CHARLES-HENRI BRUNEAU

Résolution des équations d'Euler compressibles par une méthode variationnelle en éléments finis : schémas aux différences décentrées pour la résolution des équations de Navier-Stokes incompressibles

Thèses d'Orsay, 1989

ORSAY
n° d'ordre :

UNIVERSITÉ de PARIS-SUD

Centre d'ORSAY

THÈSE

présentée
pour obtenir

le grade de Docteur ès Sciences Mathématiques

par

Charles-Henri BRUNEAU

Sujet de la thèse :

Résolution des équations d'Euler compressibles par une méthode variationnelle en éléments finis.

Schémas aux différences finies décentrées pour la résolution des équations de Navier-Stokes incompressibles.

soutenue le 21 décembre 1989 devant le Jury composé de :

MM. Roger TEMAM, Président
Jean-Paul BOUJOT
Alain DERVIEUX
Etienne FOUVRY
Marcel LESIEUR
Roger PEYRET
Jean-Claude SAUT
Pieter ZANDBERGEN

à Henri Bruneau

Je tiens à remercier tout d'abord :

Monsieur R. Temam qui m'a accueilli au Laboratoire d'Analyse Numérique, a encouragé ce travail et m'a poussé à multiplier les contacts avec l'extérieur.

Monsieur A. Dervieux, rapporteur, de l'intérêt qu'il a manifesté pour ce travail.

Messieurs J.P. Boujot, M. Lesieur, R. Peyret, J.C. Saut et P. Zandbergen qui me font l'honneur et le plaisir de participer à ce Jury.

Monsieur E. Fouvry qui, en me proposant un second sujet, m'a initié et intéressé à la théorie des nombres.

Je remercie tout particulièrement mes collègues de travail :

J.J. Chattot, qui est à l'origine de la première partie et qui, par de nombreuses et fructueuses discussions, m'a fait partager ses connaissances en mécanique des fluides.

J. Guiu-Roux avec laquelle j'ai eu la joie de collaborer il y a quelques années.

C. Jouron, avec qui j'ai entrepris la deuxième partie de cette thèse.

J. Laminie, compagnon de longue date, qui a su partager les bons comme les mauvais moments.

Mes remerciements vont aussi à :

B. Héron et J.M. Ghidaglia, avec qui j'ai eu d'intéressantes discussions liées à mon travail.

Tous mes collègues du Laboratoire d'Analyse Numérique qui m'ont soutenu toutes ces années.

H. Almoatassine, F. Pascal et M. Tajchman pour le très bon et très utile travail qu'ils ont accompli pendant leur stage de D.E.A.

Mes proches et amis qui m'ont manifesté un constant encouragement.

Madame Le Meur, enfin, qui a réalisé avec grand soin la frappe de cette thèse.

Les moyens informatiques nécessaires à la réalisation de cette thèse ont été fournis par :

Le C.N.R.S. auquel le Laboratoire est rattaché.

L'ONERA et la DRET qui ont montré leur intérêt aux études sur les équations d'Euler par des contrats successifs.

Le CCVR qui a alloué de nombreuses heures sur CRAY-1 et CRAY-2 pour ces travaux.

L'université du Minnesota et le Supercomputer Institute qui m'ont accueilli au printemps 1986 et m'ont permis de travailler en avant-première sur CRAY-2.

L'ESA, le CNES et DASSAULT qui ont supporté une partie des études hypersoniques dans le cadre du projet HERMES.

L'ICASE au Centre NASA de Langley qui m'a fourni des moyens de travail exceptionnels au printemps 1989.

Le CIRCE qui a alloué sur Siemens VP200 les heures nécessaires au calcul des solutions des équations de Navier-Stokes incompressibles.

Charles-Henri Bruneau.

Resolution of compressible Euler equations through a variational method and a finite elements approximation.
Finite differences schemes for solving the incompressible Navier-Stokes equations.

Abstract.

This work is divided into two parts, which both concern the solution of problems in fluid dynamics : compressible Euler flows in two and three dimensions and incompressible Navier-Stokes flows in two dimensions. The methods of solution are very different from one application to another. In the first case the solution of various problems is obtained by a fixed point algorithm, Newton linearization, a least-squares minimization and a finite element approximation of the conservative variables ; moreover an artificial viscosity term or an entropy corrector with a mesh adaptation are required to get transonic or hypersonic solutions. In the second case a finite differences scheme with an antidiffusion term is used to approximate the Navier-Stokes operator and the solution is achieved by means of the multigrid method coupled to a cell by cell relaxation smoother.

In addition to efficient methods, a structured programming including optimization and vectorization yields good performances in terms of CPU time. The work of this thesis gives an outlook on a broad range of knowledge and techniques developed over the last few years in numerical analysis.

Key words. Compressible Euler equations. Incompressible Navier-Stokes equations. Computational fluid dynamics. Least squares. Conservative finite elements. Finite differences with an antidiffusion term. Multigrid method. Entropy corrector. Turbulence.

AMS classification. 35F30. 35K60. 35L65. 65N10. 65N30. 76D05. 76G15. 76K05.

PRESENTATION GENERALE

*** Ce travail est en deux parties qui concernent toutes deux la résolution numérique de problèmes issus de la mécanique des fluides : calcul d'écoulements compressibles en dimension deux et trois à partir du modèle d'Euler stationnaire et étude des solutions des équations de Navier-Stokes incompressibles en dimension deux. Les méthodes utilisées sont très différentes d'une application à l'autre. Dans le premier cas les solutions de différents problèmes sont obtenues par méthode de point fixe, linéarisation par la méthode de Newton, minimisation par moindres carrés et approximation par éléments finis conformes des groupes de variables conservatifs ; de plus un terme de décentrement artificiel ou un correcteur d'entropie avec adaptation de maillage est nécessaire pour capturer des solutions transsoniques ou hypersoniques. Dans le deuxième cas les équations sont approchées par un schéma aux différences finies décentrées contenant un terme d'antidiffusion, puis résolues par méthode multigrille avec procédure de relaxation maille à maille.

Outre des méthodes numériques efficaces, une mise en oeuvre importante incluant programmation structurée et vectorisation a été nécessaire pour atteindre de bonnes performances. L'ensemble donne un assez large aperçu des méthodes utilisées et du savoir faire acquis ces dernières années en analyse numérique.

*** La première partie regroupe diverses études, menées au cours des années quatre vingts, qui utilisent la méthode des moindres carrés pour résoudre des équations de type mixte elliptique-hyperbolique. Elles ont commencé par la résolution de modèles simplifiés comme l'équation de Tricomi et le modèle isentropique puis se sont poursuivies par l'étude du système complet des équations d'Euler stationnaires pour un fluide compressible. Le but de ce travail est de montrer que la méthode variationnelle implicite, introduite et développée par J.J. Chattot, J. Guiu-Roux, J. Laminie et l'auteur, est à même de capturer toutes les solutions représentées par ce système. Au moment où nous avons entrepris ce

travail cela semblait être une gageure car si la méthode des moindres carrés était déjà utilisée dans de nombreux cas, personne ne songeait à résoudre l'écoulement d'un fluide parfait compressible sans différencier la zone subsonique de la zone supersonique. Encore aujourd'hui, de fervents défenseurs des schémas décentrés ont du mal à admettre le succès des méthodes centrées, même si les ingrédients utilisés pour calculer les solutions les plus complexes en hypersonique sont à peu près les mêmes dans les deux cas. Une autre originalité est de résoudre directement le système stationnaire au lieu de calculer la limite asymptotique du problème d'évolution. En particulier les méthodes quasi-stationnaires sont très prisées car elles intègrent les propriétés des écoulements stationnaires (équation de Bernoulli par exemple) au modèle d'évolution et permettent donc d'utiliser un des nombreux schémas développés pour résoudre les lois de conservation de la forme $\frac{\partial U}{\partial t} + div\ F(U) = 0$. Dernièrement un jeune chercheur allemand nous confiait son étonnement car on lui avait dit qu'il était impossible de résoudre directement le système stationnaire!

Dans cette partie sont expliquées en détail les évolutions de la méthode, en particulier celles survenues ces deux dernières années pour calculer des écoulements hypersoniques. Parallèlement à J.J. Chattot, J. Guiu-Roux et J. Laminie qui testaient la méthode des moindres carrés sur une équation linéaire du premier ordre, nous avons étudié l'équation scalaire de Tricomi :

$$\frac{\partial^2 u}{\partial y^2} + \alpha(y)\frac{\partial^2 u}{\partial x^2} = f$$

où $\alpha(y)$ est une fonction croissante qui change de signe avec y. Ici le principal inconvénient des moindres carrés est de transformer un problème du deuxième ordre en un problème du quatrième ordre, ce qui nécessite une approximation par éléments finis de degrés élevé et complique beaucoup les calculs. Cependant l'approche moindres carrés s'est révélée très intéressante car la solution est calculée avec la même précision quel que soit le type de l'opérateur. Par ailleurs ce modèle simplifié a permis de tester différents algorithmes de linéarisation en rajoutant le terme $u\frac{\partial u}{\partial x}$ à l'opérateur de Tricomi ; là encore de très bons résultats ont été obtenus avec la méthode de Newton.

Le premier modèle étudié se rapprochant du système d'Euler est le modèle isen-

tropique suivant :

$$\begin{cases} \dfrac{\partial \rho u}{\partial x} + \dfrac{\partial \rho v}{\partial y} = 0 \\ \dfrac{\partial v}{\partial x} - \dfrac{\partial u}{\partial y} = 0 \\ \rho = \left((\gamma - 1) M_\infty^2 \left(H - \dfrac{u^2 + v^2}{2} \right) \right)^{\frac{1}{\gamma - 1}} \end{cases}$$

où l'équation de conservation de la masse est couplée à l'irrotationalité du champ de vitesse et la masse volumique ρ est fonction du champ $\vec{q} = (u, v)^T$ et de trois constantes : γ le rapport de chaleurs spécifiques, M_∞ le nombre de Mach à l'infini et H l'enthalpie totale qui est conservée à travers les chocs. Si l'on écrit un algorithme de point fixe sur ρ il reste à résoudre le système constitué des deux premières équations qui est linéaire à ρ_n fixé. C'est à la résolution de ce système qu'est appliquée la minimisation par moindres carrés et l'on montre que le système du second ordre ainsi obtenu est équivalent au système du premier ordre. Il est bon de remarquer que le passage au second ordre ne complique en rien l'approximation qui peut être faite avec des éléments finis de bas degrés, $\mathbb{P}_1$ ou $\mathbb{Q}_1$ par exemple. De plus ce sont les groupes de variables conservatifs $\rho_n u$ et $\rho_n v$ de la première équation qui sont approchés par éléments finis et non ρ_n, u et v séparément. En effet ce sont ces groupes de variables qui sont conservés à travers les chocs et on peut donc les approcher par des éléments finis C^0 conformes.

Enfin le système d'Euler complet est étudié pour différents écoulements internes subsoniques et transsoniques en dimension deux puis des écoulements externes subsoniques, supersoniques et hypersoniques en dimension deux et trois. Plusieurs formulations sont proposées en considérant l'équation de conservation de l'énergie sous la forme de Bernoulli

$$H = \frac{\gamma p}{(\gamma - 1)\rho} + \frac{q^2}{2}$$

ou sous la forme générale aux dérivées partielles :

$$\frac{\partial(\gamma \rho e + \frac{1}{2}\rho q^2)u}{\partial x} + \frac{\partial(\gamma \rho e + \frac{1}{2}\rho q^2)v}{\partial y} + \frac{\partial(\gamma \rho e + \frac{1}{2}\rho q^2)w}{\partial z} = 0.$$

Suivant la méthodologie utilisée pour les écoulements isentropiques, on utilise l'équation de Bernoulli pour tirer ρ en fonction des autres variables et écrire un algorithme de

point fixe ; ensuite le système restant, constitué de l'équation de conservation de la masse et des équations de conservation de la quantité de mouvement, est linéarisé par la méthode de Newton et résolu par moindres carrés. Ceci donne de très bons résultats en régime subsonique mais n'est pas suffisant en régime transsonique car le point fixe sur ρ n'est plus contractant en supersonique. Un terme de décentrement rajouté à ρ permet d'introduire de la viscosité artificielle et de stabiliser le processus pour des écoulements transsoniques faibles mais ne permet pas de calculer des écoulements hypersoniques. Dans ce cas deux formulations sont envisagées ; d'une part en conservant l'équation de l'énergie sous forme générale, ce qui introduit l'énergie interne e comme nouvelle inconnue et, d'autre part, en utilisant l'équation de Bernoulli pour éliminer la pression. Les deux formulations, suivies d'une linéarisation globale par la méthode de Newton et d'une minimisation par moindres carrés, conduisent à des processus stables pour tous les types d'écoulement, du subsonique à l'hypersonique. La première formulation, utilisée pour calculer des écoulements tourbillonnaires en dimension trois, met en évidence la précision de la méthode quand on calcule l'enthalpie totale qui est constante dans tout le domaine avec une erreur maximale de trois pour cent à l'intérieur de l'enroulement tourbillonnaire. La deuxième formulation a l'avantage de supprimer une inconnue et d'être par conséquent beaucoup plus performante du point de vue du temps de calcul. Cette formulation est utilisée pour calculer des écoulements hypersoniques autour d'un corps émoussé en dimension deux. Pour des nombres de Mach à l'infini élevés ($M_\infty = 8$ par exemple) il faut rajouter au système d'Euler la condition d'entropie pour sélectionner la bonne solution et éviter des zones de pression négative. Cette difficulté commune à, à peu près, toutes les méthodes, est résolue ici en rajoutant à la fonctionnelle des moindres carrés un terme de pénalisation basé sur un correcteur d'entropie. Ceci, couplé à une procédure d'adaptation de maillage, donne des solutions hypersoniques avec un choc détaché capturé sur une maille et bien dessiné.

Nous présentons d'abord les résultats obtenus sur les modèles simplifiés mentionnés ci-dessus ([B1] et [BCLG1]) et ensuite les résultats concernant le système d'Euler complet en deux paragraphes "Ecoulements internes" et "Ecoulements externes". Grâce au terme de décentrement artificiel introduit dans l'algorithme de point fixe des solutions transsoniques avec choc sont calculées dans une tuyère convergente-divergente et un canal

contenant un profil symétrique. Les résultats sont tout à fait comparables aux résultats obtenus par ailleurs avec des schémas décentrés et une comparaison met en évidence l'erreur introduite par le terme de viscosité artificielle ([BCLG2]). Les écoulements externes considérés sont beaucoup plus nombreux, écoulements avec séparation en dimension deux et trois et écoulements hypersoniques avec choc détaché en dimension deux. Le premier cas de séparation étudié est autour d'un cylindre en dimension deux ; le décollement n'est obtenu avec le modèle d'Euler que si l'on impose une condition au bord du cylindre simulant un obstacle ou un orifice. Dans [BCLT1] plusieurs essais montrent l'influence de cette condition sur la taille de la zone de recirculation obtenue ; en particulier si l'on simule un jet à travers un orifice en imposant en un seul point une pression supérieure à la pression d'arrêt de l'écoulement générateur amont cette zone de recirculation fermée se déploie largement. Ce test montre la précision de la méthode qui est sensible à une variation de pression en un seul point. Ensuite sont étudiés des écoulements tourbillonnaires en dimension trois autour d'une plaque rectangulaire ([BCLT2] et [BCL1]) ou de forme delta ([BCL2]) en incidence. Là, le décollement est provoqué par l'arête vive de l'extrémité de plaque ou du bord d'attaque et conduit à des enroulements tourbillonnaires de forme cônique. Il est à noter que ces phénomènes sont capturés spontanément par la méthode sans ajout de condition de Kutta ou de viscosité artificielle. La taille de l'enroulement tourbillonnaire croît avec l'angle d'incidence et décroît, par contre, quand le nombre de Mach à l'infini augmente. Enfin, nous présentons des résultats d'écoulements hypersoniques autour d'un corps émoussé en dimension deux. Les solutions obtenues autour d'un cylindre sont en parfaite concordance avec la théorie et les expériences physiques tant sur la distance du choc à l'obstable que sur la forme du choc. D'autres résultats autour d'une ellipse montrent le déplacement du choc qui se rapproche du profil quand le nombre de Mach à l'infini croît.

*** La deuxième partie, démarrée en collaboration avec C. Jouron en 1987, porte sur l'étude des solutions des équations de Navier-Stokes pour un fluide incompressible. Elle vise à mettre en évidence la transition vers la turbulence en trouvant les nombres de Reynolds pour lesquels la solution stationnaire n'est plus stable. Le cas test choisi est le problème de la cavité entraînée en dimension deux, pour lequel de nombreux résultats sont disponibles dans la littérature soit à partir d'une formulation vitesse-pression, soit

pour une formulation fonction de courant-vorticité. Nous avons choisi une méthode suffisamment générale et performante pour être appliquée en dimension trois. A partir d'une formulation vitesse-pression on utilise une approximation par différences finies sur grilles décalées où la vitesse est donnée au milieu des côtés et la pression au centre des cellules. Ensuite une méthode multigrille est mise en oeuvre sur des grilles successives allant de 4×4 mailles à 512×512 mailles et chaque itération comporte un cycle en V. La résolution sur une grille est faite par une procédure de relaxation maille à maille, ce qui signifie que sur chaque maille l'on résout le système complet de Navier-Stokes pour les cinq inconnues constituées de deux vitesses horizontales, deux vitesses verticales et la pression ; puis l'on passe à la maille suivante. Ainsi, à chaque itération de relaxation, la vitesse est recalculée deux fois et la pression une fois. Cette procédure a l'avantage de régler le problème de la pression qui se cale autour d'une valeur dépendant de l'initialisation ; si l'on veut imposer cette valeur il suffit de l'imposer en un point de la grille la plus fine mais cela a pour effet de ralentir légèrement la convergence.

Dans ce travail, l'effort essentiel porte sur le schéma aux différences finies utilisé pour approcher les termes de convection. En effet, la discrétisation de ces termes a une influence capitale à la fois sur la stabilité de la méthode et sur la représentation de la solution. Pour des nombres de Reynolds faibles ($Re \leq 100$) tous les schémas donnent la même solution sans sous-relaxer ($\omega = 1$), y compris le schéma centré du deuxième ordre. Par contre, dès que le nombre de Reynolds atteint 1000 de grandes différences apparaissent d'un schéma à un autre ; avec le schéma décentré du premier ordre qui contient beaucoup de viscosité artificielle le processus global reste stable en prenant un coefficient ω proche de 1 mais la solution obtenue est très diffuse (les extrema sont rabotés et étalés) alors que pour obtenir la stabilité avec le schéma centré il faut prendre ω inférieur à 0.1 mais la solution calculée est très bonne si l'on se réfère aux résultats obtenus par d'autres auteurs. En fait tous les résultats disponibles sont voisins tant que le nombre de Reynolds n'excède pas mille et donc l'on dispose de données très précieuses pour évaluer la qualité des schémas. Le bon schéma est celui qui réalise l'équilibre entre la stabilité et la précision. Il est connu qu'un schéma décentré revient à introduire un terme de diffusion artificielle dans l'opérateur ; dans le cas d'un problème visqueux ce terme artificiel est en concurrence avec le terme existant dans le modèle et s'il est trop

important la viscosité réelle du problème approché n'a plus rien à voir avec la viscosité du problème continu. C'est cette remarque qui va nous conduire à étudier en détail une classe de schémas décentrés (schémas TVD, à variation totale décroissante, du premier ordre) et à proposer l'ajout d'un terme d'antidiffusion pour amoindrir les effets de la viscosité artificielle du schéma.

Dans un premier paragraphe nous partons des schémas TVD utilisés pour résoudre une équation scalaire de conservation de la forme $\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0$, pour construire un schéma adapté à la résolution du problème visqueux correspondant $\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} = 0$. L'idée est de combiner un terme d'antidiffusion avec le terme $-\nu \frac{\partial^2 u}{\partial x^2}$ de l'équation et le terme de viscosité artificielle introduit par le schéma numérique. De cette façon on peut construire un schéma convergent qui diffuse moins que le schéma initial. Dans le second paragraphe nous montrons d'abord comment discrétiser le terme d'antidiffusion sans altérer la stabilité de la procédure de relaxation, puis appliquons le schéma aux équations de Navier-Stokes stationnaires ([BJ1] et [BJ2]). Les résultats obtenus sont très proches de ceux disponibles pour $Re \leq 1000$ bien que pour $Re = 1000$ les extrema des tourbillons secondaires que nous obtenons sont supérieurs à ceux des autres auteurs. Nous pensons que cela est dû au correcteur d'antidiffusion qui donne des solutions plus précises. Ensuite pour $Re = 5000$ le schéma permet de calculer une solution stationnaire stable dans la cavité entraînée carrée $(0,1)\times(0,1)$ sur une grille fine 512×512 mailles ; cette solution exhibe des tourbillons secondaires et tertiaires bien représentés et la convergence est très régulière avec comme paramètre de sous-relaxation $\omega = 0.3$. Enfin pour $Re = 7500$ et $Re = 10000$ il n'y a plus convergence vers une solution stationnaire stable suivant la grille où l'on se trouve. En particulier pour $Re = 10000$ il n'y a pas convergence sur les grilles grossières puis convergence sur la grille 256×256 et non convergence à nouveau sur la grille 512×512. Nous pensons que cela s'explique par le fait que sur les grilles grossières il n'y a pas assez de points dans la couche limite pour représenter la solution, que sur la grille 256×256 les effets de couche limite sont pris en compte mais pas les effets turbulents et que sur la grille la plus fine les effets turbulents commencent à apparaître et déstabilisent la solution de la grille précédente. Ceci correspond à la théorie car sur la grille la plus fine le pas de discrétisation est de l'ordre de la longueur de Kolmogorov en $Re^{3/4}$. En considérant la procédure de relaxation comme une sorte de

temps on peut voir l'évolution de la solution au cours des itérations et nous conjecturons que la transition vers la turbulence a lieu au moment où de petits tourbillons se forment le long des parois et déstabilisent les tourbillons secondaires. Notons en conclusion que, sur ce problème de la cavité carrée entraînée, c'est la première fois que des résultats numériques mettent en évidence la transition vers la turbulence et exhibent un nombre de Reynolds critique compris entre 5000 et 7500 pour lequel il y a perte de stabilité de la solution stationnaire. Par ailleurs des résultats récents sur l'étude du problème linéarisé corroborent cette estimation du nombre de Reynolds critique car le premier exposant du Lyapunov s'annule dans la même zone. Cette étude fait l'objet du dernier paragraphe de cette partie où nous montrons comment calculer le premier exposant de Lyapunov à partir des solutions stationnaires calculées dans [BJ2].

PARTIE A

RESOLUTION DES EQUATIONS D'EULER COMPRESSIBLES

PAR UNE METHODE VARIATIONNELLE

EN ELEMENTS FINIS

Introduction

Modèles simplifiés

Ecoulements internes

Ecoulements externes

Conclusion

INTRODUCTION

Le but de cette première partie est le calcul de solutions stationnaires des équations d'Euler compressibles. L'originalité de l'approche proposée ci-dessous est de calculer ces solutions directement à partir du modèle stationnaire et non comme limite asymptotique de solutions du problème d'évolution. C'est-à-dire que nous considérons les équations d'Euler stationnaires écrites sous forme conservative (en dimension deux pour simplifier l'écriture) comme suit :

$$\begin{aligned} \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} &= 0 \\ \frac{\partial \rho u^2 + p}{\partial x} + \frac{\partial \rho uv}{\partial y} &= 0 \\ \frac{\partial \rho uv}{\partial x} + \frac{\partial \rho v^2 + p}{\partial y} &= 0 \\ \frac{\gamma p}{(\gamma - 1)\rho} + \frac{u^2 + v^2}{2} &= H \end{aligned} \qquad (A.0)$$

où l'on reconnaît pour la masse volumique ρ, la vitesse $\vec{q} = (u, v)^T$ et la pression p l'équation de conservation de la masse, les deux équations de conservation de la quantité de mouvement et l'équation de l'énergie sous la forme algébrique de Bernoulli dans le cas des écoulements stationnaires (γ est le rapport de chaleurs spécifiques et H l'enthalpie totale ; dans l'air $\gamma = 1,4$ et l'on tire $H = 3$ en prenant 1 comme nombre de Mach de référence).

Les équations sont écrites sous forme conservative et sont traitées comme telles, en particulier l'approximation par éléments finis est faite sur les groupes de variables conservés à travers les chocs par les conditions de Rankine-Hugoniot. Ceci permet d'utiliser pour approcher ces variables des éléments finis C^0 conformes quel que soit l'écoulement ; une telle approche ne serait pas cohérente pour les variables ρ, u, v et p séparément qui peuvent présenter de fortes discontinuités.

L'idée de départ pour résoudre le système d'Euler est d'utiliser une méthode de moindres carrés pour transformer le système du premier ordre en un système du deuxième ordre que l'on approche par une méthode centrée sur tout le domaine, que l'écoulement soit subsonique ou supersonique.

Ainsi, avec les deux remarques ci-dessus, nous nous proposons de calculer des solutions

stationnaires des équations d'Euler sans tenir compte de la nature de l'écoulement (et donc des équations) et de capturer les ondes de choc soit par une méthode de viscosité artificielle soit en introduisant un correcteur d'entropie. En effet le système ci-dessus a des solutions non physiques qui sont éliminées en ajoutant l'inéquation d'entropie (notée s) suivante :

$$div\ \rho\vec{q}s \geq 0$$

qui impose la croissance de s à travers les discontinuités et supprime donc les chocs de détente.

Dans cette partie nous présentons d'abord deux études préliminaires sur des modèles simplifiés : équation de Tricomi et équation du potentiel pour les écoulements isentropiques irrotationnels en vue de tester les principales caractéristiques de la méthode. Ensuite nous discutons les résultats obtenus pour le modèle d'Euler complet d'une part sur des écoulements internes dans une tuyère convergente-divergente et un canal contenant un profil, et, d'autre part, sur des écoulements externes, autour de profils présentant des recirculations et des ondes de choc.

MODELES SIMPLIFIES

Les premières études présentées dans ce paragraphe ont permis de tester sur des modèles simples les principaux ingrédients utilisés pour résoudre les équations d'Euler : minimisation des moindres carrés, algorithme de point fixe et linéarisation par la méthode de Newton.

Un premier travail ([B1]) montre que l'équation mixte de Tricomi peut être résolue globalement par moindres carrés sans se soucier du changement de type de l'opérateur elliptique-hyperbolique. C'est là une nouveauté car jusqu'alors toutes les méthodes mises en oeuvre pour résoudre ce problème ou transformaient l'opérateur ou utilisaient des schémas d'approximation (en différences finies par exemples) différents suivant le type de l'opérateur, centrés dans la partie elliptique et décentrés dans la partie hyperbolique. D'autre part une extension non linéaire du problème de Tricomi nous a permis de montrer la supériorité de la méthode de Newton sur les algorithmes standards de linéarisation. Ces premiers résultats sont de bonne augure ; cependant ce modèle est assez loin du système d'Euler, l'opérateur est un opérateur scalaire du deuxième ordre et la non linéarité introduite n'est que du premier ordre. L'équation de Tricomi aurait pu être transformée en un système du premier ordre, ce qui aurait grandement simplifié l'approximation du problème de moindres carrés. En effet nous sommes ici en présence d'un problème du quatrième ordre qui nécessite une approximation par éléments finis de haut degré. Alors qu'un système du premier ordre est transformé en un système du deuxième ordre qui peut être facilement approché par des éléments finis de bas degré comme il est montré ci-dessous.

Le deuxième problème, étudié initialement par J.J. Chattot, J. Laminie et J. Guiu-Roux, concerne le modèle isentropique du potentiel. Cette fois l'équation du potentiel est réécrite sous forme d'un système du premier ordre ([BCLG1]) comme suit en $2D$:

$$\begin{aligned} \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} &= 0 \\ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} &= 0 \\ \rho = \left[(\gamma - 1) M_\infty^2 \left(H - \frac{u^2 + v^2}{2} \right) \right]^{\frac{1}{\gamma - 1}} & \end{aligned} \qquad (A.1)$$

où la première équation est l'équation de conservation de la masse que l'on retrouve dans le système d'Euler, la deuxième équation traduit l'irrotationalité du champ de vitesse $\vec{q} =$

$(u,v)^T$ et la dernière équation algébrique exprime la masse volumique ρ comme une fonction de la vitesse, du nombre de Mach à l'infini M_∞, du rapport de chaleurs spécifiques γ et de l'enthalpie H.

Le système global (A. 1) non linéaire en (ρ,u,v) est linéarisé et résolu par un algorithme de point fixe en ρ. Ainsi à chaque itération il reste à résoudre le système linéaire (à ρ_n fixé) constitué des deux premières équations en u et v :

$$\begin{aligned} \frac{\partial \rho_n u}{\partial x} + \frac{\partial \rho_n v}{\partial y} &= 0 \\ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} &= 0 \end{aligned} \qquad (A.2)$$

Considérons maintenant un ouvert Ω sur lequel le problème (A.1) est posé et notons $\partial\Omega = \Gamma_p \cup \Gamma_p$ où Γ_p représente les parois rigides de la géométrie (canal, tuyère, aile, etc.) ; alors le système (A.2) est associé à la condition de glissement :

$$\vec{q}\cdot\vec{n} = 0 \text{ sur } \Gamma_p \qquad (A.3)$$

où $\vec{n}$ est la normale unitaire extérieure à Ω. Soient

$$A = \begin{pmatrix} \dfrac{\partial \rho_n \cdot}{\partial x} & \dfrac{\partial \rho_n \cdot}{\partial y} \\ -\dfrac{\partial \cdot}{\partial y} & \dfrac{\partial \cdot}{\partial x} \end{pmatrix}$$

l'opérateur dans (A.2), et

$$\mathcal{D}(A) = \{\, \vec{q} = (u,v)^T \ tq\ \rho_n \vec{q} \in H_{div},\ rot\ \vec{q} \in L^2(\Omega)\ et\ \vec{q}\cdot\vec{n} = 0 \text{ sur } \Gamma_p \}$$

où

$$H_{div} = \{ \vec{r} \in (L^2(\Omega))^2\ tq\ div\ \vec{r} \in L^2(\Omega) \};$$

alors le problème (A.2),(A.3) peut s'écrire

$$\begin{aligned} A\vec{q} &= 0 \text{ dans } \Omega \\ \vec{q} &\in \mathcal{D}(A) \end{aligned} \qquad (A.4)$$

Bien que l'opérateur A soit elliptique le problème (A.4) est résolu par moindres carrés en minimisant sur l'espace $\mathcal{D}(A)$ la fonctionnelle :

$$J(\vec{q}) = \frac{1}{2}\int_\Omega \left(\frac{\partial \rho_n u}{\partial x} + \frac{\partial \rho_n v}{\partial y}\right)^2 dxdy + \frac{1}{2}\int_\Omega \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)^2 dxdy;$$

ce qui correspond au problème variationnel :

$$\begin{cases} \text{Trouver } \vec{q} \in \mathcal{D}(A) \text{ tel que} \\ \displaystyle\int_\Omega \left(\frac{\partial \rho_n u}{\partial x} + \frac{\partial \rho_n v}{\partial y}\right)\left(\frac{\partial \rho_n \bar{u}}{\partial x} + \frac{\partial \rho_n \bar{v}}{\partial y}\right) dxdy \\ \displaystyle+ \int_\Omega \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\left(\frac{\partial \bar{v}}{\partial x} - \frac{\partial \bar{u}}{\partial y}\right) dxdy = 0 \quad \forall \bar{q} = (\bar{u}, \bar{v})^T \in \mathcal{D}(A) \end{cases}$$

et au problème interprété suivant :

$$\begin{aligned} -\rho_n \frac{\partial}{\partial x}\left(\frac{\partial \rho_n u}{\partial x} + \frac{\partial \rho_n v}{\partial y}\right) + \frac{\partial}{\partial y}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) &= 0 \text{ dans } \Omega \\ -\rho_n \frac{\partial}{\partial y}\left(\frac{\partial \rho_n u}{\partial x} + \frac{\partial \rho_n v}{\partial y}\right) - \frac{\partial}{\partial x}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) &= 0 \text{ dans } \Omega \\ \vec{q} \cdot \vec{n} &= 0 \text{ sur } \Gamma_p \\ \frac{\partial \rho_n u}{\partial x} + \frac{\partial \rho_n v}{\partial y} &= 0 \text{ sur } \Gamma_p \\ \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} &= 0 \text{ sur } \mathcal{D}\Omega \end{aligned}$$

Ce problème de moindres carrés est un problème du deuxième ordre auquel on a ajouté comme conditions aux limites la conservation de la masse sur les frontières libres et l'irrotationalité sur tout le bord du domaine. De plus il peut être réécrit sous la forme :

$$\begin{aligned} & A^*(A\vec{q}) = 0 \text{ dans } \Omega \\ & \vec{q} \in \mathcal{D}(A) \\ & A\vec{q} \in \mathcal{D}(A^*) \end{aligned} \qquad (A.5)$$

où l'opérateur adjoint est

$$A^* = \begin{pmatrix} -\rho_n \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} \\ -\rho_n \dfrac{\partial}{\partial y} & -\dfrac{\partial}{\partial x} \end{pmatrix}$$

et $\mathcal{D}(A^*) = \{\vec{\bar{q}} = (\bar{u}, \bar{v})^T \; tq \; \rho_n \vec{\bar{q}} \in H_{div}, rot \; \vec{\bar{q}} \in L^2(\Omega)$ et $\bar{u} = 0$ sur Γ_p, $\bar{v} = 0$ sur $\mathcal{D}\Omega\}$.

Enfin le problème (A.5) est équivalent au problème (A.4) car le problème adjoint :

$$\begin{aligned} & A^*\vec{\bar{q}} = 0 \text{ dans } \Omega \\ & \vec{\bar{q}} \in \mathcal{D}(A^*) \end{aligned}$$

admet une solution unique. Ceci valide l'approche moindres carrés pour résoudre le système du premier ordre ; de plus l'approximation ne nécessite aucun travail supplémentaire puisqu'elle peut être faite à partir des éléments finis $\mathbb{P}1$ ou $\mathbb{Q}1$ conformes par exemple en posant:

$$V_h = \{\vec{q}_h = (u_h, v_h)^T \ tq \ \rho_n \vec{q}_h \in W_h, rot \ \vec{q}_h \in L^2(\Omega) \text{ et } \vec{q}_h.\vec{n} = 0 \text{ sur } \Gamma_p\}$$

avec

$$W_h = \{\vec{r}_h = (r_{h1}, r_{h2})^T \in (C^0(\bar{\Omega}))^2 \ tq \ r_{h,i/K} \in \mathbb{P}1 \text{ ou } \mathbb{Q}1,$$
$$\forall \ i = 1 \text{ ou } 2 \text{ et } \forall \ K \in \mathcal{T}_h\}$$

où on note $\mathcal{T}_h$ la triangulation de l'ouvert Ω. W_h est une approximation conforme de H_{div} et V_h est une approximation conforme de $\mathcal{D}(A)$; remarquons que l'on approche par des fonctions continues les groupes de variables conservatives $\rho_n u$ et $\rho_n v$ mais pas les composantes de la vitesse qui peuvent être discontinues. Ceci est le point principal de l'approximation qui permet de capturer des solutions avec choc.

COMPUTER METHODS IN APPLIED MECHANICS AND ENGINEERING 39 (1983) 117–129
NORTH-HOLLAND

A NON-CONFORMING FINITE ELEMENT METHOD FOR SOLVING A LEAST-SQUARE FORMULATION OF TRICOMI'S PROBLEM*

Ch.H. BRUNEAU
Laboratoire d'Analyse Numérique, Université de Paris-Sud, Orsay, France

Received 23 November 1981
Revised manuscript received 9 December 1982

The purpose of this work is to solve in the plane Tricomi's equation of mixed type by means of global least-square formulation. The resolution is obtained with Zienkiewicz's finite element which verifies the patch test, and the incomplete Cholesky conjugate gradient algorithm (ICCG) yields very fast convergence. The least-square formulation is still applicable when a non-linear term is introduced, and the solution is searched for by Newton's method.

1. Introduction

Let Ω be the bounded domain of $\mathbb{R}^2$ with boundary $\partial\Omega$ shown in Fig. 1. We consider Tricomi's problem in the Frankl's generalized form:

$$\begin{cases} Au = \dfrac{\partial^2 u}{\partial y^2} + \alpha(y)\dfrac{\partial^2 u}{\partial x^2} = f \quad \text{in } \Omega, \\ u = 0 \quad \text{on } \Gamma \cup \gamma \end{cases} \tag{1.1}$$

where $\alpha(y)$, as is usually assumed (see, for instance, [1]), is an increasing function of y having the same sign as y. The linear partial differential operator A is elliptic for $y > 0$, hyperbolic for $y < 0$ and has a parabolic degeneracy along the straight line $y = 0$. In the hyperbolic region the characteristic curves are given by the equation: $dx^2 + \alpha(y)\,dy^2 = 0$, C is characteristic and γ is either characteristic or subcharacteristic.

The numerical solution of Tricomi's problem has been the subject of much research (cf. [2–5] and references therein). In this paper our purpose is to solve numerically the problem (1.1), in least-square variational form, using Zienkiewicz's non-conforming finite element. We also give some results for the following non-linear problem:

$$\begin{cases} Au + u\dfrac{\partial u}{\partial x} = f \quad \text{in } \Omega, \\ u = 0 \quad \text{on } \Gamma \cup \gamma, \end{cases} \tag{1.2}$$

the solution of which is obtained by Newton's method applied to a least-square functional.

*This article is an extract of Ch. H. Bruneau's Thèse de 3ème cycle, Université de Paris-Sud, Centre d'Orsay, France.

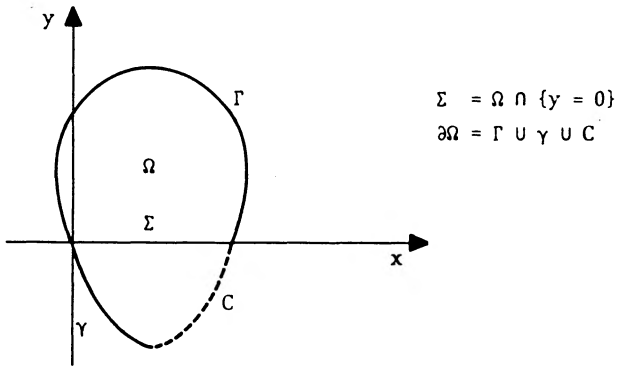


Fig. 1. Domain.

2. Least-square formulation

In this paper we assume

$$u \in H^2(\Omega). \tag{2.1}$$

In the case where $\alpha(y) = y$ and γ is subcharacteristic it is known (cf. [6–8]) that $u \in H^2(\Omega)$ but this is not always true when γ is characteristic or when $\alpha(y) = \text{sign}(y)$ (therefore (2.1) may be a restrictive assumption)

The idea is to minimize the functional

$$J(u) = \tfrac{1}{2}\int_\Omega |Au - f|^2 \mathrm{d}x\, \mathrm{d}y$$

on the space $V = \{u \in H^2(\Omega);\ u = 0 \text{ on } \Gamma \cup \gamma\}$, which yields the following variational problem:

$$\begin{cases} \text{Let } f \in H^2(\Omega) \text{ be a given function,} \\ \text{find } u \in V \text{ such that:} \\ \displaystyle\int_\Omega (Au - f)Av \,\mathrm{d}x\, \mathrm{d}y = 0 \quad \text{for all } v \in V. \end{cases} \tag{2.2}$$

Using Green's formula (where $B = (\alpha(y)\,\partial/\partial x,\ \partial/\partial y)$),

$$\int_\Omega A(Au - f)v \,\mathrm{d}x\, \mathrm{d}y - \int_\Omega (Au - f)Av \,\mathrm{d}x\, \mathrm{d}y =$$

$$= \int_{\partial\Omega} [n \cdot B(Au - f)v - (Au - f)(n \cdot Bv)] \,\mathrm{d}\tau$$

and noting that $n \cdot B = \sqrt{-\alpha(y)}\,\partial/\partial\tau$ on the characteristic curve, it is clear that the solution u

of (2.2) is also the solution of the following problem:

$$\begin{cases} A(Au-f)=0 \quad \text{in } \Omega, \\ u=0 \quad \text{on } \Gamma \cup \gamma, \\ Au-f=0 \quad \text{on } \Gamma, \\ \dfrac{\partial}{\partial \tau}[(-\alpha(y))^{1/4}(Au-f)]=0 \quad \text{on } C \\ Au-f=0 \quad \text{on } \gamma \text{ (when } \gamma \text{ is noncharacteristic)}. \end{cases} \tag{2.3}$$

Let A be the $L^2(\Omega)$ unbounded linear operator with domain $\mathcal{D}(A)=V$, the adjoint operator has as domain of definition:

$$\mathcal{D}(A^*)=\left\{u \in H^2(\Omega); \frac{\partial}{\partial \tau}[(-\alpha(y))^{1/4} u]=0 \text{ on } C, u=0 \text{ on } \Gamma \text{ or } \Gamma \cup \gamma\right\}$$

and then (2.3) is equivalent to

$$\begin{cases} A^*(Au-f)=0 \quad \text{in } \Omega, \\ u \in \mathcal{D}(A), \\ (Au-f) \in \mathcal{D}(A^*), \end{cases}$$

but

$$\begin{cases} Au-f=0 \quad \text{in } \Omega, \\ \\ u \in \mathcal{D}(A), \end{cases} \Leftrightarrow \begin{cases} A^*(Au-f)=0 \quad \text{in } \Omega, \\ u \in \mathcal{D}(A), \\ (Au-f) \in \mathcal{D}(A^*), \end{cases}$$

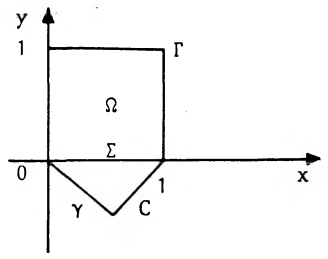
because A^* is an injective linear operator (see, for instance [1]). Hence the problem (2.2) is equivalent to the initial problem (1.1).

3. Finite dimension

We consider the model problem when $\alpha(y)=\text{sign}(y)$ in the bounded domain Ω: the union of the unit square $(0,1)\times(0,1)$ and the hyperbolic triangle bounded by the interval $\Sigma : (0,1)$ and the characteristic lines $\gamma : x+y=0$ and $C : x-y-1=0$ (Fig. 2).

On the polygon $\bar{\Omega}$ we construct a triangulation $\mathcal{T}_h : \bar{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$ where the finite elements K are triangles or rectangles with the usual assumptions that their interiors are pairwise disjoint, and that any side of an element is either a side of another element or a subset of the boundary $\partial\Omega$. We let

$$h_K = \text{diam}(K) \quad \text{for all } K \in \mathcal{T}_h \quad \text{and} \quad h = \max_{K \in \mathcal{T}_h} h_K .$$



$$\frac{\partial^2 u}{\partial y^2} + \text{sign}(y)\,\frac{\partial^2 u}{\partial x^2} = f \quad \text{in } \Omega$$

$$u = 0 \quad \text{on } \Gamma \cup \gamma$$

Fig. 2. The model problem.

Also, over each triangle K, a finite dimensional space P_K of functions is given. Then to the triangulation $\mathcal{T}_h$ and the space P_K, a finite-dimensional space V_h of functions $v_h : \bar{\Omega} \to \mathbb{R}$ is associated in such a way that

$$v_h|_K \in P_K \quad \text{for all } K \in \mathcal{T}_h .$$

Let $a(u, v) = \int_\Omega AuAv \, dx \, dy$ be a bilinear form and $(f, v) = \int_\Omega fAv \, dx \, dy$ be a linear form, the conforming method ($V_h \subset V$) is to find a function $u_h \in V_h$ such that $a(u_h, v_h) = (f, v_h)$ for all $v_h \in V_h$. This may be achieved by means of $C^1(\bar{\Omega})$ conforming finite elements, but these have numerous degrees of freedom with normal or second-order derivatives. Therefore we prefer the non-conforming method ($V_h \not\subset V$).

Thus the space V_h violates the inclusion $V_h \subseteq V$, and the problem is to find a function $u_h \in V_h$ such that

$$a_h(u_h, v_h) = \sum_{K \in \mathcal{T}_h} \int_K Au_h Av_h dx \, dy = \sum_{K \in \mathcal{T}_h} \int_K fAv_h dx \, dy$$

for all $v_h \in V_h$.

For the plate problem we have an analogous discrete problem (see [9]) and if we let

$$E_h(u, v_h) = (f, v_h) - a_h(u, v_h),$$

$$E_h(u, v_h) = E_h^0(u, v_h) + E_h^1(u, v_h)$$

where

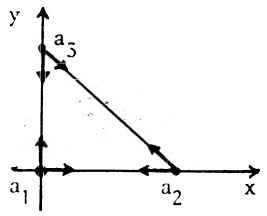
$$E_h^0(u, v_h) = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \left(\frac{\partial \Delta u}{\partial n} v_h - (1 - \sigma) \frac{\partial^2 u}{\partial n \partial t} \frac{\partial v_h}{\partial t} \right) d\tau,$$

$$E_h^1(u, v_h) = \sum_{K \in \mathcal{T}_h} \int_{\partial K} \left(-\Delta u + (1 - \sigma) \frac{\partial^2 u}{\partial t^2} \right) \frac{\partial v_h}{\partial n} d\tau,$$

we have on one hand $E_h^0(u, v_h) \equiv 0$ where $V_h \subset C^0(\bar{\Omega})$, and on the other an $O(h)$ estimation for $E_h^1(u, v_h)$ if the following patch test is verified:

$$E_h^1(u, v_h) = 0 \quad \text{for all } u \in \mathbb{P}_2 \text{ and all } v_h \in V_h$$

where $\mathbb{P}_2$ is the space of polynomials of degree 2 in two variables x, y (cf. [10]).



• value of the function at the point.

→ value of the first derivative at the point in the indicated direction.

Fig. 3. Zienkiewicz's triangle.

So we want to find finite elements which verify (see [11] for more details):
- the functions of P_K are twice continuously differentiable,
- $\mathbb{P}_2 \subseteq P_K$,
- the patch test is satisfied;

and such that:
- the functions of V_h are continuous on the degrees of freedom of the elements K,
- $V_h \subseteq C^0(\bar{\Omega}) \cap H^1(\Omega)$.

The patch test is conjectured by analogy with the plate problem but no justification is done for this discrete problem. We will discuss afterwards the numerical influence of the patch test.

Two finite elements satisfy all these conditions. They have as degrees of freedom the value of the function and the first derivatives at the vertices. These are Adini's rectangle and Zienkiewicz's triangle. But the quadrangulation of the hyperbolic triangle does not give a family of affine-equivalent finite elements and so we choose Zienkiewicz's triangle which is defined as shown in Fig. 3.

We note: $p_i = p(a_i)$; $Dp_{i,j} = Dp(a_i)(a_j - a_i)$ and G the centre of gravity.

$$\Sigma_K = \{p_i,\ 1 \leq i \leq 3;\ Dp_{i,j},\ 1 \leq i, j \leq 3,\ i \neq j\},$$

$$P_K = \left\{p \in \mathbb{P}_3;\ p(G) = \tfrac{1}{3}\sum_{i=1}^{3} p_i + \tfrac{1}{18} \sum_{\substack{1 \leq i,j \leq 3 \\ i \neq j}} Dp_{i,j}\right\}$$

and the patch test is verified by the elliptic operator of the plate problem if all sides of all triangles are parallel to three directions only (cf. [12]). In $\bar{\Omega}$ the grid is such that this condition is fulfilled either in the whole domain or separately in Ω_E (elliptic region) and in Ω_H (hyperbolic region) as in Fig. 4.

a. Condition in the whole domain.

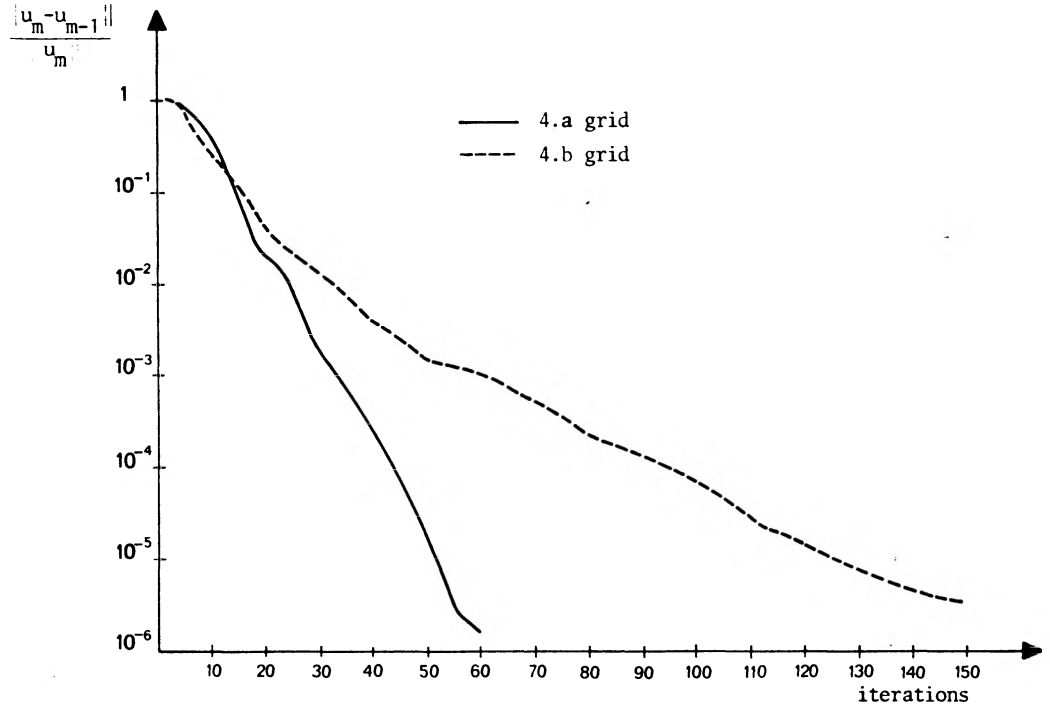
b. Condition in Ω_E and in Ω_H

Fig. 4. The meshes.

4. Numerical results

We first solve with successive overrelaxation the model problem ($\alpha(y) = \text{sign}(y)$), which solution is $u \in V_h$ (Fig. 5). The results are very good because the $H^1(\Omega)$ error is of the same order as the round-off errors in simple precision arithmetics on an IBM 168/370. Nevertheless more iterations are required for the grid shown in Fig. 4(b) than for the grid shown in Fig. 4(a). Hence it seems that the patch test condition must be respected in the whole domain, otherwise convergence is slower. We have also used the faster conjugate gradient method (cf. [13]) to improve the convergence. As the matrix has a bad conditioning, we used the incomplete Cholesky conjugate gradient method (see [14]). Now the solution is obtained four times as fast (in number of iterations) as with the relaxation method (Fig. 6). So we can use finer grids. We

a) Convergence with the two grids.

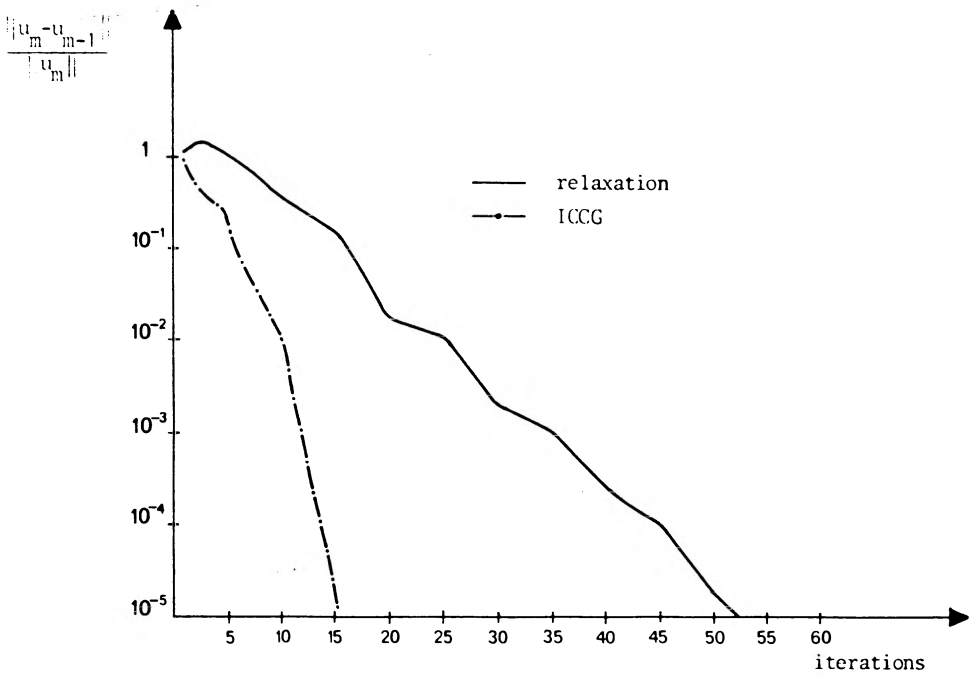


b) $\|u - u_m\|_{H^1(\Omega)}$ norm.

$$\begin{cases} \dfrac{\partial^2 u}{\partial y^2} + \text{sign}(y)\,\dfrac{\partial^2 u}{\partial x^2} = -2 \text{ in } \Omega \\ u = -y^2 - xy + y \text{ on } \Gamma \cup \gamma \end{cases}$$

	1.a grid	1.b grid
iteration m =	60	160
model difference	1.77×10^{-6}	1.56×10^{-6}
$\|u-u_m\|_1$	2.22×10^{-5}	1.86×10^{-5}

Fig. 5. The model problem with $u \in V_h$.



$$\begin{cases} \dfrac{\partial^2 u}{\partial y^2} + \operatorname{sign}(y)\,\dfrac{\partial^2 u}{\partial x^2} = -2 & \text{in } \Omega \\ u = -y^2 - xy + y & \text{on } \Gamma \cup \gamma \end{cases}$$

	relaxation	ICCG
iteration m =	60	15
$\|u-u_m\|$	2.22×10^{-5}	6.24×10^{-5}

Fig. 6. Convergence of relaxation and incomplete Cholesky conjugate gradient methods for a solution $u \in V_h$.

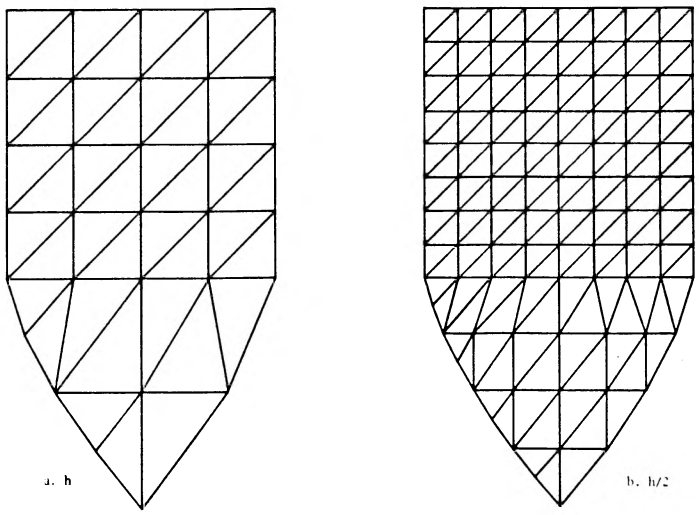


Fig. 7. Grids for Tricomi's problem.

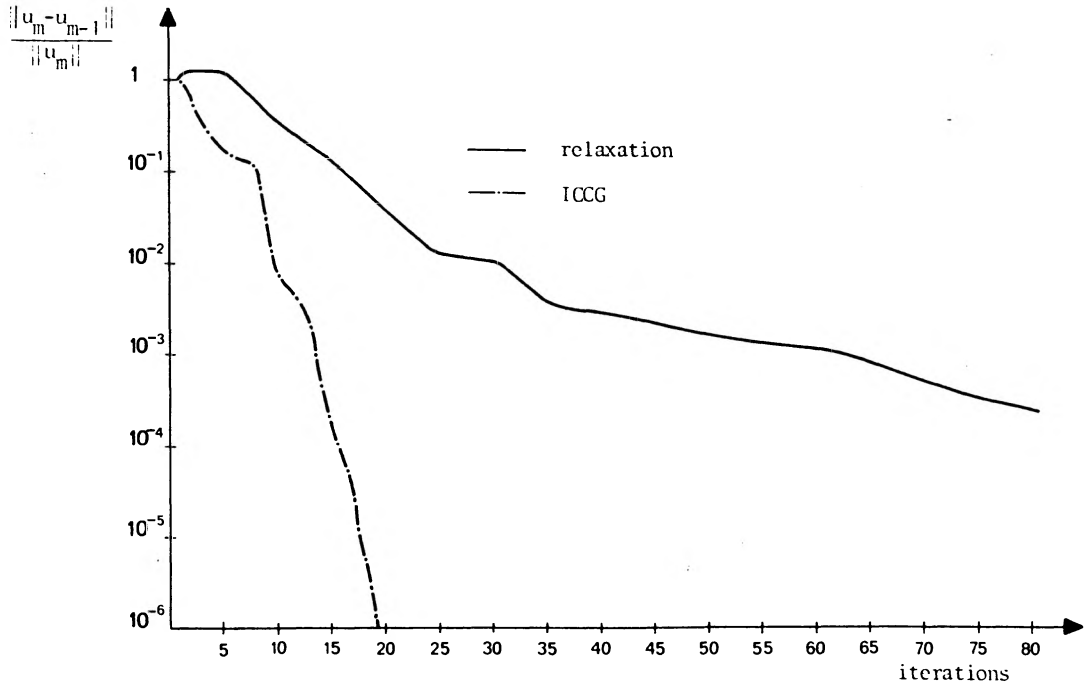
Table 1
$\partial^2 u/\partial y^2 + \text{sign}(y)\,\partial^2 u/\partial x^2 = 2\sin 2\pi y(-2\pi^2 x + \text{sign}(y))$
in Ω; $u = x^2 \sin 2\pi y$ on $\Gamma \cup \gamma$

h	$H^1(\Omega)$ error	$H^1(\Omega)$ convergence
2.5×10^{-1}	5.41×10^{-1}	1.41
1.25×10^{-1}	2.04×10^{-1}	1.69
6.25×10^{-2}	6.29×10^{-2}	

consider a solution $u \notin V_h$ and grids which have respectively h, $\frac{1}{2}h$ and $\frac{1}{4}h$ as mesh sizes. The results show that $\|u - u_h\|_{H^1(\Omega)} = O(h^s)$ as $h \to 0$ where $1.4 \leqslant s \leqslant 1.7$ (Table 1).

We now solve Tricomi's problem ($\alpha(y) = y$) where the characteristic curves are:

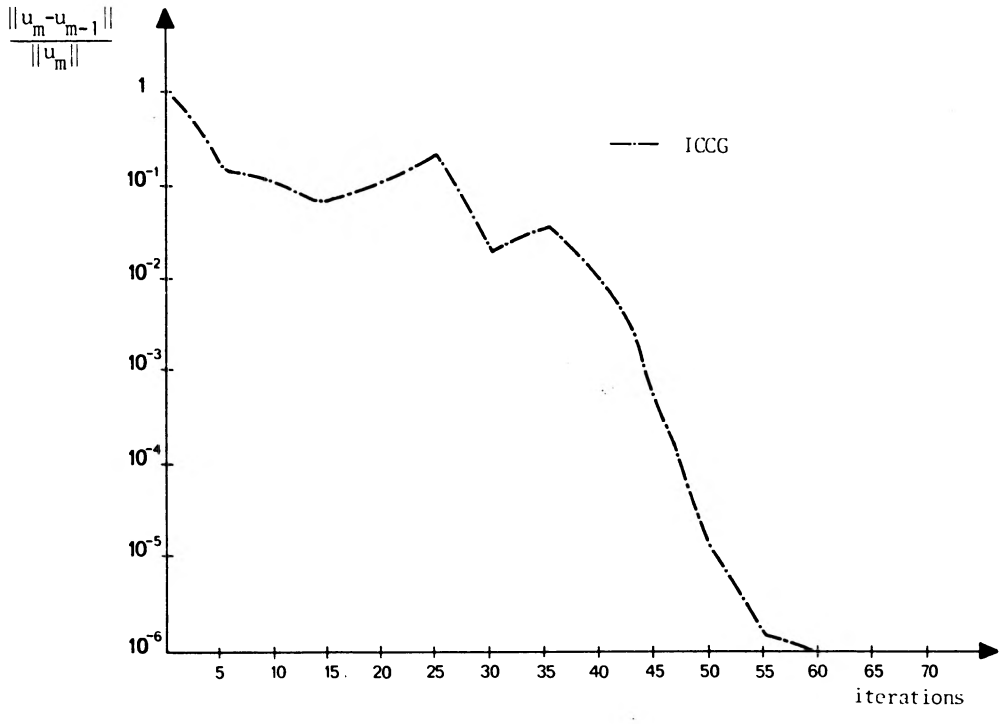
$$\gamma : x = \tfrac{2}{3}(-y)^{3/2} \quad \text{and} \quad C : x = -\tfrac{2}{3}(-y)^{3/2} + 1 .$$



$$\begin{cases} \dfrac{\partial^2 u}{\partial y^2} + y\,\dfrac{\partial^2 u}{\partial x^2} = -2 & \text{in } \Omega \\ u = -y^2 - xy + y & \text{on } \Gamma \cup \gamma \end{cases}$$

	relaxation	ICCG
iteration m =	150	20
$\|u-u_m\|_1$	1.88×10^{-4}	6.20×10^{-5}

Fig. 8. Solution of Tricomi's problem with the 7(a) grid.



$$\begin{cases} \dfrac{\partial^2 u}{\partial y^2} + y\dfrac{\partial^2 u}{\partial x^2} = -2 \text{ in } \Omega \\ u = -y^2 - xy + y \text{ on } \Gamma \cup \gamma \end{cases}$$

iteration m =	80
model difference	1.31×10^{-4}
$\|u-u_m\|_{H^1(\Omega)}$	7.02×10^{-4}

Fig. 9. Solution of Tricomi's problem with the 7(b) grid.

The grids used are shown in Fig. 7 and the patch test condition is not respected now in the whole domain. For the same solution $u \in V_h$ the results again show that the convergence is slower when the grid does not respect the patch test condition (Fig. 8). This is especially true with relaxation because the number of iterations becomes very large. We see from this example that the ICCG method is of great interest even though an ICCG iteration needs twice as much time as relaxation. Moreover we obtain the solution very fast, when the grid of Fig. 7(b) is used (Fig. 9).

5. Non-linear problem

The problem (1.2) is solved by a least-square method, corresponding to the following variational problem:

$$\begin{cases} \text{Let } f \in H^2(\Omega) \text{ be a given function,} \\ \text{find } u \in V \text{ such that} \\ \\ \displaystyle\int_\Omega \left(Au + u\frac{\partial u}{\partial x} - f\right)\left(Av + v\frac{\partial u}{\partial x} + u\frac{\partial v}{\partial x}\right)dx\,dy = 0 \\ \\ \text{for all } v \in V. \end{cases}$$

Here A is the linear operator of problem (1.1). We first solve this problem with the aid of the following algorithm of linearization (u_1 denotes successive iterations):

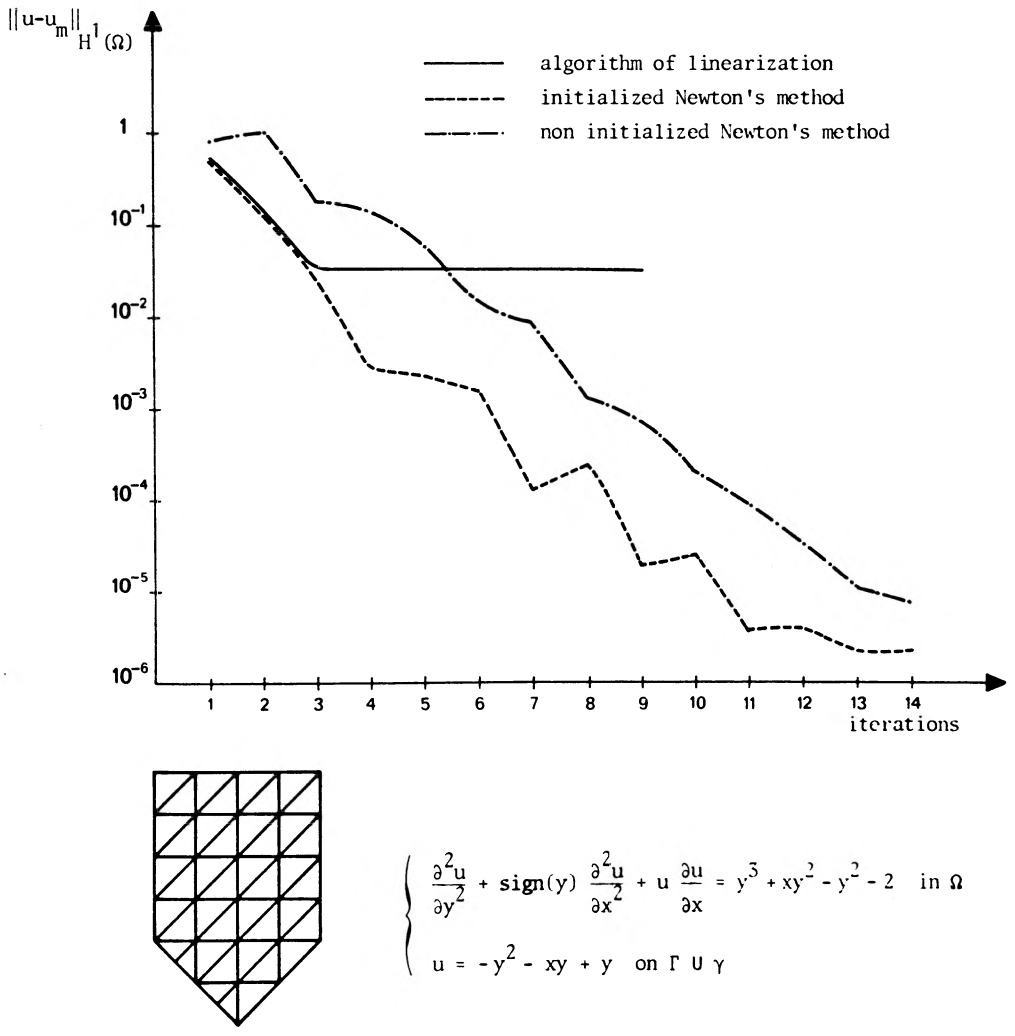


Fig. 10. Convergence of the algorithm of linearization and either initialized or non initialized Newton's method.

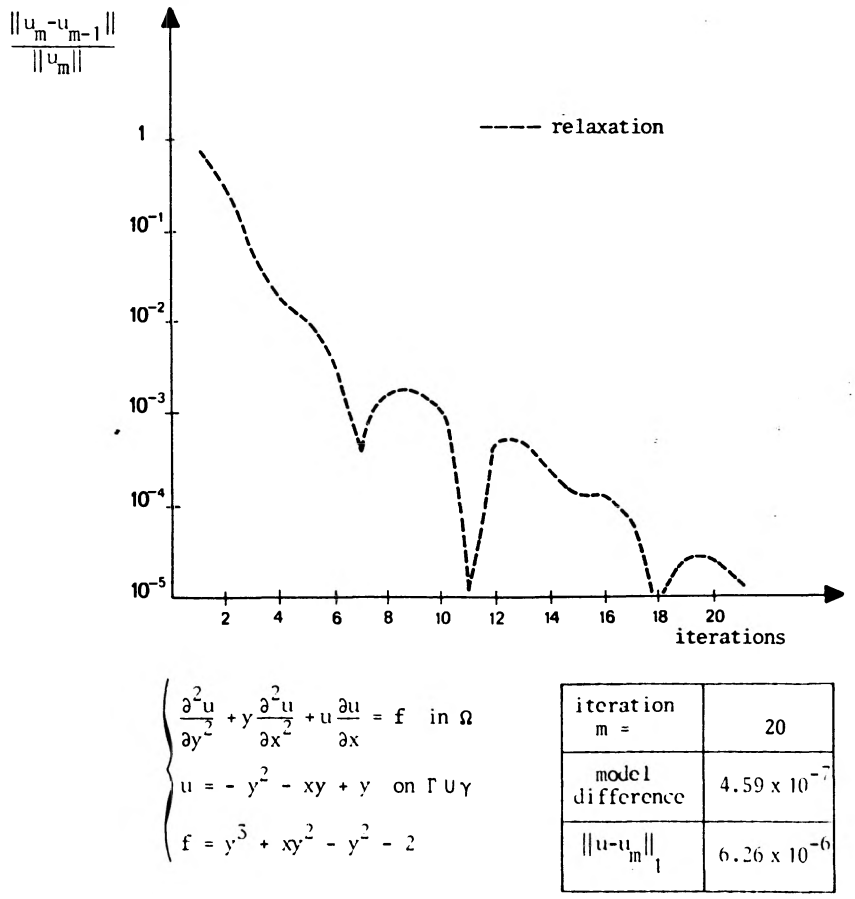
$$\begin{cases} u_0 \in V \text{ given}, \\ u_{m+1} \in V \text{ is the solution of the linear problem} \\ \int_\Omega \left(Au_{m+1} + \tfrac{1}{2}u_{m+1}\frac{\partial u_m}{\partial x} + \tfrac{1}{2}u_m\frac{\partial u_{m+1}}{\partial x} - f\right)\left(Av + v\frac{\partial u_m}{\partial x} + u_m\frac{\partial v}{\partial x}\right)dx\,dy = 0 \\ \text{for all } v \in V. \end{cases}$$

The results are not good; so we use Newton's method (cf. [15]) applied to the functional

$$J(u) = \tfrac{1}{2}\int_\Omega \left| Au + u\frac{\partial u}{\partial x} - f \right|^2 dx\,dy.$$

The algorithm is:

$$\begin{cases} u_0 \in V \text{ given}. \\ u_{m+1} = u_m - H_m^{-1}G_m. \end{cases}$$



$$\begin{cases} \frac{\partial^2 u}{\partial y^2} + y\frac{\partial^2 u}{\partial x^2} + u\frac{\partial u}{\partial x} = f \quad \text{in } \Omega \\ u = -y^2 - xy + y \quad \text{on } \Gamma \cup \gamma \\ f = y^3 + xy^2 - y^2 - 2 \end{cases}$$

iteration m =	20
model difference	4.59×10^{-7}
$\|u-u_m\|_1$	6.26×10^{-6}

Fig. 11. Convergence of the initialized Newton's method for a solution $u \in V_h$ and Tricomi's operator.

Here $G_m = G(u_m)$ and $H_m = H(u_m)$ are the gradient and the Hessian operators. It is sufficient to solve the linear problem:

$$\begin{cases} H_m \omega_m = G_m & \text{in } \Omega, \\ \omega_m = 0 & \text{on } \Gamma \cup \gamma, \end{cases}$$

for which we take an approximate solution obtained after a few iterations of relaxation.

Let A be the linear operator of the model problem and $u \in V_h$. We compare the algorithm of linearization with Newton's method (Fig. 10).

The solution is achieved in fewer iterations with Newton's method initialized with the third iteration of the algorithm of linearization. Moreover the results are as accurate as for the linear problem shown in Fig. 11 where A is Tricomi's operator.

6. Conclusion

The least-square method is perfectly adapted for solving this mixed type problem and the results obtained, in the linear case with the incomplete Cholesky conjugate gradient method, and in the non-linear case with Newton's method, are very good. The disadvantage is that we must solve a higher-order problem than the initial problem. Which, a priori, requires the use of $C^1(\bar{\Omega})$ conforming finite elements. However Zienkiewicz's element has been used without noticeable penalty in the quality of the results.

The patch test can be verified with this element and the numerical examples indicate that this condition does not necessarily have to hold in the whole domain Ω. In contrast, a resolution with Hermite's triangle does not converge to the solution of the problem.

References

[1] C.S. Morawetz, A uniqueness theorem for Frankl's problem, Comm. Pure Appl. Math. 7 (1954) 697–703.
[2] A.K. Aziz and S.H. Leventhal, Numerical solution of linear partial differential equations of elliptic–hyperbolic type, Numerical Solution of Partial Differential Equations III SYNSPADE (1975) 55–87.
[3] A.G. Deacon and S. Osher, A finite element method for a boundary value problem of mixed type, SIAM J. Numer. Anal. 16 (1979) 756–778.
[4] B. Engquist and S. Osher, Stable and entropy satisfying approximations for transonic flow calculations, Math. Comp. 34 (1980) 45–75.
[5] A. Jameson, Numerical solution of nonlinear partial differential equations of mixed type, Numerical Solution of Partial Differential Equations III SYNSPADE (1975) 275–307.
[6] K.O. Friedrichs, Symmetric positive differential equations, Comm. Pure Appl. Math. 11 (1958) 333–418.
[7] P.D. Lax and R.S. Phillips, Local boundary conditions for dissipative symmetric linear differential operators, Comm. Pure Appl. Math. 13 (1960) 427–455.
[8] R.S. Phillips and L. Sarason, Singular symmetric positive first order differential operators, J. Math. Mech. 15(2) (1966) 235–271.
[9] Ph. G. Ciarlet, Conforming and non conforming finite element methods for solving the plate problem, Conference on the Numerical Solution of Differential Equations, Dundee, 1973, Lecture Notes in Mathematics 363 (1973) 21–31.

[10] M. Crouzeix and P.A. Raviart, Conforming and non conforming finite element methods for solving the stationary Stokes equations, RAIRO Anal. Numer. R3 (1973) 33–76.
[11] Ch. Bruneau, Résolution de l'équation de Tricomi par des méthodes variationnelles en éléments finis, Thèse de 3ème cycle, Université de Paris-Sud, Orsay, France, 1980.
[12] P. Lascaux and P. Lesaint, Some non conforming finite elements for solving the plate bending problem, RAIRO Inform. Theor. 9 (1975) 9–54.
[13] M.R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, Nat. Bur. Standards J. Res. 49 (1952) 409–436.
[14] D.S. Kershaw, The incomplete Cholesky conjugate gradient method for the iterative solutions of systems of linear equations, J. Comp. Phys. 26 (1978) 43–65.
[15] J. Cea, Optimisation. Thèorie et Algorithmes (Dunod, Paris, 1971).

FINITE ELEMENT LEAST SQUARE METHODS FOR PROBLEMS OF MIXED-TYPE†

C.H. BRUNEAU*, J.J. CHATTOT**, J. LAMINIE***, J. GUIU-ROUX***

*Office National d'Etudes et de Recherches Aérospatiales (O.N.E.R.A.)
et Université de Paris-Sud.
**MATRA.
***Université de Paris-Sud.
†Work performed with financial support of DRET.

INTRODUCTION.

Two methods are currently used to solve mixed-type equations in relation to steady transonic flow computation : the time dependant method and the steady relaxation method.

The time dependant method classically embeds the mixed problem into a hyperbolic evolution problem and the steady solution is obtained as asymptotic limit for large time with time independant boundary conditions [1]. However the convergence is rather slow due to the time integration schemes available and the very weak damping associated to the hyperbolic system.

The steady relaxation method is based on a potential flow formulation and deals with a mixed second-order partial differential equation. A mixed scheme is used which introduces an upwind bias in supersonic regions. Very efficient algorithms have been devised to accelerate convergence [2].

An alternate approach is proposed to solve by relaxation mixed type problems which can be applied to first order as well as higher order systems. It is based on a least square formulation which transforms hyperbolic problems into problems of elliptic type, thus eliminating the need for mixed schemes and allowing use of accurate type-independant centered schemes as shown in linear problems. This methodology has also been employed to solve a non-linear unsteady problem of wave propagation [3]. The least square method is applied, for illustration, to four steady flow problems : a linear hyperbolic equation of convection, Tricomi's equation of mixed-type and two non-linear first order systems modelling transonic flows with isentropic and non-isentropic shocks. The discretization is achieved via a finite element approximation which is discussed in relation to conservation properties.

1. LINEAR CONVECTION PROBLEM.

The problem is to find the density ρ such that :

$$(1.1)\quad \begin{cases} \operatorname{div} \rho\vec{q} = 0 & \text{in } \Omega \\ \rho\vec{q}.\vec{n} = g & \text{on } \Gamma_1 \\ < \rho\vec{q}.\vec{n} > = 0 & \text{on } \Sigma_j \text{ ("Rankine-Hugoniot" jump condition)} \end{cases}$$

where Ω is a plane nozzle (fig. 1).

$\vec{q} = (u, v)$, $|\vec{q}| \geq \alpha \geq 0$, is the given velocity field. $\vec{q}$ is piecewise continuous except on a finite number of curves Σ_j simulating shock lines. The stream lines do not cross each other and are not tangent to the Σ_j.

$$\Gamma_1 = \{(x, y) \in \partial\Omega \ ; \ \vec{q}.\vec{n} < 0 \} \quad \text{(upstream boundary)}$$

Prop : Under the above hypotheses (1.1) has a unique solution ρ which is piecewise continuous.

1.1. Least square formulation.

(1.2) { Find $\rho \in V_1$ such that $I(\rho) = \inf_{\tau \in V_1} I(\tau)$

where $I(\tau) = \frac{1}{2} \int_\Omega (\text{div } \tau\vec{q})^2 \, dxdy$ and the space V_1 is defined as follows :

$$H_{div}(\Omega) = \{ \vec{r} \in (L^2(\Omega))^2 \; ; \; \text{div } \vec{r} \in L^2(\Omega) \}$$

$$V_{\vec{q}} = \{ \tau \in L^2(\Omega) \; ; \; \tau\vec{q} \in H_{div}(\Omega) \}$$

with norm $||\tau||^2_{V_{\vec{q}}} = |\tau|^2_{L^2(\Omega)} + |\text{div } \tau\vec{q}|^2_{L^2(\Omega)}$

$$V_1 = \{\tau \in V_{\vec{q}}(\Omega) \; ; \; \tau\vec{q}.\vec{n} = g \text{ on } \Gamma_1\}$$

We solve (1.2) by means of the variational formulation

(1.3) $\begin{cases} \text{Find } \rho \in V_1 \text{ such that} \\ \int_\Omega \text{div } \rho\vec{q}.\text{div } \tau\vec{q} \, dxdy = 0 \; ; \text{ for } \tau \in V_{\vec{q}} \text{ such that } \tau\vec{q}.\vec{n} = 0 \text{ on } \Gamma_1 \end{cases}$

Prop. : (see [4] and § 2 below) (1.3) is equivalent to (1.1).

1.2. Discrete approximation.

We denote by $\mathcal{C}_h$ the triangulation of $\Omega_h = \bigcup_{K \in \mathcal{C}_h} K$
(K, Σ_K, P_K) the finite element

$W_h = \{w_h \in C^0(\Omega_h) \; ; \; w_{h|K} \in P_K\}$

$\{\phi_i\}_{i=1,..,N}$ the basis functions.

We choose the square of type 1 (Q_1) as finite element. The space $V_h = W_h \times W_h$ is a conform approximation of $H_{div}(\Omega_h)$. Thus $\tau\vec{q}$ can be approached by a function of V_h but we cannot approach τ by regular functions. This is the main difficulty of the problem. In order to overcome this difficulty we approach the groups of variables ρu, ρv themselves by

$$(\rho u)_h = \sum_{i=1}^{N} (\rho u)_i \, \phi_i \quad \text{and} \quad (\rho v)_h = \sum_{i=1}^{N} (\rho v)_i \, \phi_i$$

Assuming that the nodes are not on the curves Σ_i we choose to write :

$$(\rho u)_i = \rho_i u_i \quad \text{and} \quad (\rho v)_i = \rho_i v_i \quad \text{for } i = 1,\ldots,N$$

This gives a meaning to the discrete problem and satisfies the conservation requirement for discontinuous solutions (cf [4]). More precise justifications will appear in a forthcoming paper.

1.3. Numerical results.

The numerical tests concern a continuous, divergence free velocity field $\vec{q}_1$ and a discontinuous velocity field $\vec{q}_2$. For $\vec{q}_1$ the solution is continuous and the accuracy is very good ($O(h^2)$). For $\vec{q}_2$ defined by

$$\vec{q}_2 = \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \begin{cases} 1 + 0.5\,(1 - x^2) \text{ if } x \geq x_{shock} \; ; \; 1 + 0.5(1 - x) \text{ if } x < x_{shock} \\ 0.4\, xy\, u_2 \end{cases}$$

$x_{shock} = 0.823532 - 0.14\, y^2$.

We obtained the discontinuous solution (fig.2) with also a good accuracy ($O(h^s)$ with $s \simeq 1.5$) and a relative error in mass flux of 10^{-3}.

2. TRICOMI'S PROBLEM.

Tricomi's equation of mixed type is solved in a plane domain Ω of boundary $\partial\Omega = \Gamma U_Y U C$ (fig. 3)

$$(2.1)\quad \begin{cases} Au = f & \text{in } \Omega \\ u = 0 & \text{on } \Gamma U_Y \end{cases} \qquad \left(A = \frac{\partial^2}{\partial y^2} + y\,\frac{\partial^2}{\partial x^2}\right)$$

2.1. Least square formulation.

To solve this mixed type problem we use a total least square formulation and minimize the functional $J(u) = \frac{1}{2}\int_\Omega (Au - f)^2 dxdy$ on the space $V = \mathcal{D}(A) = \{u \in H^2(\Omega)\ ;\ u = 0 \text{ on } \Gamma U_Y\}$

where $H^m(\Omega) = \{u \in L^2(\Omega)\ ;\ D^{|\alpha|}u \in L^2(\Omega) \text{ for } |\alpha| \leq m\}$

We can show easily that this corresponds to solving the following problem :

$$(2.2)\quad \begin{cases} A^*(Au - f) = 0 & \text{in } \Omega \ (A^* \text{ is the adjoint operator of A and } \mathcal{D}(A^*) \\ u \in \mathcal{D}(A) & \text{its domain of definition}). \\ (Au - f) \in \mathcal{D}(A^*) \end{cases}$$

which is equivalent to the initial one (2.1) because A^* is an injective linear operator (cf [5]).

2.2. Discrete approximation.

The problem of minimization is of fourth order ; so a conforming method would require a finite element with numerous degrees of freedom. Thus we prefer a non conforming finite element method and a space V_h such that :

- The functions of P_K are twice continuously differentiable
- $P_2 \subset P_K$
- The patch test is satisfied
- The functions of V_h are continuous on the degrees of freedom of the element K
- $V_h \subset C^0(\overline{\Omega_h}) \cap H^1(\Omega_h)$.

And we choose the Zienkiewicz's triangle (of type 3') which verifies all these assumptions (with regular triangulation for the patch test fig. 4).

2.3. Numerical results and extension.

The solution to the discrete system is obtained with a relaxation algorithm or the Incomplete Cholesky Conjugate Gradient method (ICCG), which greatly improves the convergence as shown fig. 5. The results are accurate and the nodal error is of the same order as the round-off errors in simple precision arithmetic on an IBM 168/370. Moreover $\|u - u_h\|_{H^1(\Omega)} = 0\,(h^s)$ for $h \to 0$ where $1.4 \leq s \leq 1.7$.

A non linear extension of Tricomi's problem is successfully solved by minimizing the functional $J(u) = \frac{1}{2}\int_\Omega (Au + u\,\frac{\partial u}{\partial x} - f)^2 dxdy$ using Newton's method with and without initialization (fig. 6).

3. ISENTROPIC TRANSONIC FLOW.

The full-potential-equation model problem is formulated as a first order system where the density ρ and the velocity field $\vec{q} = (u, v)$ are the unknowns. The equations of motion read :

$$(3.1)\quad \begin{cases} \dfrac{\partial \rho u}{\partial x} + \dfrac{\partial \rho v}{\partial y} = 0 \\ \dfrac{\partial v}{\partial x} - \dfrac{\partial u}{\partial y} = 0 \\ \rho = \left[(\gamma - 1)\, M_\infty^2 \left(h_0 - \dfrac{u^2 + v^2}{2}\right) \right]^{\frac{1}{\gamma - 1}} \end{cases} \qquad \text{in a plane nozzle } \Omega$$

and the boundary conditions are :

$$(3.2)\quad \begin{cases} p = p_{exit} = \dfrac{1}{\gamma M_\infty^2}\,\rho^{\alpha}_{exit} & \text{at the exit section if the flow is subsonic (no condition in the case of supersonic flow} \\ v = 0 \text{ on the nozzle plane of symmetry} \\ \dfrac{v}{u} = f'(x) \text{ on the nozzle wall of equation } y = f(x) \\ \dfrac{v}{u} = g(y) \text{ at the entrance section.} \end{cases}$$

3.1. Least square formulation.

The non-linear problem (3.1) is solved via a fixed point algorithm $\rho_m = (u_m, v_m)$ where u_m, v_m minimize the least-square functional (cf [6]):

$$(3.3)\quad I_m(u,v) = \frac{1}{2}\int_\Omega \left(\frac{\partial \rho_{m-1}u}{\partial x} + \frac{\partial \rho_{m-1}v}{\partial y}\right)^2 dxdy + \frac{1}{2}\int_\Omega \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)^2 dxdy$$

on the space V_0 defined as follows :

$$V_\rho = \{ \vec{r} \in (L^2(\Omega))^2 \ ; \ \rho\vec{r} \in H_{div}(\Omega) \text{ and } \mathrm{rot}\ \vec{r} \in L^2(\Omega)\}$$

$$V_0 = \{ \vec{r} \in V_\rho \text{ verifying the boundary conditions (3.2)}\}$$

Prop. : For ρ_{m-1} given the problem of minimization (3.3) is equivalent to the corresponding first order problem.

3.2. Discrete approximation.

The crucial step is the choice of the finite element for the discrete problem. The insight gained by the numerical solution of a conservation equation (section 1 above and ref [4]) indicates that the groups of variables $(\rho_{m-1}u, \rho_{m-1}v)$ and the variables themselves must be treated as a whole in each conservation law ; e.g. with square of type 1 (Q_1) approximation

$$(\rho_{m-1}u)_h = \sum_{i=1}^{N} (\rho_{m-1}u)_i\, \phi_i = \sum_{i=1}^{N} \rho_{m-1,i}\, u_i\, \phi_i \quad ; \ (\rho_{m-1,i} \text{ are known})$$

$$(u)_h = \sum_{i=1}^{N} u_i\, \phi_i$$

The flux vectors are continuous and this enforces the proper shock conditions. The new iterate is obtained by the following algorithm :

$$\begin{cases} u_m = \mathrm{Inf}\ I_m(u, v_{m-1}) & \text{for } (u, v_{m-1}) \in V_0 \\ v_m = \mathrm{Inf}\ I_m(u_m, v) & \text{for } (u_m, v) \in V_0 \\ \rho_m = \rho(u_m, v_m) \end{cases}$$

3.3. Numerical results.

Each linear system is inverted by the ICCG method. Since each equation corresponds to an elliptic problem, the matrices are very well conditionned and the convergence is achieved in 3 or 4 iterations. Besides the total process converges in about 50 fixed point iterations (see the mesh system fig. 7).

For values of the exit density of 1.4 and 1.355 fully subsonic or subsonic flow with a shock free supersonic bubble establish in the nozzle (fig. 8 and 9). By lowering the exit value of the density to 1.35 the flow chokes and the supersonic zone is terminated by a shock wave which spans the channel (fig. 10). In order to capture the shock wave and eliminate the expansion shock an artificial viscosity equivalent to that used by Jameson [2] is easily implemented in the fixed point algorithm by "upwinding" the density according to :

$$\tilde{\rho} = \rho - \mu\delta\rho \quad \text{where } \mu = \max\{0,\ 1 - \frac{1}{M^2}\}$$

4. FULL STEADY EULER EQUATIONS.

The full steady Euler equations are written in conservative form :

$$(4.1)\quad \begin{cases} \dfrac{\partial \rho u}{\partial x} + \dfrac{\partial \rho v}{\partial y} = 0 \\ \dfrac{\partial \rho u^2 + p}{\partial x} + \dfrac{\partial \rho uv}{\partial y} = 0 \\ \dfrac{\partial \rho uv}{\partial x} + \dfrac{\partial \rho v^2 + p}{\partial y} = 0 \\ \dfrac{\gamma p}{(\gamma - 1)\rho} + \dfrac{u^2 + v^2}{2} = h \end{cases} \qquad \text{in a plane nozzle } \Omega$$

where the unknowns ρ, $\vec{q} = (u, v)$ and p the pressure verify the boundary conditions (3.2) completed by the specification of entropy $s = s_0$ at the entrance section (fig. 11).

4.1. Least square formulation.

As in the previous section the problem (4.1) is solved via a fixed point algorithm $\rho_m = \rho(u_m, v_m, p_m)$ where u_m, v_m, p_m minimize the least-square functional :

$$(4.2)\quad \begin{cases} I_m(u,v,p) = \dfrac{1}{2}\displaystyle\int_\Omega \left(\frac{\partial \rho_{m-1} u}{\partial x} + \frac{\partial \rho_{m-1} v}{\partial y}\right)^2 dxdy \\ + \dfrac{1}{2}\displaystyle\int_\Omega \left(\frac{\partial \rho_{m-1} u^2 + p}{\partial x} + \frac{\partial \rho_{m-1} uv}{\partial y}\right)^2 dxdy + \frac{1}{2}\int_\Omega \left(\frac{\partial \rho_{m-1} uv}{\partial x} + \frac{\partial \rho_{m-1} v^2 + p}{\partial y}\right)^2 dxdy \end{cases}$$

4.2. Discrete approximation.

Here the non-linearity is much more complex and is still present in the minimization of (4.2), which requires the use of Newton's method. Nevertheless the discretization is the same as in section 3 and the groups of variables (ρu, ρv, $\rho u^2 + p$, ...) are approached by continuous functions

The new iterate is obtain by

$$\begin{cases} (u_m, v_m, p_m) = \underset{(u,v,p)}{\text{Inf}}\ I_m(u, v, p) \text{ where } (u_m, v_m, p_m) \text{ is a total vector} \\ \rho_m = \rho(u_m, v_m, p_m) \end{cases}$$

4.3. Numerical results.

This scheme yields very good convergence (fig. 12) and accuracy for subsonic flows and transonic flows with a free supersonic bubble near the throat (fig. 13). Numerical tests give and $O(h^s)$ convergence with s close to 2.

Shock capturing is achieved by adding an upwind term in the fixed point algorithm :

$$\tilde{\rho}_m = \rho_m(1 + \alpha\xi)$$

where ξ is a function of partial derivatives and α a coefficient depending on the mesh size.

This modified scheme gives a provisional result of subsonic-supersonic flow with shock in the divergent part of the nozzle which satisfies the entropy inequality (fig. 14).

CONCLUSION.

Finite element least square methods have been presented to solve four steady mixed-type problems of increasing complexity. The results emphasize the many advantages of this approach in term of - flexibility and accuracy, using a centered scheme regardless of the local type of the

system as opposed to other relaxation methods - rate of convergence of ICCG algorithm as opposed to time dependant methods. Work is in progress to find alternate treatment of the entropy condition.

REFERENCES.

[1]. H. Viviand, "Pseudo-unsteady methods for transonic flow computations", 7th Int. Conf. on Num. Meth. in Fluids Dynamics", W.C. Reynolds and R.W. Mac Cormack editors, Springer-Verlag, Stanford, 1980, p. 44.

[2]. A. Jameson, "Acceleration of transonic potential flow calculations on arbitrary meshes by the multiple grid method", A.I.A.A. 4th Comp. Fluid Dyn. Conf., Williamsburg, 1979, pp. 122-146.

[3]. J.J. Chattot, "Méthode variationnelle pour les problèmes hyperboliques et mixtes du premier ordre", Computing Methods in Applied Sciences & Engineering, R. Glowinski, J.L. Lions editors, North-Holland Publishing Company, 1980.

[4]. J.J. Chattot, J. Guiu-Roux, J. Laminie, "Numerical solution of a first order conservation equation by a least square method", to appear in Int. J. Num. Meth. in Fluids.

[5]. Ch.H. Bruneau, "A non conforming method for solving a least-square formulation of Tricomi's problem", to appear in Computer Methods in Applied Mechanics and Engineering.

[6]. J.J. Chattot, J. Guiu-Roux, J. Laminie, "Finite element calculation of steady transonic flow in nozzles using primitive variables", 7th Int. Conf. on Num. Meth. in Fluids Dynamics, W.C. Reynolds and R.W. Mac Cormack editors, Springer-Verlag, Stanford, 1980, p. 107.

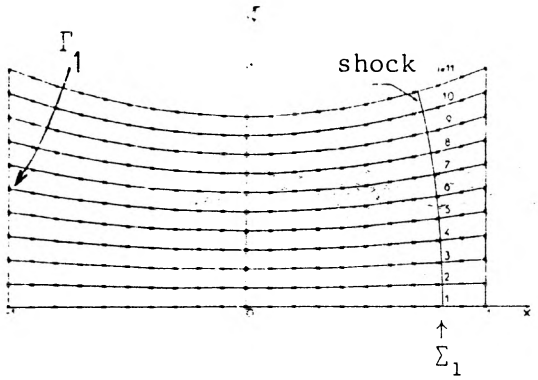


Fig.1. Nozzle geometry and position of the simulated shock.

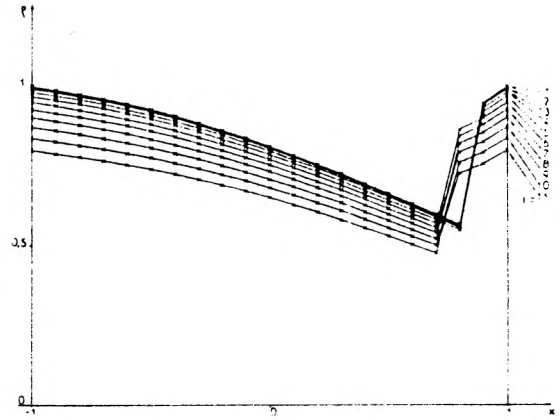


Fig.2. Distribution of density along the stream lines.

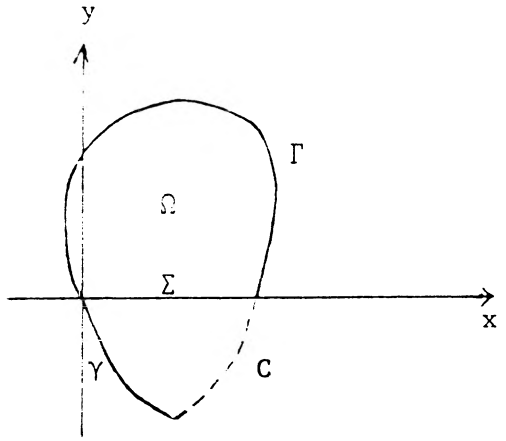


Fig.3. Tricomi's domain.

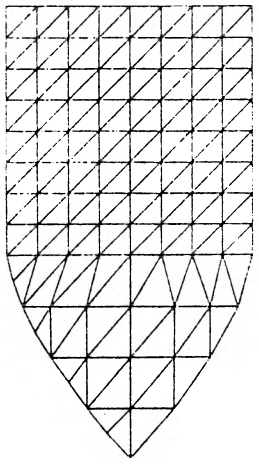


Fig.4. Triangulation.

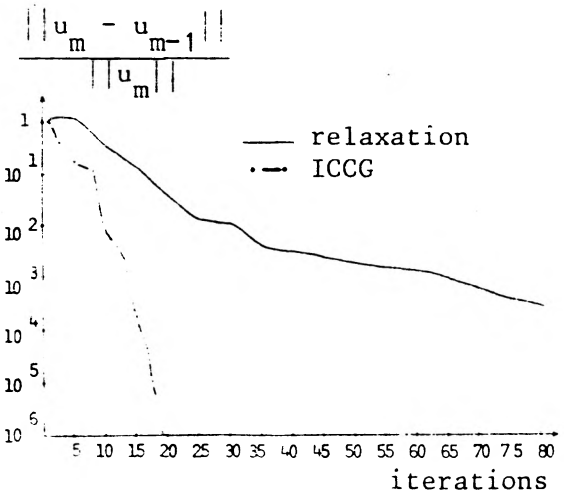


Fig.5. Comparison of relaxation and ICCG convergence history.

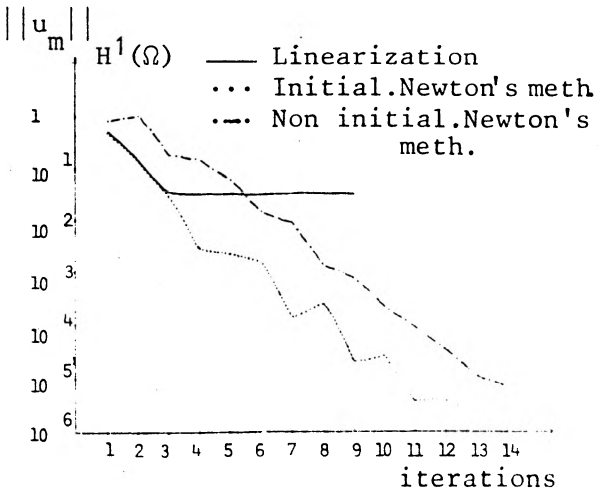


Fig.6. Comparison of convergence history in the non linear case.

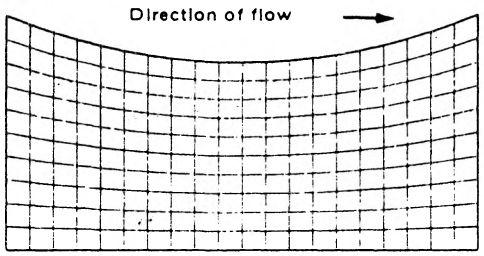


Fig.7. Nozzle geometry and mesh

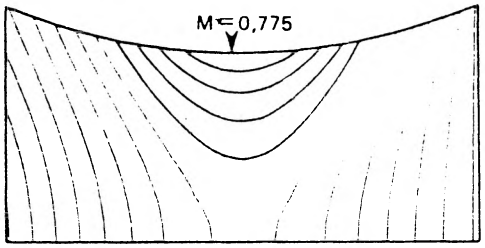


Fig.8. Iso-Mach lines for an exit density = 1.4 ($\Delta M = 0.025$)

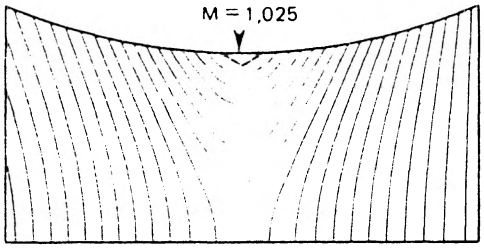


Fig.9. Iso-Mach lines for an exit density = 1.355 ($\Delta M = 0.025$).
-•- sonic line

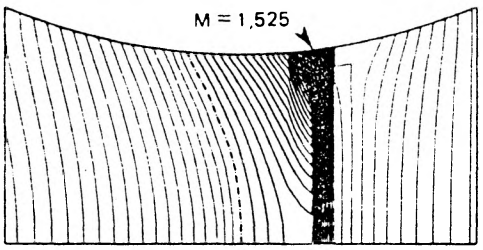


Fig.10. Iso-Mach lines for an exit density = 1.35 ($\Delta M = 0.025$)
-•- sonic line

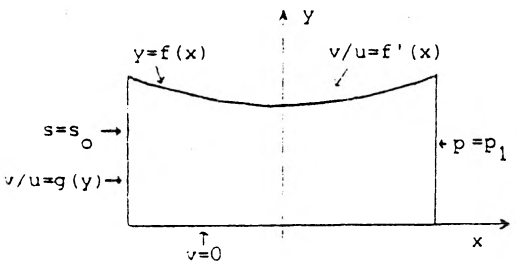


Fig.11. Boundary conditions.

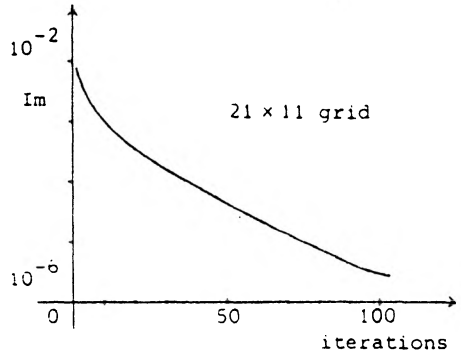


Fig.12. Convergence history ($p_1 = 1.08$).

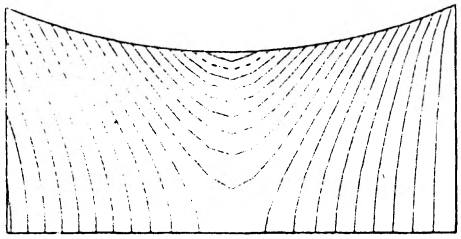


Fig.13. Iso-Mach lines for an exit pression $p_1 = 1.08$ ($\Delta M = 0.025$).
- • - sonic line

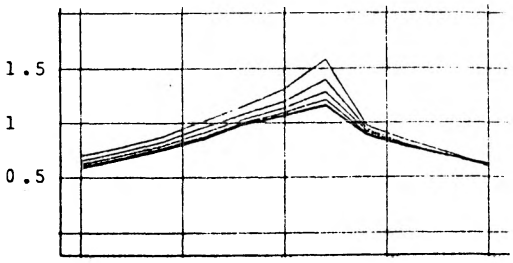


Fig.14. Mach number along the stream lines.

ECOULEMENTS INTERNES.

Dans ce paragraphe nous présentons les premiers résultats obtenus sur le système d'Euler complet (A.0) en dimension deux. Ils concernent des écoulements internes dans une tuyère convergente-divergente et un canal contenant un profil symétrique ([BCLG2]). L'équation de Bernoulli est utilisée pour écrire un algorithme de point fixe sur ρ comme dans le cas du système isentropique ; cependant cela ne suffit pas à linéariser les équations de conservation de la quantité de mouvement. Pour ce faire nous utilisons la méthode de Newton, ensuite le système linéaire du premier ordre est résolu par moindres carrés comme dans le cas précédent.

Nous mettons l'accent dans ce paragraphe sur l'interaction entre l'algorithme de point fixe et la méthode de Newton, et sur la stabilité du point fixe. En effet l'algorithme de point fixe sur la masse volumique n'est stable qu'en subsonique, ce qui va nous conduire à ajouter un terme de viscosité artificielle pour capturer des solutions transsoniques avec choc.

Pour résoudre le problème (A.0) de nombreuses possibilités s'offrent à nous ; on peut les classer en deux catégories : celles mettant en jeu un problème de point fixe et celles conduisant directement à une méthode variationnelle globale. Nous allons développer ici le premier cas (le deuxième sera traité dans le paragraphe suivant). Considérons le problème (A.0) comme un problème de point fixe ; l'équation de Bernoulli nous permet de tirer une des inconnues en fonction des trois autres. En raison des non linéarités intervenant dans les équations et les conditions aux limites (en particulier pour les écoulements internes la pression est fixée à l'aval, ce qui donne une condition de Dirichlet sur p facile à traiter) le choix le plus judicieux est d'écrire un point fixe sur la masse volumique :

$$\rho = Q(u,v,p) = \frac{\gamma p}{(\gamma-1)\left(H - \dfrac{u^2+v^2}{2}\right)} = T(\rho)$$

avec $T = QoP$ où P est la fonction qui fait correspondre à ρ la solution $U = (u,v,p)^T$ des trois premières équations.

Pour résoudre ce point fixe on pourrait par exemple appliquer la méthode de Newton à

$$F(\rho) = T(\rho) - \rho = 0$$

en construisant une suite ρ_n de la façon suivante

$$\rho_0 \text{ donné}$$

$$\rho_{n+1} = \rho_n - \tilde{\rho}_n$$

où $\tilde{\rho}_n$ satisfait $F'(\rho_n) \cdot \tilde{\rho}_n = F(\rho_n)$, et $F'(\rho_n) \cdot \tilde{\rho}_n = Q'(U_n) \cdot \tilde{U}_n - \tilde{\rho}_n$ où U_n est la solution des trois premières équations à ρ_n fixé et $(\tilde{U}_n, \tilde{\rho}_n)$ satisfait le système linéaire :

$$\begin{aligned}
&\frac{\partial}{\partial x}(\rho_n \tilde{u}_n + u_n \tilde{\rho}_n) + \frac{\partial}{\partial y}(\rho_n \tilde{v}_n + v_n \tilde{\rho}_n) = 0 \\
&\frac{\partial}{\partial x}(2\rho_n u_n \tilde{u}_n + u_n^2 \tilde{\rho}_n + \tilde{p}_n) + \frac{\partial}{\partial y}(\rho_n u_n \tilde{v}_n + \rho_n v_n \tilde{u}_n + u_n v_n \tilde{\rho}_n) = 0 \\
&\frac{\partial}{\partial x}(\rho_n u_n \tilde{v}_n + \rho_n v_n \tilde{u}_n + u_n v_n \tilde{\rho}_n) + \frac{\partial}{\partial y}(2\rho_n v_n \tilde{v}_n + v_n^2 \tilde{\rho}_n + \tilde{p}_n) = 0 \\
&Q'(U_n) \cdot \tilde{U}_n - \tilde{\rho}_n = Q(U_n) - \rho_n.
\end{aligned}$$

Ce qui conduit à un algorithme assez lourd ; de plus, à ρ_n fixé, le système des trois premières équations de (A.0) est toujours non linéaire et donc le calcul de U_n est compliqué. Nous préférons à l'algorithme ci-dessus un algorithme standard de relaxation pour résoudre le point fixe et concentrons nos efforts (et nos moyens informatiques) au calcul de U_n.

Soit l'algorithme de point fixe suivant :

$$\begin{aligned}
&\rho_0 \text{ donné} \\
&\rho_n = \frac{\gamma p_n}{(\gamma - 1)\left(H - \dfrac{u_n^2 + v_n^2}{2}\right)}
\end{aligned} \tag{A.6}$$

Il reste à résoudre à chaque itération de (A.6) le système non linéaire du premier ordre pour les variables u, v et p suivant:

$$\begin{aligned}
\frac{\partial \rho_n u}{\partial x} + \frac{\partial \rho_n v}{\partial y} &= 0 \\
\frac{\partial \rho_n u^2 + p}{\partial x} + \frac{\partial \rho_n uv}{\partial y} &= 0 \\
\frac{\partial \rho_n uv}{\partial x} + \frac{\partial \rho_n v^2 + p}{\partial y} &= 0
\end{aligned} \tag{A.7}$$

L'idée est d'utiliser, comme au paragraphe précédent, une méthode de moindres carrés pour calculer la solution $U = (u, v, p)^T$ de (A.7). Cependant, comme (A.7) est non linéaire, il y

a encore un choix à faire : soit écrire une fonctionnelle de moindres carrés puis linéariser, soit linéariser d'abord et appliquer ensuite les moindres carrés. Le premier cas est traité dans [B1] en utilisant la méthode de Newton pour linéariser et les résultats obtenus sont très satisfaisants. Néanmoins, à nouveau par souci de simplicité, nous choisissons la deuxième alternative en utilisant la méthode de Newton pour linéariser le système (A.7). Ceci conduit au système (4) de [BCLG2] qui est ensuite résolu par moindres carrés comme le système (A.2) du paragraphe précédent.

Regardons maintenant la stabilité de l'algorithme de point fixe (A.6) en dimension un d'espace, le système (A.7) s'écrit

$$\begin{aligned} \frac{\partial \rho_n u}{\partial x} = 0 \quad &\text{ou} \quad \rho_n u = A \\ \frac{\partial \rho_n u^2 + p}{\partial x} = 0 \quad &\text{ou} \quad \rho_n u^2 + p = B \end{aligned} \tag{A.8}$$

où A et B sont deux constantes représentant les quantités conservées à travers les chocs par les conditions de Rankine-Hugoniot. De l'équation de Bernoulli on tire :

$$\rho_{n+1} = Q(u,p) = \frac{\gamma p}{(\gamma-1)\left(H - \frac{u^2}{2}\right)} = \frac{2\gamma\rho_n(B\rho_n - A^2)}{(\gamma-1)(2H\rho_n^2 - A^2)}$$

en remplaçant u et p par leur valeur tirée de (A.8) ; d'où il vient

$$\begin{aligned} |\,\rho_{n+1} - \rho_n\,| &= \left| \frac{2\gamma\rho_n(B\rho_n - A^2)}{(\gamma-1)(2H\rho_n^2 - A^2)} - \frac{2\gamma\rho_{n-1}(B\rho_{n-1} - A^2)}{(\gamma-1)(2H\rho_{n-1}^2 - A^2)} \right| \\ &= \left| \frac{2\gamma\rho_n(B\rho_n - A^2)(2H\rho_{n-1}^2 - A^2) - 2\gamma\rho_{n-1}(B\rho_{n-1} - A^2)(2H\rho_n^2 - A^2)}{(\gamma-1)(2H\rho_n^2 - A^2)(2H\rho_{n-1}^2 - A^2)} \right| \\ &= \left| \frac{2\gamma(\rho_n - \rho_{n-1})[A^4 + 2HA^2\rho_{n-1}\rho_n - A^2B(\rho_{n-1} + \rho_n)}{(\gamma-1)(2H\rho_n^2 - A^2)(2H\rho_{n-1}^2 - A^2)} \right| \\ &= \frac{2\gamma A^2[A^2 + 2H\rho_{n-1}\rho_n - B(\rho_{n-1} + \rho_n)]}{(\gamma-1)(2H\rho_n^2 - A^2)(2H\rho_{n-1}^2 - A^2)} \,|\,\rho_n - \rho_{n-1}\,| \end{aligned}$$

où pour $0 < \rho \leq \rho_{\text{arrêt}}$ la fonction $C(\rho) = \frac{2\gamma A^2[A^2+2H\rho^2-2B\rho]}{(\gamma-1)(2H\rho^2-A^2)^2}$ est une fonction positive décroissante de ρ qui vaut 1 pour l'écoulement sonique ($\rho = 1$, $u = 1$, $p = \frac{1}{\gamma}$, $A = 1$ et $B = 1 + \frac{1}{\gamma}$); en effet, $C(1) = \frac{2\gamma\left[1+2H-2\left(1+\frac{1}{\gamma}\right)\right]}{(\gamma-1)(2H-1)^2} = 1$ car $\gamma = 1.4$ et $H = 3$. Ceci montre

que (A.6) est contractant quand $\rho > 1$; en d'autres termes, l'algorithme de point fixe est stable uniquement quand l'écoulement est subsonique. Ceci est vérifié numériquement car le processus global converge facilement pour des écoulements subsoniques, il converge plus lentement dans le cas d'écoulements supercritiques (avec une petite région où l'écoulement devient sonique ou légèrement supersonique) et diverge brutalement en transsonique.

Pour pallier cet inconvénient majeur nous sommes amenés à introduire un terme de viscosité artificielle qui stabilise l'algorithme en transsonique et permet de traiter des écoulements internes avec choc. Les recettes sont nombreuses et plus ou moins couronnées de succès ; notre formulation est particulièrement adaptée au décentrement de la masse volumique qui consiste à introduire :

$$\bar{\rho}_n = \rho_n - \epsilon(\Delta x, u, p)\frac{\partial f(\rho, u, p)}{\partial x} \qquad (A.9)$$

où $\epsilon(\Delta x, u, p)$ est une fonction positive qui dépend de l'écoulement et tend vers zéro quand le pas de discrétisation Δx tend vers zéro (dans la pratique $\epsilon = 0$ dans les régions où l'écoulement est subsonique) et $f(\rho, u, p)$ est généralement une fonction linéaire de l'écoulement. Ensuite on résout (A.8) en remplaçant ρ_n par $\bar{\rho}_n$, ce qui introduit clairement des termes de diffusion artificiels dans les équations. Notons que ce procédé est très courant puisque tous les schémas décentrés en différences finies sont basés là-dessus, en particulier les schémas TVD ("Total Variation Diminishing") ; nous y reviendrons dans la partie B. Dans [BCLG2] est donnée une formule explicite de (A.9) qui a permis d'obtenir des solutions satisfaisantes pour des écoulements transsoniques avec choc dans une tuyère convergente-divergente et un canal contenant un profil en dimension deux d'espace. Néanmoins ce procédé de décentrement n'est stable que pour des chocs relativement faibles et semble inadapté à des écoulements fortement supersoniques. Nous verrons dans le paragraphe suivant comment traiter les chocs forts en hypersonique.

FINITE ELEMENT LEAST SQUARE METHOD FOR SOLVING FULL STEADY EULER EQUATIONS IN A PLANE NOZZLE†

C.H. Bruneau*, J.J. Chattot**, J. Laminie***, J. Guiu-Roux****

* Université de Paris-Sud, Orsay, France, et Office National d'Etudes et de Recherches Aérospatiales (O.N.E.R.A.), Châtillon, France.
** MATRA, Vélizy, France.
*** Université de Paris-Sud, Orsay, et C.N.R.S. E.R.A. 297, Orsay, France.
**** Université de Paris-Sud, Orsay, France.

SUMMARY

A finite element least square method is applied to the steady Euler equations in a nozzle or a channel. For the capture of shock waves an artificial density formula is used. Fast convergence is achieved with I.C.C.G. algorithm.

INTRODUCTION

Progress has been accomplished towards the solution by relaxation of the full Euler equations modelling steady transonic flows of perfect fluid |1-3|. The approach is based on a least square formulation which transforms hyperbolic problems into elliptic type problems, and first order systems into second order systems with positive-definite associated matrix. This insures a proper domain of dependance of the numerical scheme, regardless of the local type of the flow, as well as favourable convergence properties.

In this paper we present the first results obtained with this method concerning the full Euler equations. In the first part, the equations, the boundary conditions and the least square formulation are described. The choice of the minimization variables, the finite element approximation and the fixed-point algorithm on ρ are discussed. For the capture of shock waves an artificial compressibility formula |4| is implemented in the numerical code, and our experience is reported in the second part. The third part is devoted to the presentation of some numerical solutions of transonic flows in nozzles and channels.

I. Variational formulation of the Euler equations.

The conservation of mass, momentum and energy (Bernoulli's equation) in divergence form are written with the usual notations :

$$(1)\qquad \begin{cases} \dfrac{\partial \rho u}{\partial x} + \dfrac{\partial \rho v}{\partial y} = 0 \\ \dfrac{\partial \rho u^2 + p}{\partial x} + \dfrac{\partial \rho uv}{\partial y} = 0 \\ \dfrac{\partial \rho uv}{\partial x} + \dfrac{\partial \rho v^2 + p}{\partial y} = 0 \end{cases}$$

$$(2)\qquad H = \frac{\gamma}{(\gamma-1)}\frac{p}{\rho} + \frac{u^2+v^2}{2} \quad ; \quad (\gamma=1.4)$$

This system is closed by the following boundary conditions : the entropy ($s = c_v \operatorname{Log} \frac{\gamma p}{\rho^\gamma}$) and the flow direction are given in the inlet section, the tangency condition is imposed on the walls and the pressure is specified downstream (subsonic

† Work performed with financial support of D.R.E.T.

exit condition assumed) (figure 1). In the system (1) and (2) of three P.D.E. and one algebraïc equation the four unknowns are ρ, u, v, p. It was chosen to get the solution of this system through a fixed point algorithm (Eq.(2)) and to solve (1) by a minimization method.

u, v and p are the minimization variables. This choice is based on the following arguments :

i) the minimization with respect to p yields a well-behaved elliptic equation regardless of the local flow type ;
ii) the pressure is specified at the exit of the nozzle or channel which gives a Dirichlet boundary condition for pressure ;
iii) ρ is computed from Bernoulli's equation and the formulation (u,v,p) implies a fixed-point algorithm on ρ which can be conveniently modified to include an artificial density formula ;
iv) this formulation allows computation of low speed flows (incompressible flows) and stagnation points are not singular points as in the unsteady Euler equations.

The fixed point algorithm reads :

$$(3)\qquad \begin{cases} \rho_1 \text{ given} \\ \rho_{n+1} = \dfrac{\gamma p_n}{(\gamma-1)\left(H - \dfrac{u_n^2 + v_n^2}{2}\right)} \end{cases}$$

where $Q_n = (u_n, v_n, p_n)$ is solution for fixed ρ_n of equations (1) linearized by Newton's method as :

$$Q_{n-1+j/q} = Q_{n-1+(j-1)/q} - \tilde{Q}_{j/q} \quad \text{where } j = 1,\ldots,q$$

and $\tilde{Q}_{j/q} = (\tilde{u}_{j/q}, \tilde{v}_{j/q}, \tilde{p}_{j/q})$ solution of

(we note $r_j = n-1+(j-1)/q$) :

$$(4)\qquad \begin{cases} A = \dfrac{\partial}{\partial x}(\rho_n \tilde{u}_{j/q}) + \dfrac{\partial}{\partial y}(\rho_n \tilde{v}_{j/q}) - \dfrac{\partial}{\partial x}(\rho_n u_{r_j}) - \dfrac{\partial}{\partial y}(\rho_n v_{r_j}) = 0 \\ B = \dfrac{\partial}{\partial x}(2\,\rho_n u_{r_j} \tilde{u}_{j/q} + \tilde{p}_{j/q}) + \dfrac{\partial}{\partial y}(\rho_n v_{r_j} \tilde{u}_{j/q} + \rho_n u_{r_j} \tilde{v}_{j/q}) \\ \qquad - \dfrac{\partial}{\partial x}(\rho_n u_{r_j}^2 + p_{r_j}) - \dfrac{\partial}{\partial y}(\rho_n u_{r_j} v_{r_j}) = 0 \\ C = \dfrac{\partial}{\partial x}(\rho_n v_{r_j} \tilde{u}_{j/q} + \rho_n u_{r_j} \tilde{v}_{j/q}) + \dfrac{\partial}{\partial y}(2\,\rho_n v_{r_j} \tilde{v}_{j/q} + \tilde{p}_{j/q}) \\ \qquad - \dfrac{\partial}{\partial x}(\rho_n u_{r_j} v_{r_j}) - \dfrac{\partial}{\partial y}(\rho_n v_{r_j}^2 + p_{r_j}) = 0 \end{cases}$$

together with homogeneous boundary conditions.

At each step j the boundary conditions for Q_{r_j} are updated according to :

. Entrance section $\quad u_{r_j} = \sqrt{\frac{2}{1+g^2}\left[H - \frac{1}{(\gamma-1)}(\gamma p_{r_j})^{(\gamma-1)/\gamma}\right]}$; (s=0)

$$v_{r_j} = g\, u_{r_j} \quad ; \text{(g is the streamline slope)}$$

. Wall condition $\quad v_{r_j} = f' u_{r_j}$; (f' is the wall slope)

. Exit section $\quad p_{r_j} = p_{exit}$.

The system (4) is not readily solvable (first order equations) and is thus embedded into least square formulation which yields a symmetric positive definite matrix with better conditioning.

The following quadratic functional

$$J_{\rho_n, r_j}(\tilde{Q}_{j/q}) = \frac{1}{2}\int_{\Omega} (A^2 + B^2 + C^2)\,dxdy$$

is minimized with respect to its 3N variables (N is the number of nodes). This is done with the use of the incomplete Cholesky conjugate gradient method (I.C.C.G.).

The discretization is performed by linear quadrilateral finite element approximation dealing with the groups of variables $(\rho u, \rho v, \rho u^2+p, \ldots)$ to enforce the conservation properties |3| ; for instance $(\rho u)_h = \sum_{i=1}^{N} (\rho u)_i \phi_i = \sum_{i=1}^{N} \rho_i u_i \phi_i$.

II. Artificial compressibility.

The fixed point algorithm on ρ is stable for subsonic flows ; however for supersonic flows an artificial viscosity term must be introduced. In a finite element formulation the most convenient way consists in adding an upwind term to the density |4| . ρ_n is replaced everywhere in (4) by $\overline{\rho_n}$ according to :

$$(5) \qquad \overline{\rho_{n,i}} = \rho_{n,i} - \alpha \mu_i \Delta s_i \frac{\partial \rho_{n,i}}{\partial s_i}$$

where i denotes the index along the streamlines,

α is a coefficient,

$$\mu_i = \text{Max}\ \frac{1}{M_c^2}\left(0, M_c^2 - \frac{M_c^2}{M_i^2}\right)$$

s_i is the curvilinear abscissa at point i,

(in numerical experiments the critical Mach number $M_c = 1$).

With $\alpha = 1$, this method does not allow the shock to move correctly and the flow overshoots before the shock.

In order to decrease the speed of the flow before the shock we take $1 < \alpha < M_i^2$ so that $\alpha \mu_i \leq 1$. Besides this, we use a method similar to those proposed by Habashi and Hafez |4| to solve the potential equation ($\mu_i = \text{Max}(\mu_{i-1}, \mu_i)$). Here we define μ_i from a mean Mach number value $M_i^2 = \frac{1}{2}(M_{i-1}^2 + M_{i+1}^2)$. In this way we add

an upwind term to the density of these subsonic points for which the new M_i^2 is greater than 1. In this case the formula (5) is applied with $\alpha \leq 1$.

III. Numerical results.

The numerical tests concern either a convergent divergent nozzle with parabolic walls of ratio 0.9 or a parallel channel having a 4.2 % thick circular arc bump on the upper wall. This second problem is a test from the GAMM workshop |5|.

The calculation of the flow in the nozzle have been done for several values of the exit pressure. We get either a subsonic or slightly supercritical shockless flow, or a chocked flow with a shock wave. Figures 2 and 3 show this latter case for an exit pressure $p_{exit} = 0.983$. On figure 4 we can see ρ and $\bar{\rho}$ on the upper wall ; the difference between these two quantities gives an estimation of the error introduced by the upwind term. This difference increases with the intensity of the shock when we lower the exit pressure. It seems that the method of artificial compressibility is not well adapted for strong shock for which the density gets close to zero.

For the parallel channel we impose an exit pressure $p_{exit} = 0.843$ corresponding to a Mach number M = 0.85 in isentropic flow |5|. The mesh and the results are shown in figures 5, 6 and 7. Notice the difference in shape of the iso-Mach curves before and after the profile due to the jump of entropy through the shock. The use of a mesh size much cruder than that given for the workshop explain the more upstream position of the shock.

CONCLUSION.

First results for the full steady Euler equations show the interest of the least square formulation and of the finite element discretization for the groups of variables. The solution of the algebraïc systems by the incomplete Choleski conjugate gradient algorithm (I.C.C.G.) yields an implicit scheme having good properties of convergence (figure 8). Further progress is necessary for the capture of strong shocks.

REFERENCES.

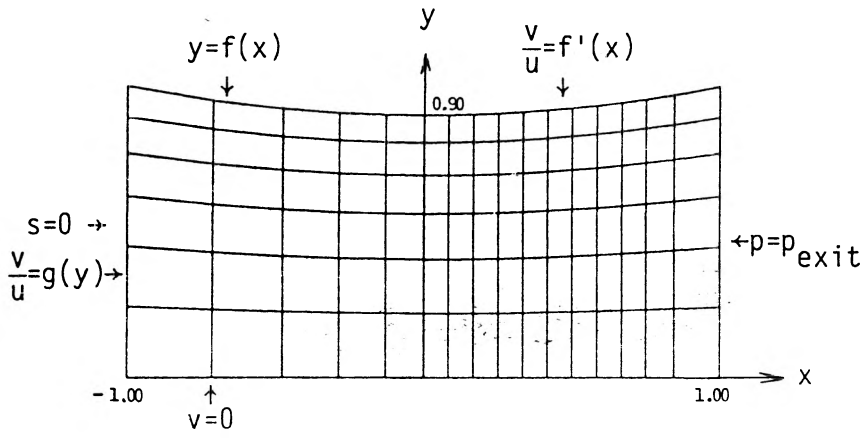
|1| . Chattot,J.J., Guiu-Roux,J., and Laminie, J., Résolution numérique d'une équation de conservation par une approche variationnelle, Proc. 6th ICNMFD, Tbilissi, USSR, June 20-25, 1978. Lecture Notes in Physics, 90, Springer-Verlag, 1979.

|2| . Chattot,J.J., Guiu-Roux,J., and Laminie,J., Finite element calculation of steady transonic flow in nozzles using primary variables, Proc. 7th ICNMFD, Stanford and NASA/Ames, USA, June 23-27, 1980, Lecture Notes in Physics, 141, Springer-Verlag, 1981.

|3| . Chattot,J.J., Guiu-Roux,J., and Laminie,J., Numerical solution of a first-order conservation equation by a least square method, Int. Journal for Num. Meth. in Fluids, vol.2, 209-219, 1982.

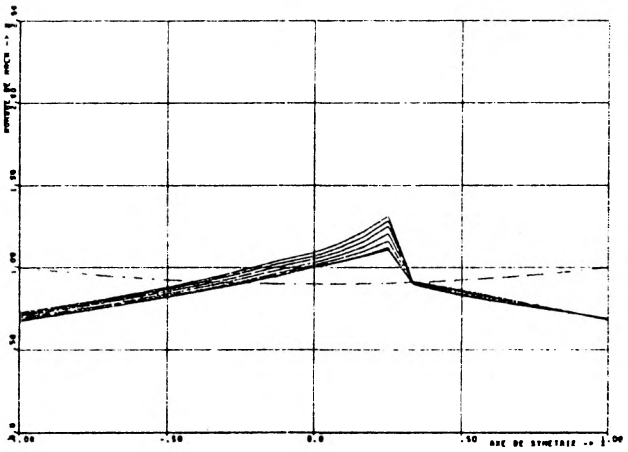
|4| . Habashi,W.G., and Hafez,M.M., Finite element method for transonic cascade flows, AIAA/SAE/ASME 17th Joint Propulsion Conference, Colorado Spring, USA, July 27-29, 1981, AIAA paper n° 81-1472.

|5| . Rizzi,A., and Viviand,H., Numerical methods for the computation of inviscid transonic flows with shock waves, a GAMM Workshop, (Notes on numerical fluid Mechanics, vol. 3), Vieweg, 1981.



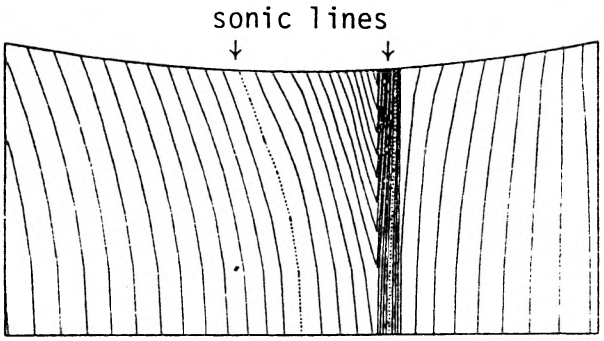
Boundary conditions and mesh for the nozzle

Figure 1



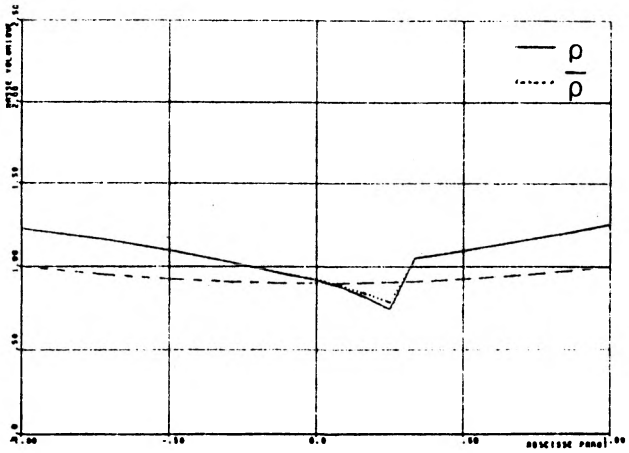
Mach number along the mesh lines

Figure 2



Iso-Mach curves

Figure 3



ρ and $\bar{\rho}$ on the upper wall

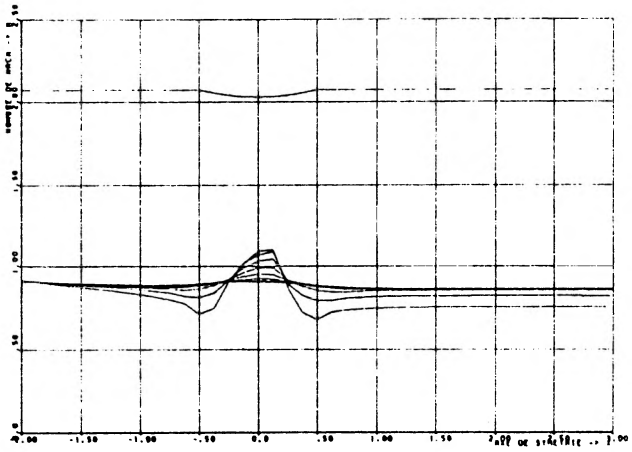
Figure 4

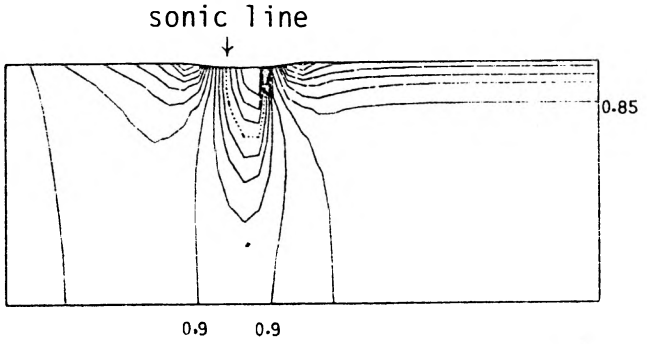
Mesh for the channel

Figure 5

Mach number along the mesh lines

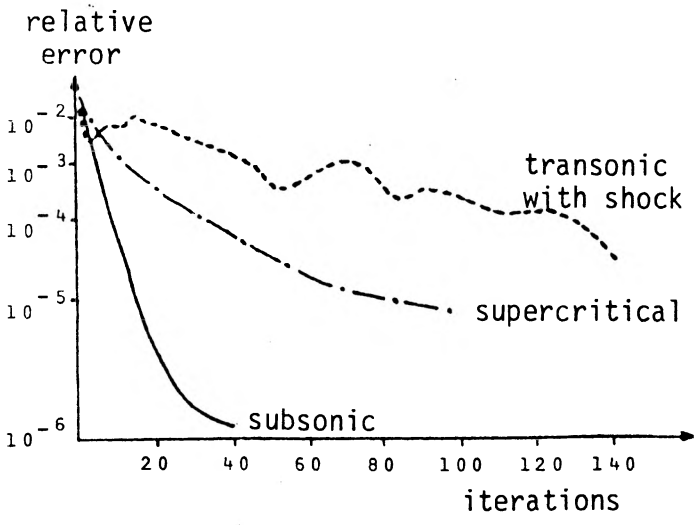
Figure 6





Iso-Mach curves

Figure 7



Convergence curves

Figure 8

ECOULEMENTS EXTERNES

Ce dernier paragraphe sur les équations d'Euler compressibles regroupe diverses études sur des écoulements avec séparation en dimension deux et trois d'espaces qui ont conduit à supprimer l'algorithme de point fixe pour traiter des écoulements hypersoniques.

La première étude concerne le calcul de solutions subsoniques autour d'un cylindre en dimension deux ; il s'agit de montrer que le modèle d'Euler peut capturer des recirculations fermées. Comme l'on ne s'intéresse qu'au cas subsonique la formulation du paragraphe précédent est conservée en totalité. La question est de savoir comment déclencher la recirculation sur une surface régulière ; il faut simuler par exemple un obstacle sur le cyclindre, ce qui est fait en général en imposant la direction (inconnue !) de l'écoulement en quelques points du maillage de façon à avoir un vecteur vitesse orthogonal à la paroi au point de décollement. Nous proposons ici d'imposer uniquement la pression d'arrêt de l'écoulement générateur en ce point ; ce qui est beaucoup plus satisfaisant d'un point de vue physique. Les résultats présentés dans [BCLT1] montrent que cela suffit à capturer une zone de recirculation dont la taille dépend de la position du point. D'autre part, en faisant varier la pression imposée au point de décollement on peut simuler un jet à travers un trou de la paroi et obtenir des zones de recirculation beaucoup plus importantes. Ceci montre la précision de la méthode qui est sensible à la moindre variation des conditions aux limites.

Forts de cette expérience nous appliquons ensuite la méthode à l'écoulement autour d'une plaque sans épaisseur en dimension trois ([BCLT1]) ; là le décollement est provoqué par le bord effilé de la plaque et la méthode capture directement l'enroulement tourbillonnaire en bout de plaque sans avoir à rajouter de viscosité artificielle ni à spécifier de condition supplémentaire. La seule limitation est la non stabilité de l'algorithme de point fixe en supersonique, autrement il semble que la méthode numérique soit capable de restituer tout ce que contient le modèle d'Euler.

Les dernières études présentées ici concernent l'extension de la méthode à des écoulements supersoniques et hypersoniques autour de plaques sans épaisseur en dimension trois et de corps émoussés en dimension deux. Dans le premier cas l'équation de Bernoulli est remplacée par l'équation aux dérivées partielles de la conservation de l'énergie et l'algorithme

de point fixe est supprimé au profit d'une linéarisation globale par la méthode de Newton. Dans [BLC1] et [BLC2] est étudié en détail l'ensemble de la méthode et sont présentés les résultats autour d'une plaque rectangulaire et d'une aile delta sans épaisseur. Les résultats coïncident avec ceux déjà obtenus en subsonique mais couvrent cette fois un éventail assez large d'écoulements allant du subsonique au fort supersonique.

Dans le deuxième cas l'équation de Bernoulli est conservée pour remplacer la pression dans les équations de la conservation de la quantité de mouvement par :

$$p = \frac{\gamma - 1}{\gamma} \rho \left(H - \frac{u^2 + v^2}{2} \right)$$

On garde ainsi un système de trois équations (en dimension deux) qui est traité comme précédemment. La difficulté est la capture et la représentation du choc détaché devant le corps en hypersonique ; en effet il est impossible de deviner a priori la forme du choc et où il se situe, donc à la fois le maillage et la solution initiale sont assez loin du résultat final. Or, en hypersonique, la pression à l'infini est très petite, et si l'on initialise par exemple par les conditions à l'infini partout, on obtient une solution non physique pour laquelle la pression est négative en certains points. Il peut même y avoir convergence sur une grille grossière dans certains cas mais cela diverge ensuite sur des grilles plus fines. Ce problème se rencontre avec d'autres approximations, en particulier pour certains schémas aux différences finies avec limiteur de flux. Ici, nous l'éliminons en incorporant au système d'Euler l'inéquation d'entropie que l'on rajoute dans la fonctionnelle des moindres carrés ([B2]). Ceci permet de capturer le choc détaché sur une maille quel que soit le nombre de Mach à l'infini, contrairement à la méthode de décentrement artificiel du paragraphe précédent. Enfin une procédure d'adaptation de maillage permet de lisser la discontinuité et d'obtenir un choc en parfait accord avec les résultats théoriques et expérimentaux.

NUMERICAL SOLUTIONS OF THE EULER EQUATIONS WITH SEPARATION BY A FINITE ELEMENT METHOD

C.H. BRUNEAU* , J.J. CHATTOT** , J. LAMINIE* , R. TEMAM*

*Laboratoire d'Analyse Numérique, Bâtiment 425, Université Paris-Sud, 91405 ORSAY, E.R.A. C.N.R.S.

**M.A.T.R.A., 37 Avenue Louis Bréguet, 78140 VELIZY.

SUMMARY.

A finite element least squares method is applied to the steady Euler equations in order to capture separated flows around a cylinder. This is achieved without using artificial viscosity but only by giving a "Kutta condition" at the separation point. The sensitivity of the solution to this condition and mesh size is discussed in this paper.

INTRODUCTION.

In perfect fluid motions, separation occurs in various circumstances and is followed by vortical phenomena which are often of primary importance, for instance in term of global efforts on obstacles.

Separation along the sharp tailing edge of a wing or a rotor blade is a typical example. A vortex sheet trails the lifting surface and the accurate prediction of its location and strength requires a high degree of sophistication [1] .

More critical is the separation occuring at a sharp leading edge since the vortex sheet interacts strongly with the lifting surface, and produces non linearities on the lift characteristic [2].

Separation can also occur on smooth surfaces as in the case of the transonic flow at Mach number M = 0.5 past a cylinder, where the numerical solutions, computed independently by the participants of a workshop held at NASA Langley Research Center on September 1, 1981 [3] exhibit a recirculating bubble. Here separation is triggered by the stagnation pressure losses and the vorticity generated at the recompression shock.

The prediction of these flows is best handled with the Euler equations which describe the general motion of perfect fluid with vorticity. However it is not yet clear as to what kind of boundary condition can be imposed ("Kutta condition") and what accuracy can be expected in the numerical solution.

This paper addresses these last two questions by way of "numerical experiments". The tests concern separated flows past a circular cylinder at low subsonic Mach number.

I. NUMERICAL METHOD .

The steady Euler equations written in conservative form read in two dimension :

$$(1)\qquad \begin{cases} \dfrac{\partial \rho u}{\partial x} + \dfrac{\partial \rho v}{\partial y} = 0 \\ \dfrac{\partial \rho u^2 + p}{\partial x} + \dfrac{\partial \rho uv}{\partial y} = 0 \\ \dfrac{\partial \rho uv}{\partial x} + \dfrac{\partial \rho v^2 + p}{\partial y} = 0 \end{cases}$$

$$(2)\qquad H = \frac{\gamma}{\gamma - 1}\frac{p}{\rho} + \frac{u^2 + v^2}{2} \quad ; \quad (\gamma = 1.4)$$

with the following boundary conditions (Figure 1) :

On Γ_0 : the symmetry condition ($v = 0$),

On Γ_1 : the tangency condition ($\vec{q}.\vec{n} = 0$, $\vec{q} = (u,v)$),

On Γ_2 : the free stream condition.

The method of reference [4] is used :

- Fixed point algorithm on ρ .
- Linearization by Newton's method.
- Least squares embedding.
- Use of finite element discretization for the components of the flux vector.
- Resolution by ICCG method.

Moreover the tangency condition is treated in variational form. The least squares functional is modified by addition of the following term :

$$(3)\qquad \frac{1}{2}\int_{\Gamma_1} \left[\left(\tilde{\vec{q}}_{j/q} - \vec{q}_{r_j}\right).\vec{n}\right]^2 d\Gamma$$

where $\left(\tilde{\vec{q}}_{j/q}, \tilde{\rho}_{j/q}\right)$ is the solution of the linearized equations and $\vec{q}_{r_j}$ corresponds to the previous iterate.

II. "KUTTA CONDITION".

To provoke the separation we need to impose some extra condition to the flow at a point on the cylinder. In a first attempt a zero tangential velocity component is given at the separation point $S = (1,\theta)$. A very small separation bubble occurs within a mesh row. It can be conjectured that there is not a unique solution of the Euler equations for a given separation point and that the numerical viscosity of the scheme plays a major role in the selection of the solution. By construction our scheme is centered and second order accurate and does not require the addition of artificial viscosity in subsonic regions, in contrast to unsteady methods. Thus

the pressure at the separation point is imposed and yields an extra degree of freedom to control the size of the bubble.

III. NUMERICAL RESULTS.

The first tests are relevent to fully attached flow in order to check the new tangency condition (3). The results are summarized in Table 1 and show at least quadratic convergence with mesh refinement.

Three types of tests are performed in the case of separated flow to evaluate the sensitivity of the solution to the pressure at the separation point S, to the mesh size and to the location of S (θ).

For $\theta = 129°$ and mesh of 16×32 points the pressure at point S, when not imposed, is at convergence $p_S = 1.05$, with a very small separation bubble. If we impose $p_S \leqq 1.05$ we find a smaller bubble and even no bubble at all for $p = 0.9$, where the flow is fully attached. On the contrary if we impose $p_S > 1.05$ we find a larger bubble which increases in size with p_S (Figure 2). Even if we take $p_S > 1.35209$ (stagnation pressure of incoming flow) we get larger separation zone (Figure 3). However we note that the stagnation pressure in the separated region (point B) is always less than that of the incoming flow. It is believed that the discrepancy between the separation pressure and the stagnation pressure in the separated region comes from the dissipative effects of the scheme at the separation point where the solution exhibits large gradient.

The sensitivity to the mesh is emphasized on Figures 4 and 6 which correspond to the conditions of Figure 2 and 3 respectively with a finer 31×61 mesh. For a given pressure and separation location the mesh refinement yields a larger bubble. Figure 5 illustrates the vortex when the pressure of the incoming flow is imposed at point S for the finer 31×61 points mesh. When the separation point S is moved upstream the separated region grows as expected (Figures 6 and 7).

CONCLUSION.

The numerical results for the capture of separated flow regions with the Euler equations indicate a great sensitivity of the solution to mesh size and conditions at separation ("Kutta condition"). The proposed scheme does not need an artificial viscosity term explicitely but requires to impose both the location and the pressure of the separation point.

Preliminary tests are carried out in three dimensions to simulate separated flows past a flat plate at incidence.

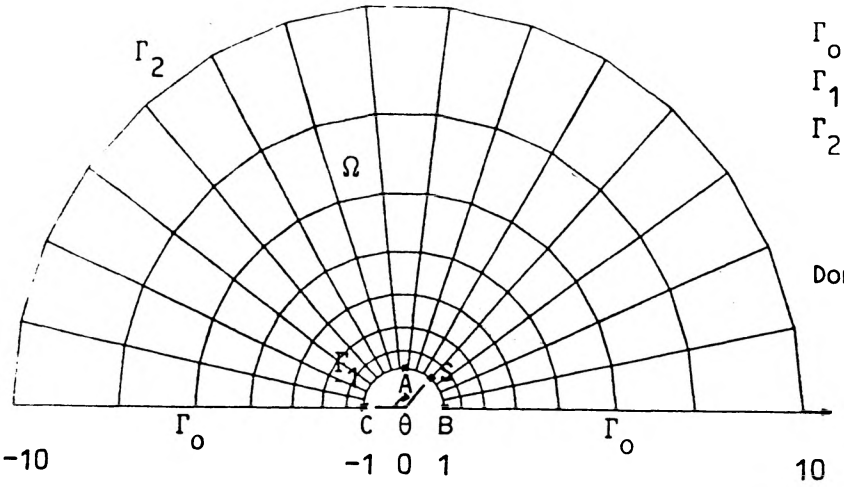
REFERENCES.

[1] J.J. CHATTOT, M. BOSCHIERO and C. KOECK, Méthodes numériques de prédiction de l'aérodynamique des missiles, AGARD-CP-336, feb. 1983.

[2] C. KOECK and J.J. CHATTOT, Computation of three-dimensional Vortex flows past wings using the Euler equations and a multiple-grid scheme, Proceedings 9th Int. Conf. Meth. in Fluid Dyn., June 25-29, 1984, CEN Saclay, France, Springer-Verlag.

[3] M.D. SALAS, Recent developments in transonic Euler flow over a circular cylinder, Proceedings 10th IMACS World Congress, Aug. 8-13, 1982, Montréal, Canada, Vol. 1, p. 175-177.

[4] C.H. BRUNEAU, J.J. CHATTOT, J. LAMINIE and J. GUIU-ROUX, Finite element least square method for solving full steady Euler equations in a plane nozzle. Proceedings 8th Int. Conf. Meth. in Fluid Dyn., June 28-July 2, 1982, Aachen, Germany, Springer-Verlag.



Γ_o : Symmetry condition
Γ_1 : Tangency condition
Γ_2 : Free stream condition

Domain and boundary conditions

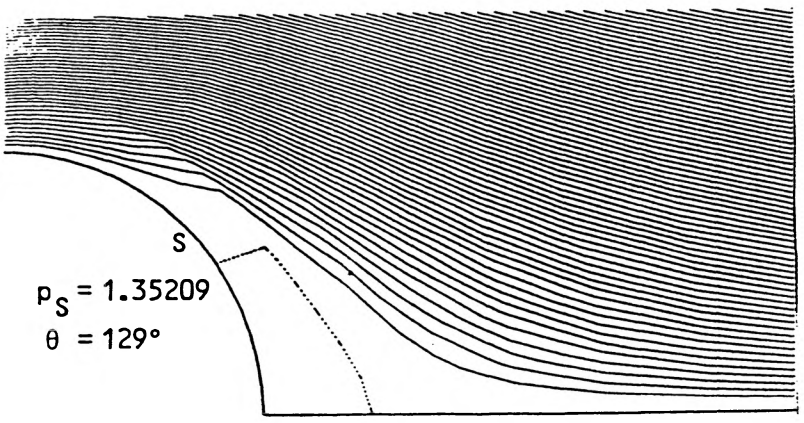
Figure 1

Comparison of accuracy and convergence for different meshs.

Table 1

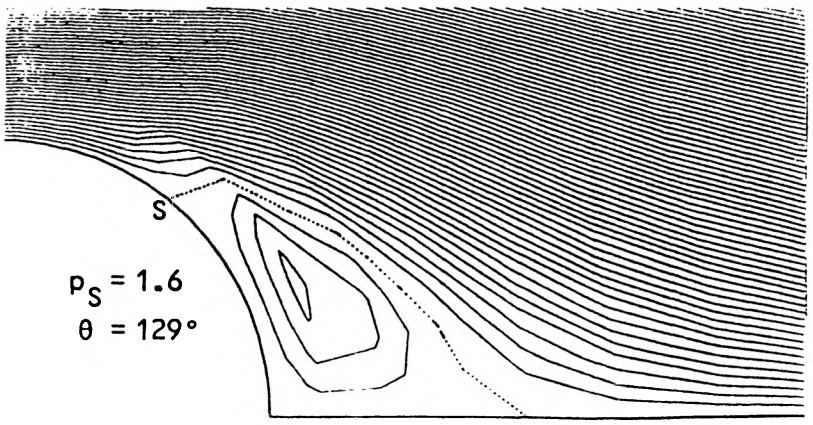
IM x JM points mesh	Velocity at the stagnation points B and C	Pressure at points B and C	Velocity at point A	Variation of entropy at point A	Value of the least-square functional
8x16	0.93×10^{-2}	1.24	1.7	11 %	1.3×10^{-1}
16x32	0.17×10^{-3}	1.32	0.85	4.5 %	2.6×10^{-3}
31x61	0.8×10^{-5}	1.35	0.93	1.7 %	4.0×10^{-4}

Stagnation pressure (1.35209) and velocity (0.425) of the incoming flow (M = 0.4).



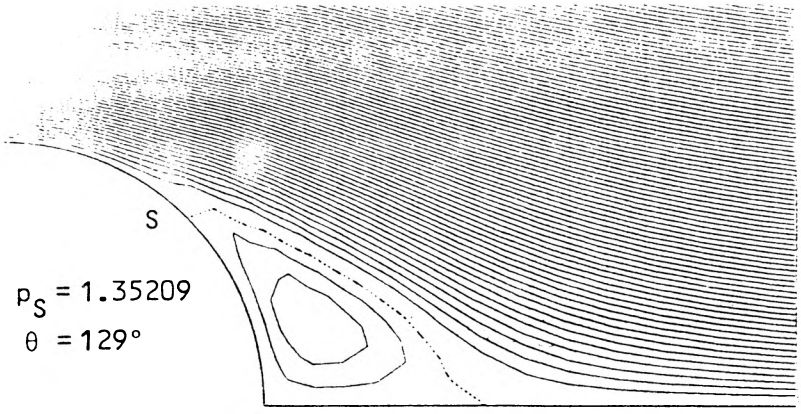
Stream lines for the 16 x 32 points mesh

Figure 2.



Stream lines for the 16 x 32 points mesh

Figure 3

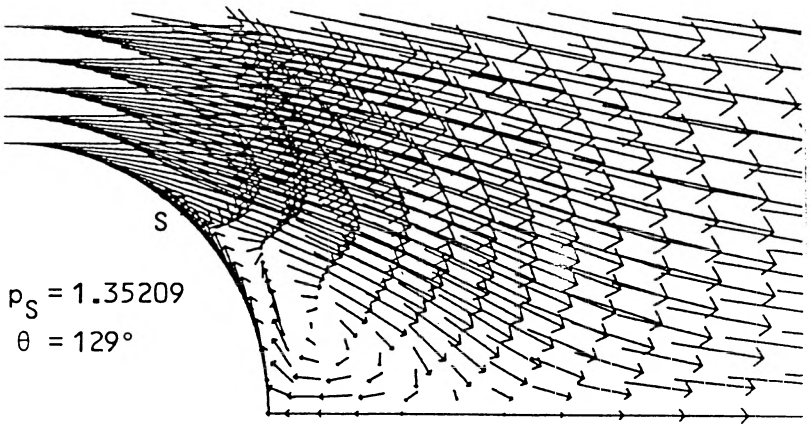


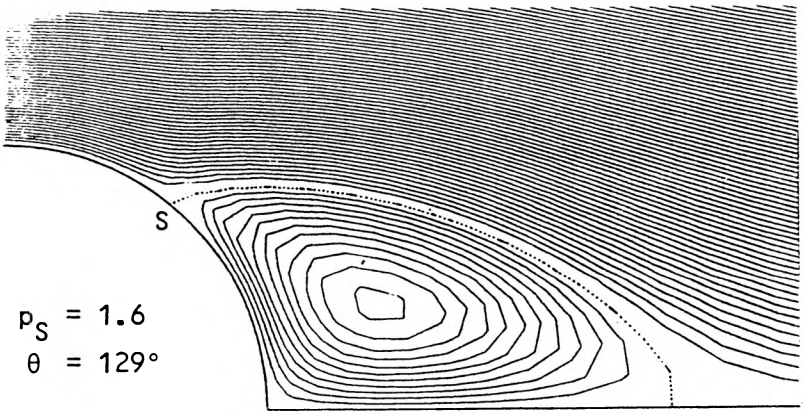
Stream lines for the
31x61 points mesh

Figure 4

Velocity field for the
31x61 points mesh

Figure 5



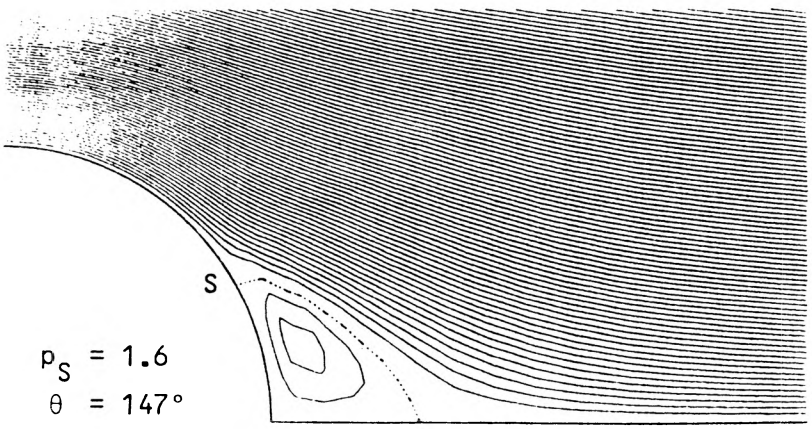


Stream lines for the
31x61 points mesh

Figure 6

Stream lines for the
31x61 points mesh

Figure 7



Computation of Vortex Flows past a Flat Plate at High Angle of Attack. †

C.H. Bruneau*, J.J. Chattot**, J. Laminie*, R. Temam*

*Laboratoire d'Analyse Numérique, C.N.R.S. et Université Paris-Sud, 91405 Orsay (France)

**MATRA, 37, Avenue Louis Bréguet, 78140 Vélizy (France)

(†Work performed with financial support of D.R.E.T. and under grants from the CCVR Ecole Polytechnique, France, and the Minnesota Supercomputer Institute, U.S.A.).

Summary.

The finite element method described in a former paper presented at the 9th ICNMFD ([1]) is applied here to solve the full, steady Euler equations in 3D. The steady flow past a rectangular flat plate of small aspect ratio at high angle of attack is computed for several meshes and angles of attack. These calculations have been made without either adding artificial viscosity or using a Kutta condition. They show how vortices develop spontaneously around the tip of the plate and propagate after the trailing edge.

Introduction.

Vortical phenomena occur in flows around wings in different ways. We are interested in flows around a thin plate of small aspect ratio for high angles of attack. Actually in this application the plate has no thickness and we expect a strong vortex structure to develop at the tip of the plate. So this does not allow the use of the potential equation and needs a more general model like Euler or Navier-Stokes equations.

It is not yet clear what is the best model for the calculation of steady, three-dimensional separated vortex flows. Our aim is to show that the Euler equations are able to represent open separation with large vortices, although they do not contain viscous terms. Numerically an

explicit viscosity term is often added or/and a Kutta-Joukowski condition is specified at sharp edges of a lifting surface. In this paper we present a method that does not require either condition for subsonic flows. The vortices are captured, develop spontaneously at the tip of the plate and are convected past the trailing edge. Particular care must be taken in imposing boundary conditions, especially downstream where the pressure can not be fixed as is usually done for two-dimensional subsonic calculations.

Physical problem.

The aim is to compute the flow around a flat plate at incidence for subsonic mach number M_∞. The plate of small aspect ratio $\mathcal{R} = 0.5$ without thickness is set at an angle α with the incoming flow. The plate is imbedded in an orthogonal domain as shown Figure 1, and the incoming flow is given by : $\vec{q}_\infty = (u_\infty, v_\infty, w_\infty) = (q_\infty \cos\alpha, 0, q_\infty \sin\alpha)$ with $M_\infty^2 = \dfrac{q_\infty^2}{a_\infty^2}$

where $a_\infty^2 = (\gamma - 1)\left(H - \dfrac{q_\infty^2}{2}\right)$. By symmetry we compute only the flow around the right half plate. In addition there is a tangency condition on both sides of the plate.

In this work we take $M_\infty = 0.7$ and a high angle of attack $\alpha = 15°$ or $\alpha = 30°$.

Outline of the method.

We solve the full, steady Euler equations in conservative form :

$$(1)\quad \begin{cases} \dfrac{\partial \rho u}{\partial x} + \dfrac{\partial \rho v}{\partial y} + \dfrac{\partial \rho w}{\partial z} = 0 & \text{Conservation of mass} \\ \dfrac{\partial \rho u^2 + p}{\partial x} + \dfrac{\partial \rho uv}{\partial y} + \dfrac{\partial \rho uw}{\partial z} = 0 & \\ \dfrac{\partial \rho uv}{\partial x} + \dfrac{\partial \rho v^2 + p}{\partial y} + \dfrac{\partial \rho vw}{\partial z} = 0 & \text{Conservation of momentum} \\ \dfrac{\partial \rho uw}{\partial x} + \dfrac{\partial \rho vw}{\partial y} + \dfrac{\partial \rho w^2 + p}{\partial z} = 0 & \end{cases}$$

$$(2)\quad \frac{\gamma p}{(\gamma - 1)\rho} + \frac{q^2}{2} = H \qquad \text{Bernoulli's equation}$$

for ρ, $\vec{q} = (u,v,w)$, p the density, the velocity and the pressure respectively where H is the total enthalpy and γ the ratio of specific heats.

- A fixed point algorithm is based on equation (2), computing the density from u,v,w,p solution of system (1) with given ρ.

- The non-linear system (1) for u,v,w,p (ρ given from the previous iteration of the fixed point algorithm) is transformed into a linear system for $\tilde{u}$, $\tilde{v}$, $\tilde{w}$, $\tilde{p}$ by means of Newton's method (cf. [2]).

- The linear system for $\tilde{u}$, $\tilde{v}$, $\tilde{w}$, $\tilde{p}$ is solved through the minimization of a least-square functional. This transforms the first order system into an equivalent symmetric second order system (cf. [3]) for the same variables.

- The unknowns $\tilde{u}$, $\tilde{v}$, $\tilde{w}$, $\tilde{p}$ and the conservative variables ($\rho\vec{q}$, etc.) are approximated through a Q1 linear finite element method.

Boundary conditions.

Because of the geometry of the plate and the domain the physical boundary conditions are easy to implement. The tangency condition on the plate (plane z = 0) reads w = 0 and the symmetry condition is v = 0 on y = 0.

On the contrary the boundary conditions at the limits of the domain are not so easy especially downstream, where we must allow the vortex to go through the exit plane. In particular we cannot impose a constant value of the pressure. With the unsteady model one generally takes a non reflecting boundary condition or sometimes impose a mean value for the pressure [4]. We have made attempts to extrapolate the pressure or both the pressure and the velocity from inside the domain in the direction of the flow, but we found that global extrapolation in the main flow direction eliminates the vortex downstream.

Finally the flow variables have been set to the freestream values at incoming boundaries ($x = x_0$, $z = z_0$) and no condition was imposed at exit boundaries ($x = x_1$, $z = z_1$), so that the values there are computed from the variational formulation. It corresponds to requiring that the first order equations be satisfied on those boundaries.

At $y = y_1$, we use the far field condition v = 0.

Numerical aspects.

Use is made of the structured nature of the mesh system (i,j,k). Thus the global matrix of the discrete linearized equations has a symmetric block structure. However it is too large to be stored on a CRAY-1S with 1 million words of memory and a relaxation method is used. The degrees of freedom are numbered plane by plane in the x direction so that the corresponding matrix is a block matrix where each block corresponds to a plane i = const. A block Gauss-Seidel iterative method is used and the blocks are inverted by means of an Incomplete Cholesky Conjugate Gradient algorithm.

The plate is contained in a double plane z = 0. On the plate the field variables are double valued, and single valued outside and at the tip of the plate, the first single valued point has 10 neighbors instead of 9 in a plane i = const. (Fig. 2).

At the exit plane, i = IMax, the conditionning of the block matrix is poor and the ICCG method is stabilized upon adding the main diagonal in the Gauss-Seidel algorithm, which can be interpreted as an extra time derivative term.

In order to take advantage of the performance of the CRAY, a particular effort is made to vectorize the matrix assembly and the solution algorithm which account for at least 70 % of the CPU time.

For matrix assembly, the interaction of a degree of freedom with a neighboring one is computed for all the elements of the plane, instead of assembling the elementary matrix corresponding to one element. Thus the innermost vectorizable loop is of length JMax × KMax (320 and 900 in our application).

In a plane, all the degrees of freedom related to $\tilde{p}$, $\tilde{u}$, $\tilde{v}$ and $\tilde{w}$ are numbered successively. Thus the matrix can be subdivised in 4×4 blocs. This allows vectorization of the Cholesky factorization when it operates on different blocks.

Numerical results.

The first case corresponds to an incidence of 15°. Two different mesh systems are used with respectively 28×18×19 ≅ 11000 points (medium mesh) and 50×30×31 ≅ 50000 points (fine mesh). The plate tip is embedded in a uniformly spaced grid and away from the tip, the grid is stretched (Fig. 3). The computing time for 30 fixed point iterations in the medium grid is 1 hour CPU, and for 50 iterations in the fine grid is 15 hours.

The vortex structure develops with a more or less conical growth in x as can be seen on the cross-flow velocity plots at the various locations in x of 70 %, 90 % and 110 % (Fig. 4-6). The last figure corresponds to the exit plane which is crossed by the vortex. The spreading of the vortex structure is strongly dependent on the mesh steps. The fine grid gives a better capture of the vortex roll-up (Fig 7-8).

The second case corresponds to an incidence of 30°. Only the medium mesh is used. The presence of the plate creates a very large perturbation to the flow and near the trailing edge the local mach number exceeds unity. The fixed point algorithm does not converge under supersonic flow condition and the flow variables oscillate with the iterations. The qualitative behaviour of the flow is very similar to the 15° case, but the vortex structure is stronger and lifts off more rapidly behind the plate, according to the angle of incidence (Fig. 9-10).

Conclusion.

The first applications of the finite element least square method in three dimension to the computation of the separated flows past a flat plate at incidence indicate that the vortex structure is well captured by the numerical scheme without the help of artificial viscosity. The quality of the steady solution is related to the total number of mesh points in the discretization, which is limited by the cost of the computation (CPU time). Progress in the solution procedure for the linear system is expected using different techniques available with the present memory size of computer (multigrid or frontal methods), or large extended memory size.

References.

[1] Ch.H.Bruneau, J.J.Chattot, J.Laminie, R.Temam : Numerical solutions of the Euler equations with separation by a finite element method. 9th ICNMFD Saclay 1984, Lecture Notes in Physics n° 218, Springer-Verlag.

[2] Ch.H.Bruneau, J.J.Chattot, J.Laminie, J.Guiu-Roux : Finite element least square method for solving full steady Euler equations in a plane nozzle. 8th ICNMFD Aachen 1982, Lecture Notes in Physics n° 170, Springer-Verlag.

[3] S.H.Chiang, G.M.Johnson : An embedding method for the steady Euler equations. J. Comp. Phys., n° 63, 1986.

[4] Ch.Koeck : Computation of three dimensional flow using the Euler equations and multiple-grid scheme. Int. J. Num. Meth. Fluids n° 5, 1985.

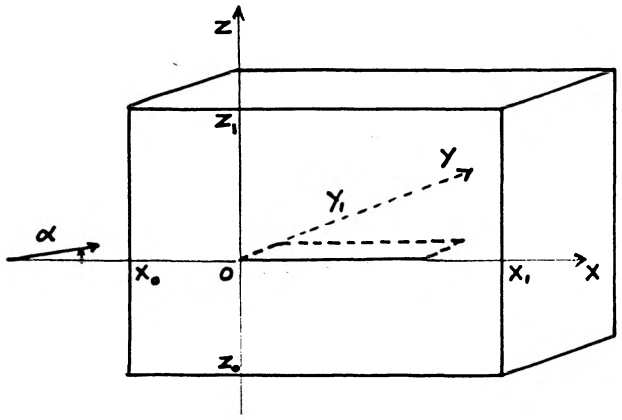


Figure 1 : Domain.

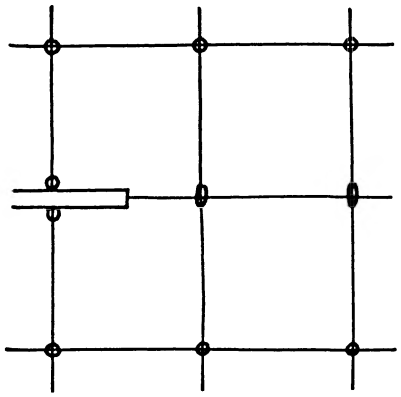
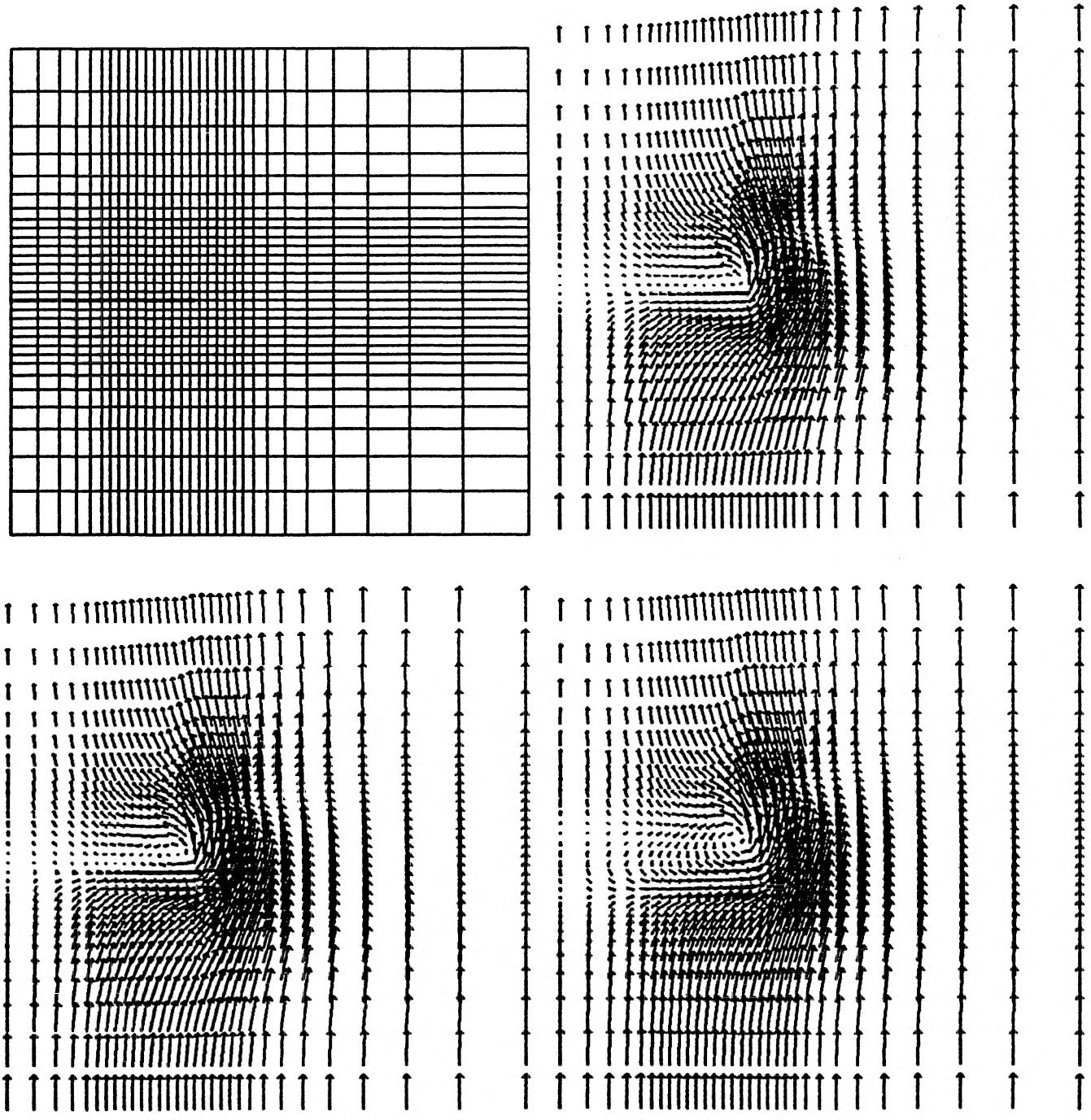


Figure 2 : Tip of the plate.



Figures 3 to 6

Mesh and cross - flow velocities at 70 %, 90 % and 110 % of the plate.

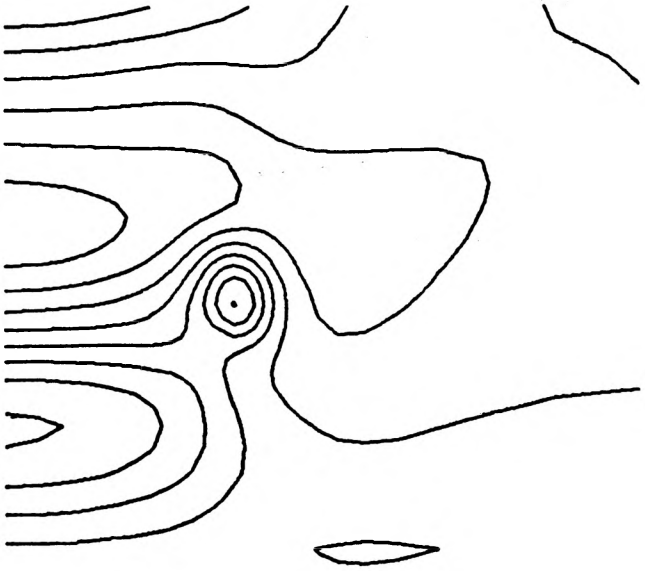


Figure 7
Isobars in a cross plane at the trailing edge.

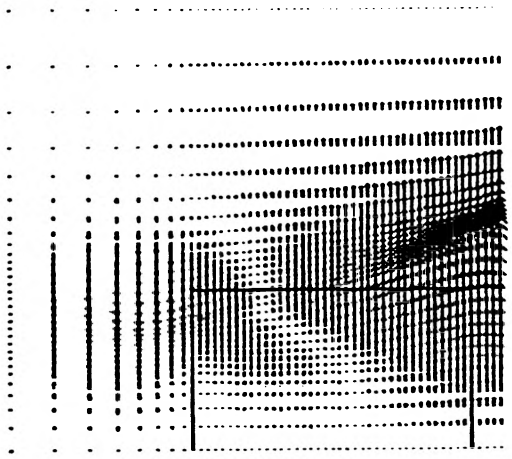


Figure 8
Component v of the velocity on the upper plate.

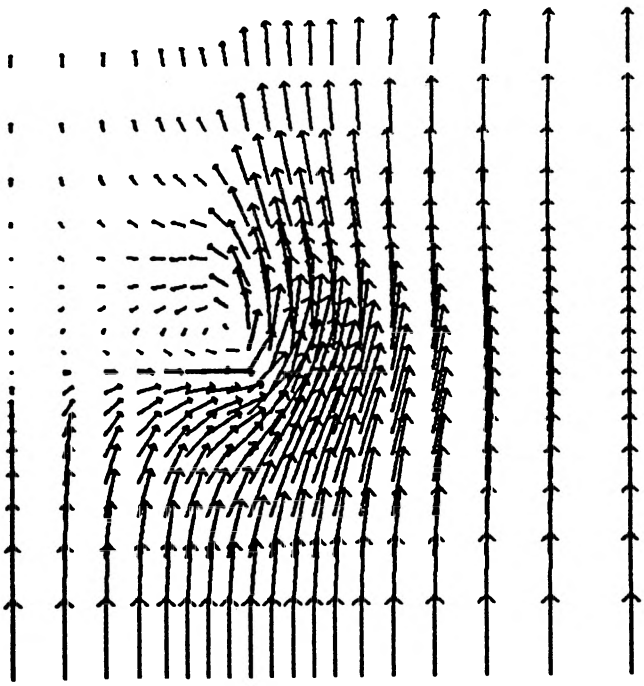


Figure 9
Cross-flow velocities at 90 % of the plate.

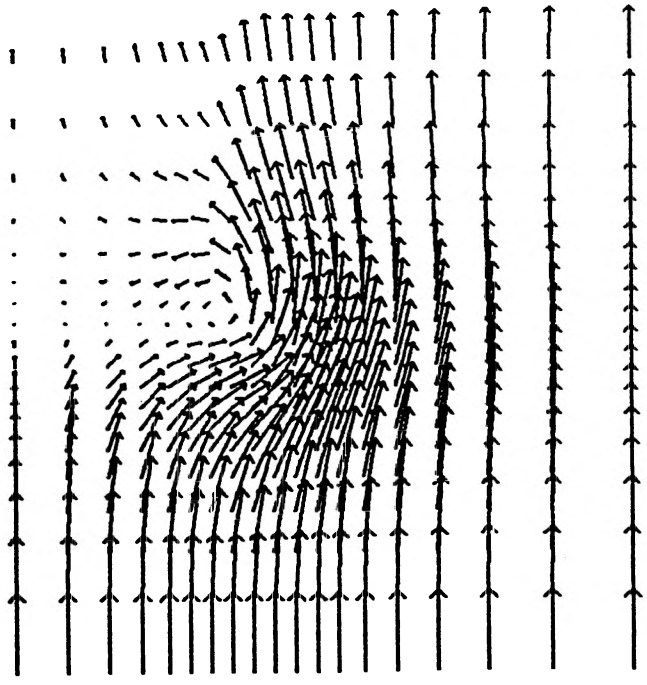


Figure 10
Cross-flow velocities at 110 % of the plate.

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS, VOL. 9, 305–323 (1989)

COMPUTATION OF 3D VORTEX FLOWS PAST A FLAT PLATE AT INCIDENCE THROUGH A VARIATIONAL APPROACH OF THE FULL STEADY EULER EQUATIONS

CHARLES-HENRI BRUNEAU AND JACQUES LAMINIE
Laboratoire d'Analyse Numérique, Université Paris-Sud, 91405 Orsay Cédex, France

AND

JEAN-JACQUES CHATTOT
Aérospatiale, 78133 Les Mureaux, France

SUMMARY

A variational method for solving directly the full steady Euler equations is presented. This method is based on both Newton's linearization and a least squares formulation. The validity of the Euler model and boundary conditions to capture the vortex sheet is discussed. A finite element approximation of the groups of conservative variables is described and results are given for 3D subsonic flows as well as supersonic flows past a flat plate at high angle of attack.

KEY WORDS Steady Euler equations Variational method 3D vortex flows Finite element approximation Newton linearization Least-squares formulation Subsonic and supersonic flows

INTRODUCTION

In the last few years intensive studies have been done by many researchers on numerical computation of vortical flows around wings; in particular, around swept and delta wings using steady or unsteady Euler or Navier–Stokes models (see for instance References 1–3 and references therein). Vortex-dominated flows are of primary interest in aeronautics and space technologies; indeed vortices developing above the wing induce low pressure on the upper surface and consequently yield an additional lift. Numerous questions still remain; among them we can point out:

1. Where and when does separation occur, in particular on smooth surfaces?
2. Is the Euler model able to capture separation on a sharp or on a smooth leading edge without the help of artificial viscosity?
3. How do vortex sheets develop in relation to the Mach number at infinity and the angle of attack?

0271–2091/89/030305–19$09.50

Received 2 February 1988
Revised 13 May 1988

At the First International Conference on Industrial and Applied Mathematics (Paris 1987) several numerical solutions of vortex-dominated flows around wings were presented; in particular A. Rizzi showed solutions about a swept wing with a rounded leading edge calculated from Euler and Navier–Stokes equations. He noticed that separation occurs in both cases but is located more upstream with the Navier–Stokes model, and that the vortex sheets are quite different in shape and growth. We believe that, even with the same mathematical model, artificial viscosity and numerical dissipation yield quite different vortex sheets from one method to another, especially when a smooth leading edge and complex geometry are concerned.

In this paper we choose to study a very simple geometry, a rectangular flat plate of zero thickness. This choice eliminates the difficulty of separation on smooth surfaces and allows us to use the Euler model because the separation does not result from viscous effects. Indeed separation occurs at the leading edge and the vortex structure rolls up at the tip of the plate, grows until the trailing edge and is convected beyond. Our aim is to show that such inviscid separation can be obtained from Euler equations which allow rotational effects, without the help of artificial viscosity or the Kutta condition, and to give some partial answers to the main questions above.

In the next section we present the variational method we use to study the full steady Euler equations; we then discuss the appropriate boundary conditions, the finite element approximation, the numerical method to solve the linear system and the implementation. The last section is devoted to numerical results for several Mach numbers at infinity, angles of attack and meshes.

DESCRIPTION OF THE METHOD

In previous works in 2D and 3D[4, 5] we used Bernoulli's theorem for steady flows to reduce the equation of conservation of energy to its algebraic form and to write consequently a fixed point algorithm on the density. Here we write the whole system as conservative partial differential equations in a 3D domain Ω as follows:

$$
\begin{aligned}
\frac{\partial \rho u}{\partial x}+\frac{\partial \rho v}{\partial y}+\frac{\partial \rho w}{\partial z}&=0,\\
\frac{\partial \rho u^2}{\partial x}+\frac{\partial \rho uv}{\partial y}+\frac{\partial \rho uw}{\partial z}+\frac{\partial(\gamma-1)\rho e}{\partial x}&=0,\\
\frac{\partial \rho uv}{\partial x}+\frac{\partial \rho v^2}{\partial y}+\frac{\partial \rho vw}{\partial z}+\frac{\partial(\gamma-1)\rho e}{\partial y}&=0,\\
\frac{\partial \rho uw}{\partial x}+\frac{\partial \rho vw}{\partial y}+\frac{\partial \rho w^2}{\partial z}+\frac{\partial(\gamma-1)\rho e}{\partial z}&=0,\\
\frac{\partial(\gamma\rho e+\frac{1}{2}\rho q^2)u}{\partial x}+\frac{\partial(\gamma\rho e+\frac{1}{2}\rho q^2)v}{\partial y}+\frac{\partial(\gamma\rho e+\frac{1}{2}\rho q^2)w}{\partial z}&=0,
\end{aligned}
\tag{1}
$$

where we recognize the equation of conservation of mass, the three equations of conservation of momentum and the equation of conservation of energy for the density ρ, the three components of the velocity $\mathbf{q}=(u, v, w)$ and the specific internal energy as unknowns. We denote by q the modulus of $\mathbf{q}$ and by γ the ratio of specific heats. One notices that the hypothesis of a perfect gas with constant specific heats is made and that the pressure and the total enthalpy are replaced by

$(\gamma-1)\rho e$ and $\gamma e+\frac{1}{2}q^2$ respectively. Here these equations are associated with Dirichlet boundary conditions on the plate and at the far-field boundaries of the domain; we impose a tangency condition (as is usual for the Euler model) on both sides of the plate ($\mathbf{q}\cdot\mathbf{n}=0$, where $\mathbf{n}$ is the unit normal pointing out from the plate), a symmetry condition on the plate $y=0$ (see Figure 1) and some other conditions at the far-field boundaries which are discussed in the next section.

For the sake of simplicity we present the method in dimension two; the first idea is to linearize the first-order hyperbolic non-linear operator L appearing in (1) by the Newton method. Indeed the strong non-linearity of the conservative terms with respect to the unknowns in (1) and our wish to keep the groups of conservative variables to enforce the conservation properties do not allow us to write directly a variational formulation. Let us denote the unknowns by $\mathbf{U}^{\mathrm{T}}=(\rho, u, v, e)$; then we solve $L(\mathbf{U})=0$ by means of the algorithm

$$\begin{cases}\mathbf{U}_0 \text{ given},\\ \mathbf{U}_{m+1}=\mathbf{U}_m-\tilde{\mathbf{U}}, \quad m=0,\dots,p,\end{cases}$$

where $\tilde{\mathbf{U}}^{\mathrm{T}}=(\tilde\rho, \tilde u, \tilde v, \tilde e)$ is the solution of the linear system

$$\begin{aligned}
&\frac{\partial\tilde\rho u_m}{\partial x}+\frac{\partial\tilde\rho v_m}{\partial y}+\frac{\partial\rho_m\tilde u}{\partial x}+\frac{\partial\rho_m\tilde v}{\partial y}=\frac{\partial\rho_m u_m}{\partial x}+\frac{\partial\rho_m v_m}{\partial y},\\
&\frac{\partial\tilde\rho u_m^2}{\partial x}+\frac{\partial\tilde\rho u_m v_m}{\partial y}+\frac{\partial(\gamma-1)\tilde\rho e_m}{\partial x}+\frac{\partial 2\rho_m u_m\tilde u}{\partial x}+\frac{\partial\rho_m\tilde u v_m}{\partial y}+\frac{\partial\rho_m u_m\tilde v}{\partial y}+\frac{\partial(\gamma-1)\rho_m\tilde e}{\partial x}\\
&\qquad=\frac{\partial\rho_m u_m^2}{\partial x}+\frac{\partial\rho_m u_m v_m}{\partial y}+\frac{\partial(\gamma-1)\rho_m e_m}{\partial x},\\
&\frac{\partial\tilde\rho u_m v_m}{\partial x}+\frac{\partial\tilde\rho v_m^2}{\partial y}+\frac{\partial(\gamma-1)\tilde\rho e_m}{\partial y}+\frac{\partial\rho_m\tilde u v_m}{\partial x}+\frac{\partial\rho_m u_m\tilde v}{\partial x}+\frac{\partial 2\rho_m v_m\tilde v}{\partial y}+\frac{\partial(\gamma-1)\rho_m\tilde e}{\partial y}\\
&\qquad=\frac{\partial\rho_m u_m v_m}{\partial x}+\frac{\partial\rho_m v_m^2}{\partial y}+\frac{\partial(\gamma-1)\rho_m e_m}{\partial y},\\
&\frac{\partial\tilde\rho(\gamma e_m+\frac{1}{2}q_m^2)u_m}{\partial x}+\frac{\partial\tilde\rho(\gamma e_m+\frac{1}{2}q_m^2)v_m}{\partial y}+\frac{\partial\gamma\rho_m e_m\tilde u}{\partial x}+\frac{\partial\frac{3}{2}\rho_m u_m^2\tilde u+\frac{1}{2}\rho_m v_m^2\tilde u}{\partial x}\\
&\quad+\frac{\partial\rho_m u_m\tilde u v_m}{\partial y}+\frac{\partial\gamma\rho_m e_m\tilde v}{\partial y}+\frac{\partial\frac{3}{2}\rho_m v_m^2\tilde v+\frac{1}{2}\rho_m u_m^2\tilde v}{\partial y}+\frac{\partial\rho_m v_m\tilde v u_m}{\partial x}+\frac{\partial\gamma\rho_m\tilde e u_m}{\partial x}+\frac{\partial\gamma\rho_m\tilde e v_m}{\partial y}\\
&\qquad=\frac{\partial(\gamma\rho_m e_m+\frac{1}{2}\rho_m q_m^2)u_m}{\partial x}+\frac{\partial(\gamma\rho_m e_m+\frac{1}{2}\rho_m q_m^2)v_m}{\partial y}
\end{aligned}\tag{2}$$

associated with some boundary conditions. Thus at each step of the Newton linearization $\tilde{\mathbf{U}}$ is the solution of a linear problem

$$\begin{cases}\mathbf{A}\tilde{\mathbf{U}}=\mathbf{F},\\ \tilde{\mathbf{U}}\in\mathscr{D}(\mathbf{A}),\end{cases}\tag{3}$$

where $\mathscr{D}(\mathbf{A})$ is a space we define more precisely later with Dirichlet boundary conditions for $\tilde{\mathbf{U}}$ which are homogeneous because $\mathbf{U}_0$ is given satisfying the Dirichlet boundary conditions of the

initial problem. In the following we drop the subscript m; thus the operator $\mathbf{A}$ reads

$$\mathbf{A}=\begin{bmatrix} \frac{\partial u\cdot}{\partial x}+\frac{\partial v\cdot}{\partial y} & \frac{\partial \rho\cdot}{\partial x} & \frac{\partial \rho\cdot}{\partial y} & 0 \\ \frac{\partial[u^2+(\gamma-1)e]\cdot}{\partial x}+\frac{\partial uv\cdot}{\partial y} & \frac{\partial 2\rho u\cdot}{\partial x}+\frac{\partial \rho v\cdot}{\partial y} & \frac{\partial \rho u\cdot}{\partial y} & \frac{\partial(\gamma-1)\rho\cdot}{\partial x} \\ \frac{\partial uv\cdot}{\partial x}+\frac{\partial[v^2+(\gamma-1)e]\cdot}{\partial y} & \frac{\partial \rho v\cdot}{\partial x} & \frac{\partial \rho u\cdot}{\partial x}+\frac{\partial 2\rho v\cdot}{\partial y} & \frac{\partial(\gamma-1)\rho\cdot}{\partial y} \\ \frac{\partial(\gamma eu+\frac{1}{2}q^2u)\cdot}{\partial x} & \frac{\partial \gamma\rho e\cdot}{\partial x}+\frac{\partial(\frac{3}{2}\rho u^2+\frac{1}{2}\rho v^2)\cdot}{\partial x} & \frac{\partial \gamma\rho e\cdot}{\partial y}+\frac{\partial(\frac{3}{2}\rho v^2+\frac{1}{2}\rho u^2)\cdot}{\partial y} & \frac{\partial \gamma\rho u\cdot}{\partial x} \\ +\frac{\partial(\gamma ev+\frac{1}{2}q^2v)\cdot}{\partial y} & +\frac{\partial \rho uv\cdot}{\partial y} & +\frac{\partial \rho vu\cdot}{\partial x} & +\frac{\partial \gamma\rho v\cdot}{\partial y} \end{bmatrix}$$

and the vector $\mathbf{F}$ is the left-hand side of (1) written in dimension two. Thus Newton linearization leads us to solve p times a linear problem with a first-order hyperbolic operator still which has the same characteristic curves as the initial one. Indeed, let φ be a function constant on the characteristic curves and φ_x, φ_y be the components of grad φ; then the characteristic polynomial is given by the following determinant simplified by linear combinations of rows and columns (since in system (3), $\mathbf{U}$ is assumed to be regular enough to write $\mathbf{A}$ in its non-conservative form):

$$\begin{bmatrix} u\varphi_x+v\varphi_y & \rho\varphi_x & \rho\varphi_y & 0 \\ 0 & \rho(u\varphi_x+v\varphi_y) & 0 & (\gamma-1)\rho\varphi_x \\ 0 & 0 & \rho(u\varphi_x+v\varphi_y) & (\gamma-1)\rho\varphi_y \\ -\gamma e(u\varphi_x+v\varphi_y) & 0 & 0 & \rho(u\varphi_x+v\varphi_y) \end{bmatrix}$$
$$=\rho^3(u\varphi_x+v\varphi_y)^2[(u\varphi_x+v\varphi_y)^2-\gamma e(\gamma-1)(\varphi_x^2+\varphi_y^2)].$$

As $\mathrm{d}\varphi=0$ on the characteristic curves, by setting $\tau=\mathrm{d}y/\mathrm{d}x=-\varphi_x/\varphi_y$ we get

$$\det=\rho^3\varphi_y^4(-\tau u+v)^2[(-\tau u+v)^2-\gamma e(\gamma-1)(\tau^2+1)]$$

and thus the streamlines are zeros of multiplicity two. Moreover, we have two additional real roots when the Mach number is greater than or equal to one. Finally, we find that the linear operator has the well known characteristic curves of the Euler system and thus the same hyperbolic behaviour. This remark leads us to the second idea, which is to solve the linear system (3) by means of a least squares method. Thus we minimize over the space $\mathcal{D}(\mathbf{A})$ the functional

$$J(\tilde{\mathbf{U}})=\frac{1}{2}\int_\Omega\left(\frac{\partial\tilde\rho u}{\partial x}+\frac{\partial\tilde\rho v}{\partial y}+\frac{\partial\rho\tilde u}{\partial x}+\frac{\partial\rho\tilde v}{\partial y}-\frac{\partial\rho u}{\partial x}-\frac{\partial\rho v}{\partial y}\right)^2\mathrm{d}x\mathrm{d}y$$
$$+\frac{1}{2}\int_\Omega\left(\frac{\partial\tilde\rho u^2}{\partial x}+\frac{\partial\tilde\rho uv}{\partial y}+\frac{\partial(\gamma-1)\tilde\rho e}{\partial x}+\frac{\partial 2\rho u\tilde u}{\partial x}+\frac{\partial\rho\tilde uv}{\partial y}+\frac{\partial\rho u\tilde v}{\partial y}+\frac{\partial(\gamma-1)\rho\tilde e}{\partial x}\right.$$
$$\left.-\frac{\partial\rho u^2}{\partial x}-\frac{\partial\rho uv}{\partial y}-\frac{\partial(\gamma-1)\rho e}{\partial x}\right)^2\mathrm{d}x\mathrm{d}y$$
$$+\frac{1}{2}\int_\Omega\left(\frac{\partial\tilde\rho uv}{\partial x}+\frac{\partial\tilde\rho v^2}{\partial y}+\frac{\partial(\gamma-1)\tilde\rho e}{\partial y}+\frac{\partial\rho\tilde uv}{\partial x}+\frac{\partial\rho u\tilde v}{\partial x}+\frac{\partial 2\rho v\tilde v}{\partial y}+\frac{\partial(\gamma-1)\rho\tilde e}{\partial y}\right.$$

$$-\frac{\partial \rho uv}{\partial x}-\frac{\partial \rho v^2}{\partial y}-\frac{\partial(\gamma-1)\rho e}{\partial y}\Bigg)^2 \mathrm{d}x\mathrm{d}y$$
$$+\frac{1}{2}\int_\Omega\Bigg(\frac{\partial\tilde{\rho}(\gamma e+\frac{1}{2}q^2)u}{\partial x}+\frac{\partial\tilde{\rho}(\gamma e+\frac{1}{2}q^2)v}{\partial y}+\frac{\partial\gamma\rho e\tilde{u}}{\partial x}+\frac{\partial\frac{3}{2}\rho u^2\tilde{u}+\frac{1}{2}\rho v^2\tilde{u}}{\partial x}$$
$$+\frac{\partial\rho u\tilde{u}v}{\partial y}+\frac{\partial\gamma\rho e\tilde{v}}{\partial y}+\frac{\partial\frac{3}{2}\rho v^2\tilde{v}+\frac{1}{2}\rho u^2\tilde{v}}{\partial y}+\frac{\partial\rho v\tilde{v}u}{\partial x}+\frac{\partial\gamma\rho\tilde{e}u}{\partial x}+\frac{\partial\gamma\rho\tilde{e}v}{\partial y}$$
$$-\frac{\partial(\gamma\rho e+\frac{1}{2}\rho q^2)u}{\partial x}-\frac{\partial(\gamma\rho e+\frac{1}{2}\rho q^2)v}{\partial y}\Bigg)^2 \mathrm{d}x\mathrm{d}y$$

and, computing the Gateau derivative G of J with respect to the unknowns, we get the variational formulation

$$\begin{cases}\text{find } \tilde{\mathbf{U}}\in\mathscr{D}(\mathbf{A}) \text{ such that}\\ \langle G(\tilde{\mathbf{U}}),\bar{\mathbf{U}}\rangle=0 \quad \forall\bar{\mathbf{U}}\in\mathscr{D}(\mathbf{A}),\end{cases} \tag{4}$$

where

$$\langle G(\tilde{\mathbf{U}}),\bar{\mathbf{U}}\rangle=\int_\Omega\Bigg(\frac{\partial\tilde{\rho}u}{\partial x}+\frac{\partial\tilde{\rho}v}{\partial y}+\frac{\partial\rho\tilde{u}}{\partial x}+\frac{\partial\rho\tilde{v}}{\partial y}-\frac{\partial\rho u}{\partial x}-\frac{\partial\rho v}{\partial y}\Bigg)\Bigg(\frac{\partial\bar{\rho}u}{\partial x}+\frac{\partial\bar{\rho}v}{\partial y}+\frac{\partial\rho\bar{u}}{\partial x}+\frac{\partial\rho\bar{v}}{\partial y}\Bigg)\mathrm{d}x\mathrm{d}y$$
$$+\int_\Omega\Bigg(\frac{\partial\tilde{\rho}u^2}{\partial x}+\frac{\partial\tilde{\rho}uv}{\partial y}+\frac{\partial(\gamma-1)\tilde{\rho}e}{\partial x}+\frac{\partial 2\rho u\tilde{u}}{\partial x}+\frac{\partial\rho\tilde{u}v}{\partial y}+\frac{\partial\rho u\tilde{v}}{\partial y}+\frac{\partial(\gamma-1)\rho\tilde{e}}{\partial x}$$
$$-\frac{\partial\rho u^2}{\partial x}-\frac{\partial\rho uv}{\partial y}-\frac{\partial(\gamma-1)\rho e}{\partial x}\Bigg)$$
$$\times\Bigg(\frac{\partial\bar{\rho}u^2}{\partial x}+\frac{\partial\bar{\rho}uv}{\partial y}+\frac{\partial(\gamma-1)\bar{\rho}e}{\partial x}+\frac{\partial 2\rho u\bar{u}}{\partial x}+\frac{\partial\rho\bar{u}v}{\partial y}+\frac{\partial\rho u\bar{v}}{\partial y}+\frac{\partial(\gamma-1)\rho\bar{e}}{\partial x}\Bigg)\mathrm{d}x\mathrm{d}y$$
$$+\int_\Omega\Bigg(\frac{\partial\tilde{\rho}uv}{\partial x}+\frac{\partial\tilde{\rho}v^2}{\partial y}+\frac{\partial(\gamma-1)\tilde{\rho}e}{\partial y}+\frac{\partial\rho\tilde{u}v}{\partial x}+\frac{\partial\rho u\tilde{v}}{\partial x}+\frac{\partial 2\rho v\tilde{v}}{\partial y}+\frac{\partial(\gamma-1)\rho\tilde{e}}{\partial y}$$
$$-\frac{\partial\rho uv}{\partial x}-\frac{\partial\rho v^2}{\partial y}-\frac{\partial(\gamma-1)\rho e}{\partial y}\Bigg)$$
$$\times\Bigg(\frac{\partial\bar{\rho}uv}{\partial x}+\frac{\partial\bar{\rho}v^2}{\partial y}+\frac{\partial(\gamma-1)\bar{\rho}e}{\partial y}+\frac{\partial\rho\bar{u}v}{\partial x}+\frac{\partial\rho u\bar{v}}{\partial x}+\frac{\partial 2\rho v\bar{v}}{\partial y}+\frac{\partial(\gamma-1)\rho\bar{e}}{\partial y}\Bigg)\mathrm{d}x\mathrm{d}y$$
$$+\int_\Omega\Bigg(\frac{\partial\tilde{\rho}(\gamma e+\frac{1}{2}q^2)u}{\partial x}+\frac{\partial\tilde{\rho}(\gamma e+\frac{1}{2}q^2)v}{\partial y}+\frac{\partial\gamma\rho e\tilde{u}}{\partial x}+\frac{\partial\frac{3}{2}\rho u^2\tilde{u}+\frac{1}{2}\rho v^2\tilde{u}}{\partial x}$$
$$+\frac{\partial\rho u\tilde{u}v}{\partial y}+\frac{\partial\gamma\rho e\tilde{v}}{\partial y}+\frac{\partial\frac{3}{2}\rho v^2\tilde{v}+\frac{1}{2}\rho u^2\tilde{v}}{\partial y}+\frac{\partial\rho v\tilde{v}u}{\partial x}+\frac{\partial\gamma\rho\tilde{e}u}{\partial x}+\frac{\partial\gamma\rho\tilde{e}v}{\partial y}$$
$$-\frac{\partial(\gamma\rho e+\frac{1}{2}\rho q^2)u}{\partial x}-\frac{\partial(\gamma\rho e+\frac{1}{2}\rho q^2)v}{\partial y}\Bigg)$$
$$\times\Bigg(\frac{\partial\bar{\rho}(\gamma e+\frac{1}{2}q^2)u}{\partial x}+\frac{\partial\bar{\rho}(\gamma e+\frac{1}{2}q^2)v}{\partial y}+\frac{\partial\gamma\rho e\bar{u}}{\partial x}+\frac{\partial\frac{3}{2}\rho u^2\bar{u}+\frac{1}{2}\rho v^2\bar{u}}{\partial x}$$
$$+\frac{\partial\rho u\bar{u}v}{\partial y}+\frac{\partial\gamma\rho e\bar{v}}{\partial y}+\frac{\partial\frac{3}{2}\rho v^2\bar{v}+\frac{1}{2}\rho u^2\bar{v}}{\partial y}+\frac{\partial\rho v\bar{v}u}{\partial x}+\frac{\partial\gamma\rho\bar{e}u}{\partial x}+\frac{\partial\gamma\rho\bar{e}v}{\partial y}\Bigg)\mathrm{d}x\mathrm{d}y.$$

All the above formulae are meaningful if we define the following spaces:

$\mathscr{V} = \{\mathbf{U} \in (L^\infty(\Omega)^4, \mathbf{U} > \mathbf{0} \text{ such that } L(\mathbf{U}) \in (L^2(\Omega))^4\}$,

$\tilde{\mathscr{V}} = \{\tilde{\mathbf{U}} \in (L^2(\Omega))^4 \text{ such that } \mathbf{A}\tilde{\mathbf{U}} \in (L^2(\Omega))^4\}$,

$\mathscr{D}(\mathbf{A}) = \{\tilde{\mathbf{U}} \in \tilde{\mathscr{V}}$ satisfying some homogeneous Dirichlet boundary conditions$\}$.

We notice that the initial variables (components or groups of components of $\mathbf{U}$) appearing in $\mathbf{A}\tilde{\mathbf{U}}$ can be considered as positive weights in $\tilde{\mathscr{V}}$; we shall make precise the boundary conditions of $\mathscr{D}(\mathbf{A})$ in the next section. Thus the equality in (3) is understood in the sense of $(L^2(\Omega))^4$ and thus almost everywhere in Ω; and all the terms in $J(\tilde{\mathbf{U}})$ and $(G(\tilde{\mathbf{U}}), \bar{\mathbf{U}})$ are in $L^1(\Omega)$.

Now if we take instead of $\mathscr{V}$ the space $\mathscr{V}_1$ defined like $\mathscr{V}$ by replacing $L^2(\Omega)$ by $H^1(\Omega)$, i.e. we request more regularity on the groups of conservative variables, we can give an interpretation of (4). Indeed, let $\bar{\mathbf{U}}$ be in $(\mathscr{D}(\Omega))^4$; then we get, in the distribution sense,

$$\mathbf{A}^*(\mathbf{A}\tilde{\mathbf{U}}) = \mathbf{A}^*\mathbf{F}, \tag{5}$$

where the linear adjoint operator $\mathbf{A}^*$ is given by

$$\mathbf{A}^* = \begin{bmatrix}
-u\frac{\partial\cdot}{\partial x} - v\frac{\partial\cdot}{\partial y} & -[u^2 + (\gamma-1)e]\frac{\partial\cdot}{\partial x} - uv\frac{\partial\cdot}{\partial y} & -uv\frac{\partial\cdot}{\partial x} - [v^2(+\gamma-1)e]\frac{\partial\cdot}{\partial y} & -(\tfrac{1}{2}q^2u + \gamma eu)\frac{\partial\cdot}{\partial x} \\
 & & & -(\tfrac{1}{2}q^2v + \gamma ev)\frac{\partial\cdot}{\partial y} \\
-\rho\frac{\partial\cdot}{\partial x} & -2\rho u\frac{\partial\cdot}{\partial x} - \rho v\frac{\partial\cdot}{\partial y} & -\rho v\frac{\partial\cdot}{\partial x} & -(\tfrac{3}{2}\rho u^2 + \tfrac{1}{2}\rho v^2)\frac{\partial\cdot}{\partial x} \\
 & & & -\gamma\rho e\frac{\partial\cdot}{\partial x} - \rho uv\frac{\partial\cdot}{\partial y} \\
-\rho\frac{\partial\cdot}{\partial y} & -\rho u\frac{\partial\cdot}{\partial y} & -\rho u\frac{\partial\cdot}{\partial x} - 2\rho v\frac{\partial\cdot}{\partial y} & -\rho vu\frac{\partial\cdot}{\partial x} - \gamma\rho e\frac{\partial\cdot}{\partial y} \\
 & & & -(\tfrac{1}{2}\rho u^2 + \tfrac{3}{2}\rho v^2)\frac{\partial\cdot}{\partial y} \\
0 & -(\gamma-1)\rho\frac{\partial\cdot}{\partial x} & -(\gamma-1)\rho\frac{\partial\cdot}{\partial y} & -\gamma\rho u\frac{\partial\cdot}{\partial x} - \gamma\rho v\frac{\partial\cdot}{\partial y}
\end{bmatrix}$$

Then, with the additional regularity, this equality is meaningful in the sense of $(L^2(\Omega))^4$ and consequently almost everywhere in Ω. Thus we have transformed the first-order hyperbolic linear system (2) into a second-order linear system of parabolic type; for instance, if $\mathbf{A}$ reduces to its main diagonal, we find that the characteristic polynomial of $\mathbf{A}^*\mathbf{A}$ is

$$\det = \gamma^2\rho^6\varphi_y^8(-\tau u + v)^4(-2\tau u + v)^2(\tau u + 2v)^2$$

and thus $\mathbf{A}^*\mathbf{A}$ is a degenerated elliptic operator. In fact $\mathbf{A}$ and $\mathbf{A}^*$ have the same characteristic polynomial, and that of $\mathbf{A}^*\mathbf{A}$ is equal to the square of it. In the next section we show how we can set the boundary conditions to get a well posed problem equivalent to problem (3).

BOUNDARY CONDITIONS

For the Euler model one generally imposes different boundary conditions for subsonic flows and supersonic flows according to the number of characteristic curves entering in the associated unsteady problem. So, on one hand, the whole flow is given upstream and no condition is imposed downstream for supersonic flows and, on the other hand, one less condition is given upstream and one condition is imposed downstream for subsonic flows. For example, in two space dimensions the entropy and the flow direction are imposed in the entrance section and the pressure is specified in the exit section in subsonic flow. In three space dimensions this is not always possible, in particular for vortex flows where the vortex sheet must propagate downstream. For unsteady flows we can use non-reflecting boundary conditions as explained in Reference 6, but the situation seems to be less clear in the steady case. For the steady Euler equations we have chosen in our methodology to impose all the flow variables upstream and no variable downstream in supersonic as well as in subsonic flows. This is consistent with the fact that, on one hand, upstream influence is possible in subsonic flow via the centred discretization used and, on the other hand, the conservation equations are satisfied near the downstream boundary. Furthermore, the flow domain is unbounded and the obstacle has a finite length, so there is no explicit physical boundary condition to apply.

We show now how the least squares formulation associated with these boundary conditions is equivalent to solving the first-order system. We first define parts of the boundary as follows:

Γ_i, the entrance section where $\mathbf{q}_\infty \cdot \mathbf{n} < 0$

Γ_o, the exit section where $\mathbf{q}_\infty \cdot \mathbf{n} > 0$

Γ_w, the wing contour or surface where $\mathbf{q} \cdot \mathbf{n} = 0$ is imposed

Γ_s, the part of the exterior boundary where $\mathbf{q} \cdot \mathbf{n} = 0$ is imposed (for instance by symmetry).

We then set the conditions

$$\begin{cases} \tilde{\mathbf{q}} \cdot \mathbf{n} = 0 & \text{on } \Gamma_w \cup \Gamma_s, \\ \tilde{\mathbf{U}} \equiv 0 & \text{on } \Gamma_i \end{cases} \tag{6}$$

and define the space

$$\mathscr{D}(\mathbf{A}) = \{\tilde{\mathbf{U}} \in \tilde{\mathscr{V}} \text{ satisfying (6)}\}.$$

Now with $\tilde{\mathbf{U}} \in \mathscr{D}(\mathbf{A})$ and $\bar{\mathbf{U}} \in (H^1(\Omega))^4$ we get from (3)

$$\begin{aligned}
(\mathbf{A}\tilde{\mathbf{U}}, \bar{\mathbf{U}}) = (\tilde{\mathbf{U}}, \mathbf{A}^*\bar{\mathbf{U}}) &+ \int_{\partial\Omega} [\tilde{\rho}\bar{\rho}(\mathbf{q}\cdot\mathbf{n}) + \rho\bar{\rho}(\tilde{\mathbf{q}}\cdot\mathbf{n})]\,\mathrm{d}\sigma \\
&+ \int_{\partial\Omega} [\tilde{\rho}u\tilde{u}(\mathbf{q}\cdot\mathbf{n}) + \tilde{\rho}v\bar{v}(\mathbf{q}\cdot\mathbf{n}) + (\gamma-1)\tilde{\rho}e(\bar{\mathbf{q}}\cdot\mathbf{n}) \\
&+ \rho\tilde{u}\bar{u}(\mathbf{q}\cdot\mathbf{n}) + \rho\tilde{v}\bar{v}(\mathbf{q}\cdot\mathbf{n}) + (\gamma-1)\rho\tilde{e}(\bar{\mathbf{q}}\cdot\mathbf{n}) \\
&+ \rho\tilde{u}\bar{u}(\tilde{\mathbf{q}}\cdot\mathbf{n}) + \rho v\bar{v}(\tilde{\mathbf{q}}\cdot\mathbf{n})]\,\mathrm{d}\sigma \\
&+ \int_{\partial\Omega} [\tilde{\rho}(\gamma e + \tfrac{1}{2}q^2)\bar{e}(\mathbf{q}\cdot\mathbf{n}) + \gamma\rho e\bar{e}(\tilde{\mathbf{q}}\cdot\mathbf{n}) + \gamma\rho\tilde{e}\bar{e}(\mathbf{q}\cdot\mathbf{n}) \\
&+ \tfrac{1}{2}\rho u^2\bar{e}(\tilde{\mathbf{q}}\cdot\mathbf{n}) + \tfrac{1}{2}\rho v^2\bar{e}(\tilde{\mathbf{q}}\cdot\mathbf{n}) \\
&+ \rho u\tilde{u}\bar{e}(\mathbf{q}\cdot\mathbf{n}) + \rho v\tilde{v}\bar{e}(\mathbf{q}\cdot\mathbf{n})]\,\mathrm{d}\sigma.
\end{aligned}$$

Thus with $\mathbf{q}\cdot\mathbf{n}=0$ on $\Gamma_w \cup \Gamma_s$, in addition to (6) we get the adjoint boundary conditions

$$\begin{cases} \tilde{\mathbf{q}}\cdot\mathbf{n}=0 & \text{on } \Gamma_w \cup \Gamma_s, \\ \tilde{\mathbf{U}}=0 & \text{on } \Gamma_o \end{cases} \tag{7}$$

and define the space

$$\mathcal{D}(\mathbf{A}^*)=\{\tilde{\mathbf{U}}\in(H^1(\Omega))^4;\quad \tilde{\mathbf{U}} \text{ satisfies (7)}\}.$$

Then from equation (5) with $\tilde{\mathbf{U}}\in\mathcal{D}(\mathbf{A})$ we get

$$(\mathbf{A}^*(\mathbf{A}\tilde{\mathbf{U}}-\mathbf{F}), \tilde{\mathbf{U}})=(G(\tilde{\mathbf{U}}), \tilde{\mathbf{U}})$$

if $\mathbf{A}\tilde{\mathbf{U}}-\mathbf{F}\in\mathcal{D}(\mathbf{A}^*)$ (which needs the additional regularity $\mathbf{U}\in\mathcal{V}_1$). Finally, (4) is equivalent to solving

$$\begin{cases} \mathbf{A}^*(\mathbf{A}\tilde{\mathbf{U}})=\mathbf{A}^*\mathbf{F} & \text{in } \Omega, \\ \tilde{\mathbf{U}}\in\mathcal{D}(\mathbf{A}), \\ \mathbf{A}\tilde{\mathbf{U}}-\mathbf{F}\in\mathcal{D}(\mathbf{A}^*) \end{cases} \tag{8}$$

and this problem is equivalent to the initial linear problem (3) under the condition that $\mathbf{A}^*$ is a one-to-one operator and admits a unique solution under the condition that $\mathbf{A}$ is a one-to-one operator.

We see that on Γ_o, where no condition is given, the least squares formulation forces the equations to be satisfied and yields a well posed linear problem, the solution of which can be computed without artifacts (artificial viscosity or Kutta condition).

Remark. Here the condition $\mathbf{q}\cdot\mathbf{n}=0$ reduces to $v=0$ or $w=0$, but for more complex geometries it is not so easy to impose such a non-linear condition; once again the least squares formulation is helpful because we can add a boundary term in the functional $J(\tilde{\mathbf{U}})$ as follows:[4]

$$\frac{1}{2}\int_{\Gamma_w}[(\mathbf{q}-\tilde{\mathbf{q}})\cdot\mathbf{n}]^2 d\sigma,$$

and thus the next iterate $\mathbf{U}_{m+1}$ of the Newton method satisfies the tangency condition on the wing.

APPROXIMATION

In order that this paper be relatively self-contained, we give a short description of the finite element approximation. Let W be a Hilbert space on which is set a variational problem

$$\begin{cases} \text{find } u\in W \text{ such that} \\ a(u, v)=l(v) \quad \forall v\in W, \end{cases} \tag{9}$$

where a is a W-elliptic continuous bilinear form and l is a continuous linear form on W; the purpose of the finite element method is to construct a finite-dimensional space W_h approximating W such that either $W_h\subset W$ or $W_h\not\subset W$. We restrict our study to the conforming case ($W_h\subset W$) for which the approximated variational problem reads

$$\begin{cases} \text{find } u_h\in W_h \text{ such that} \\ a(u_h, v_h)=l(v_h) \quad \forall v_h\in W_h, \end{cases} \tag{10}$$

where a and l have the same definition. If we define W_h and construct its basis, the problem (10) is completely determined and yields a linear system we can solve by any inversion method.

Now we see how to construct a conforming approximation of the space $H^1(\Omega)$ by means of a $\mathbb{Q}^1$ finite element in a regular domain of $\mathbb{R}^3$. Let K be a regular hexaedron in $\mathbb{R}^3$ with vertices (a_k),

$k=1, \ldots, 8$; we define the set of degrees of freedom

$$\Sigma_K=\{l_k : p \in P_K \to p(a_k) \in \mathbb{R};\ k=1, \ldots, 8\}$$

and the space of polynomials

$$P_K=\text{span}\{1, x, y, z, xy, xz, yz, xyz\};$$

then $\mathbb{Q}1$ denotes the finite element (K, Σ_K, P_K). In a regular domain as shown in Figure 1 we can easily construct a mesh made of regular hexahedra and define the space

$$W_h=\{w_h \in C^0(\bar{\Omega}) \text{ such that } w_h/K \in \mathbb{Q}1 \text{ for each } K\},$$

the dimension of which is equal to the number of vertices (a_i), $i=1, \ldots, I$, of the mesh and the basis of which is given by $\phi_i(a_j)=\delta_{ij}$, $1 \leqslant i, j \leqslant I$, where δ_{ij} is the Kronecker symbol. Moreover, this space is included in $H^1(\Omega)$ and then (10) is a conforming approximation of (9).

Here we introduce the spaces

$\mathcal{V}_h=\{\mathbf{U}_h \in (L^\infty(\Omega))^4,\ \mathbf{U}_h>\mathbf{0}$ such that every group of approximate variables $(\rho_h u_h, \ldots)$ belongs to $W_h\}$,

$\tilde{\mathcal{V}}_h=\{\tilde{\mathbf{U}}_h \in (L^2(\Omega))^4$ such that every group of mixed approximate variables $(\tilde{\rho}_h u_h, \ldots)$ belongs to $W_h\}$,

$$\mathcal{D}_h(\mathbf{A})=\{\tilde{\mathbf{U}}_h \in \tilde{\mathcal{V}}_h \text{ satisfying (6)}\}$$

and the approximate problem of (4) reads

$$\begin{cases} \text{find } \tilde{\mathbf{U}}_h \in \mathcal{D}_h(\mathbf{A}) \text{ such that} \\ (G(\tilde{\mathbf{U}}_h), \bar{\mathbf{U}}_h)=0 \quad \forall \bar{\mathbf{U}}_h \in \mathcal{D}_h(\mathbf{A}). \end{cases} \tag{11}$$

This problem is a conforming approximation of (4) with a centred scheme where not the unknowns separately but the groups of variables are approximated by the finite element method; for instance, we write

$$\tilde{\rho}_h u_h=\sum_{i=1}^{I}(\tilde{\rho}u)_i \phi_i,$$

where $(\tilde{\rho}u)_i=\tilde{\rho}_i u_i$ is the value of the group $\tilde{\rho}_h u_h$ at the vertex a_i of the mesh. This is the main point of the approximation; indeed, requiring $\tilde{\rho}_h$ and u_h separately to belong to W_h would impose too much regularity on the variables and would not allow the shocks and the contact discontinuities. Furthermore, this is the more natural way to realize a conforming approximation of the space $\mathcal{D}(\mathbf{A})$ and we have the following result:

problem (11) admits a unique solution $\tilde{\mathbf{U}}_h \in \mathcal{D}_h(\mathbf{A})$.

Proof. Obviously the result will be achieved if we prove that $|\mathbf{A}\tilde{\mathbf{U}}_h|_{(L^2(\Omega))^4}$ is a norm over the space $\mathcal{D}_h(\mathbf{A})$. In order to simplify the proof, we consider a mesh in a 2D tube and $K=[0, 1] \times [0, 1]$; we denote by $\tilde{\rho}_h u_h/K$ the restriction of $\tilde{\rho}_h u_h$ to K. As a first step we show that if $\mathbf{A}\tilde{\mathbf{U}}_h$ vanishes on K, then $\tilde{\mathbf{U}}_h/K \equiv \mathbf{0}$. A polynomial of $\mathbb{Q}1$ is written in its general form $a+bx+cy+dxy$, but with the boundary conditions upstream it reduces to $bx+dxy$; moreover, as v_h and $\tilde{v}_h$ vanish at point $(1, 0)$, we have

$$\tilde{\rho}_h u_h/K=p_1=b_1 x+d_1 xy, \qquad \tilde{\rho}_h v_h/K=p_2=d_2 xy,$$
$$\rho_h \tilde{u}_h/K=p_3=b_3 x+d_3 xy, \qquad \rho_h \tilde{v}_h/K=p_4=d_4 xy.$$

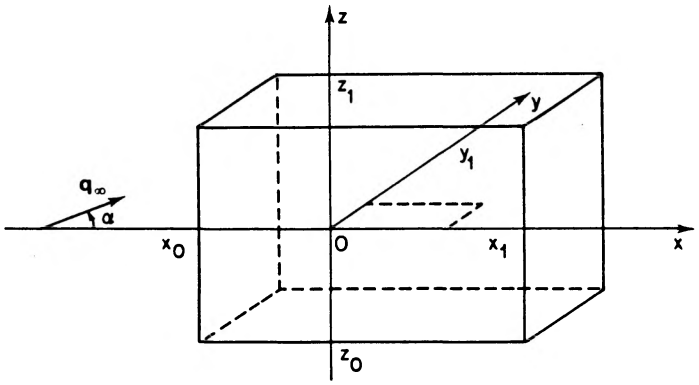


Figure 1. Domain

But from the first equation we have

$$\frac{\partial p_1}{\partial x}+\frac{\partial p_2}{\partial y}+\frac{\partial p_3}{\partial x}+\frac{\partial p_4}{\partial y}=0 \Leftrightarrow b_1+b_3+(d_2+d_4)x+(d_1+d_3)y=0;$$

then we get $p_1+p_3=0$ and $p_2+p_4=0$. But also we have

$$(\gamma-1)\tilde{\rho}_h e_h/K=p_5=b_5x+d_5xy, \qquad \rho_h u_h \tilde{u}_h/K=p_6=b_6x+d_6xy,$$
$$(\gamma-1)\rho_h \tilde{e}_h/K=p_7=b_7x+d_7xy, \qquad \rho_h \tilde{u}_h v_h/K=p_8=d_8xy,$$

and the second equation

$$\left(\frac{\partial \tilde{\rho}_h u_h^2}{\partial x}+\frac{\partial \tilde{\rho}_h u_h v_h}{\partial y}+\frac{\partial(\gamma-1)\tilde{\rho}_h e_h}{\partial x}+\frac{\partial 2\rho_h u_h \tilde{u}_h}{\partial x}+\frac{\partial \rho_h \tilde{u}_h v_h}{\partial y}+\frac{\partial \rho_h u_h \tilde{v}_h}{\partial y}+\frac{\partial(\gamma-1)\rho_h \tilde{e}_h}{\partial x}\right)\Big/ K$$
$$=u_h\left(\frac{\partial p_1}{\partial x}+\frac{\partial p_2}{\partial y}+\frac{\partial p_3}{\partial x}+\frac{\partial p_4}{\partial y}\right)+(p_1+p_3)\frac{\partial u_h}{\partial x}+(p_2+p_4)\frac{\partial u_h}{\partial y}+\frac{\partial p_5}{\partial x}+\frac{\partial p_6}{\partial x}+\frac{\partial p_7}{\partial x}+\frac{\partial p_8}{\partial y}$$
$$=\frac{\partial p_5}{\partial x}+\frac{\partial p_6}{\partial x}+\frac{\partial p_7}{\partial x}+\frac{\partial p_8}{\partial y}=0$$
$$\Leftrightarrow b_5+b_6+b_7+d_8x+(d_5+d_6+d_7)y=0;$$

then $d_8=0$ and $\rho_h \tilde{u}_h v_h/K\equiv 0$ on K.

Proceeding in the same way with the third equation, we get $\rho_h u_h \tilde{v}_h/K\equiv 0$ on K and, as $\mathbf{q}_h$ cannot vanish on an element, we have either $\tilde{u}_h/K\equiv 0$ or $\tilde{v}_h/K\equiv 0$ and consequently $p_5+p_7=0$ (indeed $\tilde{u}_h/K\equiv 0$ implies $p_6=0$); thus we get $\tilde{u}_h/K\equiv 0$ from the equalities $p_1+p_3=p_2+p_4=p_5+p_7=0$. Now marching element by element on the first row from left to right, we get that $\tilde{u}_h$ vanishes on this first row. On the first element of the second row a polynomial of $\mathbb{Q}1$ reduces now to dxy and we can apply the same method to show that $\tilde{\mathbf{U}}_h$ vanishes on this element, then on the second row and finally everywhere on Ω; so $|\mathbf{A}\tilde{\mathbf{U}}_h|_{(L^2(\Omega))^4}$ is a norm over $\mathscr{D}_h(\mathbf{A})$.

Now the unique solution of problem (11) will be obtained by solving a linear system which, unfortunately, is not so easy to invert as we shall see in the next section.

STUDY AND INVERSION OF THE LINEAR SYSTEM

Let (11) be written in its matrix form:

$$\begin{cases} \text{find the vector } \mathbf{X} \text{ such that} \\ \mathbf{SX} = \mathbf{Y}. \end{cases} \tag{12}$$

We first look at the properties of the matrix S. From the least squares minimization and the proof above we know that S is a symmetric positive definite matrix. But if A is a one-to-one operator on $\mathscr{D}_h(\mathbf{A})$, it is not true in general on $\mathscr{D}(\mathbf{A})$; then A can have a kernel and a zero in its spectrum, so we can expect the matrix S to have small eigenvalues. Indeed, if the kernel is not included in $\mathscr{D}_h(\mathbf{A})$, its eigenvectors can be approached by the conforming method and thus the smallest eigenvalue of S should become smaller and smaller as the mesh size h goes to zero. We used the algorithm developed in Reference 7 based on a Lanczos tridiagonalization without reorthogonalization to compute the eigenvalues of S and found that the smallest eigenvalue decreases and the conditioning increases when h becomes smaller. This is illustrated in Table I at the first step of the Newton method with $\mathbf{U}_0$ uniformly defined equal to the values of the incoming flow for Mach number $M = 0{\cdot}7$ and an angle of attack $\alpha = 15°$ (the results are about the same for $M = 2$, $\alpha = 10°$ and if $\mathbf{U}_0$ is replaced by an iterated solution); so we have to find the right method to solve (12).

For our regular domain (see figure 1) S is a block-structured symmetric matrix where each block represents the contribution of a vertical plane parallel to $(0, y, z)$ and (12) can be rewritten as a block tridiagonal system:

$$\begin{pmatrix} \mathbf{E}_1 & \mathbf{F}_1 & & & \\ \mathbf{F}_1^{\mathrm{T}} & \mathbf{E}_2 & \mathbf{F}_2 & & \\ & & & \mathbf{F}_{N-1} & \\ & & \mathbf{F}_{N-1}^{\mathrm{T}} & \mathbf{E}_N & \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \\ \mathbf{X}_N \end{pmatrix} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \\ \mathbf{Y}_N \end{pmatrix}. \tag{13}$$

Until now, the most efficient way we have found to solve this system is to use a block successive over-relaxation (BSOR) algorithm where the diagonal blocks $\mathbf{E}_n$ are inverted by the incomplete Cholesky conjugate gradient (ICCG) method. However, the $\mathbf{E}_n$ blocks are so ill conditioned that we need an additional shifting term to stabilize the whole process; thus the shifted block successive over-relaxation (SBSOR) algorithm reads

$$\begin{cases} (\mathbf{E}_n + \beta \mathbf{D}_n)\bar{X}_n^{j+1} = \mathbf{Y}_n - \mathbf{F}_{n-1}^{\mathrm{T}} X_{n-1}^{j+1} - \mathbf{F}_n X_{n+1}^{j} + \beta \mathbf{D}_n X_n^{j}, \\ X_n^{j+1} = \omega \bar{X}_n^{j+1} + (1-\omega) X_n^{j}, \end{cases} \tag{14}$$

where β, $0 < \beta < 1$, is the shift parameter, ω, $1 \leqslant \omega < 2$, is the over-relaxation parameter and $\mathbf{D}_n$ is the main diagonal of $\mathbf{E}_n$. Of course if $\beta = 0$ we have the usual BSOR algorithm, and if we add the whole block $\mathbf{E}_n$ instead of $\mathbf{D}_n$ in (14) we have also the usual BSOR algorithm with a new relaxation parameter $\delta = \omega/(1+\beta)$. What we want to do is to correct the conditioning of the blocks we have to

Table I

h	Number of unknowns	Smallest eigenvalue	Largest eigenvalue	Conditioning
0·0475	1470	$0{\cdot}4248 \times 10^{-4}$	8·63	$0{\cdot}20316 \times 10^{6}$
0·0211	7700	$0{\cdot}8401 \times 10^{-5}$	6·86	$0{\cdot}81694 \times 10^{6}$
0·01	53200	$0{\cdot}1294 \times 10^{-5}$	8·24	$0{\cdot}63671 \times 10^{7}$

invert; this is achieved by the consistent algorithm (14) for which we do not know how to prove the convergence compared with that of the BSOR algorithm. Nevertheless, we show numerically on the Laplace problem in the square $(0, 1) \times (0, 1)$ that the algorithm (14) converges as illustrated in Table II and in Figure 2 where the convergence history for several values of β is plotted.

Thus we have to take β as small as possible; in practice we set $\beta = 0{\cdot}1$ and $\omega = 1{\cdot}5$ for our problem. As solving (14) is embedded into the iterations of the Newton linearization , it is not

Table II

Mesh size	ω	β	Error	Number of iterations
		0	$0{\cdot}31 \times 10^{-3}$	13
0·1	1·41	0·1	$0{\cdot}36 \times 10^{-3}$	25
		0·5	$0{\cdot}46 \times 10^{-3}$	65
		1	$0{\cdot}48 \times 10^{-3}$	114
		0	$0{\cdot}15 \times 10^{-3}$	30
0·05	1·64	0·1	$0{\cdot}17 \times 10^{-3}$	78
		0·5	$0{\cdot}20 \times 10^{-3}$	243
		1	$0{\cdot}20 \times 10^{-3}$	449
		0	$0{\cdot}95 \times 10^{-4}$	83
0·02	1·84	0·1	$0{\cdot}98 \times 10^{-4}$	388
		0·5	$0{\cdot}99 \times 10^{-4}$	1486
		1	—	—

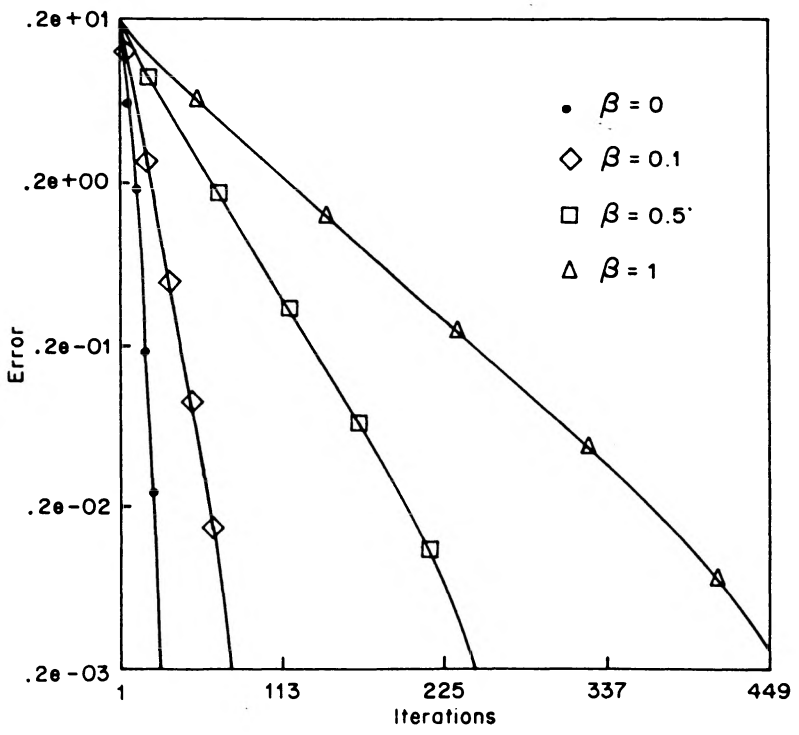


Figure 2. Convergence history of SBSOR algorithm with respect to BSOR algorithm for mesh size $h = 0{\cdot}05$

necessary to compute **X** very accurately, so a few iterations (5–10) of SBSOR is enough. Moreover, setting the parameter β to a good value improves the efficiency of the ICCG method and yields a very good solution rapidly; otherwise (β too small) the ICCG method cannot converge and the approximated $\mathbf{X}_n$ are spoiled.

VECTORIZATION OF PROGRAMS

The programs run on a CRAY-2 and the main part of the CPU time is devoted to the construction and inversion of system (13), so we shall speak only of the vectorization of these two actions. With the approximation of groups of variables described above, the assembling of the matrix is reduced to a product of an initial matrix and a vector; for instance, the term

$$\int_\Omega \frac{\partial \tilde{\rho}_h u_h}{\partial x} \frac{\partial \bar{\rho}_h u_h}{\partial x} \, dxdy$$

yields the matrix coefficients

$$u_i u_j \int_\Omega \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} \, dxdy$$

for $1 \leqslant i \leqslant I$ and some j corresponding to the basis functions ϕ_j whose support intersects with the support of ϕ_i. For a regular domain these indices j are always the same with respect to i and **S** is a multidiagonal matrix which can be easily computed knowing the initial matrices **DXX**, **DXY**, etc. representing the terms

$$\int_\Omega \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} \, dxdy, \qquad \int_\Omega \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial y} \, dxdy, \qquad \text{etc.;}$$

so we use the huge memory of CRAY-2 to store these initial matrices and then the assembling of the matrix at each iteration of the Newton linearization is written as follows.

For *i*diag = 1 to the number of diagonals of a block do:

nj = the position of the diagonal *i*diag with respect to the main diagonal.
For *i*dlib = 1 to *I* do:
*j*dlib = *i*dlib + *nj*,
E$_n$(*i*dlib, *i*diag) = *u*(*i*dlib) × (**DXX**(*i*dlib, *i*diag) × *u*(*j*dlib)
+ **DXY**(*i*dlib, *i*diag) × *v*(*j*dlib)) + . . . ,

End.
End

Thus the inner loop of large size is vectorized. For an irregular domain we can do the same with an outer loop to the number of non-zero terms and the use of the function GATHER to avoid the indirect addressing of *j*dlib in the sequence above. So in both cases the assembling is highly vectorizable.

For the inversion of the blocks in (14) the vectorization is *a priori* more difficult because the Cholesky factorization is a typical operation which involves dependency, but we benefit from the 5×5 block structure related to the five unknowns in dimension three. Indeed, each block $\mathbf{E}_n$ or $\mathbf{F}_n$ is divided into 25 equal sub-blocks and the Cholesky factorization is always vectorizable except when it occurs on the same sub-block; so about 80% of the factorization is vectorized and we reach a very good performance in terms of Mflops on the whole program.

NUMERICAL RESULTS

In this section we present results illustrating the influence of mesh size, angle of attack and Mach number at infinity on the solution. In Figure 3 we show the two meshes in a cross-plate section; the coarse one is only $14 \times 9 \times 9$ and the second one is $28 \times 18 \times 18$ cells around the same flat plate without thickness. The numbers of variables are respectively 8250 and 55100.

The solution for $M_\infty = 0{\cdot}7$ and $\alpha = 15°$ is shown in Figure 4–7 in cross-flow sections and on the upper side of the plate; we see that the vortex structure rolls up at the tip of the plate with a typical

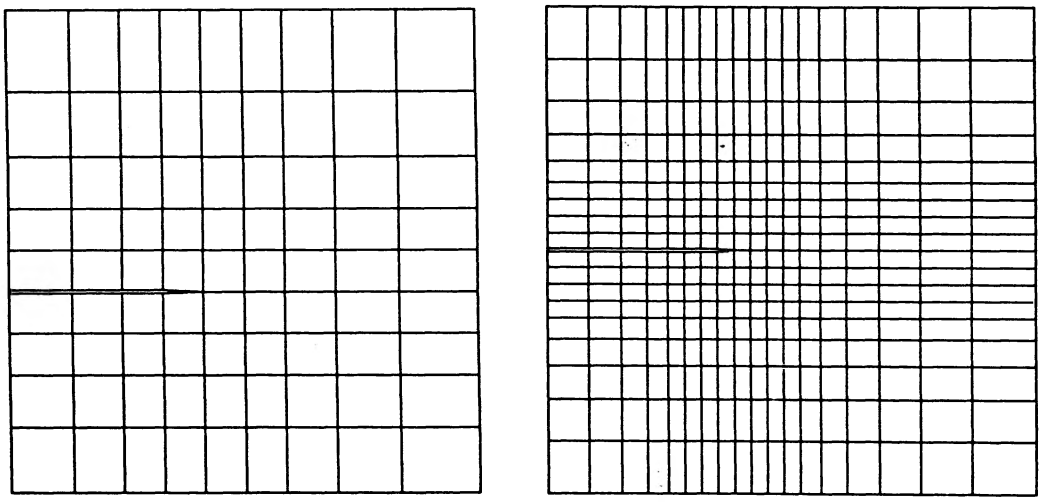


Figure 3. Meshes in a cross-plate section

(a) $14 \times 9 \times 9$ cells

(b) $28 \times 28 \times 18$ cells

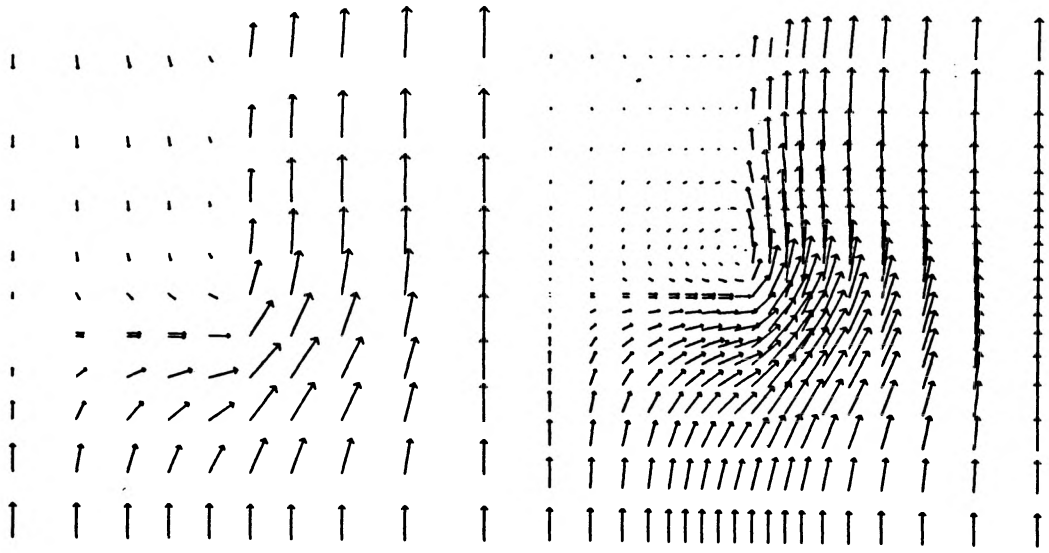


Figure 4. Comparison of the cross-flow velocities at 60% of the plate for the two meshes ($M_\infty = 0{\cdot}7$, $\alpha = 15°$)

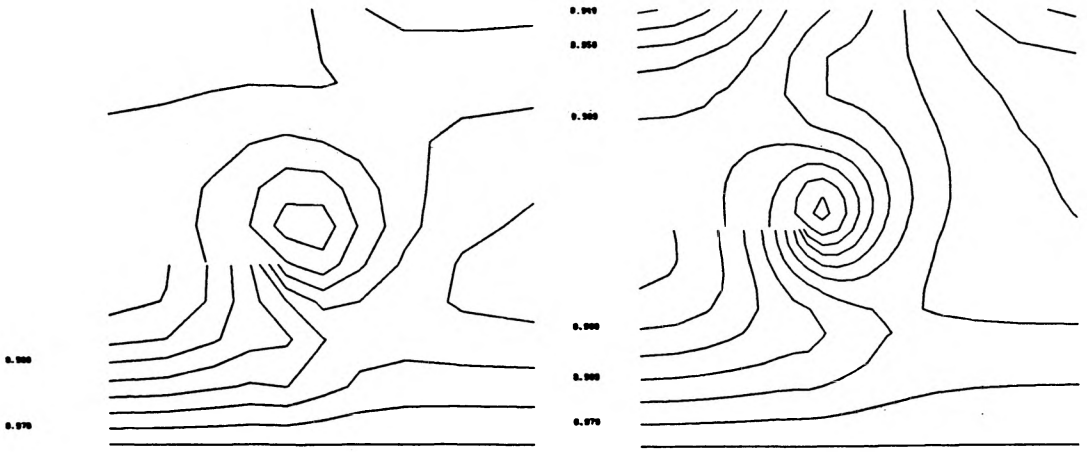


Figure 5. Comparison of isobars in a cross-plane at 60% of the plate for the two meshes ($M_\infty = 0{\cdot}7$, $\alpha = 15°$)

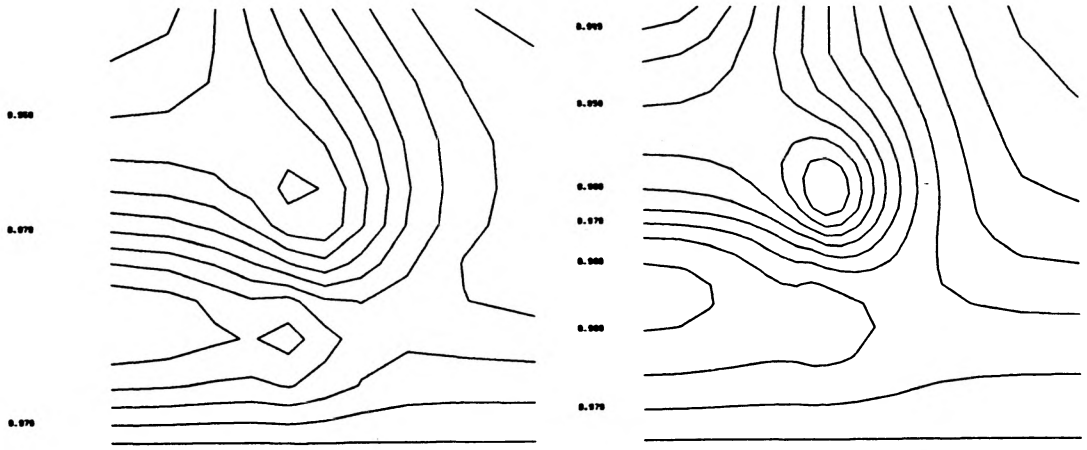


Figure 6. Comparison of isobars in a cross-plane at the trailing edge for the two meshes ($M_\infty = 0{\cdot}7$, $\alpha = 15°$)

aspect of the isobar lines as shown in Figure 7. Of course the solution is better represented by the finer mesh, although the vortex centre is located at the same place in both cases. Thus our method is able to capture the main behaviour of the flow even with a coarse mesh.

We now see the influence of the angle of attack by comparing Figures 4–7 for the finer mesh with Figures 8–10 where the solution for a higher angle of attack, $\alpha = 30°$, is plotted. One can see clearly that the size of the vortex structure and the amplitude of the pressure jump on both sides of the plate are linked directly to the angle of attack.

Finally, we have done computations for supersonic flows with $M_\infty = 2$ and $\alpha = 10°$; we notice that the vortical phenomenon is again well captured, as shown in Figures 11 and 12. In this case the vortex is closer to the plate and perturbations in pressure and flow direction are visible near the upper boundary which is located not very far away.

It is interesting to note that solving the steady energy equation does not imply that the total enthalpy is constant everywhere, but the results indicate that the error on total enthalpy does not exceed 3% and that the maximum is reached in the vortex core.

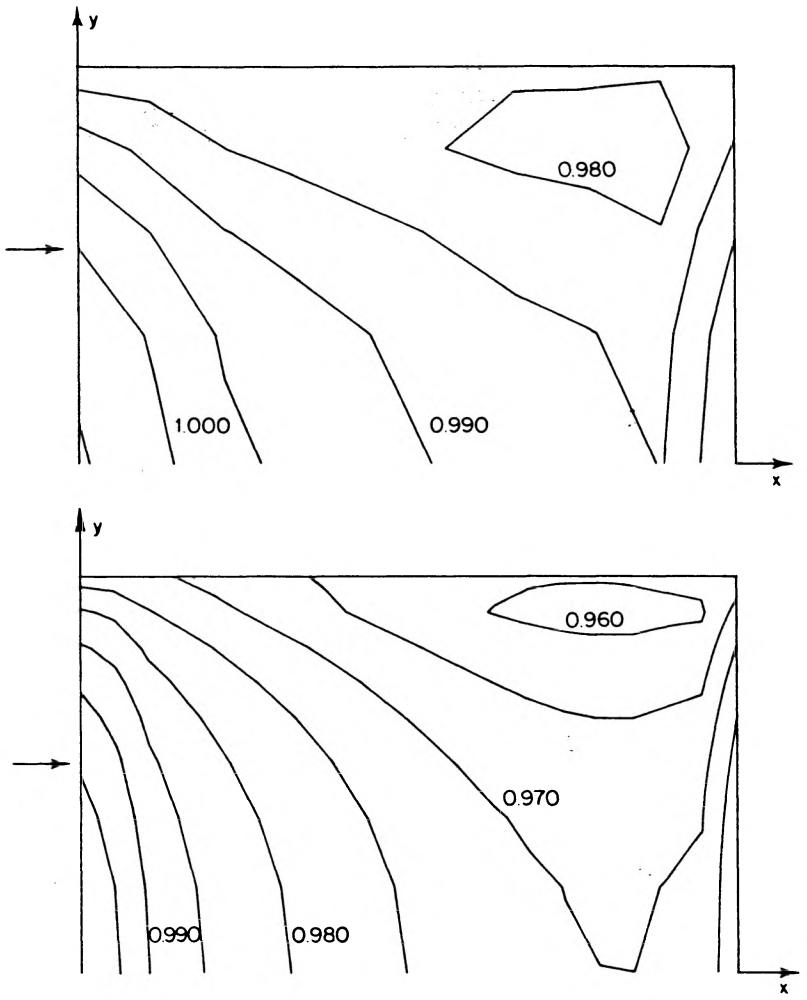


Figure 7. Comparison of isobars on the upper plate for the two meshes ($M_\infty = 0.7$, $\alpha = 15°$); pressure of the incoming flow, $p_\infty = 0.975$

CONCLUSIONS

The method presented in this paper is able to capture vortical effects in subsonic flows as well as in supersonic flows. We have shown from a mathematical point of view that the least squares formulation is equivalent to the original first-order Euler system. We have imposed all the flow variables at the entrance boundary and no condition at the exit boundary in both subsonic and supersonic flows, and we were able to compute solutions for various Mach numbers and angles of attack. However, in supersonic flow, when the incidence is larger than the characteristic cone half-angle, the boundary conditions coincide with the usual ones.

All the computations have been performed without the use of artificial viscosity. The centred finite element scheme of second-order accuracy employed in the solution procedure is stable owing to its inherent dissipative properties, which may explain why there is no need for a Kutta condition at the plate edges to provoke the shedding of vorticity.

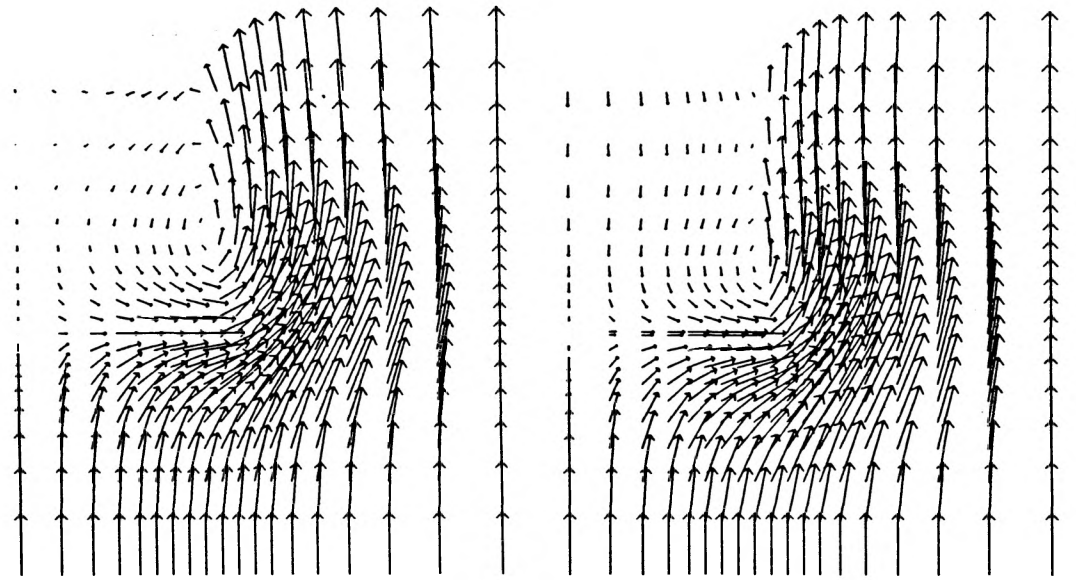


Figure 8. Cross-flow velocities at 60% of the plate and at the trailing edge for $M_\infty = 0{\cdot}7$ and $\alpha = 30°$

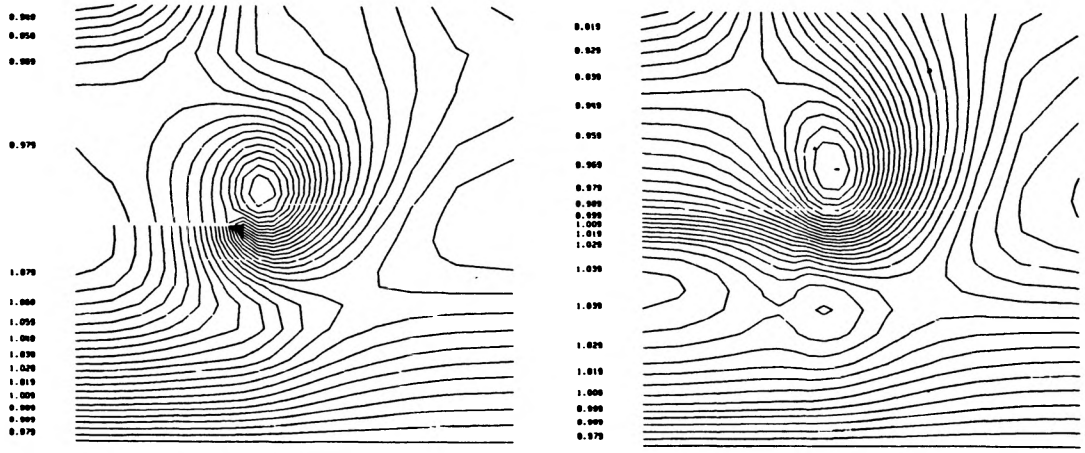


Figure 9. Isobars in cross-planes at 60% of the plate and at the trailing edge for $M_\infty = 0{\cdot}7$ and $\alpha = 30°$

Moreover, we still get a vortex flow in a hypersonic case with $M_\infty = 5$ and $\alpha = 10°$ as illustrated in Figure 13; this shows the robustness of the method.

ACKNOWLEDGEMENTS

This work was performed with financial support from DRET and under grants from C^2VR, Ecole Polytechnique, France.

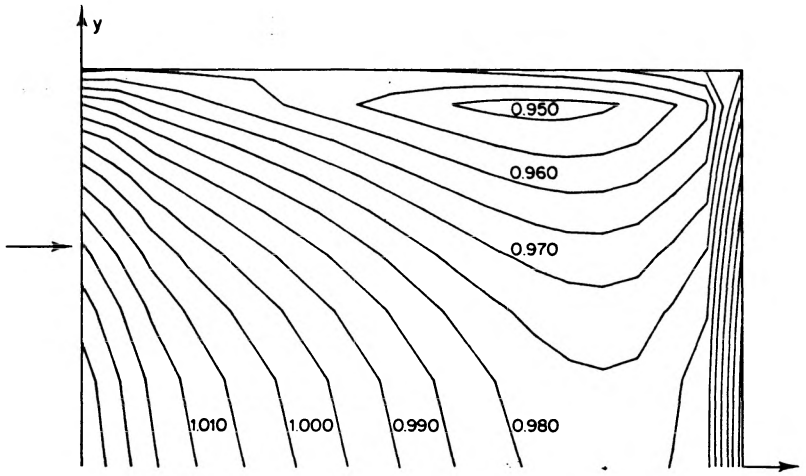


Figure 10. Isobars on the upper plate for $M_\infty = 0{\cdot}7$ and $\alpha = 30°$

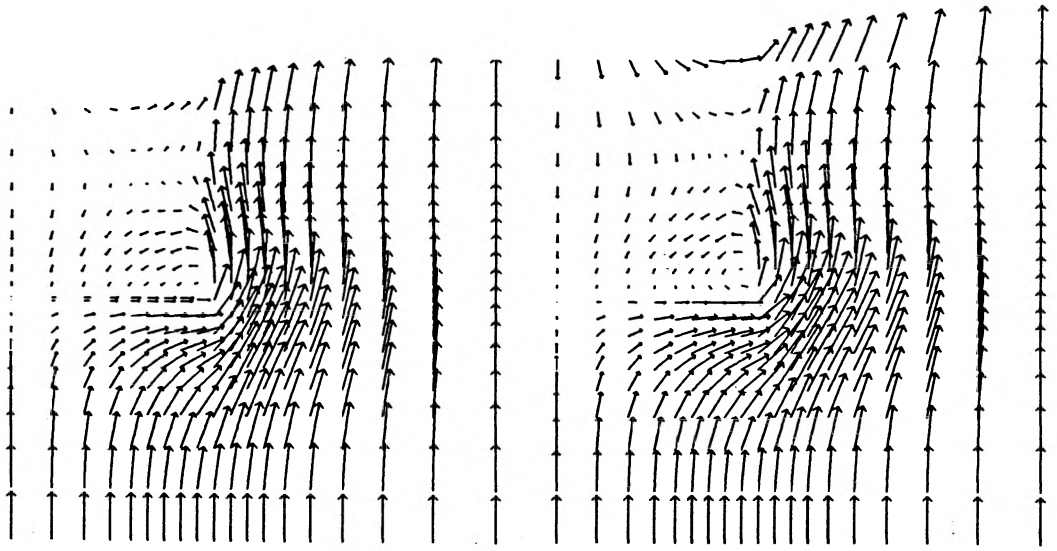


Figure 11. Cross-flow velocities at 60% of the plate and at the trailing edge for $M_\infty = 2$ and $\alpha = 10°$

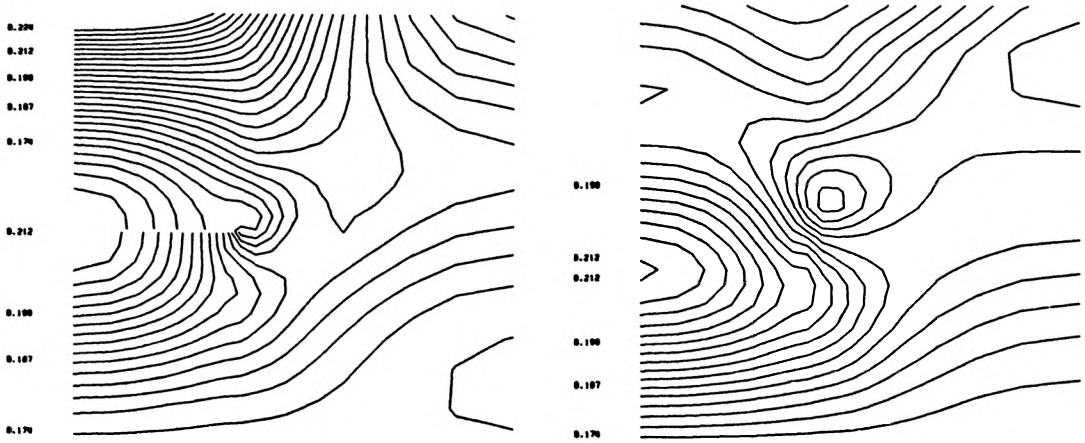


Figure 12. Isobars in cross-planes at 60% of the plate and at the trailing edge for $M_\infty = 2$ and $\alpha = 10°$

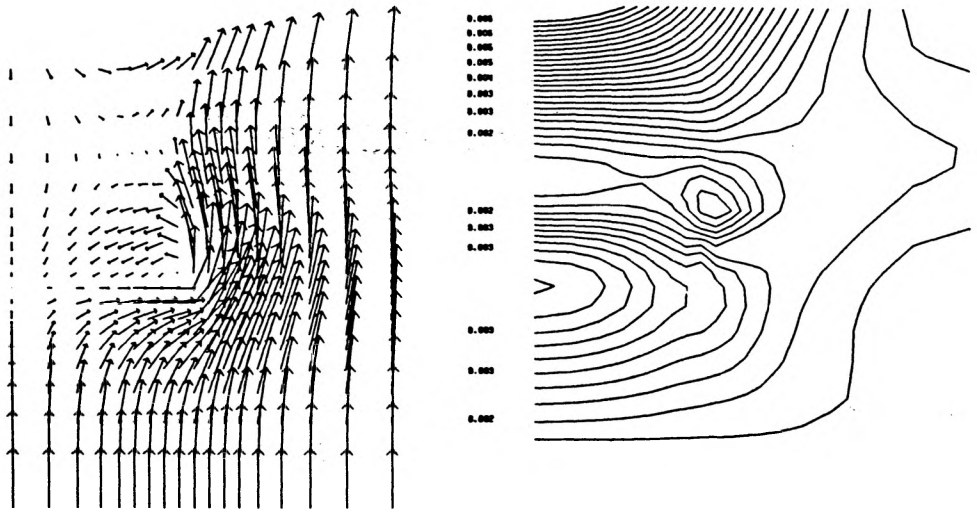


Figure 13. Hypersonic solution at the trailing edge ($M_\infty = 5$, $\alpha = 10°$)

REFERENCES

1. A. Rizzi and Ch. J. Purcell, 'On the computation of transonic leading-edge vortices using the Euler equations', *J. Fluid Mech.*, **181**, 163–195 (1987).
2. O. A. Kandil, A. H. Chuang and J. M. Shifflette, 'Finite-volume Euler and Navier–Stokes solvers for three-dimensional and conical vortex flows over delta wings', *AIAA Paper 87-0041*, 1987.
3. R. W. Newsome, 'Euler and Navier–Stokes solutions for flow over a conical delta wing', *AIAA J.* **24**, 552–561 (1986).
4. Ch. H. Bruneau, J. J. Chattot, J. Laminie and R. Temam, 'Numerical solutions of the Euler equations with separation by a finite element method', *9th ICNMFD*, Saclay; *Lecture Notes in Physics 218*, 1984, pp. 121–126.
5. Ch. H. Bruneau, J. J. Chattot, J. Laminie and R. Temam, 'Computation of vortex flows past a flat plate at high angle of attack', *10th ICNMFD*, Beijing; *Lecture Notes in Physics 264*, 1986, pp. 134–140.
6. G. W. Hedstrom, 'Non reflecting boundary conditions for nonlinear hyperbolic systems', *J. Comput. Phys.*, **30**, 222–237 (1979).
7. J. Cullum and R. A. Willoughby, 'Computing eigenvalues of very large symmetric matrices—an implementation of a Lanczos algorithm with no reorthogonalization', *J. Comput. Phys.*, **44**, 329–358 (1981).

COMPUTATION OF HYPERSONIC VORTEX FLOWS WITH AN EULER MODEL *

Charles-Henri Bruneau, Jacques Laminie
Laboratoire d'Analyse Numérique, Bâtiment 425
Université Paris-Sud, 91405 Orsay, France

Jean-Jacques Chattot
Aérospatiale, 78133 Les Mureaux, France

* Work performed with financial support of DRET and under grants from the CCVR Ecole Polytechnique, France.

ABSTRACT

The variational approach of the steady Euler equations presented at the 10th ICNMFD [1] is extended to the treatment of supersonic and hypersonic flows by introducing the energy equation in the least-squares formulation. The approximation is made with cubic or prismatic linear finite elements and the results are presented for flows around a rectangular flat plate or a thin delta wing for various Mach numbers and angles of attack. They show the occurrence of vortical flows on the upper surface of the wings due to the sharp edges.

The simulation code is run on a CRAY2 and the computing time is kept to a minimum by using the large storage capacity available, an efficient vectorization and a full multigrid procedure.

INTRODUCTION

The purpose of this work is to capture the vortex structure developing around wings with sharp leading edges for a wide range of Mach numbers and angles of attack, to see how the vortex sheet is related to the Mach number at infinity, the angle of attack and the geometry.

This study concerns only simplified geometries without thickness : a rectangular flat plate of small aspect ratio and a thin delta wing with sweep angle $\beta = 70°$. A structured mesh with cubic or prismatic cells is used to improve the efficiency of the code and due to the symmetry only half of the flow is computed.

The flow at infinity is given at the entrance sections of the domain and the tangency condition is imposed on both sides of the wings as done in [2].

DESCRIPTION OF THE METHOD

The full steady Euler equations are written in conservative partial differential form in a 3D domain Ω as follows :

$$
\begin{aligned}
A &= \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} + \frac{\partial \rho w}{\partial z} = 0 \\
B &= \frac{\partial \rho u^2}{\partial x} + \frac{\partial \rho uv}{\partial y} + \frac{\partial \rho uw}{\partial z} + \frac{\partial (\gamma-1)\rho e}{\partial x} = 0 \\
C &= \frac{\partial \rho uv}{\partial x} + \frac{\partial \rho v^2}{\partial y} + \frac{\partial \rho vw}{\partial z} + \frac{\partial (\gamma-1)\rho e}{\partial y} = 0 \\
D &= \frac{\partial \rho uw}{\partial x} + \frac{\partial \rho vw}{\partial y} + \frac{\partial \rho w^2}{\partial z} + \frac{\partial (\gamma-1)\rho e}{\partial y} = 0 \\
E &= \frac{\partial (\gamma\rho e + \tfrac{1}{2}\rho q^2)u}{\partial x} + \frac{\partial (\gamma\rho e + \tfrac{1}{2}\rho q^2)v}{\partial y} + \frac{\partial (\gamma\rho e + \tfrac{1}{2}\rho q^2)w}{\partial z} = 0
\end{aligned}
\qquad (1)
$$

where the density ρ, the velocity $\vec{q} = (u,v,w)^T$ and the internal energy e are the unknowns ; γ is the ratio of specific heats and $q = |\vec{q}|$. The main difference with [1] is that the last equation, the conservation of energy, is solved as a partial differential equation instead of using the algebraic Bernoulli's equation. Consequently e is chosen as an unknown and the pressure p is replaced by $(\gamma-1)\rho e$ on one hand, and the whole system is linearized by Newton's method for the five unknowns on the other hand. This also means that the total enthalpy H is not imposed everywhere and that its value $H = \gamma e + \frac{1}{2} q^2$ is computed at each point. We show the variation of this quantity in the numerical results.

Starting with the non linear problem (1) for the unknowns $U=(\rho,u,v,w,e)^T$ we solve at each iteration of Newton's method a first order linear system for the unknowns $\tilde{U} = (\tilde{\rho},\tilde{u},\tilde{v},\tilde{w},\tilde{e})^T$ and we update the solution in the formula $U^{n+1} = U^n - \tilde{U}$ where n is the index of Newton step and the sequence U^n converges to the solution of (1).

This linear system is solved through the minimization of a least-squares functional which transforms the first order hyperbolic system into a second order parabolic one. If $A_\ell, B_\ell, C_\ell, D_\ell$ and E_ℓ are the five linear equations the functional reads

$$J(\tilde{U}) = \int_\Omega [A_\ell^2 + B_\ell^2 + C_\ell^2 + D_\ell^2 + E_\ell^2]dxdydz.$$

Then the resulting variational problem is symmetric and also the linear system we need to inverse at each Newton iteration.

For both geometries the domain is a rectangular hexaedron and a block-successive over-relaxation (BSOR) method is used to solve the linear system where each block represents the unknowns of a vertical plane (y,z) with the same number of points. One block is inverted by

an incomplete Cholesky preconditioned conjugate gradient (ICCG) method.

APPROXIMATION

The groups of conservatives variables are approximated by a finite element method, for instance the approximation of ρu is given by

$$\rho_h u_h = \sum_{i=1}^{I} (\rho u)_i \phi_i$$

where ϕ_i is the basis of the approximate space. The finite element is either a linear cubic element spanned by the set {1,x,y,z,xy,xz,yz,xyz} or a linear prismatic element spanned by the set {1,x,y,z,xz,yz}. That means in this last case that the triangular faces are in the planes (x,y) and that the other faces are rectangular. This is very convenient to approximate the flow around the delta wing. Nevertheless, a test with the rectangular flat plate shows that the results are about the same with both approximations.

NUMERICAL ASPECTS AND EFFICIENCY

To improve the efficiency of the whole process we use a full multigrid procedure (FMG) on a set of successive grids (for instante 14×16×12, 28×32×24 and 56×64×48 around the wing). This procedure consists in interpolating the results from a coarse grid to the finer one, which divides the computing time by a factor two.

Then we use the large memory of the CRAY2 to store the matrices, independent of the solution, representing terms like

$$\int_\Omega \frac{\partial \phi_i}{\partial x} + \frac{\partial \phi_j}{\partial x} \, dxdydz, \quad 1 \leqslant i,j \leqslant I.$$

So at each iteration of Newton's linearization the assembling of the linear system is reduced to a product of these matrices by some vectors. This last operation is highly vectorizable, and this part of the calculus is kept to a minimum.

Finally 80% of the inversion of the linear system is also vectorized and a very good efficiency is reached for the whole computation in terms of CPU time. Indeed the solution for 80,000 unknowns at Mach number at infinity $M_\infty = 0.7$ and angle of attack $\alpha = 10°$ converges in 50 iterations and is obtained in about half an hour on one processor of the CRAY2.

NUMERICAL RESULTS

The results show the occurrence of a vortex sheet that rolls up at the tip of the plate or at the leading edge of the delta wing with a conical shape. For a given Mach number at infinity this vortex grows with the angle of attack. On the contrary, for a given angle of attack, the vortex tightens as the Mach at infinity increases.

The results on the thin delta wing for Mach numbers at infinity $M_\infty = 0.7$ or $M_\infty = 2$ and angle of attack $\alpha = 10°$ show the same characteristics than those available in the literature (for instance [3] to [5]) as shown on the figure.

The accuracy of the resolution of the system including the energy equation is indicated by the variation on the total enthalpy that does not exceed 3% ; moreover the maximum is reached in the vortex core.

CONCLUSION

The extended method presented above is able to capture vortex flows around wings with sharp edges in subsonic as well as in supersonic or hypersonic regime (the tests include a computation at $M_\infty = 5$).

Nevertheless the computation with 200,000 points requires about 100 megawords of storage and 8 hours of CPU time.

REFERENCES

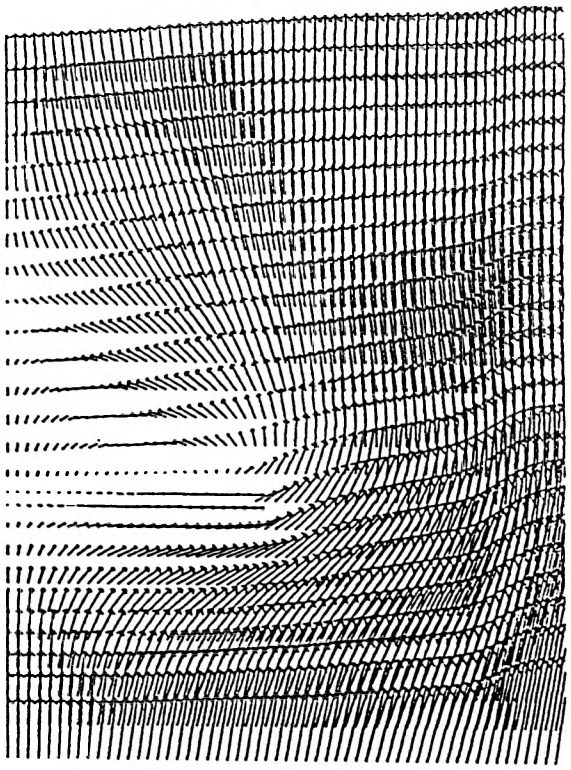
[1] Ch.H. Bruneau, J.J. Chattot, J. Laminie and R. Temam, Computation of vortex flow past a flat plate at high angle of attack, 10th ICNMFD, Beijing, 1986, Lecture Notes in Physics 264.

[2] Ch.H. Bruneau, J. Laminie and J.J. Chattot, Computation of 3D vortex flows past a flat plate at incidence through a variational approach of the full steady Euler equations, to appear in Int. J. Num. Meth. in Fluids.

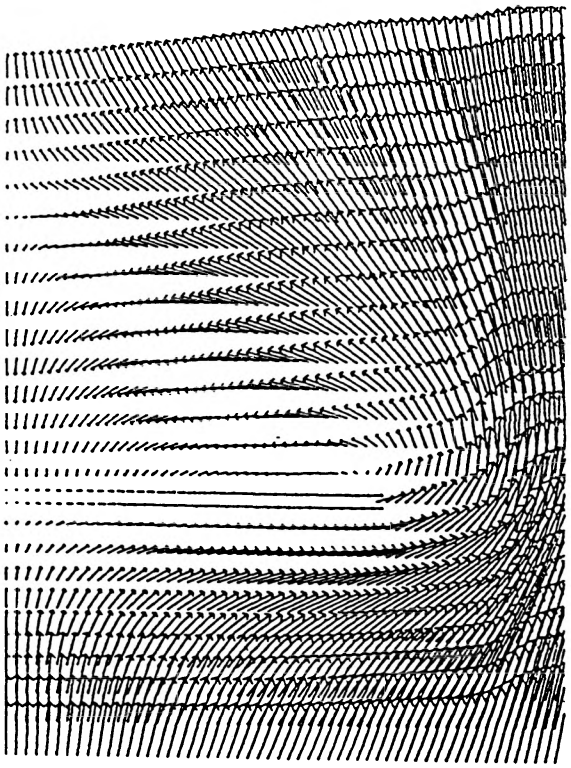
[3] K. Powell, E. Murman, E. Perez and J. Baron, Total pressure loss in vortical solutions of the conical Euler equations, AIAA paper 85-1701.

[4] C. Koeck, Computation of three-dimensional flow using the Euler equations and a multiple-grid scheme, Int. J. Num. Meth. in Fluids, 5, 1985.

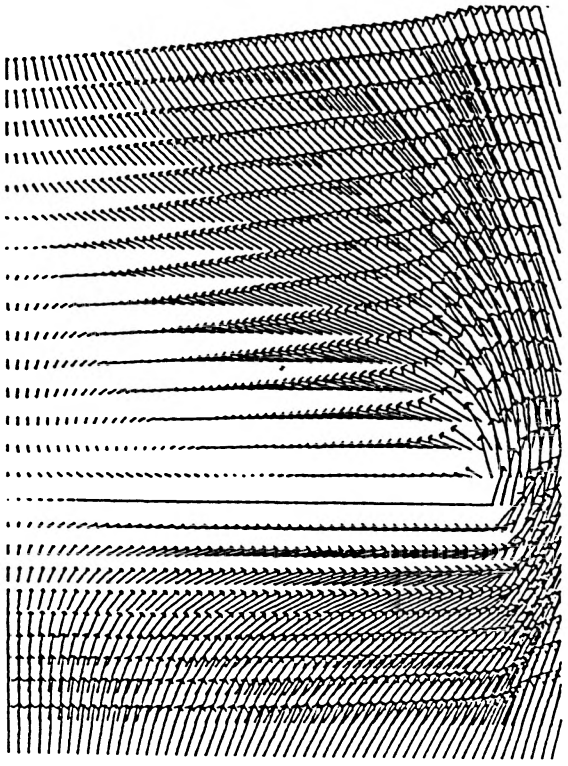
[5] L.E. Eriksson and A. Rizzi, Computation of vortex flow around wings using the Euler equations, 4th GAMM Conf. on Num. Meth. in Fluid Mech., 1981.



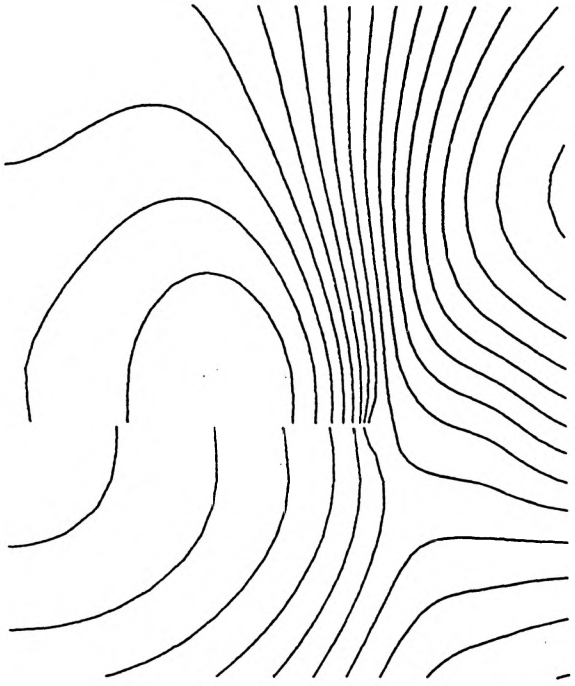
50%



75%



100%



Pressure contours

75%

$M_\infty = 2.0 \quad \alpha = 10°$

Solution around the delta wing with the 56x64x48 grid

Computation of Hypersonic Flows Round a Blunt Body in 2D with Euler Model

Laboratoire d'Analyse Numérique
Université Paris-Sud, 91405 Orsay (France)

Abstract.

The 2D steady Euler equations in density and velocity variables are solved by a variational method based on a Newton linearization and a least-squares embedding. Flows around a blunt body are computed for hypersonic regimes. The strong bow shock occuring in front of the nose is captured in a first step by means of an entropy corrector added to the least-squares functional. Then a mesh adaptation and displacement procedure is used to fit the mesh to the shock and to give the good location of the shock.

Numerical results are presented for Mach number at infinity equal to 3 and 8 round a circular cylinder and an ellipse.

Introduction.

In former works we have studied the full steady Euler equations written in conservative form by two ways. In the first approach Bernoulli's equation is used to write a fixed point algorithm on the density ; then a Newton linearization and a least-squares minimization yield the solution of the remaining system in velocity and pressure at each iteration of the fixed point algorithm. The problem of this approach is that the fixed point algorithm on the density is stable only in subsonic regime. Consequently the method is restricted to subsonic flows or weakly transonic flows by adding an artificial viscosity term (cf. [1]). In the second approach we proposed to keep the whole system in its partial differential form and to solve directly the system in density, velocity and internal energy by Newton method and least-

squares embedding ; this second method is stable in subsonic as well as in supersonic or hypersonic regime (cf. [2]).

In this paper we study a third approach using Bernoulli's equation to eliminate the pressure in the equations of conservation of momentum. This method is as stable as the previous one and its advantage is to reduce the system of four equations in two dimension to a system of three equations. Consequently it is cheaper in terms of memory and CPU time.

The purpose of this work is to show that the embedding method with its centered approximation is able to capture and represent strong shocks. To validate this affirmation we choose to compute hypersonic flows round blunt bodies. The numerical tests concern the flow round a circular cylinder or an ellipse for Mach numbers at infinity equal to 3 and 8. In the case of the circular cylinder the results are compared with those given by the theory and experiments ; in both cases the results are compared with other numerical investigations.

In the first section we present the third approach with its properties and its approximation. The following section is devoted to the study of an entropy corrector that permits to select the physical solution. Indeed an entropy inequality must be added to the set of Euler equations to avoid expansion shocks ; starting with this inequality we construct an entropy corrector that is included in the least-squares functional and eliminates non physical solutions. Then a mesh adaptation and displacement is presented in the third section. The aim of this procedure is, on one hand, to fit the mesh to the shock front and, on the other hand, to position the shock with respect to the body by reducing the diffusion in the incoming region of the flow before the shock. Finally we discuss numerical results for high Mach numbers up to 8 for which the Euler model seems to be still valid (cf. [3]).

1. Description of the method.

Starting with the full steady Euler equations written in conservative form we use Bernoulli's equation to get the pressure p as a function of the density ρ and the velocity

$\vec{q} = (u, v)^T$. So we get :

$$p = \frac{\gamma - 1}{\gamma}\rho\left(H - \frac{q^2}{2}\right)$$
$$= \gamma_1\rho - \gamma_2\rho q^2$$

where q is the modulus of $\vec{q}$, H is the total enthalpy, γ is the ratio of the specific heats and γ_1 and γ_2 are two positive constants ; for the air we have $\gamma = 1.4$, $H = 3.0$, $\gamma_1 = 6/7$ and $\gamma_2 = 1/7$. Replacing $\vec{\nabla}p$ in the equations of conservation of momentum we get linear terms on the density and the same nonlinear terms than those already present in the equations. The three equations of conservation of mass and momentum read :

$$\begin{aligned} A &= \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} = 0 \\ B &= \frac{\partial \gamma_1 \rho}{\partial x} + \frac{\partial (1-\gamma_2)\rho u^2}{\partial x} + \frac{\partial \rho uv}{\partial y} - \frac{\partial \gamma_2 \rho v^2}{\partial x} = 0 \\ C &= \frac{\partial \gamma_1 \rho}{\partial y} - \frac{\partial \gamma_2 \rho u^2}{\partial y} + \frac{\partial \rho uv}{\partial x} + \frac{\partial (1-\gamma_2)\rho v^2}{\partial y} = 0 \end{aligned} \tag{1.1}$$

For any function φ constant on the characteristic curves the characteristic polynomial is given by

$$det = \frac{\rho^2 \varphi_y^3}{\gamma}(-\tau u + v)[(-\tau u + v)^2 - a^2(\tau^2 + 1)]$$

where τ is the slope of the characteristic curves, φ_y the second component of $\vec{\nabla}\varphi$, and a is the local speed of sound. The discriminant of the second order polynomial is $\Delta' = a^4(M^2 - 1)$ where M is the local Mach number. Thus, in addition to the streamline $\tau = \frac{v}{u}$ there are two other zeros of *det* when the flow is supersonic.

We solve the first order nonlinear hyperbolic system (A, B, C) for the unknown $U = (\rho, u, v)^T$ by means of Newton linearization as follows :

$$\begin{aligned} &U_0 \ given \\ &U_{m+1} = U_m - \tilde{U} \ ; \ m = 0, ..., p \end{aligned} \tag{1.2}$$

where $\tilde{U} = (\tilde{\rho}, \tilde{u}, \tilde{v})^T$ is the solution of the first order linear hyperbolic system (A_m, B_m, C_m) deduced from (A, B, C) at each Newton step.

Then we use a least-squares embedding which consists in minimizing the functional :

$$J(\tilde{U}) = \frac{1}{2}\int_\Omega (A_m^2 + B_m^2 + C_m^2) dxdy$$

where Ω is the computational domain around the blunt body. We refer the reader to [2] for more details on the method ; in this work we focus specially on the new difficulties occuring when strong shocks are involved.

The domain Ω is limited by the body Γ_b and by the external boundary $\Gamma_i \cup \Gamma_o$ where Γ_i is the entrance section ($\vec{q}_\infty \cdot \vec{n} < 0$) and Γ_o is the exit section ($\vec{q}_\infty \cdot \vec{n} > 0$); $\vec{n}$ denotes always the unit normal vector pointing outside of the domain. On the body we impose the usual tangency condition $\vec{q} \cdot \vec{n} = 0$ that is added in the functoinal $J(\tilde{U})$ as :

$$\frac{1}{2} \int_{\Gamma_b} [(\vec{q}_m - \vec{q}) \cdot \vec{n}]^2 d\sigma \tag{1.3}$$

Indeed, as we use a least-squares minimization, the first order equations are implicitly imposed on the boundary but the tangency condition is not contained in the variational formulation. The condition (1.3) is very efficient and the solution of the whole process satisfies the tangency condition as soon as U_0 in (1.2) satisfies : $\vec{q}_0 \cdot \vec{n} = 0$ on Γ_b. On the external boundary we use the classical conditions for supersonic flows : U is given equal to the flow at infinity on Γ_i and is left free on Γ_o; what yields in our formulation $U_0 = U_\infty$ and $\tilde{U} = 0$ on Γ_i.

The approximation is done, like in [2], by means of linear quadrilateral finite elements applied to the groups of conservative variables, for instance we write :

$$\rho_h u_h = \sum_{i=1}^{I} (\rho u)_i \Phi_i$$

where the subscript h denotes the approximate variables, I is the number of free nodes of the discretization and $(\Phi_i)_{i=1,\ldots,I}$ is the basis of the approximate space.

The inversion of the resulting linear system is achieved by means of the incomplete Cholesky conjugate gradient method (ICCG). As we use a structured mesh and the $Q1$ finite element, each block of the matrix has nine diagonals. To improve the efficiency of the ICCG method the four neighbouring diagonals inside the profile are added in the factorization (cf. [4]). Consequently the solution of the linear system is reached with accuracy after a few iterations. In this section we have presented the main lines of our method, more details can be found in [2]. In the two next sections we propose an algorithm to capture and fit strong shocks.

2. Entropy corrector.

Applying the method above we can select bad solutions. For instance round a cylinder with a coarse mesh of 5×5 cells the computation for Mach numbers at infinity $M_\infty = 8$ gives a converged solution in 14 iterations of Newton method ; but the values of the density and the velocity are such that the pressure is negative at a few points ! This is due to the fact that the initial data U_0 is far from the solution, then Newton linearization select a nonphysical solution. It is well-known that hyperbolic systems need an entropy inequality to avoid expansion shocks and select the good solution. For Euler system this inequality can be written in the form

$$\frac{\partial \rho s}{\partial t} + div\ \rho \vec{q} s \geq 0 \tag{2.1}$$

where $s = \frac{\gamma p}{\rho^\gamma}$ is the entropy we use because it is more convenient to normalize the flow at infinity ; we always take $s_\infty = 1$. For steady flows equation (2.1) reduces to

$$div\ \rho \vec{q} s \geq 0$$

and using the conservation of mass ($div\ \rho \vec{q} = 0$) we finally get as ρ is always positive :

$$\vec{q} \cdot \vec{\nabla} s \geq 0 \tag{2.2}$$

This formula completes the system (1.1) ; the difficulty is to include such a condition in the method. What is important is to control the entropy in the hypersonic region Ω_∞ before the shock to avoid expansion phenomena. In this region the flow must be constant and equal to the flow at infinity ; by taking $\vec{q}_\infty$ in the x direction we can reduce (2.2) to :

$$\frac{\partial s}{\partial x} \geq 0\ in\ \Omega_\infty \tag{2.3}$$

where s is the nonlinear function of U given by :

$$s(U) = \frac{\gamma p}{\rho^\gamma} = (\gamma - 1)\rho^{1-\gamma}\left(H - \frac{u^2 + v^2}{2}\right)$$

As U is regular in Ω_∞ we can rewrite (2.3) in nonconservative form as :

$$S(U) \geq 0\ in\ \Omega_\infty \tag{2.4}$$

where

$$S(U) = -(\gamma-1)^2\ \rho^{-\gamma}\frac{\partial \rho}{\partial x}\left(H - \frac{u^2+v^2}{2}\right) - (\gamma-1)\ \rho^{1-\gamma}\left(u\frac{\partial u}{\partial x} + v\frac{\partial v}{\partial x}\right)$$

To include condition (2.4) in Newton method we introduce at each step m the linearized function :

$$S_m(U) = -(\gamma-1)^2\ \rho_m^{-\gamma}\frac{\partial \rho}{\partial x}\left(H - \frac{u_m^2+v_m^2}{2}\right) - (\gamma-1)\ \rho_m^{1-\gamma}\left(u_m\frac{\partial u}{\partial x} + v_m\frac{\partial v}{\partial x}\right)$$

and require that the new iterate U_{m+1} satisfies :

$$S_m(U_{m+1}) \geq 0 \ in\ \Omega_\infty$$

that is

$$S_m(U_m - \tilde{U}) = 0 \ in\ \Omega_\infty.$$

This last condition is then included in the least-squares minimization by adding to the functional $J(\tilde{U})$ the term

$$\frac{C_s}{2}\int_{\Omega_\infty}[Y(s_m)S_m(U_m-\tilde{U})]^2 dxdy$$

where

$$C_s \ is\ a\ positive\ constant$$

$$s_m = s(U_m)$$

$$Y(s_m) = \begin{cases} 1 \ \ if\ s_m < s_\infty \\ \\ 0 \ \ otherwise \end{cases}$$

We see that this entropy corrector is applied only when the entropy is less that s_∞ to avoid the expansion shock. It seems that local decreases of the entropy during the iterations do not play a significant role in the selection of the solution as long as the entropy is greater than s_∞. Thus the bilinear term :

$$C_s\int_{\Omega_\infty} Y(s_m)\left[(\gamma-1)^2\rho_m^{-\gamma}\frac{\partial\tilde{\rho}}{\partial x}\left(H-\frac{u_m^2+v_m^2}{2}\right)+(\gamma-1)\rho_m^{1-\gamma}\left(u_m\frac{\partial\tilde{u}}{\partial x}+v_m\frac{\partial\tilde{v}}{\partial x}\right)\right]$$
$$\left[(\gamma-1)^2\rho_m^{-\gamma}\frac{\partial\bar{\rho}}{\partial x}\left(H-\frac{u_m^2+v_m^2}{2}\right)+(\gamma-1)\rho_m^{1-\gamma}\left(u_m\frac{\partial\bar{u}}{\partial x}+v_m\frac{\partial\bar{v}}{\partial x}\right)\right]dxdy$$

and the linear term :

$$C_s \int_{\Omega_\infty} Y(s_m) \left[(\gamma-1)^2 \rho_m^{-\gamma} \frac{\partial \rho_m}{\partial x} \left(H - \frac{u_m^2 + v_m^2}{2} \right) + (\gamma-1) \rho_m^{1-\gamma} \left(u_m \frac{\partial u_m}{\partial x} + v_m \frac{\partial v_m}{\partial x} \right) \right]$$
$$\left[(\gamma-1)^2 \rho_m^{-\gamma} \frac{\partial \bar{\rho}}{\partial x} \left(H - \frac{u_m^2 + v_m^2}{2} \right) + (\gamma-1) \rho_m^{1-\gamma} \left(u_m \frac{\partial \bar{u}}{\partial x} + v_m \frac{\partial \bar{v}}{\partial x} \right) \right] dxdy$$

are added to the variational formulation ($\bar{U} = (\bar{\rho}, \bar{u}, \bar{v})^T$ denotes the test functions vector).

The only point remaining is the determination of C_s, unfortunately we do not have any explicit formula. C_s depends on both the Mach number at infinity M_∞ and the mesh size h $(C_s(M_\infty, h))$· C_s decreases with h almost linearly ; when h is divided by two C_s is divided by about thirty. The dependence on M_∞ is less clear but C_s decreases drastically as M_∞ increases ; for instance on a coarse mesh around the cylinder (5×5 cells) we take $C_s = 10^{-2}$ for $M_\infty = 3$ and $C_s = 10^{-5}$ for $M_\infty = 8$.

In the numerical computations we use a multigrid procedure to improve the efficiency of the method ; we solve the problem on a coarse mesh, then we extend the solution to a finer grid with a mesh size twice as fine as the previous one to give a good predictor on this finer grid. This procedure is repeated until the finest grid is reached. The remarks on the constants C_s above show that the entropy corrector is particularly important on the coarsest mesh to select the good solution and that its influence vanishes when very fine grids are concerned.

On figures 1 and 2 are plotted the solution for $M_\infty = 3$ around the circular cylinder for two consecutive grids ; these results show that the shock is captured within a cell. However we notice some local perturbations due to the fact that the shock front is not aligned with the mesh. Consequently the shock has to jump from one row of cells to another ; these jumps imply small recirculation bubbles which disturb the solution and sometimes introduce some oscillations in the convergence history. To avoid this problem we propose a mesh adaptation to fit the shock once it has been captured by the entropy corrector.

3. Mesh adaptation.

In this section we present the adaptive procedure used to fit the shock ; it is quite

simple and well suited for a structured mesh. We do not give here a general method and refer to [5] for the last developments on adaptive mesh.

We adapt the mesh in two steps : at first we localize the captured shock and construct a regular front shock by means of a global polynomial, then we move this front to reduce the diffusion in the Ω_∞ region according to the variations of the entropy.

For the sake of simplicity we present the method for the cylinder whose center is located at the origin. As the captured shock jumps from one row of cells to another (figure 1) there is a kink in the shock front and to eliminate this kink we smooth the curve by a third order polynomial :

$$x(y) = ay^3 + by^2 + cy + d \tag{3.1}$$

and we impose the value of x and its first derivative (equal to zero in this case) at the first point of the shock front on the x axis $(x_1, 0)$; then (3.1) becomes :

$$x(y) = ay^3 + by^2 + x_1 \tag{3.2}$$

Let n be the number of points of the shock front (here n is equal to the number of radius of the mesh) then we take the projection of each point (x_i, y_i) on the curve (3.2) parallel to the x axis, and minimize the square of the distance between (x_i, y_i) and its projection to determine a and b ; that is :

$$Minimize \sum_{i=2}^{n} (ay_i^3 + by_i^2 + x_1 - x_i)^2$$

Deriving this expression with respect to a and b we get a linear system we can solve easily; finally we have :

$$a = \frac{\left[\sum_{i=2}^n y_i^3(x_1 - x_i)\right]\left(\sum_{i=2}^n y_i^4\right) - \left[\sum_{i=2}^n y_i^2(x_1 - x_i)\right]\left(\sum_{i=2}^n y_i^5\right)}{\left(\sum_{i=2}^n y_i^6\right)\left(\sum_{i=2}^n y_i^4\right) - \left(\sum_{i=2}^n y_i^5\right)^2}$$

and

$$b = \frac{\left[\sum_{i=2}^n y_i^2(x_1 - x_i)\right]\left(\sum_{i=2}^n y_i^6\right) - \left[\sum_{i=2}^n y_i^3(x_1 - x_i)\right]\left(\sum_{i=2}^n y_i^5\right)}{\left(\sum_{i=2}^n y_i^6\right)\left(\sum_{i=2}^n y_i^4\right) - \left(\sum_{i=2}^n y_i^5\right)^2}$$

Thus we get a global smooth curve representing the shock front and we can reconstruct the mesh in order that curve be located in the middle of a row of cells. Consequently a

computation with this new mesh yields a better solution with good representation of the shock and no kinks.

We can then improve the location of the shock using the variation of the entropy in the Ω_∞ region. Indeed in this region before the shock the entropy should be equal everywhere to the entropy at infinity s_∞ but, due to the dissipation of the method, there are some small gradients. These gradients are not the same on all the radius of the mesh, so we can use them to improve the shape of the shock wave. The main idea is the following : if at one point before the shock we have $s_i < s_\infty$ then there is an expansion and we move the shock upsteam, on the contrary if $s_i > s_\infty$ there is a compression and the shock must be moved downstream. So on each radius we move the position of the shock according to $(s_i - s_\infty)$; then we use a smoothing polynomial to get the shock front. In conclusion the mesh adaptation proposed above gives the good shape of the shock which is captured within a cell and reduces the dissipation before the shock. The effect can be seen also on the value of the least-squares functional $J(\tilde{U})$ which is divided by a factor 3 to 10 depending on the numerical test.

4. Numerical results.

Let us start this section by showing how the mesh adaptation improves the quality of the solution for $M_\infty = 3$ around the circular cylinder. Figures 3 and 4 must be compared to figures 1 and 2 respectively ; we see on one hand that the disappearance of the kinks in the shock fronts suppresses local perturbations and on the other hand that the location of the shock with respect to the cylinder is better. Indeed the distance between the shock and the cylinder is 0.68 with the 21×21 points mesh ; this is in very good agreement with the theory and experiments (see for instance [6] and [7]). Moreover the shape of the shock is very closed to those obtained in [7]. On figure 5 is plotted the pressure on the axis of symmetry and on the cylinder ; it shows a very good representation of the bow shock.

Other numerical tests are presented around the cylinder for $M_\infty = 8$, the results are closed to those computed in [8] for a 21×21 equivalent mesh. However the representation of the shock is more accurate in our computation because there are no points in the shock structure (figures 6 and 7). Furthermore we see the influence of the number of points in the

mesh ; on figure 7 the shock is much closer to the profile.

Then we apply the method above to the flow round an ellipse for the same Mach numbers at infinity $M_\infty = 3$ and $M_\infty = 8$. The remarks on the coefficients C_s of the entropy corrector are about the same than for the circular cylinder. The results plotted on figures 8 and 9 show, on one hand, that the bow shock is well captured and, on the other hand, that there are no oscillations in the field after the shock. The results are very similar to the ones in [9]. Indeed, the work of Lerat et al. shows also that a centered scheme ([10]) can be well-suited to compute transonic flows with shocks, and even hypersonic flows with strong shocks when an entropy corrector is added (11). Their method developped from a completely different approach can be quite close to ours in some cases and yields the same kind of results.

Let us conclude this section by a few remarks on the efficiency of the method :

First, the Newton linearization and the least-squares embedding yield a very robust method that can capture strong phenomena without the help of artificial viscosity.

Second, the incomplete Cholesky conjugate gradient method with addition of some terms in the factorization is very efficient to solve linear systems.

Third, the full multigrid procedure improves considerably the convergence on fine grids.

Fourth, the mesh adaptation, by fitting the grid to the solution, accelerates the capture of a good solution.

Fifth, the programs are highly vectorized (cf. [2]).

Finally, in spite of a totally implicit method for solving the steady Euler model, we can reach acceptable performances in terms of CPU time. The solution around the cylinder for $M_\infty = 3$ is achieved on the 21×21 points grid in 67 seconds of one processor of the CRAY2.

Conclusion.

In this paper we show that, introducing an entropy corrector which vanishes when the mesh size decreases, the centered method presented in former works is able to compute

hypersonic flows with strong shocks. The computed solutions exhibit a very accurate representation of the bow shock which is captured within a cell. Moreover, by using efficient techniques, we can make the method nearly competitive.

Acknowledgements.

I would like to thank my colleagues, C. Jouron and J. Laminie with whom I had fruitful discussions. Furthermore I am really grateful to M. Tajchman who provided me with all the meshes used before adaptation.

This work was partly supported by ESA-CNES-DASSAULT under contract RDAN3/87.

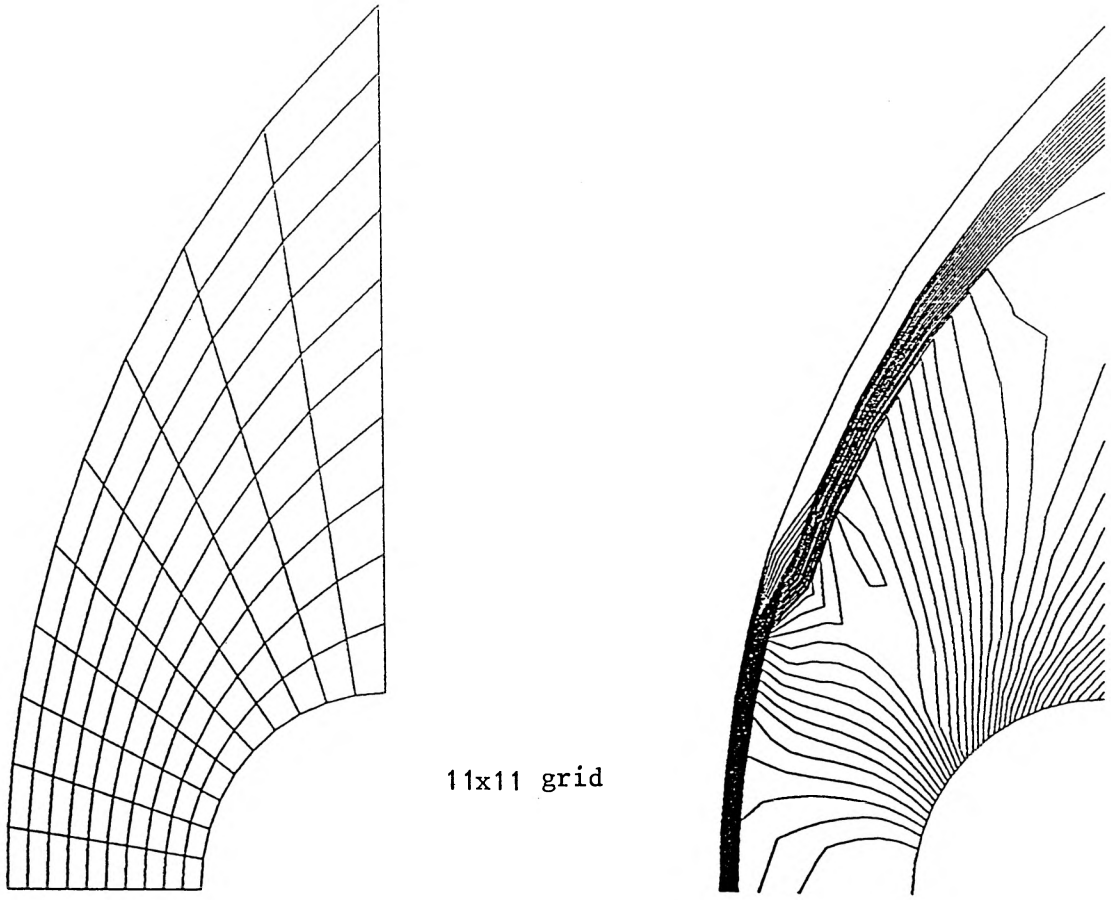


Figure 1 : Mesh and pressure contours for $M_\infty = 3$.

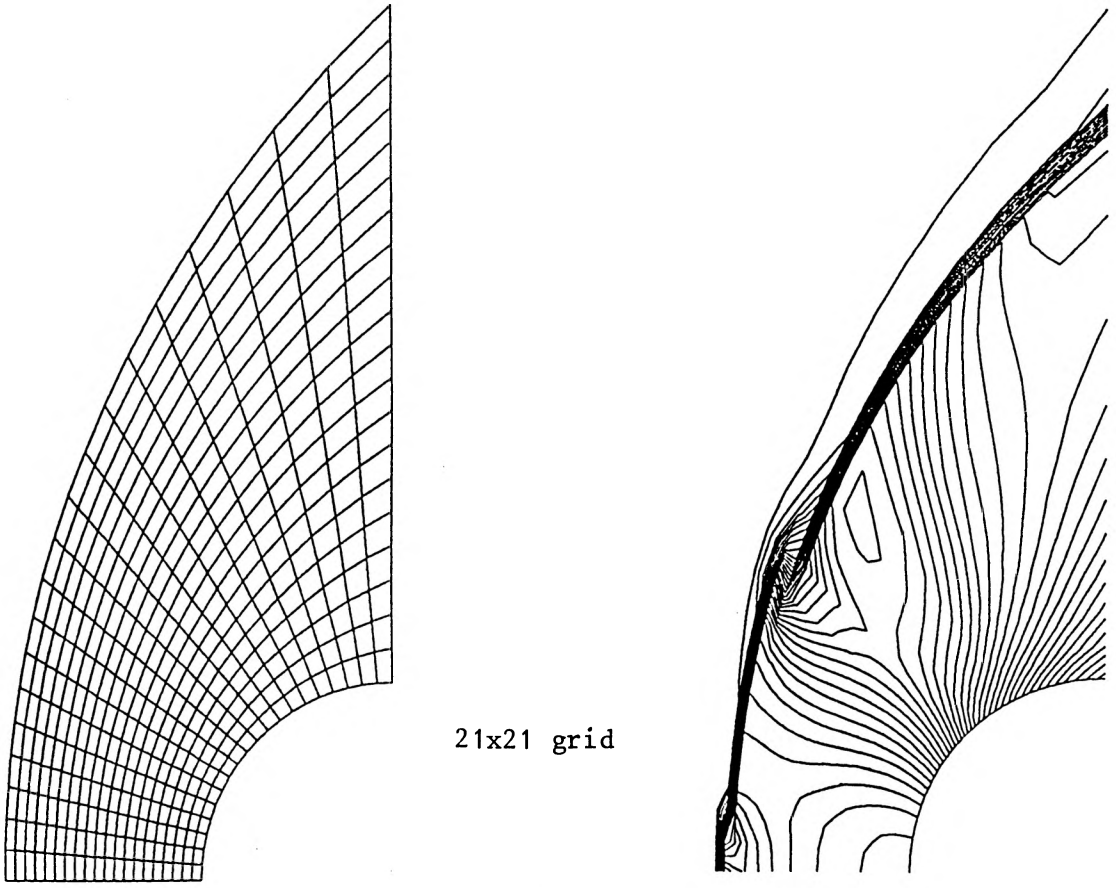
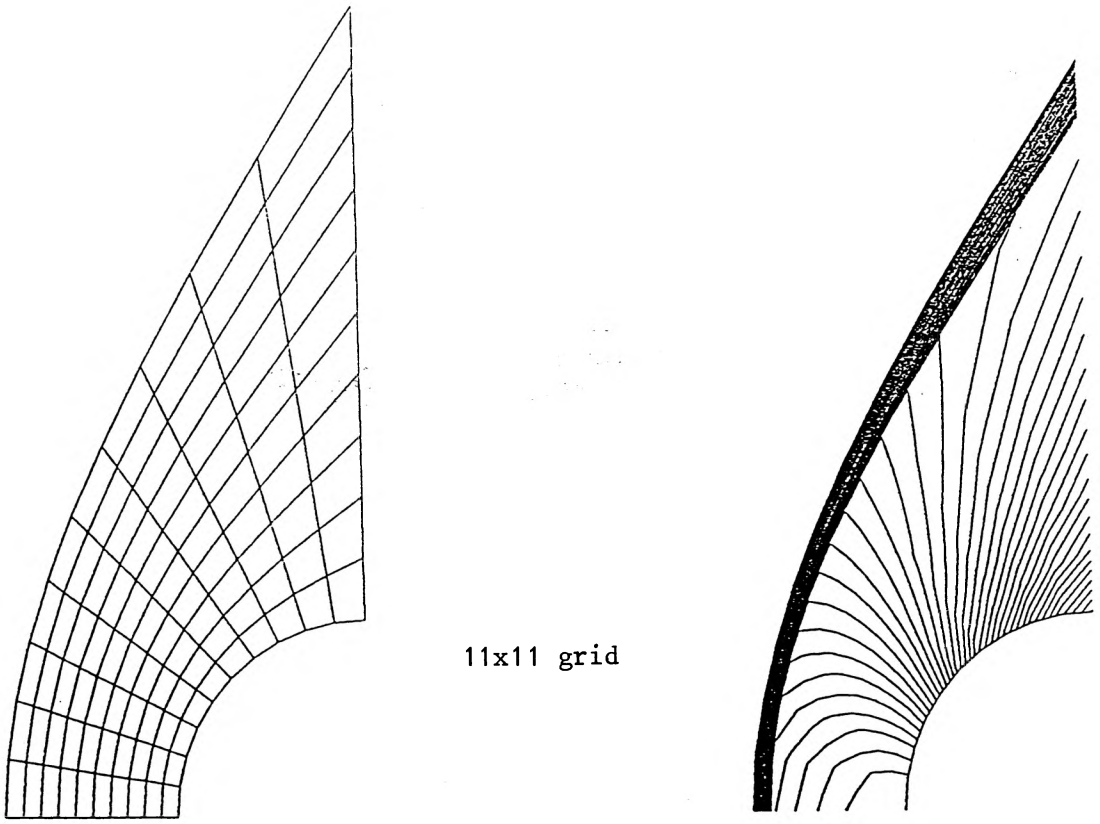
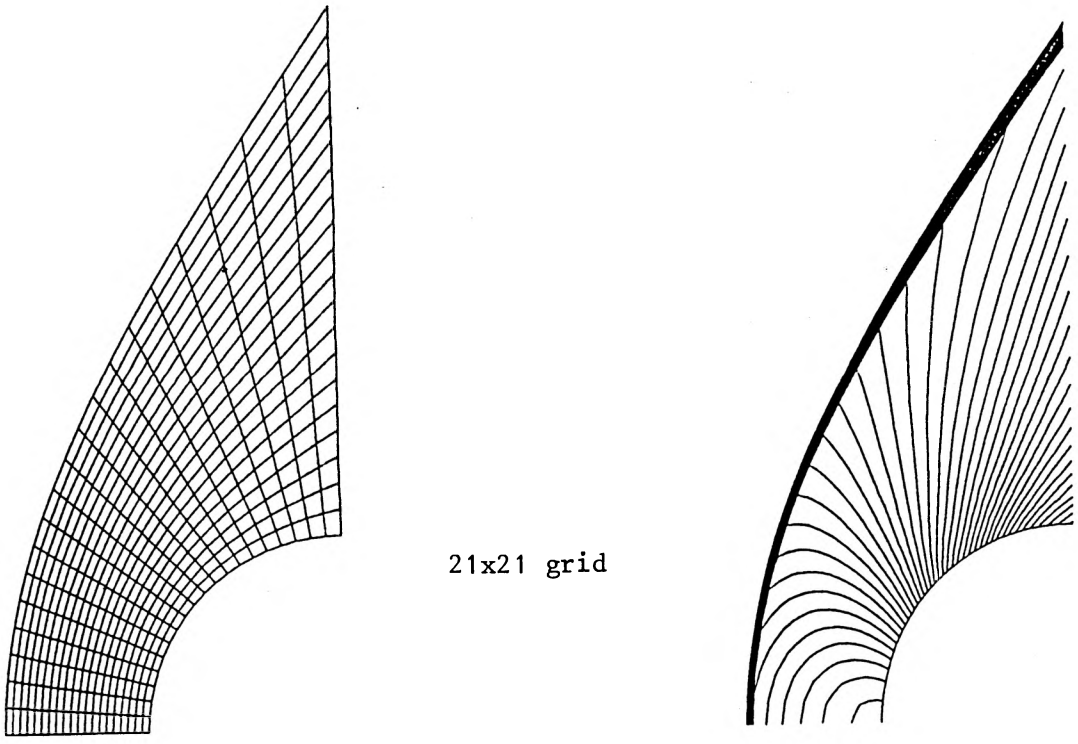


Figure 2 : Mesh and pressure contours for $M_\infty = 3$.



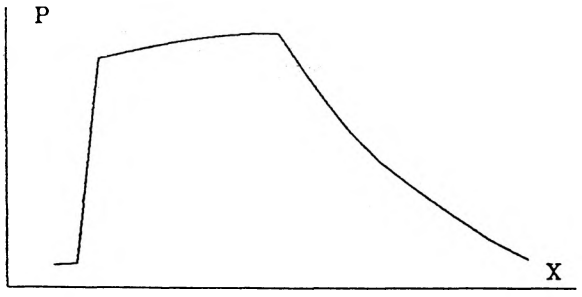
11x11 grid

Figure 3 : Adapted mesh and pressure contours for $M_\infty = 3$.

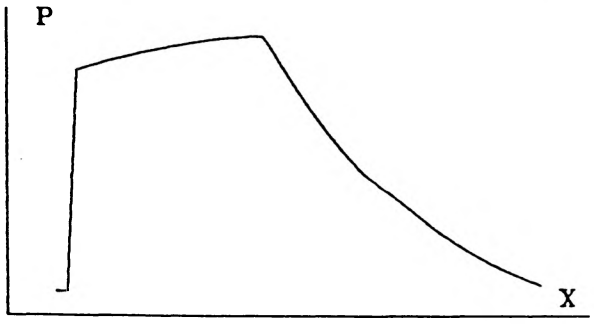


21x21 grid

Figure 4 : Adapted mesh and pressure contours for $M_\infty = 3$.



11x11 grid



21x21 grid

Figure 5 : Pressure on the axis of symmetry and on the profile for the $M_\infty = 3$ flow around the circular cylinder.

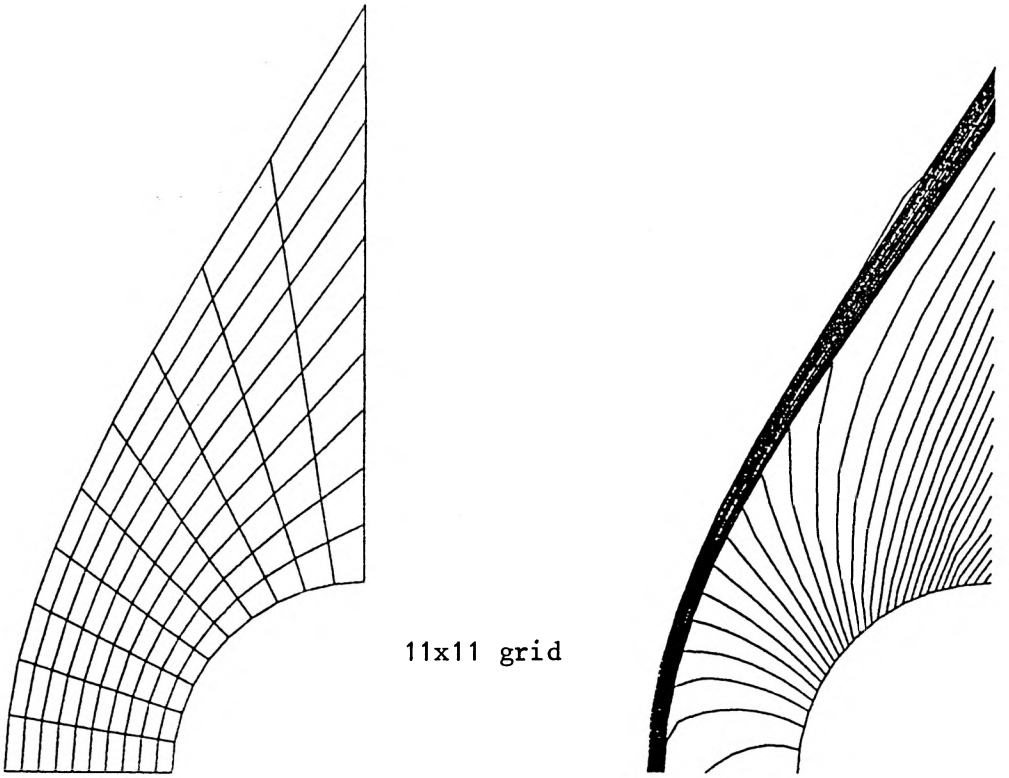


Figure 6 : Adapted mesh and pressure contours for $M_\infty = 8$.

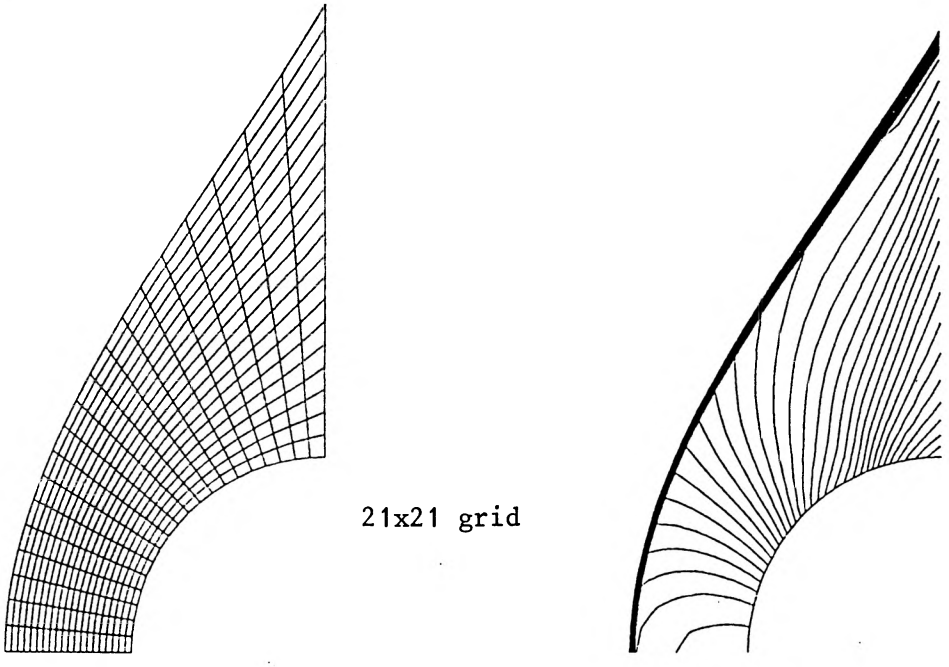


Figure 7 : Adapted mesh and pressure contours for $M_\infty = 8$.

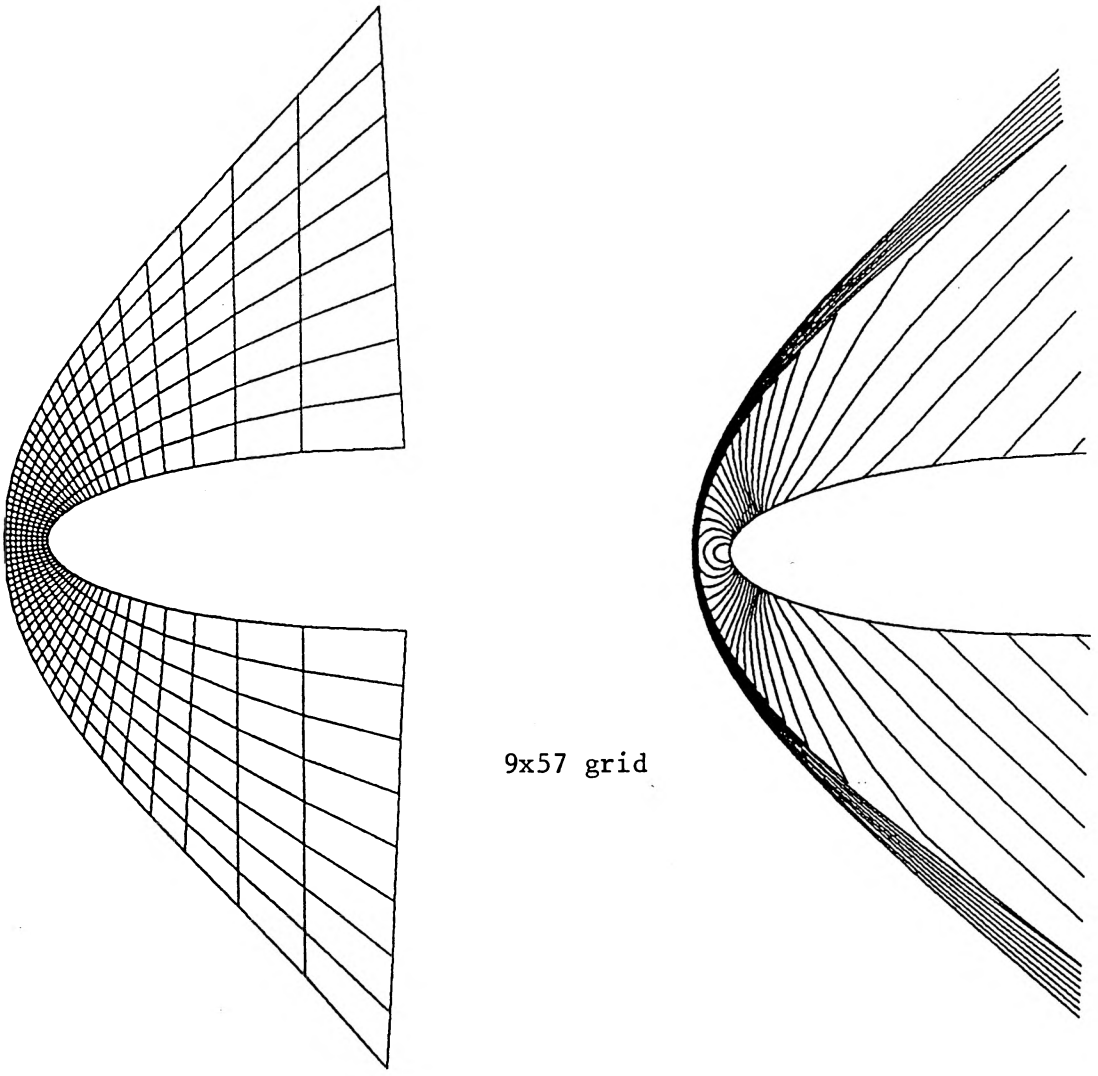


Figure 8 : Adapted mesh and pressure contours for the $M_\infty = 3$ flow around the ellipse

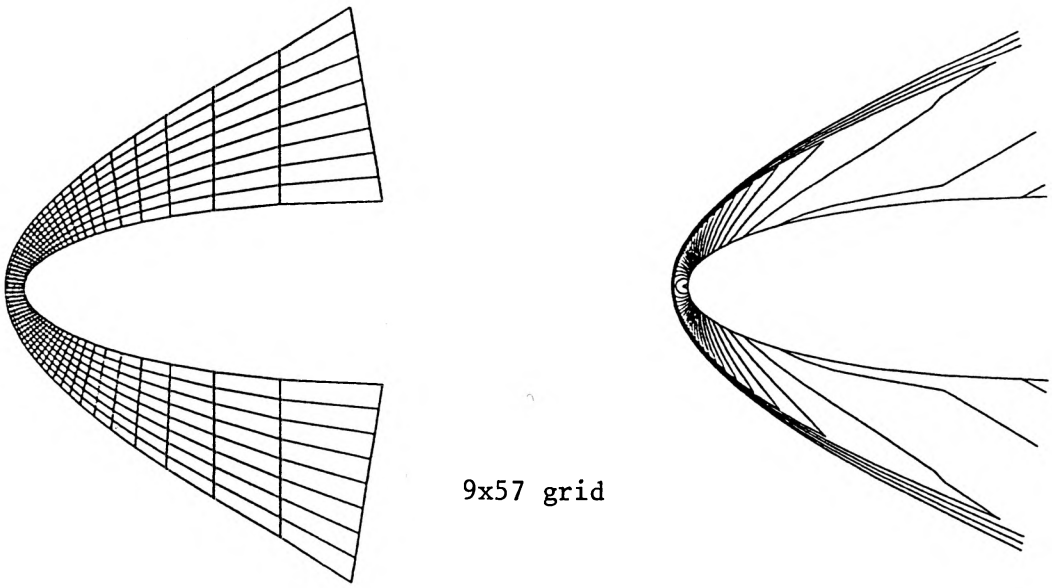


Figure 9 : Adapted mesh and pressure contours for the $M_\infty = 8$ flow around the ellipse

References.

[1] Ch.-H. Bruneau, J.J. Chattot, J. Laminie and J. Guiu-Roux, Finite element least-squares method for solving full steady Euler equations in a plane nozzle, 8th ICNMFD, Aachen 1982, Lecture Notes in Physics 170.

[2] Ch.-H. Bruneau, J. Laminie and J.J. Chattot, Computation of 3D vortex flows past a flat plate at incidence through a variational approach of the full steady Euler equations, to appear in Int. J. for Num. Meth. in Fluids.

[3] Workshop on hypersonic flows for reentry problems, Antibes, 1990.

[4] G.H. Golub and G.A. Meurant, Résolution numérique des grands systèmes linéaires, Ecole d'Eté d'Analyse Numérique, CEA-EDF-INRIA, 1983.

[5] 7th International Conference of finite element methods in flow problems, Huntsville, 1989.

[6] P. Rebuffet, Aérodynamique Expérimentale, Tomes I and II, Dunod, 1966.

[7] P. Diringer, Calcul d'ondes de choc détachées, La Recherche Aéronautique 88, 1962.

[8] V. Selmin and L. Quartapelle, Finite element solution to the Euler equations, 10th ICNMFD, Beijing 1986, Lecture Notes in Physics 264.

[9] K. Khalfallah and A. Lerat, Communication to Hermes Meeting, Rome 1988.

[10] A. Lerat and J. Sides, Efficient solution of the steady Euler equations with a centered implicit method, Numerical Methods for Fluid Dynamics, Oxford, 1988.

[11] K. Khalfallah and A. Lerat, Correction d'entropie pour des schémas numériques approchant un système hyperbolique, C. R. Acad. Sc. Paris, 308, II, n° 9, 1989.

CONCLUSION.

La méthode développée tout au long de ce chapitre a permis de calculer la solution stationnaire des équations d'Euler compressibles dans des conditions très variées. Les dernières modifications montrent qu'une approximation centrée peut capturer des solutions hypersoniques avec des enroulements tourbillonnaires et de forts chocs détachés contrairement aux idées en cours. De plus la méthode proposée permet de calculer directement la solution stationnaire, ce qui est, là encore, peu usuel.

La méthode variationnelle utilisée conduit à une approximation complètement implicite qui est de ce fait assez coûteuse ; néanmoins une gestion optimale des capacités du CRAY2 a permis de la rendre compétitive.

PARTIE B

SCHEMAS AUX DIFFERENCES FINIES DECENTREES POUR LA RESOLUTION DES EQUATIONS DE NAVIER-STOKES INCOMPRESSIBLES

Introduction

Schémas décentrés avec terme d'antidiffusion

Résolution numérique des équations de Navier-Stokes incompressibles dans une cavité entraînée

Etude du problème linéarisé

Conclusion.

INTRODUCTION.

Nous présentons dans cette deuxième partie un schéma numérique adapté à la résolution de problèmes visqueux et l'appliquons aux équations de Navier-Stokes incompressibles. L'idée est de mettre au point une méthode numérique robuste et performante capable de calculer correctement la solution et de capturer des phénomènes transitoires. Ce travail s'inscrit comme un premier pas vers la représentation de solutions turbulentes et notre but est de trouver le nombre de Reynolds critique pour lequel il y a perte de stabilité de la solution stationnaire. Le problème physique choisi est celui de la cavité entraînée en dimension deux pour lequel de nombreux résultats sont disponibles, en particulier sur la cavité carrée $(0,1)\times(0,1)$. Les travaux passés sur ce problème ont permis soit de capturer des solutions stationnaires stables jusqu'à $Re = 10000$, soit ont fait apparaître des difficultés, pour des nombres de Reynolds inférieurs, que les auteurs attribuent à la présence d'un régime transitoire vers la turbulence. Nous espérons apporter ici une partie de la réponse en étudiant d'abord un schéma capable de supporter le régime transitoire et de représenter correctement la solution puis en essayant de calculer le nombre de Reynolds critique.

Le premier paragraphe est consacré à l'étude d'une classe de schémas adaptés à la résolution de problèmes visqueux en dimension un du type

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} = 0 \qquad (B.0)$$

où f est une fonction non linéaire de u (par exemple $u^2/2$ pour l'équation de Burgers) et ν est un nombre réel positif. Ensuite sont présentés les résultats obtenus sur le problème de la cavité entraînée en utilisant une variante de l'un des schémas, puis le calcul du premier exposant de Lyapunov du système de Navier-Stokes linéarisé.

SCHEMAS DECENTRES AVEC TERME D'ANTIDIFFUSION

Avant d'attaquer le problème (B.0.) présentons le cadre dans lequel nous nous plaçons sur l'équation scalaire de conservation :

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0. \qquad (B.1)$$

Pour résoudre cette équation de nombreux schémas ont été introduits ces trente dernières années dont la plupart (Godunov, Lax-Friedrichs, Murman, etc.) appartiennent à la catégorie des schémas TVD. ("Total Variation Diminishing") introduite par Harten. Dans cette présentation nous nous restreignons aux schémas TVD à trois points, c'est-à-dire que la valeur approchée de u en un point est déduite par différences finies sur trois points.

Supposons que l'on cherche $u(x)$ sur l'ouvert (a,b), alors on note v_j la valeur approchée au point $a+\frac{j}{N}(b-a)$ en espace et v_j^n la valeur approchée au même point au temps $t = n\Delta t$ où N est le nombre de segments de discrétisation $(\Delta x = \frac{(b-a)}{N})$ et Δt est le pas de temps. Alors, on peut résoudre l'équation (B.1) par un schéma conservatif à flux numérique de la forme :

$$v_j^{n+1} = v_j^n - \frac{\Delta t}{\Delta x}\left(g_{j+1/2}^n - g_{j-1/2}^n\right) \qquad (B.2)$$

où v_j^{n+1} est déduit explicitement en fonction du flux numérique g qui est consistant si $g(u,u) = f(u)$ (dans le cas d'un schéma à trois points on a $g_{j+1/2}^n = g\left(v_j^n, v_{j+1}^n\right)$). Parmi ces schémas conservatifs Harten a introduit les schémas TVD tels que $TV(v^{n+1}) \leq TV(v^n)$ où $TV(v) = \sum_j \mid v_{j+1} - v_j \mid$.

Si l'on réécrit le schéma (B.2) sous la forme $\left(\lambda = \frac{\Delta t}{\Delta x}\right)$:

$$v_j^{n+1} = v_j^n + C_{j+1/2}^n\left(v_{j+1}^n - v_j^n\right) - D_{j-1/2}^n\left(v_j^n - v_{j-1}^n\right)$$

avec

$$C_{j+1/2}^n = \lambda\frac{f(v_j^n) - g_{j+1/2}^n}{v_{j+1}^n - v_j^n} \text{ et } D_{j-1/2}^n = \lambda\frac{f(v_j^n) - g_{j-1/2}^n}{v_j^n - v_{j-1}^n}$$

ou encore :

$$\begin{aligned} v_j^{n+1} = v_j^n &- \frac{\lambda}{2}(f(v_{j+1}^n) - f(v_{j-1}^n)) + \frac{\lambda}{2}(Q_{j+1/2}^n\left(v_{j+1}^n - v_j^n\right) \\ &- Q_{j-1/2}^n\left(v_j^n - v_{j-1}^n\right) \end{aligned} \qquad (B.3)$$

avec

$$Q_{j+1/2}^n = \frac{1}{\lambda}\left(C_{j+1/2}^n + D_{j+1/2}^n\right) = \frac{f(v_j^n) + f(v_{j+1}^n) - 2g_{j+1/2}^n}{v_{j+1}^n - v_j^n},$$

on peut donner une double caractérisation des schémas TVD à 3 points déduite directement de la définition :

Un schéma conservatif à trois points est TVD si et seulement si l'une des deux conditions suivantes est satisfaite

(i) $C_{j+1/2} \geq 0$, $D_{j+1/2} \geq 0$ et $C_{j+1/2} + D_{j+1/2} \leq 1$ pour tous les j
(ii) $\left| \dfrac{f(u_{j+1}) - f(u_j)}{u_{j+1} - u_j} \right| \leq Q_{j+1/2} \leq \dfrac{1}{\lambda}$ pour tous les j.

Avant d'exploiter toutes ces formules nous rappelons le résultat essentiel de convergence:

Un schéma TVD construit à partir d'un flux numérique localement lipschitzien est convergent si et seulement si il est L^∞ stable.

Remarquons maintenant que l'écriture (B.3) du schéma correspond à une discrétisation en espace par un schéma centré du deuxième ordre de l'équation :

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} - \frac{\Delta x}{2} \frac{\partial}{\partial x} \left(Q(u) \frac{\partial u}{\partial x} \right) = 0 \qquad (B.4)$$

donc, en discrétisant (B.1) par un schéma TVD à trois points, on résout non pas (B.1) mais un problème visqueux obtenu en rajoutant à (B.1) le terme de diffusion ci-dessus faisant intervenir la fonction Q dont les bornes sont données par la caractérisation (ii) :

$$| a(u) | \leq Q(u) \leq \frac{1}{\lambda} \quad (a(u) = f'(u)).$$

De plus ces bornes correspondent à des schémas classiques, le schéma de Murman pour la borne inférieure et le schéma de Lax-Friedrichs pour la borne supérieure (le schéma de Godunov est aussi une borne pour les schémas entropiques mais ce n'est pas ici l'objet de notre propos). Donc parmi ces schémas, celui de Lax-Friedrichs est celui qui tasse le plus les phénomènes (et en particulier étale le plus les chocs) et celui de Murman est le plus près du problème (B.1) que l'on veut résoudre. Voyons comment ce dernier schéma est construit. A l'instant initial on définit la fonction constante par morceaux $v^0(x)$ par

$$v_j^0(x) = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u_0(y) dy, \quad x_{j-1/2} < x < x_{j+1/2} \text{ pour tous les } j,$$

ensuite à chaque pas de temps on résout sur chaque maille le problème de Riemann suivant, déduit de la forme non conservative de l'équation (B.1) $\frac{\partial u}{\partial t} + a(u)\frac{\partial u}{\partial x} = 0$ que l'on linéarise:

$$\begin{cases} \dfrac{\partial v_{Tj}}{\partial t} + a^n_{j+1/2}\dfrac{\partial v_{Tj}}{\partial x} = 0 \\ v_{Tj}(x,0) = \begin{cases} v^n_j \quad \text{si } x_j < x < x_{j+1/2} \\ v^n_{j+1} \text{ si } x_{j+1/2} < x < x_{j+1} \end{cases} \quad \text{pour tous les } j, \end{cases}$$

où $a^n_{j+1/2} = \dfrac{f(v^n_{j+1}) + f(v^n_j)}{v^n_{j+1} - v^n_j}$ remplace $a(v_{Tj})$; donc la donnée initiale est transportée à la vitesse $a^n_{j+1/2}$ et la solution en un point (x,t) est donnée par

$$v_{Tj}\left(\frac{x - x_{j+1/2}}{t}, v^n_j, v^n_{j+1}\right) = \begin{cases} v^n_j \quad \text{si } \dfrac{x - x_{j+1/2}}{t} < a^n_{j+1/2} \\ v^n_{j+1} \text{ si } \dfrac{x - x_{j+1/2}}{t} > a^n_{j+1/2} \end{cases}$$

Soit v_T la fonction définie sur (a,b) égale aux v_{Tj} sur chaque intervalle (x_j, x_{j+1}) ; on définit la fonction v par projection sur les fonctions constantes par morceaux :

$$v(x,t) = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} v_T(y,t)dy \ ; \ x_{j-1/2} < x < x_{j+1/2} \text{ pour tous les } j;$$

cette fonction est bien définie si les ondes transportées de l'instant $n\Delta t$ n'interfèrent pas avant l'instant $(n+1)\Delta t$, ce qui est vrai sous la condition CFL (f convexe) $\lambda \leq \dfrac{1}{max_j |a^n_{j+1/2}|}$ qui traduit le fait qu'en un sommet du maillage le cône de dépendance numérique contient le cône de dépendance théorique. Enfin on calcule v^{n+1}_j en intégrant l'équation (B.1) comme suit :

$$\int_{n\Delta t}^{(n+1)\Delta t}\int_{x_{j-1/2}}^{x_{j+1/2}} \left(\frac{\partial v}{\partial t} + \frac{\partial f(v)}{\partial x}\right) dxdt = 0$$

$$\int_{x_{j-1/2}}^{x_{j+1/2}} [v(x,(n+1)\Delta t) - v(x,n\Delta t)]dx +$$

$$+ \int_{n\Delta t}^{(n+1)\Delta t} [f(v^-(x_{j+1/2},t)) - f(v^+(x_{j-1/2},t))]dt = 0$$

$$\Delta x\left(v^n_{j+1} - v^n_j\right) + \int_{n\Delta t}^{(n+1)\Delta t} [f(v_T(o^-, v^n_j, v^n_{j+1})) - f(v_T(o^+, v^n_{j-1}, v^n_j))]dt = 0$$

et comme la fonction $\xi \to f(v_T(\xi, v^n_j, v^n_{j+1}))$ est continue en $\xi = 0$ on tire

$$v^n_{j+1} = v^n_j - \lambda(f(v_T(0, v^n_j, v^n_{j+1})) - f(v_T(0, v^n_{j-1}, v^n_j)).$$

Ceci donne pour le flux numérique :

$$g^n_{j+1/2} = f(v_T(0, v^n_j, v^n_{j+1})) = \begin{cases} f(v^n_j) & \text{si } a^n_{j+1/2} > 0 \\ f(v^n_{j+1}) & \text{si } a^n_{j+1/2} < 0 \end{cases}$$

et permet de réécrire le schéma sous la forme :

$$v^{n+1}_j = v^n_j - \frac{\lambda}{2}(f(v^n_{j+1}) - f(v^n_{j-1}))$$
$$+\frac{\lambda}{2}\left\{ \mid a^n_{j+1/2} \mid (v^n_{j+1} - v^n_j) - \mid a^n_{j-1/2} \mid (v^n_j - v^n_{j-1}) \right\}.$$

On a alors un schéma TVD à trois points, construit à partir d'un flux numérique localement lipschitzien, L^∞ stable sous la condition CFL et donc convergent. On peut encore traduire ce schéma comme suit

$$v^{n+1}_j = v^n_j - \begin{cases} \lambda(f(v^n_j) - f(v^n_{j-1})) & \text{si } a^n_{j+1/2} > 0 \text{ et } a^n_{j-1/2} > 0 \\ \lambda(f(v^n_{j+1}) - f(v^n_j)) & \text{si } a^n_{j+1/2} < 0 \text{ et } a^n_{j-1/2} < 0 \\ \lambda(f(v^n_{j+1}) - f(v^n_{j-1})) & \text{si } a^n_{j+1/2} < 0 \text{ et } a^n_{j-1/2} > 0 \\ 0 & \text{sinon} \end{cases} \qquad (B.5)$$

qui montre clairement la faiblesse de ce schéma qui n'est pas entropique et n'élimine pas les chocs de détente (son terme de diffusion minimal est inférieur à cleui du schéma de Godunov qui est la borne inférieure des schémas TVD à trois points entropiques). Néanmoins, pour l'étude présente cela n'intervient pas et si l'on s'intéresse à ce schéma c'est justement parce que son terme de diffusion est minimal.

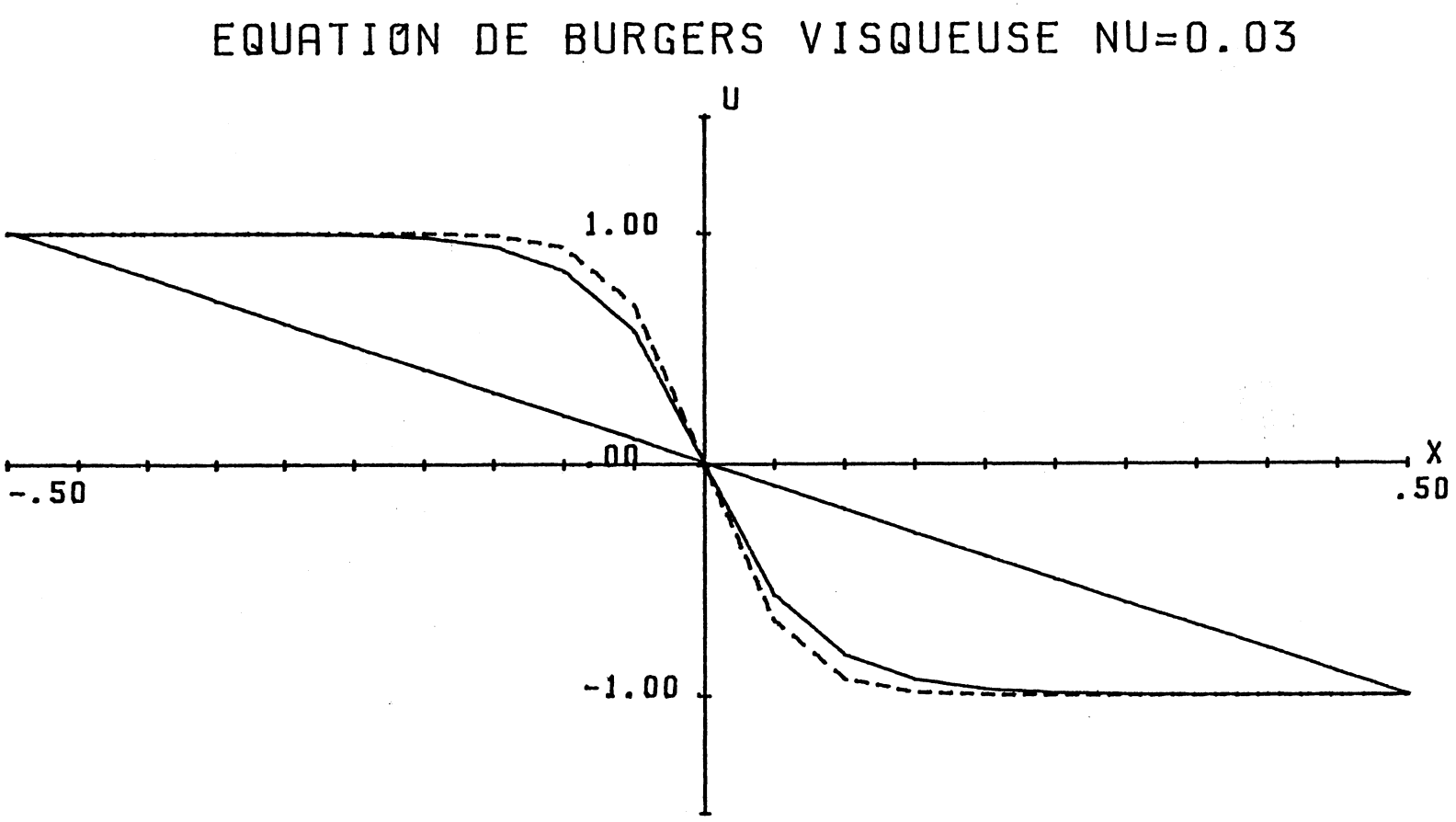
Revenons maintenant au problème (B.0) que l'on résout à l'aide du schéma ci-dessus pour la partie (B.1) plus une discrétisation centrée en espace du deuxième ordre pour le terme de diffusion alors on résout en fait le problème (voir (B.4)) :

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} - \frac{\Delta x}{2}\frac{\partial}{\partial x}\left(\mid a(u) \mid \frac{\partial u}{\partial x} \right) - \nu \frac{\partial^2 u}{\partial x^2} = 0$$

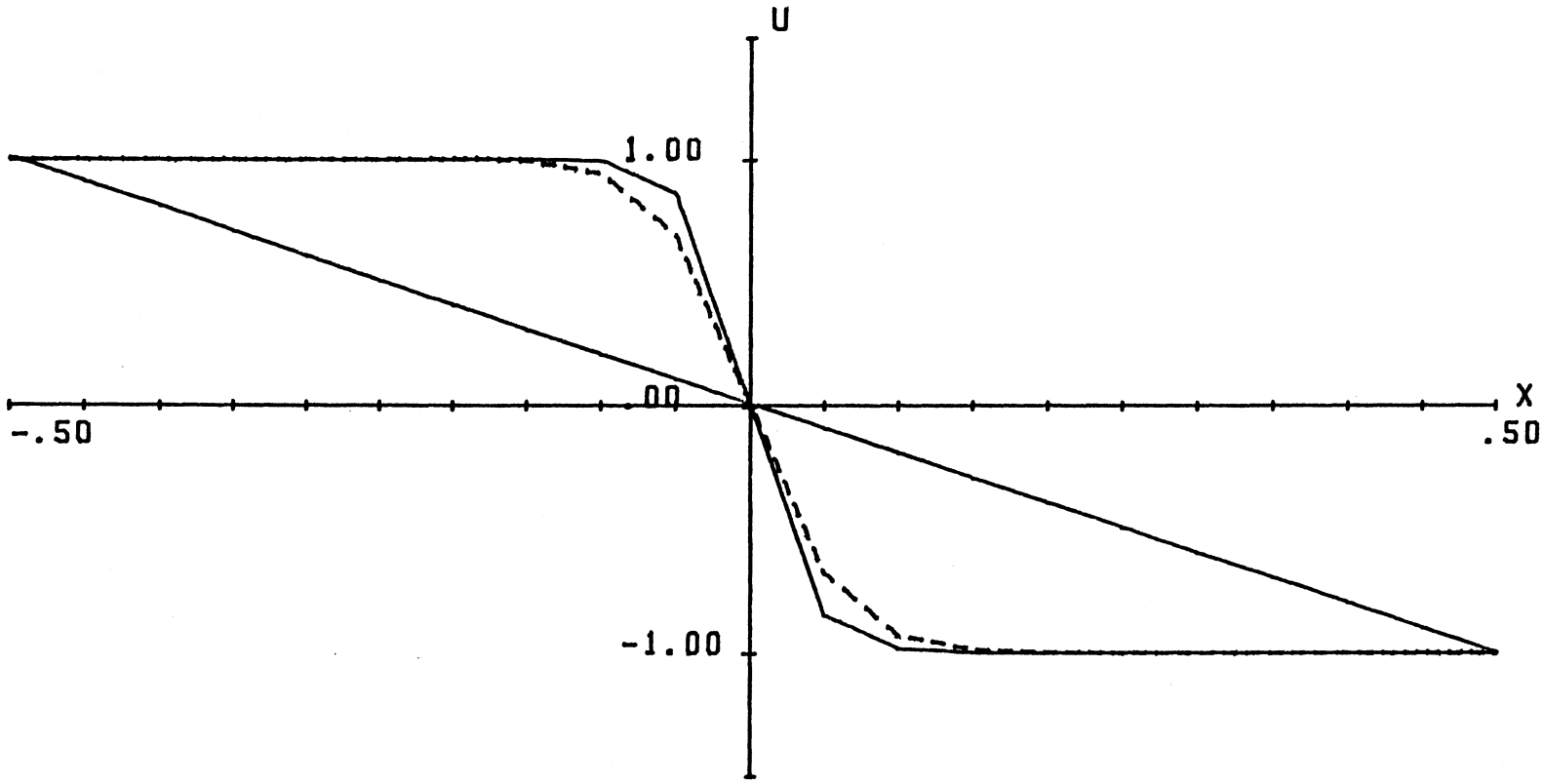
avec une approximation du deuxième ordre en espace. Cela signifie que, quand $\nu < \frac{\Delta x}{2} \mid a(u) \mid$, le terme visqueux dans (B.0) est dominé par le terme artificiel de diffusion $-\frac{\Delta x}{2} \mid a(u) \mid \frac{\partial^2 u}{\partial x^2}$ introduit par le schéma TVD ; or on veut calculer la solution pour $\nu << 1$. Donc le schéma contient trop de viscosité et la solution calculée est très amortie dès que ν est petit ; en fait elle correspond davantage à la solution pour $\nu' = \nu + \frac{\Delta x}{2} \mid a(u) \mid$.

Néanmoins le problème (B.0) ne peut être approché par un schéma centré pour de petits ν ; l'idée est de prendre un intermédiaire pour lequel d'une part on est plus près de la solution et d'autre part on conserve la stabilité, en introduisant un terme d'antidiffusion $\epsilon \mid a(u) \mid \frac{\partial^2 u}{\partial x^2}$. Il y a de nombreux choix pour ϵ, la limite de stabilité est obtenue pour $\epsilon = min\left(\frac{\nu}{\mid a(u) \mid}, \frac{\Delta x}{2}\right)$ qui permet de conserver les propriétés du schéma TVD initial et donc la convergence. Cela veut dire que tant que ν domine le coefficient du terme artificiel on élimine ce terme et que sinon on supprime le terme visqueux lui-même puisque le terme introduit artificiellement par le schéma lui est supérieur.

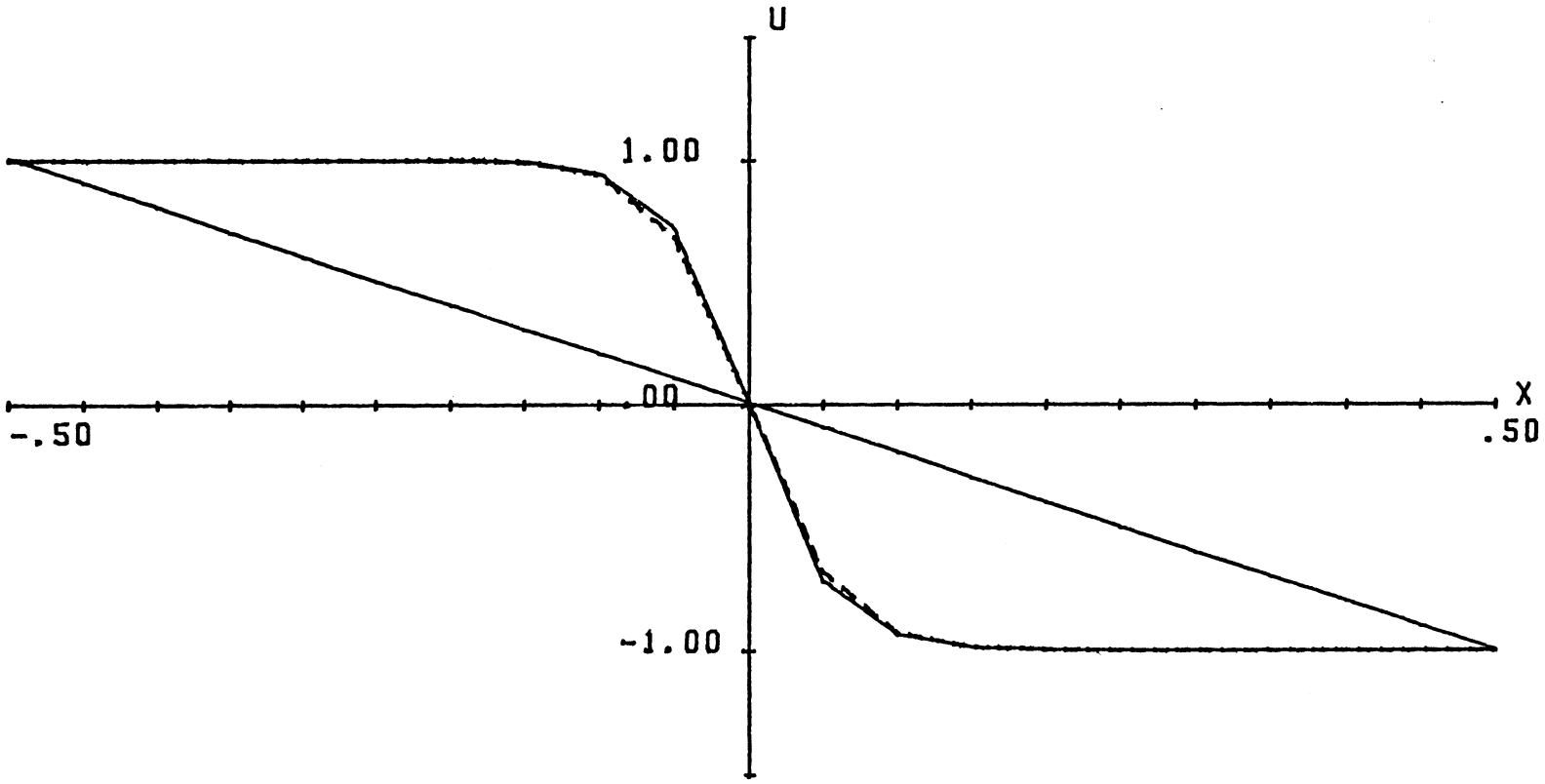
En conclusion il faut être prudent quand on approche un problème visqueux du type (B.0) par un schéma aux différences finies décentrées car on croit résoudre le problème pour un certain paramètre ν et en fait on le résout pour un autre paramètre ν' qui peut être très différent, en particulier si le pas de discrétisation n'est pas assez fin. Nous illustrons ceci, pour l'équation de Burgers visqueuse, en comparant la solution calculée à la solution exacte stationnaire $u_{ex} = \frac{1-e^{x/\nu}}{1+e^{x/\nu}}$ tracée en pointillés sur les figures ci-dessous. La solution calculée sans terme de correction pour $\nu = 0.03$ est plus tassée que la solution exacte et correspond en fait à u_{ex} pour $\nu = 0.04$. Ceci donne dans ce cas particulier une estimation de ν' qui vérifie a priori $\nu \leq \nu' \leq 0.055$ selon les valeurs de u ($\Delta x = 0.05$). La correction apportée par le terme d'antidiffusion avec $\epsilon = \min\left(\frac{\nu}{|a(u)|}, \frac{\Delta x}{2}\right) = \frac{\Delta x}{2}$ est trop importante et la solution calculée est aussi éloignée de la solution exacte que sans correcteur mais cette fois-ci de l'autre côté comme le montre la deuxième figure. Par contre la valeur $\epsilon = \frac{\Delta x}{3}$ permet d'obtenir une excellent approximation de la solution exacte.



SCHEMA DE MURMAN


EQUATION DE BURGERS VISQUEUSE NU=0.03 E=DX/2
U
1.00
.00
-1.00
-.50
.50
X
SCHEMA DE MURMAN

EQUATION DE BURGERS VISQUEUSE NU=0.03 E=DX/3



SCHEMA DE MURMAN

RESOLUTION NUMERIQUE DES EQUATIONS DE NAVIER-STOKES INCOMPRESSIBLES DANS UNE CAVITE ENTRAINEE.

Ici nous voyons comment appliquer le schéma ci-dessus aux équations de Navier-Stokes incompressibles en dimension deux. Le cadre de la résolution est très largement développé dans [BJ2] ; pour atteindre les meilleures performances possibles nous utilisons une méthode multigrille FMG-FAS ("Full Multi Grid"-"Full Approximation Storage") avec relaxation maille à maille sur chaque grille et une discrétisation par variables décalées. Ce qui signifie que l'on résout sur une maille le système complet de Navier-Stokes pour les cinq inconnues de la maille (voir figure 1 dans [BJ2]) ; ce système est linéarisé par relaxation et l'ensemble de la résolution est stabilisé par sous-relaxation dès que le nombre de Reynolds dépasse 100 environ sur la cavité entraînée $(0,1)\times(0,1)$.

Tout le succès de l'approximation réside dans l'équilibre entre la représentation de la solution et la stabilité ; le schéma centré par exemple qui représente le mieux la solution puisqu'il ne rajoute pas de termes de diffusion dans les équations, est instable dès que le nombre de Reynolds dépasse quelques centaines et donc la solution explose quel que soit le paramètre de relaxation ω que l'on choisisse (même pour $\omega < 0.1$). Ceci peut être illustré facilement en observant le système linéaire 5×5 que l'on obtient dans ce cas pour l'inconnue $U = (u_{i-1/2,j}, u_{i+1/2,j}, v_{i,j-1/2}, v_{i,j+1/2}, p_{i,j})^T$ sur la maille (i,j). Regardons la ligne correspondant à la première inconnue ; l'opérateur de diffusion donne une contribution sur la diagonale principale $\frac{2}{Re(\Delta x)^2} + \frac{2}{Re(\Delta y)^2}$ et le terme de convection $u\frac{\partial u}{\partial x}$ amène sur cette même diagonale $\frac{u^n_{i+1/2,j} - u^n_{i-3/2,j}}{2\Delta x}$ qui dépend de la solution à l'itération précédente et dont le signe est variable selon la maille où l'on se trouve. Si l'on a un pas uniforme $h = \Delta x = \Delta y$ alors il faut comparer $\frac{4}{Reh^2}$ avec $\frac{1}{2h}(u^n_{i+1/2,j} - u^n_{i-3/2,j})$; or si l'on regarde par exemple les profils médians de la solution pour $Re = 1000$ (figure 4 dans [BJ2]) on s'aperçoit que les gradients sont déjà importants, en particulier près des bords. Ce qui veut dire que si le pas de discrétisation n'est pas assez fin (c'est le cas sur les grilles grossières) on peut même obtenir un terme diagonal négatif et dans tous les cas on affecte suffisamment (quand la dérivée est négative) le terme diagonal dominant issue de la discrétisation de $-\frac{1}{Re}\Delta u$ pour perdre la stabilité. Par contre, si l'on discrétise $u\frac{\partial u}{\partial x}$ par un schéma décentré, on amène sur la diagonale principale une contribution toujours positive mais qui n'est pas forcément dominante ; le mieux que l'on puisse obtenir est avec le schéma décentré du

premier ordre :

$$u^n_{i-1/2}\frac{u_{i-1/2}-u^n_{i-3/2}}{\Delta x} \text{ si } u^n_{i-1/2}>0$$
$$u^n_{i-1/2}\frac{u^n_{i+1/2}-u_{i-1/2}}{\Delta x} \text{ si } u^n_{i-1/2}<0$$

pour lequel la contribution à la diagonale principale $\frac{|u^n_{i-1/2}|}{\Delta x}$ est de même module que les autres termes. Ce schéma, comme celui de Murman d'ailleurs, est très stable ; il permet de calculer une solution pour $Re = 10^5$ par exemple sur des grilles relativement grossières ($h = 1/64$ par exemple) mais c'est seulement, comme on l'a vu au paragraphe précédent, parce que le terme artificiel $\frac{h}{2} \mid u \mid$ domine $\frac{1}{Re}$ et en fait, la solution calculée correspond à un nombre de Reynolds de l'ordre de 10^3 sur la majeure partie du domaine. Le remède à ce problème d'équilibre entre diffusion et stabilité est d'utiliser le schéma avec terme d'antidiffusion du paragraphe précédent ; malheureusement ce schéma ne peut être utilisé tel quel car la discrétisation centrée de $\epsilon \mid u \mid \frac{\partial^2 u}{\partial x^2}$ amène une contribution négative sur la diagonale principale et affecte beaucoup la stabilité.

Nous présentons maintenant une variante du schéma d'antidiffusion qui reste stable, sur l'équation scalaire (B.0) modifiée :

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}-\nu\frac{\partial^2 u}{\partial x^2}+\epsilon \mid a(u) \mid \frac{\partial^2 u}{\partial x^2}=0. \qquad (B.6)$$

Le terme visqueux est discrétisé par différences finies centrées car c'est lui qui amène un terme fixe positif sur la diagonale principale et donne la stabilité. Voyons maintenant comment discrétiser le dernier terme, tout en renforçant la stabilité, sur l'équation (B.1) modifiée :

$$\frac{\partial u}{\partial t}+\frac{\partial f(u)}{\partial x}+\epsilon \mid a(u) \mid \frac{\partial^2 u}{\partial x^2}=0 \qquad (B.7)$$

En fait nous montrons ici comment coupler la discrétisation du terme d'antidiffusion à celle du terme de conservation traité par le schéma TVD de Murman en introduisant le schéma :

$$v^{n+1}_j = v^n_j - \lambda(g^n_{j+1/2}-g^n_{j-1/2}) - \epsilon b^n_j \lambda(h^n_{j+1/2}-h^n_{j-1/2}) \qquad (B.8)$$

avec

$$g^n_{j+1/2} = g(v^n_j, v^n_{j+1}) = \begin{cases} f(v^n_j) & \text{si } a^n_{j+1/2}>0 \\ f(v^n_{j+1}) & \text{sinon} \end{cases}$$

$$b^n_j = b(v^n_{j-1}, v^n_j, v^n_{j+1}) = \max(a^n_{j-1/2}, 0) - \min(a^n_{j+1/2}, 0)$$

$$h^n_{j+1/2} = h(v^n_{j-1}, v^n_j, v^n_{j+1}, v^n_{j+2}) = \begin{cases} (v^n_j - v^n_{j-1})/\Delta x & \text{si } a^n_{j+1/2} > 0 \\ (v^n_{j+2} - v^n_{j+1})/\Delta x \text{ sinon} \end{cases}$$

où la dérivée $\frac{\partial u}{\partial x}$ est traitée comme $f(u)$ avec décentrement des différences finies (des différences centrées donnent zéro dans le cas de la compression), le couplage se faisant à travers les tests sur $a^n_{j+1/2}$ et $a^n_{j-1/2}$. De plus les termes rajoutés sont consistants, en effet on a bien $b(u,u,u) =\mid a(u) \mid$ et $h(u,u,u,u)$ converge vers $\frac{\partial u}{\partial x}$ quand Δx tend vers zéro. Enfin, le schéma décentré (B.8) correspond à la discrétisation centrée du deuxième ordre en espace de l'équation :

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} - \frac{\Delta x}{2}\frac{\partial}{\partial x}\left(\mid a(u) \mid \frac{\partial u}{\partial x}\right) + \epsilon \mid a(u) \mid \frac{\partial^2 u}{\partial x^2} \\ - \epsilon a(u)\Delta x \frac{\partial^3 u}{\partial x^3} + \epsilon \mid a(u) \mid \frac{(\Delta x)^2}{2}\frac{\partial^4 u}{\partial x^4} = 0 \end{aligned} \tag{B.9}$$

Car

$$\begin{aligned} h^n_{j+1/2} - h^n_{j-1/2} &= \frac{v^n_{j+1} - v^n_j}{\Delta x} - sg(a^n_{j+1/2})\frac{v^n_{j+2} - v^n_{j+1} - v^n_j + v^n_{j-1}}{2\Delta x} \\ &+ \frac{v^n_{j+2} - 3v^n_{j+1} + 3v^n_j - v^n_{j-1}}{2\Delta x} - \frac{v^n_j - v^n_{j-1}}{\Delta x} \\ &+ sg(a^n_{j-1/2})\frac{v^n_{j+1} - v^n_j - v^n_{j-1} + v^n_{j-2}}{2\Delta x} - \frac{v^n_{j+1} - 3v^n_j + 3v^n_{j-1} - v^n_{j-2}}{2\Delta x} \\ h^n_{j+1/2} - h^n_{j-1/2} &= \Delta x\left(\frac{v^n_{j+1} - 2v^n_j + v^n_{j-1}}{(\Delta x)^2}\right) \\ &+ \frac{(\Delta x)^3}{2}\left(\frac{v^n_{j+2} - 4v^n_{j+1} + 6v^n_j - 4v^n_{j-1} + v^n_{j-2}}{(\Delta x)^4}\right) \\ &- (\Delta x)^2\frac{1}{\Delta x}\left(sg(a^n_{j+1/2})\frac{v^n_{j+2} - v^n_{j+1} - v^n_j + v^n_{j-1}}{2(\Delta x)^2}\right. \\ &\left. -sg(a^n_{j-1/2})\frac{v^n_{j+1} - v^n_j - v^n_{j-1} + v^n_{j-2}}{2(\Delta x)^2}\right) \end{aligned}$$

Ceci montre que la discrétisation décentrée du terme d'antidiffusion introduit un terme de dispersion du troisième ordre et un terme de dissipation du quatrième ordre qui sont tous deux négligeables devant $\epsilon \mid a(u) \mid \frac{\partial^2 u}{\partial x^2}$ dès que Δx est suffisamment petit. En conclusion, quand on approche (B.6) par le schéma (B.8) plus une discrétisation centrée du terme de diffusion cela correspond à l'équation :

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} - \frac{\Delta x}{2}\frac{\partial}{\partial x}\left(\mid a(u) \mid \frac{\partial u}{\partial x}\right) - \nu\frac{\partial^2 u}{\partial x^2} + \epsilon \mid a(u) \mid \frac{\partial^2 u}{\partial x^2} \\ - \epsilon a(u)\Delta x \frac{\partial^3 u}{\partial x^3} + \epsilon \mid a(u) \mid \frac{(\Delta x)^2}{2}\frac{\partial^4 u}{\partial x^4} = 0 \end{aligned} \tag{B.10}$$

avec $\epsilon = \min\left(\frac{\nu}{|a(u)|}, \frac{\Delta x}{2}\right)$; et donc les termes supplémentaires d'ordres supérieurs sont aussi négligeables devant le terme visqueux.

Revenons au schéma (B.8) qui peut s'écrire sous la forme:

$$v_j^{n+1} = v_j^n - \begin{cases} \lambda(f(v_j^n) - (f(v_{j-1}^n)) \left[1 + \frac{\epsilon}{\Delta x}\left(\frac{v_j^n - 2v_{j-1}^n + v_{j-2}^n}{v_j^n - v_{j-1}^n}\right)\right] \\ \quad \text{si } a_{j+1/2}^n > 0 \text{ et } a_{j-1/2}^n > 0 \\ \lambda(f(v_{j+1}^n) - (f(v_j^n)) \left[1 - \frac{\epsilon}{\Delta x}\left(\frac{v_{j+2}^n - 2v_{j+1}^n + v_j^n}{v_{j+1}^n - v_j^n}\right)\right] \\ \quad \text{si } a_{j+1/2}^n < 0 \text{ et } a_{j-1/2}^n < 0 \\ \lambda(f(v_{j+1}^n) - (f(v_j^n)) \left[1 - \frac{\epsilon}{\Delta x}\left(\frac{v_{j+2}^n - v_{j+1}^n - v_{j-1}^n + v_{j-2}^n}{v_{j+1}^n - v_j^n}\right)\right] \\ + \lambda(f(v_j^n) - (f(v_{j-1}^n)) \left[1 + \frac{\epsilon}{\Delta x}\left(\frac{v_{j+2}^n - v_{j+1}^n - v_{j-1}^n + v_{j-2}^n}{v_j^n - v_{j-1}^n}\right)\right] \\ \quad \text{si } a_{j+1/2}^n < 0 \text{ et } a_{j-1/2}^n > 0 \\ 0 \qquad \text{sinon} \end{cases} \tag{B.11}$$

qui apparaît clairement comme une perturbation de (B.5) ou encore sous la forme :

$$v_j^{n+1} = v_j^n - \begin{cases} \lambda a_{j-1/2}^n \left[v_j^n - v_{j-1}^n + \frac{\epsilon}{\Delta x}(v_j^n - 2v_{j-1}^n + v_{j-2}^n)\right] \\ \quad \text{si } a_{j+1/2}^n > 0 \text{ et } a_{j-1/2}^n > 0 \\ \lambda a_{j+1/2}^n \left[v_{j+1}^n - v_j^n + \frac{\epsilon}{\Delta x}(v_{j+2}^n - 2v_{j+1}^n + v_j^n)\right] \\ \quad \text{si } a_{j+1/2}^n < 0 \text{ et } a_{j-1/2}^n < 0 \\ \lambda a_{j+1/2}^n \left[v_{j+1}^n - v_j^n + \frac{\epsilon}{\Delta x}(v_{j+2}^n - v_{j+1}^n - v_{j-1}^n + v_{j-2}^n)\right] \\ + \lambda a_{j-1/2}^n \left[v_j^n - v_{j-1}^n + \frac{\epsilon}{\Delta x}(v_{j+2}^n - v_{j+1}^n - v_{j-1}^n + v_{j-2}^n)\right] \\ \quad \text{si } a_{j+1/2}^n < 0 \text{ et } a_{j-1/2}^n > 0 \\ 0 \qquad \text{sinon} \end{cases} \tag{B.12}$$

Ces deux écritures montrent que le schéma (B.8) n'est pas conservatif à cause de l'expression de la dérivée seconde qui diffère dans le troisième cas des expressions obtenues dans les deux premiers cas (ceci en dépit de l'approche conservative utilisée). Pour le rendre conservatif il suffit de remplacer le troisième cas de (B.12) par :

$$\begin{aligned} v_j^{n+1} = v_j^n &- \lambda a_{j+1/2}^n \left[v_{j+1}^n - v_j^n + \frac{\epsilon}{\Delta x}(v_{j+2}^n - 2v_{j+1}^n + v_j^n)\right] \\ &- \lambda a_{j-1/2}^n \left[v_j^n - v_{j-1}^n + \frac{\epsilon}{\Delta x}(v_j^n - 2v_{j-1}^n + v_{j-2}^n)\right] \end{aligned} \tag{B.13}$$

qui a en plus l'avantage de faire apparaître le terme v_j^n avec le bon signe, ce qui va encore renforcer la stabilité. La différence entre le troisième cas de (B.12) et (B.13) est en $O(\Delta x)$ par rapport aux termes du deuxième ordre ; donc, quand Δx est suffisamment petit, le schéma (B.8)-(B.13) est un schéma conservatif équivalent à (B.8). Ces deux schémas appliqués à l'équation (B.6) avec $\epsilon = \min\left(\frac{\nu}{|a(u)|}, \frac{\Delta x}{2}\right)$ conservent les propriétés du schéma de Murman appliqué à l'équation (B.1) et convergent dès que Δx est suffisamment petit.

L'avantage essentiel de ces schémas (qui utilisent des différences finies décentrées pour discrétiser le terme d'antidiffusion) par rapport au schéma avec terme d'antidiffusion centré du paragraphe précédent, est qu'ils amènent des contributions positives sur la diagonale principale quand on les applique aux équations de Navier-Stokes. Discutons maintenant la valeur du paramètre qui théoriquement doit être égal à $\epsilon = \min\left(\frac{1}{Re|u|}, \frac{\Delta x}{2}\right)$ ou bien $\epsilon = \min\left(\frac{1}{Re|v|}, \frac{\Delta y}{2}\right)$ en dimension deux pour Navier-Stokes. Comme on peut difficilement envisager de prendre des pas de discrétisation aussi petits que l'on veut, cela signifie que dans une partie de l'écoulement on ne peut pas éliminer le terme de viscosité artificielle introduit par le schéma. Par exemple, si l'on estime que dans le cas de la cavité carrée $(0,1)\times(0,1)$ le pas le plus fin que l'on puisse prendre (pour des raisons évidentes de temps de calcul) est de l'ordre de $1/500$ cela veut dire que pour $Re = 10000$, le terme artificiel n'est pas éliminé dès que le module de la vitesse est supérieur à 0,1 ; or c'est le cas pour le tourbillon central qui occupe la majeure partie du domaine. Par contre, dans la couche limite et dans les coins, c'est bien le terme artificiel qui est éliminé ; néanmoins, la question est de savoir si numériquement on ne peut pas remplacer ϵ par $\epsilon' = \frac{\Delta x}{2}$ par exemple qui éliminerait toujours le bon terme.

Voyons ce que donne le schéma (B.8)-(B.13) avec $\epsilon' = \frac{\Delta x}{C}$ où C est une constante

positive:

$$v_j^{n+1} = v_j^n - \begin{cases} \Delta t\ a_{j-1/2}^n \left[\dfrac{(C+1)v_j^n - (C+2)v_{j-1}^n + v_{j-2}^n}{C\Delta x}\right] \\ \quad \text{si } a_{j+1/2}^n > 0 \text{ et } a_{j-1/2}^n > 0 \\ \Delta t\ a_{j+1/2}^n \left[\dfrac{-(C+1)v_j^n + (C+2)v_{j+1}^n - v_{j+2}^n}{C\Delta x}\right] \\ \quad \text{si } a_{j+1/2}^n < 0 \text{ et } a_{j-1/2}^n < 0 \\ \Delta t\ a_{j+1/2}^n \left[\dfrac{-(C+1)v_j^n + (C+2)v_{j+1}^n - v_{j+2}^n}{C\Delta x}\right] \\ + \Delta t\ a_{j-1/2}^n \left[\dfrac{(C+1)v_j^n - (C+2)v_{j-1}^n + v_{j-2}^n}{C\Delta x}\right] \\ \quad \text{si } a_{j+1/2}^n < 0 \text{ et } a_{j-1/2}^n > 0 \\ 0 \qquad \text{sinon} \end{cases} \tag{B.14}$$

Il apparaît clairement la discrétisation du terme $a(u)\frac{\partial u}{\partial x}$ avec un schéma de type Murman où la dérivée $\frac{\partial u}{\partial x}$ est discrétisée par différences finies décentrées ; les valeurs entières de C correspondent à des approximations du deuxième ordre et du troisième ordre de la dérivée au point $a + \frac{j-1/2}{N}(b-a)$ ou au point $a + \frac{j+1/2}{N}(b-a)$ et donc le schéma ci-dessus est un schéma du premier ordre correspondant à l'équation (B.1) écrite sous forme non conservative. Appliqué à (B.1) ce schéma est instable et la solution approchée explose rapidement mais il devient stable sur le problème visqueux (B.0) pour un bon choix de C. Malheureusement le schéma obtenu avec $C = 2$ est instable pour les nombres de Reynolds élevés, la valeur $C = 3$ apparaît comme la valeur optimale réalisant l'équilibre entre stabilité et représentation de la solution ; le schéma (B.8)-(B.13) avec $\epsilon' = \frac{\Delta x}{3}$ est celui présenté dans [BJ1] et utilisé dans [BJ2] pour calculer les solutions dans la cavité entraînée. Rappelons que pour l'équation de Burgers traitée au paragraphe précédent, la correction avec $C = 2$ a tendance à disperser la solution et que pour $C = 3$ on obtient quasiment la solution exacte. Le tableau ci-dessous permet de comparer les schémas (B.8) et (B.8)-(B.13) avec ϵ et $\epsilon' = \frac{\Delta x}{3}$ sur la cavité carrée (0,1)×(0,1) pour $Re = 1000$.

Grille		*Schéma* $(B.8)$ *avec* ϵ	*Schéma* $(B.8)$ *avec* ϵ'	*Schéma* $(B.8) - (B.13)$ *avec* ϵ	*Schéma* $(B.8) - (B.13)$ *avec* ϵ'
16×16	U min (position)	-0.2142 (0.1875)	-0.3258 (0.125)	-0.1907 (0.1875)	-0.2633 (0.125)
16×16	V min (position)	-0.3228 (0.875)	-0.3979 (0.875)	-0.2877 (0.875)	-0.3381 (0.875)
16×16	Nombre d' itérations	13	20	12	17
16×16	ω	0.5	0.5	0.5	0.5
32×32	U min (position)	-0.2985 (0.125)	-0.3614 (0.125)	-0.2880 (0.1562)	-0.3314 (0.125)
32×32	V min (position)	-0.4473 (0.9063)	-0.4923 (0.9063)	-0.4327 (0.9063)	-0.4584 (0.9063
32×32	Nombre d' itérations	11	13	9	13
32×32	ω	0.7	0.7	0.7	0.7
64×64	U min (position)	-0.3512 (0.1406)	-0.3624 (0.1406)	-0.3435 (0.1406)	-0.3556 (0.1406)
64×64	V min (position)	-0.4978 (0.9219)	-0.5068 (0.9219)	-0.4855 (0.9219)	-0.4959 (0.9219)
64×64	Nombre d' itérations	13	10	11	8
64×64	ω	0.6	0.5	0.7	0.7

Ces résultats sur trois grilles consécutives amènent plusieurs remarques ; les résultats sont très différents sur la grille la plus grossière et tout à fait comparables sur la grille 64×64 où la solution (voir tableau comparatif dans [BJ2]) est presque atteinte. La correction (B.13) rend le schéma conservatif plus diffusif et le paramètre ϵ' donne de meilleurs résultats que ϵ surtout quand le pas de discrétisation est gros (dans ces cas-là le minimum est atteint par $\frac{1}{Re|u|}$ dans la majeure partie du domaine et donc le correcteur d'antidiffusion est beaucoup plus faible que le terme de viscosité artificielle). Une dernière remarque est que le schéma (B.8)-(B.13) avec ϵ' converge le plus vite avec le paramètre de relaxation le plus élevé sur la grille fine ; il est plus stable que le schéma (B.8) avec ϵ' qui donne pourtant de meilleurs résultats.

Dans [BJ2] ce schéma a permis de calculer des solutions transitoires dans la cavité carrée pour des nombres de Reynolds supérieurs à 5000 sur des grilles très fines ($\Delta x = \Delta y = \frac{1}{512}$). Ces résultats indiquent un nombre de Reynolds critique R_c, $5000 < R_c \leq 7500$ dans le cas de la cavité carrée et nous conjecturons que la transition a lieu au moment où de petits tourbillons se développent le long des parois et détruisent les tourbillons secondaires.

Analyse numérique/*Numerical Analysis*

Un nouveau schéma décentré pour le problème de la cavité entraînée

Charles-Henri BRUNEAU et Claude JOURON

Résumé — Les équations de Navier-Stokes incompressibles stationnaires sont résolues numériquement dans la cavité entraînée $(0,1)\times(0,1)$. Les variables physiques, vitesse et pression, sont discrétisées sur des grilles décalées et une méthode multigrille avec relaxation maille à maille est utilisée pour obtenir une bonne convergence. Un nouveau schéma aux différences finies, décentrées, est proposé pour approcher les termes de convection. Il permet d'obtenir des solutions stables avec une bonne représentation des tourbillons secondaires. La transition vers la turbulence est observée numériquement pour $\mathrm{Re}=10\,000$.

A new upwind scheme for the driven cavity flow

Abstract — *An efficient numerical method is proposed for solving steady Navier-Stokes equations in the driven cavity* $(0,1)\times(0,1)$. *The primitive variables, pressure and velocity components, are discretized on staggered grids and a multigrid procedure with a cell by cell relaxation is used to achieve a good convergence. A new upwind differences scheme is proposed for the convection terms. Stable solutions exhibiting good representation of secondary vortices are obtained. The transition to turbulence is observed numerically for* $\mathrm{Re}=10{,}000$.

1. LE PROBLÈME DE LA CAVITÉ ENTRAÎNÉE. — Dans la cavité $\Omega=(0,1)\times(0,1)$ les équations de Navier-Stokes stationnaires pour un fluide incompressible s'écrivent

$$(1)\qquad -\frac{1}{\mathrm{Re}}\Delta u+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+\frac{\partial p}{\partial x}=0,\qquad -\frac{1}{\mathrm{Re}}\Delta v+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+\frac{\partial p}{\partial y}=0$$

$$(2)\qquad \operatorname{div}\mathbf{q}=0$$

où $\mathbf{q}=(u, v)$ est la vitesse, p la pression et Re le nombre de Reynolds. Le problème de la cavité entraînée consiste à associer aux équations (1)-(2) les conditions aux limites sur le bord $\partial\Omega$:

$$(3)\qquad \mathbf{q}=(1,0)\quad \textit{sur}\ \ \mathrm{I}=(0,1)\times\{1\},\qquad \mathbf{q}=\mathbf{0}\quad \textit{sur}\ \ \partial\Omega\setminus\mathrm{I}.$$

Les difficultés numériques du problème sont, d'une part de représenter correctement la pression qui est définie à une constante près et, d'autre part, de discrétiser les termes non linéaires de convection en tenant compte du sens de propagation des informations.

De nombreuses méthodes numériques ont été développées ces deux dernières décennies, parmi lesquelles on peut citer [1] à [5]. On trouvera des bibliographies plus exhaustives dans les ouvrages [6] et [7].

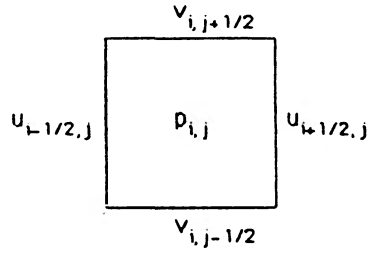


Fig. 1. — Les cinq inconnues de la maille (i, j).
Fig. 1. — The five unknowns in the cell (i, j).

Note présentée par Jacques-Louis LIONS.

0249-6291/88/03070359 $ 2.00

TABLEAU

Extremums de la fonction de courant.

Extrema of stream function.

	Tourbillon principal		Tourbillon secondaire en bas à droite		Tourbillon secondaire en bas à gauche	
	Φ	(localisation)	Φ	(localisation)	Φ	(localisation)
Ghia, Ghia, Shin	−0,1179	(0,5313, 0,5625)	0,175E-2	(0,8594, 0,1094)	0,231E-3	(0,0859, 0,0781)
Schreiber, Keller	−0,1160	(0,5286, 0,5643)	0,170E-2	(0,8643, 0,1071)	0,217E-3	(0,0857, 0,0714)
Vanka	−0,1138	(0,5313, 0,5664)	0,164E-2	(0,8672, 0,1133)	0,238E-3	(0,0820, 0,0781)
Méthode proposée	−0,1163	(0,5313, 0,5586)	0,191E-2	(0,8711, 0,1094)	0,325E-3	(0,0859, 0,0820)

2. MÉTHODE DE RÉSOLUTION. — Le problème approché de (1)-(2)-(3) est résolu par une méthode multigrille FMG-FAS [8] avec, sur chaque grille, une procédure de sous-relaxation maille à maille [3]. Cela consiste à résoudre implicitement sur une maille un système mettant en jeu les inconnues de la maille (composantes de la vitesse et pression), à sous-relaxer la solution obtenue puis à passer à la maille suivante. Outre son efficacité cette procédure permet d'ignorer la première difficulté signalée au paragraphe précédent.

L'ensemble de cette méthode est très performant mais la solution obtenue dépend presque uniquement de la capacité qu'a le schéma numérique à traiter correctement les termes de convection. C'est l'objet de cette Note qui présente un schéma à la fois stable et capable de représenter la solution avec un minimum de points de discrétisation.

3. DISCRÉTISATION. — Les inconnues du problème sont discrétisées sur des grilles décalées, la pression au centre des mailles et les composantes de la vitesse au milieu des côtés (*fig.* 1).

Tous les opérateurs linéaires de (1) et (2) sont approchés par différences finies centrées aux points considérés, la quantité de mouvement en u aux points $(i-1/2, j)$ et $(i+1/2, j)$, la quantité de mouvement en v aux points $(i, j-1/2)$ et $(i, j+1/2)$ et l'incompressibilité au centre de la maille. Par contre les termes de correction sont approchés en des points intermédiaires suivant les valeurs locales de la vitesse au voisinage du point considéré.

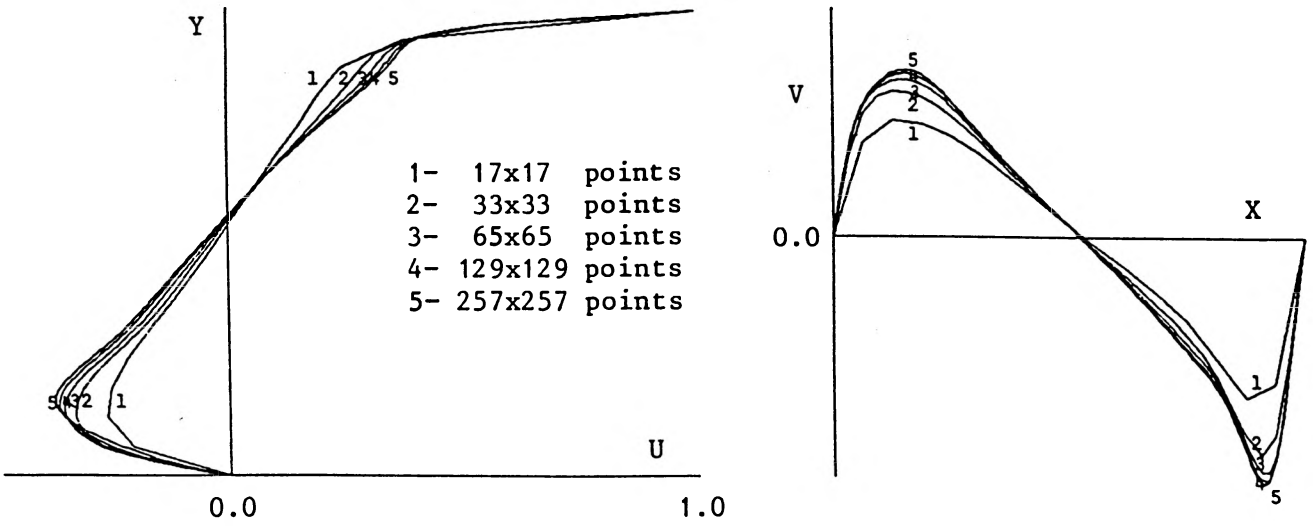


Fig. 2. — Re = 1 000. Profils des vitesses le long des axes médians pour différentes grilles.

Fig. 2. — *Re* = 1,000. *Velocity profiles along centerlines with different grids.*

Par exemple le terme $u(\partial u/\partial x)$ intervenant pour l'équation en $u_{i-1/2,j}$ est approché par :

$$\frac{u_{i-1,j}^{n-1}}{3\Delta x}(4u_{i-1/2,j}-5u_{i-3/2,j}^{n-1}+u_{i-5/2,j}^{n-1}) \quad \text{si} \quad u_{i-1,j}^{n-1}>0$$

$$-\frac{u_{i,j}^{n-1}}{3\Delta x}(4u_{i-1/2,j}-5u_{i+1/2,j}^{n-1}+u_{i+3/2,j}^{n-1}) \quad \text{si} \quad u_{i,j}^{n-1}<0$$

où Δx est le pas de discrétisation uniforme en x, n est l'indice de l'itération de relaxation et les valeurs $u_{i-1,j}^{n-1}$ et $u_{i,j}^{n-1}$ sont calculées par interpolation linéaire de la solution obtenue à l'itération précédente. L'approximation du terme $u(\partial u/\partial x)$ est donc un cumul (parfois réduit à zéro) de différences finies au point $(i-1, j)$ et au point (i, j); la dérivée $\partial u/\partial x$ est approchée en ces mêmes points par différences finies décentrées du deuxième ordre [9]. Ce schéma tient compte du sens de propagation des termes de convection. De plus il rajoute des termes positifs sur la diagonale principale du système 5×5 à résoudre sur la maille (i, j), ce qui a pour effet de renforcer la stabilité de la méthode.

4. RÉSULTATS NUMÉRIQUES. — Pour Re = 1 000 les résultats sont en accord avec ceux de la littérature [3], [4] et [5] (tableau) et la solution est déjà pratiquement atteinte avec un maillage 64×64 comme le montre la figure 2.

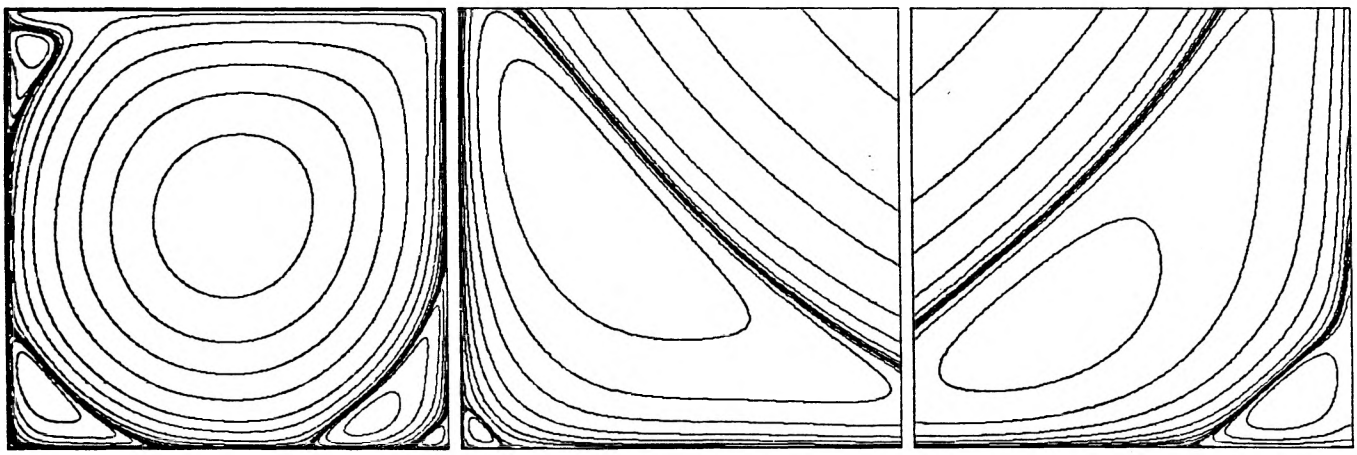


Fig. 3. — Re = 5 000. Lignes de courant pour la grille 512×512.
Fig. 3. — Re = 5,000. Streamlines with the 512×512 *grid.*

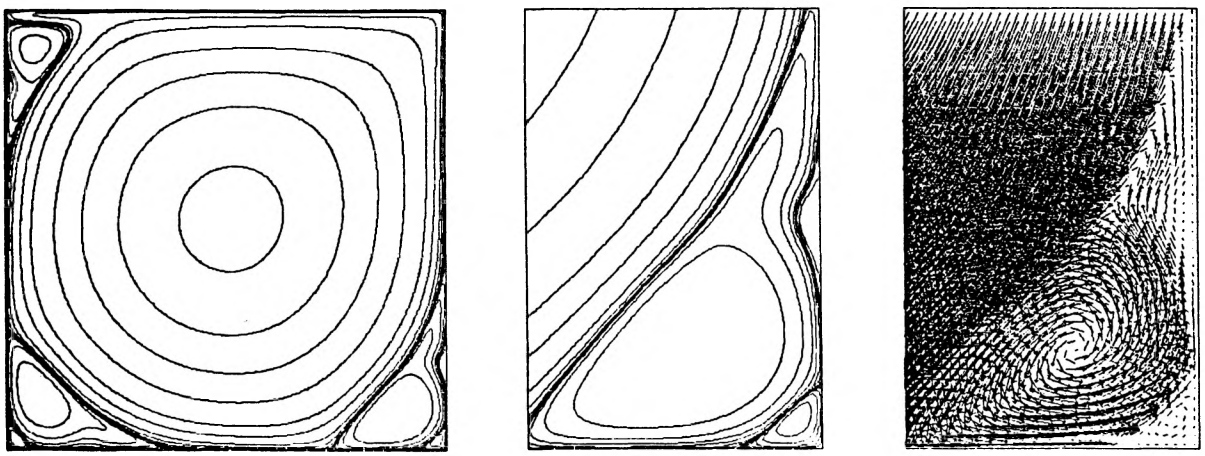


Fig. 4. — Re = 10 000. Lignes de courant et champ des vitesses pour la grille 256×256.
Fig. 4. — Re = 10,000. Streamlines and velocity field with the 256×256 *grid.*

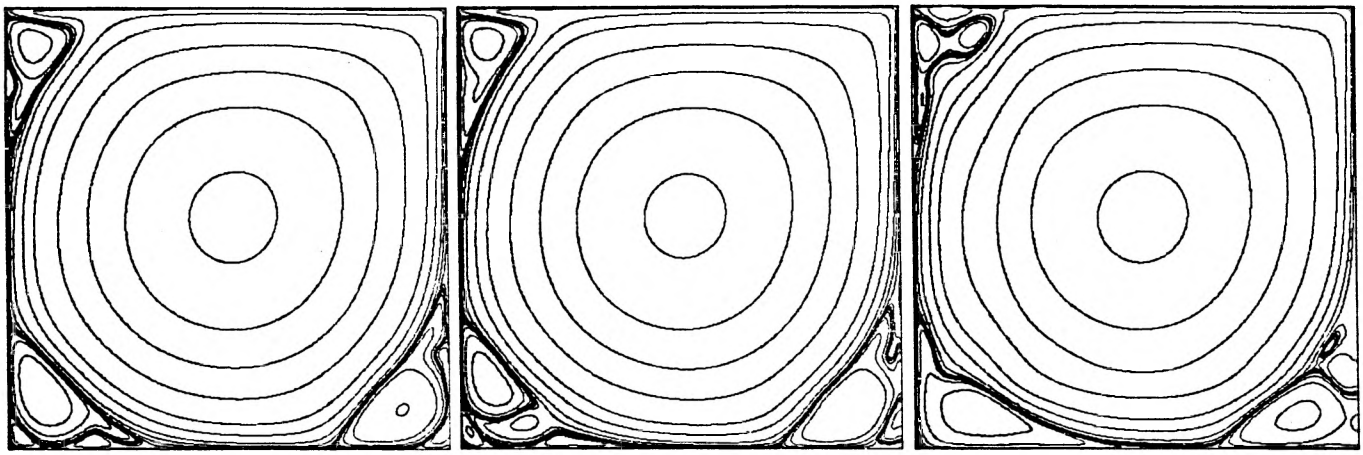


Fig. 5. – Re = 10 000. Évolution de la solution au cours des itérations pour la grille 512 × 512.
Fig. 5. – *Re* = 10,000. *Evolution of the solution during iterations with the* 512 × 512 *grid.*

Pour Re = 5 000 une solution laminaire stable est obtenue sur une grille 512 × 512 avec une représentation très fine des tourbillons secondaires et tertiaires (*fig.* 3). Finalement pour Re = 10 000 un phénomène particulièrement intéressant est observé. En effet la méthode ne converge pas sur les grilles grossières (jusqu'à 128 × 128) tant qu'il n'y a pas de points dans la couche limite. Puis elle converge très bien sur la grille 256 × 256 vers une solution laminaire stable présentant des tourbillons de parois supplémentaires (*fig.* 4). Enfin elle ne converge plus sur la grille 512 × 512 sur laquelle les phénomènes turbulents commencent à être pris en compte. La transition vers la turbulence est illustrée par l'évolution de la solution en fonction des itérations (*fig.* 5). Ces résultats montrent la présence d'un nombre de Reynolds critique qui se situe, pour le problème de la cavité entraînée (0,1) × (0,1), entre 5 000 et 10 000.

Note reçue le 26 mai 1988, acceptée le 6 juin 1988.

RÉFÉRENCES BIBLIOGRAPHIQUES

[1] M. FORTIN, R. PEYRET et R. TEMAM, Résolution numérique des équations de Navier-Stokes pour un fluide incompressible, *J. Mécanique*, 10, 1971, p. 357-390.
[2] K. GUSTAFSON et K. HALASI, Vortex dynamics of cavity flows, *J. Comput. Phys.*, 64, 1986, p. 279-319.
[3] S. P. VANKA, Block-implicit multigrid solution of Navier-Stokes equations in primitive variables, *J. Compt. Phys.*, 65, 1986, p. 138-158.
[4] U. GHIA, K. N. GHIA et C. T. SHIN, High-Re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method, *J. Comput. Phys.*, 48, 1982, p. 387-411.
[5] R. SCHREIBER et J. B. KELLER, Driven cavity flows by efficient numerical techniques, *J. Comput. Phys.*, 49, 1983, p. 310-333.
[6] R. PEYRET et T. D. TAYLOR, *Computational methods for fluid flow*, Springer, 1983.
[7] F. THOMASSET, *Implementation of finite element methods for Navier-Stokes equations*, Springer, 1981.
[8] A. BRANDT et N. DINAR, *Multigrid solutions to elliptic flow problems, Numerical Methods for Partial Differential Equations*, V. PARTER éd., Madison, 1978, p. 53-147.
[9] C. H. BRUNEAU et C. JOURON, *An efflcient scheme for solving steady incompressible Navier-Stokes equations*, en préparation.

Laboratoire d'Analyse numérique, Bât. n° 425, *Université Paris-Sud*, 91405 *Orsay Cedex.*

AN EFFICIENT SCHEME FOR SOLVING STEADY INCOMPRESSIBLE NAVIER-STOKES EQUATIONS

Charles-Henri BRUNEAU and Claude JOURON

Université Paris-Sud

Laboratoire d'Analyse Numérique, Bâtiment 425

91405 Orsay, France.

ABSTRACT.

The steady incompressible Navier-Stokes equations in a 2D driven cavity are solved in primitive variables by means of the multigrid method. The pressure and the components of the velocity are discretized on staggered grids, a block-implicit relaxation technique is used to achieve a good convergence and a simplified FMG-FAS algorithm is proposed.

Special focus on the finite differences scheme used to approach the convection terms is made and a large discussion with other schemes is given. Results in a square driven cavity are obtained for Reynolds numbers as high as 15000 on fine uniform meshes and the solution is in good agreement with other studies. For Re = 5000 the secondary vortices are very well represented showing the robustness of the method. For Reynolds numbers higher than 5000 the loss of stability for the steady solution is discussed. Moreover some computations on a rectangular cavity of aspect ratio equal to two are presented.

Subject classification : 65N05, 76D05

Key words : Incompressible flow, Steady Navier-Stokes equations, Multigrid method, Driven cavity.

Running head : Navier-Stokes equations solver.

In addition the method is very efficient as far as CPU time is concerned ; for instance the solution for Re = 1000 on a 128×128 grid is obtained within 24 seconds on a SIEMENS VP 200.

INTRODUCTION.

In the seventies, following some first works such as [1], many numerical methods were studied to solve the incompressible Navier-Stokes equations. These methods gave first indications on the shape of the solutions for Reynolds numbers up to 1000 but where limited by the lack of capacity of the computers. A quite exhaustive survey can be found in [2] and [3]. Today with the development of powerful computers with large memory and vector processor a new era of scientific computing is coming. In particular one can expect to make significant progress involving Navier-Stokes equations, to observe the transition to turbulence as done in [4] and in the future to compute turbulent flows ; according to Kolmogorov and [5] this will need as many degrees of freedom as $Re^{9/4}$ in three dimension. Such an amount of grid points for high Reynolds number requires huge memory which is not available yet but some work can already be done, in particular in two dimension. In this work we solve steady flows in a driven cavity and our aim is to observe, on one hand, solutions exhibiting good representation of secondary vortices for large Reynolds numbers and, on the other hand, the transition to turbulence.

Many approaches have been proposed this last decade to compute solutions of Navier-Stokes equations, among them the spectral method, the linearization methods and the multigrid method. The spectral or pseudo-spectral methods give good results for natural convection (see for instance [6] and references therein). They can also be used for the

regularized driven cavity problem [7]. The linearization by Newton's method either with a velocity-pressure formulation [4] or with a stream function-vorticity formulation [8] and finally the multigrid method in primitive variables [9] or in flow variables [10] are currently used. All these methods yield about the same steady solution for Re = 1000 and some good solutions for Re as high as 5000 or 10000 ; nevertheless for these high Reynolds numbers there are some noticeable discrepancies from one method to an other.

In brief we can say that everyone agrees on the main characteristics of the flow in a square cavity : the presence of a large primary vortex the center of which moves to the center of gravity of the square cavity as Re increases, the apparition of secondary vortices in the bottom corners and then in the corner of the incoming flow, the formation of quite large tertiary corner vortices for high Reynolds numbers [11]. But there is quite poor representation of these small structures in the corners and along the walls of the cavity. In [12] a local grid refinement procedure permits to capture, in a bottom corner, Moffat vortices with geometric progression scaling in good agreement with the theory ; this shows that small phenomena occuring in the boundary layer can be obtained by numerical methods if there are enough points to represent them. The questions are : are there other eddies than Moffat eddies ? And if there are other eddies, where are they located, how do they appear when Reynolds number increases and what is their role in the route to turbulence ?

In this work we try to give an answer to this question by computing the solution on fine grids for high Reynolds numbers in a square driven cavity and in a rectangular one of aspect ratio two ; the comparison of these two cases shows that some difficulties, occuring in the square driven cavity for Re > 5000, can appear already for Re = 1000 in the rectangular

one where the flow is much more unstable [13].

After some recalls on the equations we present in this paper a FMG-FAS (Full Multigrid-Full Approximation Storage) algorithm with simplified control tests and use the SCGS (Symmetrical Coupled Gauss-Seidel) smoothing procedure proposed in [9]. Then we study several finite difference schemes approximating the convection terms and show the influence of local testing of velocity sign on one hand and second order uncentered approximation of derivatives on the other hand to get an efficient scheme. Here by efficient we mean that the scheme is able to represent the solution as soon as there are enough points in the boundary layer. Finally we give results on grids as fine as 512×512 and a range of Reynolds number between 100 and 15000 for the square cavity and between 100 and 1000 for the rectangular cavity of aspect ratio two. The results are compared with those obtained in [8] [9] [10] and [14] for Reynolds numbers less or equal to 5000. For higher Reynolds numbers we observe the appearance of small eddies along the walls and conjecture that they play a significant role in the transition to turbulence.

1. MATHEMATICAL PROBLEM AND APPROXIMATION.

The non-dimensional Navier-Stokes equations of an incompressible viscous fluid in two-dimension are written as follows in primitive variables :

$$\begin{cases} -\frac{1}{Re}\Delta u + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{\partial p}{\partial x} = 0 \\ -\frac{1}{Re}\Delta v + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \frac{\partial p}{\partial y} = 0 \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \end{cases} \quad (1.1)$$

where $\vec{q} = (u,v)^T$ and p denote the velocity and the pressure respectively.

The Reynolds number we speak of in this paper is the one appearing in equations (1.1) although it is representative neither of the macro structures nor of the micro structures ; but it is better to take this one to compare with other authors. This work concerns only the well-known driven cavity problem, so equations (1.1) are set in a domain $\Omega = (0,1)\times(0,b)$ with b = 1 or 2 and are associated to the following Dirichlet boundary conditions :

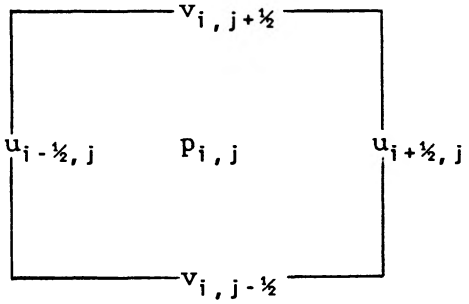
$$\begin{cases} u = 1 \text{ on } I = (0,1)\times\{b\} \\ u = 0 \text{ on } \partial\Omega\backslash I \\ v = 0 \text{ on } \partial\Omega. \end{cases} \tag{1.2}$$

Once again one could contest such a discontinuous boundary condition for u, in particular with spectral method u = 1 would be replaced by $u = (1 - (2x-1)^2)^2$ for instance, but for comparison motive we take condition (1.2) even if it is not the best one; moreover this discontinuity is not a main difficulty. Indeed the difficulties are, as it is well known, the representation of the pressure defined within an arbitrary constant and the treatment of the non linear convection terms. We shall see in the next sections how to avoid the first difficulty and how to treat the second one.

Let us rewrite equations (1.1) in the form :

$$\begin{cases} N(u,v) + \nabla p = 0 \\ \text{Div } \vec{q} = 0 \end{cases} \tag{1.3}$$

where N(u,v) is the sum of linear diffusion terms D(u,v) and non linear convection terms C(u,v) ; we approximate these operators by finite differences on staggered grids with the pressure given at the center of a cell and the components of the velocity given at the middles of the sides as usual (see figure 1).



Pressure and velocity in the cell (i,j)
Figure 1

Let G_ℓ be the fine grid on which we are looking for the solution we use centered finite differences to approximate the linear terms ; so on the cell (i,j) we have for the variables $u_{i-\frac{1}{2}j}$, $v_{ij-\frac{1}{2}}$ and p_{ij} the following formulas :

$$[D_\ell(u,v)]_{(i,j)} = \begin{cases} \dfrac{1}{Re(\Delta x)^2}(2u_{i-\frac{1}{2},j} - u_{i+\frac{1}{2},j} - u_{i-3/2,j}) + \\ \qquad + \dfrac{1}{Re(\Delta y)^2}(2u_{i-\frac{1}{2},j} - u_{i-\frac{1}{2},j+1} - u_{i-\frac{1}{2},j-1}) \\ \dfrac{1}{Re(\Delta x)^2}(2v_{i,j-\frac{1}{2}} - v_{i+1,j-\frac{1}{2}} - v_{i-1,j-\frac{1}{2}}) + \\ \qquad + \dfrac{1}{Re(\Delta y)^2}(2v_{i,j-\frac{1}{2}} - v_{i,j+\frac{1}{2}} - v_{i,j-3/2}) \end{cases}$$

$$[\nabla_\ell p]_{(i,j)} = \begin{cases} \dfrac{1}{\Delta x}(p_{i,j} - p_{i-1,j}) \\ \dfrac{1}{\Delta y}(p_{i,j} - p_{i,j-1}) \end{cases} \qquad (1.4)$$

$$[Div_\ell \vec{q}]_{(i,j)} = \frac{1}{\Delta x}(u_{i+\frac{1}{2},j} - u_{i-\frac{1}{2},j}) + \frac{1}{\Delta y}(v_{i,j+\frac{1}{2}} - v_{i,j-\frac{1}{2}})$$

where u,v,p are the discretized variables on the grid G_ℓ and Δx and Δy are the size of the uniform mesh in the x and y direction respectively.

In addition we propose in this paper to approximate the nonlinear convection term as follows :

$$
\begin{cases}
\dfrac{u^{n-1}_{i-1\,j}}{3\Delta x}\left(4u_{i-\frac{1}{2}j} - 5u^{n-1}_{i-3/2\,j} + u^{n-1}_{i-5/2\,j}\right) & \text{if } u^{n-1}_{i-1\,j} > 0 \\
-\dfrac{u^{n-1}_{i\,j}}{3\Delta x}\left(4u_{i-\frac{1}{2}j} - 5u^{n-1}_{i+\frac{1}{2}j} + u^{n-1}_{i+3/2\,j}\right) & \text{if } u^{n-1}_{i\,j} < 0
\end{cases}
$$

$$
+\begin{cases}
\dfrac{v^{n-1}_{i-\frac{1}{2}j-\frac{1}{2}}}{3\Delta y}\left(4u_{i-\frac{1}{2}j} - 5u^{n-1}_{i-\frac{1}{2}j-1} + u^{n-1}_{i-\frac{1}{2}j-2}\right) & \text{if } v^{n-1}_{i-\frac{1}{2}j-\frac{1}{2}} > 0 \\
-\dfrac{v^{n-1}_{i-\frac{1}{2}j+\frac{1}{2}}}{3\Delta y}\left(4u_{i-\frac{1}{2}j} - 5u^{n-1}_{i-\frac{1}{2}j+1} - u^{n-1}_{i-\frac{1}{2}j+2}\right) & \text{if } v^{n-1}_{i-\frac{1}{2}j+\frac{1}{2}} < 0
\end{cases}
$$

$$[C_k(u,v)]_{(i\,j)} = \tag{1.5}$$

$$
\begin{cases}
\dfrac{u^{n-1}_{i-\frac{1}{2}j-\frac{1}{2}}}{3\Delta x}\left(4v_{i\,j-\frac{1}{2}} - 5v^{n-1}_{i-1\,j-\frac{1}{2}} + v^{n-1}_{i-2\,j-\frac{1}{2}}\right) & \text{if } u^{n-1}_{i-\frac{1}{2}j-\frac{1}{2}} > 0 \\
-\dfrac{u^{n-1}_{i+\frac{1}{2}j-\frac{1}{2}}}{3\Delta x}\left(4v_{i\,j-\frac{1}{2}} - 5v^{n-1}_{i+1\,j-\frac{1}{2}} + v^{n-1}_{i+2\,j-\frac{1}{2}}\right) & \text{if } u^{n-1}_{i+\frac{1}{2}j-\frac{1}{2}} < 0
\end{cases}
$$

$$
+\begin{cases}
\dfrac{v^{n-1}_{i\,j-1}}{3\Delta y}\left(4v_{i\,j-\frac{1}{2}} - 5v^{n-1}_{i\,j-3/2} + v^{n-1}_{i\,j-5/2}\right) & \text{if } v^{n-1}_{i\,j-1} > 0 \\
-\dfrac{v^{n-1}_{i\,j}}{3\Delta y}\left(4v_{i\,j-\frac{1}{2}} - 5v^{n-1}_{i\,j+\frac{1}{2}} + v^{n-1}_{i\,j+3/2}\right) & \text{if } v^{n-1}_{i\,j} < 0
\end{cases}
$$

We shall see in details in section 4 how this scheme is constructed. We can not do it now because we need to describe the relaxation solver first.

In summary we get a nonlinear finite dimensional operator L_ℓ and we have to solve :

$$L_\ell(u,v,p) = 0 = f_\ell. \tag{1.6}$$

2. MULTIGRID PROCEDURE.

In order to solve (1.6) we use a FMG-FAS algorithm. So that this paper be relatively self-contained we describe briefly these multigrid procedures ; for more details the reader is refered to [15] and [16]. The Full Approximation Storage algorithm consists in evaluating a residu, projecting this residu on a coarser grid, computing a corrector on this coarse grid, then extending this corrector to the fine grid and correct the previous solution. One step can be summarized for a linear problem on two consecutive grids G_k and G_{k-1} as follow :

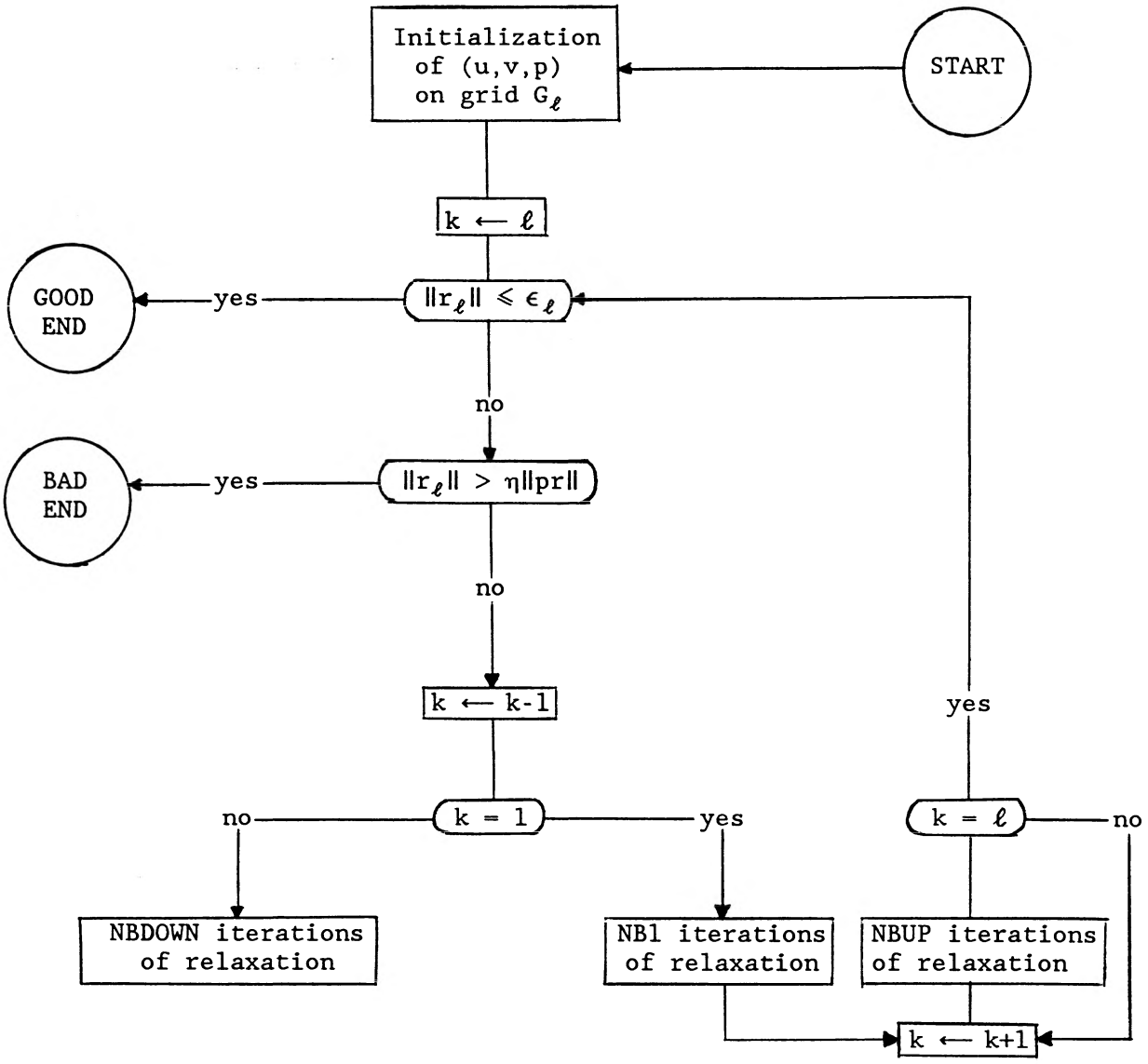
$$\begin{cases} \text{Solve } L_k(u_k, v_k, p_k) = f_k \\ r_k = f_k - L_k(u_k, v_k, p_k) \\ f_{k-1} = P_{k-1}^k \, r_k \\ \text{Solve } L_{k-1}(u_{k-1}, v_{k-1}, p_{k-1}) = f_{k-1} \\ (u_k, v_k, p_k) \longleftarrow (u_k, v_k, p_k) + E_k^{k-1}(u_{k-1}, v_{k-1}, p_{k-1}) \end{cases} \tag{2.1}$$

where (u_k, v_k, p_k) denote the variables on the grid G_k, P_{k-1}^k is the projection operator from the grid G_k to the grid G_{k-1} and E_k^{k-1} is the extension operator from the grid G_{k-1} to the grid G_k. For a nonlinear problem the algorithm is slightly different and can be written in the form :

$$\begin{cases} \text{Solve } L_k(u_k, v_k, p_k) = f_k \\ r_k = f_k - L_k(u_k, v_k, p_k) \\ f_{k-1} = P_{k-1}^k \, r_k + L_{k-1}(P_{k-1}^k(u_k, v_k, p_k)) \\ \text{Solve } L_{k-1}(u_{k-1}, v_{k-1}, p_{k-1}) = f_{k-1} \\ (u_k, v_k, p_k) \longleftarrow (u_k, v_k, p_k) + E_k^{k-1}((u_{k-1}, v_{k-1}, p_{k-1}) - P_{k-1}^k(u_k, v_k, p_k)) \end{cases} \tag{2.2}$$

For a set of grids $(G_k)_{k=1,\ldots,\ell}$ the control of the algorithm to go down and up to the coarser and finer grids can be done in different ways ; the usual one described in [9] and [15] is driven by precision tests of the

residual on the current grid. We notice that, with such a control, the FAS algorithm can go down and up between two levels for a long time and even never go up to the finest grid for high Reynolds numbers. So, as we solve $L_k(u_k, v_k, p_k) = f_k$ by a relaxation procedure, we use a much simpler control of the algorithm which consists in going down to the coarsest grid before going up, doing a fixed number of iterations on the intermediate grids, then going up to the finest grid. This FAS algorithm is illustrated on figure 2.



ϵ_ℓ : required precision on grid G_ℓ

r_ℓ : residual on grid G_ℓ

pr : previous residual on grid G_ℓ

η : coefficient of allowed divergence rate.

Organization chart of simplified FAS

Figure 2

The full Multi Grid algorithm consists in solving the problem with the FAS algorithm on a starting grid G_k $(k > 1)$ with the test $r_k < \epsilon_k$

before going up to a finer grid, then extend this solution to the grid G_{k+1} and so on until the finest grid. The only control on successive grids is the value of ϵ_k that is set to :

$$\epsilon_k = \delta\ \epsilon_{k+1}, \quad \delta > 1.$$

In practice for the simplified FMG-FAS procedure we take $\eta = 1.2$ and $\delta = 4$. That means that we allow small local divergence between two iterates and require on coarse grids less precision than on fine grids. In addition we take NBDOWN = 2, NBUP = 1 and NB1 = 10 for a 4×4 coarsest grid (obviously this last number depends on the choice of the coarsest grid). Our experience is that it is not necessary to do numerous iterations of relaxation during the going down or the going up and that 2 and 1 seems to be the best choice in term of CPU time. Nevertheless we need to do more iterations on the coarsest grid to have a good predictor ; we can either impose such a fixed number of iterations or require a precision test as follows :

$$\|r_1\| \leqslant \epsilon_1 = 10^{-2}\|r_\ell\|.$$

In other words, when we approach the convergence on the finest grid we require $\|r_1\| \leqslant 10^{-2}\epsilon_\ell$ on the coarsest grid.

Let us complete this section with the definition of P_{k-1}^{k} and E_{k}^{k-1}. These operations represent a tiny part of the CPU time ; so we can take good interpolations in order to decrease the number of global iterations on the finest grid. The projection operator is defined by :

$$
\begin{cases}
(u_{k-1})_{i-\frac{1}{2},j} = \frac{1}{4}\left((u_k)_{2i-3/2,2j} + (u_k)_{2i-3/2,2j-1}\right) \\
\qquad\qquad + \frac{1}{8}\left((u_k)_{2i-5/2,2j}+(u_k)_{2i-\frac{1}{2},2j}+(u_k)_{2i-5/2,2j-1}+(u_k)_{2i-\frac{1}{2},2j-1}\right) \\
(v_{k-1})_{i,j-\frac{1}{2}} = \frac{1}{4}\left((v_k)_{2i,2j-3/2} + (v_k)_{2i-1,2j-3/2}\right) \\
\qquad\qquad + \frac{1}{8}\left((v_k)_{2i,2j-5/2}+(v_k)_{2i,2j-\frac{1}{2}}+(v_k)_{2i-1,2j-5/2}+(v_k)_{2i-1,2j-\frac{1}{2}}\right) \\
(p_{k-1})_{i,j} = \frac{1}{4}\left((p_k)_{2i-1,2j}+(p_k)_{2i,2j}+(p_k)_{2i-1,2j-1}+(p_k)_{2i,2j-1}\right)
\end{cases}
$$

For the extension operator it is more important to have an accurate interpolation, specially when we extend a converged solution in the FMG procedure ; otherwise we lose some information on the finer grid and need more iterations. So the extension operator is defined by :

$$
\begin{aligned}
(u_k)_{2i-3/2,2j-1} = &-\frac{3}{128}\left((u_{k-1})_{i-3/2,j+1}+2(u_{k-1})_{i-\frac{1}{2},j+1}+(u_{k-1})_{i+\frac{1}{2},j+1}\right) \\
&+\frac{15}{64}\left((u_{k-1})_{i-3/2,j}+2(u_{k-1})_{i-\frac{1}{2},j}+(u_{k-1})_{i+\frac{1}{2},j}\right) \\
&+\frac{5}{128}\left((u_{k-1})_{i-3/2,j-1}+2(u_{k-1})_{i-\frac{1}{2},j-1}+(u_{k-1})_{i+\frac{1}{2},j-1}\right)
\end{aligned}
$$

$$
\begin{aligned}
(u_k)_{2i-\frac{1}{2},2j-1} = &-\frac{3}{64}\left((u_{k-1})_{i-\frac{1}{2},j+1}+(u_{k-1})_{i+\frac{1}{2},j+1}\right) \\
&+\frac{15}{32}\left((u_{k-1})_{i-\frac{1}{2},j}+(u_{k-1})_{i+\frac{1}{2},j}\right) \\
&+\frac{5}{64}\left((u_{k-1})_{i-\frac{1}{2},j-1}+(u_{k-1})_{i+\frac{1}{2},j-1}\right)
\end{aligned}
$$

$$
\begin{aligned}
(v_k)_{2i-1,2j-3/2} = &-\frac{3}{128}\left((v_{k-1})_{i+1,j-3/2}+2(v_{k-1})_{i+1,j-\frac{1}{2}}+(v_{k-1})_{i+1,j+\frac{1}{2}}\right) \\
&+\frac{15}{64}\left((v_{k-1})_{i,j-3/2}+2(v_{k-1})_{i,j-\frac{1}{2}}+(v_{k-1})_{i,j+\frac{1}{2}}\right) \\
&+\frac{5}{128}\left((v_{k-1})_{i-1,j-3/2}+2(v_{k-1})_{i-1,j-\frac{1}{2}}+(v_{k-1})_{i-1,j+\frac{1}{2}}\right)
\end{aligned}
$$

$$(v_k)_{2i-1,2j-\frac{1}{2}} = -\frac{3}{64}\left((v_{k-1})_{i+1,j-\frac{1}{2}} + (v_{k-1})_{i+1,j+\frac{1}{2}}\right) + \frac{15}{32}\left((v_{k-1})_{i,j-\frac{1}{2}} + (v_{k-1})_{i,j+\frac{1}{2}}\right) + \frac{5}{64}\left((v_{k-1})_{i-1,j-\frac{1}{2}} + (v_{k-1})_{i-1,j+\frac{1}{2}}\right)$$

$$(p_k)_{2i,2j} = \frac{1}{16}\left(3(p_{k-1})_{i,j+1} + (p_{k-1})_{i+1,j+1} + 9(p_{k-1})_{i,j} + 3(p_{k-1})_{i+1,j}\right)$$

3. RELAXATION PROCEDURE

In the multigrid technique the choice of the relaxation procedure plays an important role in the efficiency of the method. We use the smoothing operator, proposed in [9], called the Symmetrical Coupled Gauss-Seidel procedure (SCGS) which consists in a cell by cell under-relaxation procedure. This technique is satisfying because it leads us to solve the problem in a control volume, here a cell, and to update the solution in this control volume. That means, with staggered grids, that the five unknowns of a cell (i,j) are computed simultaneously by solving a 5×5 linear system of the form :

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 & 1/\Delta x \\ 0 & a_{22} & 0 & 0 & -1/\Delta x \\ 0 & 0 & a_{33} & 0 & 1/\Delta y \\ 0 & 0 & 0 & a_{44} & -1/\Delta y \\ -1/\Delta x & 1/\Delta x & -1/\Delta y & 1/\Delta y & 0 \end{pmatrix} \begin{pmatrix} u_{i-\frac{1}{2},j} \\ u_{i+\frac{1}{2},j} \\ v_{i,j-\frac{1}{2}} \\ v_{i,j+\frac{1}{2}} \\ p_{i,j} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix} \qquad (3.1)$$

where the diagonal terms contain the diagonal contributions of D(u,v) and C(u,v) and the second members contain all the remaining terms of L_k and the

second member f_k. We could have put some other terms in the matrix to intensify the interaction between the unknowns but it seems that such a choice with dominant positive terms on the main diagonal is more efficient. We shall see in the next section how to approximate the convection terms to improve this property. After computing the five unknowns on a cell we update the five corresponding variables by under relaxation, for instance we have :

$$u^n_{i-\frac{1}{2},j} = \omega\, u_{i-\frac{1}{2},j} + (1-\omega)u^{n-1}_{i-\frac{1}{2},j} \quad ; \; 0 < \omega \leqslant 1. \tag{3.2}$$

where n is the index of relaxation iterations. Then we solve the problem on the next cell. So, during one iteration, the pressure is updated once and the velocity twice because the components are given on the sides of the cells. In addition we choose a global cell by cell computation from the top to the bottom of the cavity to take immediately into account the driven condition. We conclude this section by giving in table 1 the value of the optimal under-relaxation parameter ω for various Reynolds numbers in the square cavity and in the rectangular one. In fact this parameter is closely related to the solution itself, so for a given geometry it depends only on Re but there is a large difference from one geometry to another.

Re	100	400	1000	5000	10000
Domain (0.1)×(0.1)	1	0.9	0.7	0.3	0.2
Domain (0.1)×(0.2)	0.8	0.3	0.1	-	-

Value of the relaxation parameter with the scheme (1.4)-(1.5)
Table 1

4. UNCENTERED SCHEMES FOR CONVECTION TERMS.

In this section we analyze in details the influence of the

discretization of the convection terms to the computed solution as announced in section 1. All the multigrid procedure with smoothing operator described above gives a good efficiency but it does not have any influence on the solution. On the contrary the approximation of convection terms has a huge influence on the obtained solution except for small Reynolds numbers (Re $\leqslant$ 100) where every scheme gives a good solution within a few iterations ; even the centered scheme written as follows for $u_{i-\frac{1}{2},j}$ and $v_{i,j-\frac{1}{2}}$:

$$C_k(u,v)]_{(i,j)} = \begin{cases} \dfrac{u_{i-\frac{1}{2}j}}{2\Delta x}\left(u^{n-1}_{i+\frac{1}{2},j} - u^{n-1}_{i-3/2,j}\right) + \dfrac{v^{n-1}_{i-\frac{1}{2}j}}{2\Delta y}\left(u^{n-1}_{i-\frac{1}{2},j+1} - u^{n-1}_{i-\frac{1}{2},j-1}\right) \\ \dfrac{u^{n-1}_{ij-\frac{1}{2}}}{2\Delta x}\left(v^{n-1}_{i+1,j-\frac{1}{2}} - v^{n-1}_{i-1,j-\frac{1}{2}}\right) + \dfrac{v_{ij-\frac{1}{2}}}{2\Delta y}\left(v^{n-1}_{i,j+\frac{1}{2}} - v^{n-1}_{i,j-3/2}\right) \end{cases} \qquad (4.1)$$

where the unknown values $u^{n-1}_{i,j-\frac{1}{2}}$ and $v^{n-1}_{i-\frac{1}{2},j}$ are computed by linear interpolation of the four known neighbours. Replacing formula (1.5) this formula completes formulas (1.4) ; we see that it adds terms on the main diagonal of matrix (3.1) with positive or negative sign according to the sign of $u^{n-1}_{i+\frac{1}{2},j} - u^{n-1}_{i-3/2,j}$ and $v^{n-1}_{i,j+\frac{1}{2}} - v^{n-1}_{i,j-3/2}$. This plays a significant role in the behaviour of the centered scheme that is good for small Reynolds numbers but very unstable and dispersive as soon as Re increases. Moreover the centered scheme does not take into account the hyperbolic behaviour of the convection terms. So the first idea is to use for these terms an upwind scheme given for the first unknown $u_{i-\frac{1}{2},j}$ by :

$$\begin{cases} \dfrac{u^{n-1}_{i-\frac{1}{2},j}}{\Delta x}\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-3/2,j}\right) & \text{if } u^{n-1}_{i-\frac{1}{2},j} > 0 \\ \dfrac{u^{n-1}_{i-\frac{1}{2},j}}{\Delta x}\left(u^{n-1}_{i+\frac{1}{2},j} - u_{i-\frac{1}{2},j}\right) & \text{otherwise} \end{cases} \qquad (4.2)$$

$$+ \begin{cases} \dfrac{v^{n-1}_{i-\frac{1}{2},j}}{\Delta y}\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-\frac{1}{2},j-1}\right) & \text{if } v^{n-1}_{i-\frac{1}{2},j} > 0 \\ \dfrac{v^{n-1}_{i-\frac{1}{2},j}}{\Delta y}\left(u^{n-1}_{i-\frac{1}{2},j+1} - u_{i-\frac{1}{2},j}\right) & \text{otherwise} \end{cases}$$

Now this scheme is sensitive to the propagation of the informations with hyperbolic terms and, besides, the additional terms on the main diagonal of the matrix (3.1) are always positive. Consequently this scheme is stable for high Reynolds numbers but, as it is often the case for uncentered first order schemes (see for instance [17]) it is very diffusive. So both the location and the amplitude of the extrema of velocity profiles are uncorrect as shown on figure 3 for Re = 1000 on a medium 64×64 grid .

At this stage we see that we have to find an intermediate scheme both stable and able to capture correctly the solution as soon as there are enough points to represent it. In order to improve the location of the phenomena we write an other uncentered first order scheme sensitive to the solution in the neighbourhood of the cell (i,j) ; that is for $u_{i-\frac{1}{2},j}$:

$$
\begin{cases}
\dfrac{u^{n-1}_{i-1,j}}{\Delta x}\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-3/2,j}\right) & \text{if } u^{n-1}_{i-1,j} > 0 \\
+\dfrac{u^{n-1}_{i,j}}{\Delta x}\left(u^{n-1}_{i+\frac{1}{2},j} - u_{i-\frac{1}{2},j}\right) & \text{if } u^{n-1}_{i,j} < 0
\end{cases}
\qquad (4.3)
$$

$$
+ \begin{cases}
\dfrac{v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}}}{\Delta y}\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-\frac{1}{2},j-1}\right) & \text{if } v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}} > 0 \\
\dfrac{v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}}}{\Delta y}\left(u^{n-1}_{i-\frac{1}{2},j+1} - u_{i-\frac{1}{2},j}\right) & \text{if } v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}} < 0
\end{cases}
$$

where $u^{n-1}_{i-1,j}$, $u^{n-1}_{i,j}$, $v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}}$ and $v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}}$ can be seen as the mean component of the velocity in a left, right, bottom and top cell of the current point $(i-\frac{1}{2},j)$ respectively. This scheme, inspired by Murman's, is much more stable than the previous one and gives a quite good location of the extrema of velocity profiles (see figure 3) but, unfortunately, it needs many points to get the right amplitude as the scheme (4.2). We can remark that the convection terms in (4.3) are not written at point $(i-\frac{1}{2},j)$ but at points $(i-1,j)$, (i,j), $(i-\frac{1}{2},j-\frac{1}{2})$ and $(i-\frac{1}{2},j+\frac{1}{2})$ instead ; then the derivatives in these terms are approximated by centered second order differences at these points.

We propose finally to replace these derivatives by uncentered second order differences at these neighbouring points. Let us denote by x one of these points we can write Taylor's formulas :

$$
\begin{cases}
u(x-h/2) = u(x) - h/2\ u'(x) + h^2/8\ u''(x) + O(h^3) \\
u(x-3h/2) = u(x) - 3h/2\ u'(x) + 9h^2/8\ u''(x) + O(h^3)
\end{cases}
\qquad (4.4)
$$

From (4.4) we have : $u'(x) = \dfrac{8u(x) - 9u(x-h/2) + u(x-3h/2)}{3h} + O(h^2)$, and with linear interpolation of the unknown value u(x) we find for the first derivative :

$$u'(x) = \frac{4u(x+h/2) - 5u(x-h/2) + u(x-3h/2)}{3h} \qquad (4.5)$$

Instead of (4.4) or in addition to (4.4) we could have written Taylor's series for u(x-h) and u(x+h/2) ; then we get many possible combinations to determine the first derivative with a second or third order of accuracy. Among all these possibilities the good formulas have the form :

$$u'(x) = \frac{\alpha u(x+h/2) - (\alpha+1)u(x-h/2) + u(x-3h/2)}{(\alpha-1)h} \text{ with } \alpha = 3,4,5,7 \text{ and } 13 \ ;$$

the two latest are of third order. These formula correspond to introducing an anti-diffusion term $\frac{h|u|}{(\alpha-1)} u''(x)$, which reduces the artificial diffusion contained in the first order scheme (4.3). For $\alpha = 3$ the resulting scheme is very unstable even with small relaxation parameters. On the contrary all the other values of α yield a stable scheme that gives better results than (4.3). Thus, as α increases, the scheme is more stable and the results less good because there is more and more artificial diffusion in the scheme ; the case $\alpha = 4$ can be seen as a limit of stability and, as it is often the case, gives the best results. Finally the value $\alpha = 4$ yields the proposed scheme (1.5).

On figure 3 we compare this scheme with (4.2) and (4.3) schemes and those used by Vanka in [9]. We see that for Re = 1000 and a medium mesh of 64×64 cells the scheme (4.5) gives a good location and amplitude of the extrema of velocity profiles. The results on a finer mesh are in good agreement with the results presented in [8], [9], [10] and [14] as shown in table 2 although the secondary vortices are larger with the present scheme, we think that this is due to the fact that our scheme is less diffusive. Moreover we see on figure 4 that the results obtained on the 64×64 mesh are

	u_{min}	y_{min}	v_{max}	x_{max}	v_{min}	x_{min}
Ghia,Ghia,Shin	-0.3829	0.1719	0.3709	0.1563	-0.5155	0.9063
Vanka (*)	-0.3798	0.1680	0.3669	0.1563	-0.5186	0.9102
Zhang	-0.3901	0.1699	0.3785	0.1582	-0.5284	0.9082
Present method(*)	-0.3764	0.1602	0.3665	0.1523	-0.5208	0.9102

Extrema of velocity profiles along centerlines

	Primary vortex		Secondary vortex bottom right		Secondary vortex bottom left	
	Φ	(location)	Φ	(location)	Φ	(location)
Ghia,Ghia,Shin	-0.1179	(0.5313,0.5625)	0.175E-2	(0.8594,0.1094)	0.231E-3	(0.0859,0.0781)
Schreiber,Keller	-0.1160	(0.5286,0,5643)	0.170E-2	(0.8643,0.1071)	0.217E-3	(0.0857,0.0714)
Vanka (*)	-0.1138	(0.5313,0.5664)	0.164E-2	(0.8672,0.1133)	0.238E-3	(0.0820,0.0781)
Zhang	-0.1193	(0.5313,0.5664)	0.174E-2	(0.8633,0.1133)	0.235E-3	(0.0820,0.0781)
Present method (*)	-0.1163	(0.5313,0.5586)	0.191E-2	(0.8711,0.1094)	0.325E-3	(0.0859,0.0820)

Extrema of stream function

(*) Results computed in this work on a 256×256 grid.

Comparison of solutions for Re = 1000

Table 2

already closed to the solution, and closer than those obtained with Vanka's scheme (figure 3). In this last scheme the diffusion term and the convection term are closely linked and the diffusion term is suppressed as soon as the convection term prevails. We can write $[D_k(u,v)]_{(i,j)} + [C_k(u,v)]_{(i,j)}$ for the unknown $u_{i-\frac{1}{2},j}$ as follows :

$$\begin{cases} \left(\frac{\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-3/2,j}\right)}{Re(\Delta x)^2} + \frac{u^{n-1}_{i-1,j}}{2\Delta x}\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-3/2,j}\right)\right. & \text{if } \left|\frac{u^{n-1}_{i-1,j}}{2\Delta x}\right| < \frac{1}{Re(\Delta x)^2} \\ \left.\left(\left|\frac{u^{n-1}_{i-1,j}}{2\Delta x}\right| + \frac{u^{n-1}_{i-1,j}}{2\Delta x}\right)\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-3/2,j}\right)\right. & \text{otherwise} \end{cases}$$

$$+ \begin{cases} \left(\frac{\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i+\frac{1}{2},j}\right)}{Re(\Delta x)^2} + \frac{u^{n-1}_{i,j}}{2\Delta x}\left(u^{n-1}_{i+\frac{1}{2},j} - u_{i-\frac{1}{2},j}\right)\right. & \text{if } \left|\frac{u^{n-1}_{i,j}}{2\Delta x}\right| < \frac{1}{Re(\Delta x)^2} \\ \left.\left(\left|\frac{u^{n-1}_{i,j}}{2\Delta x}\right| - \frac{u^{n-1}_{i,j}}{2\Delta x}\right)\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i+\frac{1}{2},j}\right)\right. & \text{otherwise} \end{cases}$$

$$+ \begin{cases} \left(\frac{\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-\frac{1}{2},j-1}\right)}{Re(\Delta y)^2} + \frac{v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}}}{2\Delta y}\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-\frac{1}{2},j-1}\right)\right. & \text{if } \left|\frac{v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}}}{2\Delta y}\right| < \frac{1}{Re(\Delta y)^2} \\ \left.\left(\left|\frac{v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}}}{2\Delta y}\right| + \frac{v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}}}{2\Delta y}\right)\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-\frac{1}{2},j-1}\right)\right. & \text{otherwise} \end{cases}$$

$$+ \begin{cases} \left(\frac{\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-\frac{1}{2},j+1}\right)}{Re(\Delta y)^2} + \frac{v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}}}{2\Delta y}\left(u^{n-1}_{i-\frac{1}{2},j+1} - u_{i-\frac{1}{2},j}\right)\right. & \text{if } \left|\frac{v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}}}{2\Delta y}\right| < \frac{1}{Re(\Delta y)^2} \\ \left.\left(\left|\frac{v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}}}{2\Delta y}\right| - \frac{v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}}}{2\Delta y}\right)\left(u_{i-\frac{1}{2},j} - u^{n-1}_{i-\frac{1}{2},j+1}\right)\right. & \text{otherwise} \end{cases}$$

where $u^{n-1}_{i-1,j}$, $u^{n-1}_{i,j}$, $v^{n-1}_{i-\frac{1}{2},j-\frac{1}{2}}$ and $v^{n-1}_{i-\frac{1}{2},j+\frac{1}{2}}$ are defined as in (4.3). We see that this scheme is a good intermediate between the centered scheme (4.1) and the scheme (4.3), which could be the key of its success. Nevertheless the suppression of the diffusion term in some cases can decrease its stability and explain the difficulties to get the solution for Reynolds number equals to 5000. To close this section we add that the scheme

presented in [14] is the same than (4.2) with a second order approximation of the derivatives of the form : $u'(x) = \dfrac{3u(x) - 4u(x-h) + u(x-2h)}{2h}$. It gives good results but is not very stable either as we have seen above for $\alpha = 3$.

We do not pretend to make an exhaustive discussion about the schemes but hope to shed some light on the difficult problem of the approximation of convection terms.

We close this section by giving the extrapolation at the boundary ; as the scheme (1.4)-(1.5) uses second order derivatives we take a parabolic extrapolation written for the unknown u as follow on the left hand side :

$$u_{-\frac{1}{2},j} = 3\, u_{\frac{1}{2},j} - 3\, u_{3/2,j} + u_{5/2,j}$$

and at the bottom :

$$u_{i-\frac{1}{2},0} = \frac{8}{3}\, u_{i-\frac{1}{2},\frac{1}{2}} - 2\, u_{i-\frac{1}{2},1} + \frac{1}{3}\, u_{i-\frac{1}{2},2}$$

$$u_{i-\frac{1}{2},-1} = 8\, u_{i-\frac{1}{2},\frac{1}{2}} - 9\, u_{i-\frac{1}{2},1} + 2\, u_{i-\frac{1}{2},2}$$

This extrapolation improves the representation of the boundary conditions and gives better results on coarse grids. But on fine meshes the first order extrapolation yields about the same results.

5. NUMERICAL RESULTS.

In this section we comment some results in the square driven cavity and in the rectangular cavity of aspect ratio two. These results

are obtained with the simplified FMG-FAS algorithm and the finite differences schemes (1.4)-(1.5). Generally the solution is initialized by zero on the starting grid of the FMG algorithm but in some cases, particularly for high Reynolds numbers, it is initialized on this grid with the solution obtained for a lower Reynolds number. The pressure is also initialized by zero everywhere and is not fixed at any point. Indeed we know that the pressure is determined within a constant but this constant is given by the initialization ; the relaxation procedure allows to let the pressure free without getting overshoots. In fact the results show that the pressure, in this case of zero initialization, is always less than one.

First of all we compare the efficiency of the present method with other methods. For Re = 1000 we get the solution on a 128×128 grid in only 24 seconds of SIEMENS VP 200. In [10] with a stream function-vorticity formulation the same test requires 1.5 minutes on the AMDAHL 470V/6, which is of the same order of magnitude. But we can compare more precisely the efficiency of our scheme with Vanka's scheme because we have tested both of them. In table 3 we give a comparison for Re = 1000 on a 128×128 grid with the simplified FMG-FAS algorithm and on a 256×256 grid with FAS algorithm and initialization by the solution obtained on the previous grid. We notice that the scheme (1.4)-(1.5) is approximately twice as fast as Vanka's scheme on the 128×128 grid and that the optimal relaxation parameter is larger.

	128×128 grid			256×256 grid		
	ω optimal	number of iterations	CPU time	ω optimal	number of iterations	CPU time
Vanka	0.6	20	53"	0.5	10	1'41"
Present method	0.7	7	24"	0.7	6	1'10"

Comparison of efficiency with Vanka's scheme

Table 3

Re	100	1000	5000
u_{min} y_{min}	-0.2106 0.4531	-0.3764 0.1602	-0.4359 0.0664
v_{max} x_{max}	0.1786 0.2344	0.3665 0.1523	0.4259 0.0762
v_{min} x_{min}	-0.2521 0.8125	-0.5208 0.9102	-0.5675 0.9590
Primary vortex - Φ (location)	-0.1026 (0.6172,0.7344)	-0.1163 (0.5313,0.5586)	-0.1142 (0.5156,0.5313)
Second. vortex bottom right Φ (location)	0.123E-4 (0.9453,0.0625)	0.191E-2 (0.8711,0.1094)	0.465E-2 (0.8301,0.0703)
Second. vortex bottom left Φ (location)	0.163E-5 (0.0313,0.0391)	0.325E-3 (0.0859,0.0820)	0.222E-2 (0.0664,0.1484)
Second. vortex top left Φ (location)	-	0.755E-3 (0.0039,1.0000)	0.175E-2(0.0625,0.9102)
Tertiary vortex bottom right Φ (location)	-	-	-0.247E-4 (0.9668,0.0293)
Tertiary vortex bottom left Φ (location)	-	-0.306E-8 (0.0039,0.0039)	-0.233E-6 (0.0117,0.0098)

Extrema of velocity profiles along centerlines and stream function

Table 4

In table 4 are given the main characteristics of the solution for Re = 100 on a 128×128 grid, Re = 1000 on a 256×256 grid and Re = 5000 on a 512×512 grid. In every case we observe a good convergence to the steady laminar solution with a residual r_ℓ less that 10^{-4} . Let us notice that, in the three cases, we represent the solution on the finest grid reached by FMG algorithm that stops when the solution does not change noticeably. The solution depends on the grid when the grids are coarse and is independent when the grids are finer. For instance this is illustrated on the figure 4 where we see that the solution is quite different on the coarse grids until 64×64 and then is the same on the grids 128×128 and 256×256. In this case the grid 256×256 is very fine with respect to Reynolds number 1000 but confirms the stability of the solution. We can affirm that we get a steady laminar solution until Re = 5000.

Moreover, in accordance with [8], the minimum of the stream function increases between Re = 1000 and Re = 5000. On figures 5 and 6 are plotted these solutions in the square cavity with details of secondary and tertiary vortices for Re = 5000. The streamlines plotted on the figures are either equidistant with a step of 10^{-3} (the separation lines represented by level zero are denoted by 0) or fixed to forty given values. These values are listed in table 5.

Contour letter	Value of Φ	Contour letter	Value of Φ	Contour letter	Value of Φ	Contour letter	Value of Φ
y	-0.1	t	-1.E-2	e	-1.E-10	ζ	1.E-9
x	-0.08	s	-3.E-3	d	-1.E-11	η	1.E-8
w	-0.06	r	-1.E-3	c	-1.E-12	ω	1.E-7
v	-0.04	p	-3.E-4	b	-1.E-13	ι	1.E-6
u	-0.02	n	-1.E-4	a	-1.E-14	κ	3.E-6
		m	-3.E-5	o	0	λ	1.E-5
		l	-1.E-5	α	1.E-14	μ	3.E-5
		k	-3.E-6	β	1.E-13	ν	1.E-4
		i	-1.E-6	γ	1.E-12	ψ	3.E-4
		h	-1.E-7	δ	1.E-11	ρ	1.E-3
		g	-1.E-8	ϵ	1.E-10	σ	3.E-3
		f	-1.E-9			τ	1.E-2

Values of the streamlines
Table 5

Now we notice a very interesting behaviour of the computation process for Re = 7500 and Re = 10000. Indeed on coarse grids (until 128×128 for Re = 10000) there is no convergence of the residual ; on the 256×256 grid there is a good convergence to a steady laminar solution (figure 7) and on the finer 512×512 grid there is no convergence any more (the residual decreases at the beginning and then oscillates around a value) and apparition of new small eddies (figure 8). In this late case the relaxation procedure can be seen like a time procedure, and then we observe an evolution of the small eddies along the walls. Our explanation for this behaviour is the following : if there is no points in the boundary layer, the solution cannot be represented properly and the method does not converge. We know that the thickness of the boundary layer is of the order of $1/\sqrt{Re}$; so this explains the non convergence on coarse grids. For smaller values of the mesh size the process converges to the steady laminar solution if it exists; this is the case for Re = 5000 for instance. But, as we use a steady model, the method can converge artificially on quite fine grids as long as the mesh size is not fine enough to take into account the instabilities. This is the case for Re = 7500 and Re = 10000 on the 256×256 grid. The Kolmogorov length is of the order of $1/Re^{3/4}$ and according to

Kolmogorov and [5] we need a mesh size as small as this to capture the turbulence. So it seems that there is no laminar solution any more for these two Reynolds numbers and that, as soon as the mesh size is small enough to represent the phenomena, the method can not converge. We see that the converged solutions on the 256×256 grid exhibit small eddies along the wall in addition to the usual secondary and tertiary vortices in the corners (the main characteristics of these solutions are listed in table 6). Then unstable flow on the grid 512×512 for Re = 10000 is plotted on figure 9.

Re	7500	10000
u_{min} y_{min}	-0.4379 0.0508	-0.4373 0.0430
v_{max} x_{max}	0.4179 0.0625	0.4141 0.0547
v_{min} x_{min}	-0.5640 0.9688	-0.5610 0.9727
Primary vortex - Φ (location)	-0.1113 (0.5156,0.5234)	-0.1053 (0.5156,0.5234)
Second. vortex bottom right Φ (location)	0.832E-2 (0.8828,0.0820)	0.986E-2 (0.8945,0.0820)
Second. vortex bottom left Φ (location)	0.476E-2 (0.0703,0.1289)	0.623E-2 (0.0781,0.1133)
Second. vortex top left Φ (location)	0.314E-2 (0.0664,0.9141)	0.403E-2 (0.0664,0.9141)

Extrema of velocity profiles along centerlines and stream function of the solution with the 256×256 grid

Table 6

Now starting with the solution obtained on the 256×256 grid for Re = 10000, we run the code on this grid for higher Reynolds numbers up to 15000. Then we see small eddies developing along the walls which move during the iterations as shown on figure 10 for Re = 11000 and the apparition of strange structures like a double center vortex for Re = 15000 (figure 11). Finally, we conjecture that beyond Re = R_c with $5000 < R_c < 7500$ there is not a steady laminar solution any more and the transition to turbulence occurs when small eddies develop along the walls.

Let us conclude this section with some results in the rectangular cavity of aspect ratio equal to two. The main results are listed in table 7 for Re = 100 and 400 on a uniform 128×256 grid and for Re = 1000 on a 256×512 grid. Illustrations are plotted on figures 12 and 13. We notice, as indicated in table 1, that the optimal relaxation parameter is very small with respect to the square cavity. We believe that this limit of stability is due to the complexity of the solution in this case and that the transition to turbulence should appear for R_c quite close to 1000. In [13] a Hopf bifurcation is expected in this cavity for $2000 \leqslant R_c \leqslant 10000$.

CONCLUSION.

In this work a new scheme for the discretization of convection terms of Navier-Stokes equations is proposed. Coupled with a simplified FMG-FAS algorithm and a cell by cell relaxation procedure this scheme is very stable and efficient. The steady laminar solution is computed for Re = 1000 on a 128×128 grid in 24 seconds and for Re = 5000 on a 512×512 grid in half an hour of CPU time on a SIEMENS VP 200.

The results are in good agreement with other published results in the square driven cavity for Reynolds number less or equal to 1000. The solutions computed in this work exhibit a very good representation of secondary and tertiary vortices. It appears that the solution is laminar for $Re \leqslant 5000$ and becomes unstable for the steady model with Re = 7500 on the fine 512×512 grid when small eddies develop along the walls in the corners. We hope to confirm these results with further works on the unsteady Navier-Stokes equations.

Finally some results are given in the rectangular driven cavity of

Re	100	400	1000
u_{min} y_{min}	-0.1968 1.4531	-0.3146 1.2578	-0.3851 1.1680
u_{max} y_{max}	0.00195 0.3438	0.01781 0.5703	0.03102 0.5
Primary vortex top Φ (location)	-0.1033 (0.6172,1.7344)	-0.1124 (0.5547,1.5938)	-0.1169 (0.5273,1.5625)
Primary vortex bottom Φ (location)	0.783E-3 (0.5391,0.5859)	0.909E-2 (0.4297,0.8125)	0.0148 (0.3516,0.7891)
Second. vortex bottom left Φ (location)	-0.149E-7 (0.0313,0.0313)	-0.257E-6 (0.0391,0.0469)	-0.108E-4 (0.0977,0.1094)
Second. vortex bottom right(*) Φ (location)	-0.335E-6 (1.0000,0.3750)	-0.146E-6 (0.9688,0.0391)	-0.459E-4 (1.0000,0.3750)

(*) This last vortex is quite unstable during iterations.

Extrema of u velocity profile along centerline and stream function

Table 7

aspect ratio equal to two and a parallel is made with results obtained in the square cavity. It seems that the flow is much more complex in the rectangular cavity and the limit of stability is reached in this case for Re = 1000 although a steady laminar solution is still captured for this Reynolds number.

REFERENCES.

[1] M. FORTIN, R. PEYRET and R. TEMAM, J. Mécanique 10, 357 (1971).

[2] F. THOMASSET, Implementation of finite element methods for Navier-Stokes equations (Springer-Verlag, New York Heidelberg Berlin, 1981).

[3] R. PEYRET and T.D. TAYLOR, Computational methods for fluid flow. (Springer-Verlag, New York Heidelberg Berlin, 1983).

[4] A. FORTIN, M. FORTIN and J.J. GERVAIS, J. Comput. Phys. 70, 295 (1987).

[5] P. CONSTANTIN, C. FOIAS and R. TEMAM, Attractors representing turbulent flows, Memoirs of A.M.S. 53, 314 (1985).

[6] P. LE QUERE and T. ALZIARY DE ROQUEFORT, J. Comput. Phys. 57, 210 (1985).

[7] J. SHEN, Thesis, University of Paris-Sud, Orsay, 1987 (unpublished).

[8] R. SCHREIBER and H.B. KELLER, J. Comput. Phys. 49, 310 (1983).

[9] S.P. VANKA, J. Comput. Phys. 65, 138 (1986).

[10] U. GHIA, K.N. GHIA and C.T. SHIN, J. Comput. Phys. 48, 387 (1982).

[11] H.K. MOFFATT, J. Fluid Mech. 18, 1 (1964).

[12] K. GUSTAFSON and R. LEBEN, Applied. Math. and Comput. 19, 89 (1986).

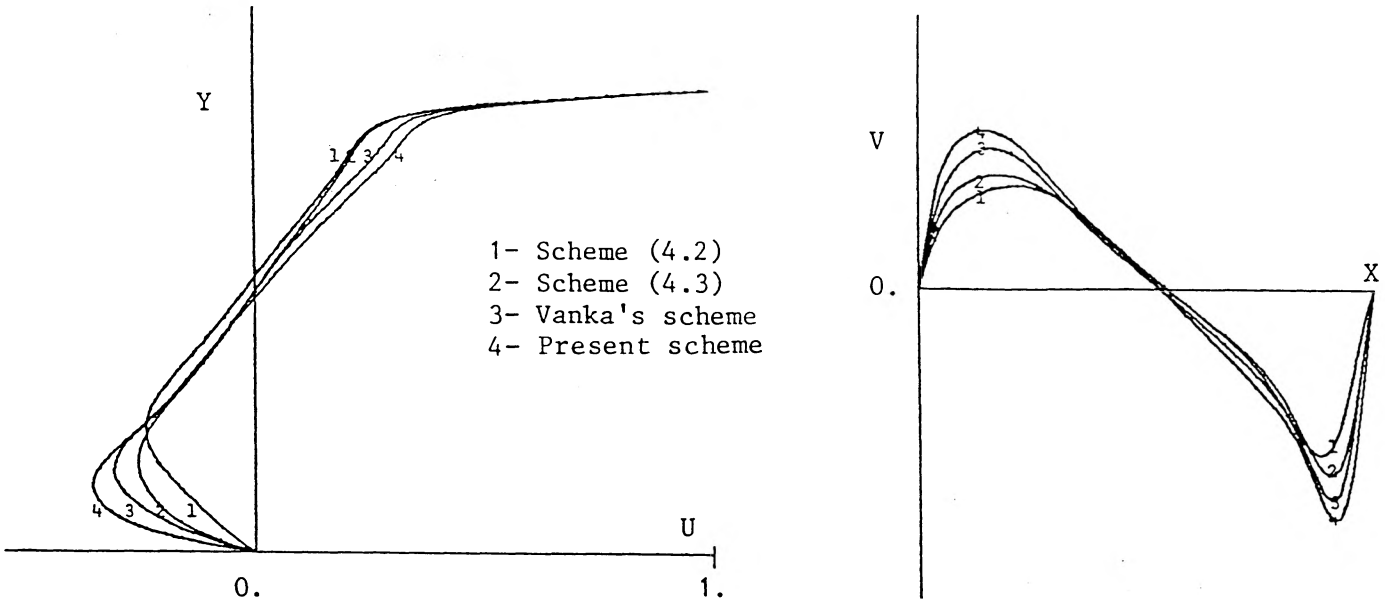
[13] K. GUSTAFSON and K. HALASI, J. Comput. Phys. 70, 271 (1987).

[14] L.B. ZHANG, Thesis, University of Paris-Sud, Orsay, 1987 (unpublished).

[15] A. BRANDT and N. DINAR, in Proceedings of an Advanced Seminar Conducted by the Mathematical Research Center, The University of Wisconsin-Madison, 1978, edited by S.V. Parter (Academic Press, 1979) p.53.

[16] W. HACKBUSCH, Multigrid Methods and Applications (Springer-Verlag, New York Heidelberg Berlin, 1985).

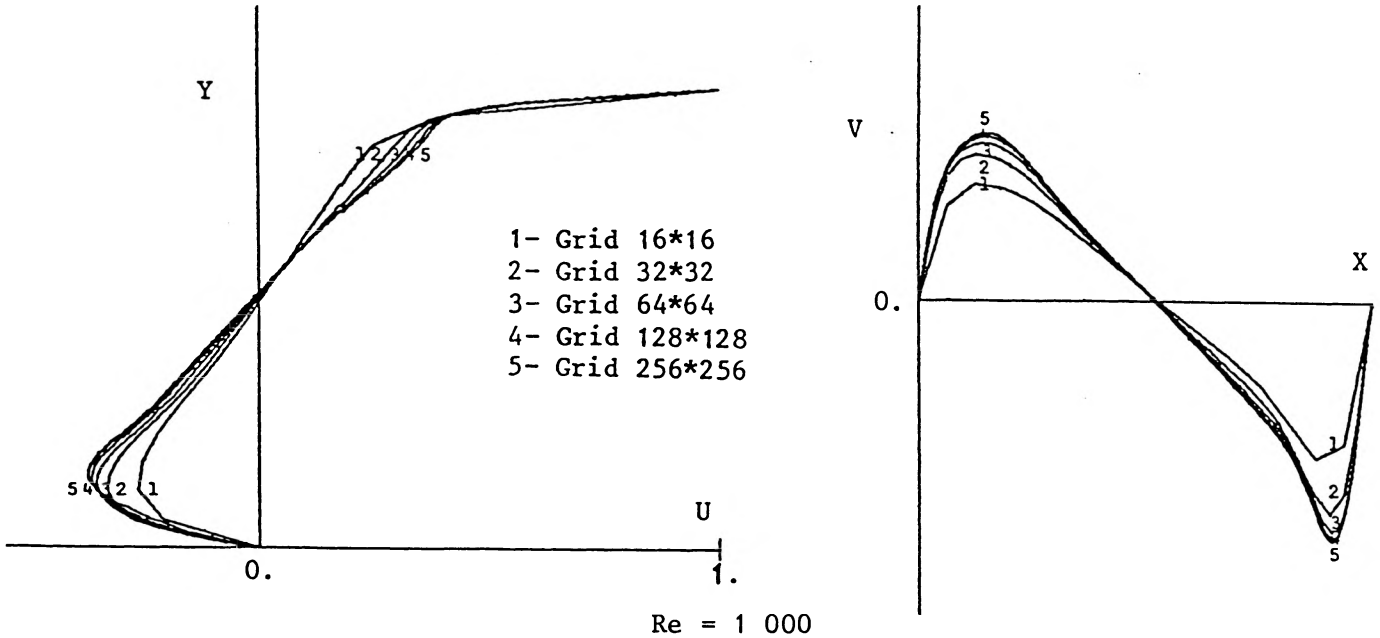
[17] D.B. SPALDING, Int. J. for Num. Meth. in Eng. 4, 551 (1972).



Re = 1 000 Grid 64*64

Comparison of different schemes on velocity profiles

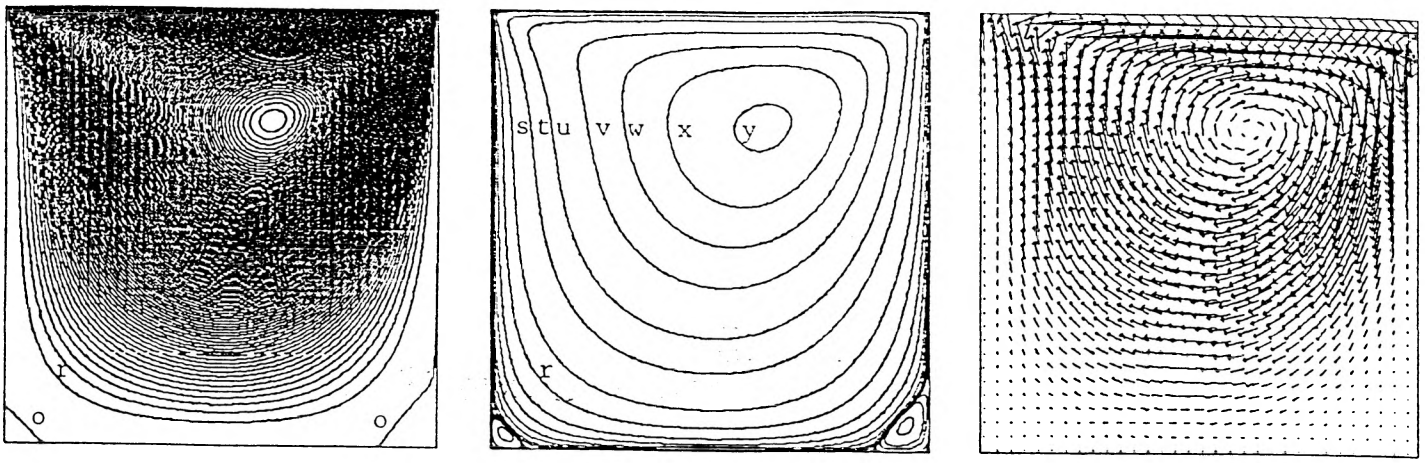
FIGURE 3



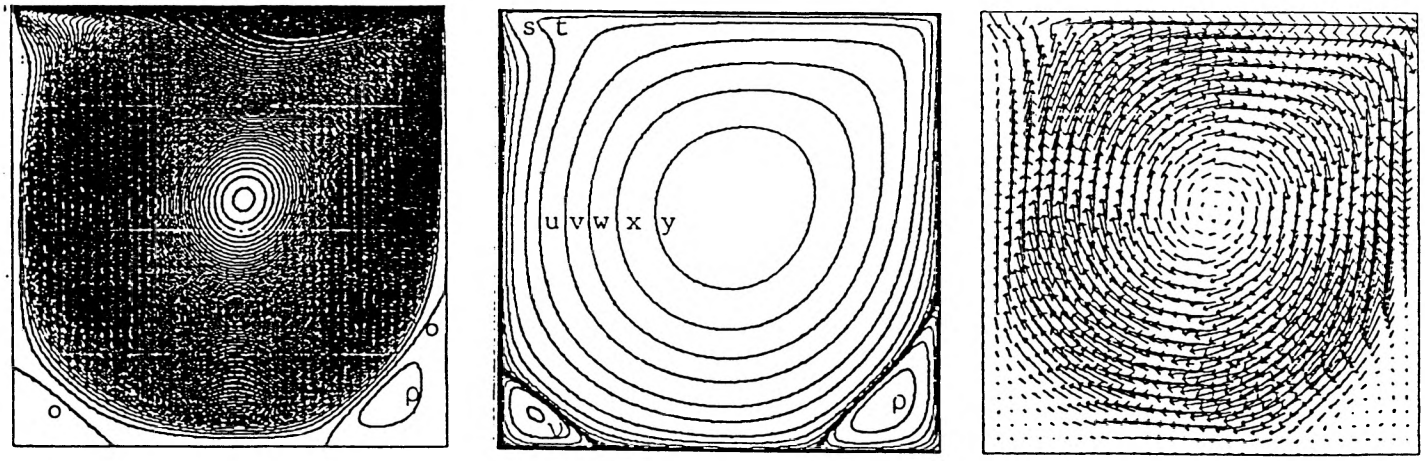
Re = 1 000

Evolution of velocity profiles along centerlines with different grids

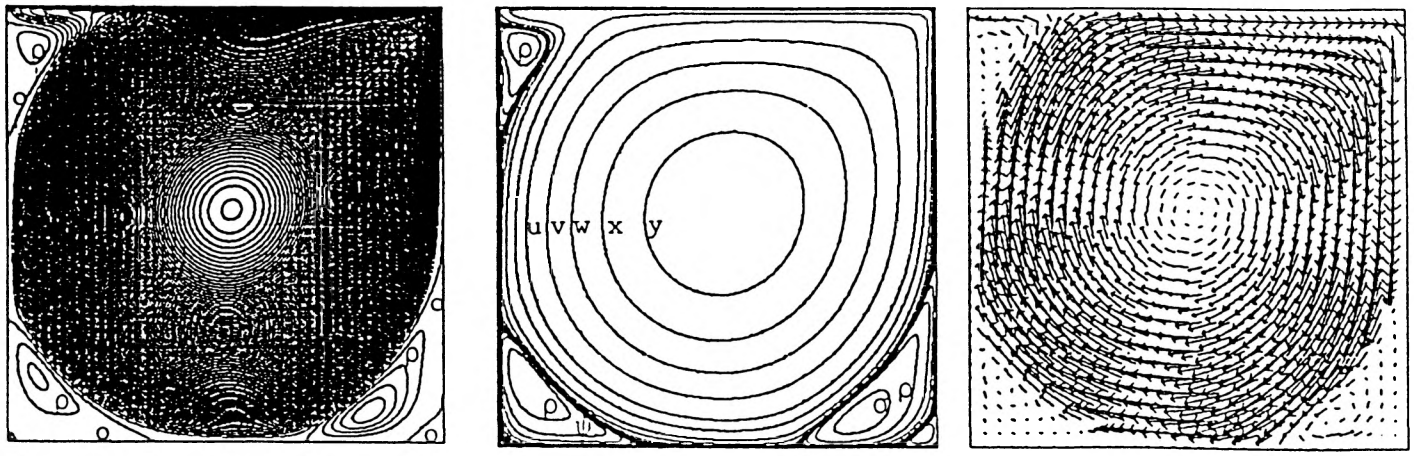
FIGURE 4



Re = 100



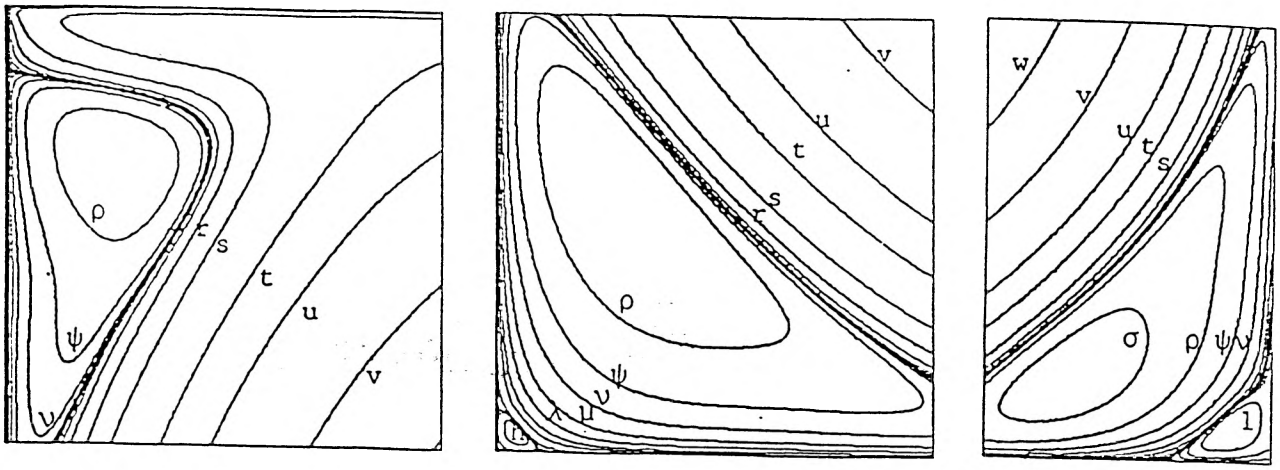
Re = 1000



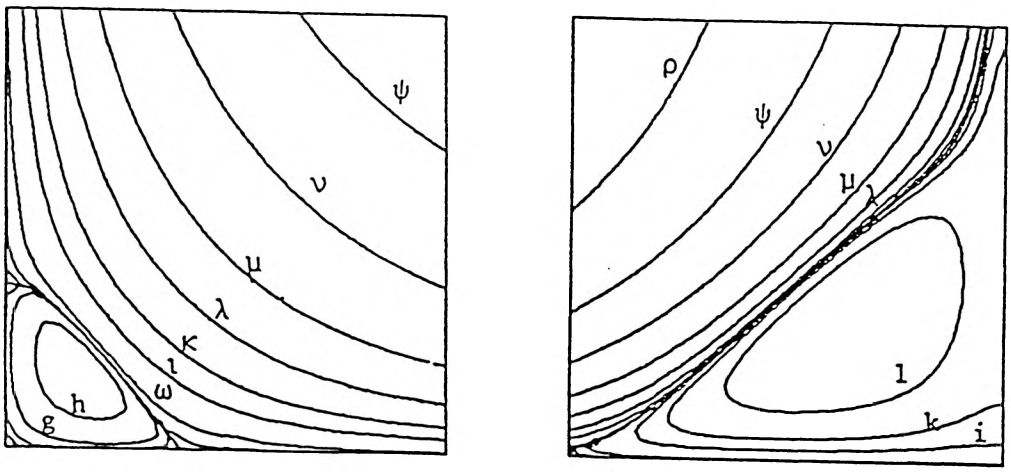
Re = 5000

Equidistant streamlines, tabulated streamlines and velocity fields in the square cavity.

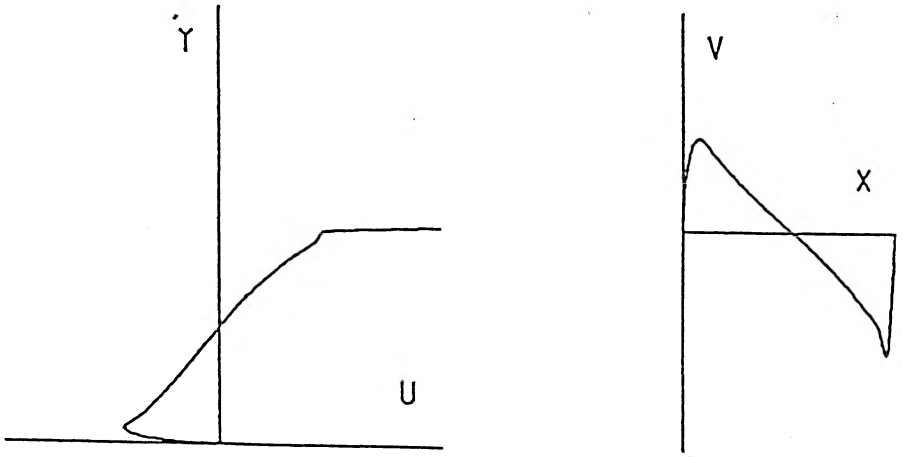
FIGURE 5



Secondary vortices



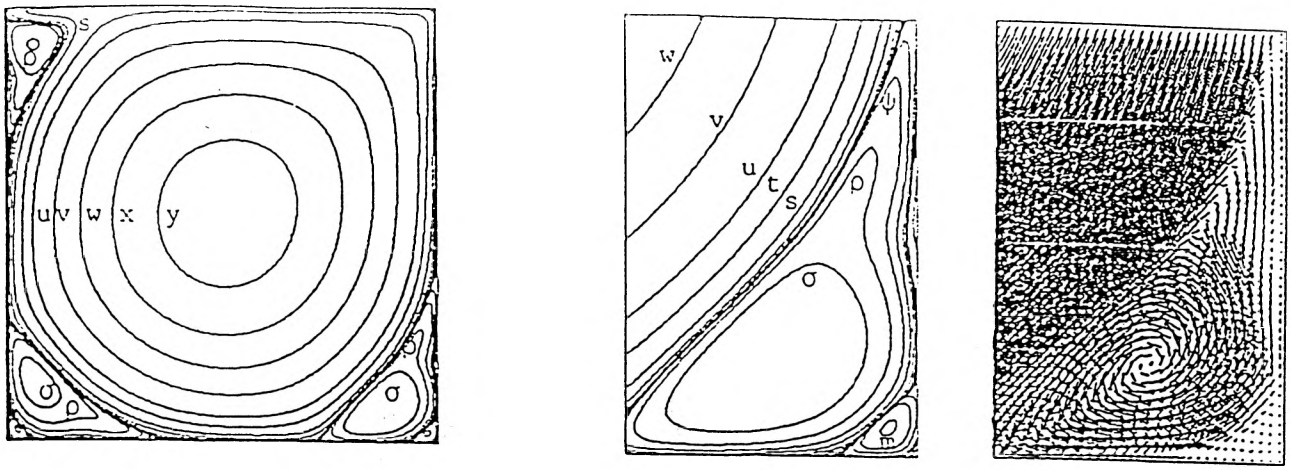
Tertiary vortices



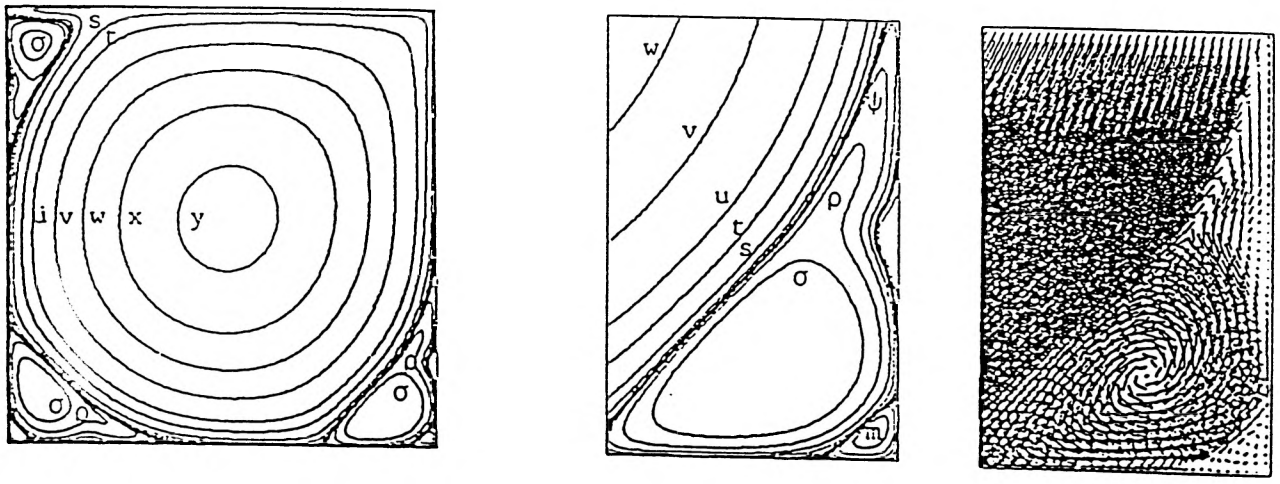
Velocity profiles along centerlines

Details of the solution for Re = 5000

FIGURE 6



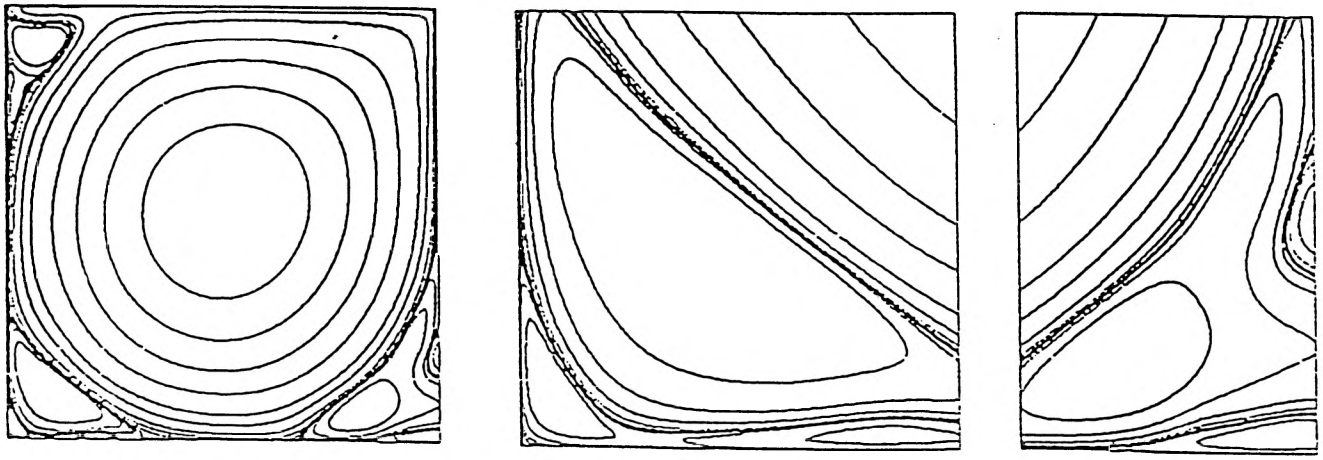
Re = 7500



Re = 10.000

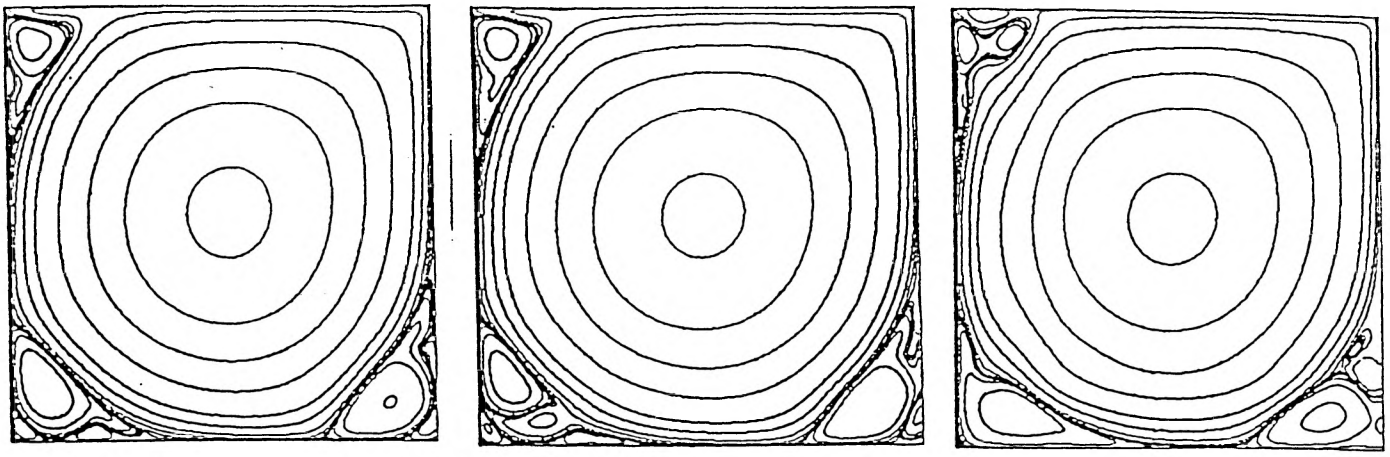
Converged solutions with the 256 x 256 grid

FIGURE 7



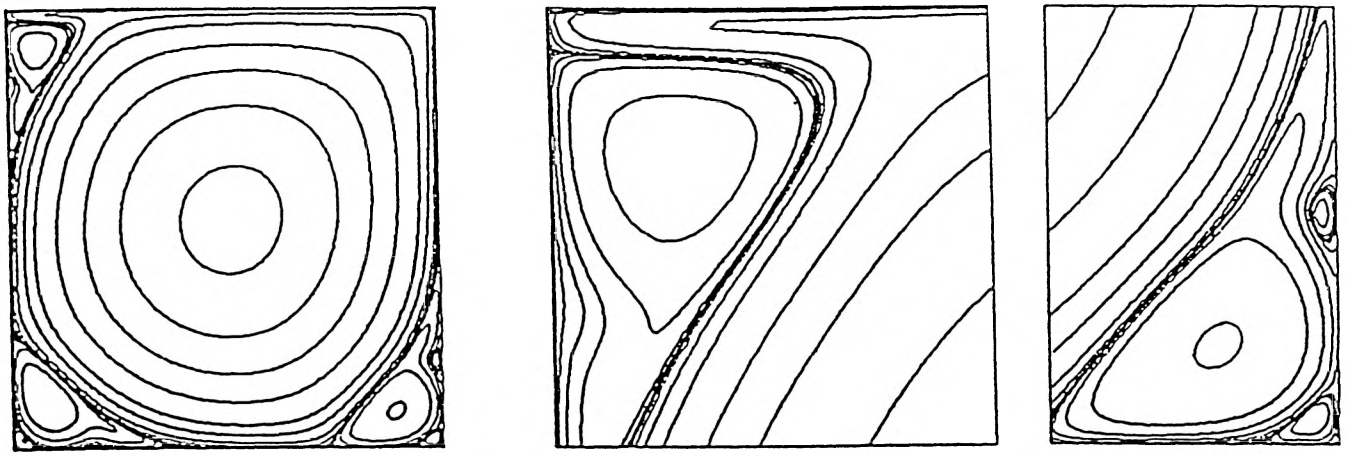
Small perburbations along the walls in the corners for Re = 7500
with the 512 x 512 grid

FIGURE 8

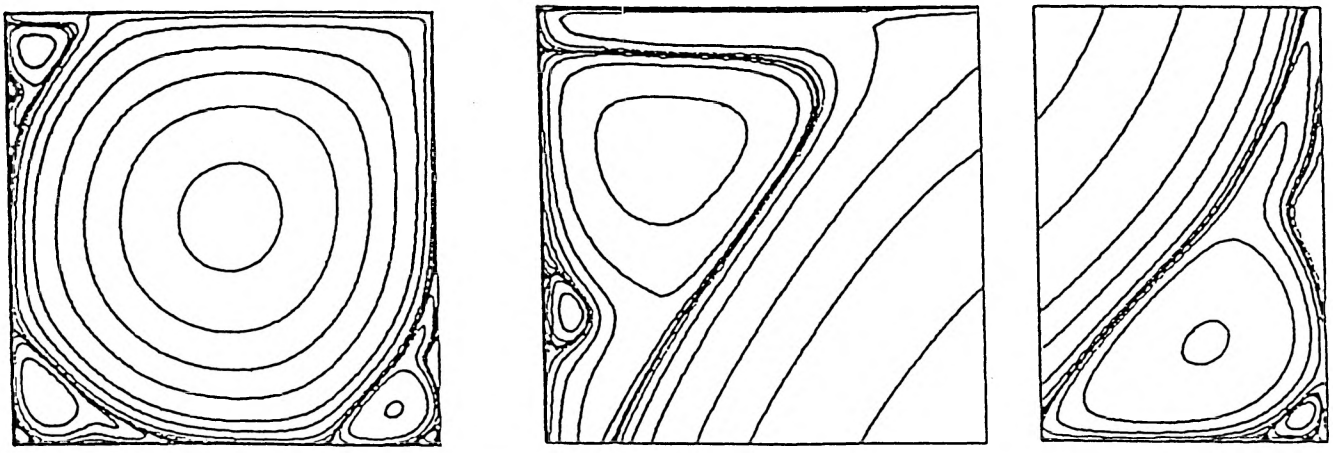


Evolution of the solution during iterations for
Re = 10000 with the 512x512 grid

FIGURE 9



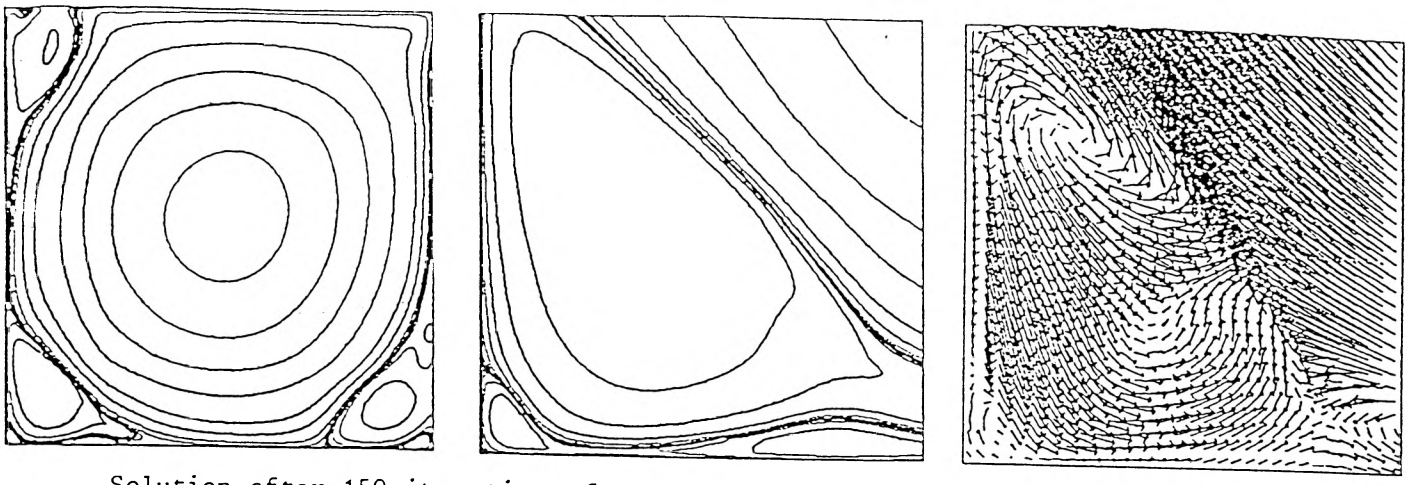
50 iterations



150 iterations

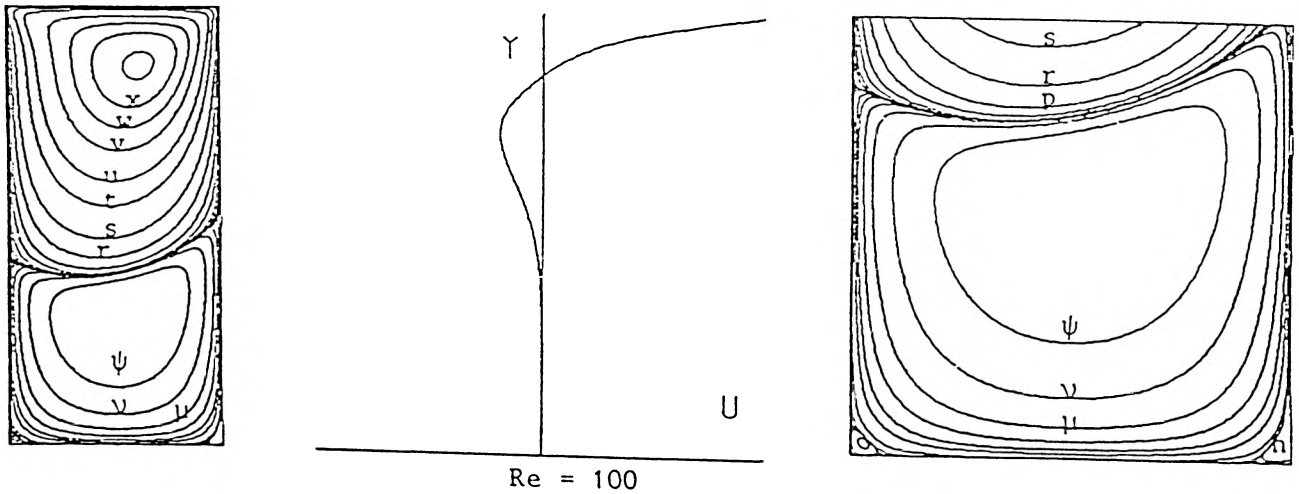
Appearance and disappearance of small eddies for Re = 11000 with
the 256 x 256 grid

FIGURE 10

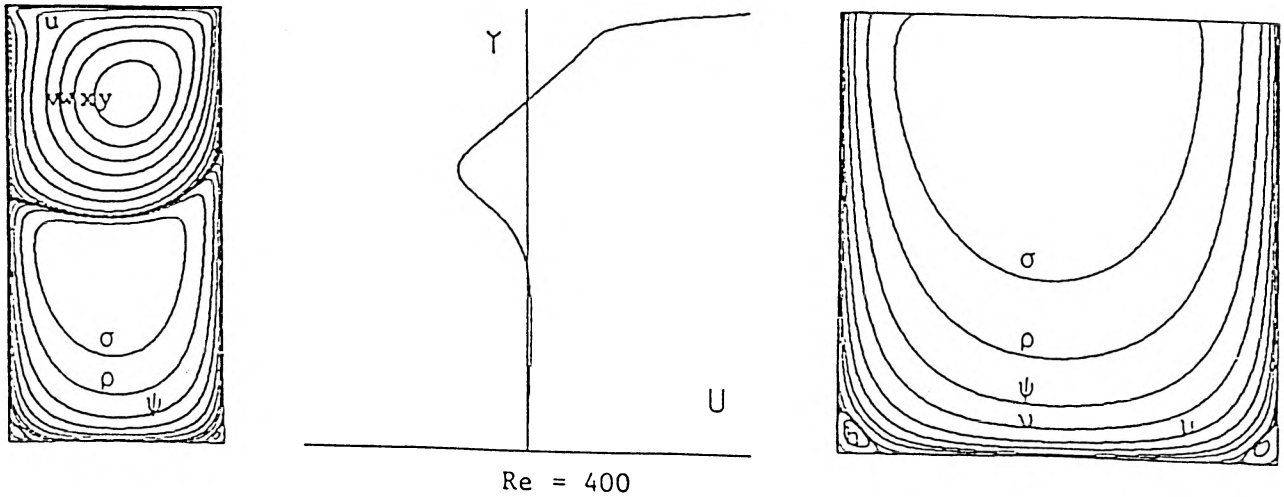


Solution after 150 iterations for Re = 15000 with the 256x256 grid

FIGURE 11



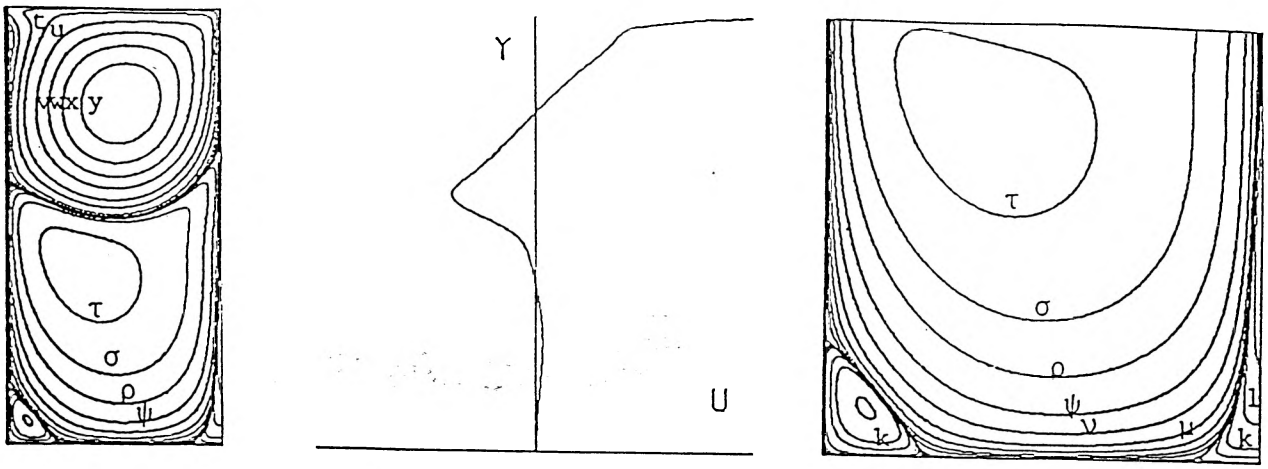
Re = 100

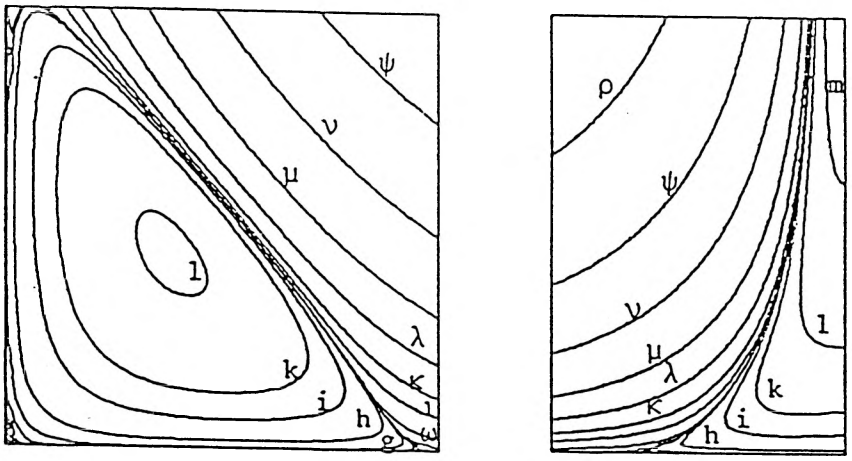


Re = 400

Streamlines and u velocity profile along centerline in the rectangular cavity

FIGURE 12





Streamlines and u velocity profile along centerline in the rectangular cavity

FIGURE 13

ETUDE DU PROBLEME LINEARISE

Soient Ω la cavité (0,1)×(0,1), $\partial\Omega$ son bord et $U = (u_1, u_2)^T$, $V = (v_1, v_2)^T$, il s'agit de résoudre le double système :

$$\begin{aligned}
&\frac{\partial U}{\partial t} - \frac{1}{Re}\Delta U + (U \cdot \nabla)U + \nabla p = 0 \text{ dans } \Omega \\
&\text{div } u = 0 \text{ dans } \Omega \\
&U = 1 \text{ sur } I = (0,1) \times \{1\} \\
&U = 0 \text{ sur } \partial\Omega \backslash I \\
&\frac{\partial V}{\partial t} - \frac{1}{Re}\Delta V + (U \cdot \nabla)V + (V \cdot \nabla)U + \nabla q = 0 \text{ dans } \Omega \\
&\text{div } V = 0 \text{ dans } \Omega \\
&V = 0 \text{ sur } \partial\Omega
\end{aligned} \tag{B.15}$$

où le problème linéarisé en V est associé à des conditions aux limites de Dirichlet homogènes.

En fait on s'intéresse au comportement de la solution pour les grands temps et au calcul du premier exposant de Lyapunov $\mu_1 = \lim_{t \to +\infty} \frac{\|V(t)\|}{t}$ ce qui entraîne de très gros calculs pour les grands nombres de Reynolds car, d'une part, il faut des grilles très fines pour bien représenter la solution et, d'autre part, les temps significatifs sont plus grands. Pour cette étude qui est motivée par le calcul du nombre de Reynolds critique R_c pour lequel il y a perte de stabilité de la solution stationnaire, nous proposons de simplifier beaucoup le système (B.15) en remplaçant U par les solutions stationnaires U_S calculées dans [BJ2]. En effet, quand $Re < R_c$ la solution du premier système converge vers la solution stationnaire stable U_S et donc le comportement de U pour les grands temps est donné par U_S. Ceci permet de remplacer U par U_S dans le deuxième système pour trouver le comportement de V pour les grands temps ; on se borne à résoudre :

$$\begin{aligned}
&\frac{\partial V}{\partial t} - \frac{1}{Re}\Delta V + (U_S \cdot \nabla)V + (V \cdot \nabla)U_S + \nabla q = 0 \text{ dans } \Omega \\
&\text{div } V = 0 \text{ dans } \Omega \\
&V = 0 \text{ sur } \partial\Omega
\end{aligned} \tag{B.16}$$

dont la solution V tend vers zéro quand t tend vers l'infini comme $e^{\mu_1 t} = e^{-Re(\lambda_1)t}$ où μ_1 est le premier exposant de Lyapunov et λ_1 la plus petite valeur propre de l'opérateur. Quand

le nombre de Reynolds est petit cette valeur propre est réelle positive ($\lambda_1 = -\mu_1 > 0$), en particulier pour $Re = 1$ la plus petite valeur propre est proche de celle du problème de Stokes ou du Laplacien en dimension deux. Ensuite elle décroît vers zéro quand le nombre de Reynolds augmente et ce qui nous intéresse est de déterminer à quel moment elle atteint l'axe, c'est-à-dire la valeur du nombre de Reynolds critique R_c.

Nous proposons pour résoudre (B.16) d'utiliser un schéma explicite en temps comme suit :

$$\begin{aligned} &\frac{V^{n+1} - V^n}{\Delta t} - \frac{1}{Re}\Delta V^n + (U_S \cdot \nabla)V^n + (V^n \cdot \nabla)U_S \\ &\qquad + \nabla q^n = 0 \text{ dans } \Omega \\ &\frac{q^{n+1} - q^n}{\Delta t} + \text{ div } V^n = 0 \text{ dans } \Omega \\ &V^{n+1} = 0 \text{ sur } \partial\Omega \end{aligned} \tag{B.17}$$

où l'on introduit la dérivée $\frac{\partial q}{\partial t}$ pour calculer la pression ; la discrétisation en espace est la même que celle présentée dans [BJ1]. L'ensemble donne un schéma très simple qui a pour seule restriction la condition CFL de stabilité. Pour le problème de Stokes cette condition est donnée par $\Delta t \leq \frac{Re(\Delta x)^2}{4}$ et donc elle est moins stricte pour les grands nombres de Reynolds ; cependant l'ajout des termes dépendant de U_S perturbe la stabilité et oblige à prendre des pas de temps plus petits. Dans la pratique il faut prendre Δt de l'ordre de $\frac{1}{1000}$ quand le nombre de Reynolds est supérieur à 1000 pour $\Delta x = \frac{1}{32}$; ce qui est beaucoup plus restrictif que la condition CFL ci-dessus qui donne $\Delta t \simeq \frac{1}{4}$ ou la condition CFL du terme de transport $\Delta t \leq \frac{\Delta x}{|U_S|}$ qui est vraie pour $\Delta t \leq \Delta x = \frac{1}{32}$.

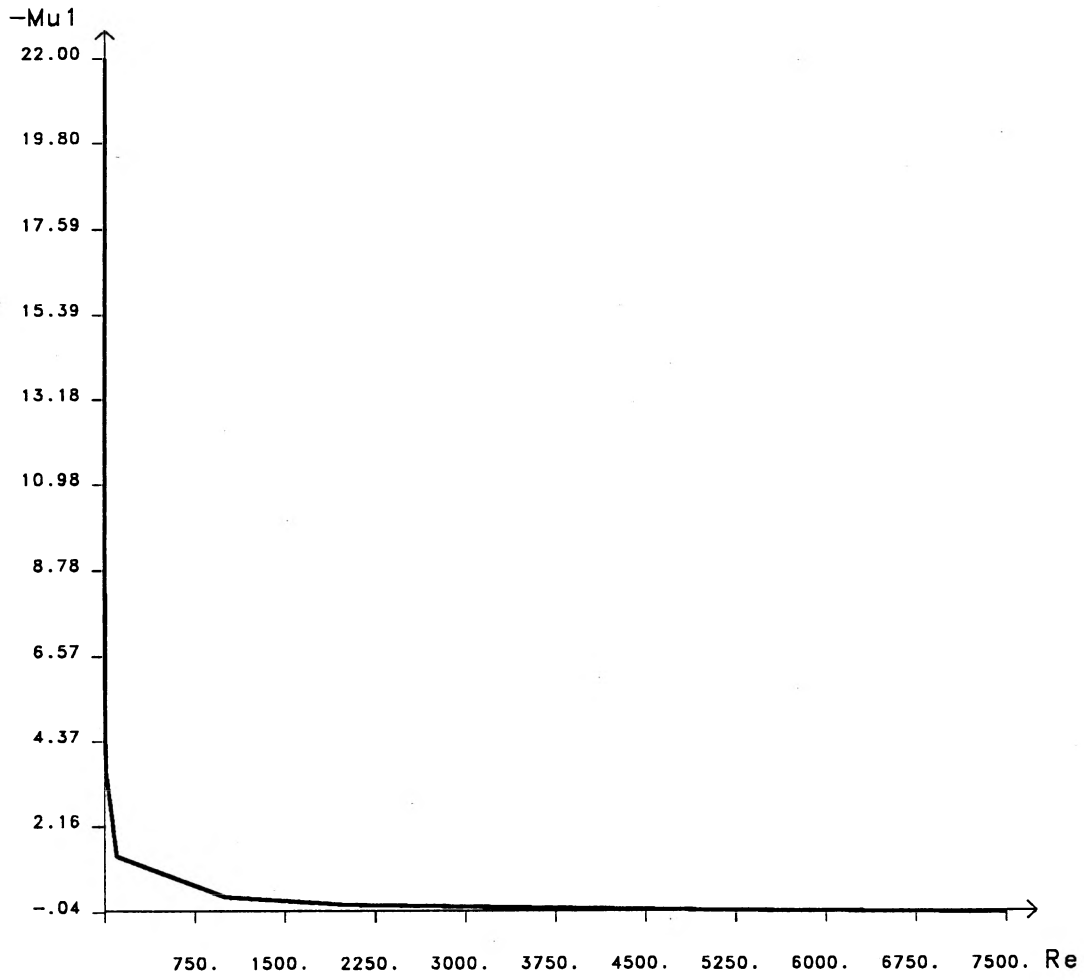
La difficulté numérique essentielle est la représentation de V^{n+1} qui devient très petit quand $t = (n+1)\Delta t$ est grand. Pour pallier cet inconvénient on normalise le problème à chaque itération en temps, ce qui revient à calculer $V^{n+1}/\parallel V^n \parallel = W^{n+1}$ qui est de norme proche de 1 et, pour calculer l'exposant de Lyapunov μ_1, on construit une suite μ_1^n donnée par :

$$\mu_1^{n+1} = \frac{\sum_{i=0}^{n} \text{Log } \parallel W^{i+1} \parallel}{t} \; ; \quad n = 0, 1, \ldots$$

si l'on prend $V^0 = W^0$ de norme 1. La convergence de la suite μ_1^n est lente mais très régulière quand Re n'est pas trop grand (par exemple 100) mais devient beaucoup plus anarchique pour les grands Reynolds quand μ_1 est proche de zéro. Des premiers résultats sur une grille 32×32 sont donnés dans le tableau ci-joint ; ils ont été obtenus en projetant

sur cette grille les solutions U_S calculées dans [BJ2] sur une grille 256×256 où la solution stationnaire reste stable numériquement jusqu'à $Re = 10000$. Cet artifice permet de calculer avec précision μ_1 sur une grille grossière dans des temps raisonnables (de l'ordre d'une minute de CRAY2). La courbe d'évolution de $-\mu_1$ en fonction du Reynolds ressemble à une branche d'hyperbole ayant pour asymptotes les axes. La valeur obtenue pour $Re = 7500$ est légèrement négative bien que $\mu_1 \leq 0$ a priori ; cela s'explique probablement par les difficultés numériques rencontrées pour les grands Reynolds. Ces résultats permettent de donner une fourchette pour R_c entre 5000 et 7500 comme on l'a déjà vu dans [BJ2]. Une analyse plus fine pour des valeurs au-delà de 5000 devrait permettre de mieux localiser R_c.

Re	1	2	5	10	100	1000	2000	5000	7500
$-\mu_1$	22	12	5.7	3.6	1.4	0.36	0.16	0.015	-0.04



Evolution de $-\mu_1$ en fonction de Re

CONCLUSION.

L'étude de schémas avec terme d'antidiffusion appliqués à un problème visqueux a permis de mettre au point un schéma réunissant à la fois deux qualités essentielles : la diffusion minimale et la stabilité. Ce schéma, couplé à un algorithme multigrille et une relaxation maille à maille, donne d'excellentes performances sur le problème de la cavité entraînée en incompressible ; d'une part au niveau des temps de calculs très compétitifs et d'autre part en ce qui concerne la représentation de la solution. La comparaison avec les résultats existants pour $Re = 1000$, montre de légères différences de niveau de la fonction de courant pour les tourbillons secondaires qui semblent mieux capturés. De plus, des résultats à haut nombre de Reynolds montrent la transition vers la turbulence au moment où de nouvelles structures apparaissent le long des parois. Enfin l'étude du problème linéarisé a permis de confirmer la présence d'un nombre de Reynolds critique compris entre 5000 et 7500 comme cela est apparu sur les solutions numériques calculées dans la cavité carrée $(0,1)\times(0,1)$.

PLAN

BIBLIOGRAPHIE

[B1] Ch.H. Bruneau : A non conforming finite element method for solving a least-square formulation of Tricomi's problem, *Comp. Meth. Appl. Mech. Eng. 39,* 1983.

[BCLG1] Ch.H. Bruneau, J.J. Chattot, J. Laminie, J. Guiu-Roux : Finite element least square methods for problem of mixed type, 4th Int. Symp. Fin. El. Meth. Flow, Tokyo 1982, Finite Element Flow Analysis, 1982.

[BCLG2] Ch.H. Bruneau, J.J. Chattot, J. Laminie, J. Guiu-Roux : Finite element least square method for solving full steady Euler equations in a plane nozzle, 8th ICNMFD, Aachen 1982, *Lecture Notes in Physics 170,* 1983.

[BCLT1] Ch.H. Bruneau, J.J. Chattot, J. Laminie, R. Temam : Numerical solutions of the Euler equations with separation by a finite element method, 9th ICNMFD, Saclay 1984, *Lecture Notes in Physics 218,* 1985.

[BCLT2] Ch.H. Bruneau, J.J. Chattot, J. Laminie, R. Temam : Computation of vortex flows past a flat plate at high angle of attack, 10th ICNMFD, Beijing 1986, *Lecture Notes in Physics 264,* 1987.

[BLC1] Ch.H. Bruneau, J. Laminie, J.J. Chattot : Computation of 3D vortex flows past a flat plate at incidence through a variational approach of the full steady Euler equations, *Int. J. Num. Meth. Fluids 9,* 1989.

[BLC2] Ch.H. Bruneau, J. Laminie, J.J. Chattot : Computation of hypersonic vortex flows with an Euler model, 11th ICNMFD, Williamsburg 1988, *Lecture Notes in Physics 323,* 1989.

[B2] Ch.H. Bruneau : Computation of hypersonic flows round a blunt body in $2D$ with Euler model, soumis à *Computers and Fluids.*

[BJ1] Ch.H. Bruneau, C. Jouron : Un nouveau schéma décentré pour le problème de la cavité entraînée, *C. R. Acad. Sci. Paris 301,* série I, 1988.

[BJ2] Ch.H. Bruneau, C. Jouron, An efficient scheme for solving steady incompressible Navier-Stokes equations, à paraître dans *J. Comp. Phys.*