Mathématiques pour l'électricien

Méthodes numériques

par **Jacques-Hervé SAÏAC**
Ingénieur de l'École Centrale de Paris
Docteur et titulaire d'une habilitation à diriger des recherches
de l'Université Pierre-et-Marie-Curie (Paris VI)
Maître de Conférences au Conservatoire national des arts et métiers (CNAM)

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

Les méthodes numériques utiles à l'ingénieur sont nombreuses. Cependant, les algorithmes de base sont déjà exposés dans le traité Sciences fondamentales. Nous avons donc choisi de présenter ici les méthodes de discrétisation des équations de la physique en nous concentrant sur le modèle de la conduction.

Les mathématiques utilisent couramment les notions d'infini et de continu. La solution exacte d'un problème d'équations différentielles ou aux dérivées partielles est une fonction continue. Les ordinateurs ne connaissent que le fini et le discret. Les solutions approchées seront calculées en définitive comme des collections de valeurs discrètes sous la forme de composantes d'un vecteur solution d'un problème matriciel.

*En vue du passage d'un problème exact (**continu**) au problème approché (**discret**), on dispose de plusieurs techniques concurrentes : les différences finies, les éléments finis et les volumes finis. Chacune de ces trois méthodes correspond à une formulation différente des équations de la physique :*

— équilibre des forces en chaque point pour les différences finies ;
— minimisation de l'énergie ou principe des travaux virtuels pour les éléments finis ;
— loi de conservation et calcul des flux pour la méthode des volumes finis.

Nota : le lecteur pourra se reporter aux articles :

[A 1 220] **Méthodes numériques de base**.

[A 1 207] **Modèles et modélisation en électrotechnique**

et également aux articles [A 550] **Approximation des équations aux dérivées partielles. Méthodes aux différences finies** et [A 656] **Méthode des éléments finis**.

1. Principes généraux des méthodes numériques

1.1 Généralités

Examinons rapidement les avantages et les inconvénients de chacune de ces trois méthodes mentionnées dans l'Introduction.

1.1.1 Différences finies

La méthode des différences finies consiste à remplacer les dérivées apparaissant dans le problème continu par des différences divisées ou combinaisons de valeurs ponctuelles de la fonction en un nombre fini de points discrets ou nœuds du maillage.

Avantages :
— grande simplicité d'écriture ;
— faible coût de calcul.

Inconvénients :
— limitation de la géométrie des domaines de calculs ;
— difficultés de prise en compte des conditions aux limites ;
— en général, absence de résultats de majoration d'erreurs.

1.1.2 Éléments finis

La méthode des éléments finis consiste à approcher, dans un sous-espace de dimension finie, un problème écrit sous forme variationnelle (comme minimisation de l'énergie, en général) dans un espace de dimension infinie. La solution approchée est, dans ce cas, une fonction déterminée par un nombre fini de paramètres comme, par exemple, ses valeurs en certains points (les nœuds du maillage).

Avantages :
— traitement possible de géométries complexes ;
— détermination plus naturelle des conditions aux limites ;
— possibilité de démonstrations mathématiques de convergence et de majoration d'erreurs.

Inconvénients :
— complexité de mise en œuvre ;
— coût en temps de calcul et en mémoire.

1.1.3 Volumes finis

La méthode des volumes finis intègre, sur des volumes élémentaires de forme simple, les équations écrites sous forme de loi de conservation. Elle fournit ainsi de manière naturelle des approximations discrètes conservatives et est donc particulièrement bien adaptée aux équations de la mécanique des fluides : équation de conservation de la masse, équation de conservation de la quantité de mouvement, équation de conservation de l'énergie.

Sa mise en œuvre est simple si les « volumes » élémentaires sont des rectangles (ou des parallélépipèdes rectangles en dimension 3). Cependant, la méthode des volumes finis permet d'utiliser des volumes élémentaires de forme quelconque, donc de traiter des géométries complexes, ce qui est un **avantage** sur les différences finies. Il existe une grande variété de méthodes selon le choix de la géométrie des volumes élémentaires et des formules de calcul des flux.

Par contre, on dispose de peu de résultats théoriques de convergence.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

1.2 Un exemple de problème en dimension 1

Soit le problème :

$$\left.\begin{aligned} -u''(x) &= f(x) \qquad a < x < b \\ u(a) &= \alpha \\ u(b) &= \beta \end{aligned}\right\} \tag{1}$$

Interprétations physiques :
— barre élastique sous un chargement axial ;
— corde élastique soumise à un chargement transverse ;
— conduction électrique ou thermique dans une barre.

1.3 Approche différences finies

Toutes les méthodes numériques présupposent la discrétisation du domaine géométrique afin de passer d'un problème continu à une infinité d'inconnues à un problème discret ne comptant qu'un nombre fini d'inconnues.

Dans le cas des différences finies, on discrétise l'intervalle continu $[a, b]$ en un nombre fini de points x_i (figure **1**).

On remplace ainsi le problème continu par celui de la recherche de valeurs approchées u_i des solutions exactes $u(x_i)$ aux points x_i de la discrétisation. Mais on ne peut plus, dans ce cas, conserver les opérateurs de dérivation qui s'appliquent à des fonctions continues. On les remplace par des analogues discrets, les différences divisées ou différences finies.

Quelques formules simples d'approximation de la dérivée première par des différences divisées

- **Différence divisée progressive**

Le développement limité

$$u'(x_i) = \frac{u(x_i + h) - u(x_i)}{h} - \frac{h}{2} u''(\xi_i) \tag{2}$$

conduit à l'approximation suivante :

$$u'(x_i) = \frac{\mathrm{d}u}{\mathrm{d}x}(x_i) \approx \frac{u_{i+1} - u_i}{x_{i+1} - x_i} \tag{3}$$

- **Différence divisée régressive**

De même, le développement limité

$$u'(x_i) = \frac{u(x_i) - u(x_i - h)}{h} + \frac{h}{2} u''(\eta_i) \tag{4}$$

donne :

$$u'(x_i) = \frac{\mathrm{d}u}{\mathrm{d}x}(x_i) \approx \frac{u_i - u_{i-1}}{x_i - x_{i-1}} \tag{5}$$

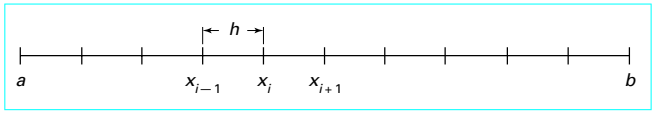


Figure 1 – Discrétisation en différences finies

- **Différence divisée centrée**

Le développement limité

$$u'(x_i) = \frac{u\left(x_i + \frac{h}{2}\right) - u\left(x_i - \frac{h}{2}\right)}{h} - \frac{h^2}{24} u'''(\theta_i) \tag{6}$$

conduit, dans le cas de discrétisations uniformes de pas constant h, à :

$$u'(x_i) = \frac{\mathrm{d}u}{\mathrm{d}x}(x_i) \approx \frac{u_{i+1} - u_{i-1}}{2h} \quad \text{ou} \quad \frac{u_{i+1/2} - u_{i-1/2}}{h} \tag{7}$$

On a noté $u_{i+1/2}$ et $u_{i-1/2}$ les valeurs approchées de u aux points $x_i + \frac{h}{2}$ et $x_i - \frac{h}{2}$ respectivement.

Formule simple d'approximation d'une dérivée seconde par une différence divisée centrée

Dans le cas particulier de points x_i régulièrement espacés d'un pas h uniforme, on retrouve, en utilisant :

$$u''(x_i) = \frac{u'\left(x_i + \frac{h}{2}\right) - u'\left(x_i - \frac{h}{2}\right)}{h} - \frac{h^2}{24} u^{(4)}(\theta_i) \tag{8}$$

et la relation (6), la discrétisation centrée classique de la dérivée seconde :

$$u''(x_i) = \frac{\mathrm{d}^2 u}{\mathrm{d}x^2}(x_i) \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} \tag{9}$$

On obtient ainsi le système d'équations linéaires suivant dont la résolution donne les valeurs u_i de la solution approchée du problème (1) :

$$\left.\begin{aligned} -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} &= f_i \qquad \text{pour } i = 1, N-1 \\ \text{avec } u_0 = \alpha \quad &\text{et} \quad u_N = \beta, \end{aligned}\right\} \tag{10}$$

ce qui s'écrit, sous forme matricielle :

$$\frac{1}{h^2}\begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & \cdots & -1 & 2 \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_i \\ \vdots \\ u_{N-2} \\ u_{N-1} \end{pmatrix} = \begin{pmatrix} f_1 + \alpha/h^2 \\ f_2 \\ \vdots \\ f_i \\ \vdots \\ f_{N-2} \\ f_{N-1} + \beta/h^2 \end{pmatrix} \tag{11}$$

Il ne reste alors plus qu'à résoudre ce système linéaire par des techniques standard de factorisation (méthodes de Gauss LU ou méthode de Choleski LL^T, § 3.1).

1.4 Approche éléments finis

La présentation très succincte faite ici n'est qu'une première introduction à la méthode des éléments finis. Elle a pour but de donner les idées de base dans un cas extrêmement simple.

On introduit tout d'abord un produit scalaire de deux fonctions selon :

$$(v, w) = \int_a^b v(x)\, w(x)\, \mathrm{d}x$$

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

et l'espace $L^2[a, b]$ des fonctions de carré sommable sur $[a, b]$, c'est-à-dire telles que l'intégrale suivante existe :

$$\int_a^b v^2(x)\ \mathrm{d}x$$

Soit $H^1[a, b]$ l'espace des fonctions v, de carré sommable et dont la dérivée est également de carré sommable :

$$\left.\begin{array}{l} v \in L^2[a, b] \\ v' \in L^2[a, b] \end{array}\right\}$$

et soit $H_0^1[a, b]$ l'espace des fonctions v de $H^1[a, b]$ nulles en a et b :

$$\left.\begin{array}{l} v \in L^2[a, b] \\ v' \in L^2[a, b] \\ v(a) = v(b) = 0 \end{array}\right\}$$

En posant, par exemple :

$$u(x) = \alpha\ \frac{b-x}{b-a} + \beta\ \frac{x-a}{b-a} + \tilde{u}(x) \qquad (12)$$

Le problème différentiel donné par l'équation (1):

$$\left.\begin{array}{ll} -u''(x) = f(x) & a < x < b \\ u(a) = \alpha & u(b) = \beta \end{array}\right\}$$

se ramène très simplement, par translation de fonction inconnue, au problème homogène suivant :

$$\left.\begin{array}{l} -\tilde{u}''(x) = f(x) \qquad \forall x \in [a, b] \\ (\ \tilde{u}(a) = \tilde{u}(b) = 0). \end{array}\right\} \qquad (13)$$

C'est sur ce problème que nous allons présenter, par souci de simplicité, la méthode. Multiplions l'équation (1) par $v(x)$ et intégrons sur $[a, b]$:

$$-\int_a^b \tilde{u}''(x)\ v(x)\ \mathrm{d}x = \int_a^b f(x)\ v(x)\ \mathrm{d}x \qquad (14)$$

Par intégration par parties, il vient :

$$-\int_a^b \tilde{u}''(x)\ v(x)\ \mathrm{d}x = \int_a^b \tilde{u}'(x)\ v'(x)\ \mathrm{d}x + \tilde{u}'(a)\ v(a) - \tilde{u}'(b)\ v(b) \qquad (15)$$

On obtient une nouvelle formulation du problème (13), dite **formulation variationnelle**, qui, en prenant en compte les conditions sur l'espace $H_0^1[a, b]$,

$$v(a) = v(b) = 0$$

s'écrit :

$$\left.\begin{array}{l} \text{Chercher la fonction } \tilde{u} \text{ appartenant à } H_0^1[a, b] \text{ telle que :} \\ \displaystyle\int_a^b \tilde{u}'(x)\ v'(x)\ \mathrm{d}x = \int_a^b f(x)\ v(x)\ \mathrm{d}x \qquad \forall v \in H_0^1[a, b] \end{array}\right\} \qquad (16)$$

Cette formulation est équivalente à la **minimisation** d'une forme quadratique représentant **l'énergie du système** qui s'écrit :

$$\left.\begin{array}{l} \text{Chercher la fonction } u \text{ vérifiant } u(a) = \alpha \quad u(b) = \beta \\ \text{qui réalise le minimum de la forme } J \text{ définie par :} \\ \displaystyle J(v) = \frac{1}{2}\int_a^b v'^2\ \mathrm{d}x - \int_a^b fv\ \mathrm{d}x \end{array}\right\} \qquad (17)$$

Pour s'en convaincre, il suffit de calculer $J(u + \lambda v)$

avec u solution du problème variationnel,
λ réel quelconque,
$v \in H_0^1[a, b]$ quelconque.

On obtient simplement, pour tout $\lambda \in \mathbb{R}$ et pour tout $v \in H_0^1[a, b]$, l'équivalence :

$$J(u + \lambda v) \geq J(u) \Leftrightarrow \int_a^b \tilde{u}'(x)\ v'(x)\ \mathrm{d}x = \int_a^b f(x)\ v(x)\ \mathrm{d}x \qquad (18)$$

Il y a ainsi trois formes équivalentes du problème :
— une forme différentielle (1) ;
— une forme variationnelle (16) (principe des travaux virtuels) ;
— une forme minimisation de l'énergie (17).

1.5 Un premier exemple simple : les éléments P1

1.5.1 Présentation

On approche l'espace $H_0^1[a, b]$ par l'espace $V_{0,h} \subset H_0^1[a, b]$ construit de la manière suivante.

On choisit une discrétisation de l'intervalle $[a, b]$ en N sous-intervalles ou éléments

$$K_i = [x_{i-1}, x_i]\ ;$$

les éléments K_i n'ont pas forcément même longueur (figure **2**) $V_{0,h}$ est alors l'espace des fonctions continues affines par morceaux (c'est-à-dire affines sur les segments K_i) et nulles aux extrémités a et b.

L'utilisation de fonctions affines, fonctions polynomiales de degré un, justifie la dénomination d'éléments P1. Chaque fonction $v_h \in V_{0,h}$ est déterminée de manière unique par la donnée de ses valeurs aux points x_i pour $i = 1, \ldots, N-1$ (figure **3**). L'espace $V_{0,h}$ est de dimension $N-1$.

1.5.2 Base de Lagrange

Considérons les $N-1$ fonctions $w_i \in V_{0,h}$ (figure **4**) définies par les $N-1$ conditions suivantes :

$$w_i(x_j) = \delta_{ij} \qquad \forall i = 1, \ldots, N-1 \text{ et } \forall j = 1, \ldots, N-1 \qquad (19)$$

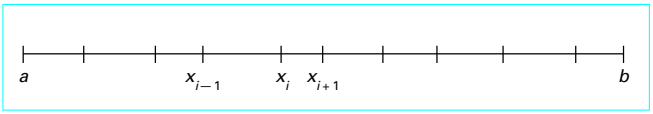


Figure 2 – Discrétisation (maillage) du segment [*a*, *b*] en éléments finis

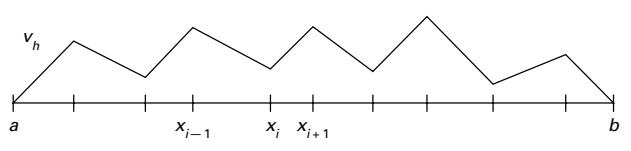


Figure 3 – Une fonction affine par morceaux

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

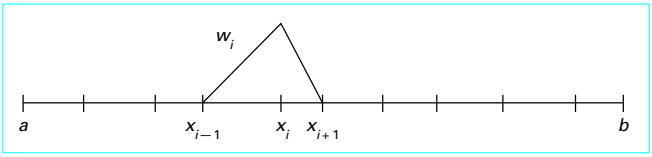


Figure 4 – Fonction de base de Lagrange

Ces $N-1$ fonctions forment une base de $V_{0,h}$; une fonction v_h quelconque s'écrit dans cette base :

$$v_h(x) = \sum_{i=1}^{i=N-1} v_i w_i(x) \qquad (20)$$

avec $v_i = v_h(x_i)$. Les coefficients v_i sont donc les valeurs de v_h aux points (x_i).

1.5.3 Écriture du problème approché

Écrivons le problème approché dans $V_{0,h}$:

$$\int_a^b u_h'(x)\, v_h'(x)\, \mathrm{d}x = \int_a^b f(x)\, v_h(x)\, \mathrm{d}x \qquad \forall v_h \in V_{0,h} \qquad (21)$$

Le problème étant linéaire, l'égalité est vraie pour tout v_h si elle est vraie pour une base de l'espace vectoriel $V_{0,h}$.

On peut donc remplacer dans la relation (21) :

$$\forall v_h \in V_{0,h} \text{ par } \forall w_i \text{ pour } i = 1, \ldots, N-1 .$$

D'autre part, écrivons u_h, solution du problème approché dans $V_{0,h}$, dans la base des w_i :

$$u_h(x) = \sum_{j=1}^{j=N-1} u_j w_j(x) \qquad (22)$$

avec $u_j = u_h(x_j)$ valeur approchée de la solution exacte au point (x_j).

On obtient l'écriture suivante du problème approché :

Trouver $u_1, u_2, \ldots, u_{N-1}$ tels que :

$$\int_a^b \left(\sum_{j=1}^{j=N-1} u_j w_j'(x) \right) w_i'(x)\mathrm{d}x = \int_a^b f(x) w_i(x)\mathrm{d}x \quad \forall i = 1, \ldots, N-1 \qquad (23)$$

ce que l'on peut récrire :

$$\sum_{j=1}^{j=N-1} \left(\int_a^b w_j'(x) w_i'(x)\mathrm{d}x \right) u_j = \int_a^b f(x) w_i(x)\mathrm{d}x \quad \forall i = 1, \ldots, N-1 \qquad (24)$$

En posant :

$$\int_a^b f(x) w_i(x)\mathrm{d}x = F_i \qquad (25)$$

et

$$\int_a^b w_j'(x) w_i'(x)\mathrm{d}x = A_{ij} \qquad (26)$$

il vient :

$$\sum_{j=1}^{j=N-1} A_{ij}\, u_j = F_i \qquad \forall i = 1, \ldots, N-1 \qquad (27)$$

On a ainsi obtenu un système linéaire de $N-1$ équations à $N-1$ inconnues, qui peut s'écrire sous la forme matricielle :

$$A\, U = F \qquad (28)$$

Chaque ligne d'indice i du système linéaire correspond au choix d'une fonction de base d'indice i de l'espace des fonctions tests $V_{0,h}$. Il y a autant d'équations que d'inconnues puisque la partie inconnue de la solution approchée s'exprime, elle aussi, dans la base des w_i.

1.5.4 Calcul des coefficients du système

On calcule les coefficients A_{ij} [relation (26)] de la matrice en sommant les contributions des différents éléments selon :

$$A_{ij} = \int_a^b w_j'(x) w_i'(x)\mathrm{d}x = \sum_{k=1}^{k=N} \int_{x_{k-1}}^{x_k} w_j'(x) w_i'(x)\mathrm{d}x \qquad (29)$$

Considérons l'élément $K_i = [x_{i-1}, x_i]$. Sur cet élément, il n'y a que deux fonctions de base non nulles : w_{i-1} et w_i :

$$\left.\begin{aligned} w_{i-1}|_{K_i} &= \frac{x_i - x}{x_i - x_{i-1}} \\ w_i|_{K_i} &= \frac{x - x_{i-1}}{x_i - x_{i-1}} \end{aligned}\right\} \qquad (30)$$

$$\left.\begin{aligned} w'_{i-1}|_{K_i} &= \frac{-1}{x_i - x_{i-1}} \\ w'_i|_{K_i} &= \frac{1}{x_i - x_{i-1}} \end{aligned}\right\} \qquad (31)$$

L'élément K_i produira donc effectivement une contribution non nulle aux quatre coefficients $A_{i-1,i-1}$, $A_{i-1,i}$, $A_{i,i}$ et $A_{i,i-1}$ de la matrice globale A.

Calculons les contributions élémentaires de K_i et disposons-les sous la forme d'une matrice élémentaire 2×2 :

$$\mathrm{Elem}_i = \begin{pmatrix} e^i_{1,1} & e^i_{1,2} \\ e^i_{2,1} & e^i_{2,2} \end{pmatrix} \qquad (32)$$

On trouve sans difficulté :

$$\mathrm{Elem}_i = \begin{pmatrix} \dfrac{1}{x_i - x_{i-1}} & \dfrac{-1}{x_i - x_{i-1}} \\ \dfrac{-1}{x_i - x_{i-1}} & \dfrac{1}{x_i - x_{i-1}} \end{pmatrix} = \frac{1}{x_i - x_{i-1}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \qquad (33)$$

Chaque composante F_i du vecteur second membre global [relation (25)]

$$F_i = \int_a^b f(x) w_i(x)\, \mathrm{d}x \qquad (34)$$

est calculée également par assemblage de contributions élémentaires :

$$F_i = \sum_{k=1}^{k=N} \int_{x_{k-1}}^{x_k} f(x) w_i(x)\, \mathrm{d}x \qquad (35)$$

Tout dépend alors de la donnée de f. Si f est donnée analytiquement, on peut parfois calculer les intégrales exactement à la main. Mais en général, f n'est connue que par ses valeurs aux points x_i pour $i = 0, N$.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

On écrit donc f dans la base des w_i selon :

$$f(x) = \sum_{j=0}^{j=N} f_j \, w_j(x) \mathrm{d}x \qquad (36)$$

Le problème se ramène alors au calcul des intégrales :

$$\int_{x_{k-1}}^{x_k} w_j(x) \, w_i(x) \mathrm{d}x \qquad (37)$$

On utilise des formules d'intégration numérique, par exemple :

— la **formule des trapèzes** :

$$\int_{x_{k-1}}^{x_k} F(x) \, \mathrm{d}x = \frac{x_k - x_{k-1}}{2} \left[F(x_{k-1}) + F(x_k) \right] \qquad (38)$$

exacte pour des polynômes de degré 1 ;

— la **formule de Simpson** :

$$\int_{x_{k-1}}^{x_k} F(x) \, \mathrm{d}x = \frac{x_k - x_{k-1}}{6} \left[F(x_{k-1}) + 4F\left(x_{k-\frac{1}{2}}\right) + F(x_k) \right] \qquad (39)$$

exacte pour des polynômes de degré inférieur ou égal à 3.

Avec des fonctions tests $w_i \in P1$, la méthode des trapèzes conduit à une valeur approchée de l'intégrale, qui, dans le cas de points équidistribués de pas h, redonne le résultat :

$$F_i = h \, f_i \qquad (40)$$

obtenu en différences finies.

La méthode de Simpson permet un calcul exact et donne dans le même cas :

$$F_i = \frac{h}{6} \left[f_{i-1} + 4 f_i + f_{i+1} \right] \qquad (41)$$

1.5.5 Assemblage de la matrice globale

Chaque coefficient A_{ij} est obtenu en sommant les contributions des éléments K_i. Remarquons que, seuls deux éléments produisent une contribution non nulle à chaque coefficient (§ 1.5.4). Ainsi, on vérifie que sur chaque ligne i de la matrice A il n'y a que 3 coefficients non nuls : $A_{i,\,i-1}$, $A_{i,\,i}$, $A_{i,\,i+1}$.

La règle d'assemblage est la suivante : on passe en revue les éléments K_i et on distribue leurs contributions à la matrice globale A aux positions appropriées. Dans notre cas simple monodimensionnel, le calcul complet peut être fait à la main.

Le coefficient $A_{i,i-1}$ provient uniquement de la contribution de l'élément K_i ; on a ainsi :

$$A_{i,\,i-1} = -\frac{1}{x_i - x_{i-1}} \qquad (42)$$

Le coefficient $A_{i,i+1}$ provient uniquement de la contribution de l'élément K_{i+1} :

$$A_{i,\,i+1} = -\frac{1}{x_{i+1} - x_i} \qquad (43)$$

Le coefficient $A_{i,i}$ est la somme des contributions des deux éléments adjacents au point x_i, K_i et K_{i+1} :

$$A_{i,\,i} = \frac{1}{x_i - x_{i-1}} + \frac{1}{x_{i+1} - x_i} \qquad (44)$$

Cela est valable $\forall i = 1$ à $N-1$, car tous les points x_i pour $i = 1$ à $N-1$ correspondant à des valeurs inconnues de la solution appartiennent, dans ce cas de conditions aux limites de Dirichlet, à deux segments adjacents.

Cas particulier : si les points sont régulièrement espacés d'un pas h, on retrouve, à un coefficient multiplicatif h près, la matrice obtenue dans la méthode des différences finies [relation (11)] :

$$A = \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & \cdots & -1 & 2 \end{pmatrix} \qquad (45)$$

1.6 Approche volumes finis

La méthode des volumes finis, comme celle des éléments finis, utilise une formulation intégrale des équations. Mais au lieu d'utiliser un produit scalaire de L^2 par des fonctions tests, on se contente d'intégrer les équations différentielles sur des volumes élémentaires ou volumes de contrôle. Cela peut s'interpréter comme l'utilisation de fonctions tests, fonctions indicatrices des volumes élémentaires.

À partir d'un maillage en volumes finis, où l'on prend les inconnues au centre x_i des volumes de contrôle qui sont, dans ce cas monodimensionnel, les segments $[x_{i-1/2}, x_{i+1/2}]$, on passe ainsi de :

$$-u''(x) = f(x) \Leftrightarrow -(u'(x))' = f(x) \qquad (46)$$

à

$$-\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (u'(x))' \mathrm{d}x = u'\left(x_{i-\frac{1}{2}}\right) - u'\left(x_{i+\frac{1}{2}}\right) = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x) \mathrm{d}x \qquad (47)$$

Les $u'\left(x_{i-\frac{1}{2}}\right)$ et $u'\left(x_{i+\frac{1}{2}}\right)$ apparaissent comme des flux aux interfaces des volumes élémentaires, les segments $[x_{i-1/2}, x_{i+1/2}]$.

Si on approche les dérivées selon :

$$\left.\begin{aligned} u'\left(x_{i+\frac{1}{2}}\right) &= \frac{u(x_i + h) - u(x_i)}{h} \\ \text{et} \qquad u'\left(x_{i-\frac{1}{2}}\right) &= \frac{u(x_i) - u(x_{i-h})}{h} \end{aligned}\right\} \qquad (48)$$

dans le cas particulier de points x_i régulièrement espacés d'un pas h uniforme, on retrouve la discrétisation centrée classique de la dérivée seconde [relation (9)] :

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

L'intérêt de cette technique de volumes finis n'apparaît que dans les applications aux problèmes en dimension 2 et 3.

De nombreuses variantes existent, selon la géométrie des volumes élémentaires, le choix du positionnement des inconnues aux centres ou aux sommets des volumes de contrôle, et les diverses formules de calcul des flux aux interfaces. Un exposé des techniques de volumes finis sortant du cadre de cet article, nous renvoyons en bibliographie au livre de C. Hirsch [1] et à la référence [2].

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

2. Problèmes en dimension deux et trois

2.1 Rappels

2.1.1 Opérateurs différentiels en dimension deux (et trois)

■ Soit u une fonction de 2 variables (x, y) à valeurs réelles définie sur un domaine Ω de $\mathbb{R}^2$. On appelle **gradient** (en dimension 2) de u et on note **grad**(u) ou ∇u, le vecteur :

$$\textbf{grad}\ (u) = \nabla u = \begin{pmatrix} \dfrac{\partial u}{\partial x} \\ \dfrac{\partial u}{\partial y} \end{pmatrix} \tag{49}$$

■ Soit $\boldsymbol{V}$ une fonction vectorielle de 2 variables (x, y) définie sur un domaine Ω de $\mathbb{R}^2$ et à valeurs V_1, V_2 dans $\mathbb{R}$. On appelle **divergence** du vecteur $\boldsymbol{V}$ et on note div$(\boldsymbol{V})$ ou $\nabla \cdot \boldsymbol{V}$, le scalaire :

$$\text{div}\ (\boldsymbol{V}) = \nabla \cdot \boldsymbol{V} = \frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} \tag{50}$$

■ Soit u une fonction de 2 variables (x, y) à valeurs réelles définie sur un domaine Ω de $\mathbb{R}^2$. On appelle **laplacien** de u et on note Δu ou $\nabla^2 u$, le scalaire :

$$\Delta u = \text{div}\ (\textbf{grad}\ (u)) = \nabla \cdot \nabla (u) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \tag{51}$$

■ Tout cela passe simplement en dimension trois. Nous nous plaçons pour la suite en dimension deux. Mais le passage à la dimension trois ne pose pas de problème conceptuel majeur. Il est évidemment techniquement plus difficile et pratiquement plus coûteux.

2.1.2 Formules de Green

■ **Formule de base**

Soit $\Omega \subset \mathbb{R}^2$ de frontière Γ C^1 par morceaux et u et v fonctions de $H^1(\Omega)$; on a l'égalité suivante pour chaque composante $i = 1, 2$:

$$\int\int_\Omega \frac{\partial u}{\partial x_i}\, v\, \mathrm{d}x_1 \mathrm{d}x_2 = -\int\int_\Omega \frac{\partial v}{\partial x_i}\, u\, \mathrm{d}x_1 \mathrm{d}x_2 + \int_\Gamma u n_i\, v\, \mathrm{d}\gamma \tag{52}$$

avec n_i i^{e} composante du vecteur normal unitaire à Γ orienté vers l'extérieur de Ω et noté n,

γ abscisse curviligne sur Γ orientée dans le sens direct.

■ **Formules déduites**

● En notant en caractère gras les vecteurs (en particulier $\boldsymbol{u}$ désigne ici un vecteur), on déduit aisément les formules suivantes :

$$\int\int_\Omega \text{div}\ \boldsymbol{u}\ v\ \mathrm{d}x\mathrm{d}y = -\int\int_\Omega \boldsymbol{u}\ \textbf{grad}\ v\ \mathrm{d}x\mathrm{d}y + \int_\Gamma \boldsymbol{un}\ v\mathrm{d}\gamma \tag{53}$$

et en posant $\boldsymbol{u} = a\ \textbf{grad}(u)$ où a est une fonction $C^1(\bar{\Omega})$

et $\dfrac{\partial u}{\partial n} = \textbf{grad}\ (u)\boldsymbol{n}$:

$$\int\int_\Omega \text{div}\ (a\ \textbf{grad}\ u) v\ \mathrm{d}x\mathrm{d}y$$

$$= -\int\int_\Omega a\ \textbf{grad}\, u\ \ \textbf{grad}\, v\ \mathrm{d}x\mathrm{d}y + \int_\Gamma a \frac{\partial u}{\partial n}\, v\ \mathrm{d}\gamma \tag{54}$$

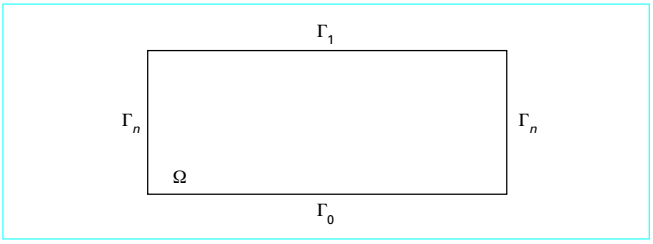


Figure 5 – Domaine de calcul pour le modèle de conduction électrique

● Enfin, dans le cas $\boldsymbol{a = 1}$, on obtient la formule classique :

$$-\int\int_\Omega \Delta u\ v\ \mathrm{d}x\mathrm{d}y = -\int\int_\Omega \textbf{grad}\ u\ \textbf{grad}\, v\ \mathrm{d}x\mathrm{d}y - \int_\Gamma \frac{\partial u}{\partial n}\, v\ \mathrm{d}\gamma \tag{55}$$

2.2 Modèle de la conduction

■ Nous renvoyons à l'article [A 1 207] *Modèles et modélisation en électrotechnique* pour la justification du modèle de la conduction électrique. Soit Ω le domaine de calcul (figure **5**).

On considère le **potentiel électrique** ψ. Il vérifie l'équation :

$$\left.\begin{aligned} -\text{div}\ (\sigma\ \textbf{grad}\ \psi)(x, y) &= 0 \qquad \forall x, y \in \Omega \\ \psi|_{\Gamma_0}(x, y) &= 0 \\ \psi|_{\Gamma_1}(x, y) &= 1 \\ \left.\frac{\partial \psi}{\partial n}\right|_{\Gamma_n}(x, y) &= 0 \end{aligned}\right\} \tag{56}$$

σ étant la conductivité électrique.

■ Le même type de problème se pose dans de nombreux domaines, en particulier en thermique, en élasticité, en mécanique des fluides ou en électromagnétisme. Nous regrouperons tous les problèmes précédents sous la **forme générale** :

$$\left.\begin{aligned} -\text{div}\ (\sigma\ \textbf{grad}\ u)(x, y) &= f(x, y) \qquad \forall x, y \in \Omega \\ u|_{\Gamma_d} &= u_d \\ \left.\frac{\partial u}{\partial n}\right|_{\Gamma_n} &= g \end{aligned}\right\} \tag{57}$$

avec Ω, ouvert borné de $\mathbb{R}^2$ de frontière $\Gamma = \Gamma_d \cup \Gamma_n$.

2.3 Approximation par différences finies

Comme en dimension un, la première étape consiste à discrétiser le domaine. C'est dans l'application aux problèmes bidimensionnels et tridimensionnels que la méthode des différences finies présente sa plus sévère limitation. En effet elle n'est bien adaptée qu'à la discrétisation de domaines rectangulaires ou parallélépipédiques par des maillages formés de grilles perpendiculaires (figure **6**), les dérivées partielles dans chaque direction d'axe étant approchées comme les dérivées en dimension un.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

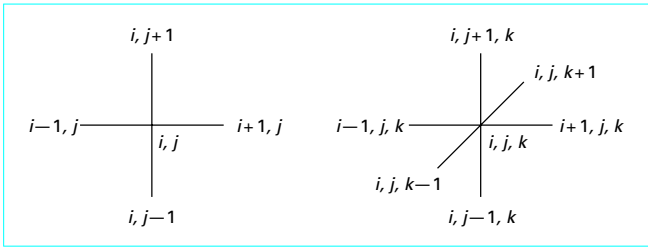


Figure 6 – Grilles différences finies bidimensionnelles et tridimensionnelles

2.3.1 Discrétisation géométrique

Dans le cas de domaines rectangulaires (ou parallélépipédiques en dimension 3) de côtés parallèles aux axes, on construit une grille de discrétisation en différences finies par quadrillage selon les deux (trois) directions d'axes. On notera Δx le pas de discrétisation selon x et de même Δy et Δz les pas de discrétisation en y et z.

On obtient ainsi aux intersections des lignes du quadrillage les nœuds de coordonnées (x_i , y_j , z_k) du maillage en différences finies.

Cette technique de maillage est généralisable aux assemblages de rectangles (ou de parallélépipèdes) ainsi qu'aux domaines se ramenant par bijection régulière à un rectangle (ou un parallélépipède). Par contre, dans le cas de géométries complexes, les discrétisations par éléments finis (§ 2.4) sont mieux adaptées.

2.3.2 Quelques formules simples d'approximation des dérivées partielles

Notons, en dimension deux, $u_{i,j}$ l'approximation de la valeur exacte $u(x_i , y_j)$ pour le point d'indice i, j de la grille.

■ Pour les **dérivées partielles premières**, on obtient les approximations suivantes :

— **différences divisées progressives** :

$$\left.\begin{aligned} \frac{\partial u}{\partial x}(x_i , y_j) &\approx \frac{u_{i+1,j} - u_{i,j}}{\Delta x} \\ \frac{\partial u}{\partial y}(x_i , y_j) &\approx \frac{u_{i,j+1} - u_{i,j}}{\Delta y} \end{aligned}\right\} \quad \textbf{(58)}$$

— **différences divisées régressives** : on considère alors :

$$\left.\begin{aligned} \frac{\partial u}{\partial x}(x_i , y_j) &\approx \frac{u_{i,j} - u_{i-1,j}}{\Delta x} \\ \frac{\partial u}{\partial y}(x_i , y_j) &\approx \frac{u_{i,j} - u_{i,j-1}}{\Delta y} \end{aligned}\right\} \quad \textbf{(59)}$$

— **différences divisées centrées** : on obtient (comme en dimension un [relation (7)] une approximation du second ordre :

$$\left.\begin{aligned} \frac{\partial u}{\partial x}(x_i , y_j) &\approx \frac{u_{i+1/2,j} - u_{i-1/2,j}}{\Delta x} \\ \frac{\partial u}{\partial y}(x_i , y_j) &\approx \frac{u_{i,j+1/2} - u_{i,j-1/2}}{\Delta y} \end{aligned}\right\} \quad \textbf{(60)}$$

■ On en déduit, par double différentiation, l'**approximation centrée d'ordre deux du laplacien** :

$$-\Delta u\,(x_i , y_j) \approx \frac{-u_{i+1,j} + 2u_{i,j} - u_{i-1,j}}{\Delta x^2} + \frac{-u_{i,j+1} + 2u_{i,j} - u_{i,j-1}}{\Delta y^2} \quad \textbf{(61)}$$

et, plus généralement, l'approximation de l'équation (57) :

$$-\operatorname{div}(\sigma\,\mathbf{grad}\,u)(x_i , y_j)$$

$$\approx -\frac{\sigma(x_{i+1/2} , y_j)(u_{i+1,j} - u_{i,j})}{\Delta x^2} - \frac{\sigma(x_{i-1/2}, y_j)(u_{i,j} - u_{i-1,j})}{\Delta x^2}$$

$$-\frac{\sigma(x_i , y_{j+1/2})(u_{i,j+1} - u_{i,j})}{\Delta y^2} - \frac{\sigma(x_i , y_{j-1/2})(u_{i,j} - u_{i,j-1})}{\Delta y^2} \quad \textbf{(62)}$$

On obtient ainsi le système d'équations linéaires dont la résolution donne les valeurs $u_{i,j}$ de la solution approchée du problème (57) selon :

$$\frac{-\sigma(x_{i+1/2}, y_j)u_{i+1,j} + (\sigma(x_{i+1/2}, y_j) + \sigma(x_{i-1/2}, y_j))u_{i,j} - \sigma(x_{i-1/2}, y_j)u_{i-1,j}}{\Delta x^2}$$
$$+\frac{-\sigma(x_i , y_{j+1/2})u_{i+j+1} + (\sigma(x_i , y_{j+1/2}) + \sigma(x_i , y_{j-1/2}))u_{i,j} - \sigma(x_i , y_{j-1/2})u_{i,j-1}}{\Delta y^2}$$
$$= f_{i,j} \quad \textbf{(63)}$$

On doit y ajouter la prise en compte des **conditions aux limites** :

— pour les conditions de Dirichlet, il suffit de fixer les valeurs de $u_{i,j}$ correspondant aux valeurs données sur la frontière Γ_d ;

— pour les conditions de Neumann, on doit discrétiser :

$$\left.\frac{\partial u}{\partial n}\right|_{\Gamma_n} = g$$

Il y a plusieurs choix possibles pour approcher la dérivée normale en différences finies. Le bon choix, qui conserve la symétrie de la matrice du système linéaire global et qui s'interprète de manière naturelle en éléments finis, consiste à remplacer, selon le côté de frontière concerné, la dérivée normale $\left.\frac{\partial u}{\partial n}\right|_{\Gamma_n}$ par une des quatre expressions

$$\frac{u_{i+1,j} - u_{i,j}}{\Delta x} , \quad \frac{u_{i,j} - u_{i-1,j}}{\Delta x} , \quad \frac{u_{i,j+1} - u_{i,j}}{\Delta y} , \quad \frac{u_{i,j} - u_{i,j-1}}{\Delta y} \quad \textbf{(64)}$$

Il est alors nécessaire de numéroter les $u_{i,j}$ pour qu'elles constituent les composantes U_I d'un vecteur inconnu U. La numérotation influe sur la structure de la matrice. On utilise des algorithmes de numérotation optimale afin de minimiser le stockage (« profil ») de la matrice. Il ne reste alors plus qu'à résoudre le système linéaire obtenu par des méthodes directes de factorisation (méthodes de Gauss LU ou méthode de Choleski LL^T) ou par des méthodes itératives (voir ci-après).

2.4 Approximation par éléments finis

2.4.1 Présentation

On introduit, comme en dimension un (§ 1.4), un produit scalaire de deux fonctions selon :

$$(v, w) = \int\int_{\Omega} v(x, y)\, w(x, y)\, \mathrm{d}x\mathrm{d}y$$

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

et l'espace $L^2[\Omega]$ des fonctions de carré sommable sur Ω, c'est-à-dire telles que l'intégrale suivante existe :

$$\int\int_{\Omega} v^2(x, y)\ dxdy$$

Soit $H^1[\Omega]$ l'espace des fonctions v, de carré sommable et dont les dérivées partielles sont également de carré sommable, et soit $H_0^1[\Omega]$ l'espace des fonctions v de $H^1[\Omega]$ nulles sur la frontière Γ de Ω.

On utilisera la formule de Green (équivalente à l'intégration par parties en dimension un) [relation (55)] :

$$-\int\int_{\Omega} \Delta u\ v dxdy = \int\int_{\Omega} \mathbf{grad}\, u\ \mathbf{grad}\, v dxdy - \int_{\Gamma} \frac{\partial u}{\partial n}\ v d\gamma \qquad \forall v \in H^1(\Omega) \quad \textbf{(65)}$$

Ici encore, nous avons choisi de présenter les méthodes dans le cas de la dimension deux, mais tout ce qui précède s'étend sans difficulté au cas de la dimension trois d'espace.

Soit Ω le domaine borné polygonal de $\mathbb{R}^2$ de frontière Γ. Supposons Γ constituée de deux parties Γ_d et Γ_n :

$$\Gamma = \Gamma_d \cup \Gamma_n .$$

Sur Γ_d sont imposées des conditions de Dirichlet.

Sur Γ_n sont imposées des conditions de Neumann.

Nous rappelons le problème (57)

$$-\text{div}\,(\sigma\ \mathbf{grad}\ u)(x, y) = f(x, y) \qquad \forall x, y \in \Omega$$
$$u|_{\Gamma_d} = u_d$$
$$\left.\frac{\partial u}{\partial n}\right|_{\Gamma_n} = g$$

2.4.2 Formulation variationnelle

La formulation du problème s'écrit dans l'espace V des fonctions de $H^1(\Omega)$ nulles sur la partie Γ_d de la frontière où la solution u est fixée. On peut se ramener aux conditions de Dirichlet homogènes sur Γ_d en introduisant une fonction auxiliaire simple prenant les valeurs imposées sur la frontière Γ_d :

$$u_0|_{\Gamma_d} = u_d$$

et en posant

$$u = \tilde{u} + u_0 .$$

On obtient le problème variationnel :

$$\left.\begin{array}{l} \text{Trouver } \tilde{u} \in V \text{ telle que : } \forall v \in V \\ \displaystyle\int\int_{\Omega} \mathbf{grad}\,\tilde{u}\ \mathbf{grad}\, v\ dxdy \\ \displaystyle= \int\int_{\Omega} fv\ dxdy - \int\int_{\Omega} \mathbf{grad}\, u_0\ \mathbf{grad}\, v\ dxdy + \int_{\Gamma_n} gv d\gamma \end{array}\right\} \quad \textbf{(66)}$$

2.4.3 Maillage

Dans la méthode des éléments finis, la construction du sous-espace V_h nécessite la discrétisation préalable du domaine Ω en éléments géométriques simples.

En dimension un (§ 1.5), la discrétisation préalable du domaine, un intervalle de $\mathbb{R}$, en éléments, ne posait pas de difficultés.

En dimension deux et, plus encore, en dimension trois, la discrétisation du domaine Ω est un problème technique difficile. La qualité du maillage est cruciale pour la qualité de l'approximation. Le problème de la réalisation du maillage se pose à la fois en amont de la résolution numérique, qui s'appuie sur une discrétisation a priori, et en aval dans les techniques de maillages adaptatifs par lesquelles on s'efforce d'améliorer la qualité de la discrétisation en fonction des résultats obtenus.

L'exposé des techniques mises en œuvre pour construire un maillage en éléments finis dépasse le cadre de cet article et nous renvoyons à la littérature (notamment [3] et [4]). Disons simplement que l'on peut distinguer deux **types de maillages** :

— les **maillages structurés** (figure **7**), qui fournissent une discrétisation « régulière » obtenue par transformation d'une grille régulière sur un domaine rectangulaire ;

— les **maillages non structurés** (figure **8**), principalement construits par la méthode de Voronoï, qui peuvent s'appliquer aux géométries les plus générales.

En dimension deux, les éléments sont des triangles ou des quadrangles de côtés droits ou curvilignes. En dimension trois, ce sont des tétraèdres, pentaèdres ou hexaèdres.

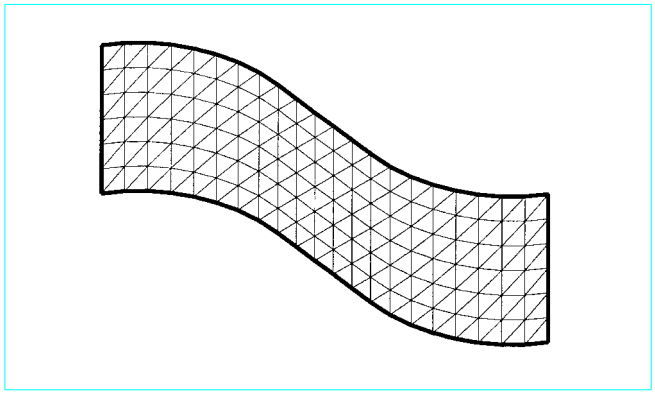


Figure 7 – Maillage structuré

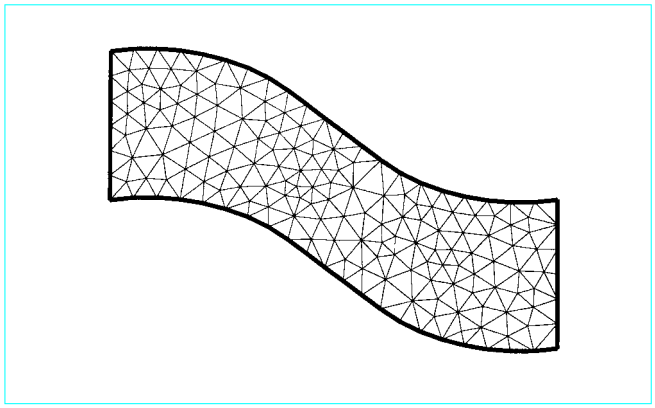


Figure 8 – Maillage non structuré

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

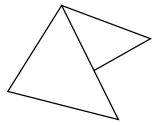


Figure 9 – Configuration interdite

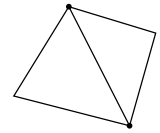


Figure 10 – Configuration conforme d'éléments adjacents

■ Un maillage en éléments finis doit satisfaire les **critères** suivants :

1) Les éléments K_I du maillage doivent recouvrir le domaine Ω :

$$\bigcup_i K_i = \bar{\Omega}$$

Cela implique que ce domaine soit polygonal ou approché par un polygône si l'on utilise des éléments droits.

2) L'intersection de deux éléments distincts ne peut être que :
— l'ensemble vide,
— un sommet,
— un côté,
— une face (en dimension trois).

Cela a pour but d'assurer la continuité des fonctions de V_h. En particulier, une disposition telle que celle présentée figure **9** est interdite.

2.4.4 Éléments finis de Lagrange triangulaires de degré un : les éléments finis P1

Dans ce cas, l'espace d'approximation V_h est un espace de fonctions continues affines par éléments triangulaires.

Dans chaque triangle, la restriction des fonctions de V_h est donc un polynôme de degré un de la forme $a_0 + a_1 x + a_2 y$ qui est donc déterminé de façon unique par ces valeurs en trois points distincts. On choisit les trois sommets du triangle. Cela assure la continuité globale des fonctions de V_h sur le domaine polygonal Ω. En effet, sur chaque arête commune à deux triangles adjacents (figure **10**), les restrictions des fonctions de V_h sont des fonctions affines fixées de manière unique par la donnée de leurs valeurs aux deux sommets sur l'arête.

Globalement, les fonctions de V_h seront uniquement déterminées par leurs valeurs aux sommets ou nœuds de la triangulation. **La dimension totale de V_h est donc égale au nombre de nœuds du maillage**.

Dans le cas de conditions aux limites de Dirichlet sur une partie de la frontière, la dimension de V_h est évidemment réduite au nombre des nœuds associés à une valeur inconnue de la solution. Elle est donc égale au nombre de nœuds total moins le nombre de nœuds où la solution est fixée par une condition de Dirichlet.

■ Les fonctions de base P1

La construction d'une base de V_h se fait selon la technique classique de Lagrange. On prend comme fonctions de base les N fonctions w_I de V_h définies par les N conditions suivantes aux N nœuds (x_I, y_I) du maillage :

$$w_I(x_J, y_J) = \delta_{IJ}$$

On remarquera que ces fonctions ont un support réduit à l'union des triangles dont le point (x_I, y_I) est un sommet (figure **11**).

Dans cette base, une fonction de V_h s'écrit :

$$v_h(x, y) = \sum_I v_I w_I(x, y)$$

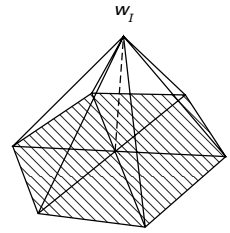


Figure 11 – Graphe d'une fonction w_I

■ Les fonctions de forme P1

On appellera **fonctions de forme** les restrictions des fonctions de base dans un élément.

Dans chaque triangle T de sommets A_1, A_2, A_3, il n'y a que 3 fonctions de base non nulles. Les restrictions de ces fonctions de base notées $w_{I_1}, w_{I_2}, w_{I_3}$ sont les trois fonctions polynomiales de degré un prenant la valeur 1 en un des sommets et nulles aux deux autres sommets. Notons-les respectivement $\lambda_1, \lambda_2, \lambda_3$.

λ_1 est le polynôme de degré un prenant la valeur 1 en A_1 et nul en A_2 et A_3 :

$$\lambda_1(x, y) = a_0 + a_1 x + a_2 y \qquad \textbf{(67)}$$

λ_1 est donc déterminé par le système linéaire suivant :

$$\left.\begin{aligned} \lambda_1(x_1, y_1) &= a_0 + a_1 x_1 + a_2 y_1 = 1 \\ \lambda_1(x_2, y_2) &= a_0 + a_1 x_2 + a_2 y_2 = 0 \\ \lambda_1(x_3, y_3) &= a_0 + a_1 x_3 + a_2 y_3 = 0 \end{aligned}\right\} \qquad \textbf{(68)}$$

Le déterminant de ce système

$$\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} = 2 \text{ Aire } (T) \qquad \textbf{(69)}$$

est différent de zéro si les points A_1, A_2, A_3 ne sont pas alignés.

En résolvant le système 3×3 ci-dessus et les systèmes analogues pour λ_2 et λ_3, on obtient les formules suivantes :

$$\lambda_1(x, y) = \frac{x_2 y_3 - x_3 y_2 + x(y_2 - y_3) + y(x_3 - x_2)}{2 \text{ Aire } (T)} \qquad \textbf{(70)}$$

$$\lambda_2(x, y) = \frac{x_3 y_1 - x_1 y_3 + x(y_3 - y_1) + y(x_1 - x_3)}{2 \text{ Aire } (T)} \qquad \textbf{(71)}$$

$$\lambda_3(x, y) = \frac{x_1 y_2 - x_2 y_1 + x(y_1 - y_2) + y(x_2 - x_1)}{2 \text{ Aire } (T)} \qquad \textbf{(72)}$$

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

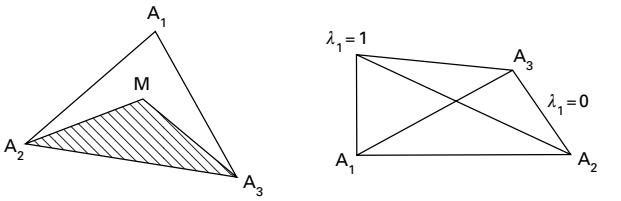


Figure 12 – Deux représentations de la fonction λ_1

■ On en déduit les expressions des **gradients** :

$$\left.\begin{aligned}\frac{\partial \lambda_1}{\partial x} &= \frac{y_2 - y_3}{2\ \text{Aire}\ (T)}\\ \frac{\partial \lambda_1}{\partial y} &= \frac{x_3 - x_2}{2\ \text{Aire}\ (T)}\end{aligned}\right\} \qquad (73)$$

$$\left.\begin{aligned}\frac{\partial \lambda_2}{\partial x} &= \frac{y_3 - y_1}{2\ \text{Aire}\ (T)}\\ \frac{\partial \lambda_2}{\partial y} &= \frac{x_1 - x_3}{2\ \text{Aire}\ (T)}\end{aligned}\right\} \qquad (74)$$

$$\left.\begin{aligned}\frac{\partial \lambda_3}{\partial x} &= \frac{y_1 - y_2}{2\ \text{Aire}\ (T)}\\ \frac{\partial \lambda_3}{\partial y} &= \frac{x_2 - x_1}{2\ \text{Aire}\ (T)}\end{aligned}\right\} \qquad (75)$$

■ Les trois fonctions λ_1, λ_2, λ_3 s'appellent les **coordonnées barycentriques** du triangle T. On les désigne sous le nom d'« ***area coordinates*** » en anglais car elles représentent en chaque point M de coordonnées x, y le rapport des aires algébriques des triangles $M A_i A_j$ et T. Par exemple, on a (figure **12**) :

$$\lambda_1(M) = \frac{\text{Aire}\ (A_2 A_3 M)}{\text{Aire}\ (T)}$$

Une fonction quelconque de V_h prenant les valeurs v_1, v_2, v_3 aux sommets A_1, A_2, A_3 du triangle T s'écrit dans T sous la forme :

$$v_h(x, y) = v_1\lambda_1(x, y) + v_2\lambda_2(x, y) + v_3\lambda_3(x, y) \qquad (76)$$

2.4.5 Écriture du problème approché en éléments finis P1

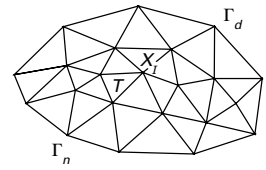
Le problème approché s'écrit dans l'espace V_h des fonctions continues affines par triangles et nulles sur la partie Γ_d de la frontière (figure **13**).

Notons I l'ensemble des indices des nœuds du maillage correspondant à une valeur inconnue de la solution u, c'est-à-dire ici l'ensemble des nœuds n'appartenant pas à Γ_d.

Notons J l'ensemble des indices des sommets du maillage appartenant à Γ_d.

■ La solution $\tilde{u}_h$ s'écrira dans la base des w_J pour $J \in I$ selon :

$$\tilde{u}_h(x, y) = \sum_{J \in I} u_J\, w_J(x, y) \qquad (77)$$



X_I nœud du maillage (donc sommet d'un triangle) de coordonnées (x_I, y_I)

Figure 13 – Maillage en triangles du domaine de calcul

■ La fonction auxiliaire u_0 sera approchée par une fonction $u_{0,h}$ continue et affine par morceaux prenant les valeurs imposées sur Γ_d et nulle sur tous les nœuds d'indices $J \in I$:

$$u_{0,h}(x, y) = \sum_{J \in I} u_d(x_J,\ y_J)\ w_J(x, y) \qquad (78)$$

Ce choix de $u_{0,h}$ présente deux **avantages**.

1) La solution cherchée u_h et la solution calculée $\tilde{u}_h$ prennent les mêmes valeurs aux points où la solution u est inconnue. Il n'y a donc pas de transformation a posteriori à effectuer sur les résultats.

2) Les conditions aux limites ne produisent qu'une modification limitée du système linéaire qui n'intervient que sur quelques composantes du second membre.

■ En intégrant, dans la formulation variationnelle, tous ces éléments, on obtient en définitive le problème approché dans V_h :

$$\left.\begin{aligned}&\text{Trouver les valeurs } u_J \text{ pour } J \in I \text{ telles que :}\\ &\sum_{J \in I}\left(\int\int_\Omega \mathbf{grad}\, w_J\ \mathbf{grad}\, w_I\ \mathrm{d}x\mathrm{d}y\right)u_J\\ &= \int\int_\Omega f\ w_I\ \mathrm{d}x\mathrm{d}y + \int_{\Gamma_n} g\ w_I\ \mathrm{d}\gamma\\ &\quad - \sum_{J \in J}\left(\int\int_\Omega \mathbf{grad}\, w_J\ \mathbf{grad}\, w_I\ \mathrm{d}x\mathrm{d}y\right)u_d(x_J, y_J) \qquad \forall I \in I\end{aligned}\right\} \qquad (79)$$

On obtient un système de N_I équations à N_I inconnues où N_I désigne le nombre de points de maillage d'indices $I \in I$, donc le nombre de nœuds correspondant à des valeurs inconnues de la solution.

Ce système s'écrit sous la forme matricielle :

$$K\,U = F \qquad (80)$$

où K est la matrice de raideur de coefficients :

$$K_{I,J} = \int\int_\Omega \mathbf{grad}\, w_J\ \mathbf{grad}\, w_I\ \mathrm{d}x\mathrm{d}y \qquad (81)$$

et F le vecteur second membre de composantes :

$$F_I = \int\int_\Omega f\ w_I\ \mathrm{d}x\mathrm{d}y + \int_{\Gamma_n} g\ w_I\ \mathrm{d}\gamma - \sum_{J \in J}\left(\int\int_\Omega \mathbf{grad}\, w_J\ \mathbf{grad}\, w_I\ \mathrm{d}x\mathrm{d}y\right)u_d(x_J, y_J) \qquad (82)$$

où l'on reconnaît :

— un premier terme représentant les efforts surfaciques (correspondant au second membre du problème différentiel) ;

— un deuxième terme, intégrale curviligne, provenant des conditions aux limites de Neumann ;

— un dernier terme, expression des conditions de Dirichlet non homogènes.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

$$\begin{pmatrix} \dfrac{(y_2-y_3)^2+(x_2-x_3)^2}{4\ \text{Aire}\ (T)} & \dfrac{(y_2-y_3)(y_3-y_1)+(x_2-x_3)(x_3-x_1)}{4\ \text{Aire}\ (T)} & \dfrac{(y_2-y_3)(y_1-y_2)+(x_2-x_3)(x_1-x_2)}{4\ \text{Aire}\ (T)} \\ \dfrac{(y_2-y_3)(y_3-y_1)(x_2-x_3)(x_3-x_1)}{4\ \text{Aire}\ (T)} & \dfrac{(y_3-y_1)^2+(x_3-x_1)^2}{4\ \text{Aire}\ (T)} & \dfrac{(y_3-y_1)(y_1-y_2)+(x_3-x_1)(x_1-x_2)}{4\ \text{Aire}\ (T)} \\ \dfrac{(y_2-y_3)(y_1-y_2)+(x_2-x_3)(x_1-x_2)}{4\ \text{Aire}\ (T)} & \dfrac{(y_3-y_1)(y_1-y_2)+(x_3-x_1)(x_1-x_2)}{4\ \text{Aire}\ (T)} & \dfrac{(y_1-y_2)^2+(x_1-x_2)^2}{4\ \text{Aire}\ (T)} \end{pmatrix}$$

Le calcul des coefficients $K_{I,J}$ de la matrice de raideur et des composantes F_I du second membre se fait par une procédure d'assemblage des contributions apportées par chacun des éléments T_k de la triangulation.

Par exemple pour la matrice de raideur K, on a :

$$K_{I,J} = \int\int_{\Omega} \mathbf{grad}\, w_J\ \mathbf{grad}\, w_I\ \mathrm{d}x\mathrm{d}y = \sum_k \int\int_{T_k} \mathbf{grad}\, w_J\ \mathbf{grad}\, w_I\ \mathrm{d}x\mathrm{d}y \qquad \textbf{(83)}$$

On observe que la matrice K est très « creuse » ; un grand nombre de ses coefficients sont nuls, en raison du choix de fonctions w_I de support limité.

2.4.6 Calcul de la matrice de raideur élémentaire P1

Un des outils de base essentiels à la programmation de la méthode des éléments finis est un tableau de correspondances entre les nœuds X_I du maillage global et les points d'un élément particulier : ici, les sommets A_1, A_2, A_3 des éléments triangulaires.

Dans chaque élément triangulaire T_k de sommets A_1, A_2, A_3 correspondant aux nœuds X_I, X_J, X_K, les seules fonctions de base non nulles sont les fonctions w_I, w_J, w_K. Leurs restrictions dans le triangle sont respectivement les trois coordonnées barycentriques λ_1, λ_2, λ_3 calculées précédemment (§ 2.4.4).

La matrice élémentaire relative au triangle T_k est donc une matrice 3×3 de coefficients :

$$\text{Elem}\ K_{i,j} = \int\int_{T_k} \mathbf{grad}\,\lambda_j\ \mathbf{grad}\,\lambda_i\ \mathrm{d}x\mathrm{d}y \qquad \forall i, j = 1, 2, 3 \qquad \textbf{(84)}$$

Comme, dans ce cas particulier simple d'éléments P1, les gradients sont constants par triangles, les intégrales à calculer sont des intégrales de fonctions constantes. Il suffit d'en multiplier la valeur par l'aire de l'élément.

On obtient ainsi la matrice élémentaire P1 de coefficients (voir encadré en haut de page) :

En particulier, dans le cas du triangle rectangle isocèle de sommets :

$$\left.\begin{aligned} A_1 &= (0, 0) \\ A_2 &= (1, 0) \\ A_3 &= (0, 1) \end{aligned}\right\} \qquad \textbf{(85)}$$

souvent utilisé comme triangle de référence, on a la matrice élémentaire de raideur suivante :

$$\text{Elem}\,K = \begin{pmatrix} 1 & -\dfrac{1}{2} & -\dfrac{1}{2} \\ -\dfrac{1}{2} & \dfrac{1}{2} & 0 \\ -\dfrac{1}{2} & 0 & \dfrac{1}{2} \end{pmatrix} \qquad \textbf{(86)}$$

2.4.7 Calcul des seconds membres élémentaires

Le second membre se compose de 3 termes.

2.4.7.1 Un premier terme surfacique

Il s'écrit :

$$F_{S_I} = \int\int_{\Omega} f\ w_I\ \mathrm{d}x\mathrm{d}y \qquad \forall I \qquad \textbf{(87)}$$

Son calcul s'effectue en sommant les contributions de chaque élément triangulaire :

$$F_{S_I} = \sum_k F_{S_I,\,T_k} = \sum_k \int\int_{T_k} f\ w_I\ \mathrm{d}x\mathrm{d}y \qquad \textbf{(88)}$$

Si f est connue analytiquement et suffisamment simple, on peut calculer exactement, à la main, les intégrales. Mais, en général, le second membre f est connu par ses valeurs aux nœuds. On le représente dans la base des w_I et on est ramené aux calculs suivants :

$$F_{S_I} = \int\int_{\Omega} \sum_J f_J\ w_J\ w_I\ \mathrm{d}x\mathrm{d}y \qquad \forall I, \qquad \textbf{(89)}$$

c'est-à-dire aux calculs de :

$$\int\int_{\Omega} w_J\ w_I\ \mathrm{d}x\mathrm{d}y \qquad \forall I, J \qquad \textbf{(90)}$$

Dans chaque élément, on doit donc calculer la matrice de masse élémentaire.

Matrices élémentaires de masse

On obtient la matrice élémentaire de masse, en utilisant les formules d'intégration exactes suivantes :

$$\int\int_T \lambda_i\ \mathrm{d}x\mathrm{d}y = \frac{\text{Aire}\ (T)}{3} \qquad \textbf{(91)}$$

$$\int\int_T \lambda_i^2\ \mathrm{d}x\mathrm{d}y = \frac{\text{Aire}\ (T)}{6} \qquad \textbf{(92)}$$

$$\int\int_T \lambda_j\ \lambda_i\ \mathrm{d}x\mathrm{d}y = \frac{\text{Aire}\ (T)}{12} \qquad \text{si } i \text{ différent de } j, \qquad \textbf{(93)}$$

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

ce qui donne, pour la matrice élémentaire de masse, le résultat :

$$\text{Elem}\, M_{T_k} = \text{Aire}(T_k) \begin{pmatrix} \frac{1}{6} & \frac{1}{12} & \frac{1}{12} \\ \cdots & \frac{1}{6} & \frac{1}{12} \\ \cdots & \cdots & \frac{1}{6} \end{pmatrix} \quad (94)$$

Matrice de masse condensée (*lumping*)

Il peut être avantageux de calculer la matrice de masse élémentaire de manière approchée en utilisant la formule suivante (exacte sur P1) :

$$\iint_T \phi \, \mathrm{d}x\mathrm{d}y = \frac{\text{Aire}\,(T)}{3} \sum_{i=1}^{3} \phi(A_i) \quad (95)$$

On obtient, dans ce cas, une matrice de masse diagonale égale à :

$$\text{Elem}\, M_{T_k} = \frac{\text{Aire}(T_k)}{3} I \quad (96)$$

On en déduit le second membre élémentaire surfacique :

$$\begin{pmatrix} F_{S_{I,T_k}} \\ F_{S_{J,T_k}} \\ F_{S_{K,T_k}} \end{pmatrix} = \text{Elem}\, M_{T_k} \begin{pmatrix} f_I \\ f_J \\ f_K \end{pmatrix} \quad (97)$$

2.4.7.2 Un terme de bord provenant des conditions de Neumann

Il s'écrit :

$$F_{N_I} = \int_{\Gamma_n} g \, w_I \, \mathrm{d}\gamma \quad (98)$$

Son calcul s'effectue sur les éléments ayant un côté sur Γ_d. Il se ramène à une intégrale simple sur un côté A d'un triangle. Si la fonction g est donnée par ses valeurs aux nœuds du maillage et si A a pour extrémités X_I et X_J, on a sur A :

$$g = g_I \, w_I + g_J \, w_J$$

et on doit donc calculer :

$$g_J \int_A w_J \, w_I \, \mathrm{d}\gamma \quad (99)$$

$$g_J \int_A w_J^2 \, \mathrm{d}\gamma \quad (100)$$

et

$$g_I \int_A w_I^2 \, \mathrm{d}\gamma \quad (101)$$

Ces calculs s'effectuent exactement par la formule de Simpson :

$$g_J \int_A w_J \, w_I \, \mathrm{d}\gamma = \frac{\text{Longueur (A)}}{6} g_J \quad (102)$$

$$g_J \int_A w_J^2 \, \mathrm{d}\gamma = \frac{\text{Longueur (A)}}{3} g_J \quad (103)$$

etc.

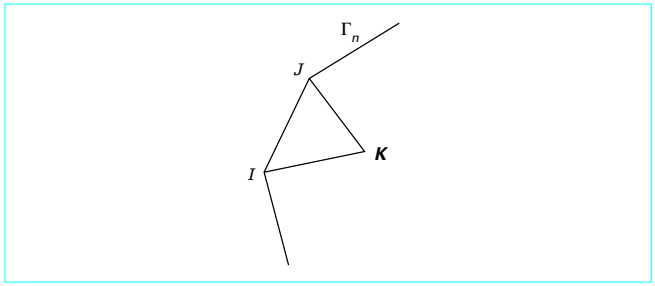


Figure 14 – Un élément ayant une arête sur la partie de frontière Neumann

Dans le cas de la figure **14**, on obtient :

$$\begin{pmatrix} F_{N_{T,I}} \\ F_{N_{T,J}} \\ F_{N_{T,K}} \end{pmatrix} = \frac{\text{Longueur (A)}}{6} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} g_I \\ g_J \\ g_K \end{pmatrix} \quad (104)$$

On peut également calculer ce terme de façon approchée par la formule des trapèzes, ce qui donne :

$$\begin{pmatrix} F_{N_{T,I}} \\ F_{N_{T,J}} \\ F_{N_{T,K}} \end{pmatrix} = \frac{\text{Longueur (A)}}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} g_I \\ g_J \\ g_K \end{pmatrix} \quad (105)$$

Dans tous les cas, le terme provenant des conditions de Neumann n'induit de contributions non nulles que pour les composantes du second membre relatives à des nœuds du maillage situés sur Γ_n.

2.4.7.3 Un troisième terme provenant des conditions de Dirichlet non homogènes

Il s'écrit :

$$F_{D_I} = -\sum_{J \in \mathcal{J}} \left(\iint_\Omega \mathbf{grad}\, w_J \, \mathbf{grad}\, w_I \, \mathrm{d}x\mathrm{d}y \right) u_d(x_J, y_J) \quad (106)$$

Son calcul se ramène encore à une somme de contributions des triangles T.

Pour chaque triangle dont les points d'indice $J \in \mathcal{J}$ et le point I sont sommets on obtient une contribution à la I^{e} composante du second membre égale à :

$$-\sum_{J \in \mathcal{J}} \left(\iint_T \mathbf{grad}\, \lambda_j \, \mathbf{grad}\, \lambda_i \, \mathrm{d}x\mathrm{d}y \right) u_d(x_J, y_J) \quad (107)$$

On retrouve des coefficients déjà calculés pour la matrice de raideur. Pour un triangle I, J, K comme celui de la figure **15** on obtient une contribution unique à la composante K :

$$F_{D_{T,K}} = -(a_{i,k}, a_{j,k}, a_{k,k}) \begin{pmatrix} u_{d_i} \\ u_{d_j} \\ 0 \end{pmatrix} \quad (108)$$

où les i, j, k sont les numéros internes au triangle T correspondant aux indices globaux I, J, K.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

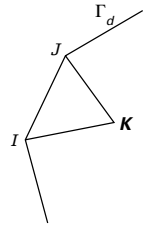


Figure 15 – Un élément ayant une arête sur la partie de frontière Dirichlet

2.4.8 Algorithme d'assemblage

Supposons un maillage en N éléments T_k pour $k = 1, 2 \ldots N$. Notons A la matrice globale à assembler (matrice de raideur ou de masse globale, second membre) et a_k les matrices élémentaires correspondantes relatives à chaque élément T_k. L'algorithme d'assemblage est très simple, dès lors que l'on dispose d'un tableau *me* associant les sommets d'un élément T_k et les nœuds du maillage global. Dans ce cas simple d'éléments triangulaires P1, chaque élément T_k comprend trois nœuds X_I, X_J, X_K correspondant aux sommets A_1, A_2, A_3 du triangle T_k. D'où l'algorithme :

POUR $k = 1$, N FAIRE
  POUR $i = 1, 3$ FAIRE
    $I = me\,(k,i)$
  POUR $j = 1, 3$ FAIRE
    $J = me\,(k, j)$
      $A\,(I, J) = A\,(I, J) + a_k\,(i, j)$
FIN DES 3 BOUCLES

2.5 Généralisation

Tout ce qui précède se généralise sans difficultés, sinon techniques, à des approximations par éléments finis de degré plus élevé et de formes triangulaires, ou quadrangulaires, tétraédriques ou prismatiques à côtés (ou faces) droites (planes) ou courbes (gauches).

3. Méthodes de résolution des systèmes linéaires

3.1 Méthodes directes

Les méthodes directes de résolution des systèmes linéaires sont des méthodes dans lesquelles la solution est obtenue de façon exacte en un nombre fini d'opérations. De façon exacte s'entend, sur un ordinateur, aux erreurs « d'arrondis machine » près.

■ Le prototype de méthode directe est la méthode du **pivot de Gauss**. Cette méthode permet de ramener la résolution d'un système général à la résolution d'un système triangulaire supérieur, résolution qui se fait explicitement par un processus de remontée. On commence par calculer la dernière composante du vecteur inconnu en utilisant la dernière équation et on remonte équation par équation en déterminant les composantes correspondantes. La rencontre de pivot nul peut nécessiter la permutation de lignes du système. Cependant, pour certaines classes de matrices, en particulier les matrices symétriques définies positives, on est assuré de pouvoir trianguliariser le système par Gauss sans permutation. La méthode du pivot équivaut alors à une factorisation de type

$$A = LU \qquad (109)$$

de la matrice A. L est une matrice triangulaire inférieure à diagonale unité et U une matrice triangulaire supérieure.

On peut utiliser la symétrie de A pour obtenir une factorisation de type

$$A = LDL^T \qquad (110)$$

avec D diagonale.

■ Dans le cas d'une matrice A symétrique définie positive, la **méthode de Choleski** conduit à une factorisation :

$$A = LL^T \qquad (111)$$

On trouve la matrice L, qui cette fois n'est plus à diagonale unité, par un algorithme d'identification de coefficients.

De

$$A_{ij} = \sum_{k=1,\,i} L_{ik}\, L_{jk} \quad \text{pour} \quad j \leqslant i$$

on déduit pour tout i :

$$L_{ii} = \sqrt{A_{ii} - \sum_{k=1,\,i-1} L_{ik}^2}$$

et pour tout $j < i$

$$L_{ij} = \frac{A_{ij} - \sum_{k=1,\,j-1} L_{ik}\, L_{jk}}{L_{jj}}$$

Les méthodes directes présentent l'**avantage** de fournir la solution exacte (aux erreurs d'arrondis machine près) en un nombre fini d'opérations (de l'ordre de $n^3/3$ pour Gauss). La méthode du pivot de Gauss s'applique à tout système inversible.

Par contre, les méthodes directes ont un coût important en stockage mémoire (bien que des techniques de stockage minimal associées à des algorithmes de numérotation optimale des inconnues permettent de le réduire sensiblement). Cela rend leur application pratiquement impossible, en l'état actuel de la technologie, pour de gros systèmes à plus de 10^5 inconnues, et donc, en particulier, pour la résolution de problèmes industriels en dimension 3 d'espace.

Nous renvoyons au livre de P. Lascaux et R. Théodor (tome 1) [6] pour plus de précisions sur les méthodes directes.

3.2 Méthodes itératives

Le principe général des méthodes itératives est le suivant. Le vecteur solution du système est obtenu comme limite (quand elle existe) d'une suite itérative de vecteurs définie par une récurrence linéaire de la forme :

$$\left.\begin{aligned} X^{(0)} &\quad \text{donné} \\ X^{(k+1)} &= M\,X^{(k)} + C \end{aligned}\right\} \qquad (112)$$

où M est une matrice $n \times n$ dite matrice d'itération et C un vecteur de $\mathbb{R}^n$.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

À la convergence, on a

$$X = MX + C \tag{113}$$

donc cette équation de type *point fixe* doit, évidemment, être équivalente à l'équation initiale :

$$AX = B \tag{114}$$

3.2.1 Conditions de convergence

On a convergence de la suite vectorielle $\{X^{(k)}\}$ vers la solution X à condition que, pour une norme vectorielle, on ait :

$$\lim_{k \to \infty} \| X^{(k)} - X \| = 0 \tag{115}$$

On peut également introduire le vecteur **résidu** :

$$R^{(k)} = B - AX^{(k)} = A\,(X - X^{(k)}) \tag{116}$$

On a de manière équivalente (A est inversible) convergence des itérations vers la solution si :

$$\lim_{k \to \infty} \| R^{(k)} \| = 0 \tag{117}$$

Définition 1 (Norme matricielle). On appelle norme matricielle induite par une norme vectorielle $\| \cdot \|$ l'application de l'espace des matrices dans $\mathbb{R}^+$ définie pour toute matrice A par :

$$\| A \| = \max_{X \neq 0} \frac{\| AX \|}{\| X \|} \tag{118}$$

Définition 2 (Rayon spectral). On appelle rayon spectral d'une matrice A le nombre positif

$$\rho (A) = \max_i |\lambda_i| \tag{119}$$

où les λ_i sont les valeurs propres de la matrice A.

Si A est symétrique, son rayon spectral $\rho (A)$ est égal à sa norme induite par la norme vectorielle euclidienne.

Théorème 1. L'itération (112) :

$$\left.\begin{array}{l} X^{(0)} \quad \text{donné} \\ X^{(k+1)} = M\,X^{(k)} + C \end{array}\right\}$$

converge vers la solution X si pour une norme matricielle donnée :

$$\| M \| < 1 \tag{120}$$

Démonstration

De

$$X^{(k+1)} = M\,X^{(k)} + C$$

et

$$X = M\,X + C$$

on déduit :

$$X - X^{(k+1)} = M\,(X - X^{(k)})$$

soit

$$X - X^{(k)} = M^k\,(X - X^{(0)})$$

donc

$$\| X - X^{(k)} \| = \| M^k\,(X - X^{(0)}) \| \leqslant \| M \|^k\, \| X - X^{(0)} \|$$

D'où la convergence de la suite $\{X^k\}$ vers la solution X avec $\| M \| < 1$.

Dans le cas d'une matrice M symétrique, on aura donc convergence si

$$\rho (M) < 1 \tag{121}$$

On admettra que ce résultat est vrai quelle que soit la matrice d'itération M. On voit que le problème de l'estimation des valeurs propres de la matrice d'itération est crucial pour l'étude de la convergence des méthodes.

3.2.2 Méthode de Jacobi

Soit

$$AX = B$$

le système linéaire à résoudre.

La méthode de Jacobi correspond à la décomposition de la matrice A sous la forme

$$A = D - E - F \tag{122}$$

avec D matrice diagonale constituée des éléments diagonaux a_{ii} de A ;

$-E$ matrice triangulaire inférieure stricte, constituée des éléments strictement sous-diagonaux de A : a_{ij} pour $i > j$;

$-F$ matrice triangulaire supérieure stricte, constituée des éléments strictement surdiagonaux de A : a_{ij} pour $i < j$.

On définit alors la méthode de Jacobi comme la méthode itérative :

$$\left.\begin{array}{l} X^{(0)} \quad \text{donné} \\ X^{(k+1)} = D^{-1}\,(E + F)\,X^{(k)} + D^{-1}\,B \end{array}\right\} \tag{123}$$

ou ce qui revient au même :

$$\left.\begin{array}{l} X^{(0)} \quad \text{donné} \\ X^{(k+1)} = [I - D^{-1}\,A]\,X^{(k)} + D^{-1}\,B \end{array}\right\} \tag{124}$$

La matrice d'itération de Jacobi est donc :

$$J = D^{-1}\,(E + F) = I - D^{-1}\,A \tag{125}$$

et la condition de convergence s'exprime par

$$\rho (J) < 1 \tag{126}$$

On observe que la méthode de Jacobi correspond à l'écriture ligne par ligne suivante :

$$x_i^{k+1} = \frac{b_i - \sum_{j \neq i} a_{ij}\, x_j^k}{a_{ii}} \tag{127}$$

3.2.3 Méthode de Gauss-Seidel ou de relaxation

Soit

$$AX = B$$

le système linéaire à résoudre.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique

La méthode de Gauss-Seidel correspond aussi à la décomposition de la matrice A sous la forme de la relation (122) :

$$A = D - E - F$$

Mais on définit cette fois la méthode de Gauss-Seidel comme la méthode itérative :

$$\left.\begin{array}{l} X^{(0)} \quad \text{donné} \\ X^{(k+1)} = (D-E)^{-1} F\, X^{(k)} + (D-E)^{-1} B \end{array}\right\} \tag{128}$$

La matrice d'itération de Gauss-Siedel est donc

$$\mathcal{L} = (D-E)^{-1} F \tag{129}$$

et la condition de convergence s'exprime par

$$\rho(\mathcal{L}) < 1 \tag{130}$$

On observe que la méthode de Gauss-Seidel correspond à l'écriture ligne par ligne suivante :

$$x_i^{k+1} = \frac{b_i - \sum_{j<i} a_{ij}\, x_j^{k+1} - \sum_{j>i} a_{ij}\, x_j^k}{a_{ii}} \tag{131}$$

3.2.4 Méthodes de descente. Méthode du gradient

■ Les méthodes de descente sont des méthodes itératives qui utilisent l'équivalence entre les problèmes suivants :

$$\left.\begin{array}{l} \text{Trouver } X \in \mathbb{R}^N \quad \text{tel que :} \\ AX = B \end{array}\right\} \tag{132}$$

et

$$\left.\begin{array}{l} \text{Trouver } X \in \mathbb{R}^N \quad \text{tel que :} \\ J(X) = \dfrac{1}{2}(AX, X) - (B, X) \text{ soit minimal} \end{array}\right\} \tag{133}$$

et qui sont donc limitées au cas des systèmes dont la matrice A est symétrique définie positive.

Les méthodes de descente sont basées sur le calcul de la solution comme limite d'une suite minimisante de la forme quadratique J. Cette suite est construite comme une suite récurrente :

$$\left.\begin{array}{l} X^{(0)} \quad \text{donné} \\ X^{(k+1)} = X^{(k)} + \alpha_k\, d^{(k)} \end{array}\right\} \tag{134}$$

avec $d^{(k)} \in \mathbb{R}^N$ vecteur donnant la direction de descente à l'étape k et $\alpha_k \in \mathbb{R}$ coefficient déterminé de manière à minimiser J dans la direction $d^{(k)}$:

$$J(X^{(k)} + \alpha_k\, d^{(k)}) \leqslant J(X^{(k)} + \alpha\, d^{(k)}) \qquad \forall \alpha \in \mathbb{R} \tag{135}$$

Après développement, on obtient α_k comme valeur annulant la dérivée de J par rapport à α, soit :

$$\alpha_k = -\frac{(G^{(k)}, d^{(k)})}{(Ad^{(k)}, d^{(k)})} \tag{136}$$

avec

$$G^{(k)} = \mathbf{grad}(J(X^{(k)})) = AX^{(k)} - B \tag{137}$$

■ Dans la **méthode du gradient**, on choisit comme direction de descente la direction du vecteur gradient de J au point $X^{(k)}$. Cette direction est la direction de variation maximale de J. On dit encore direction de plus profonde descente, d'où le nom « *steepest descent method* », employé dans certains ouvrages en anglais.

L'itération de la méthode du gradient s'écrit :

$$\left.\begin{array}{l} X^{(0)} \quad \text{donné} \\ X^{(k+1)} = X^{(k)} + \alpha_k\, G^{(k)} \end{array}\right\} \tag{138}$$

avec

$$\alpha_k = -\frac{\|G^{(k)}\|^2}{(AG^{(k)}, G^{(k)})} \tag{139}$$

3.2.5 Vitesse de convergence de la méthode du gradient. Conditionnement

À partir de

$$G^{(k+1)} = AX^{(k+1)} - B$$

et de

$$X^{(k+1)} = X^{(k)} + \alpha_k\, G^{(k)},$$

on obtient la récurrence suivante sur les gradients :

$$G^{(k+1)} = G^{(k)} + \alpha_k\, AG^{(k)}$$

On en déduit

$$\|G^{(k+1)}\|^2 = \|G^{(k)}\|^2 + 2\alpha_k (AG^{(k)}, G^{(k)}) + \alpha_k^2 \|AG^{(k)}\|^2$$

Avec la relation (139) :

$$\alpha_k = -\frac{\|G^{(k)}\|^2}{(AG^{(k)}, G^{(k)})}$$

on obtient :

$$\|G^{(k+1)}\|^2 = \|G^{(k)}\|^2 \left[\frac{\|G^{(k)}\|^2 \|AG^{(k)}\|^2}{(AG^{(k)}, G^{(k)})^2} - 1\right]$$

En utilisant les valeurs propres et les vecteurs propres de la matrice A supposée définie positive, on déduit :

$$\|G^{(k+1)}\|^2 \leq \|G^{(k)}\|^2 \left[\frac{\lambda_{\max}}{\lambda_{\min}} - 1\right] \tag{140}$$

Le nombre

$$K(A) = \frac{\lambda_{\max}}{\lambda_{\min}} \tag{141}$$

rapport des plus grandes et plus petites valeurs propres de A est appelé **nombre de conditionnement de la matrice *A***. Plus il est proche de 1, plus vite la méthode du gradient converge.

Les techniques de préconditionnement, qui conduisent à l'algorithme du gradient préconditionné ont, entre autres, pour but de rapprocher les valeurs propres extrêmes afin d'accélérer la convergence des méthodes itératives.

Il existe d'autres méthodes itératives plus performantes : méthode du gradient conjugué préconditionné, méthode du double gradient (pour résoudre les systèmes non symétriques), méthodes de résidus minimaux et de quasi Newton avec en particulier l'algorithme GMRES. Nous renvoyons le lecteur à la bibliographie [4], [5] pour des développements sur ces méthodes.

Toute reproduction sans autorisation du Centre français d'exploitation du droit de copie est strictement interdite.
© Techniques de l'Ingénieur, traité Génie électrique