1967).
- [4] GOLUB (G.H.) et MEURANT (G.A.). – *Résolution numérique des grands systèmes linéaires*. Eyrolles (1983).
- [5] MEURANT (G.). – *Computer solution of large linear systems*. North Holland (1999).
- [6] CIARLET (P.-G.). – *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques Appliquées, Masson (1982).
- [7] LASCAUX (P.) et THEODOR (R.). – *Analyse numérique matricielle appliquée à l'art de l'ingénieur*. Tomes 1 et 2, Masson (1986).
- [8] ORTEGA (J.) et RHEINBOLDT (W.). – *Iterative solution of nonlinear equations in several variables*. Academic Press (1970).
- [9] CUVELIER (C.), DESCLOUX (J.) et RAPPASZ (J.). – *Éléments d'équations aux dérivées partielles pour ingénieurs*. Tomes 1 et 2, Presses Polytechniques Romandes (1988).
- [10] EUVRARD (D.). – *Résolution numérique des équations aux dérivées partielles de la physique, de la mécanique et des sciences de l'ingénieur*. Masson (1988).
- [11] LE POURHIET (A.). – *Résolution numérique des équations aux dérivées partielles : une première approche*. Cepadues Edition (1988).
- [12] SIBONY (M.) et MARDON (J.C.). – *Approximations et équations différentielles*. Hermann (1982).
- [13] SAINSAULIEU (L.). – *Calcul scientifique*. Masson (1996).
- [14] DEBEAUMARCHÉ (G.). – *Introduction aux équations aux dérivées partielles linéaires*. AF 162 p. 1-21 (1999).

Dans ce traité des Techniques de l'Ingénieur