
Méthodes d'éléments finis pour les problèmes de convection-diffusion

Jean Donea

*Division mécanique appliquée
Centre commun de recherche
Commission des communautés européennes
I-21020 Ispra (Va), Italie*

RÉSUMÉ. *Cet article est consacré à une présentation succincte des méthodes de Galerkin généralisées adaptées à la résolution par éléments finis de problèmes de transport par convection. La première partie traite le cas de problèmes stationnaires de transport par convection et diffusion. L'accent est mis sur l'utilisation de méthodes du type Petrov-Galerkin pour reproduire dans le cadre de la méthode des éléments finis l'effet upwind utilisé antérieurement avec succès en différences finies. La deuxième partie est consacrée au cas de problèmes transitoires de convection pure et, de manière plus générale, aux problèmes gouvernés par des équations hyperboliques du premier ordre. On examine successivement la mise en œuvre de méthodes faisant explicitement usage des courbes caractéristiques, de techniques du type Petrov-Galerkin, de méthodes de Taylor-Galerkin, pour conclure par une présentation de méthodes de moindres carrés.*

ABSTRACT. *This paper presents a brief overview of generalized Galerkin methods for the finite element solution of transport problems governed by convection and diffusion. The first part is concerned with steady problems. The emphasis is placed upon the use of Petrov-Galerkin methods to reproduce in the finite element context the upwind effect previously used with success in finite differences. The second part deals with transient problems describing pure convection and, more generally, with situations governed by first-order hyperbolic equations. The methods discussed include techniques making an explicit use of the characteristic curves, Petrov-Galerkin and Taylor-Galerkin methods, as well as least-squares methods.*

MOTS-CLÉS : *modélisation, éléments finis, convection, diffusion, caractéristiques, Petrov-Galerkin, Taylor-Galerkin, moindres carrés.*

KEY WORDS : *modelling, finite elements, convection, diffusion, characteristics, Petrov-Galerkin, Taylor-Galerkin, least-squares.*

1 Introduction

Introduite vers la fin des années 1950, la méthode des éléments finis s'est désormais imposée comme une des méthodes numériques les plus puissantes

conçues à ce jour. Parmi les caractéristiques de base de la méthode qui ont conduit à sa popularité actuelle, on mentionnera la facilité de modéliser des géométries de forme complexe, le traitement naturel des conditions aux limites de type différentiel, ainsi que la possibilité d'être programmée sous la forme de logiciels adaptables au traitement d'une vaste gamme de problèmes.

La formulation classique de la méthode des éléments finis repose sur la méthode des résidus pondérés de Galerkin. Celle-ci consiste à choisir comme fonctions de pondération les fonctions de forme retenues pour la description du champ local de la variable inconnue du problème. La formulation de Galerkin s'est révélée extrêmement efficace en application aux problèmes de mécanique des structures et dans d'autres situations, telle que le transfert conductif de la chaleur, gouvernées par des équations du type diffusion. La raison de ce succès est qu'en application à des problèmes gouvernés par des équations aux dérivées partielles elliptiques ou paraboliques, la méthode des éléments finis de Galerkin conduit à des matrices de rigidité symétriques. Or dans ce cas on peut démontrer mathématiquement que la solution numérique possède la propriété de meilleure approximation. Ceci signifie que la différence entre la solution obtenue par la méthode des éléments finis et la solution exacte est minimisée par rapport à une certaine norme (voir par exemple Strang et Fix [STF 73]).

Ce succès de la méthode de Galerkin a suscité, au début des années 1970, une forte impulsion en vue de son utilisation dans la simulation numérique de problèmes de la mécanique des fluides. On pouvait, en effet, penser que les avantages de la méthode des éléments finis qui s'étaient manifestés dans la résolution de problèmes de structures et de transport par diffusion seraient directement exploitables pour la modélisation des problèmes d'écoulement de fluides. En fait, ceci se révéla une vue trop optimiste, particulièrement en ce qui concerne la simulation des phénomènes de transport dominés par la convection.

Une difficulté importante apparut en raison de la présence de termes convectifs dans la formulation mathématique des lois de conservation de la masse, de la quantité de mouvement et de l'énergie lorsqu'on fait appel, comme c'est pratiquement toujours le cas en mécanique des fluides, à une description cinématique autre que la description Lagrangienne. Les opérateurs de convection sont, en effet, non symétriques, si bien que la propriété de meilleure approximation de la méthode de Galerkin, base de son succès dans les cas symétriques, est perdue lorsque la convection domine le processus de transport.

En pratique, les solutions numériques délivrées par la méthode de Galerkin pour des problèmes à convection dominante sont fréquemment polluées par des oscillations non-physiques que l'on ne peut supprimer qu'en affinant considérablement le maillage, ou, dans le cas transitoire, en réduisant fortement le pas de temps. Ceci compromet, bien entendu, l'utilité même de la méthode de Galerkin et a motivé le développement de méthodes alternatives excluant la présence d'oscillations non-physiques indépendamment de tout raffinement du maillage ou du pas de temps.

Nous débiterons par un exposé des méthodes de Galerkin généralisées

développées pour le traitement de problèmes stationnaires de convection-diffusion (voir également la revue présentée en [DON 91]).

2 Problèmes stationnaires

2.1 Inconvénients des méthodes standards

Un problème simple sera utilisé pour illustrer la nécessité d'une généralisation de la méthode de Galerkin standard pour pouvoir traiter correctement les problèmes dominés par la convection. Considérons l'équation de convection-diffusion monodimensionnelle

$$-au_x + ku_{xx} = 0 \quad (1)$$

où la vitesse d'écoulement a et le coefficient de diffusion k sont tous deux supposés positifs. Pour un maillage uniforme d'éléments linéaires et en numérotant les noeuds consécutivement, l'approximation de Galerkin relative à l'Eq.(1) fournit l'équation discrète suivante pour tout noeud intérieur I :

$$(1 - \gamma)u_{I+1} - 2u_I + (1 + \gamma)u_{I-1} = 0 \quad (2)$$

où

$$\gamma = \frac{ah}{2k} \quad (3)$$

est le nombre de Péclet de maille et h indique la longueur d'un élément. L'équation (2) indique clairement que la méthode des éléments finis de Galerkin engendre des approximations discrètes des opérateurs différentiels qui sont du type différences centrées. Comme toute équation discrète, l'équation (2) possède une erreur de troncature. Pour analyser cette erreur, nous développons l'équation (2) en série de Taylor autour de u_I et utilisons l'équation différentielle originale (1) pour remplacer les dérivées d'ordre supérieur en termes de u_{xx} . Ce développement indique que l'équation discrète (2) est équivalente à

$$(-au_x + ku_{xx})_I + \frac{k}{2\gamma} \left(\frac{1}{\gamma} (\cosh 2\gamma - 1) - \sinh 2\gamma \right) (u_{xx})_I = 0 \quad (4)$$

si bien que le second terme de l'équation (4) représente exactement l'erreur de troncature de l'équation discrète (2). Cette erreur se présente sous la forme d'un opérateur de diffusion et on montre aisément que la fonction dans l'Eq.(4) faisant intervenir le nombre de Péclet γ est systématiquement négative quelle que soit la valeur de γ . En conséquence, il apparaît que l'équation discrète (2) résultant de l'approximation de l'équation (1) résout en fait une équation modifiée à *diffusion réduite* que l'on peut écrire sous la forme:

$$-au_x + (k - \bar{k})u_{xx} = 0 \quad (5)$$

où k est la diffusion physique et \bar{k} une diffusion artificielle d'origine purement numérique. La dissipation numérique négative inhérente à l'approximation de

l'équation (1) par la méthode de Galerkin (et par différences centrées) est la cause des difficultés numériques (oscillations) rencontrées dans la résolution de problèmes dominés par la convection. Lorsque le nombre de Péclet de maille est supérieur à l'unité, la diffusion artificielle \bar{k} est plus grande que la diffusion physique k , ce qui, bien entendu, rend la solution dépourvue de sens.

Pour se convaincre de l'apparition d'oscillations lorsque le nombre de Péclet est supérieur à l'unité, on écrira l'équation (2) sous la forme:

$$(1 - \gamma)(u_{I+1} - u_I) = (1 + \gamma)(u_I - u_{I-1})$$

qui montre bien que les pentes de la solution à gauche et à droite du noeud I sont effectivement de signe opposé lorsque $\gamma \geq 1$.

Puisque l'approximation de Galerkin (et celle des différences centrées) relative à l'équation (1) introduit une diffusion artificielle négative, un remède justifié pour améliorer la solution numérique consiste à ajouter de la diffusion de manière à contrebalancer la diffusion numérique négative.

2.2 Différences amont ou upwind

Dans le cadre de la méthode des *différences finies*, on peut introduire de la diffusion numérique en remplaçant l'approximation discrète centrée du terme convectif de l'équation (1) par une approximation décentrée, ou dérivée amont, qui s'écrit:

$$a(u_x)_I = \frac{a}{h}(u_I - u_{I-1}) \quad (6)$$

Toutefois, dans la plupart des cas, la discrétisation amont du terme convectif donnée par l'équation (6) conduit à des solutions numériques exagérément dissipatives, ce qui a donné lieu à de nombreuses critiques de cette technique.

Etant donné que la méthode des différences centrées introduit une dissipation négative, tandis que la méthode des différences amont produit une diffusion numérique excessive, on conçoit qu'une combinaison linéaire de différences centrées et amont représente une solution optimale. En fait, pour le cas simple de l'équation (1), il est possible de formuler une technique amont optimale conduisant à la solution exacte aux noeuds du maillage quelle que soit la valeur du nombre de Péclet. Cette méthode optimale peut également être construite en ajoutant à la méthode des différences centrées une diffusion artificielle donnée par

$$\bar{k} = (ah/2)\bar{\xi} \quad (7)$$

où

$$\bar{\xi} = \coth(\gamma) - 1/\gamma \quad (8)$$

Toutefois, des problèmes de précision subsistent dans des cas plus compliqués, tels que les situations multidimensionnelles, ou encore en présence de termes source ou de champs de vitesse variables dans l'espace. On consultera les références [SHY 85] et [CHR 85] pour une discussion de ces problèmes et une présentation de méthodes aptes à les surmonter.

2.3 Éléments Finis de type Upwind

Diverses techniques ont été proposées pour réaliser dans le cadre de la *méthode des éléments finis* l'effet produit en différences finies par le décentrage de la dérivée première. Une caractéristique commune à toutes ces techniques est l'abandon de la formulation de Galerkin standard et le recours à une méthode du type *Petrov-Galerkin* où la fonction de pondération est choisie dans une classe de fonctions différente de celle utilisée pour représenter la solution approchée du problème.

2.3.1 Premières Méthodes

Les premières méthodes d'éléments finis basées sur une formulation du type Petrov-Galerkin ont été présentées par les équipes de recherche de Dundee et Swansea (voir, par exemple, Christie et al. [CGM 76], Heinrich et al. [HHZ 77], Heinrich et Zienkiewicz [HZI 79] et Griffiths et Mitchell [GRM 79]). Ces méthodes reposent sur le choix de fonctions de pondération aptes à accentuer le poids de l'élément en amont d'un noeud par rapport à l'élément en aval. Quoique capables de reproduire la solution exacte de l'équation linéaire de convection-diffusion (1), ces premières méthodes d'éléments finis présentent les mêmes inconvénients que les premières méthodes de différences finies du type amont dans la résolutions de problèmes plus compliqués. En outre, la formulation requiert le choix de fonctions de pondération d'ordre polynomial élevé, ce qui complique sa mise en oeuvre par rapport à la méthode de Galerkin classique.

2.3.2 Méthode SUPG

Un progrès important pour le traitement par éléments finis des problèmes stationnaires de convection-diffusion a été réalisé avec l'introduction par Hughes et Brooks [HBR 79, BRH 80, BRH 82] de la méthode dite *SUPG (Streamline Upwind Petrov-Galerkin)*. L'idée de base de cette méthode est de modifier la fonction de pondération de la méthode standard de Galerkin en y ajoutant un terme de perturbation de type amont agissant uniquement dans la direction de l'écoulement et non transversalement. La fonction de pondération ainsi modifiée est appliquée à tous les termes de l'équation, y compris le terme source éventuel, donnant ainsi lieu à une méthode de résidus pondérés *consistante*.

La méthode SUPG possède les qualités de robustesse des techniques de différences finies de type amont, en ce sens qu'elle prévient l'apparition d'oscillations parasites. En outre, étant basée sur une technique de résidus pondérés consistante, la méthode SUPG n'est pas sujette aux critiques de diffusion artificielle manifestées à l'égard des premières méthodes d'éléments finis du type amont.

Pour le problème de convection-diffusion décrit par l'équation (1) avec des conditions aux limites de Dirichlet en $x = 0$ et $x = L$, Hughes et Brooks [HBR 79] remplacent la forme faible classique

$$\int_0^L (wau_x + w_x k u_x) dx = 0 \quad (9)$$

où w désigne la fonction de pondération, par la forme faible suivante:

$$\int_0^L (wau_x + w_x(k + \bar{k})u_x) dx = 0 \quad (10)$$

où

$$\bar{k} = \beta ah/2 \quad (11)$$

représente une diffusion ajoutée dont l'importance est réglée par le paramètre libre β , ($0 \leq \beta \leq 1$). La valeur $\beta = 1$ correspond à un décentrage complet du terme convectif comme dans la relation (6).

En vertu de la définition (11) de la diffusion ajoutée, on constate que la forme faible (10) peut être réécrite sous la forme:

$$\int_0^L \left[\left(w + \beta \frac{h}{2} w_x \right) au_x + w_x k u_x \right] dx = 0 \quad (12)$$

L'équation (12) indique que la méthode de la diffusion ajoutée peut être considérée comme une méthode de Petrov-Galerkin dans laquelle une fonction de pondération *modifiée* donnée par

$$\tilde{w} = w + \beta \frac{h}{2} w_x \quad (13)$$

est appliquée au terme convectif. On notera que cette fonction modifiée est discontinue à la frontière entre les éléments.

Hughes et ses collaborateurs ont ensuite généralisé la méthode SUPG décrite ci-dessus pour traiter le cas d'équations comportant un terme source et pouvoir aborder les problèmes multidimensionnels (voir Brooks et Hughes [BRH 82] et Hughes et Mallet [HMA 86] pour un exposé complet). Ils ont tout d'abord fait remarquer que lorsque l'équation de convection-diffusion comporte un terme source, ce dernier doit être discrétisé, comme le terme convectif, en utilisant la fonction de pondération modifiée donnée par la relation (13). Ensuite, pour étendre la méthode de la diffusion ajoutée au traitement des problèmes multidimensionnels, Hughes et Brooks ont proposé de construire l'opérateur de diffusion ajoutée sous forme tensorielle, de manière qu'il agisse uniquement dans la direction de l'écoulement et non transversalement. Pour ce faire, la diffusion artificielle scalaire de l'équation (11) est remplacée par la diffusion tensorielle

$$\bar{k}_{ij} = \bar{k} a_i a_j / \|a\|^2 \quad (14)$$

où \bar{k} est une diffusion artificielle scalaire et a_i indique la composante de la vitesse selon la direction x_i . On vérifiera aisément que l'équation (14) représente une diffusion agissant uniquement dans la direction de l'écoulement et non transversalement. La méthode SUPG a été appliquée avec succès par Hughes

et ses collaborateurs pour la résolution de problèmes 2D et 3D de convection-diffusion, ainsi que pour la solution des équations de Navier-Stokes en régime incompressible (cf. Hughes et Brooks [BRH 82]). La méthode SUPG peut être considérée comme une méthode de Petrov-Galerkin dans laquelle des fonctions de pondération modifiées données par

$$\tilde{w} = w + \left(\frac{\bar{k}}{\|a\|^2} \right) a \cdot \nabla w \quad (15)$$

sont appliquée de manière consistante au terme convectif et au terme source de l'équation. La méthode SUPG a également été appliquée avec succès au cas de problèmes transitoires. Dans ce cas, la fonction (15) est également appliquée pour la pondération du terme de dérivée temporelle.

La seule faiblesse de la méthode SUPG réside dans la présence d'un paramètre libre gouvernant l'amplitude de la diffusion ajoutée. Mis à part le cas de l'équation linéaire de convection-diffusion, il n'existe, en effet, pas de recette universelle permettant de fixer la valeur de ce paramètre.

2.3.3 Méthode Galerkin-Moindres Carrés

Plus récemment, Hughes et ses collaborateurs [HFH 89] ont introduit la méthode Galerkin-Moindres Carrés (GMC) en tant que technique alternative à la méthode SUPG pour accroître la stabilité de la méthode de Galerkin pour la résolution de problèmes de convection-diffusion.

La méthode GMC consiste à ajouter à la forme intégrale résultant de la méthode de Galerkin standard des termes liés à la minimisation de la fonctionnelle quadratique associée à l'équation de convection-diffusion considérée. On peut démontrer que ces termes additionnels renforcent la stabilité de la méthode standard de Galerkin sans en dégrader la précision. La méthode GMC représente en fait une simplification conceptuelle de la méthode SUPG et est applicable à une gamme de problèmes plus étendue. On consultera, par exemple, la référence [HUG 89] pour une description de l'application de la méthode GMC en mécanique du solide.

3 Problèmes transitoires

Nous passerons à présent en revue quelques méthodes par éléments finis adaptées aux problèmes transitoires de transport convectif. Nous examinerons successivement la mise en oeuvre de méthodes faisant explicitement usage des courbes caractéristiques, de techniques du type Petrov-Galerkin et Taylor-Galerkin, pour conclure par une présentation de méthodes basées sur des techniques de moindres carrés.

3.1 Rôle des caractéristiques

Pour illustrer dans un contexte simple le rôle des courbes caractéristiques dans

la solution des problèmes transitoires de convection pure, considérons l'équation linéaire de convection transitoire à une dimension

$$u_t + au_x = 0 \quad (16)$$

et supposons que la distribution initiale de la quantité scalaire u soit donnée par

$$u(x, 0) = g(x) \quad (17)$$

Dans ces conditions, la solution exacte de l'Eq.(16) est donnée par

$$u(x, t) = g(x - at) \quad (18)$$

ce qui signifie qu'au cours du temps la distribution initiale $g(x)$ de la quantité u subit une translation simple le long de l'axe x avec une vitesse constante a .

Le résultat (18) peut se vérifier en reformulant l'Eq. (16) à l'aide de nouvelles variables X, t , où la coordonnée X se déplace par rapport à x à la vitesse a , soit:

$$X = x - at \quad (19)$$

On obtient ainsi

$$u = u(X(x, t), t) \quad (20)$$

si bien que dans les nouvelles coordonnées (19), l'Eq.(16) s'écrit

$$u_t + (X_t + aX_x)u_X = 0 \quad (21)$$

ou encore, compte-tenu de la relation (19) :

$$u_t(X, t) = 0 \quad (22)$$

Ce résultat confirme que la solution exacte de l'Eq.(16) est bien donnée par l'expression (18).

Les droites d'équation $X = \text{const.}$, qui dans le cas présent coïncident avec les trajectoires des particules, sont appelées *caractéristiques*. Comme le montre l'Eq.(22), la quantité transportée u est constante le long des caractéristiques et cette propriété est largement utilisée dans la construction d'algorithmes pour la résolution numérique des équations hyperboliques dont l'Eq.(16) est le prototype le plus simple.

Après ces considérations simples, nous aborderons une étude plus détaillée des propriétés mathématiques de l'équation transitoire de convection. Dans ce but, considérons encore le cas d'une quantité scalaire u transportée par un champ de vitesse multidimensionnel $\mathbf{a}(\mathbf{x}, t)$ en présence d'un terme *source* prescrit $s(\mathbf{x}, t)$. Dans ce contexte multidimensionnel, l'équation linéaire de convection s'écrit

$$u_t + \mathbf{a} \cdot \nabla u = s \quad (23)$$

et la condition aux limites la plus fréquemment associée à l'Eq. (23) est de la forme

$$u(\mathbf{x}, t)|_{\Gamma_{in}} = u_{\Gamma}(\mathbf{x}_{\Gamma}, t), \quad \mathbf{x}_{\Gamma} \in \Gamma_{in} \quad (24)$$

où Γ_{in} désigne la frontière à flux entrant, c.a.d. la portion de la frontière Γ du domaine Ω considéré sur laquelle $\mathbf{n} \cdot \mathbf{a}(\mathbf{x}_{\Gamma}, t) < 0$, \mathbf{n} désignant le vecteur unitaire extérieur normal à Γ , tandis que $u_{\Gamma}(\mathbf{x}_{\Gamma}, t)$ est une fonction donnée définie sur Γ_{in} pour $t \geq 0$. Strictement parlant, Γ_{in} dépend en général du temps, c.a.d., $\Gamma_{in} = \Gamma_{in}(t)$, mais cette dépendance n'est pas indiquée ici pour simplifier la notation. Pour compléter la définition du problème de convection, une condition initiale est prescrite sous la forme

$$u(\mathbf{x}, 0) = u^0(\mathbf{x}) \quad (25)$$

où $u^0(\mathbf{x})$ est la distribution initiale de l'inconnue u sur le domaine Ω .

La solution analytique du problème linéaire (23-25) en un point donné $\mathbf{x} \in \Omega$ peut ici encore être déterminée à l'aide du concept de *caractéristiques*. La Fig. 1 illustre le cas bidimensionnel: Pour tout point spatio-temporel (\mathbf{x}, t) , on détermine la caractéristique $\mathbf{X}(\tau)$ passant par (\mathbf{x}, t) en résolvant l'équation différentielle (vectorielle) ordinaire

$$\frac{d\mathbf{X}}{d\tau} = \mathbf{a}(\mathbf{X}(\tau), \tau), \quad \tau < t \quad (26)$$

sujette à la condition finale

$$\mathbf{X}(t) = \mathbf{x} \quad (27)$$

Une notation plus précise pour $\mathbf{X}(\tau)$ consisterait à écrire $\mathbf{X}(\mathbf{x}, t; \tau)$. Il convient de noter que l'Eq. (26) est *non linéaire*, sauf lorsque le champ de vitesse

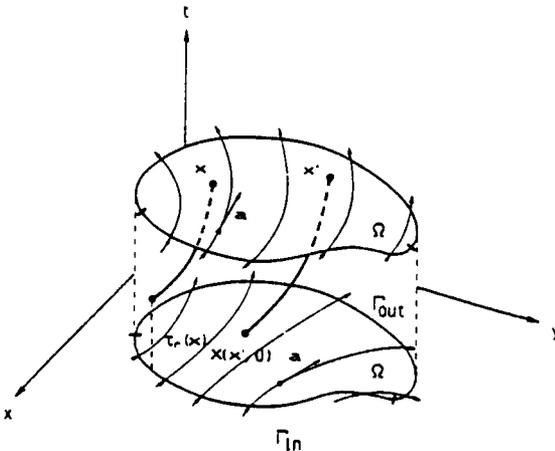


Figure 1: Représentation schématique des courbes caractéristiques pour l'équation linéaire de convection transitoire en présence d'un champ de vitesse variable en fonction du temps.

possède une dépendance spatiale particulière de la forme $\mathbf{a}(\mathbf{x}, t) = \alpha(t) + \beta(t) \mathbf{x}$, où $\alpha(t)$ et $\beta(t)$ sont des fonctions arbitraires. Excluant une telle situation, on peut affirmer que la présence de la non-linéarité est le prix à payer pour la réduction du problème aux dérivées partielles en un problème différentiel ordinaire.

L'intégration de l'Eq. (26) procède de $\tau = t$ en arrière, jusqu'à ce que la caractéristique soit intersecte la frontière Γ_{in} ou atteigne le temps initial $t = 0$ (cf. [PIR 88]).

i) L'intersection de la caractéristique avec Γ_{in} se produira à un instant qui dépend de la caractéristique considérée et que nous noterons $\tau_{\Gamma}(\mathbf{x}, t)$, tel que $\mathbf{X}(\mathbf{x}, t; \tau_{\Gamma}(\mathbf{x}, t)) \in \Gamma_{in}$. Si ce point est dénoté par $\mathbf{X}_{\Gamma}(\mathbf{x}, t)$, on aura

$$\mathbf{X}_{\Gamma}(\mathbf{x}, t) = \mathbf{X}(\mathbf{x}, t; \tau_{\Gamma}(\mathbf{x}, t)) \tag{28}$$

ou encore, omettant l'indication explicite de la dépendance par rapport à (\mathbf{x}, t) pour simplifier l'écriture:

$$\mathbf{X}_{\Gamma} = \mathbf{X}(\tau_{\Gamma}) \tag{29}$$

Dans le cas présent, la solution en (\mathbf{x}, t) dépend par conséquent de la valeur prescrite au point \mathbf{X}_{Γ} de la frontière. Pour déterminer $u(\mathbf{x}, t)$, l'équation de départ est réécrite sous la forme caractéristique. Dénotant la valeur de l'inconnue le long de la caractéristique $\mathbf{X}(\tau)$ par $U(\tau) = u(\mathbf{X}(\tau), \tau)$, l'Eq. (23) se réduit le long de $\mathbf{X}(\tau)$ à l'équation différentielle ordinaire suivante

$$\frac{dU}{d\tau} = S(\tau) \tag{30}$$

dans laquelle $S(\tau) = s(\mathbf{X}(\tau), \tau)$. L'Eq. (30) doit être intégrée jusqu'à $\tau = t$ avec la condition initiale en $\tau = \tau_{\Gamma}$

$$U(\tau_{\Gamma}) = u_{\Gamma}(\mathbf{X}_{\Gamma}, \tau_{\Gamma}) \tag{31}$$

On vérifie immédiatement que la solution s'écrit

$$u(\mathbf{x}, t) = U(t) = u_{\Gamma}(\mathbf{X}_{\Gamma}, \tau_{\Gamma}) + \int_{\tau_{\Gamma}}^t S(\tau) d\tau \tag{32}$$

ii) Par ailleurs, lorsque la caractéristique atteint le plan $t = 0$, on a $\mathbf{X}(0) \in \Omega$. Dans ce cas, la solution $u(\mathbf{x}, t)$ est déterminée univoquement en intégrant l'équation caractéristique (30) avec la condition initiale $U(0) = u^0(\mathbf{X}(0))$ provenant de la condition initiale originale (25), ce qui donne

$$u(\mathbf{x}, t) = U(t) = u^0(\mathbf{X}(0)) + \int_0^t S(\tau) d\tau \tag{33}$$

3.2 Méthodes Semi-Lagrangiennes et Caractéristiques

Nous décrivons à présent deux méthodes par éléments finis pour résoudre l'équation linéaire (23) en présence d'un champ d'advection $\mathbf{a}(\mathbf{x}, t)$ dépendant des variables spatiales et du temps. Si ce champ était connu à l'avance pour tout le domaine temporel considéré $[0, t]$, la solution du problème pourrait être déterminée comme suit. Pour une triangulation donnée du domaine Ω , on pourrait construire les courbes caractéristiques passant par tous les noeuds du maillage en les approximant par des lignes continues constituées par des segments de droites. Les valeurs nodales de l'inconnue u au temps t seraient alors obtenues en évaluant les intégrales figurant dans les équations (32) et (33) par quadrature numérique. Une description détaillée d'une méthode de ce type est donnée par Pironneau [PIR 82]. En pratique, le champ $\mathbf{a}(\mathbf{x}, t)$ dépend de la solution elle-même et il est, dès lors, nécessaire d'avancer la solution par incréments successifs d'un instant t^n à l'instant suivant $t^{n+1} = t^n + \Delta t$, où Δt désigne un pas de temps approprié.

3.2.1 Méthode semi-lagrangienne

La première méthode par éléments finis basée sur la notion de caractéristique que nous décrivons ici est la méthode *semi-lagrangienne* [ROB 81, ROB 82, PUS 84, STP 85].

Soit \mathbf{x} un point du maillage et $u^n(\mathbf{x})$ la solution approchée de l'Eq. (23) calculée à l'instant t^n . La solution $u^{n+1}(\mathbf{x})$ relative à l'instant successif t^{n+1} sera obtenue en approximant l'équation (26) définissant la courbe caractéristique pendant l'intervalle $[t^n, t^{n+1}]$ avec une précision du second ordre, en utilisant des différences centrées sur l'espace et le temps, ce qui donne

$$\Delta \mathbf{X}(\mathbf{x}) = \mathbf{a}\left(\mathbf{x} - \frac{1}{2}\Delta \mathbf{X}, t^{n+\frac{1}{2}}\right) \Delta t \quad (34)$$

Cette équation *non linéaire* pour le vecteur déplacement $\Delta \mathbf{X}$ peut être résolue par itérations à l'aide de la relation

$$\Delta \mathbf{X}_{(k+1)}(\mathbf{x}) = \mathbf{a}\left(\mathbf{x} - \frac{1}{2}\Delta \mathbf{X}_{(k)}, t^{n+\frac{1}{2}}\right) \Delta t \quad (35)$$

et en faisant usage d'une formule d'*interpolation* pour évaluer \mathbf{a} entre les noeuds du maillage. On approximera ensuite l'équation caractéristique (30) avec une précision du second ordre par rapport au temps en utilisant la relation

$$\frac{u^{n+1}(\mathbf{x}) - u^n(\mathbf{x} - \Delta \mathbf{X})}{\Delta t} = \frac{1}{2} [s^{n+1}(\mathbf{x}) + s^n(\mathbf{x} - \Delta \mathbf{X})] \quad (36)$$

Ici $s^n(\mathbf{x}) = s(\mathbf{x}, t^n)$, etc., et les quantités $u^n(\mathbf{x} - \Delta \mathbf{X})$ et $s^n(\mathbf{x} - \Delta \mathbf{X})$ sont évaluées à partir de $u^n(\mathbf{x})$ et $s^n(\mathbf{x})$ à l'aide d'*interpolations*. Dans la pratique, des *splines bi-cubiques* ont été utilisées pour les deux interpolations ci-dessus dans le cadre de calculs en deux et trois dimensions. La valeur de $\Delta \mathbf{X}(\mathbf{x})$ à l'intervalle de temps précédent a été utilisée comme première approximation

pour résoudre le système non linéaire (34) en chaque noeud [PUS 84, TSS 89, STC 90]. La convergence de la procédure itérative (35) est assurée pour autant que $\mathbf{a}(\mathbf{x}, t^{n+1/2})$ et ses dérivées partielles premières soient continues, que le pas de temps Δt soit suffisamment petit, et que la première approximation soit relativement proche de la solution recherchée.

Le schéma ci-dessus est inconditionnellement stable et sa phase ne dépend que de la partie fractionnaire du nombre de Courant. Les propriétés numériques et les performances de la méthode semi-lagrangienne ont été étudiées en détail et discutées de manière extensive par Pudykiewicz et Staniforth [PUS 84] dans le cas général de l'équation de convection - diffusion à une et deux dimensions. En raison de sa précision et de son efficacité, la méthode semi-lagrangienne est des plus intéressantes pour les prévisions météorologiques et la simulation de problèmes d'écoulement liés à la protection de l'environnement. On consultera les références [TSS 89, STC 90] pour la description d'applications pratiques utilisant des maillages cartésiens non uniformes.

3.2.2 Méthodes des caractéristiques

La notion de caractéristiques peut également être exploitée en liaison plus étroite avec la structure d'une formulation variationnelle faible, comme c'est le cas pour la méthode *caractéristique* par éléments finis proposée par Benqué *et al.* [BER 81, BIK 82]. Dans cette méthode, l'équation de convection $u_t + \mathbf{a} \cdot \nabla u = s$ est tout d'abord écrite sous forme faible en la multipliant par une fonction de pondération appropriée $v(\mathbf{x}, t)$ et en intégrant ensuite sur le domaine spatio-temporel $\Omega \times [t^n, t^{n+1}]$, ce qui donne

$$\int_{t^n}^{t^{n+1}} \int_{\Omega} [u_t + \mathbf{a} \cdot \nabla u] v \, d\Omega \, dt = \int_{t^n}^{t^{n+1}} \int_{\Omega} s v \, d\Omega \, dt \tag{37}$$

Après intégration par parties, à la fois sur l'espace et sur le temps, et dans l'hypothèse où $\nabla \cdot \mathbf{a} = 0$, l'Eq. (37) prend la forme

$$\begin{aligned} - \int_{t^n}^{t^{n+1}} \int_{\Omega} u [v_t + \mathbf{a} \cdot \nabla v] \, d\Omega \, dt + \int_{\Omega} [u^{n+1} v^{n+1} - u^n v^n] \, d\Omega \\ + \int_{t^n}^{t^{n+1}} \int_{\Gamma} \mathbf{n} \cdot \mathbf{a} u v \, d\Gamma \, dt = \int_{t^n}^{t^{n+1}} \int_{\Omega} s v \, d\Omega \, dt \end{aligned} \tag{38}$$

On choisit alors les fonctions de pondération $v(\mathbf{x}, t)$ de manière à satisfaire l'équation de convection homogène

$$v_t + \mathbf{a} \cdot \nabla v = 0 \tag{39}$$

complétée par la condition 'finale'

$$v(\mathbf{x}, t^{n+1}) = w(\mathbf{x}) \tag{40}$$

où $w(\mathbf{x})$ sont les fonctions de pondération purement spatiales associées à l'espace

choisi pour l'approximation par éléments finis sur le domaine Ω . La solution de l'Eq. (39) au sens de cette approximation est déterminée en introduisant les courbes caractéristiques définies par les Eqs. (26)-(27) adaptées à la présente situation, soit,

$$\frac{d\mathbf{X}}{d\tau} = \mathbf{a}(\mathbf{X}(\tau), \tau), \quad \mathbf{X}(t^{n+1}) = \mathbf{x} \quad (41)$$

Sous la forme caractéristique, l'équation (39) s'écrit $dV/d\tau = 0$, avec $t^n \leq \tau \leq t^{n+1}$, et la solution satisfaisant la condition finale (40) est simplement $V(\tau) = \text{const} = w(\mathbf{x})$. Il s'en suit que, pour chaque point nodal \mathbf{x} , on a $v(\mathbf{X}(t^n), t^n) = V(t^n) = V(t^{n+1}) = w(\mathbf{x})$. On admet ensuite que la transformation $\mathbf{x} \rightarrow \mathbf{X}(t^n)$ définit une transformation isoparamétrique T des éléments finis du niveau temporel t^{n+1} à t^n , soit $\mathbf{X}(t^n) = T\mathbf{x}$, et on introduit des fonctions de pondération transformées en t^n à l'aide de la définition $w_*(T\mathbf{x}) = w(\mathbf{x})$. L'équation variationnelle (38) est ainsi transformée en

$$\int_{\Omega} u^{n+1} w \, d\Omega = \int_{\Omega} u^n w_* \, d\Omega - \int_{t^n}^{t^{n+1}} \int_{\Gamma} \mathbf{n} \cdot \mathbf{a} u v \, d\Gamma \, dt + \int_{t^n}^{t^{n+1}} \int_{\Omega} s v \, d\Omega \, dt \quad (42)$$

L'équation (42) est un système linéaire dont la matrice est symétrique et définie positive (*matrice de masse consistante*). La difficulté technique de cette méthode réside dans le calcul du membre de droite de l'équation (42), en particulier celui du terme comprenant les fonctions transformées w_* .

Dans le cadre des méthodes faisant usage des caractéristiques, il convient également de mentionner la méthode Caractéristique-Galerkin introduite par Morton [MOR 85] et [MOR 87]. Celle-ci fait appel à une fonction de pondération modifiée

$$W^n(\mathbf{x}) = \frac{1}{|\mathbf{a}(u^n(\mathbf{x}))| \Delta t} \int_{\mathbf{x}}^{\mathbf{x} + \mathbf{a}(u^n(\mathbf{x})) \Delta t} w(\mathbf{z}) |d\mathbf{z}| \quad (43)$$

qui représente une moyenne de la fonction de pondération originale sur le segment correspondant à la projection de la caractéristique sur le domaine spatial. La Fig. 2 montre en représentation polaire l'erreur de phase rela-

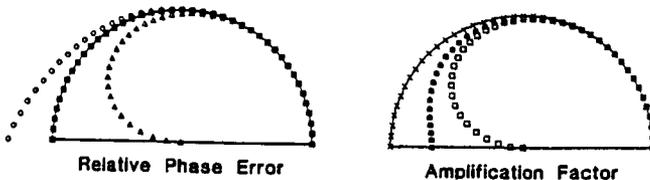


Figure 2: Propriétés numériques des méthodes Galerkin Caractéristique et Semi-Lagrangienne pour diverses valeurs du nombre de Courant: (Δ) $\nu = 0.25$, (\square) $\nu = 0.5$, (\circ) $\nu = 0.75$, (\times) $\nu = 1$.

tive $\phi_{\text{numer}}/\phi_{\text{exact}}$ et le module du facteur d'amplification des méthodes Caractéristique Galerkin et Semi-Lagrangienne pour diverses valeurs du nombre de Courant ν . La vitesse de phase numérique est exacte à toutes les longueurs d'onde pour les deux valeurs $\nu = \frac{1}{2}$ et $\nu = 1$. On remarque également que l'erreur d'amplitude atteint un maximum pour $\nu = \frac{1}{2}$ à toutes les longueurs d'onde et qu'une dissipation identique est obtenue lorsque ν est plus grand ou plus petit que $\frac{1}{2}$ par une même quantité.

3.3 Méthode Petrov-Galerkin

Une méthode de type Petrov-Galerkin conceptuellement très intéressante a été proposée par Morton et Parrott [MPA 80] pour résoudre l'équation linéaire de convection transitoire à une dimension

$$u_t + au_x = 0 \quad (44)$$

où a est une constante. Considérons en premier lieu l'approximation standard de Galerkin relative à l'Eq. (44). En laissant la variable temps continue et en représentant la dépendance spatiale par éléments finis linéaires, la formulation de Galerkin produit l'équation semi-discrète suivante en tout noeud intérieur j

$$\left(1 + \frac{1}{6}\delta^2\right) \frac{dU_j}{dt} = -\frac{a}{h}\Delta_0 U_j \quad (45)$$

Comme l'ont montré en premier Swartz et Wendroff [SWE 74], le terme de couplage $(1 + \frac{1}{6}\delta^2)$ dans le membre de gauche de l'Eq. (45), qui représente la *matrice de masse consistante*, assure une précision spatiale du quatrième ordre. Ce résultat remarquable est également maintenu pour certaines équations non linéaires, comme l'ont fait remarquer Cullen et Morton [CUM 80]). Cette précision spatiale élevée obtenue en discrétisant par éléments finis linéaires représente un gain significatif par rapport à la méthode des différences finies du second ordre. A ce gain s'ajoutent les propriétés de conservation inhérentes à la méthode de projection de Galerkin. Il convient toutefois de noter que le système d'équations différentielles (45) doit encore être intégré par rapport au temps pour produire la solution transitoire recherchée. C'est à ce stade que la précision spatiale du quatrième ordre pourra être rapidement érodée si l'algorithme de discrétisation temporelle n'est pas de précision comparable. En particulier, l'utilisation de schémas standards, tels que les méthodes d'Euler, leap-frog et Crank-Nicolson, aura pour conséquence de ne pouvoir maintenir la précision spatiale élevée obtenue par éléments finis que pour des valeurs faibles du pas de temps Δt . En fait, la précision et les bonnes propriétés de phase de la formulation de Galerkin disparaissent pour des valeurs relativement modestes du pas de temps Δt . En particulier, la propriété CFL unitaire (solution exacte aux noeuds pour un nombre de Courant unitaire) n'est pas obtenue en combinant éléments finis linéaires et schémas temporels explicites ou implicites d'ordre deux. Cette dégradation des performances des éléments

finis dès que le nombre de Courant croît est particulièrement décevante si l'on compare éléments finis et différences finies. On sait, en effet, que la plupart des schémas par différences finies de précision temporelle d'ordre deux possèdent la propriété CFL unitaire.

Morton et Parrott ont été les premiers à analyser en détail les raisons de ce mariage difficile entre éléments finis linéaires et schémas temporels du second ordre. Ils ont, en particulier, bien mis en évidence le fait qu'un aspect critique de la résolution de problèmes transitoires de convection par éléments finis est l'obtention d'un *couplage* adéquat entre la discrétisation spatiale et la discrétisation temporelle.

Pour réaliser ce couplage, Morton et Parrott ont repris l'approche *Petrov-Galerkin* qui s'était révélée si fructueuse pour stabiliser les problèmes stationnaires de convection-diffusion et ont suggéré de considérer, en association avec chaque schéma temporel, une classe particulière de fonctions de pondération modifiées de manière à fournir une bonne précision globale [MPA 80]. Prenant le cas du schéma d'Euler, et désignant par $U^n(x)$ la solution approchée recherchée, soit

$$U^n(\mathbf{x}) = \sum_j U_j^n w_j(\mathbf{x}), \quad (46)$$

la formulation Petrov-Galerkin relative à l'Eq. (44) s'écrit

$$\left\langle \frac{U^{n+1} - U^n}{\Delta t} + aU_x^n, \bar{w}_j \right\rangle = 0 \quad (47)$$

où \bar{w}_j est la fonction de pondération modifiée. Pour combiner la précision élevée de la méthode de Galerkin standard lorsque le nombre de Courant $\nu \rightarrow 0$ avec l'exigence d'obtenir la solution exacte pour $\nu \rightarrow 1$, Morton et Parrott ont proposé de définir \bar{w}_j par la combinaison linéaire

$$\bar{w}_j = (1 - \rho)w_j + \rho\chi_j \quad (48)$$

où χ_j est une fonction test spéciale ayant le même support que que la fonction de pondération de Galerkin w_j et calibrée de manière à obtenir la solution exacte de l'équation linéaire de convection à l'aide d'un maillage uniforme d'éléments linéaires lorsque $\nu = 1$. Le paramètre ρ dépend de ν et doit satisfaire les conditions $\rho(0) = 0$ and $\rho(1) = 1$. Le schéma résultant donne l'équation discrète

$$\left[1 + \frac{1}{6}(1 - \rho(\nu))\delta^2\right] (U_j^{n+1} - U_j^n) = -\nu\Delta_0 U_j^n + \frac{1}{2}\rho(\nu)\nu\delta^2 U_j^n \quad (49)$$

caractérisée par une modification de la matrice de masse qui passe de la forme Galerkin classique pour $\rho = 0$ à la matrice unitaire typique des différences finies pour $\rho = 1$. Le choix $\rho = \nu$ donne un schéma précis au second ordre, appelé EPGI, qui contient un opérateur spatial de type Lax-Wendroff et qui est stable pour $|\nu| \leq 1$.

D'autre part, le choix $\rho = \nu^2$ dans le membre de gauche de l'Eq. (49) combiné

avec $\rho = \nu$ dans le membre de droite conduit à un schéma appelé EPGII qui est précis au troisième ordre et stable pour $|\nu| \leq 1$.

Des formulations semblables ont été proposées par Morton et Parrott en relation avec les schémas de discrétisation temporelle leap-frog et Crank-Nicolson. On consultera la référence [MPA 80] pour les détails relatifs à la construction des fonctions de pondérations associées à ces schémas temporels.

L'approche Petrov-Galerkin de Morton et Parrott fournit des schémas pour la convection transitoire en une dimension qui ont des propriétés supérieures à celles des méthodes de différences finies du second ordre. Toutefois, cette approche ne possède pas les propriétés de conservation de la formulation de Galerkin. En outre, sa généralisation en vue du traitement de problèmes multidimensionnels et non linéaires n'est pas évidente.

Néanmoins, la méthode Petrov-Galerkin décrite ci-dessus a une importance historique en ce sens qu'elle a abordé pour la première fois le problème de l'intégration temporelle des équations hyperboliques du premier ordre en liaison avec une discrétisation spatiale par éléments finis, mettant ainsi bien en évidence l'interconnection entre discrétisations spatiale et temporelle de ces problèmes.

3.4 Méthode Taylor-Galerkin

Un développement ultérieur des principes de base de la méthode Petrov-Galerkin de Morton et Parrott a conduit à la méthode *Taylor-Galerkin* introduite par Donea [DON 84]. Dans cette méthode, qui représente une généralisation aux éléments finis des techniques du type Lax-Wendroff, la discrétisation temporelle précède la discrétisation spatiale et est prédisposée pour être combinée à la précision spatiale du quatrième ordre fournie par la méthode de Galerkin dans le cas d'éléments finis linéaires [DQS 87]. Nous commencerons par décrire un schéma explicite de Taylor-Galerkin à un pas et du troisième ordre par rapport au temps, pour passer ensuite à la présentation d'un schéma en deux pas également précis au troisième ordre et particulièrement intéressant pour les problèmes non linéaires multidimensionnels [SEL 87].

3.4.1 Schéma Euler-Taylor-Galerkin

Contrairement à la procédure classiquement suivie dans la méthode des éléments finis appliquée aux problèmes transitoires où la discrétisation spatiale précède la discrétisation temporelle, commençons par discrétiser par rapport au temps l'équation $u_t + \nabla \cdot f(u) = 0$, en laissant la variable spatiale x continue. Le développement en série de Taylor par rapport au temps conduisant au schéma explicite d'Euler fournit

$$u^{n+1} = u^n + \Delta t u_t^n + \frac{1}{2}(\Delta t)^2 u_{tt}^n + \frac{1}{6}(\Delta t)^3 u_{ttt}^n + O[(\Delta t)^4] \quad (50)$$

En utilisant l'équation de départ et ses dérivées successives par rapport au temps, on constate que (cf. également [MOR 87])

$$\begin{aligned}
u_{tt} &= -\partial_t [\nabla \cdot \mathbf{f}(u)] = -\nabla \cdot [\partial_t \mathbf{f}(u)] \\
&= -\nabla \cdot [\mathbf{a}(u) u_t] = -\nabla \cdot [\mathbf{a}(u) (-\nabla \cdot \mathbf{f}(u))] \\
&= \nabla \cdot [\mathbf{a}(u) \nabla \cdot \mathbf{f}(u)]
\end{aligned} \tag{51}$$

où $\mathbf{a}(u) = d\mathbf{f}(u)/du$, et, par conséquent:

$$\begin{aligned}
u_{ttt} &= (\Delta t)^{-1} [u_{tt}^{n+1} - u_{tt}^n] \\
&= (\Delta t)^{-1} \nabla \cdot [(\mathbf{a} \nabla \cdot \mathbf{f})^{n+1} - (\mathbf{a} \nabla \cdot \mathbf{f})^n]
\end{aligned} \tag{52}$$

Le terme du troisième ordre dans l'expression (52) a été écrit sous la forme *mixte temporelle-spatiale* pour éviter la présence de dérivées spatiales du troisième ordre qui exigeraient la continuité C^1 de l'espace d'interpolation. En introduisant les expressions (51) et (52) dans la série de Taylor (50) et en appliquant ensuite la formulation standard de Galerkin, on obtient après intégration par parties et quelques réarrangements le schéma

$$\begin{aligned}
\left\langle \frac{u^{n+1} - u^n}{\Delta t}, w \right\rangle + \frac{1}{6} \Delta t \langle [(\mathbf{a} \nabla \cdot \mathbf{f})^{n+1} - (\mathbf{a} \nabla \cdot \mathbf{f})^n], \nabla w \rangle \\
= -\langle \nabla \cdot \mathbf{f}^n, w \rangle - \frac{1}{2} \Delta t \langle \mathbf{a}^n \nabla \cdot \mathbf{f}^n, \nabla w \rangle \\
= -\langle \nabla \cdot \mathbf{f}^n, w + \frac{1}{2} \Delta t \mathbf{a}^n \cdot \nabla w \rangle
\end{aligned} \tag{53}$$

où $\mathbf{f}^n = \mathbf{f}(u^n)$, $\mathbf{a}^n = \mathbf{a}(u^n)$, $(\mathbf{a} \nabla \cdot \mathbf{f})^n = \mathbf{a}(u^n) \nabla \cdot \mathbf{f}(u^n)$, etc.. Le second terme dans le membre de gauche de l'Eq. (53) est le terme responsable de la précision du troisième ordre. Ce terme introduit une modification de la matrice de masse qui dépend non linéairement de u^n et u^{n+1} dans le cas d'une équation de conservation. Si on considère le cas simple de l'équation de convection linéaire à coefficient constant $u_t + \mathbf{a} \cdot \nabla u = 0$, le schéma Euler/Taylor-Galerkin (53) se réduit à

$$\begin{aligned}
\left\langle \frac{u^{n+1} - u^n}{\Delta t}, w \right\rangle + \frac{1}{6} \Delta t \langle \mathbf{a} \cdot \nabla (u^{n+1} - u^n), \mathbf{a} \cdot \nabla w \rangle \\
= -\langle \mathbf{a} \cdot \nabla u^n, w + \frac{1}{2} \Delta t \mathbf{a} \cdot \nabla w \rangle
\end{aligned} \tag{54}$$

Une intégration par parties a été utilisée pour obtenir les relations (53) et (54), si bien que des termes de surface doivent être inclus dans les relations en question.

Pour le cas de la convection à une dimension traitée à l'aide d'éléments linéaires, le schéma suivant est obtenu:

$$\left[1 + \frac{1}{6}(1 - \nu^2)\delta^2\right] (U_j^{n+1} - U_j^n) = -\nu \Delta_0 U_j^n + \frac{1}{2} \nu^2 \delta^2 U_j^n \tag{55}$$

dont le facteur d'amplification s'écrit

$$G_{TG}^E(\xi, \nu) = 1 + \frac{-i\nu \sin \xi - 2\nu^2 \sin^2 \frac{1}{2} \xi}{1 - \frac{2}{3}(1 - \nu^2) \sin^2 \frac{1}{2} \xi} \tag{56}$$

et donne la condition de stabilité $|\nu| \leq 1$. Ce schéma est également connu

sous le nom de méthode Lax–Wendroff/Taylor–Galerkin [DQS 87]. Ses propriétés numériques sont présentées à la Fig. 3, sous la forme de diagrammes polaires d’erreurs de phase et d’amplitude. La comparaison avec la méthode Galerkin Caractéristique (ou Semi-Lagrangienne) montre que la phase est exacte dans les deux méthodes pour $\nu = \frac{1}{2}$ et $\nu = 1$, et que l’erreur de phase est pratiquement identique, le schéma Taylor–Galerkin étant légèrement inférieur dans l’intervalle $\frac{1}{2} < \nu < 1$. Ainsi, en plus de leur stabilité inconditionnelle, les méthodes Semi-Lagrangienne et Galerkin Caractéristique fournissent une discrétisation du troisième ordre.

3.4.2 Schéma du troisième ordre à deux pas

Dans ce paragraphe nous décrivons un schéma Taylor–Galerkin d’ordre trois utilisable en pratique dans la résolution de problèmes hyperboliques non linéaires et multidimensionnels. L’idée de base de ce schéma consiste à obtenir une précision du troisième ordre par rapport au temps à l’aide d’une procédure à deux pas due à Selmin [SEL 87] et qui s’écrit

$$\tilde{u}^n = u^n + \frac{1}{3}\Delta t u_t^n + \alpha(\Delta t)^2 u_{tt}^n \tag{57}$$

$$u^{n+1} = u^n + \Delta t u_t^n + \frac{1}{2}(\Delta t)^2 \tilde{u}_{tt}^n \tag{58}$$

où la valeur du paramètre α est provisoirement laissée libre. En fait, tandis que les autres coefficients dans (57–58) prennent la valeur nécessaire pour obtenir une précision du troisième ordre par combinaison des deux pas, le paramètre α n’affecte que le coefficient du terme du quatrième ordre dans le développement en série global.

En conséquence, la valeur de ce paramètre affectera le facteur d’amplification du schéma résultant mais pas sa phase. On pourra donc tirer profit du degré de liberté disponible en imposant que la réponse en phase du schéma à deux pas coïncide avec celle du schéma Euler/Taylor–Galerkin (54) en application à l’équation linéaire de convection [SEL 87]. La version complètement discrétisée des Eqs. (57–58) s’écrit dans le cas de l’équation linéaire de convection à une dimension:

$$\left[1 + \frac{1}{6}\delta^2\right] (\tilde{U}^n - U^n) = -\frac{1}{3}\nu\Delta_0 U^n + \alpha\nu^2\delta^2 U^n \tag{59}$$

$$\left[1 + \frac{1}{6}\delta^2\right] (U^{n+1} - U^n) = -\nu\Delta_0 U^n + \frac{1}{2}\nu^2\delta^2 \tilde{U}^n \tag{60}$$

si bien que le facteur d’amplification du schéma paramétrisé à deux pas est donné par la relation:

$$G_{TTG}(\xi, \nu; \alpha) = 1 + \frac{-i\nu\sin\xi - 2\nu^2\sin^2\frac{1}{2}\xi \tilde{G}(\xi, \nu; \alpha)}{1 - \frac{2}{3}\sin^2\frac{1}{2}\xi} \tag{61}$$

où

$$\tilde{G}(\xi, \nu; \alpha) = 1 + \frac{-\frac{1}{3}i\nu\sin\xi - 4\alpha\nu^2\sin^2\frac{1}{2}\xi}{1 - \frac{2}{3}\sin^2\frac{1}{2}\xi} \tag{62}$$

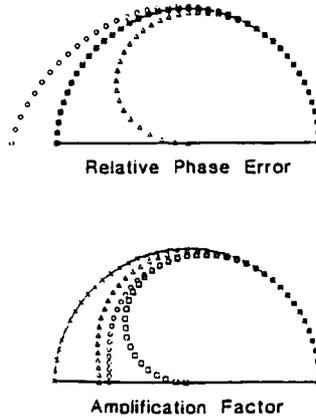


Figure 3: Propriétés numériques de la méthode Euler/Taylor-Galerkin pour diverses valeurs du nombre de Courant: (Δ) $\nu = 0.25$, (\square) $\nu = 0.5$, (\circ) $\nu = 0.75$, (\times) $\nu = 1$.

On vérifie aisément que

$$G_{\text{TTG}}(\xi, \nu; \frac{1}{9}) = G_{\text{TG}}^{\text{E}}(\xi, \nu) R(\xi, \nu) \quad (63)$$

où $R(\xi, \nu)$ est une fonction réelle donnée par

$$R(\xi, \nu) = \frac{[1 - \frac{2}{3}(1 - \nu^2)\sin^2 \frac{1}{2}\xi] [1 - \frac{2}{3}(1 + \nu^2)\sin^2 \frac{1}{2}\xi]}{[1 - \frac{2}{3}\sin^2 \frac{1}{2}\xi]^2} \quad (64)$$

Ainsi, pour $\alpha = \frac{1}{9}$ la procédure à deux pas (59–60) reproduit exactement les excellentes propriétés de phase du schéma Euler/Taylor-Galerkin à un pas. La condition de stabilité numérique du schéma à deux pas est donnée par $|\nu| \leq \sqrt{3}/2$. La figure 4 montre les diagrammes d'erreur de phase et d'amplitude du schéma du troisième ordre à deux pas. Le nouveau schéma est plus dissipatif que la version à un pas, mais la différence est négligeable pour toutes les longueurs d'onde, sauf lorsque le nombre de Courant ν est très proche de la limite de stabilité.

Un point important qu'il convient de souligner est que, même à deux et trois dimensions, les propriétés de phase du schéma Taylor-Galerkin du troisième ordre à un pas peuvent être reproduites exactement par la procédure à deux pas. En outre, la condition de stabilité du schéma à deux pas pour les cas bi- et tri-dimensionnels est pratiquement identique à celle du cas monodimensionnel. En fait, en introduisant le vecteur nombre de Courant $\nu = (\nu_x, \nu_y)$, on démontre que le domaine de stabilité numérique de la version à deux dimensions du schéma (59–60) avec $\alpha = \frac{1}{9}$ ne dépend que très marginalement de l'orientation du vecteur ν [SEL 87]: le domaine de stabilité coïncide pratiquement avec le

cercle de rayon $\sqrt{3}/2$ ($= 0.866$), à l'exception d'un léger aplatissement dans la direction de la diagonale $\nu_x = \nu_y$ pour laquelle la limite de stabilité s'écrit

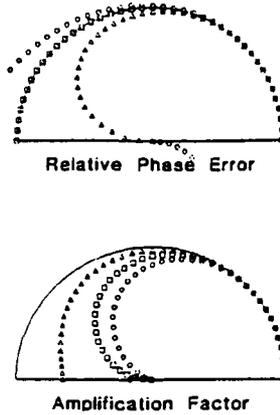


Figure 4: Propriétés numériques de la méthode Taylor-Galerkin du troisième ordre à deux pas pour diverses valeurs du nombre de Courant: (Δ) $\nu = 0.25$, (\square) $\nu = 0.5$, (\circ) $\nu = 0.75$.

$|\nu| < 0.854$ (voir Fig. 5). Ce résultat est à comparer à la forme du domaine de stabilité du schéma Euler/Taylor-Galerkin à un pas qui, comme la méthode des différences finies de Lax-Wendroff, dépend fortement de la direction de ν , et implique, en conséquence, une réduction drastique de la limite de stabilité qui se réduit à $|\nu| < 1/(2\sqrt{2})$ en 2D et $|\nu| < 1/(3\sqrt{3})$ en 3D. Ainsi, tout en rendant la méthode Taylor-Galerkin du troisième ordre praticable pour la résolution de problèmes non linéaires, le schéma à deux pas présente l'avantage ultérieur de posséder un domaine de stabilité numérique pratiquement *isotrope* pour les problèmes multidimensionnels.

Pour conclure, nous indiquerons la forme prise par le schéma à deux pas dans le cas de l'équation de conservation non linéaire $u_t + \nabla \cdot \mathbf{f}(u) = 0$. Dans ce cas, le schéma s'écrit:

$$\langle \tilde{u}^n - u^n, w \rangle / \Delta t = \frac{1}{3} \langle \mathbf{f}(u^n) - \frac{1}{3} \Delta t \mathbf{a}(u^n) \nabla \cdot \mathbf{f}(u^n), \nabla w \rangle \quad (65)$$

$$\langle u^{n+1} - u^n, w \rangle / \Delta t = \langle \mathbf{f}(u^n) - \frac{1}{2} \Delta t \mathbf{a}(\tilde{u}^n) \nabla \cdot \mathbf{f}(\tilde{u}^n), \nabla w \rangle \quad (66)$$

On notera que dans le schéma à deux pas, l'effet de la matrice de masse modifiée, qui dépend de la solution elle-même, est obtenu par une double application de la matrice de masse consistante usuelle. Bien entendu, des termes de frontière appropriés doivent être inclus dans le membre de droite des deux équations, comme déjà indiqué pour le schéma à un pas.

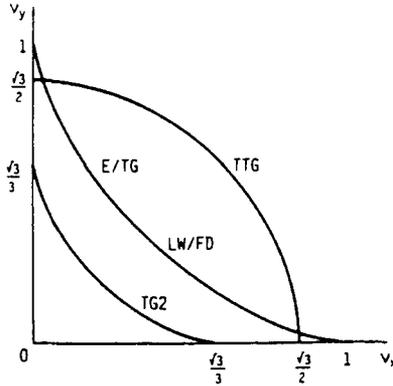


Figure 5: Domaines de stabilité numérique des schémas Taylor–Galerkin et Lax–Wendroff pour l'équation de convection à deux dimensions.

3.5 Méthodes des Moindres Carrés

3.5.1 Méthode basée sur les caractéristiques

La première méthode des moindres carrés que nous décrirons repose sur l'utilisation des caractéristiques pour résoudre l'équation *linéaire* de convection

$$u_t + \mathbf{a}(\mathbf{x}, t) \cdot \nabla u = 0 \quad (67)$$

à l'aide d'un maillage d'éléments finis non-structuré. Représentons la solution approchée $U^n(\mathbf{x})$ au temps t^n sous la forme

$$U^n(\mathbf{x}) = \sum_j U_j^n w_j(\mathbf{x}) \quad (68)$$

La solution au temps t^{n+1} est donnée par

$$U_*^{n+1}(\mathbf{x}) = U^n(\mathbf{X}) \quad (69)$$

où le déplacement \mathbf{X} obtenu par intégration de t^{n+1} à t^n de la caractéristique est approximé par

$$\mathbf{X} = \mathbf{x} - \mathbf{a}(\mathbf{x}, t^{n+1})\Delta t \quad (70)$$

L'indice \star , qui caractérise la solution à t^{n+1} , a été introduit pour indiquer explicitement que la solution obtenue en (69) par translation pure des données initiales n'appartient pas à l'espace d'interpolation sous-tendu par les fonctions de forme $\{w_j(\mathbf{x})\}$. En vue d'exprimer la solution $U^{n+1}(\mathbf{x})$ dans cet espace comme en (67), on peut utiliser la méthode des moindres carrés [CLI 90] pour minimiser la différence entre $U^{n+1}(\mathbf{x})$ et $U_*^{n+1}(\mathbf{x})$, à savoir,

$$\int_{\Omega} [U^{n+1}(\mathbf{x}) - U_*^{n+1}(\mathbf{x})]^2 d\Omega = \min \quad (71)$$

ce qui conduit par variation de l'inconnue $U^{n+1}(\mathbf{x})$ à

$$\int_{\Omega} [U^{n+1}(\mathbf{x}) - U_{*}^{n+1}(\mathbf{x})] \delta U^{n+1}(\mathbf{x}) d\Omega = 0 \tag{72}$$

Compte-tenu du développement (68) pour $U^{n+1}(\mathbf{x})$ et de l'Eq. (69), l'intégrale (72) prend la forme

$$\sum_{j'} U_{j'}^{n+1} \int w_{j'}(\mathbf{x}) w_j(\mathbf{x}) d\Omega = \sum_{j'} U_{j'}^n \int w_{j'}(\mathbf{X}) w_j(\mathbf{x}) d\Omega \tag{73}$$

où $\mathbf{X} = \mathbf{x} - \mathbf{a}(\mathbf{x}, t^{n+1})\Delta t$. Dans le cas de coefficients constants, on peut montrer que cette expression est équivalente à celle que délivre la méthode Galerkin-Characteristique, pour autant que les intégrales soient évaluées exactement. Etant donné que dans les applications pratiques, ces intégrales ne peuvent être évaluées que numériquement, il en résulte que des règles d'intégration d'ordre élevé sont requises pour le calcul des intégrales figurant au membre de droite de l'Eq. (73). Ceci représente un aspect critique de la méthode, en particulier lorsque des fonctions de forme d'ordre polynomial faible sont utilisées [CLI 90]. Il reste néanmoins le résultat intéressant que la méthode Galerkin-Characteristique peut être interprétée, tout au moins en application à l'équation linéaire de convection, comme étant une méthode de moindres carrés.

3.5.2 Méthode de Carey et Jiang

Une méthode de moindres carrés très intéressante pour les problèmes transitoires de convection a été proposée par Carey et Jiang [CAJ 88]. Dans ce paragraphe, nous présenterons une version de cette méthode adaptée à la résolution de problèmes linéaires.

Considérons l'équation linéaire de convection

$$u_t + \mathbf{a}(\mathbf{x}) \cdot \nabla u = 0 \tag{74}$$

et la fonctionnelle quadratique (instantanée) associée

$$I(t) = \int_{\Omega} [u_t + \mathbf{a} \cdot \nabla u]^2 d\Omega \tag{75}$$

Si l'équation de départ est discrétisée par rapport au temps à l'aide du schéma moyen θ , la fonctionnelle (75) est donnée à l'instant $t^{n+\theta}$ par

$$I^{n+\theta} = \int_{\Omega} \left[\frac{u^{n+1} - u^n}{\Delta t} + \theta \mathbf{a} \cdot \nabla u^{n+1} + (1 - \theta) \mathbf{a} \cdot \nabla u^n \right]^2 d\Omega \tag{76}$$

Clairement, $\theta = 0$ correspond au schéma explicite d'Euler, $\theta = \frac{1}{2}$ au schéma de Crank-Nicolson et $\theta = 1$ au schéma implicite d'Euler. Bien entendu, seul le choix $\theta = \frac{1}{2}$ donne une précision du second ordre par rapport au temps. En minimisant $I^{n+\theta}$ par rapport à la variation de l'inconnue u^{n+1} , on obtient

l'équation aux moindres carrés (toujours continue par rapport aux variables d'espace)

$$\int_{\Omega} \left[\frac{u^{n+1} - u^n}{\Delta t} + \theta \mathbf{a} \cdot \nabla u^{n+1} + (1 - \theta) \mathbf{a} \cdot \nabla u^n \right] \times [\delta u^{n+1} + \theta \Delta t \mathbf{a} \cdot \nabla (\delta u^{n+1})] d\Omega = 0 \quad (77)$$

En exprimant la variation δu^{n+1} en termes de la fonction de pondération w , l'Eq. (77) fournit

$$\left\langle \frac{u^{n+1} - u^n}{\Delta t} + \theta \mathbf{a} \cdot \nabla u^{n+1}, w + \theta \Delta t \mathbf{a} \cdot \nabla w \right\rangle = -(1 - \theta) \langle \mathbf{a} \cdot \nabla u^n, w + \theta \Delta t \mathbf{a} \cdot \nabla w \rangle \quad (78)$$

L'équation (78) peut également s'écrire sous la forme incrémentielle compacte suivante

$$\langle [1 + \theta \Delta t \mathbf{a} \cdot \nabla](u^{n+1} - u^n), [1 + \theta \Delta t \mathbf{a} \cdot \nabla]w \rangle = -\Delta t \langle \mathbf{a} \cdot \nabla u^n, [1 + \theta \Delta t \mathbf{a} \cdot \nabla]w \rangle \quad (79)$$

qui montre bien le caractère symétrique de l'opérateur implicite engendré par la méthode des moindres carrés. En une dimension et pour des éléments finis linéaires, l'équation (79) fournit le schéma discret suivant

$$[1 + (\frac{1}{6} - \theta^2 \nu^2) \delta^2] (U_j^{n+1} - U_j^n) = -\nu \Delta_0 U_j^n + \theta \nu^2 \delta^2 U_j^n \quad (80)$$

et on démontre que ce schéma est inconditionnellement stable pour $\frac{1}{2} \leq \theta \leq 1$ [CAJ 88]. La Fig. 6 illustre les propriétés du schéma (80) pour $\theta = \frac{1}{2}$ et $\theta = 1$.

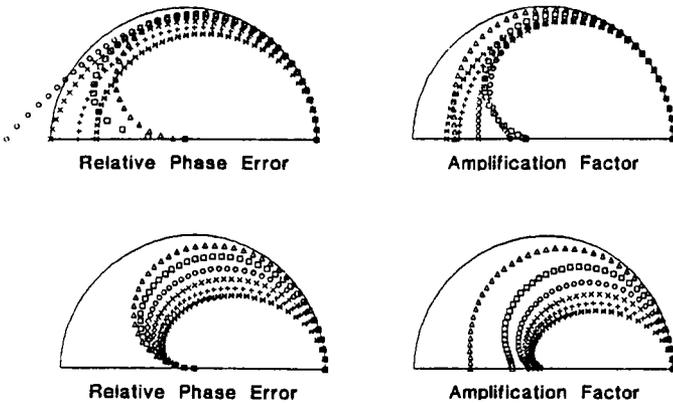


Figure 6: Propriétés de la méthode moindres carrés /Crank-Nicolson pour $\theta = 1/2$ (haut) et $\theta = 1$ (bas), pour diverses valeurs du nombre de Courant: $(\Delta) \nu = 0.25$, $(\square) \nu = 0.5$, $(\circ) \nu = 0.75$, $(\times) \nu = 1$, $(+)$ $\nu = 1.25$, $(*) \nu = 1.5$.

Le schéma moindres carrés/Crank–Nicolson (MC/CN) possède une réponse en phase précise pour autant qu'il soit utilisé dans l'intervalle $\nu \leq 1$. Par ailleurs, le schéma complètement implicite (MC/CI) possède une réponse en phase très médiocre, sauf pour des valeurs faibles du nombre de Courant et son utilisation n'est envisageable que pour le calcul de solutions stationnaires.

3.5.3 Autres méthodes de moindres carrés

Une méthode très précise pour le traitement de problèmes transitoires de convection a été introduite récemment par Park et Liggett [PAL 90]. Cette méthode combine les concepts Taylor–Galerkin et moindres carrés et est du quatrième ordre par rapport au temps. Pour la discrétisation spatiale, Park et Liggett utilisent des éléments finis à interpolation cubique (polynômes de Hermite).

Nous terminerons cette section consacré aux méthodes de moindres carrés pour le traitement de problèmes transitoires de convection en mentionnant une méthode d'éléments finis de type espace–temps proposée par Nguyen et Reynen [NGR 84]. Cette méthode est fondamentalement différente des méthodes de moindres carrés décrites précédemment en ce sens que la fonctionnelle quadratique associée comporte une intégrale étendue à l'espace et au temps. Le schéma résultant de l'utilisation d'éléments linéaires, est inconditionnellement stable et précis au second ordre. Les propriétés de la méthode moindres carrés espace–temps sont illustrées à la Fig. 7. En comparant les figures 6 et 7, on constate que l'approche espace–temps donne un schéma moins précis et plus dissipatif que le schéma moindres carrés obtenu en utilisant dès le départ la discrétisation de Crank–Nicolson. Pour cette raison, la méthode des moindres carrés espace–temps n'est pas recommandée pour la résolution de problèmes de convection pure. Par contre, elle donne de bons résultats pour les problèmes de convection–diffusion.

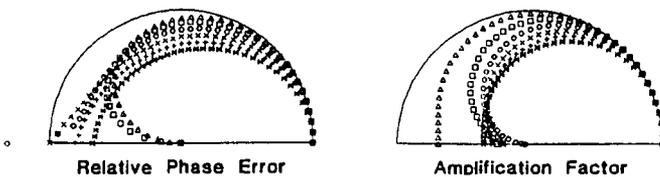


Figure 7: Propriétés de la méthode moindres carrés/ espace–temps pour diverses valeurs du nombre de Courant: (Δ) $\nu = 0.25$, (\square) $\nu = 0.5$, (\circ) $\nu = 0.75$, (\times) $\nu = 1$, ($+$) $\nu = 1.25$, ($*$) $\nu = 1.5$.

Références

- [BER 81] J. P. Benqué and J. Ronat, Quelques difficultés des modèles numériques en hydraulique, *Fifth Intern. Symp. on Computing Methods in Applied Sciences and Engineering*, INRIA, December 1981, Versailles, France.
- [BIK 82] J. P. Benqué, B. Ibler, A. Keramsi and G. Labadie, A new finite element method for the Navier–Stokes equations coupled with a temperature equation, in *Proc. Fourth Int. Symp. on Finite Element in Flow Problem*, Ed. T. Kawai, North-Holland, 1982, 295–301.
- [BRH 80] A.N. Brooks and T.J.R. Hughes, Streamline upwind/Petrov-Galerkin methods for advection dominated flows, *Proc. Third Internat. Conf. on Finite Element Methods in Fluid Flow*, Banff, Canada, 1980.
- [BRH 82] A.N. Brooks and T.J.R. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, *Comp. Meths. Appl. Mech. Engng.*, 32 (1982), 199–259.
- [CAJ 88] G. F. Carey and B. N. Jiang, Least-squares finite elements for first-order hyperbolic systems, *Int. J. Numer. Meths. Eng.* 26 (1988) 81–93.
- [CGM 76] I. Christie, D.F. Griffiths, A.R. Mitchell and O.C. Zienkiewicz, Finite element methods for second order differential equations with significant first derivatives, *Internat. J. Numer. Methods Engng.*, 10 (1976), 1389–1396.
- [CHR 85] I. Christie, Upwind compact finite difference schemes, *J. Comput. Physics*, 59 (1985), 353–368.
- [CLI 90] C. W. Li, Least square-characteristics and finite elements for advection–dispersion simulation, *Int. J. Numer. Meths. Eng.* 29 (1990), 1343–1364.
- [CUM 80] M. J. P. Cullen and K. W. Morton, Analysis of evolutionary error in finite element and other methods, *J. Comput. Phys.* 34 (1980) 245–267.
- [DON 84] J. Donea, A Taylor–Galerkin method for convective transport problems, *Int. J. Numer. Meths. Eng.* 20 (1984) 101–120.
- [DON 91] J. Donea, Generalized Galerkin methods for convection dominated transport phenomena, *Appl. Mech. Rev.*, 44 (1991), 205–214.

- [DOQ 92] J. Donea and L. Quartapelle, An introduction to finite element methods for transient advection problems, *Comp. Meths. Appl. Mech. Engrg.*, 95 (1992), 169-203.
- [DQS 87] J. Donea, L. Quartapelle and V. Selmin, An analysis of time discretization in the finite element solution of hyperbolic problems, *J. Comput. Phys.* 70 (1987) 463-499.
- [GRM 79] D.F. Griffiths and A.R. Mitchell, On generating upwind finite element methods, *Finite Element Methods for Convection Dominated Flows*, ed. T.J.R. Hughes, AMD Vol. 34, ASME, New York, 1979.
- [HBR 79] T.J.R. Hughes and A. Brooks, A multidimensional upwind scheme with no crosswind diffusion, *Finite Element Methods for Convection Dominated Flows*, ed. T.J.R. Hughes, AMD Vol 34, ASME, New York, 1979.
- [HFH 89] T.J.R. Hughes, L.P. Franca and G.M. Hulbert, A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/Least-squares method for advective-diffusive equations, *Comput. Meths. Appl. Mech. Eng.* 73 (1989) 173-189.
- [HHZ 77] J.C. Heinrich, P.S. Huyakorn, O.C. Zienkiewicz and A.R. Mitchell, An 'upwind' finite element scheme for two-dimensional convective transport equation, *Int. J. Numer. Meths. Engrg.*, 11 (1977), 134-143.
- [HMA 86] T. J. R. Hughes and M. Mallet, A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advection-diffusion systems, *Comput. Meths. Appl. Mech. Eng.* 58 (1986) 305-328.
- [HUG 89] T.J.R. Hughes, New directions in computational mechanics, *Nuclear Eng. and Design* 114 (1989) 197-210.
- [HZI 79] J.C. Heinrich and O.C. Zienkiewicz, The finite element method and 'upwinding' techniques in the numerical solution of convection dominated flow problems, *Finite Element Methods for Convection Dominated Flows*, ed. T.J.R. Hughes, AMD Vol. 34, ASME, New York, 1979.
- [MOR 85] K. W. Morton, Generalized Galerkin methods for hyperbolic problems, *Comput. Meths. Appl. Mech. Eng.* 52 (1985) 847-871.
- [MOR 87] K. W. Morton, Finite element methods for hyperbolic equations, in *Finite Elements in Physics*, R. Gruber, Ed., North-Holland, Amsterdam, 1987.

- [MPA 80] K. W. Morton and A. K. Parrott, Generalized Galerkin methods for first-order hyperbolic equations, *J. Comput. Phys.* 36 (1980) 249-270.
- [NGR 84] H. Nguyen and J. Reynen, A space-time least-squares finite element scheme for advection-diffusion equations, *Comput. Meths. Appl. Mech. Eng.* 42 (1984) 331-342.
- [PAL 90] N.-S. Park and J. A. Liggett, Taylor-least squares finite element for two-dimensional advection dominated advection-diffusion problems, *Int. J. Numer. Meths. Fluids* 11 (1990) 21-38.
- [PIR 82] O. Pironneau, On the transport-diffusion algorithm and its application to the Navier-Stokes equations, *Numer. Math.* 38 (1982) 309-332.
- [PIR 88] O. Pironneau, *Méthodes des Éléments Finis pour les Fluides*, Masson, Paris, 1988.
- [PUS 84] J. Pudykiewicz and A. Staniforth, Some properties and comparative performance of semi-Lagrangian method of Robert in the solution of the advection-diffusion equation, *Atmos. Ocean* 22 (1984) 283-308.
- [ROB 81] A. Robert, A stable numerical integration scheme for the primitive meteorological equations, *Atmos. Ocean* 19 (1981) 35-46.
- [ROB 82] A. Robert, A semi-Lagrangian and semi-implicit numerical integration scheme for the primitive meteorological equations, *J. Meteorol. Soc. (Jpn.)* 60 (1982) 319-325.
- [SEL 87] V. Selmin, Third-order finite element schemes for the solution of hyperbolic problems, *INRIA Report 707*, 1987.
- [SHY 85] W. Shyy, A study of finite difference approximations to steady-state, convection-dominated flow problems, *J. Comput. Physics*, 57 (1985), 415-438.
- [STC 90] A. Staniforth and J. Côté, Semi-Lagrangian integration schemes and their application to environmental flows, *Lecture Notes in Physics*, Springer-Verlag, New York, 1990.
- [STF 73] G. Strang and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall, New Jersey, 1973.
- [STP 85] A. Staniforth and J. Pudykiewicz, Reply to comments on and addenda to "Some properties and comparative performance of semi-Lagrangian method method of Robert in the solution of the advection-diffusion equation", *Atmos. Ocean* 23 (1985) 195-200.

- [SWE 74] B. Swartz and B. Wendroff, The relation between the Galerkin and collocation methods using smooth splines, *SIAM J. Numer. Anal.* 11 (1974) 1059–1068.
- [TSS 89] M. Tanguay, A. Simard and A. Staniforth, A three-dimensional semi-Lagrangian integration scheme for the Canadian regional finite-element forecast model, *Mon. Weather Rev.* 117 (1989) 1861–1871.