

LA MÉTHODE DE GALERKIN GÉNÉRALISÉE UNIDIMENSIONNELLE
AVEC DES ÉLÉMENTS FINIS ET DES FONCTIONS COMBINÉES
D'INTERPOLATION

par

Doina Brădeanu

1. La méthode de Galerkin généralisée (méthode des résidus pondérés). On considère l'application de la méthode de Galerkin généralisée, discrétisée sur des éléments finis, à propos d'un problème aux limites relatif à l'équation opératorielle

$$(1) \quad Au + Bu = f(x) \quad , \quad x \in \Omega \quad , \quad f \in L_2(\Omega)$$

où A et B sont des opérateurs linéaires ou nonlinéaires dans l'espace de Hilbert H, aux domaines de définition D(A) et D(B). On supposera que $D(A) \subseteq D(B)$ et que D(A) est dense en H.

On cherche, pour la solution exacte u, une solution approchée par interpolation, de la forme

$$U(x) = u_N(x) = \sum_{j=0}^N u_j \varphi_j(x) \quad , \quad x \in \Omega = (a, b) \subset \mathbb{R} \quad ,$$

dans laquelle $u_j = u(x_j)$, où $x_0, x_1, \dots, x_N$ sont des points par lesquels on discrétise l'intervalle $[a, b]$ en N sous-intervalles (éléments finis) et φ_j sont des fonctions d'approximation (d'interpolation, fonctions de base), définies par morceaux, par

$$(2) \quad \varphi_j(x) = \begin{cases} 1 - \frac{x_j - x}{h} \exp(r(1 - \frac{x_j - x}{h})) & , x \in [x_{j-1}, x_j] \\ (1 - \frac{x - x_j}{h}) \exp(r \frac{x - x_j}{h}) & , x \in [x_j, x_{j+1}] \\ 0 & , x \notin (x_{j-1}, x_{j+1}) \end{cases}$$

$$(r \in \mathbb{R} , j = 1, 2, \dots, N-1)$$

Le système de fonctions $\{\varphi_j\}$ introduites en [3] remplit toutes les conditions requises par l'application de la méthode de Galerkin discrétisée sur des éléments finis unidimensionnels.

La méthode des résidus pondérés introduit également les fonctions de poids (fonctions test) $\psi_j(x)$, $x \in \Omega$, associées aux fonctions de base φ_j , avec $\psi_j(x) = \psi(\frac{x-x_j}{h})$, $\psi_j(x_j) = 1$, définies par morceaux, fonctions qui s'annulent à l'extérieur de l'intervalle (x_{j-1}, x_{j+1}) , c'est-à-dire telles que $\psi(s) = 0$ pour $|s| > 1$, $s = (x - x_j) / h$.

Les fonctions ψ_j , censées au début arbitraires, se choisissent conformément à la nature du problème différentiel (1); par exemple, dans la méthode de Galerkin on choisit $\psi_j = \varphi_j$.

Pour résoudre le problème on introduit un nouvel opérateur $\tilde{K}$ au domaine de définition $D(\tilde{K}) \supseteq D(A)$ et on supposera que toutes les fonctions test ψ_j appartiennent à l'ensemble $D(\tilde{K})$.

Dans la méthode des résidus pondérés, les inconnues u_j sont déterminées par le système algébrique

$$(3) \quad (A u_N + B u_N - f, \tilde{K} \psi_j) = 0, j = 0, 1, \dots, N,$$

où $(\dots)$ est le produit scalaire en H muni de la norme $\|u\| = \sqrt{(u, u)}$.

La variété des méthodes résiduelles et la résolution ultérieure du système (3), ainsi que la convergence de u_N vers u lorsque

$N \rightarrow \infty$ (dans un sens à préciser), dépendent du type des opérateurs donnés A et B et du choix des fonctions $\{\varphi_j\}$, $\{\psi_j\}$ et de l'opérateur $\tilde{K}$.

2. Produits scalaires en $L_2(\Omega)$ avec des fonctions de base φ_j combinées. Nous choisirons à titre de fonctions de base dans le problème résiduel de l'élément fini unidimensionnel, les fonctions continues φ_j données par (2).

La construction du système (3) pour le problème aux limites différentiel exige le calcul de certains produits scalaires en $L_2(\Omega)$.

Nous supposons que les fonctions de poids ψ_j sont arbitraires, mais qu'elles jouissent de certaines propriétés, dont certaines ont été spécifiées et les autres seront mentionnées ultérieurement.

Calcul du produit scalaire $(U, \varphi_j)_{L_2(\Omega)}$. Nous avons

$$\begin{aligned} (U, \varphi_j)_{L_2(\Omega)} &= \int_a^b U(x) \varphi_j(x) dx = \int_a^b \sum_{k=0}^N U_k(x) \varphi_j(x) dx = \\ &= U_{j-1} \int_{x_{j-1}}^{x_j} \varphi_{j-1} \varphi_j dx + U_j \int_{x_{j-1}}^{x_{j+1}} \varphi_j \varphi_j dx + U_{j+1} \int_{x_j}^{x_{j+1}} \varphi_{j+1} \varphi_j dx. \end{aligned}$$

Pour simplifier les calculs on passera à la variable $s = \frac{x - x_j}{h}$.

En utilisant les formules données en [3]

$$\begin{aligned} (4) \quad \varphi_{j-1}(s) &= 1 - \varphi_j(s) \quad , \quad s \in [-1, 0] \\ \varphi_{j+1}(s) &= 1 - \varphi_j(s) \quad , \quad s \in [0, 1] \end{aligned}$$

nous obtenons

$$(U, \varphi_j)_{L_2(\Omega)} = h \left[U_{j-1} \int_{-1}^0 (1 - \varphi_j) \varphi_j ds + U_j \int_{-1}^1 \varphi_j \varphi_j ds + U_{j+1} \int_0^1 (1 - \varphi_j) \varphi_j ds \right]$$

On imposera à la fonction de poids ψ la condition de

normalisation

$$(5) \quad \int_{-1}^1 \Psi(s) \, ds = 1$$

et on introduira les nombres réels

$$(6) \quad \alpha_1 = 1 - \int_{-1}^1 \varphi(s) \Psi(s) \, ds, \quad \beta_1 = \int_{-1}^1 (1 - \varphi) \Psi \operatorname{sgn}(s) \, ds$$

et les opérateurs (opérateurs des différences finies)

$$(7) \quad \mu U_j = \frac{1}{2} (U_{j+1} - U_{j-1}), \quad \delta^2 U_j = U_{j+1} - 2U_j + U_{j-1}.$$

On obtient la délimitation opératorielle

$$(8) \quad (U, \Psi_j)_{L_2(\Omega)} = h(1 + \beta_1 \mu + \frac{1}{2} \alpha_1 \delta^2) U_j.$$

Pour α_1 et β_1 nous avons les expressions

$$\alpha_1 = a_1 + a_2, \quad \beta_1 = -(a_1 - a_2)$$

où

$$(9) \quad \begin{aligned} a_1 &= - \int_{-1}^0 s \exp(r(1+s)) \Psi(s) \, ds \\ a_2 &= \int_0^1 [1 - (1-s) \exp(rs)] \Psi(s) \, ds \end{aligned}$$

Calcul du produit scalaire $(U', \Psi_j)_{L_2(\Omega)}$. Nous désignons par

U' la dérivée de la fonction d'interpolation U , et nous obtenons

$$\begin{aligned} (U', \Psi_j)_{L_2(\Omega)} &= \int_a^b U' \Psi_j \, dx = \int_a^b \sum_{k=0}^N U_j \varphi'_k \Psi_j \, dx = \\ &= U_{j-1} \int_{x_{j-1}}^{x_j} \varphi'_{j-1} \Psi_j \, dx + U_j \int_{x_{j-1}}^{x_{j+1}} \varphi'_j \Psi_j \, dx + U_{j+1} \int_{x_j}^{x_{j+1}} \varphi'_{j+1} \Psi_j \, dx. \end{aligned}$$

On introduit la variable s et l'on a les formules

$$\varphi'_{j-1}(s) = - \varphi'_j(s) \quad , \quad s \in [-1, 0]$$

$$\varphi'_{j+1}(s) = - \varphi'_j(s) \quad , \quad s \in [0, 1]$$

déduites de (4). On trouve ainsi la formule

$$\begin{aligned} (U', \psi_j)_{L_2(\Omega)} &= - U_{j-1} \int_{-1}^0 \varphi'_j(s) \psi_j(s) ds + U_j \int_{-1}^1 \varphi'_j(s) \psi_j(s) ds - \\ &- U_{j+1} \int_0^1 \varphi'_j(s) \psi_j(s) ds. \end{aligned}$$

On introduit les opérateurs (7) et les nombres réels

$$(10) \quad \alpha_2 = \int_{-1}^1 \varphi'(s) \psi(s) ds \quad , \quad \beta_2 = \int_{-1}^1 \varphi'(s) \psi(s) \operatorname{sgn}(s) ds .$$

On trouve la formule opératorielle

$$(11) \quad (U', \psi_j)_{L_2(\Omega)} = - \left(\beta_2 \mu + \frac{1}{2} \alpha_2 \delta^2 \right) U_j$$

Pour les nombres α_2 et β_2 nous obtenons les expressions

$$\alpha_2 = b_1 + b_2 \quad , \quad \beta_2 = - (b_1 - b_2) \quad ,$$

où

$$\begin{aligned} b_1 &= \int_{-1}^0 (1 + rs) \exp(r(1+s)) \psi(s) ds \\ (12) \quad b_2 &= \int_0^1 (-1 + r - rs) \exp(rs) \psi(s) ds . \end{aligned}$$

Calcul du produit scalaire $(U', \psi'_j)_{L_2(\Omega)}$. Nous avons

$$\begin{aligned} (U', \psi'_j)_{L_2(\Omega)} &= \int_a^b U'(x) \psi'_j(x) dx = \int_a^b \sum_{k=0}^N U_k \varphi'_k(x) \psi'_j(x) dx = \\ &= U_{j-1} \int_{x_{j-1}}^{x_j} \varphi'_{j-1} \psi'_j dx + U_j \int_{x_{j-1}}^{x_{j+1}} \varphi'_j \psi'_j dx + U_{j+1} \int_{x_j}^{x_{j+1}} \varphi'_{j+1} \psi'_j dx . \end{aligned}$$

En procédant de manière similaire à celle des cas précédents,

en introduisant les opérateurs (7) et les nombres réels

$$(13) \quad \begin{aligned} \alpha_3 &= \int_{-1}^1 \varphi'(s) \psi'(s) \, ds \\ \beta_3 &= \int_{-1}^1 \varphi'(s) \psi'(s) \operatorname{sgn}(s) \, ds \end{aligned}$$

on obtient la formule opératorielle

$$(14) \quad (U', \psi_j')_{L_2(\Omega)} = -\frac{1}{h} \left(\beta_3 \mu + \frac{1}{2} \alpha_3 \delta^2 \right) u_j.$$

Pour les nombres α_3 et β_3 nous avons les expressions

$$\alpha_3 = c_1 + c_2, \quad \beta_3 = -(c_1 - c_2)$$

où

$$(15) \quad \begin{aligned} c_1 &= \int_{-1}^0 (1 + rs) \exp(r(1+s)) \psi'(s) \, ds \\ c_2 &= \int_0^1 (-1 + r - rs) \exp(rs) \psi'(s) \, ds \end{aligned}$$

Méthode de Galerkin ($\Psi = \varphi$). Supposons que les fonctions de poids (test) ψ_j sont identiques aux fonctions de base φ_j si les intervalles sont les mêmes. En ce cas, on calcule les coefficients α et β avec les formules précédentes, les fonctions φ_j données par (2) et $\psi_j \equiv \varphi_j$. On obtient les valeurs

$$(16) \quad \begin{aligned} \alpha_1 &= \alpha_{1G} = a_{1G} + a_{2G} = \frac{1}{r} \left(-1 - \frac{1}{r} + \frac{1}{2r^2} + \frac{2}{r} e^r - \frac{1}{2r^2} e^{2r} \right) \\ \beta_{1G} &= -(a_{1G} - a_{2G}) = 0 ; (a_{1G} = a_{2G}) \\ \alpha_{2G} &= b_{1G} + b_{2G} = \frac{1}{2} - \frac{1}{2} = 0 ; (b_{2G} = -b_{1G}) \\ \beta_{2G} &= -(b_{1G} - b_{2G}) = -1 \\ \alpha_{3G} &= c_{1G} + c_{2G} = 2 c_{1G} = \frac{1}{2r} \left[-1 + 2r(1 - r) + e^{2r} \right] \\ \beta_{3G} &= -(c_{1G} - c_{2G}) = 0 ; (c_{1G} = c_{2G}) \end{aligned}$$

Nous avons $\lim_{r \rightarrow 0} \alpha_{1G}(r) = \frac{1}{3}$ et $\lim_{r \rightarrow 0} \alpha_{3G}(r) = 2$.

Méthode de Galerkin généralisée ($\Psi \neq \varphi$). Les φ_j étant des fonctions arbitraires et les ψ_j étant des fonctions données par (2), on impose les conditions suivantes (condition de normalisation de Ψ et conditions supplémentaires) :

$$(17) \quad \begin{aligned} \int_{-1}^1 \Psi(s) ds &= 1 \\ \int_{-1}^1 \varphi'(s) \Psi'(s) \operatorname{sgn}(s) ds &= 0 \\ \int_{-1}^1 \varphi'(s) \Psi'(s) ds &= \alpha_{3G} \\ \int_{-1}^1 \varphi'(s) \Psi(s) \operatorname{sgn}(s) ds &= -1 \end{aligned}$$

On obtient alors les produits scalaires

$$(18) \quad \begin{aligned} (u, \varphi_j)_{L_2(\Omega)} &= h \left(1 + \beta_1 \mu + \frac{1}{2} \alpha_1 \delta^2 \right) u_j \\ (u, \psi_j)_{L_2(\Omega)} &= \left(\mu - \frac{1}{2} \alpha \delta^2 \right) u_j, \quad \alpha \equiv \alpha_2 \\ (u, \psi'_j)_{L_2(\Omega)} &= -\frac{1}{2h} \alpha_{3G} \delta^2 u_j \end{aligned}$$

Fonctions de poids (fonctions test). Nous définirons un système simple de fonctions de poids ψ par la formule

$$\psi(s) = \varphi(s) + k \bar{\sigma}(s), \quad k \in \mathbb{R}, \quad s \in [-1, 1]$$

où la constante k et la fonction $\bar{\sigma}$ se déterminent en conformité avec les conditions (17). Par conséquent, la fonction ψ se choisit comme une perturbation de la fonction de base (voir, par exemple [5])

Les conditions (17) se réduisent à

$$(a) \quad \begin{aligned} \int_{-1}^1 \bar{\sigma}(s) ds &= 0, \quad \int_{-1}^1 \varphi'(s) \bar{\sigma}'(s) \operatorname{sgn}(s) ds = 0, \\ \int_{-1}^1 \varphi'(s) \bar{\sigma}'(s) ds &= 0 \end{aligned}$$

et

$$(b) \quad \beta_2 = 2 b_2 - \alpha_2 = -1 \quad \text{ou} \quad b_2 = \frac{1}{2} (\alpha_2 - 1) \quad \text{ou}$$

$$b_{2G} + k \int_0^1 (-1 + r - rs) \exp(rs) \tilde{G}(s) ds = \frac{1}{2} (\alpha_2 - 1) \quad \text{ou}$$

$$k \int_0^1 (-1 + r - rs) \exp(rs) \tilde{G}(s) ds = \frac{1}{2} \alpha_2 .$$

Il résulte de la première condition que $\tilde{G}$ doit être une fonction impaire dans l'intervalle $[-1, 1]$. Nous choisirons $\tilde{G}(s) = \text{sgn}(s)$, $s \in [-1, 1]$ ($\tilde{G}(s) = 0$ pour $|s| > 1$) et la condition (b) nous donne $-k = \alpha_2 / 2$.

Nous obtenons la fonction de poids

$$(19) \quad \psi(s) = \varphi(s) - \frac{1}{2} \alpha \text{sgn}(s) , \quad s \in [-1, 1] \quad (\alpha \equiv \alpha_2)$$

Cas $r = 0$. (approximation linéaire par morceaux). En ce cas les fonctions φ_j données par (2) sont linéaires par morceaux. Avec $r = 0$, les expressions des coefficients α et β se réduisent à

$$(20) \quad \begin{aligned} \alpha_1 &= \int_{-1}^1 |s| \psi(s) ds , & \beta_1 &= \int_{-1}^1 s \psi(s) ds ; \\ \alpha_2 &= - \int_{-1}^1 \psi(s) \text{sgn}(s) ds = \int_0^1 [\psi(s) - \psi(-s)] ds , \\ \beta_2 &= - \int_{-1}^1 \psi(s) ds = -1 ; \\ \alpha_3 &= - \int_{-1}^1 \psi'(s) \text{sgn}(s) ds , & \beta_3 &= - \int_{-1}^1 \psi'(s) ds , \end{aligned}$$

et si l'on impose les conditions

$$(21) \quad \int_{-1}^1 \psi'(s) \text{sgn}(s) ds = -2 , \quad \int_{-1}^1 \psi'(s) ds = 0$$

nous avons $\alpha_3 = 2$, $\beta_3 = 0$.

Pour la fonction test ψ , la forme (19) subsiste.

Au cas du procédé de Galerkin ($\psi = \varphi$) on obtient les valeurs

$$(22) \quad \alpha_1 = \frac{1}{3} , \quad \beta_1 = 0 , \quad \alpha_2 = 0 , \quad \beta_2 = -1 , \quad \alpha_3 = 2 , \quad \beta_3 = 0 .$$

3. Application. Convection et diffusion unidimensionnelle stationnaire.

a) Propriétés des schémas à élément fini de type contraire au courant. Considérons un mouvement d'un fluide visqueux au dessus d'un corps solide (sémipaque plane) avec présence de couche limite. Nous désignons par Oxy le repère attaché à la plaque (l'axe Ox en direction longitudinale de la plaque et l'axe Oy en direction perpendiculaire à la plaque) et par $\vec{v} = (u, v)$ la vitesse du fluide dans la couche. Si la dérivée $\partial / \partial x \equiv 0$, la couche limite est dite asymptotique (en direction Ox) et ses équations sont (au cas incompressible, $\rho, \mu = \text{const.}$ et stationnaire, $\partial / \partial t = 0$)

$$v \frac{du}{dy} = \frac{\mu}{\rho} \frac{d^2 u}{dy^2}, \quad \frac{dv}{dy} = 0, \quad y \in (0, \infty).$$

Nous attacherons à ces équations les conditions aux limites (la surface la surface de la plaque étant perméable au fluide) à transfert de masse :

$$u(0) = 0, \quad v(0) = v_0 \quad \text{ou} \quad v(0) = -v_0 \quad (v_0 > 0 \text{ donné})$$

$$u(\infty) = u_\infty, \quad \left. \frac{du}{dy} \right|_{\infty} = 0.$$

Il résulte donc que $v(y) = \text{constant}$ et nous obtenons pour la vitesse v les formules

$$v(y) = v_0 \quad \text{ou} \quad v(y) = -v_0$$

et pour la vitesse longitudinale $u(y)$ les équations

$$v_0 \frac{du}{dy} = \frac{\mu}{\rho} \frac{d^2 u}{dy^2} \quad \text{ou} \quad -v_0 \frac{du}{dy} = \frac{\mu}{\rho} \frac{d^2 u}{dy^2}$$

($\rho > 0, \mu > 0, v_0 > 0$)

qui correspondent aux deux sens de transfert du fluide à travers la surface du corps solide (de la plaque) ; ces sont des équations du type de la diffusion (par le fait de la présence du terme $\frac{d^2 u}{dy^2}$)

et de la convection (par le fait de la présence du terme $\frac{du}{dy}$), où le coefficient de la dérivée du premier ordre est positif ou négatif et peut prendre des valeurs grandes ou petites par rapport à μ/ρ .

• On considérera l'équation de convection-diffusion unidimensionnelle stationnaire :

$$(23) \quad v \frac{du}{dx} - k \frac{d^2 u}{dx^2} = 0, \quad x \in (a, b) \quad v, k \in \mathbb{R}_+$$

$$u(a) = u_0, \quad u(b) = u_N$$

La méthode résiduelle pondérée à fonctions de poids ψ_j au noeud x_j conduit aux équations algébriques :

$$(24) \quad v (U^i, \psi_j) + k (U^i, \psi_j') = 0, \quad j = 1, 2, \dots, N-1.$$

Après remplacement des produits scalaires par leurs expressions (18), le système (24) se transforme en le schéma de type contraire au courant (upwind scheme) à élément fini (μ et δ^2 étant les opérateurs des différences finies introduits par les formules (7))

$$2hv \mu U_j - (k(\alpha_{3G} + \frac{vh}{k} \alpha)) \delta^2 U_j = 0, \quad j = 1, 2, \dots, N-1$$

ou

$$(25) \quad [\alpha_{3G} + 2P(1+\alpha)] U_{j-1} - 2(\alpha_{3G} + 2P\alpha) U_j + [\alpha_{3G} - 2P(1-\alpha)] U_{j+1} = 0$$

$$P = \frac{vh}{2k}, \quad j = 1, 2, \dots, N-1$$

où P est le nombre de Péclet pour l'élément de longueur h et U_0 et U_N sont des valeurs connues (grâce aux conditions aux limites). Le schéma (25) représente une famille à un paramètre α à déterminer.

On cherche une solution de (25) de la forme $U_j = \lambda^j$, $\lambda \in \mathbb{R}$. L'équation caractéristique en λ admet les racines λ_1 et λ_2 et l'équation (25) a la solution générale

$$(26) \quad U_j = A \lambda_1^j + B \lambda_2^j = A \left[\frac{\alpha_{3G} + 2P(1+\alpha)}{\alpha_{3G} - 2P(1-\alpha)} \right]^j + B, \quad j = 1, 2, \dots, N-1$$

où A et B sont des constantes arbitraires.

On remarque que si le dénominateur de la fraction est négatif (avec $\alpha_{3G} > 0$, $P > 0$), les valeurs U_j , $j = 1, 2, \dots, N-1$ ont des signes alternatifs. Par conséquent, pour obtenir une solution non-oscillatoire (stable, sans oscillations fausses), il faudra remplir la condition de stabilité

$$(27) \quad |\alpha| > 1 - \frac{\alpha_{3G}}{2|P|}, \quad \text{pour } 2|P| \geq \alpha_{3G}; \quad (P\alpha > 0, \alpha_{3G} > 0)$$

On a considéré ici le cas plus général où P peut être positif ou négatif. Si α ne remplit pas cette condition, le signe de U_j changera d'un noeud à l'autre.

On voit en (27) que le paramètre α est du même signe que P, c'est-à-dire du même signe que la vitesse v. Ce paramètre, choisi de manière à remplir la condition ci-dessus, a le rôle d'éliminer les oscillations de U_j au cas où l'on utilise la méthode de Galerkin généralisée (Galerkin-Petrov); mais il faut le choisir de telle manière que la solution numérique U_j soit obtenue avec une certaine précision.

Procédé Galerkin ($\psi = \varphi$). En ce cas nous avons $\alpha = \alpha_{2G} = 0$. Le système (25), sa solution (26) et la condition de stabilité (27) deviennent

$$(\alpha_{3G} + 2P)U_{j-1} - 2\alpha_{3G}U_j + (\alpha_{3G} - 2P)U_{j+1} = 0, \quad j = 1, 2, \dots, N-1$$

$$(28) \quad U_j = A \left[\frac{\alpha_{3G} + 2P}{\alpha_{3G} - 2P} \right]^j + B$$

$$2|P| < \alpha_{3G}$$

Considérons le problème bilocal $u(a) = U_0$ et $u(b) = U_N$. En ces conditions, la solution (28) devient

$$\frac{U_j - U_0}{U_N - U_0} = \frac{1 - \left[\frac{\alpha_{3G} + 2P}{\alpha_{3G} - 2P} \right]^j}{1 - \left[\frac{\alpha_{3G} + 2P}{\alpha_{3G} - 2P} \right]^N} ; \quad \left(\frac{U_j - U_0}{U_N - U_0} = Q(\alpha_{3G} - 2P)^{N-j} \right)$$

On voit que pour $2P = \alpha_{3G}$ nous avons $U_j = U_0$, $\forall j \neq N$.
 Pour $j = N$ nous trouvons $U_j = U_N$.

Considérons le cas $2P \gg \alpha_{3G}$.

En ce cas, en écrivant $\alpha_{3G} / 2P = \varepsilon$, nous obtenons immédiatement, exactement comme en [4]

$$\frac{U_j - U_0}{U_N - U_0} \approx \frac{1 - (-1)^j (1 + \varepsilon)^{2j}}{1 - (-1)^N (1 + \varepsilon)^{2N}} \approx \frac{1 - (-1)^j (1 + 2j\varepsilon)}{1 - (-1)^N (1 + 2N\varepsilon)}$$

Le comportement oscillatoire de la solution (les oscillations numériques) se décide par rapport à la parité de N et de j , notamment

1°) N est impair, nous avons la solution

$$\frac{U_j - U_0}{U_N - U_0} \approx \begin{cases} -\frac{j\varepsilon}{1 + N\varepsilon} & , \text{ si } j \text{ est pair} \\ \frac{1 + j\varepsilon}{1 + N\varepsilon} & , \text{ si } j \text{ est impair} \end{cases}$$

2°) N est pair, nous avons la solution, [4]

$$\frac{U_j - U_0}{U_N - U_0} \approx \begin{cases} \frac{j}{N} & , \text{ si } j \text{ est pair} \\ -\left(\frac{1}{N\varepsilon} + \frac{j}{N} \right) & , \text{ si } j \text{ est impair} \end{cases}$$

Cas $r = 0$ (approximation linéaire par morceaux). En ce cas nous avons $\alpha_{3G} = 2$ (et $\alpha_3 = 2$) et la solution du système (25) est

$$(29) \quad U_j = A \left[\frac{1 + P(1 + \alpha)}{1 - P(1 - \alpha)} \right]^j + B, \quad j = 1, 2, \dots, N-1$$

à la condition de stabilité

$$|\alpha| > 1 - \frac{1}{|P|}, \quad \text{pour } |P| > 1; \quad (P\alpha > 0)$$

Dans le procédé de Galerkin ($\psi = \varphi$) avec $r = 0$ nous avons $\alpha = 0$ ($\alpha_{3G} = 2$) ; l'équation (25), la solution (26) et la condition de stabilité se réduisent à, [4]

$$(1 + P)U_{j-1} - 2U_j + (1 - P)U_{j+1} = 0, \quad j = 1, 2, \dots, N-1$$

$$(30) \quad U_j = A \left[\frac{1+P}{1-P} \right]^j + B$$

$$|P| < 1 \quad \text{ou} \quad h < \frac{2k}{|v|}$$

Remarques. Le schéma (25) associé à l'équation (23) reproduit :

1°) l'équation à différences finies centrales (l'équation (28)) si $\alpha = 0$, $\alpha_{3G} = 2$; (avec une solution qui a des oscillations numériques)

2°) les schémas à différences finies de type contraire au courant (en amont) si $\alpha_{3G} = 2$ et $\alpha = \pm 1$ ($\alpha = -1$ correspond à $v < 0$ et $\alpha = 1$ correspond à $v > 0$) ; en ce cas la solution n'a pas des oscillations numériques. En effet, si en (23) on utilise la différence centrale pour la dérivée seconde et les différences latérales

$$(dU / dx)_j = (U_{j+1} - U_j) / h \quad \text{si } v < 0$$

$$(dU / dx)_j = (U_j - U_{j-1}) / h \quad \text{si } v > 0$$

on obtient les schémas

$$(31) \quad U_{j-1} - 2(1 + |P|)U_j + (1 + 2|P|)U_{j+1} = 0, \quad v < 0$$

$$(32) \quad (1 + 2|P|)U_{j-1} - 2(1 + |P|)U_j + U_{j+1} = 0, \quad v > 0$$

$$(|P| = \frac{|v|h}{2k}, \quad v > 0, \quad h > 0)$$

Les schémas (31)-(32) s'obtiennent directement de (25) si l'on y remplace α par -1 respectivement par 1 .

b) Procédé de l'assemblage (matrices de convection et diffusion)

I. Matrices et équation de Galerkin pour un élément. Considérons la division $a = x_0 < x_1 < \dots < x_N = b$ au pas h dans l'intervalle $[a, b]$ et l'élément $\Omega^{(e)} = [x_j, x_{j+1}]$ dont nous désignerons les noeuds par 1 et 2 (numérotation locale).

Nous choisissons sur l'élément $\Omega^{(e)}$ l'approximation

$$U(x) = u_N(x) = \sum_{k=1}^2 u_k \varphi_k(x)$$

et les fonctions de poids ψ_1 et ψ_2 .

Pour l'élément $\Omega^{(e)}$, la méthode des résidus pondérés (méthode de Galerkin généralisée) conduit selon (24) au système (1, indice de sommation)

$$(33) \quad (c_{ji} + d_{ji}) U_i = 0, \quad j = 1; 2 \quad i = 1; 2$$

$$(34) \quad c_{ji} = v(\psi_j, \varphi'_i), \quad d_{ji} = k(\psi'_j, \varphi'_i)$$

ou à l'équation matricielle

$$(35) \quad (C + D) U = 0, \quad C = [c_{ji}], \quad D = [d_{ji}], \quad U = (U_1, U_2)^T$$

Nous avons pour la colonne des fonctions de base et des fonctions de poids

$$\begin{aligned} \underline{\varphi} = \begin{Bmatrix} \varphi_1 \\ \varphi_2 \end{Bmatrix} &= \begin{Bmatrix} (1-s)e^{+\frac{x}{h}} \\ 1-(1-s)e^{+\frac{x}{h}} \end{Bmatrix}; \quad \underline{\psi} = \begin{Bmatrix} \psi_1 \\ \psi_2 \end{Bmatrix} = \begin{Bmatrix} \varphi_1 - \frac{1}{2}\alpha \\ \varphi_2 + \frac{1}{2}\alpha \end{Bmatrix} = \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{Bmatrix} \varphi_1 \\ \varphi_2 \end{Bmatrix} + \frac{\alpha}{2} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \begin{Bmatrix} \varphi_1 \\ \varphi_2 \end{Bmatrix} = M \underline{\varphi}; \quad M = I + \frac{\alpha}{2} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Avec ces matrices, les coefficients c_{ji} et d_{ji} ont les valeurs

$$\begin{aligned}
 c_{11} &= v(\psi_1, \varphi'_1) = v \int_{x_1}^{x_{1+1}} \psi_1 \varphi'_1 dx = v \int_0^1 (-1+r-rs) e^{rs} \psi_1 ds = \\
 &= v b_2 = v \left(-\frac{1}{2} + \frac{\alpha}{2} \right) \\
 c_{12} &= v(\psi_1, \varphi'_2) = -v(\psi_1, \varphi'_1) = -v b_2 = v \left(\frac{1}{2} - \frac{\alpha}{2} \right) \\
 c_{21} &= v(\psi_2, \varphi'_1) = v \int_0^1 (-1+r-rs) e^{rs} \psi_2 ds = v \left(-\frac{1}{2} - \frac{1}{2} \alpha \right) \\
 (36) \quad c_{22} &= v(\psi_2, \varphi'_2) = -v(\psi_2, \varphi'_1) = v \left(\frac{1}{2} + \frac{1}{2} \alpha \right) \\
 d_{11} &= k(\psi'_1, \varphi'_1) = \frac{k}{h} \int_0^1 (-1+r-rs) e^{rs} \psi'_1(s) ds = \frac{k}{h} c_2 = \frac{k}{h} c_{2G} \\
 d_{12} &= k(\psi'_1, \varphi'_2) = -k(\psi'_1, \varphi'_1) = -\frac{k}{h} c_2 = -\frac{k}{h} c_{2G} \\
 d_{21} &= k(\psi'_2, \varphi'_1) = \frac{k}{h} \int_0^1 (-1+r-rs) e^{rs} \psi'_2(s) ds = -\frac{k}{h} c_{2G} \\
 d_{22} &= k(\psi'_2, \varphi'_2) = \frac{k}{h} c_{2G}
 \end{aligned}$$

Les matrices de convection et diffusion pour $\Omega^{(e)}$ sont

$$(37) \quad C = \frac{v}{2} \begin{bmatrix} -1+\alpha & 1-\alpha \\ -1-\alpha & 1+\alpha \end{bmatrix} = M C_g \quad \text{avec} \quad C_g = \frac{v}{2} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$(38) \quad D = \frac{k \mathcal{L}_{2G}}{h} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = M D_g \quad \text{avec} \quad D_g = D$$

où α est un paramètre arbitraire et c_{2G} est donné par (16) et a la valeur

$$(39) \quad c_{2G} = \frac{1}{4r} \left[-1 + 2r(1-r) + e^{2r} \right]$$

On remarque que D est une matrice symétrique et qu'elle ne dépend pas du paramètre α . Par conséquent elle est identique à la matrice du cas $\alpha = 0$ (cas du procédé de Galerkin). D'autre part, la matrice

On ne dépend pas de r , c'est-à-dire qu'elle est identique à celle du cas $r = 0$ (cas de l'approximation linéaire).

L'équation de Galerkin sous la forme matricielle pour l'élément $\Omega^{(e)}$, c'est-à-dire l'équation (35) devient

$$(40) \quad \begin{bmatrix} \frac{v}{2}(-1+\alpha) + \frac{k c_{2G}}{h} & \frac{v}{2}(1-\alpha) - \frac{k c_{2G}}{h} \\ -\frac{v}{2}(1+\alpha) - \frac{k c_{2G}}{h} & \frac{v}{2}(1+\alpha) + \frac{k c_{2G}}{h} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$$

Nous récrivons l'équation (40) sous la forme symbolique

$$(40') \quad \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$$

II. Assemblage. Le procédé habituel d'assemblage des équations matricielles des éléments conduit à l'équation matricielle tridiagonale globale

$$(41) \quad \begin{bmatrix} k_{12}+k_{11} & k_{12} & & & \\ k_{21} & k_{12}+k_{11} & k_{12} & & \\ & k_{21} & k_{12}+k_{11} & k_{12} & \\ & & k_{21} & k_{12}+k_{11} & k_{12} \\ 0 & & & k_{21} & k_{12}+k_{11} \end{bmatrix} \begin{Bmatrix} U_1 \\ U_2 \\ \vdots \\ U_j \\ \vdots \\ U_{N-1} \end{Bmatrix} = \begin{Bmatrix} -k_{21} U_0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ -k_{12} U_N \end{Bmatrix}$$

Cette équation matricielle peut s'écrire sous la forme du système suivant

$$(41') \quad k_{21} U_{j-1} + (k_{22}+k_{11}) U_j + k_{12} U_{j+1} = 0, \quad j = 1, 2, \dots, N-1.$$

Ayant, selon (16), $\alpha_{3G} = 2c_{2G}$ et en posant $P = \frac{v h}{2k}$, le système (41') s'écrit, après remplacement des éléments k_{ji} , sous la forme (25).

La solution de ce système est donnée par (26) et par la condition de stabilité (27) relative à α et P.

c) Exemple numérique (convection dominante).

$$v \frac{du}{dx} - k \frac{d^2u}{dx^2} = 0, \quad x \in (0,1)$$

$$u(0) = 1, \quad u(1) = 0$$

Table 1.

($v/k = 60$; $h = 1/40$)

x	Solution exacte	Méthode de Galerkin à fonctions de base combinées		Méthode de Galerkin à fonctions de base linéaires,[6]
		$r = 2/3$	$r = -2/3$	$r = 0$
0.90	0.9975	0.9969	0.9981	0.9996
0.925	0.9889	0.9868	0.9909	0.9971
0.95	0.9502	0.9441	0.9564	0.9796
0.975	0.7769	0.7736	0.7911	0.8571
1.00	0	0	0	0

Table 2.

($v/k = 60$; $h = 1/20$)

x	Solution exacte	Méthode de Galerkin à fonctions de base combinées	Méthode de Galerkin à fonctions de base linéaires,[6]
		$r = 1$	$r = 0$
0.90	0.9975	0.9990	0.9600
0.95	0.9502	0.9686	1.2000
1.00	0	0	0

BIBLIOGRAPHIE

1. AHUËS M., TÉLIAS M., Méthodes d'éléments finis pour l'équation de diffusion-convection,
Séminaire IMAG n°. 370 - 1981.
2. BELYTSKO T., ELDIB I., Analysis of a finite element upwinding scheme,
Finite Element Methods for Convection Dominated Flows,
AMD - vol.34, 1979, p.195-200.
3. BRADEANU D., Fonctions combinées d'interpolation par morceaux et leurs propriétés , (dans ce Preprint)
4. GRESHO F.M., LEE R.L., Don't suppress the wigles - they're telling you something,
Finite Element Methods for Convection Dominated Flows,
AMD - vol.34, 1979, p.37-61.
5. GRIFFITS D.F., MITCHELL A.R., On generating upwind finite element upwinding scheme,
Finite Element Methods for Convection Dominated Flows,
AMD - vol.34, 1979, p.91-104.
6. MITCHELL A.R., WAIT R., The finite element method in partial differential equations, (russian)
Izd. "Mir" Moskva, 1981.
7. МАРЧУК Г.И., АГОШКОВ В.И., Введение в проекционно-сеточные методы,
„Наука“, Москва, 1981.