FRÉDÉRIC PASCAL

Méthodes de Galerkin non linéaires en discrétisation par éléments finis et pseudo-spectrale : application à la mécanique des fluides

Thèses d'Orsay, 1992

ORSAY
n° d'ordre :

63800

UNIVERSITÉ de PARIS-SUD

Centre d'ORSAY

THÈSE

présentée
pour obtenir

le grade de Docteur en Sciences

Spécialité : Mathématiques

par

Frédéric PASCAL

Sujet :

Méthodes de Galerkin non linéaires en discrétisation par éléments finis et pseudo-spectrale. Application à la mécanique des fluides.

soutenue le 16 janvier 1992 devant la Commission d'examen

MM. R. TEMAM, Président
C. BASDEVANT, Rapporteur
J.-F. MAITRE, Rapporteur
Mme M. FARGE
M. J. LAMINIE
Mme M. MARION
MM. G. MEURANT
J.-C. SAUT

A la mémoire de Philippe Rolland.

A mes copains rugbymen à Courbevoie.

Je tiens à exprimer ma profonde gratitude à Monsieur R. Temam qui a dirigé mes travaux de recherches. Ses conseils, son expérience et sa confiance m'ont permis de mener à bien ce travail au Laboratoire d'Analyse Numérique d'Orsay et à l'Institut de Mathématiques Appliquées et de Calcul Scientifique à Bloomington.

Je suis infiniment reconnaissant à Monsieur C. Basdevant qui m'a accueilli au Laboratoire de Météorologie Dynamique à l'E.N.S. de la rue d'Ulm en tant que Scientifique du Contingent et qui m'a fort bien conseillé.

Je remercie Madame M. Farge qui m'a consacré un peu de son temps et de son enthousiasme, ainsi que Madame M. Marion et Monsieur J.F. Maître avec qui j'ai eu d'intéressantes discussions.

Ce travail doit beaucoup à la complicité et à la disponibilité de Jacques Laminie. Ma dette est grande et j'espère qu'il sait à quel point je le remercie.

Je tiens particulièrement à remercier Messieurs G. Meurant et J.C. Saut, membres du jury.

Merci à Monsieur C. Jouron pour ses conseils avisés, à Madame Lemeur pour sa disponibilité et sa gentillesse, et à l'ensemble des membres du Laboratoire d'Analyse Numérique.

Merci à tous ceux qui m'ont encouragé et en particulier à mes parents.

Les travaux numériques ont été réalisés :

- au Centre de Calcul vectoriel pour la Recherche, CCVR, à Palaiseau (Sun et Cray II) dont je remercie l'équipe assistance qui réalise un travail remarquable,
- au LANOR à Orsay (Bull, Sun, Stardent, Macintosh),
- au LMD à Paris (Sun, Macintosh),
- au IAMSC à Bloomington (Sun, Alliant).

Frédéric Pascal :

Nonlinear Galerkin methods in finite element and pseudo-spectral discretizations. Application to fluid mechanics.

Abstract :

For computing the solution of partial differential equations in fluid mechanics and physics, some new numerical schemes, the nonlinear Galerkin methods, motivated by concepts of the dynamical systems theory were developped by Marion and Temam.

In the case of a pseudo-spectral discretization, the first part is setting out theoretical motivations and principles of the new schemes with models of the turbulence. It is proved that the evolution of small scales can be frozen if the cut-off number is near dissipation range : some criterions for the selection of small scales and some characteristic times are introduced and analyzed.

The second part concerns the description of the classical mixed, conforming P1-4P1 finite element method for computing the solution of the Navier-Stokes equations. The numerical difficulties due to the incompressibility, the nonlinearity and time evolution are described. A code which permits to obtain some data bases for the driven cavity, has been completely developped.

In the third part, small and large structures induced by the hierarchical basis are shown and studied (particularly in the large gradient region) in the perspective of implementing nonlinear Galerkin methods. Then a preconditionned conjugated gradient method for the Dirichlet problem is developped based on the multigrid structure of the hierarchical basis.

After a study of the numerical stability of the schemes with a linear case, the description of algorithms which consist of determining separately the different structures are proposed in the fourth part of this thesis. Some numerical results are presented for Burgers and Navier-Stokes problems in dimension two.

Key words :

Fluid mechanics ; incompressible Navier-Stokes equations ; generalized Burgers' equations ; nonlinear Galerkin method ; finite elements ; hierarchical basis ; multigrid ; pseudo-spectral discretization ; simulation of the turbulence ; iterated laplacian ; numerical stability.

Table des matières

I METHODE DE GALERKIN NON LINEAIRE EN SPECTRAL ET EN DIMENSION 2 **3**

I.1 Les équations de Navier–Stokes. 3
I.1.1 Formulation du problème. 3
I.1.2 Cadre fonctionnel et résultats théoriques. 7
I.2 Motivations et principaux résultats théoriques. 11
I.2.1 Introduction. 11
I.2.2 V.I.A. : grandes et petites structures. 13
I.2.3 La méthode de Galerkin non linéaire. 17
I.3 Motivations numériques. 26
I.3.1 Simulation numérique des écoulements turbulents dans une cavité 2Π périodique. 26
I.3.2 Analyse des expériences numériques. 32
I.4 Les schémas numériques. 50
I.4.1 Première version. 50
I.4.2 Deuxième version. 52
I.4.3 Conclusion. 56
I.5 Annexe A 60

II RESOLUTION DES EQUATIONS DE NAVIER–STOKES PAR LA METHODE DES ELEMENTS FINIS 4P1-P1. (GALERKIN CLASSIQUE) **83**

II.1 Présentation des éléments finis. 84
II.2 Un problème modèle : la cavité entrainée. 86
II.3 Résolution du problème de Stokes. 90
II.3.1 Introduction des éléments 4P1-P1. 90
II.3.2 Méthode d'Uzawa et préconditionnement. 94
II.3.3 Résultats numériques. 100
II.4 Résolution du problème de Navier–Stokes. 105
II.4.1 Discrétisation en espace et en temps. 105
II.4.2 Méthode des moindres carrés. 109
II.4.3 Programmation et résultats numériques. 113

III LA BASE HIERARCHIQUE. **137**
III.1 Introduction de la base hiérarchique. 138
III.1.1 Définition. 138
III.1.2 Propriétés de la base. 143
III.1.3 Démonstration des propriétés. 143
III.2 Equations de Navier-Stokes et base hiérarchique. 147
III.2.1 Décomposition sur la base hierarchique. 147
III.2.2 Décomposition de la solution stationnaire. 150
III.2.3 Evolution des petites et grandes structures. 157
III.3 Etude préliminaire : Base hiérarchique et problème de Poisson. . . 166
III.3.1 Le problème. 166
III.3.2 Formulation variationnelle et discrétisation. 166
III.3.3 Systèmes linéaires associés. 167
III.3.4 Description des solveurs. 169
III.3.5 Résultats numériques. 174

IV BASE HIERARCHIQUE ET PROBLEMES D'EVOLUTION : IMPLEMENTATION DE LA METHODE DE GALERKIN NON LINEAIRE 2D **191**
IV.1 Equation de la chaleur. 192
IV.1.1 Le problème et méthode d'intégration classique. 192
IV.1.2 Schémas induits de la base hiérarchique. 193
IV.1.3 Résultats numériques. 197
IV.2 Un problème non linéaire : équations de Burgers 2D. 202
IV.2.1 Tests avec la solution exacte connue. 202
IV.2.2 Tests avec la solution exacte inconnue. 208
IV.3 Problème de Navier-Stokes. 219
IV.3.1 Description des algorithmes. 219
IV.3.2 Les résultats numériques. 222

INTRODUCTION.

M. Marion et R. Temam ont développé en 1988 de nouvelles méthodes pour résoudre les problèmes aux dérivées partielles issus de la mécanique des fluides. Ces schémas semblent adaptés à la résolution des problèmes évolutifs sur de grands intervalles de temps et à la simulation de phénomènes tels que la turbulence que les ordinateurs actuels abordent de façon insatisfaisante et ce malgré des progrès techniques fulgurants. Ces méthodes qui s'appuient sur des concepts de la théorie des systèmes dynamiques, consistent à chercher une solution approchée du problème appartenant à une variété inertielle approximative ; objet introduit par Foias-Manley-Temam approchant l'attracteur universel. Ces variétés sont généralement non linéaires d'où le nom de méthodes de Galerkin non linéaires.

Pour mettre en place ces méthodes, l'idée est d'introduire une décomposition de la solution approchée en petites et grandes structures. En discrétisation pseudo-spectrale cette décomposition est induite du développement en séries de Fourier de la solution ; les grands modes sont associés aux petites échelles ou structures et les petits modes aux grandes échelles. En discrétisation par éléments finis, l'utilisation de la base hiérarchique et de plusieurs niveaux de maillage (multigrille) fait apparaître en des points distincts de la triangulation des grandes et petites structures. Ces différents objets interagissent non linéairement et il a été prouvé que les équations de ces interactions définissent des variétés inertielles approximatives. L'objectif des méthodes de Galerkin non linéaires est alors d'intégrer les grandes structures que l'on corrige "séparément" en tenant compte des petites structures.

Le but de mon travail est d'implémenter ces méthodes dans le cadre d'une discrétisation par éléments finis pour des problèmes du type Burgers généralisé ou Navier-Stokes et d'étudier leur comportement avec des modèles de turbulence tels que le laplacien itéré employé par C. Basdevant et R. Sadourny dans le cas de discrétisation pseudo-spectrale.

Dans un premier temps, après une introduction des équations de Navier-Stokes, je rappelle les motivations et principaux résultats théoriques des nouveaux schémas. Des motivations numériques dans le cas de simulation de la turbulence libre et forcée sont ensuite explicitées. Il est montré que l'évolution des petites échelles peut être figée à condition que la troncature soit proche de la zone de dissipation et que ces échelles soient relaxées sur une variété inertielle approxima-

tive. Le schéma est par la suite décrit et analysé : des critères de détermination dynamique des petites échelles et des temps caractéristiques pendant lesquels elles sont figées sont introduits et discutés.

La deuxième partie décrit la méthode de Galerkin classique dans le cas d'une discrétisation par éléments finis, mixtes, conformes 4P1-P1 pour le problème de Navier-Stokes incompressible, évolutif en formulation vitesse-pression. La discrétisation en temps (une méthode de splitting) fait apparaître un problème de Stokes généralisé et un problème non linéaire à chaque pas de temps. Le premier problème est résolu par la méthode d'Uzawa et le deuxième par la méthode des moindres carrés. La réalisation du code numérique a permis d'obtenir des résultats (comparés avec des publications récentes) pour la cavité entrainée avec les nombres de Reynolds 1000 et 5000 pour lesquels la solution évolue vers une solution stationnaire.

Les objectifs ce cette implémentation sont d'une part de permettre l'étude des petites et grandes structures introduites par la base hiérarchique, ensuite de servir de base de travail pour la mise en place des méthodes de Galerkin non linéaires et enfin de fournir des résultats qualitatifs et quantitatifs de référence pour la cavité entrainée par exemple.

Dans la troisième partie, la base hiérarchique dont quelques propriétés essentielles sont rappelées est introduite et appliquée aux résultats obtenus précédemment. A l'inverse de la base nodale, cette base induit une décomposition du champ de vitesse en grandes et petites structures. On se propose d'étudier la forme de ces structures dans les zones de forts gradients, leur évolution au cours du temps, l'influence du pas de discrétisation en espace et du nombre de Reynolds et les effets de cette décomposition sur les différents termes des équations (en particulier les termes de diffusion et de transport). Des schémas s'appuyant sur la structure multigrille de la base hiérarchique sont ensuite mis en place pour le problème de Dirichlet : ces schémas permettent des gains de temps non négligeables et se présentent comme d'excellents préconditionneurs de schémas itératifs tels que les gradients conjugués.

Dans la dernière partie, après une étude de la stabilité numérique des schémas sur un problème linéaire, la description des algorithmes qui diffèrent par le traitement des termes d'évolution, et les premiers résultats numériques similaires à ceux obtenus par une méthode classique sont proposés pour les problèmes de Burgers et Navier-Stokes en dimension 2.

Chapitre I

METHODE DE GALERKIN NON LINEAIRE EN SPECTRAL ET EN DIMENSION 2

I.1 Les équations de Navier–Stokes.

I.1.1 Formulation du problème.

Formulation vitesse-pression

Les équations de Navier–Stokes proviennent de la mécanique des fluides : un fluide newtonien vérifie l'équation de conservation de la masse ou équation de continuité, l'équation de conservation de la quantité de mouvement ou conservation du moment et la loi de comportement des fluides newtoniens.

On considère Ω (un ouvert de $\mathbb{R}^2$) le domaine occupé par le fluide et les 3 fonctions ρ, u, p pour décrire l'état du fluide :

$\rho(x,t)$ est la densité du fluide,

$u(x,t) = (u_1(x,t), u_2(x,t))$ est la vitesse de la particule au point x,

$p(x,t)$ représente la pression du fluide en x à l'instant t,

avec $x = (x_1, x_2) \in \Omega$ et $t \in \mathbb{R}^+$.

Remarque I.1 *Un écoulement bidimensionnel n'est pas un écoulement infiniment mince, mais un écoulement invariant par translation dans une direction d'espace et tel que les vitesses soient perpendiculaires à la direction d'invariance.*

On se place dans le cas incompressible, c'est-à-dire que l'on suppose que la densité ρ varie peu : par exemple la densité de l'eau ρ est égale à 1 g/cm^3 et pour l'air à faible vitesse ρ est de l'ordre de 1.2 $10^{-3} g/cm^3$. Les équations de Navier–Stokes régissant l'écoulement d'un fluide newtonien, visqueux, incompressible dans Ω soumis aux densités de forces extérieures $f = (f_1(x,t), f_2(x,t))$ sont :

$$\begin{aligned} \rho\frac{\partial u}{\partial t} - \nu\Delta u + \rho(u.\nabla)u + \nabla p = f \qquad & \text{dans } Q = \Omega \times \mathbb{R}^+ \\ \nabla.u = 0 \qquad & \text{dans } Q \end{aligned} \tag{I.1}$$

ν étant le coefficient de viscosité. La première équation traduit la conservation du mouvement et la loi de comportement, la deuxième la conservation de masse pour les fluides incompressibles.

Sous une forme indicée, (I.1) s'écrit :

$$\begin{aligned} \rho\frac{\partial u_1}{\partial t} - \nu(\frac{\partial^2 u_1}{\partial x_1^2} + \frac{\partial^2 u_1}{\partial x_2^2}) + \rho(u_1\frac{\partial}{\partial x_1} + u_2\frac{\partial}{\partial x_2})u_1 + \frac{\partial p}{\partial x_1} = f_1 \\ \rho\frac{\partial u_2}{\partial t} - \nu(\frac{\partial^2 u_2}{\partial x_1^2} + \frac{\partial^2 u_2}{\partial x_2^2}) + \rho(u_1\frac{\partial}{\partial x_1} + u_2\frac{\partial}{\partial x_2})u_2 + \frac{\partial p}{\partial x_2} = f_2 \\ \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} = 0. \end{aligned} \tag{I.2}$$

$(u.\nabla)u$ est le terme de transport ou convection, ∇p représente les forces de pressions et $-\nu\Delta u$ est le terme de diffusion due aux contraintes de viscosité et par conséquent à la déformation du fluide.

Dans la suite pour simplifier les notations p, f et ν représenteront la pression, les forces et la viscosité réduites ie p/ρ, f/ρ et ν/ρ.

Si l'on note V une vitesse caractéristique de l'écoulement, L une longueur caractéristique du domaine et T un temps caractéristique (typiquement $T = \frac{L}{V}$) alors le système (I.1) se met sous la forme adimensionnée suivante (N.S.) :

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{1}{Re}\Delta u + (u.\nabla)u + \nabla p = f \qquad & \text{dans } Q \\ \nabla.u = 0 \qquad & \text{dans } Q \end{aligned} \tag{I.3}$$

où

$$\frac{1}{Re} = \frac{\nu}{LV}$$

et x, t, u, p, f sont les fonctions adimensionnées :

$$u = \frac{u}{V}, \quad x = \frac{x}{L}, \quad t = \frac{t}{T}, \quad p = \frac{pL^2}{V^2}, \quad f = \frac{fL}{V^2}.$$

Re est le nombre de Reynolds qui dépend non seulement de la viscosité (et par conséquent de la densité) du fluide, mais aussi des caractéristiques de l'écoulement considéré.

D'autres problèmes comme le problème de Stokes et Stokes généralisé sont à envisager. En effet, dans le cas de très faibles nombres de Reynolds ($Re \ll 1$), les équations de Stokes stationnaires (S.) :

$$\begin{aligned} -\frac{1}{Re}\Delta u + \nabla p &= f \qquad \text{dans } Q \\ \nabla . u &= 0 \qquad \text{dans } Q \end{aligned} \tag{I.4}$$

donnent une excellente approximation de (N.S.).

De nombreux schémas en temps pour les équations de Navier–Stokes (par exemple la méthode des pas fractionnaires, les méthodes de projection, les méthodes de directions alternées...) font apparaître les équations de Stokes généralisées (S.G.) :

$$\begin{aligned} \alpha u - \frac{1}{Re}\Delta u + \nabla p &= f \qquad \text{dans } Q \\ \nabla . u &= 0 \qquad \text{dans } Q \end{aligned} \tag{I.5}$$

avec $\alpha = \frac{1}{\Delta t} > 0$.

Il reste à fermer les systèmes ; pour cela les équations de Stokes (S.), (S.G.) et Navier–Stokes (N.S.) sont complétées par :

- des conditions initiales (uniquement pour les problèmes instationnaires) du type :
$$u(x,0) = u_0(x) \quad \forall x \in \Omega,$$
la vitesse initiale u_0 étant donnée avec $div u_0 = 0$.

- des conditions aux limites
  - de Dirichlet
  $$u(x,t) = g(x,t) \quad \forall x \in \partial\Omega,$$
  g vérifiant la condition de flux nul imposée par l'incompressibilité :
  $$\int_{\partial\Omega} g.n \, d\sigma = 0$$
  où l'on note n la normale extérieure à Ω.
  - périodiques : par exemple si $\Omega = [0, L_1] \times [0, L_2]$
  $$u(x + L_i e_i, t) = u(x,t) \quad \forall x \in \Omega, \quad \forall i = 1,2$$
  avec $e_1 = (1,0)$ et $e_2 = (0,1)$. On supposera dans ce cas que l'écoulement est de moyenne nulle ie
  $$\int_\Omega f(x)\, dx = 0 \text{ et } \int_\Omega u(x,t)\, dx = 0.$$

Formulation fonction de courant-vorticité

Dans certains cas, il est préférable d'utiliser des variables différentes de la pression et de la vitesse pour simuler l'écoulement. Par exemple l'incompressibilité du fluide permet de définir la fonction de courant $\psi(x_1, x_2)$ reliée à la vitesse par :

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} +\frac{\partial \psi}{\partial x_2} \\ -\frac{\partial \psi}{\partial x_1} \end{pmatrix}.$$

On note ω le rotationnel de la vitesse u appelé tourbillon ou "vorticité" proportionnel à la vitesse locale de rotation du fluide et défini par :

$$\omega = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}.$$

En prenant le rotationnel de l'équation de Navier–Stokes, le système (N.S.) est équivalent au système pour la vorticité suivant :

$$\begin{array}{rl} \frac{\partial \omega}{\partial t} + J(\omega, \psi) = \nu \Delta \omega + \mathrm{rot} f & \text{dans } Q \\ \omega = -\Delta \psi & \text{dans } Q \\ \omega(x, 0) = \omega_0(x) & \text{dans } \Omega \end{array} \tag{I.6}$$

avec des conditions aux limites. $J(\omega, \psi)$ est le jacobien défini par :

$$\begin{aligned} J(\omega, \psi) &= (u.\nabla)\omega = u_1 \frac{\partial \omega}{\partial x_1} + u_2 \frac{\partial \omega}{\partial x_2} \\ &= \frac{\partial (u_1 \omega)}{\partial x_1} + \frac{\partial (u_2 \omega)}{\partial x_2} = \frac{\partial \psi}{\partial x_2}\frac{\partial \omega}{\partial x_1} - \frac{\partial \psi}{\partial x_1}\frac{\partial \omega}{\partial x_2} \end{aligned}$$

I.1.2 Cadre fonctionnel et résultats théoriques.

Cadre fonctionnel

On suppose que l'ouvert Ω est un ouvert borné de $\mathbb{R}^2$ dont la frontière $\partial\Omega$ est localement lipschitzienne. On suppose que $L^2(\Omega)$ est muni de son produit scalaire usuel et de sa norme usuelle :

$$(u,v) = \int_\Omega u(x)v(x)\,dx \quad \forall u,v \in L^2(\Omega)$$

$$|u| = (u,u)^{1/2} \quad \forall u \in L^2(\Omega)$$

et que l'espace de Sobolev $H^1(\Omega)$ est muni du produit scalaire et de la norme :

$$(u,v)_1 = (u,v) + (\nabla u, \nabla v) \quad \forall u,v \in H^1(\Omega)$$

$$|u|_1 = (u,u)_1^{1/2} \quad \forall u \in H^1(\Omega).$$

On notera $((u,v)) = (\nabla u, \nabla v), \quad \forall u,v \in H^1(\Omega)$.

Pour la suite, on supposera que les conditions au bord sont $u = g$ sur $\partial\Omega$, la fonction g vérifiant la condition de flux nul.

On pose alors $H^1_g(\Omega)$ le sous espace de $H^1(\Omega)$ tel que :

$$H^1_g(\Omega) = \{v \in H^1(\Omega),\ v = g \text{ sur } \partial\Omega\}$$

On note V le sous espace de $H^1(\Omega)\times H^1(\Omega)$ espace de Hilbert associé au problème de Stokes et muni du produit scalaire et de la norme de $H^1(\Omega) \times H^1(\Omega)$:

$$V = \{u \in H^1(\Omega) \times H^1(\Omega), \mathrm{div} u = 0,\ u = g \text{ sur } \partial\Omega\}$$

et H le sous espace de $L^2(\Omega) \times L^2(\Omega)$ défini par :

$$H = \{u \in L^2(\Omega) \times L^2(\Omega), \mathrm{div} u = 0,\ u = g \text{ sur } \partial\Omega\}$$

On note V' l'espace dual de V ; on rappelle que $V \subset H \subset V'$.

Remarque I.2 *Dans le cas où les conditions au bord sont périodiques, il faut remplacer $u = g$ par u périodique, $H^1_g(\Omega)$ par $H^1_p(\Omega)$ et $L^2(\Omega)$ par $L^2_p(\Omega)$ ie par les ensembles des fonctions de $H^1(\Omega)$ et $L^2(\Omega)$ périodiques.*

Formulation variationnelle de (S.G.)

Les chapitres I et III de Temam [24] établissent les formulations variationnelles enoncées ci-dessous.

Le problème de Stokes généralisé (S.G.) est équivalent aux 2 formulations suivantes :

$$f \in L^2(\Omega) \quad \text{donnée}$$

1ere formulation

$$\begin{aligned}&\text{Trouver } u \in V \text{ tel que}\\ &\alpha(u,v)+\tfrac{1}{Re}((u,v)) = (f,v) \quad \forall v \in \{v \in (H_0^1(\Omega))^2,\ \text{div } v = 0\}\end{aligned} \tag{I.7}$$

2eme formulation

$$\begin{aligned}&\text{Trouver } (u,p) \in (H_g^1(\Omega))^2 \times L^2(\Omega) \text{ tel que}\\ &\alpha(u,v)+\tfrac{1}{Re}((u,v))+(\nabla p,v) = (f,v) \quad \forall v \in (H_0^1(\Omega))^2\\ &(q,\text{div}u) = 0 \quad \forall q \in L^2(\Omega)\end{aligned} \tag{I.8}$$

La première formulation a l'avantage de réduire le problème à la recherche de la vitesse u seulement. Lorsque l'existence de u est prouvée, l'existence de la pression p au sens des distributions est une conséquence immédiate de la théorie des distributions.

Formulation variationnelle de (N.S.)

Notons la forme continue, trilinéaire b sur $H^1(\Omega)$ définie par :

$$b(u,v,w) = \sum_{i,j=1}^{2} \int_\Omega u_i (\frac{\partial v_j}{\partial x_i}) w_j \, dx.$$

Le problème de Navier–Stokes (N.S.) est alors équivalent à :

Etant données $f \in L^2(0,T;V')$ et $u_0 \in (H_g^1(\Omega))^2$ avec $\text{div}u_0 = 0$

1ere formulation

$$\begin{aligned}&\text{Trouver } u \in L^2(0,T;V) \text{ satisfaisant}\\ &\tfrac{d}{dt}(u,v)+\tfrac{1}{Re}((u,v))+b(u,u,v) = (f,v) \quad \forall v \in \{v \in (H_0^1(\Omega))^2,\ \text{div } v = 0\}\\ &u(0) = u_0\end{aligned} \tag{I.9}$$

2eme formulation

$$\begin{aligned}&\text{Trouver } (u,p) \in L^2(0,T;(H_g^1(\Omega))^2) \times L^2(0,T,L^2(\Omega)) \text{ tel que}\\ &\tfrac{d}{dt}(u,v)+\tfrac{1}{Re}((u,v))+b(u,u,v)+(\nabla p,v) = (f,v) \quad \forall v \in (H_0^1(\Omega))^2\\ &(q,\text{div}u) = 0 \quad \forall q \in L^2(\Omega)\\ &u(0) = u_0\end{aligned} \tag{I.10}$$

Théorèmes d'existence et d'unicité

1. **Problème de Stokes**
   Le théorème de Lax-Milgram permet d'établir (cf Temam [24] ch I.2) :

   Théorème I.1 *Pour tout $f \in L^2(\Omega)$,*

   (a) le problème (I.7) a une solution unique u dans V,

   (b) le problème (I.8) a une solution unique $(u,p) \in (H_g^1(\Omega))^2 \times L^2(\Omega)/\mathbb{R}$ ie p est déterminé à une constante additive près.

2. **Problème de Navier–Stokes**
   En dimension 2 en se référant au chapitre III de Temam [24], on démontre le résultat suivant :

   Théorème I.2 *Pour tout $f \in L^2(0,T;V')$ et $u_0 \in V$, il existe une unique solution u au problème (I.9). De plus $u \in L^\infty(0,T;H)$ et u est presque partout égale à une fonction continue de $[0,T]$ dans H, $u(0) = u_0$ ayant alors un sens.*

Formulation abstraite du problème de Navier–Stokes

On supposera ici que la condition au bord est $u = 0$ sur $\partial\Omega$. Soit A l'opérateur linéaire, positif, auto-adjoint correspondant à l'opérateur de Stokes défini par :

$$Au \in V'$$

$$(Au, v) = ((u,v)) \quad \forall u, v \in V.$$

On note $B(.,.)$ l'opérateur compact, bilinéaire, continue de $V \times V$ dans V' défini par :

$$(B(u,v),w) = b(u,v,w) \quad \forall u,v,w \in V$$

et vérifiant :

$$b(u,v,v) = 0 \quad \forall u,v, \in V \quad \textit{condition d'orthogonalité} \tag{I.11}$$

$$|B(u,v)| \leq c_1 \begin{cases} |u|^{1/2}||u||^{1/2}||v||^{1/2}|Av| \\ |u|^{1/2}|Au|^{1/2}||v|| \end{cases} \quad \forall u,v, \in V \tag{I.12}$$

$$|(B(u,v),w)| \leq c_2|u|^{1/2}||u||^{1/2}||v|||w|^{1/2}||w||^{1/2} \quad \forall u,v,w, \in V \tag{I.13}$$

$$|B(u,v)| \leq c_4 \begin{cases} ||u||||v||(1+\log\frac{|Au|^2}{\lambda_1||u||^2})^{1/2} \\ |u||Av|(1+\log\frac{|A^{3/2}v|^2}{\lambda_1|Av|^2})^{1/2} \end{cases} \quad \forall u,v, \in V \tag{I.14}$$

On peut alors mettre le problème de Navier–Stokes (N.S.) sous la forme abstraite suivante :

$$\begin{array}{c} \text{Trouver } u \in V \subset H \text{ satisfaisant} \\ \frac{\partial u}{\partial t} + \frac{1}{Re} Au + B(u,u) = f \end{array} \tag{I.15}$$

avec la condition initiale $u(0) = u_0$.

Remarque I.3 *En modifiant A, H, V et f d'autres conditions au bord telles que des conditions périodiques, des conditions non homogènes, des conditions de friction ou de non glissement,... peuvent être envisagées.*

Comme A^{-1} est compact et auto-adjoint, A possède une famille de vecteurs propres w_j orthogonaux dans H ie :

$$Aw_j = \lambda_j w_j \quad \forall j$$

où λ_j sont les valeurs propres avec

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \text{ et } \lambda_j \longrightarrow +\infty \text{ quand } j \longrightarrow +\infty$$

La solution $u \in H$ développée dans cette base s'écrit alors

$$u(x,t) = \sum_{j=1}^{\infty} \hat{u}_j(t) w_j(x).$$

On appelle nombre d'onde le nombre k_m défini par : $\lambda_m = k_m^2$, la longueur d'onde étant proportionnelle à l'inverse de k_m.

Dans le cas de conditions périodiques dans une cavité $[0, L_1] \times [0, L_2]$, les vecteurs propres sont des combinaisons des fonctions sinus et cosinus ; $u \in H$ est alors développable en série de Fourier et s'écrit :

$$u(x,t) = \sum_{j \in \mathbb{R}^2} \hat{u}_j(t) e^{ij.x} = \sum_{j=(j_1,j_2) \in \mathbb{Z}^2} \hat{u}_j(t) e^{i\{\frac{2\pi j_1}{L_1} x_1 + \frac{2\pi j_2}{L_2} x_2\}},$$

le nombre d'onde étant dans ce cas $2\pi \sqrt{\frac{j_1^2}{L_1^2} + \frac{j_2^2}{L_2^2}}$.

Remarque I.4 *En dimension 2, la solution u du problème de Navier–Stokes reste bornée : il existe 2 constantes M_0 et M_1 dépendant de ν, $|f|$ et de λ_1 et il existe t_0 dépendant de u_0, ν, $|f|$ et de λ_1 tels que :*

$$|u(t)| \leq M_0, \quad ||u(t)|| \leq M_1, \quad \forall t > t_0.$$

I.2 Motivations et principaux résultats théoriques.

I.2.1 Introduction.

On s'intéresse au comportement de la solution du problème de Navier–Stokes pour de grands temps. Dans [21], Jie Shen montre (numériquement) dans le cas de la cavité entrainée (B) (cf. chapitre II) que l'écoulement converge vers un état stationnaire pour un nombre de Reynolds (Re) inférieur à 10000.0, que l'écoulement est périodique en temps pour un nombre de Reynolds compris entre Re_1 (avec $10000.0 < Re_1 \leq 10500.0$) et Re_2 (avec $15000.0 < Re_2 \leq 15500.0$) et enfin qu'au-delà de Re_2 l'écoulement perd sa périodicité.

Un attracteur est un sous-espace compact de l'espace de Hilbert H (espace défini dans la section I.1.2) vers lequel convergent toutes les solutions (ou orbites) qui se trouvent initialement dans son voisinage. Pour des nombres de Reynolds petits, la solution converge vers une solution stationnaire après un temps de transition et ceci quelque soit la condition initiale : l'attracteur se réduit à un point fixe. Lorsque le nombre de Reynolds augmente, une "bifurcation" (de Hopf par exemple) se produit : l'écoulement devient périodique, l'attracteur est alors un cycle limite. Pour des nombres de Reynolds plus grands, l'écoulement peut devenir quasi-périodique. Puis lorsque Re est suffisamment grand (un certain nombre de bifurcations s'étant produit), l'écoulement du fluide devient, après un temps de transition, "chaotique". Il est dépendant du temps même si les conditions au bord et les forces extérieures en sont indépendantes (comportement turbulent) avec une extrême sensibilité aux conditions initiales entrainant une imprédicibilité de l'évolution temporelle. Les attracteurs ne sont plus des points fixes, ni des cycles limites mais des attracteurs étranges (notion introduite par Smale, Takens et Ruelle [19]) d'une extrême complexité (de type fractal par exemple bien que cela ne soit pas démontré).

On appelle attracteur global ou universel $\mathcal{A}$, la réunion de ces attracteurs et de quelques orbites exceptionnelles (pour un même nombre de Reynolds). $\mathcal{A}$ est un sous-espace de H compact et connexe défini par :

$$\begin{array}{rcll} S(t)\mathcal{A} & = & \mathcal{A} & \forall t \geq 0 \\ \lim_{t\to\infty} \mathrm{dist}\,(S(t)B, \mathcal{A}) & = & 0 & \forall B \text{ ensemble borné de } H \end{array} \tag{I.16}$$

où $S(t)$ est l'opérateur de semi-groupe associé à l'équation de Navier–Stokes. L'existence de $\mathcal{A}$ en dimension 2 a été prouvé par Foias et Temam [7] qui demontrent également la finitude de sa dimension. Constantin et Foias [5] en donnent une majoration :

$$cRe^{2/3} \leq \dim\mathcal{A} \leq c(\frac{k_d}{k_0})^2 \leq cRe$$

où k_0^{-1} est une longueur caractéristique de l'écoulement et ${k_d}^{-1}$ l'échelle de dissipation de Kolmogorov. Rappelons que le nombre d'onde k_d est tel que d'une part $k_d^{-1} \ll k^{-1}$ et que d'autre part les structures de taille inférieure à k_d^{-1} sont très rapidement dissipées (de façon exponentielle). Cette majoration est améliorée dans le cas d'un écoulement périodique : $dim\mathcal{A} \leq Re^{2/3}(1 + \log Re)^{1/3}$ (cf. Constantin *et al* [4]).

Une conséquence essentielle de la finitude de la dimension de $\mathcal{A}$ est que l'écoulement turbulent établi après un temps de transition peut être décrit par un nombre fini de paramètres ie de degrés de liberté. Ce nombre nécessairement grand dépend du nombre de Reynolds, les questions qui se posent étant alors :

- quel est ce nombre ?
- quels sont ces paramètres ?

La théorie phénoménologique de Kolmogorov et de Kraichnan (cf. Lesieur et ses références [15]) prévoit un nombre nécessaire de degrés de liberté pour décrire un écoulement turbulent de l'ordre de

$$(\frac{k_d}{k_0})^2 \simeq Re.$$

Rappelons que $Re = \frac{UL}{\nu}$ où U et L sont respectivement une vitesse et une longueur caractéristique. Or on peut estimer U en dimension 2 :

$$U = \frac{|f|L}{\nu} \Longrightarrow Re = \frac{|f|L^2}{\nu^2} \tag{I.17}$$

avec

$$|f| = (\int_\Omega |f(x)|^2 \, dx)^{1/2}$$

En effet en multipliant scalairement la formulation abstraite non adimensionnée de l'équation de Navier-Stokes $\frac{\partial u}{\partial t} + \nu Au + B(u,u) = f$ par u dans H, la propriété d'orthogonalité de B donne :

$$\begin{aligned} \frac{1}{2}\frac{d|u|^2}{dt} + \nu||u||^2 = (f,u) &\leq |f||u| \\ &\leq \lambda_1^{-1/2}|f|||u|| \\ &\leq \frac{\nu}{2}||u||^2 + \frac{1}{2\nu\lambda_1}|f|^2 \end{aligned}$$

d'où

$$\frac{d|u|^2}{dt} + \nu||u||^2 \leq \frac{|f|^2}{\nu\lambda_1},$$

or $||u||^2 \geq \lambda_1|u|^2$

$$\frac{d|u|^2}{dt} + \nu\lambda_1|u|^2 \leq \frac{|f|^2}{\nu\lambda_1}$$

d'où en intégrant

$$|u(t)|^2 \leq |u(0)|^2 \exp(-\nu\lambda_1 t) + \frac{|f|^2}{\nu^2\lambda_1^2}(1 - \exp(-\nu\lambda_1 t)).$$

On en conclut que :

$$|u(t)| \leq \frac{|f|}{\nu\lambda_1} \quad \forall t \geq t_\star(\nu, f, \lambda_1, u_0).$$

Or λ_1, la plus petite valeur propre de l'opérateur A est de l'ordre de l'inverse du carré de la longueur caractéristique L de l'écoulement et $\frac{(\sup_{t>t_\star}|u(t)|)}{L}$ fournit une estimation de la vitesse caractéristique.

Devant la complexité de $\mathcal{A}$, Foias *et al* [9] ont introduit la notion de variété inertielle $\mathcal{M}$; variété plus régulière de dimension finie contenant $\mathcal{A}$ et définie par :

- $\mathcal{M}$ est une variété lipschitzienne
- $S(t)\mathcal{M} \subset \mathcal{M} \quad \forall t > 0$
- $\mathcal{M}$ attire toutes les solutions de (NS) de façon exponentielle.

H étant décomposé en 2 sous-espaces ($H = P_m H \oplus Q_m H$), $\mathcal{M}$ est cherchée sous la forme du graphe d'une fonction

$$\Phi : P_m H \longrightarrow Q_m H.$$

L'existence de $\mathcal{M}$ pour les équations de Navier–Stokes n'a pas été démontré.

I.2.2 V.I.A. : grandes et petites structures.

Pour approcher l'attracteur $\mathcal{A}$, une notion plus intéressante (dans le sens où l'existence est prouvée) a été introduite par Foias, Manley et Temam [10] : il s'agit de la variété inertielle approximative (VIA). C'est une variété lipschitzienne de dimension finie dont un proche voisinage attire exponentiellement et en un temps fini toutes les orbites. $\mathcal{A}$ est donc inclus dans ce voisinage.

Soit $P_n H$ la projection de H sur l'espace engendré par les n premiers vecteurs propres de A ie sur $\mathrm{Vect}(w_1, \ldots, w_n)$ et $Q_n H = H - P_n H$ la projection de H sur $\mathrm{Vect}(w_{n+1}, \ldots)$. On a donc :

$$H = P_n H \oplus Q_n H$$

et $u \in H$ peut se mettre sous la forme suivante :

$$u = y_n + z_n$$

avec

$$y_n(x,t) = \sum_{1\leq j\leq n} \hat{u}_j(t)w_j(x) \quad \in P_nH$$

qui correspond aux petits nombres d'onde du développement de u ie aux grandes échelles de l'écoulement et avec

$$z_n(x,t) = \sum_{j>n} \hat{u}_j(t)w_j(x) \quad \in Q_nH$$

qui correspond aux petites échelles de l'écoulement.

Les équations de Navier–Stokes écrites sous la forme abstraite :

$$\frac{\partial u}{\partial t} + \nu Au + B(u,u) = f$$

sont projetées sur P_nH et Q_nH. Le système d'équations couplées est alors le suivant :

$$\frac{\partial y_n}{\partial t} + \nu Ay_n + P_nB(u,u) = P_nf \tag{I.18}$$

$$\frac{\partial z_n}{\partial t} + \nu Az_n + Q_nB(u,u) = Q_nf \tag{I.19}$$

Le terme $B(u,u)$ est composé de 2 parties :

- $B(y_n,y_n)$, terme non linéaire associé à y_n
- $B(y_n,z_n) + B(z_n,y_n) + B(z_n,z_n)$, terme de couplage et d'interactions entre petites et grandes structures.

En tenant compte des inégalités (I.12),(I.13) et (I.14), de la propriété d'orthogonalité de B et en prenant le produit scalaire de (I.19) avec z_n dans H, Foias *et al* [10] ont montré que z_n était petit au sens suivant :

$$\begin{array}{rcccll} |z_n(t)| & \leq & \kappa_0 & \mathrm{L}^{1/2} & \delta & \forall t \geq t_\star \\ ||z_n(t)|| & \leq & \kappa_1 & \mathrm{L}^{1/2} & \delta^{1/2} & \forall t \geq t_\star \end{array} \tag{I.20}$$

à condition que

$$\lambda_{n+1} \geq \frac{(2c_2 \sup_{s\in[0,\infty[} ||u(s)||)^2}{\nu^2} \tag{I.21}$$

avec

$$\delta = \frac{\lambda_1}{\lambda_n} \quad \text{et} \quad \mathrm{L} = 1 + \log\frac{\lambda_n}{\lambda_1}.$$

κ_0, κ_1 et $t_\star$ sont des constantes ne dépendant que de $\lambda_1, \nu, f, \Omega$ et des bornes supérieures de $|u|$, $||u||$ et $|Au|$.

L'analyticité en temps de la solution de (NS) démontrée dans Foias et Temam [8] et la formule de Cauchy pour estimer dans le plan complexe permettent

également d'obtenir des majorations de la dérivée en temps de z_n (cf Foias *et al* [10] et Promislow [18]) :

$$\begin{array}{rcccll} |z_n'(t)| & \leq & \kappa_2 & \mathrm{L}^{1/2} & \delta & \forall t \geq t_\star \\ |Az_n(t)|| & \leq & \kappa_3 & \mathrm{L}^{1/2} & & \forall t \geq t_\star \end{array} \tag{I.22}$$

sous la même condition (I.21).

Il est également possible de comparer les différents termes non linéaires par rapport au terme de dissipation :

$$\begin{array}{rcccll} |Q_nB(z_n(t),z_n(t))|/\nu|Az_n(t)| & \leq & \kappa_4 & \mathrm{L}^{3/4} & \delta & \forall t \geq t_\star \quad \text{si (I.21)} \\ |Q_nB(z_n(t),y_n(t))|/\nu|Az_n(t)| & \leq & \kappa_5 & \mathrm{L}^{1/2} & \frac{\delta^{1/2}}{\lambda_1^{1/2}} & \forall t \\ |Q_nB(y_n(t),z_n(t))|/\nu|Az_n(t)| & \leq & \kappa_6 & \mathrm{L}^{1/4} & \delta^{1/4} & \forall t \geq t_\star \quad \text{si (I.21)} \end{array}$$

En effet

$$|B(z_n,z_n)| \leq c_1|z_n|^{1/2}||z_n||^{1/2}||z_n||^{1/2}|Az_n|$$

d'où

$$\frac{|Q_nB(z_n,z_n)|}{\nu|Az_n|} \leq \frac{|B(z_n,z_n)|}{\nu|Az_n|} \leq \frac{c_1}{\nu}|z_n|^{1/2}||z_n||,$$

(I.20) permettant de conclure pour la première inégalité. Pour la deuxième inégalité, la relation

$$|B(z_n,y_n)| \leq c_4||z_n||.||y_n||.(1+\log\frac{|Ay_n|^2}{\lambda_1||y_n||^2})^{1/2}$$

et $|Ay_n|^2 \leq \lambda_n||y_n||^2$, $||z_n||^2 \leq \lambda_{n+1}^{-1}|Az_n|^2$ entrainent que

$$\frac{|Q_nB(z_n,y_n)|}{\nu|Az_n|} \leq \frac{|B(z_n,y_n)|}{\nu|Az_n|} \leq \frac{\tilde{c}_4\mathrm{L}^{1/2}}{\nu\lambda_{n+1}} \quad \forall t$$

d'où le résultat. Enfin

$$|B(y_n,z_n)| \leq c_1|y_n|^{1/2}||y_n||^{1/2}||z_n||^{1/2}|Az_n|$$

(I.20), et le fait que $|y_n|$ et $||y_n||$ soient bornés impliquent si (I.21)

$$\frac{|Q_nB(y_n,z_n)|}{\nu|Az_n|} \leq \frac{|B(y_n,z_n)|}{\nu|Az_n|} \leq \frac{\mathrm{L}^{1/4}\delta^{1/4}\kappa_1^{1/2}}{\nu}c_1 \sup|u|^{1/2} \sup||u||^{1/2}.$$

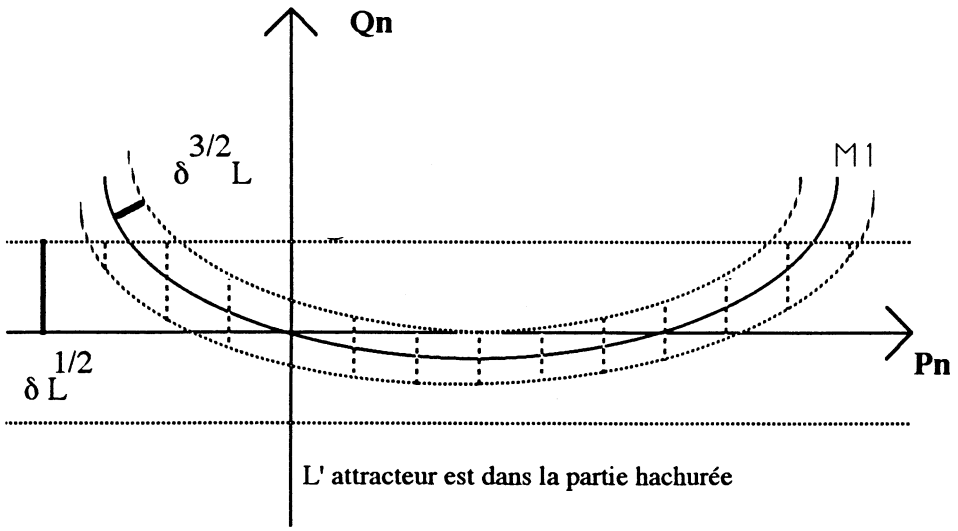


Figure I.1: *Attracteur et variétés inertielles approchées*

On constate donc que pour $t \geq t_\star$ et λ_{n+1} suffisamment grand, $|z_n(t)|$ est petit devant $|y_n(t)|$. Une conséquence de ce résultat est que le terme $|B(z_n(t), z_n(t))|$ est petit devant $|B(z_n(t), y_n(t))|$ et $|B(y_n(t), z_n(t))|$ qui sont eux même petits devant $|B(y_n(t), y_n(t))|$. De plus les termes $|B(z_n(t), z_n(t))|$, $|B(z_n(t), y_n(t))|$ et $|B(y_n(t), z_n(t))|$ sont petits devant $\nu|Az_n(t)|$. Enfin le temps de relaxation de y_n de l'ordre de $(\nu\lambda_1)^{-1}$ est bien plus grand que le temps de relaxation de z_n de l'ordre de $(\nu\lambda_{n+1})^{-1}$. On en déduit que si $t \geq t_\star$ et si λ_{n+1} est suffisamment grand, alors l'équation (I.19) peut être simplifiée et mise sous la forme :

$$\nu A z_n + Q_n B(y_n, y_n) = Q_n f \tag{I.23}$$

L'equation (I.23) traduit une loi d'interaction non linéaire entre grandes et petites structures. C'est l'équation d'une variété inertielle approximative $\mathcal{M}_1$ (cf.Foias *et al* [10]) telle que pour tout t supérieur à $t_\star$, la distance dans H de toute orbite à la VIA $\mathcal{M}_1$ soit inférieure à $\kappa \mathrm{L}\delta^{3/2}$, ($\kappa$ étant une constante dépendant uniquement de ν, λ_1, $|f|$ et de la condition initiale u_0). Si l'on remarque que $P_n H$ est une VIA noté $\mathcal{M}_0$ telle que :

$$dist(u(t), \mathcal{M}_0) \leq \kappa \mathrm{L}^{1/2}\delta \quad \forall t \geq t_\star \quad \text{dans } H$$

alors $\mathcal{M}_1$ apparait être une meilleure variété approximative de l'attracteur que $P_n H$ (cf figure I.1).

Remarque I.5 *D'autres variétés inertielles approximatives peuvent être construites :*

- *soit de façon non constructive (cf. Foias* et al *[10])*
- *soit en perturbant l'équation de Navier–Stokes (cf. Temam [26])*
- *soit à partir de développements asymptotiques (cf. Temam [25])*
- *soit de façon implicite (cf. Titi [28]).*

Remarque I.6 *Dans le cas de conditions au bord périodiques, la condition (I.21) est équivalente à (cf. [10]) :*

$$\frac{\lambda_{n+1}}{\lambda_1} \geq 8c_2^2 Re^2.$$

c_2 étant la constante de l'inégalité (I.2.3). Cette inégalité implique que le nombre de modes nécessaires est supérieur à celui prévu par la théorie phénoménologique : cela est dû au haut degré de généralité de l'étude.

I.2.3 La méthode de Galerkin non linéaire.

Le principe.

Le but est de construire un système de dimension finie plus approprié pour décrire à long terme les écoulements turbulents. Rappelons que la méthode de Galerkin classique consiste à chercher une approximation u_m de u sous la forme :

$$u_m(x,t) = \sum_{j=1}^{m} \hat{u}_j(t) w_j(x)$$

solution de l'équation de Navier-Stokes (NS) projetée sur $P_m H$:

$$\frac{\partial u_m}{\partial t} + \nu A u_m + P_m B(u_m, u_m) = P_m f. \tag{I.24}$$

En tenant compte des résultats exposés dans la section I.2.2), Marion et Teman [16] ont proposé la méthode de Galerkin non linéaire qui consiste à chercher une solution approchée de u sous la forme :

$$u_m(x,t) = u_n(x,t) + u_{m-n}(x,t)$$

avec

$$u_n(x,t) = \sum_{j=1}^{n} \hat{u}_j(t) w_j(x) \quad \in P_n H$$

et

$$u_{m-n}(x,t) = \sum_{j=n+1}^{m} \hat{u}_j(t) w_j(x) \quad \in P_{m-n} H$$

u_m appartenant à la variété non linéaire $\mathcal{M}_1$ ie u_n et u_{m-n} sont solutions du système suivant où l'on note $P_{m-n} = P_m - P_n$ le projecteur sur $\mathrm{Vect}(w_{n+1}, \ldots, w_m)$:

$$\begin{aligned} \frac{\partial u_n}{\partial t} + \nu A u_n + P_n(B(u_n,u_n) + B(u_n,u_{m-n}) + B(u_{m-n},u_n)) &= P_n f \\ \nu A u_{m-n} + P_{m-n} B(u_n,u_n) &= P_{m-n} f. \end{aligned} \tag{I.25}$$

Si l'on met (I.25) sous la forme :

$$\begin{aligned} u_{m-n} &= (\nu A)^{-1}(P_{m-n} f - P_{m-n} B(u_n,u_n)) = \Psi(u_n) \\ \frac{\partial u_n}{\partial t} + \nu A u_n + P_n(B(u_n,u_n) + B(u_n,\Psi(u_n)) + B(\Psi(u_n),u_n)) &= P_n f \end{aligned} \tag{I.26}$$

alors u_{m-n} apparait comme une correction de u_n, la dynamique de l'écoulement étant déterminée par les modes inférieur à n ; le nombre de modes nécessaire est ainsi réduit (cf. figure I.2). Il est à noter que u_{m-n} corrige de façon non négligeable cette dynamique sur de grands intervalles de temps.

Stabilité et convergence.

Marion et Temam [16] ont montré, dans le cas où $m = dn$ $(d > 1)$, la stabilité et la convergence du schéma (I.26) sans discrétisation en temps au sens suivant :

$\forall u_0 \in H$, condition initiale
$u_n \longrightarrow u$ dans $L^2(0,T;V)$ et $L^p(0,T;H)$, $\forall T > 0$, $1 \leq p < +\infty$
$u_n \rightharpoonup^{\star} u$ dans $L^\infty(\mathbb{R}^+, H)$ lorsque $n \longrightarrow +\infty$.

Remarque I.7 *Pour que la consistance du schéma soit assurée, il est essentiel de négliger $P_n B(z_n, z_n)$ dans la première équation du système (I.25) ; en particulier dans l'établissement des estimations a priori.*

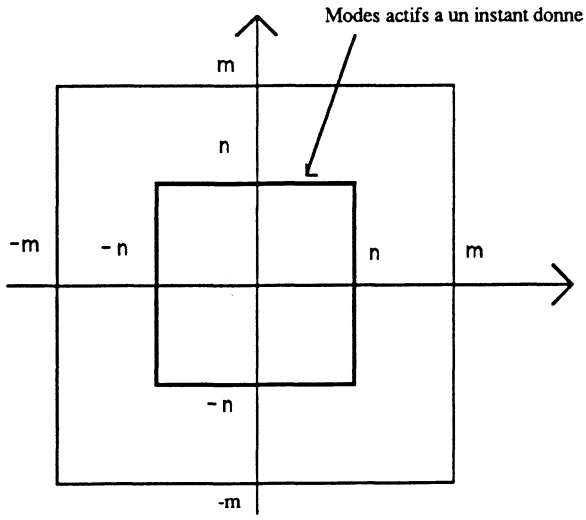


Figure I.2: *Degrés de liberté*

L'analyse de la discrétisation en temps a été effectuée par Jauberteau *el al* [13] pour $m = 2n$ et par J. Shen [22] dans le cas où $m = dn$. Considérons pour la discrétisation en temps, un schéma explicite pour les termes non linéaires et implicite pour les termes linéaires qui s'écrit (Δt étant le pas de temps) :

- pour la méthode de Galerkin classique

$$\frac{u_m^{k+1}-u_m^k}{\Delta t} + \nu A u_m^{k+1} = P_m(f^k - B(u_m^k, u_m^k)) \tag{I.27}$$

- pour la méthode de Galerkin non linéaire

$$\begin{aligned} \nu A u_{m-n}^{k+1} &= P_{m-n}(f^k - B(u_n^k, u_n^k)) \\ \frac{u_n^{k+1}-u_n^k}{\Delta t} + \nu A u_n^{k+1} &= P_n(f^k - B(u_n^k, u_n^k) \\ &\quad - B(u_{m-n}^{k+1}, u_n^k) - B(u_n^k, u_{m-n}^{k+1})) \end{aligned} \tag{I.28}$$

où u_n^k (respectivement u_m^k, u_{m-n}^k) est la valeur de u_n (respectivement u_m, u_{m-n}) au temps $k\Delta t$. Ce schéma est conditionnellement stable : sous réserve de conditions portant sur Δt du type $\Delta t \lambda_n \leq cste$, on a

$$\begin{aligned} \forall u_o \in H, \exists K(u_o) > 0 \ / \ \forall k > K(u_0) \quad & |u_n^k|^2 + |u_{m-n}^k|^2 \leq R_1(\nu, \lambda_1, f) \\ & ||u_n^k||^2 + ||u_{m-n}^k||^2 \leq R_2(\nu, \lambda_1, f). \end{aligned}$$

où R_1 et R_2 sont des constantes ne dépendant que de ν, λ_1 et f. La condition de stabilité porte sur λ_n indépendamment du nombre d permettant d'envisager l'utilisation d'un pas de temps plus grand dans le schéma de Galerkin non linéaire

que dans la méthode de Galerkin classique pour laquelle la condition de stabilité est :

$$\Delta t \lambda_{dn}^{1/2} \leq 1.$$

Il est cependant nécessaire que d soit suffisamment grand pour qu'un tel choix de pas de temps soit possible car la condition de stabilité dans le cas non linéaire est plus forte. On peut ainsi espérer réduire le temps de calcul par rapport à la méthode de Galerkin classique et par conséquent le coût des simulations numériques des écoulements turbulents.

Remarque I.8 *D'autres méthodes d'intégration telles qu'une méthode explicite d'Adams-Bashforth du second ordre ou une méthode explicite de Runge–Kutta d'ordre 3 (cf. Jauberteau* et al *[13]) peuvent être employées et conduisent au même type de résultats pour la stabilité.*

Choix de la troncature.

La difficulté essentielle réside dans le choix des modes les plus actifs pour l'écoulement, c'est-à-dire dans la détermination du nombre d'onde n séparant grandes et petites échelles. Jauberteau [11], Jauberteau *et al* [12], Dubois *et al* [6] ont proposé de déterminer ce nombre de façon dynamique ie de l'évaluer de façon régulière au cours du temps. Pour cela on fixe m le nombre d'onde maximum ($(2m)^2$ est donc la taille de la grille la plus fine sur laquelle est approchée la solution) et on se donne des nombres d'onde admissibles $n_1, \cdots, n_i, \cdots$ tels que $(4 \leq n_1 < \cdots < n_i < \cdots \leq m)$ auxquels correspondent des grilles de moins en moins grossières .

Le niveau n_i est défini dynamiquement

- soit (cf. Jauberteau [11]) en comparant l'énergie cinétique de u_{m-n_i} avec l'erreur de discrétisation en temps estimée à : $(\Delta t)^r$ où r est l'ordre du schéma en temps.
- soit (cf. Dubois *et al* [6]) en comparant l'énergie cinétique de u_{m-n_i} avec l'énergie cinétique de u_{n_i}
- soit (cf. Pascal et Basdevant [Annexe A]) en comparant l'enstrophie de u_{m-n_i} avec l'enstrophie de u_{n_i}
- soit (cf. Pascal et Basdevant [Annexe A]) en comparant les termes non linéaires de couplage.

Rappelons que l'on appelle énergie cinétique modale à l'instant t :

$$\tilde{E}(k) = \frac{1}{2}|\hat{u}_k|^2 = \frac{1}{2}k^2|\hat{\psi}_k|^2 = \frac{1}{2}\frac{|\hat{\omega}_k|^2}{k^2},$$

l'enstrophie modale étant le carré du rotationnel de la vitesse :

$$\tilde{Z}(k) = \frac{1}{2}|\hat{\omega}_k|^2$$

où ω est le rotationnel de la vitesse u et ψ la fonction de courant associée à u. Dans le cas d'une cavité périodique Ω de période L dans chaque direction, l'énergie s'écrit dans le cas continu :

$$E = \frac{1}{2} < u^2 > = \frac{1}{2L^2}\int_\Omega u^2(x)dx = \frac{L^2}{4\pi^2}\int_{\mathbb{R}^2} \tilde{E}(k)dk$$

et dans le cas discret (pour les méthodes de Galerkin, $\mathbb{Z}^2$ est remplacé par l' ensemble des modes excités $\{k \in \mathbb{Z}^2, |k| \leq m\}$) :

$$E = \frac{1}{2}\sum_{k\in\mathbb{Z}^2} |\hat{u}_k|^2,$$

l'enstrophie étant dans le cas continu :

$$Z = \frac{1}{2} < \omega^2 > = \frac{1}{2L^2}\int_\Omega \omega^2(x)dx = \frac{L^2}{4\pi^2}\int_{\mathbb{R}^2} \tilde{Z}(k)dk$$

et dans le cas discret :

$$Z = \frac{1}{2}\sum_{k\in\mathbb{Z}^2} |\hat{\omega}_k|^2.$$

E et Z sont 2 invariants des équations de Navier–Stokes à viscosité (ν) nulle et à force extérieure (f) nulle.

On appelle spectre d'énergie à l'instant t la fonction $E(k)$ telle que :

$$E(k) = \frac{1}{2\pi}\int_{|l|=k} k\tilde{E}(l)d\theta \quad \text{avec } \theta \text{ argument de } l \in \mathbb{R}^2.$$

L'énergie totale s'écrit en fonction du spectre :

$$E = \frac{L^2}{2\pi}\int_0^{+\infty} E(k)dk.$$

Dans le cas discret, $\{I_j\}_j$ étant un recouvrement de $\mathbb{Z}^2$ (ou de l'ensemble $\{k \in \mathbb{Z}^2, |k| \leq m\}$ pour les méthodes de Galerkin), le spectre d'énergie est la fonction discrète définie sur $\{I_j\}_j$ par :

$$E(k_j) = \frac{k_j}{\alpha_j}\sum_{|k|\in I_j} \tilde{E}(k)$$

où k_j est le nombre d'onde central de I_j et α_j est la multiplicité de k_j ie le nombre d'élément de I_j.

La détermination du spectre d'enstrophie est identique à celle du spectre d'énergie.

Les critères utilisés pour déterminer n_i comparant les énergies et les enstrophies sont les suivants (θ_Z et θ_E étant des paramètres sans dimension) :

$$\frac{E(u_{m-n_i})}{E(u_m)} \leq \theta_E^2$$

ou

$$\frac{Z(u_{m-n_i})}{Z(u_m)} \leq \theta_Z^2.$$

Regardons dans un cadre théorique quelles sont les conséquences d'un tel critère sur n_i.

Si l'on suppose (cf. théorie de Kraichnan) que le spectre d'énergie est en k^{-3} sur l'ensemble des modes allant des grandes échelles k_0^{-1} à l'échelle de dissipation k_d^{-1} et si l'on suppose que la troncature n_i est dans cette zone alors en notant $k_\star$ le nombre d'onde associé à la troncature n_i et en négligeant énergie et enstrophie au-delà de k_d, on peut évaluer le rapport de l'énergie cinétique des petites structures $E(u_{m-n_i})$ sur l'énergie cinétique des grandes échelles $E(u_{n_i})$ ou sur l'énergie totale du système $E(u_m)$ de la façon suivante :

$$\frac{E(u_{m-n_i})}{E(u_{n_i})} = \frac{||u_{m-n_i}||^2}{||u_{n_i}||^2} = \frac{\int_{k_\star}^{k_d} k^{-3} dk}{\int_{k_0}^{k_\star} k^{-3} dk} = \frac{{k_\star}^{-2} - {k_d}^{-2}}{{k_0}^{-2} - {k_\star}^{-2}} = \frac{1 - (\frac{k_\star}{k_d})^2}{(\frac{k_\star}{k_d})^2 (\frac{k_d}{k_0})^2 - 1}$$

ou encore

$$\frac{E(u_{m-n_i})}{E(u_{n_i})} = \frac{1-\alpha}{\alpha Re - 1}$$

où $\alpha = (\frac{k_\star}{k_d})^2$ représente le pourcentage des degrés de liberté associés à u_{n_i} sur l'ensemble des degrés de liberté. On en déduit que :

$$\frac{E(u_{m-n_i})}{E(u_m)} = \frac{E(u_{m-n_i})}{E(u_{n_i}) + E(u_{m-n_i})} = \frac{1-\alpha}{\alpha(Re-1)}.$$

De même et avec des hypothèses identiques, on peut évaluer le rapport de l'enstrophie des petites structures $Z(u_{m-n_i})$ sur l'enstrophie des grandes échelles $Z(u_{n_i})$ ou sur l'enstrophie totale du systéme $Z(u_m)$ par :

$$\frac{Z(u_{m-n_i})}{Z(u_{n_i})} = \frac{||\omega_{m-n_i}||^2}{||\omega_{n_i}||^2} = \frac{\int_{k_\star}^{k_d} k^{-1} dk}{\int_{k_0}^{k_\star} k^{-1} dk} = \frac{-\ln \alpha}{\ln \alpha + \ln Re}$$

et

$$\frac{Z(u_{m-n_i})}{Z(u_m)} = \frac{-\ln \alpha}{\ln Re}.$$

Les courbes $\frac{E(u_{m-n_i})}{E(u_m)}$ et $\frac{E(u_{m-n_i})}{E(u_{n_i})}$ des figures I.3 et I.4 montrent que les petites échelles u_{m-n_i} concentrent une petite fraction de l'énergie totale et que l'énergie $E(u_{m-n_i})$ ne représente qu'un faible pourcentage ($\simeq$ 1% avec $Re = 100$ et $\leq 0.01\%$ avec $Re = 10000$) de $E(u_m)$ et ceci quelque soit $\alpha \geq 0.5$. Ces 2 rapports sont d'autant plus petits que Re augmente. Ainsi une part non négligeable de modes (jusqu'à 50%) transporte une très faible quantité d'énergie, cependant les courbes concernant l'enstrophie montrent que ces mêmes modes transportent une part d'enstrophie importante ($\leq 15\%$ avec $Re = 100$ et $\leq 8\%$ avec $Re = 10000$) qui ne décroit qu'avec le logarithme du nombre de Reynolds. Afin de réduire la part d'enstrophie transportée par u_{m-n_i}, il est nécessaire que α soit proche de 1 ie $k_\star \simeq k_d$. Par conséquent la troncature $k_\star$, si elle se situe dans la zone inertielle, doit être proche de la zone de dissipation.

Il est à noter que si la théorie phénoménologique inclue l'intermittence spatio-temporelle (cf. Basdevant et Sadourny [2]), le spectre d'énergie est alors plus pentu ce qui entraine des rapports plus faibles et $k_\star$ légèrement plus petit.

Comme les critères de détermination de n_i comparant les énergies et les enstrophies induisent :

$$\frac{E(u_{m-n_i})}{E(u_{n_i})} \leq \theta_E^2 \Longleftrightarrow k_\star \geq k_d \sqrt{\frac{1+\theta_E^2}{1+Re\theta_E^2}}$$

$$\frac{Z(u_{m-n_i})}{Z(u_{n_i})} \leq \theta_Z^2 \Longleftrightarrow k_\star \geq k_d Re^{-\frac{\theta_Z^2}{2+2\theta_Z^2}}$$

ou

$$\frac{E(u_{m-n_i})}{E(u_m)} \leq \theta_E^2 \Longleftrightarrow k_\star \geq k_d \sqrt{\frac{1}{1+(Re-1)\theta_E^2}}$$

$$\frac{Z(u_{m-n_i})}{Z(u_m)} \leq \theta_Z^2 \Longleftrightarrow k_\star \geq k_d Re^{-\frac{\theta_Z^2}{2}},$$

les paramètres θ_Z et θ_E déterminés empiriquement vérifient les relations suivantes :

$$\theta_E^2 \ll \frac{1}{Re} \quad \text{et} \quad \theta_Z^2 \ll 1.$$

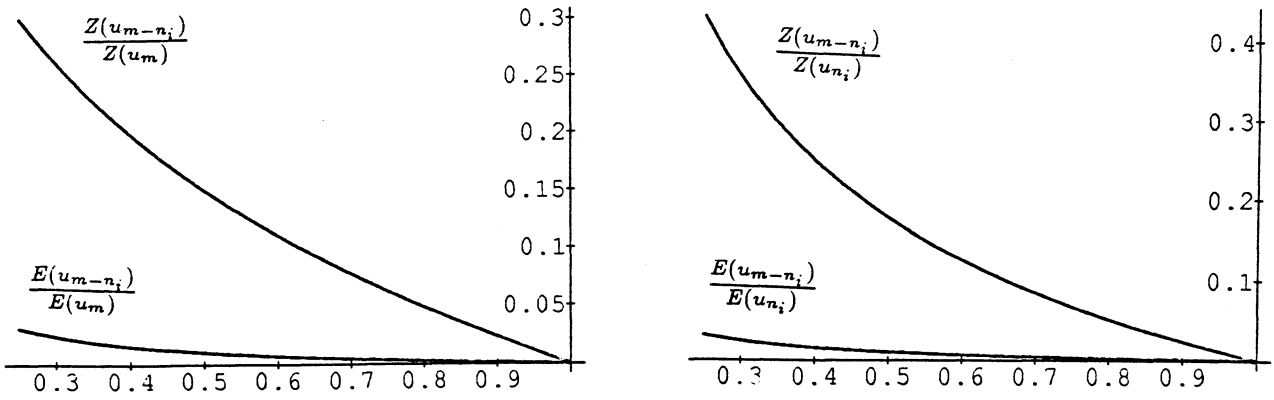


Figure I.3: $Re = 100$. a) $\frac{Z(u_{m-n_i})}{Z(u_m)}$ *vs* $\frac{E(u_{m-n_i})}{E(u_m)}$; b) $\frac{Z(u_{m-n_i})}{Z(u_{n_i})}$ *vs* $\frac{E(u_{m-n_i})}{E(u_{n_i})}$.

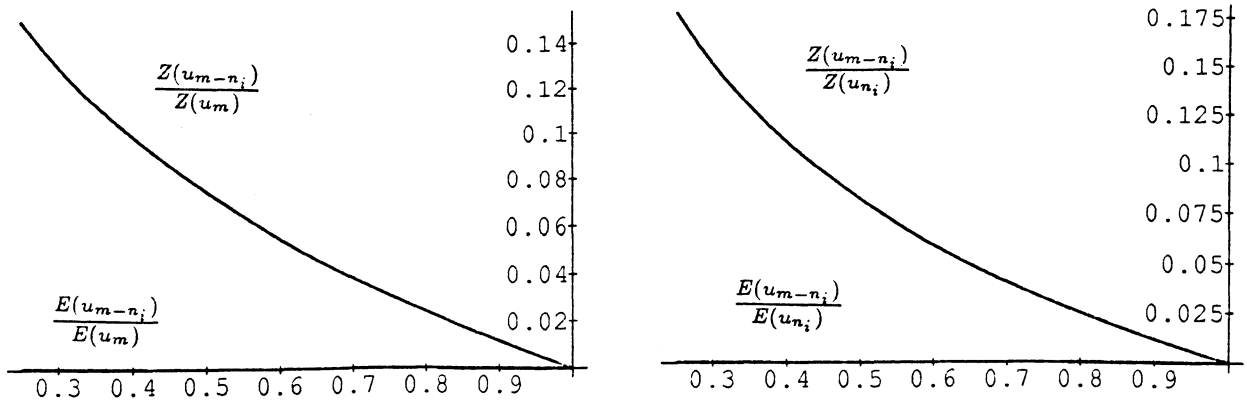


Figure I.4: $Re = 10000$. a) $\frac{Z(u_{m-n_i})}{Z(u_m)}$ *vs* $\frac{E(u_{m-n_i})}{E(u_m)}$; b) $\frac{Z(u_{m-n_i})}{Z(u_{n_i})}$ *vs* $\frac{E(u_{m-n_i})}{E(u_{n_i})}$.

Premiers tests numériques réalisés par Rosier, Jauberteau et Dubois.

Les articles [6], [11], [12], [12] exposent des résultats concernant les premiers tests numériques. Ces tests concernent essentiellement :

cas 1 des solutions exactes connues du type

$$\begin{aligned} u_1(x_1,x_2,t) &= -g(t)w\cos(wx_1)\sin(wx_2)\exp(\cos(wx_1)\cos(wx_2)) \\ u_2(x_1,x_2,t) &= +g(t)w\sin(wx_1)\cos(wx_2)\exp(\cos(wx_1)\cos(wx_2)) \end{aligned}$$

où la fonction g est donnée par

$$g(t) = \frac{1}{40}\left(\frac{1}{10}e^{(\sin 2t - 3\sin 2\pi t)} + \cos(2\sqrt{2}t)\right) + \frac{0.375}{5}$$

cas 2 des simulations directes d'écoulements.

Les principaux résultats numériques portent sur :

- la précision et la convergence : l'erreur absolue par rapport à la solution exacte dans le cas 1) et l'erreur relative dans le cas 2) sont identiques au cours du temps pour Galerkin classique et Galerkin non linéaire,
- la stabilité : le schéma G.N.L. est plus stable que le schéma G.C.,
- le gain de temps : le gain de temps CPU est de l'ordre de 40% par rapport à la méthode de Galerkin classique, ce qui est très appréciable.

Cependant d'une part les tests effectués avec des solutions exactes ne présentent pas toutes les difficultés rencontrées lors de simulations d'écoulements turbulents (en particulier le spectre d'énergie reste constant au cours du temps), et d'autre part les moyens informatiques malgré leurs progrès ne permettent pas de réaliser des simulations directes avec de grands nombres de Reynolds car il est impossible de tenir compte de toutes les échelles nécessaires. Seules des simulations avec de faibles nombres de Reynolds ont été mis en oeuvre.

Nous présentons dans la suite les motivations numériques et les résultats de la méthode de Galerkin non linéaire dans le cadre d'un modèle turbulent où sont simulées les effets des échelles tronquées.

I.3 Motivations numériques.

I.3.1 Simulation numérique des écoulements turbulents dans une cavité 2Π périodique.

Les équations.

On considère le domaine $\Omega = (0, 2\Pi) \times (0, 2\Pi)$ de $\mathbb{R}^2$. et les équations de Navier–Stokes en formulation fonction de courant ψ et vorticité ω (cf I.1.1) :

$$\begin{aligned} \frac{\partial \omega}{\partial t} + J(\omega, \psi) &= \nu \Delta \omega + rot f \quad \text{dans } Q \\ \omega &= -\Delta \psi \quad \text{dans } Q \\ \omega(x, 0) &= \omega_0(x) \quad \text{dans } \Omega \end{aligned} \tag{I.29}$$

avec des conditions aux limites périodiques.

Méthode de Galerkin classique.

La solution pouvant être développée en série de Fourier, la méthode de Galerkin classique consiste à chercher une approximation ω_m de la solution dans l'espace de dimension finie engendré par les premiers vecteurs propres de l'opérateurs de stokes ie par $w_j(x) = e^{ij.x}$ tels que $j \in \mathbb{Z}^2$ et $|j| \leq m$. On a donc :

$$\omega_m(x, t) = \sum_{|j| \leq m} \hat{\omega}_j(t) e^{ij.x}$$

Le schéma de discrétisation spatiale est la méthode de collocation ou (pseudo-spectrale) (cf. Orzag [17]) utilisant pour le calcul du jacobien $J(\omega, \psi)$ les transformées FFT (Fast Fourier Transform) qui permettent de passer très rapidement des valeurs prises par une fonction aux points d'une grille régulière (espace physique) à ses coefficients de Fourier (espace spectral) et inversement ; l'idée étant que la dérivation est simple en Fourier (multiplication des coefficients par un complexe) et que le produit de 2 fonctions est simple dans l'espace physique (multiplication de 2 réels). Le calcul du jacobien est l'étape la plus coûteuse en nombre d'opérations et les erreurs d'aliasing (cf. Paterson et Orszag [17]) sont supprimées en réduisant d'un tiers le nombre de modes.

La partie linéaire de l'équation est intégrée exactement en temps. Les termes non linéaires sont discrétisés de façon explicite avec un schéma d'Adams–Bashforth du second ordre qui résout les problèmes du type $f'(t) = g(t)$ en déterminant f au pas de temps $t_{n+1} = (n+1)\Delta t$ par :

$$f(t_{n+1}) = f(t_n) + \frac{\Delta t}{2}(3g(t_n) - g(t_{n-1})).$$

Le schéma est initialisé par une méthode de Runge–kutta d'ordre 2. Le choix du pas de temps Δt est alors soumis à une condition du type CFL :

$$\|u\frac{\Delta t}{\Delta x}\| < 1$$

Simulation directe – Simulation des grandes échelles.

L'écoulement est correctement simulé si l'ensemble des transferts d'énergie et d'enstrophie est modélisé. Par conséquent pour envisager une simulation directe d'un écoulement turbulent, il faut un nombre de points de maille de l'ordre du nombre de Reynolds (en dimension 2) car il faut tenir compte de toutes les échelles entre l'échelle caractéristique et l'échelle de dissipation de Kolmogorov. Ce n'est généralement pas envisageable pratiquement. Dans le cas de la simulation atmosphérique aux échelles planétaires (en 2D), il faut tenir compte de 5 10^6 degrés de liberté ($l_0 = 2\ 10^7\ m, l_d = 10^4\ m$) interagissant non linéairement. Les ordinateurs actuels ne permettent pas d'envisager de telles simulations et en pratique pour de très grands nombres de Reynolds la troncature m se situe dans la zone inertielle (zone non influencée par le forçage et la dissipation) ; seules les grandes échelles sont simulées.

Cependant les effets statistiques des petites échelles sur les grandes et les transferts d'énergie et d'enstrophie ne sont plus modélisés, l'enstrophie s'accumule (cf. Basdevant et Sadourny [3]) et l'effet de la troncature est de développer excessivement des petites structures (c'est ce qui se passe lors du run représenté en figure 12 de l'annexe A). Basdevant et Sadourny [2] ont proposé pour modéliser les effets des échelles virtuelles (i.e. au delà de la troncature) qui ne sont pas explicitement résolus, d'introduire un opérateur de superviscosité

$$-\frac{\nu}{(k_T)^{2p}}(\Delta)^p = -\frac{\nu}{(k_T)^{2p}}(\frac{\partial^2}{\partial {x_1}^2} + \frac{\partial^2}{\partial {x_2}^2})^p$$

avec $p = 2, 4$ ou 8, où k_T est le mode de troncature et ν, la viscosité artificielle, est choisie de telle sorte que $\nu\Delta t \simeq 0.5$. Cet opérateur dissipe l'enstrophie tout en minimisant le transfert d'énergie à travers la troncature. Dans ces modèles sous-mailles, le nombre de Reynolds définis en I.1.1 n'est plus un paramètre essentiel dès qu'il est suffisamment grand c'est à dire dès que l'échelle de dissipation est négligeable devant celle de la troncature et que les termes non linéaires dominent les termes visqueux.

On observe figure I.5, lors de l'évolution libre de 2 vortex de même signe, les effets des différentes dissipativités sur l'énergie et l'enstrophie du système. Comme il n'y a pas de forçage et de dissipation à grande échelle, l'energie est conservée au cours du temps (elle est d'autant mieux conservée que p est grand). L'enstrophie diminue au cours du temps (elle diminue d'autant plus que p est petit). En pratique $p = 4$ et $p = 8$ sont fréquemment utilisés car l'énergie est

peu dissipée et les transferts d'enstrophie simulés (ie l'échelle de dissipation est supérieure à l'échelle de la maille au voisinage de la troncature).

On voit sur les spectres d'énergie (figures I.8 et I.9) que la zone de dissipation diminue et se concentre au voisinage de la troncature lorsque p augmente.

Cependant les échelles sous-mailles n'étant pas résolues de façon parfaite, le modéle numérique ne peut décrire exactement l'évolution des grandes échelles d'un point de vue déterministe, du moins pour des temps plus grands que le temps de prédicibilité, mais prédit correctement les propriétés statistiques de la turbulence et les formes tourbillonnaires des écoulements aux échelles simulées.

Les expériences sur Galerkin classique et l'implémentation de Galerkin non linéaire ont été effectuées à partir du code numérique du Laboratoire de Météorologie Dynamique développé par C. Basdevant et B. Legras [1].

Premier type d' expériences numériques : turbulence forcée

Pour les expériences de turbulence forcée, le forçage est obtenu par maintient constant de l'amplitude d'un coefficient de Fourier du développement de la vorticité ω au nombre d'onde $k_I = 10$, nombre d'onde d'injection :

$$\hat{\omega}_{(10,0)} = \frac{10.0i}{\sqrt{2}}.$$

L'état initial étant un champ de vorticité aléatoire de moyenne nulle, une dissipation d'énergie à grande échelle du type friction de Rayleigh ($-\nu_{Ra}\psi$) permet d'atteindre un état turbulent stationnaire dont le spectre (en coordonnées Log-Log) prévu par la phénoménologie (cf Le Roy [14]) est reproduit en figure I.6a.

On peut y distinguer 4 zones (des plus grandes échelles vers les plus petites) :

- une zone affectée par la friction de Rayleigh
- une cascade inverse d'énergie où $E(k) \simeq k^{-5/3}$
- une cascade d'enstrophie où $E(k) \simeq k^{-3}$
- une zone de dissipation où $E(k)$ décroit exponentiellement.

La figure I.6b représente le spectre d'énergie obtenu après 30000 pas de temps par la méthode de Galerkin classique et avec 128^2 degrés de liberté. On retrouve numériquement les 4 zones bien que la cascade d'enstrophie n'y soit pas très développée, l'échelle de forçage étant trop proche de l'échelle de troncature. Le champ de tourbillon est reproduit sur la figure I.7 aux instants $t = 0.3$ et $t = 3.0$; on peut y constater que les structures cohérentes sont de l'ordre de l'échelle de forçage ie de l'ordre de $\frac{1}{10}$ du cadre.

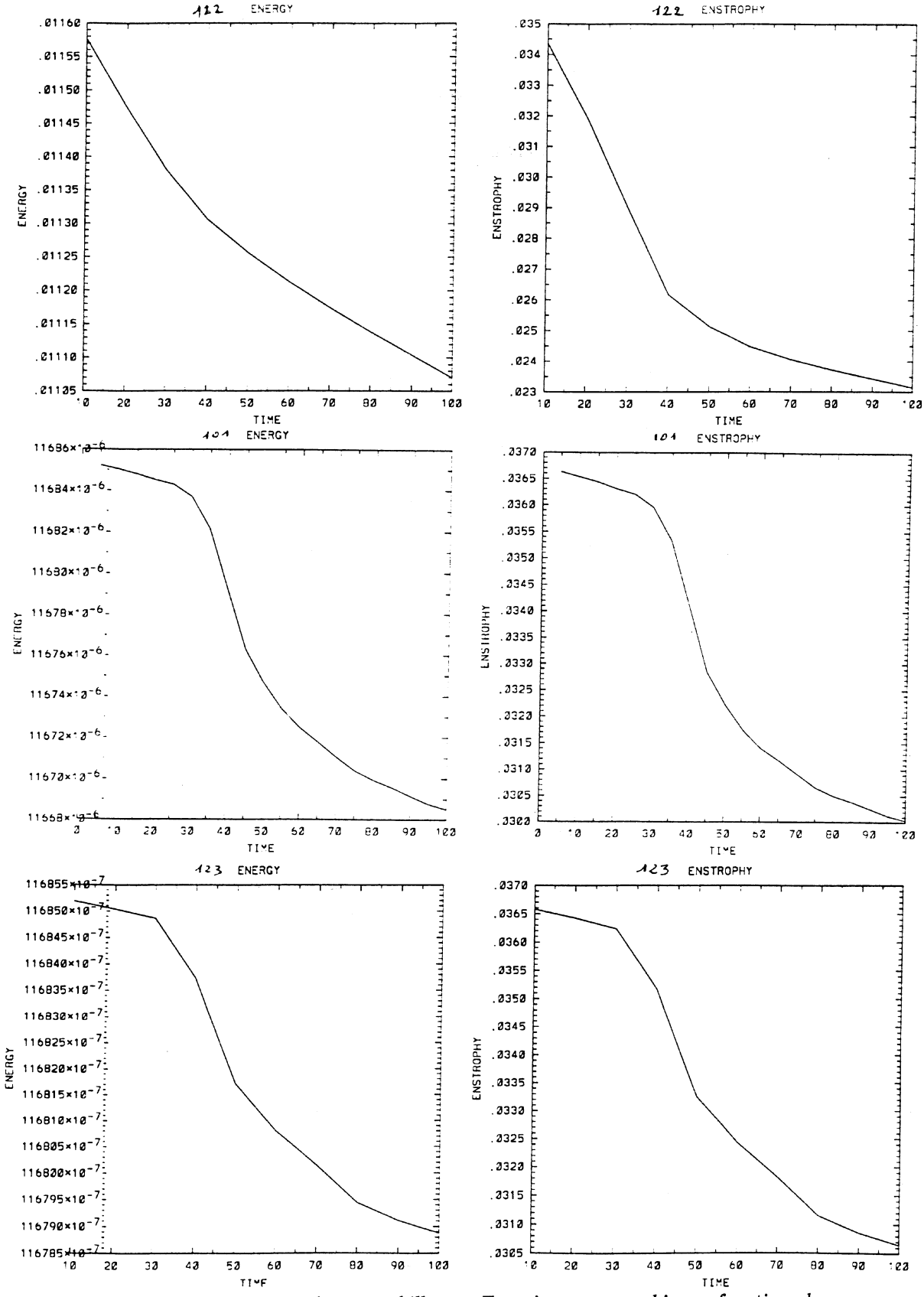


Figure I.5: *Interaction de 2 tourbillons : Energie et enstrophie en fonction du temps pour 1) $p = 2$; 2) $p = 4$; 3) $p = 8$ (p étant la puissance du laplacien itéré).*

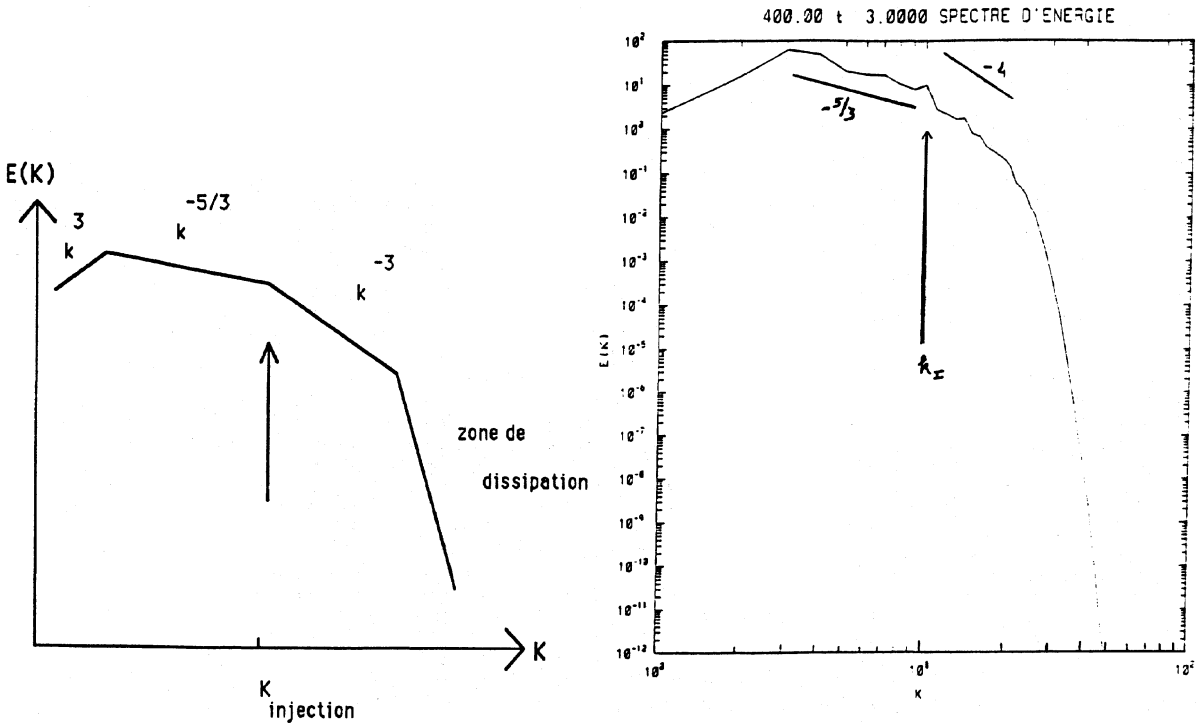


Figure I.6: *Turbulence entretenue : spectre d'énergie a) obtenu en théorie phénoménologique ; b) à $t = 3.0$ obtenu avec la méthode de Galerkin classique.*

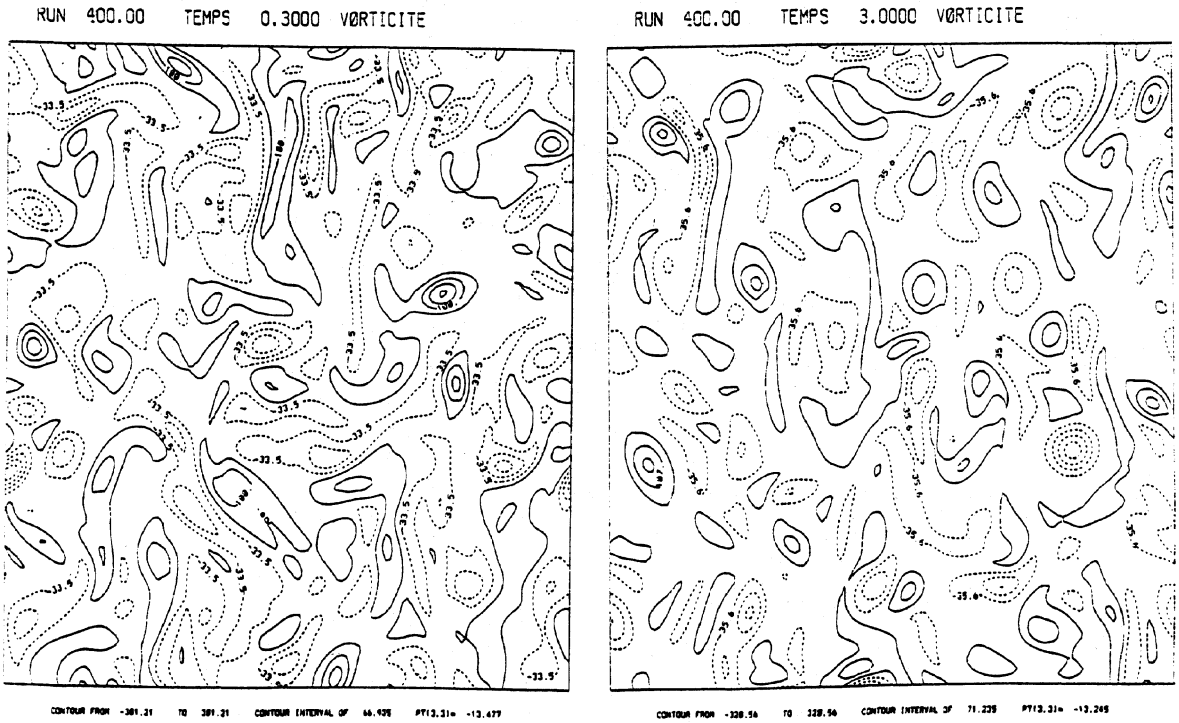


Figure I.7: *Turbulence entretenue : Champ de tourbillon à $t = 0.3$. et $t = 3.0$*

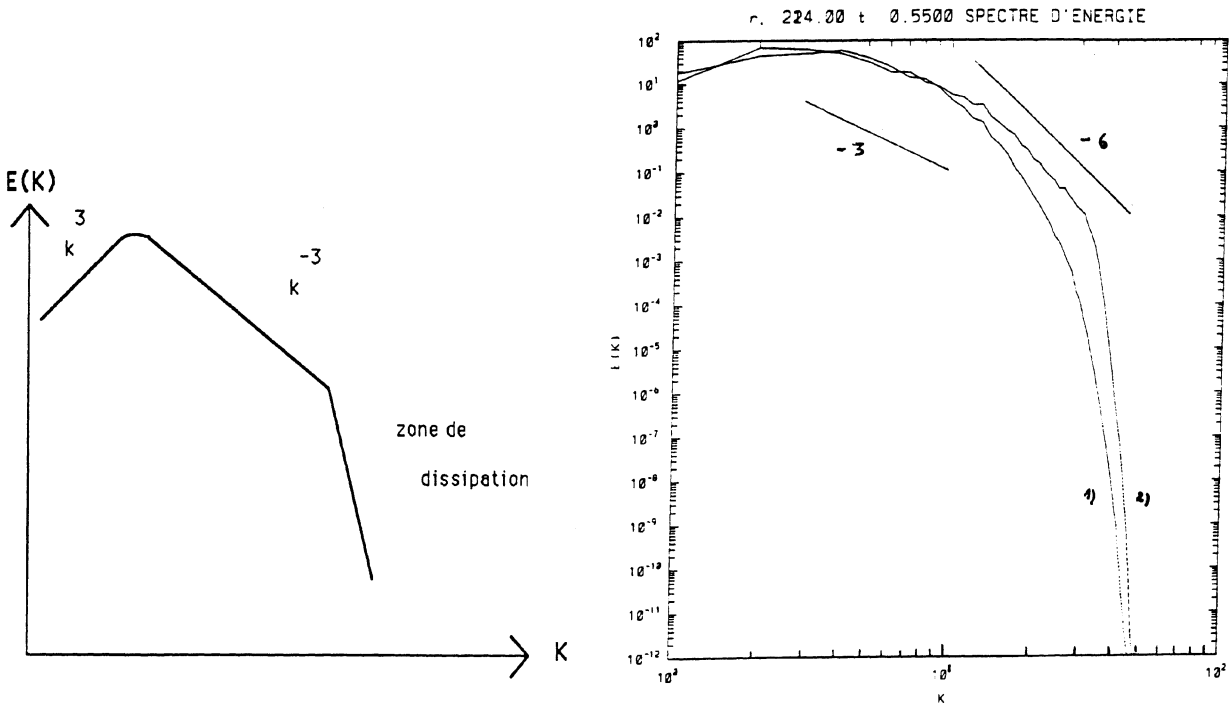


Figure I.8: *Ecoulement libre d'un champ initialement turbulent : spectre d'énergie a) prévu par la théorie phénoménologique ; b) à* $t = 30.0$ *obtenu avec la méthode de Galerkin classique et 1)* $p = 4$, *2)* $p = 8$.

Deuxième type d' expériences numériques : turbulence libre

Les 2 expériences suivantes concernent la simulation de la turbulence en régime libre ($f = 0$) :

1. Dans un premier cas (qu'on appellera "interaction de 2 tourbillons"), l'état initial est constitué de 2 vortex de même signe dans l'espace physique (la vorticité est la somme de 2 gaussiennes decentrées). Cette paire de tourbillons fusionnent s'ils ne sont pas trop éloignés l'un de l'autre. La figure I.10 représente l'histoire de la fusion de ces 2 tourbillons avec $p = 4$; le processus étant légèrement accéléré ou ralenti avec respectivement $p = 2$ et $p = 8$. Le pas de temps est dans ce cas 0.01.

2. Dans un deuxième cas (qu'on appellera "turbulence libre") l'état initial est un champ turbulent (obtenu lors de simulations forcées et fourni par A.Babiano et D. Oueslati du L.M.D.). La figure I.11 présente quelques images de l'évolution au cours du temps ; on remarquera tout particulièrement une fusion de 2 tourbillons positif à $t = 0.6$ ie vers 6000 pas de temps.

Dans le cas d'une turbulence en régime libre, le spectre d'énergie théorique (Staquet [23]) présente une cascade d'enstrophie et une zone de dissipation (figure I.8a) que l'on retrouve numériquement (figure I.8b).

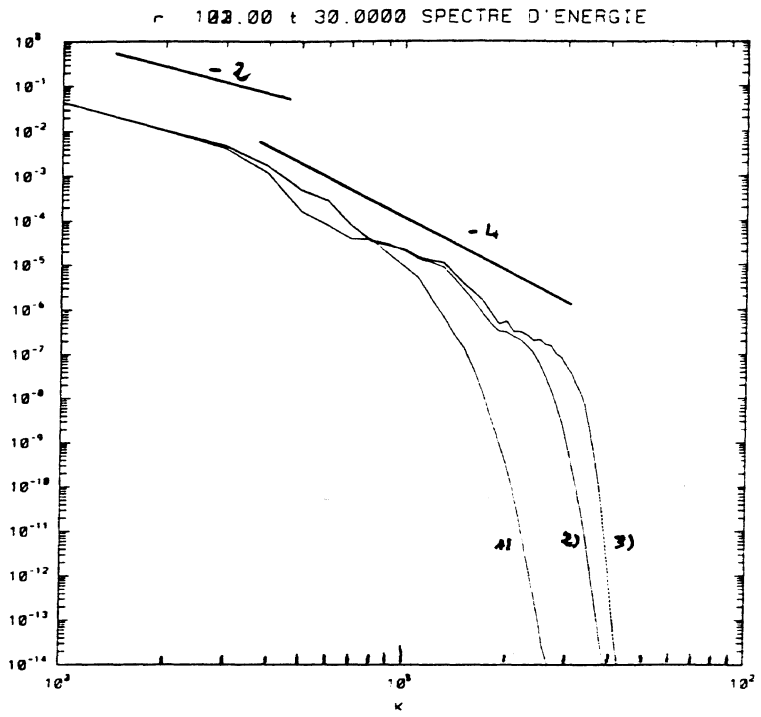


Figure I.9: *Interaction de 2 tourbillons : spectre d'énergie obtenu avec la méthode de Galerkin classique. 1) $p = 2$; 2) $p = 4$; 3) $p = 8$.*

I.3.2 Analyse des expériences numériques.

Introduction.

Tous les résultats numériques de ce paragraphe ont été réalisés avec la méthode de Galerkin classique décrite dans la section précédente. Ici nous analysons à partir des données ainsi obtenues, les différents termes et les différents rapports qui interviennent ou justifient l'algorithme de Galerkin non linéaire.

Pour définir la variété inertielle approximative décrite par une loi d'interactions entre petites et grandes structures, les équations de Navier–Stokes (I.29) sont projetées sur l'espace engendré par les premiers vecteurs propres w_j tels que $|j| \leq n_i$ et sur l'espace engendré par les vecteurs propres restant (ie tels que $n_i < |j| \leq m$) avec n_i le nombre d'onde séparant "grandes" et "petites" échelles. n_i vérifie $(4 \leq n_1 < \cdots < n_i < \cdots \leq m)$ et quelques propriétés dues aux FFT. Ces nombres $n_1, n_2, \cdots$ fixés définissent des maillages de plus en plus fins du domaine. Le système d'équations couplées est alors le suivant :

$$\frac{\partial \omega_{n_i}}{\partial t} + \nu A \omega_{n_i} + P_{n_i} J(\omega_m, \psi_m) = P_{n_i} f \tag{I.30}$$

$$\frac{\partial \omega_{m-n_i}}{\partial t} + \nu A \omega_{m-n_i} + P_{m-n_i} J(\omega_m, \psi_m) = P_{m-n_i} f \tag{I.31}$$

où

$$\omega_{n_i} = P_{n_i} \omega_m = \sum_{|j| \leq n_i} \hat{\omega}_j(t) e^{ij.x}$$

correspond aux petits nombres d'ondes et donc aux grandes échelles et où

$$\omega_{m-n_i} = P_{m-n_i} \omega_m = \sum_{n_i < |j| \leq m} \hat{\omega}_j(t) e^{ij.x}$$

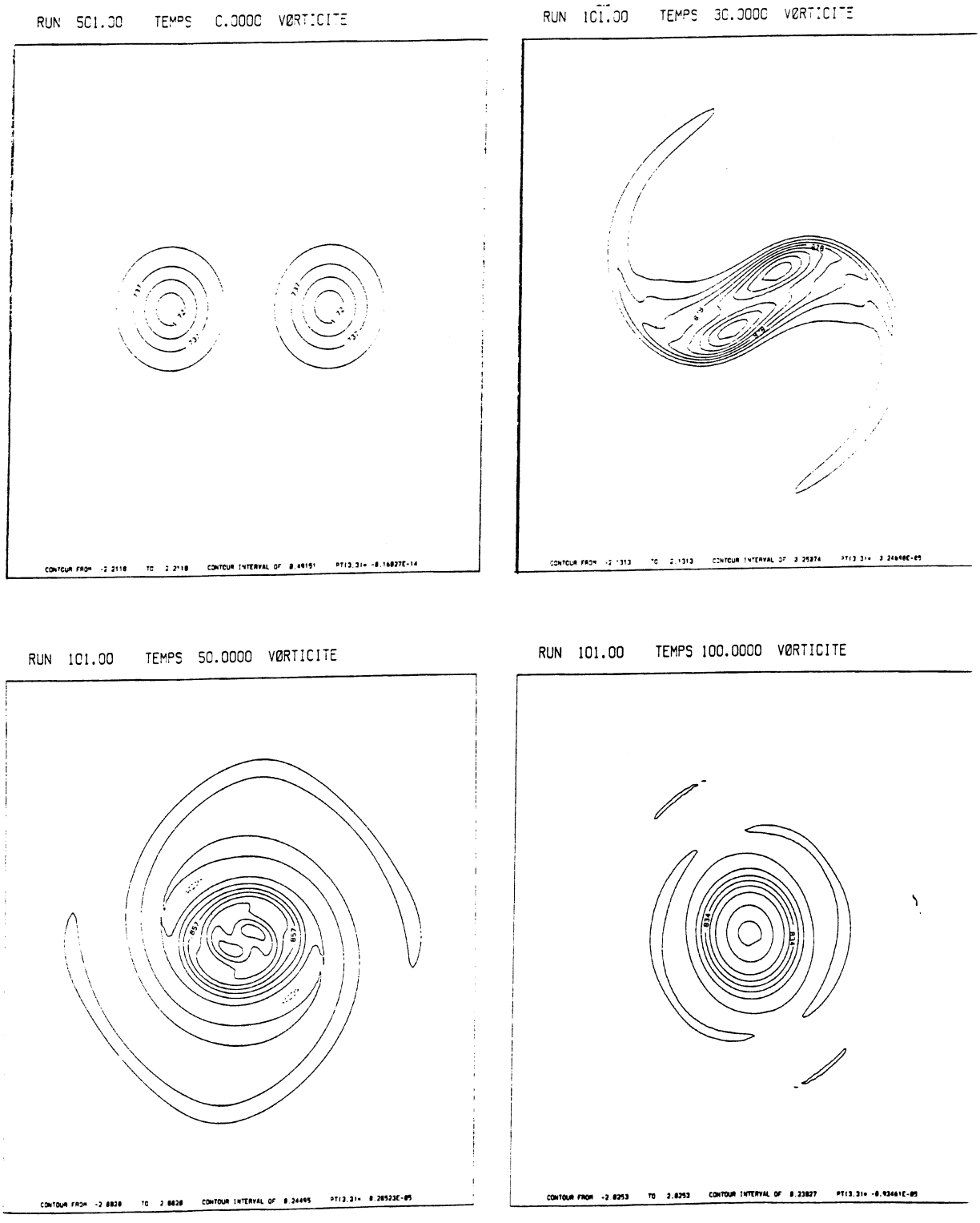


Figure I.10: *Interaction de 2 tourbillons : Histoire.*

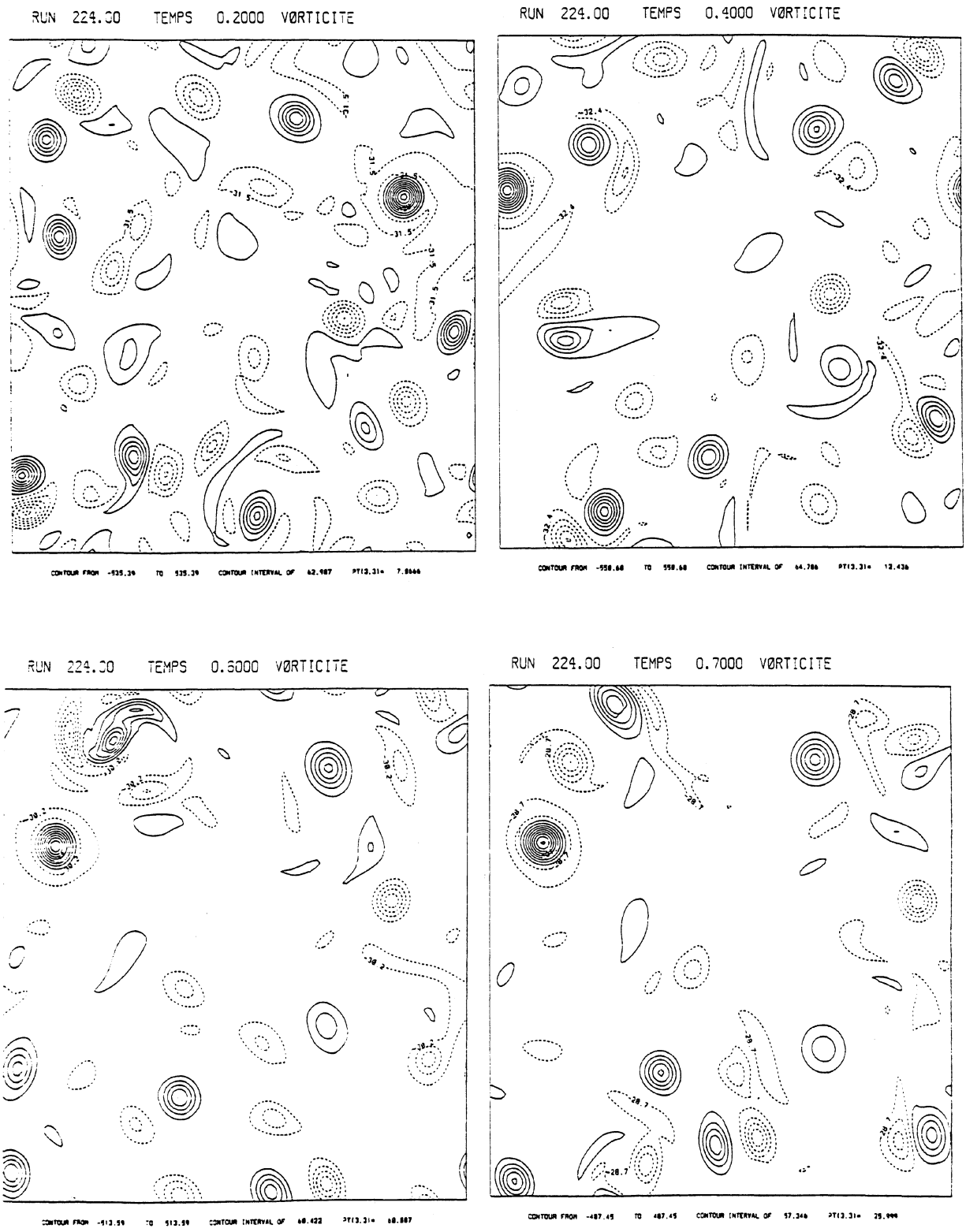


Figure I.11: *Ecoulement libre d'un champ initialement turbulent : Histoire.*

correspond aux petites longueurs d'ondes. Rappelons que $J(\omega_m, \psi_m)$ est composé de 2 parties :

- $J(\omega_{n_i}, \psi_{n_i})$, terme non linéaire associé à ω_{n_i},
- $J(\omega_{n_i}, \psi_{m-n_i}) + J(\omega_{m-n_i}, \psi_{n_i}) + J(\omega_{m-n_i}, \psi_{m-n_i})$, terme de couplage et d'interactions entre petites et grandes structures.

Etude des petites échelles et des grandes échelles.

La figure I.12 représente les champs ω_{n_i} et ω_{m-n_i} avec $n_i = 48$, $n_i = 32$ et $m = 64$. ω_{n_i} est une solution approchée de ω périodique de période 2π proche de ω_m. ω_{m-n_i} est par contre une superposition de fonctions périodiques toutes de période inférieure $\frac{2\pi}{m-n_i}$, l'ordre de grandeur de ω_{m-n_i} (environ 3.10^{-5} pour $n_i = 48$, 4.10^{-2} pour $n_i = 32$) étant inférieur à celui de ω_{n_i}. Le champ ω_m est réobtenu en sommant les 2. Sur le spectre d'énergie en figure I.13 à $t = 40.0$, sont représentées la zone de dissipation (constituée des modes supérieurs à 34) et les troncatures $n_i = 48$ et $n_i = 32$ qui se trouvent respectivement dans la zone de dissipation et dans la zone inertielle.

La figure I.15 montre le champ de vorticité ainsi que 2 coupes de ce champ lors de l'évolution en régime libre d'un champ initialement turbulent. On y remarque de nombreuses structures dont la fusion de 2 vortex dans le coin supérieur gauche.

On peut observer sur la figure I.16 que ces larges structures sont bien conservées lorsque l'on projette la solution sur $\text{Vect}(e^{ik.x}, |k| \leq 48)$ et que les coupes du champ ainsi obtenu sont très voisines des coupes du champ non tronqué (cf figure I.15). Ce n'est plus le cas lorsque l'on projette sur $\text{Vect}(e^{ik.x}, |k| \leq 16)$: la fusion des 2 vortex est floue, des structures apparaissent dans le marécage turbulent et les tailles des grandes structures sont réduites. Il est à noter que la limite zone inertielle – zone de dissipation se situe aux environs de $n_i = 32$ (cf. figure I.14).

Les figures I.18 et I.19 présentent la vorticité projetée sur les modes précédemment négligés. On constate dans le cas $n_i = 48$ que la vorticité est très faible et quasiment équirépartie en espace sauf au niveau de la fusion. Dans le cas $n_i = 16$ les structures réapparaissent.

Ces 2 exemples montrent combien il est nécessaire que la troncature n_i se situe dans la zone de dissipation ou proche de celle ci pour que ω_{n_i} soit proche de ω_m et que ω_{m-n_i} soit négligeable. Rappelons que dans le cas d'une telle troncature, il a été montré en I.2.3 (choix de la troncature) que la part d'enstrophie transportée par ω_{m-n_i} est très faible. Cela confirme donc le choix des paramètres θ_E et θ_Z des critères de détermination de n_i.

Les temps caractéristiques.

La figure I.20 qui représente les temps caractéristiques des grandes échelles ω_{n_i} estimés par l'expression :

$$\tau_{n_i} = \frac{||\omega_{n_i}||_{L^2}}{||\nu A\omega_{n_i} + P_{n_i}J(\omega_m, \psi_m)||_{L^2}}$$

et des petites échelles ω_{m-n_i} :

$$\tau_{m-n_i} = \frac{||\omega_{m-n_i}||_{L^2}}{||\nu A\omega_{m-n_i} + P_{m-n_i}J(\omega_m, \psi_m)||_{L^2}}$$

permet d'apprécier que d'une part ce temps est indépendant de la grille (ie de n_i) pour les grandes échelles et qu'il est d'autre part nettement inférieur pour les petites échelles : on retrouve donc numériquement que w_{m-n_i} évolue plus vite mais dans une amplitude moindre.

Une deuxième grandeur temporelle a été introduite : il s'agit du temps caractéristique des transferts des petites échelles ; seul le terme non linéaire intervient dans sa définition :

$$\tilde{\tau}_{m-n_i} = \frac{||\omega_{m-n_i}||_{L^2}}{||P_{m-n_i}J(\omega_m, \psi_m)||_{L^2}}$$

Comme le montre la comparaison de la figure I.21 et de la figure I.20, ces temps sont beaucoup plus courts que les temps caractéristiques de w_{m-n_i}. Ils nous permettront d'évaluer le temps pendant lequel les structures ω_{m-n_i} peuvent être figées sans modifier le comportement tourbillonnaire des grandes échelles.

$$\begin{array}{ccccc} ||\nu A\omega_{n_i}|| & \leq & ||P_{n_i}J(\omega_m,\psi_m)|| & \cong & ||\frac{\partial\omega_{n_i}}{\partial t}|| \\ \vee & & \vee & & \vee \\ ||\nu A\omega_{m-n_i}|| & \simeq & ||P_{m-n_i}J(\omega_m,\psi_m)|| & \gg & ||\frac{\partial\omega_{m-n_i}}{\partial t}|| \end{array}$$

Tableau I.1: *Relations entre les différents termes des équations ($m = 64$ et $n_i = 54$) : cas où la troncature est dans la zone de dissipation.*

$$\begin{array}{ccccc} ||\nu A\omega_{n_i}|| & \ll & ||P_{n_i}J(\omega_m,\psi_m)|| & \simeq & ||\frac{\partial\omega_{n_i}}{\partial t}|| \\ \wedge & & \vee & & \vee \\ ||\nu A\omega_{m-n_i}|| & \leq & ||P_{m-n_i}J(\omega_m,\psi_m)|| & \geq & ||\frac{\partial\omega_{m-n_i}}{\partial t}|| \end{array}$$

Tableau I.2: *Relations entre les différents termes des équations ($m = 64$ et $n_i = 32$) : cas où la troncature est dans la zone inertielle.*

Les différents termes du système.

Enfin les figures I.22, I.23 et I.24 permettent d'établir pour un temps supérieur à T_c, temps de transition pendant lequel la turbulence s'établit, et de l'ordre de 0.15, 0.20 pour l'expérience de la turbulence libre, les relations entre les différents termes des équations (I.30) et (I.31) à savoir les termes de dissipation, les termes non linéaires et les termes d'évolution. ces relations sont résumées dans le tableau (I.1) lorsque la troncature n_i se situe dans la zone de dissipation et dans le tableau (I.2) lorsque n_i est en fin de zone inertielle. On y remarque que le terme de dissipation est négligeable dans la zone inertielle et que le terme d'évolution l'est dans la zone de dissipation.

On notera que les rapports des différents termes des équations sont qualitativement les mêmes pour les 3 expériences numériques, l'ordre de grandeur et le temps de transition variant. Il est donc envisageable de figer les modes appartenant à la zone de dissipation ou de les intégrer en négligeant le terme d'évolution et certains termes de couplages non linéaires dans l'équation correspondante (I.31).

Remarque I.9 *Le rapport de $||\nu A\omega_{m-n_i}||$ sur $||\frac{\partial\omega_{m-n_i}}{\partial t}||$ lorsque la troncature n_i se situe dans la zone inertielle varie de par et d'autre de 1.0. Aucun de ces 2 termes ne reste au cours du temps plus grand ou plus petit que l'autre.*

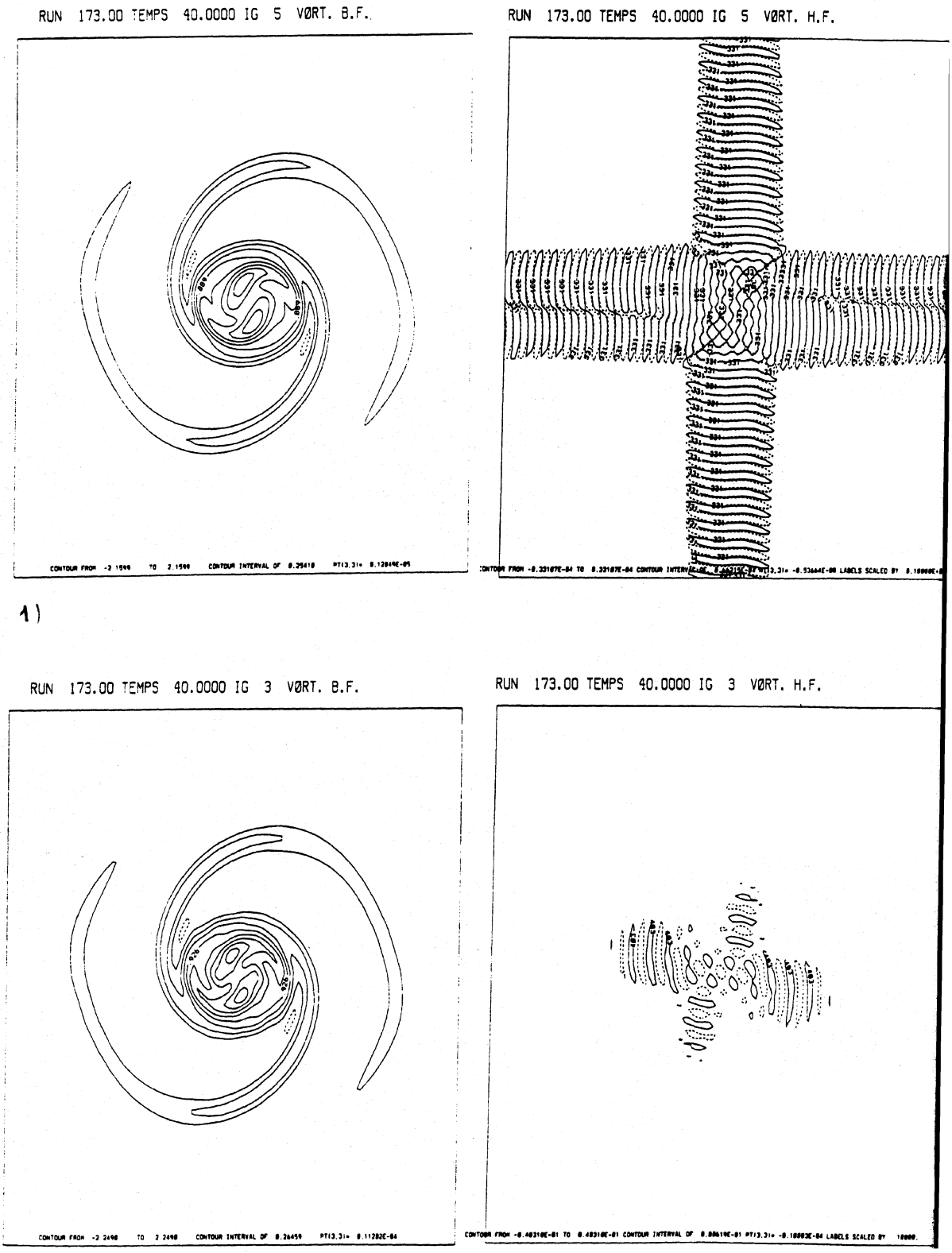


Figure I.12: *Interaction de 2 tourbillons : Projection P_{n_i} et P_{m-n_i} du rotationnel de la vitesse 1) $m = 64$ et $n_i = 48$; 2) $m = 64$ et $n_i = 32$*

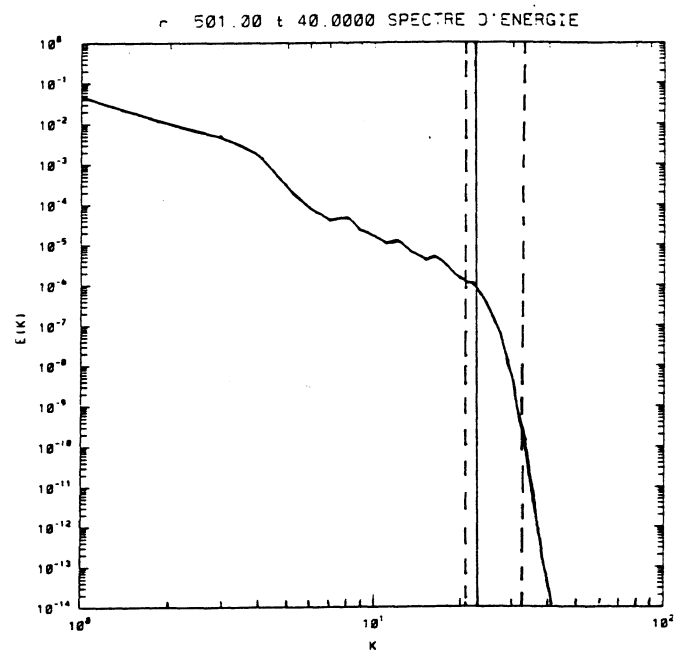


Figure I.13: *Interaction de 2 tourbillons : Spectre d'énergie à* $t = 40.0$ *avec la zone de dissipation (trait continu) et les troncatures* $n_i = 48$ *et* $n_i = 32$ *(trait en pointillé). Il faut multiplier les valeurs numériques par* $\frac{3}{2}$ *pour obtenir* n_i *sur l'axe des abscisses (cf. désaliasing).*

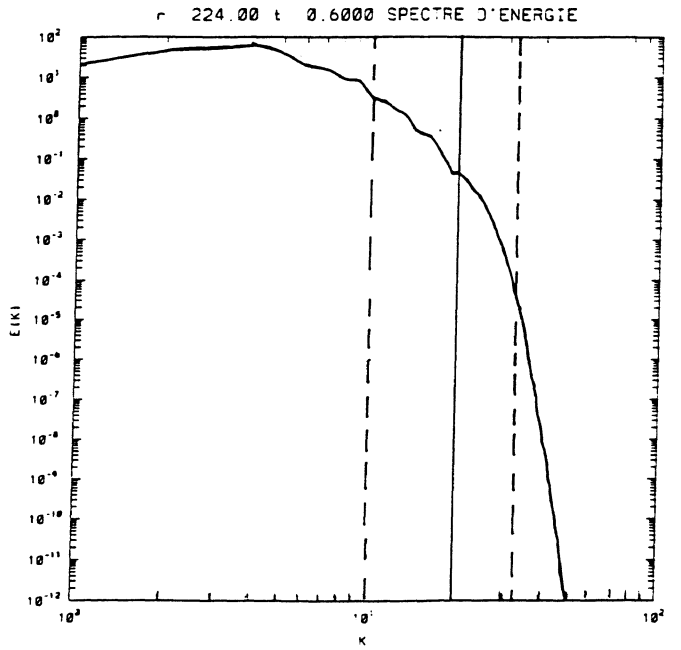


Figure I.14: *Turbulence libre : Spectre d'énergie à* $t = 0.6$ *avec la zone de dissipation (trait continu) et les troncatures* $n_i = 48$ *et* $n_i = 16$ *(trait en pointillé). Il faut multiplier les valeurs numériques par* $\frac{3}{2}$ *pour obtenir* n_i *sur l'axe des abscisses (cf. désaliasing).*

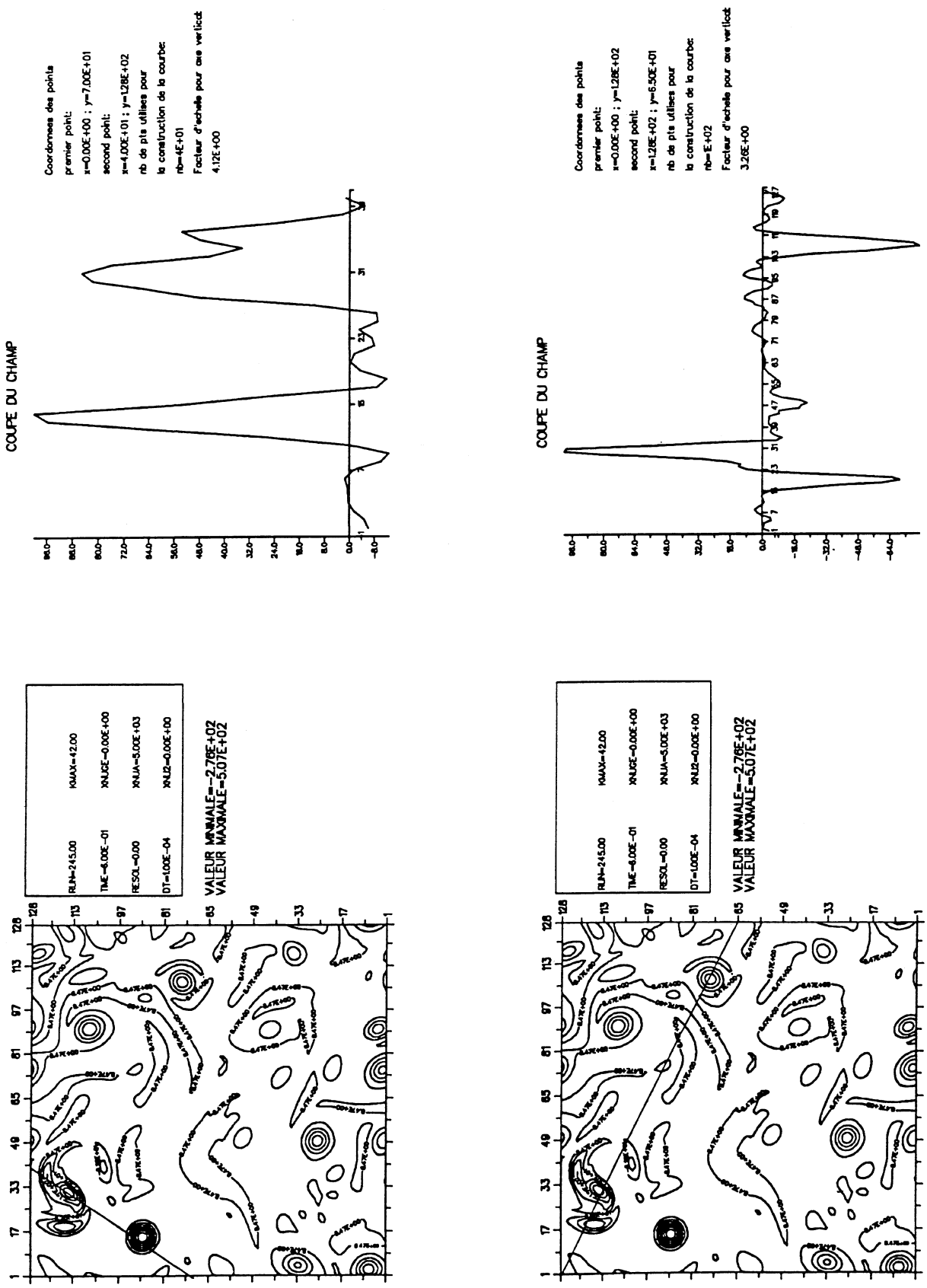


Figure I.15: *Turbulence libre. Galerkin classique ($p = 4$) avec 64^2 modes. Champ et coupes à $t = 0.6$ (programme graphique de T. Philipovitch).*

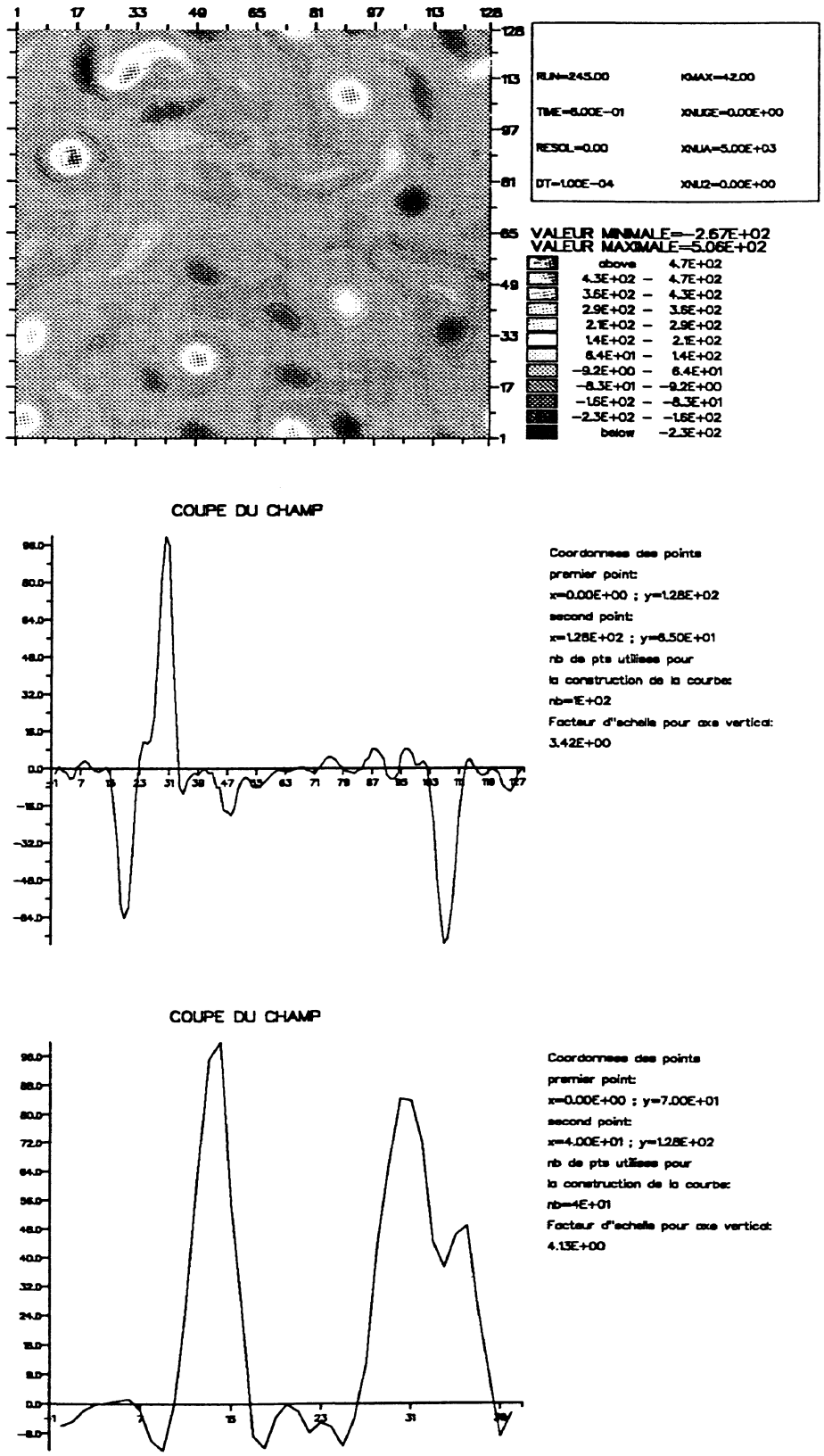


Figure I.16: *Turbulence libre. Galerkin classique ($p = 4$). Projection sur les 48^2 premiers modes : Champ et coupes à $t = 0.6$*

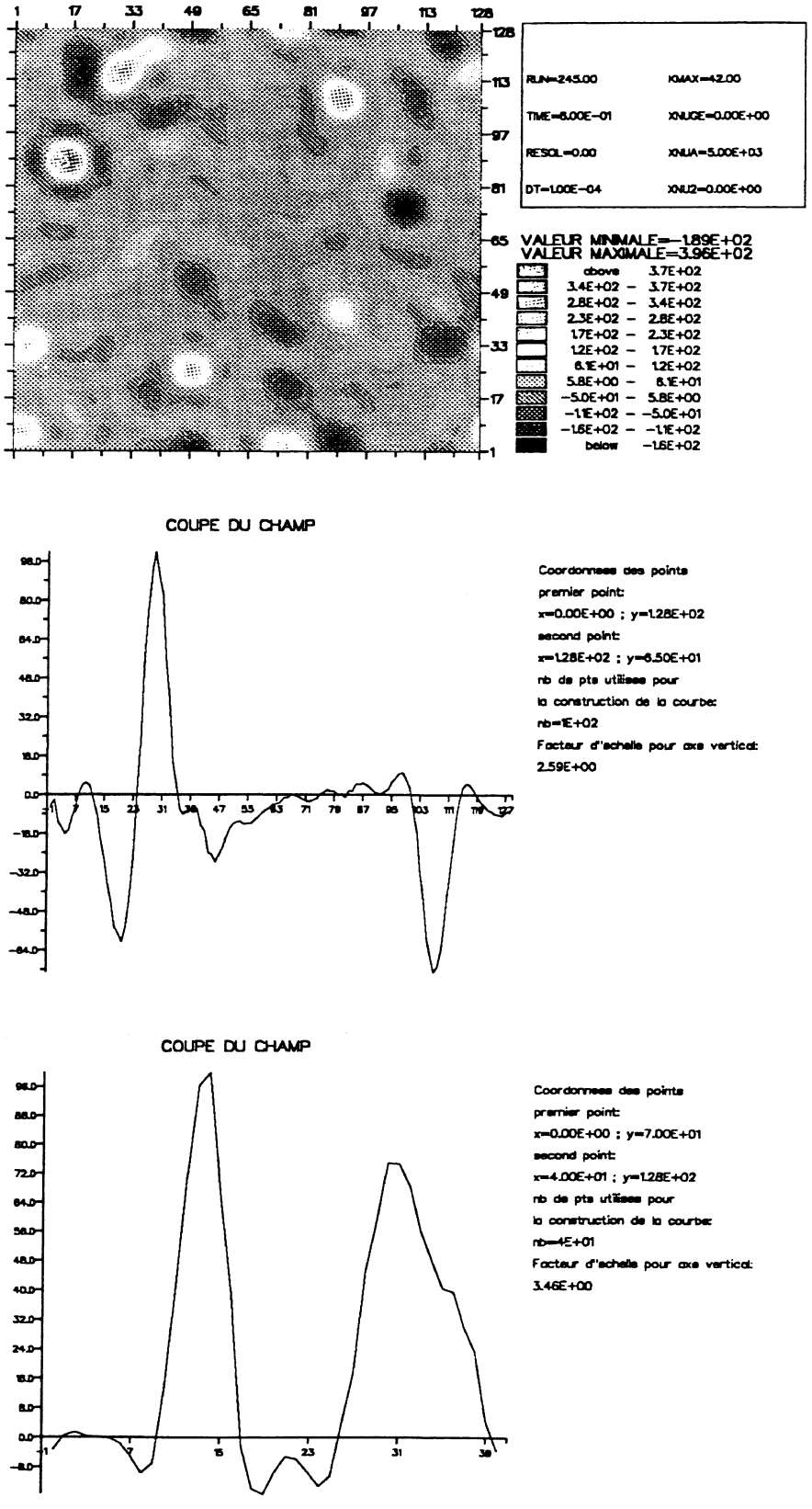


Figure I.17: *Turbulence libre. Galerkin classique ($p = 4$). Projection sur les 16^2 premiers modes : Champ et coupes à $t = 0.6$*

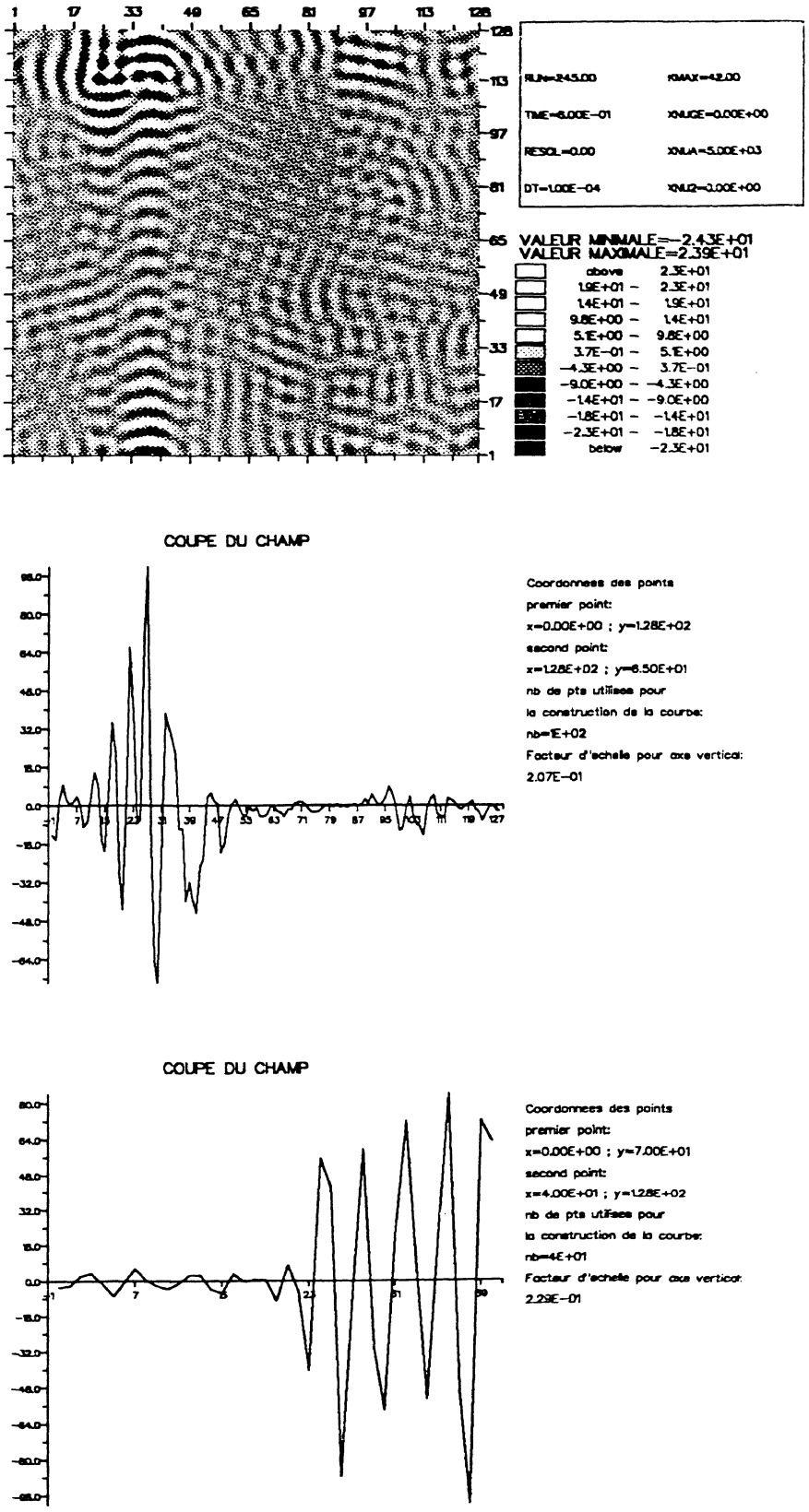


Figure I.18: *Turbulence libre. Galerkin classique ($p = 4$). Projection sur les $64^2 - 48^2$ derniers modes : Champ et coupes à $t = 0.6$*

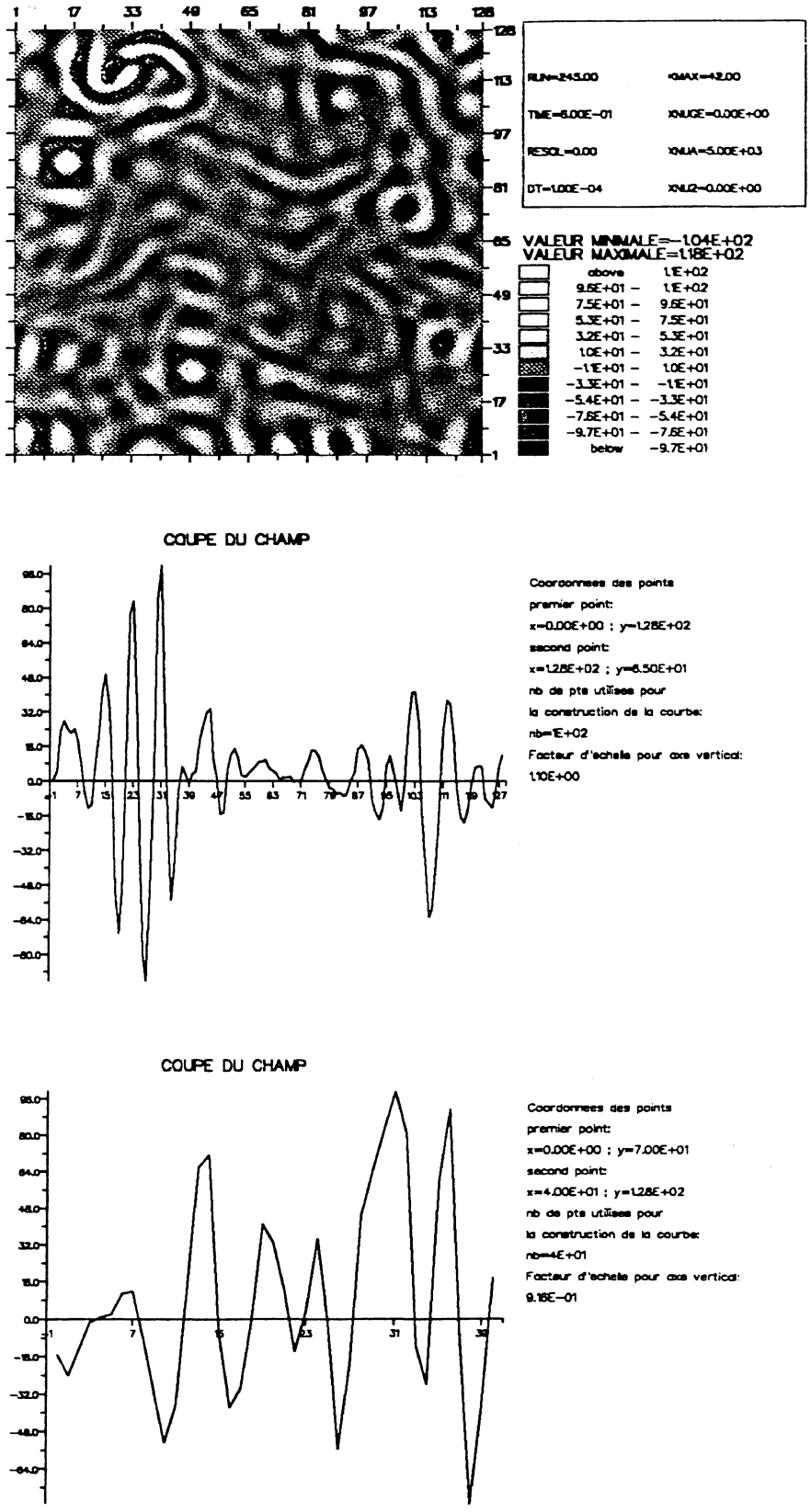
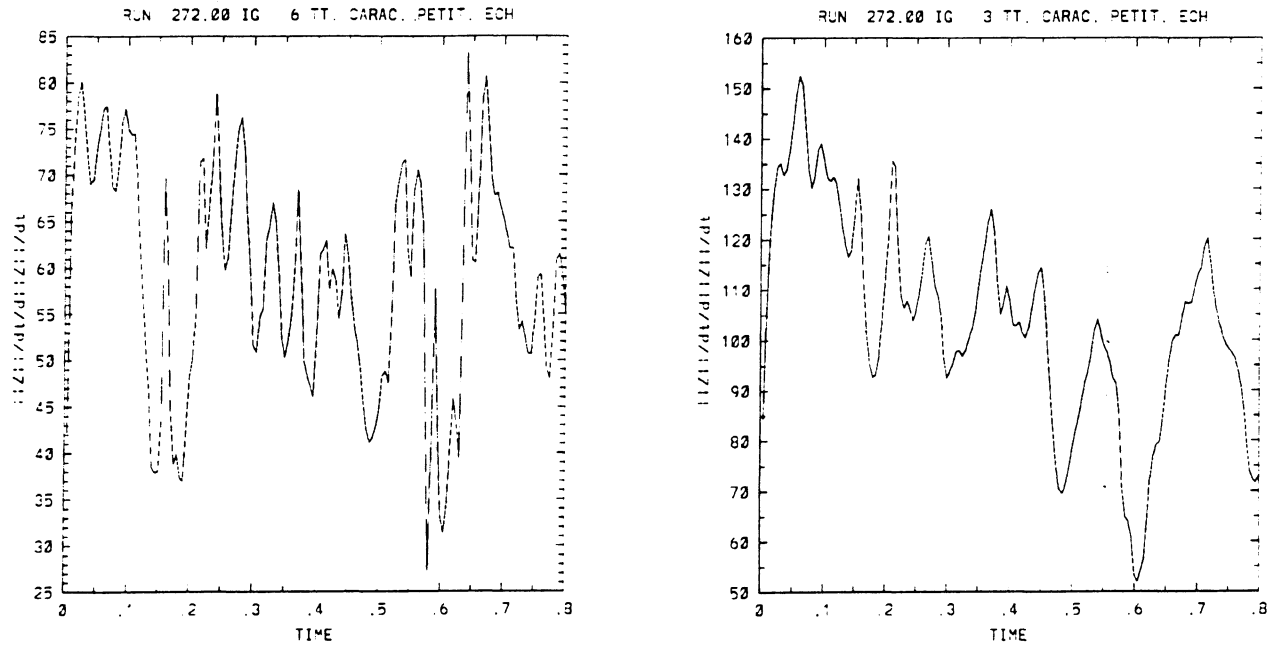
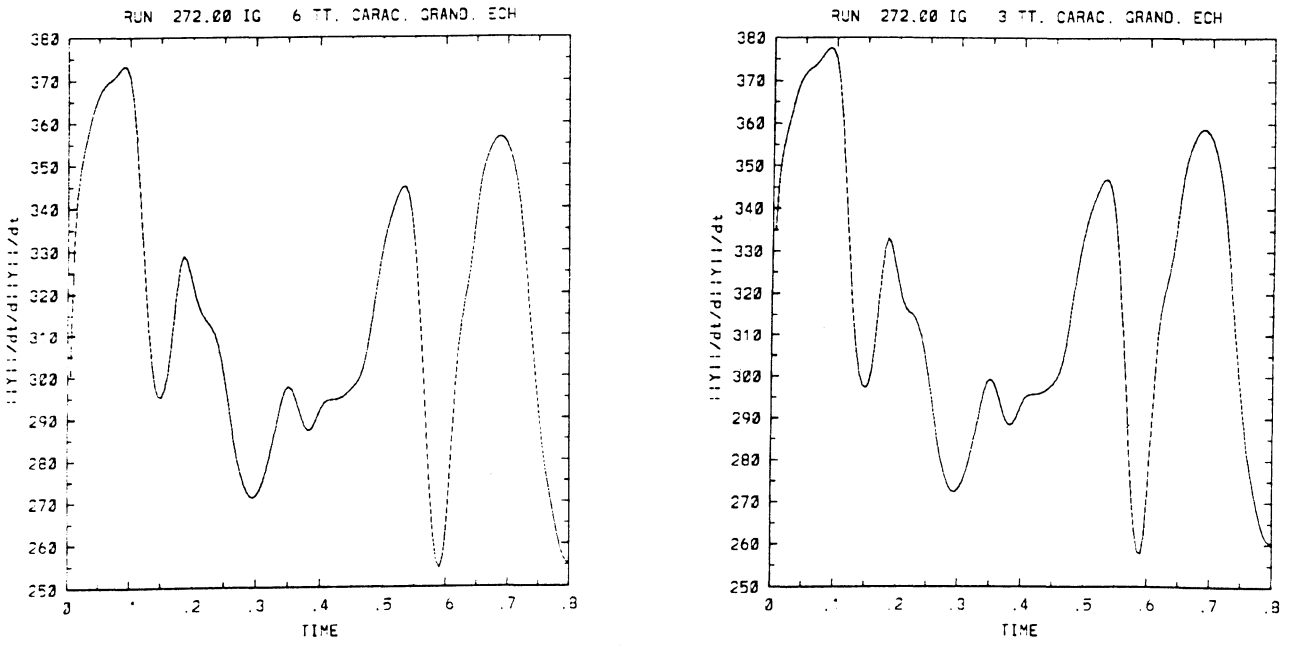


Figure I.19: *Turbulence libre. Galerkin classique ($p = 4$). Projection sur les $64^2 - 16^2$ derniers modes : Champ et coupes à $t = 0.6$*



1)



2)

Figure I.20: *Turbulence libre. Galerkin classique ($p = 4$). 1) Temps caractéristique des petites structures ω_{m-n_i} ; 2) Temps caractéristique des grandes structures ω_{n_i}. A gauche $n_i = 54$ et à droite $n_i = 32$ et $(m = 64)$.*

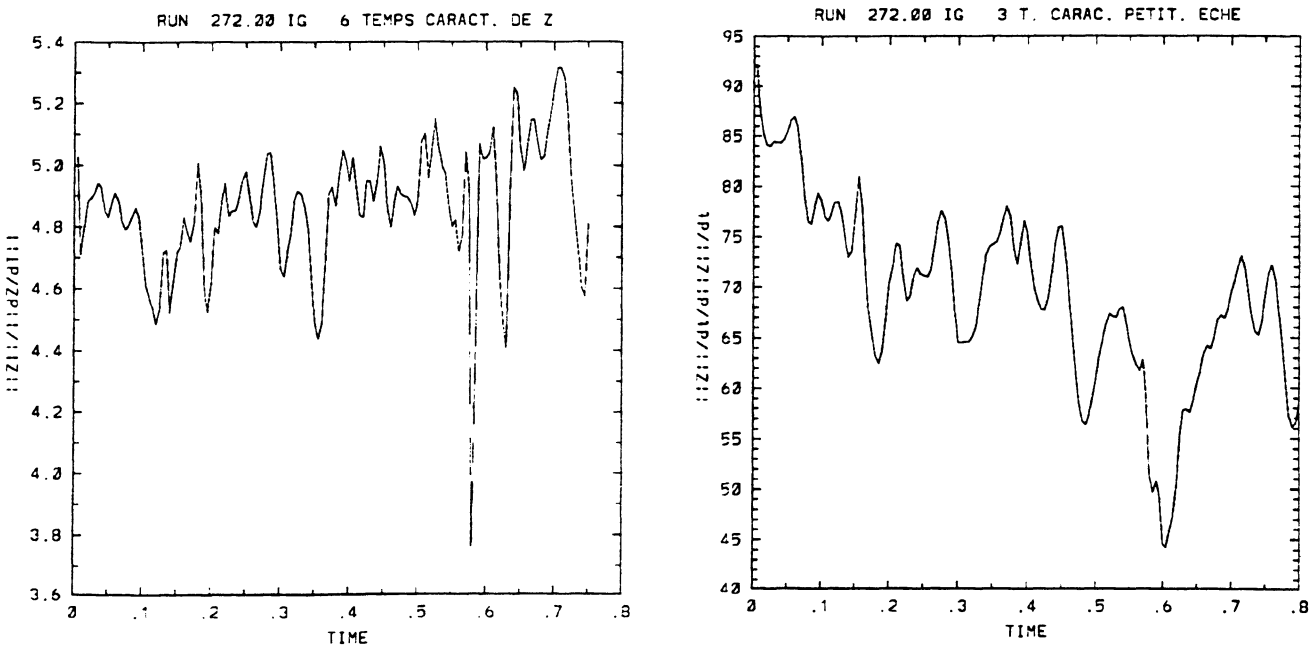
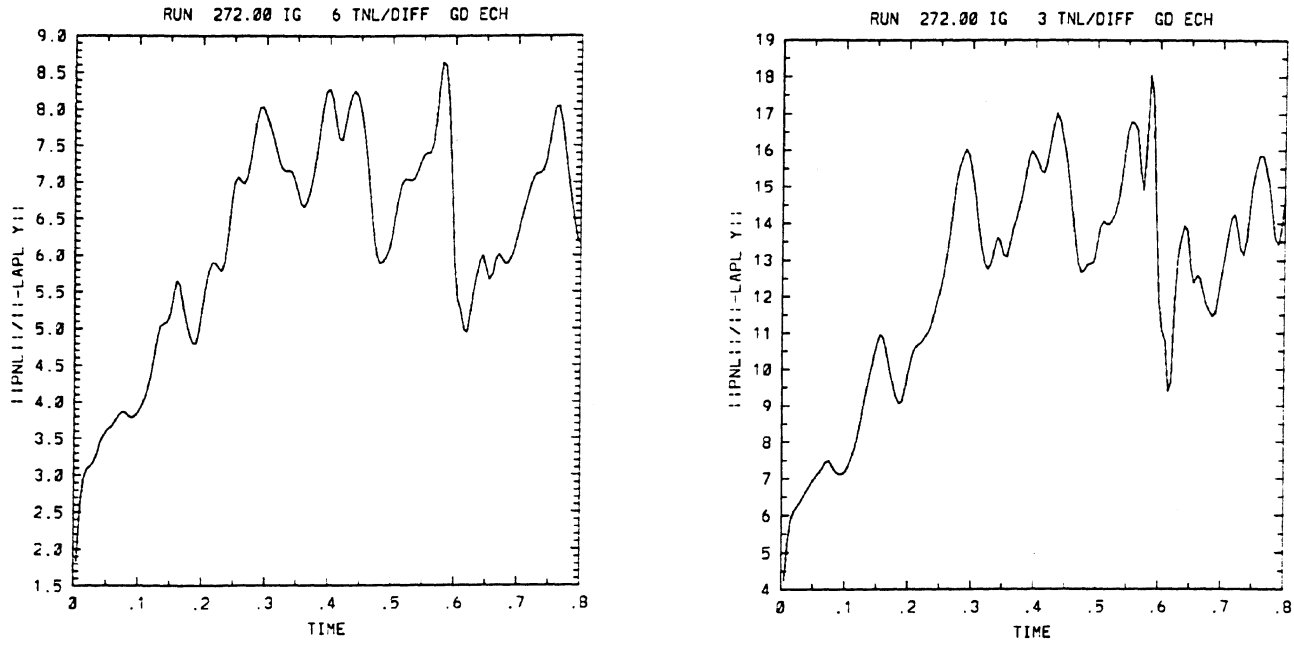
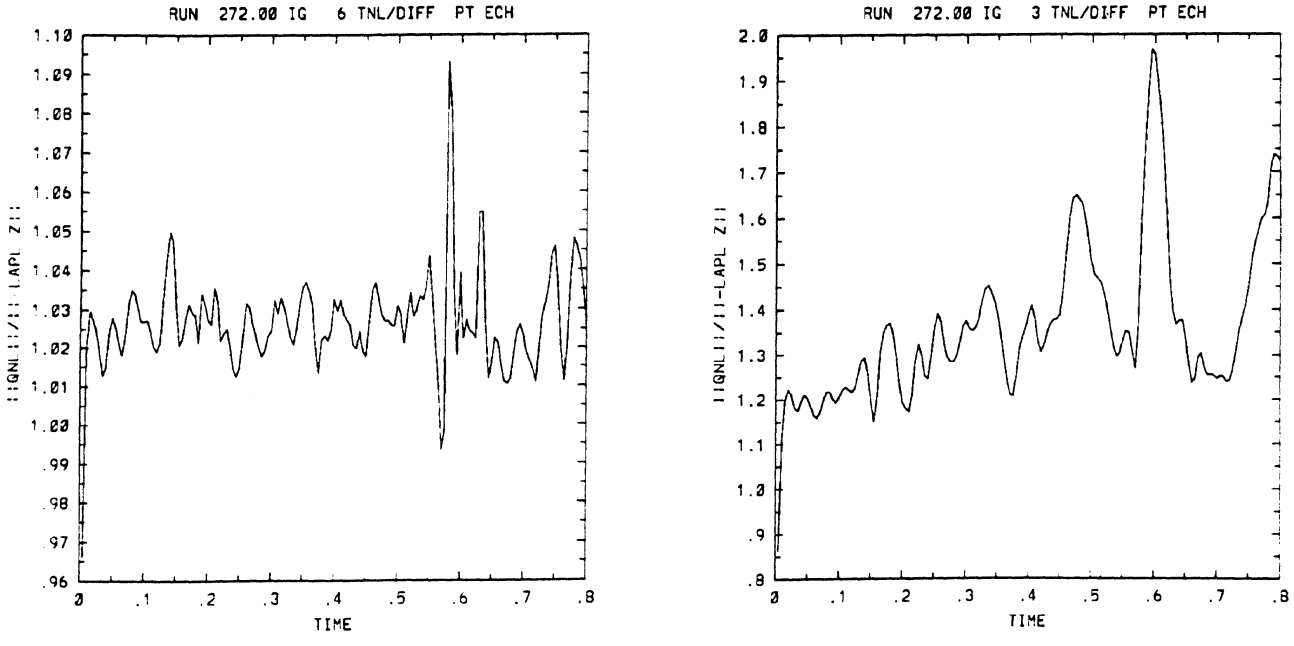


Figure I.21: *Turbulence libre. Galerkin classique ($p = 4$). Temps caractéristique de transfert de ω_{m-n_i} ; A gauche $n_i = 54$ et à droite $n_i = 32$ ($m = 64$).*

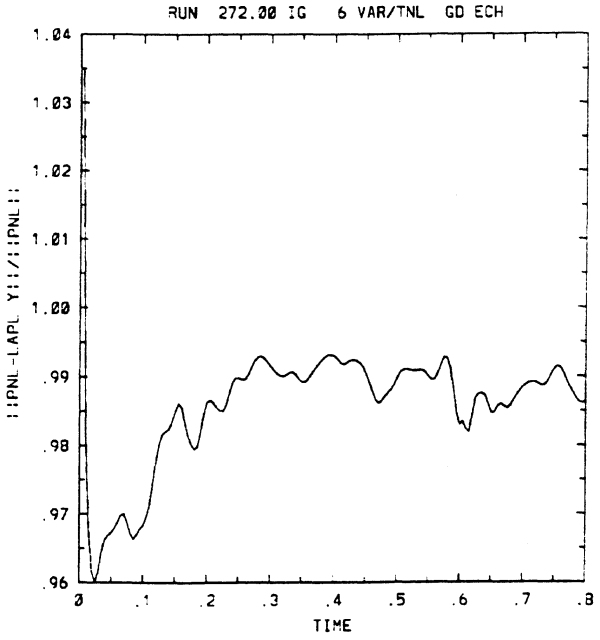


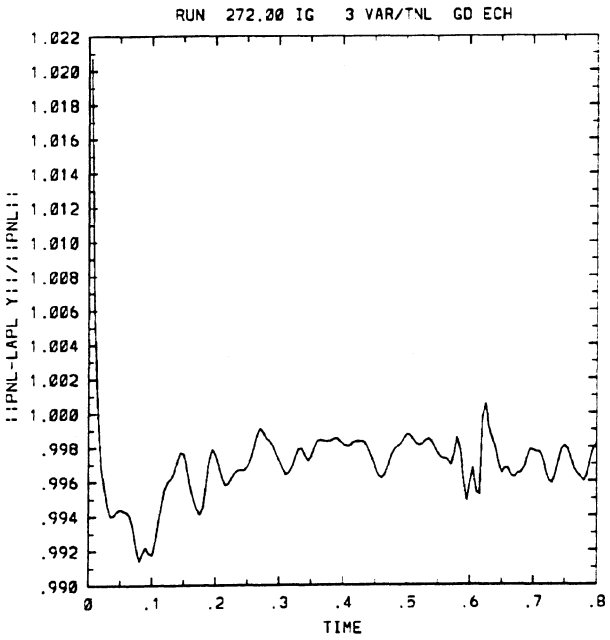
1)



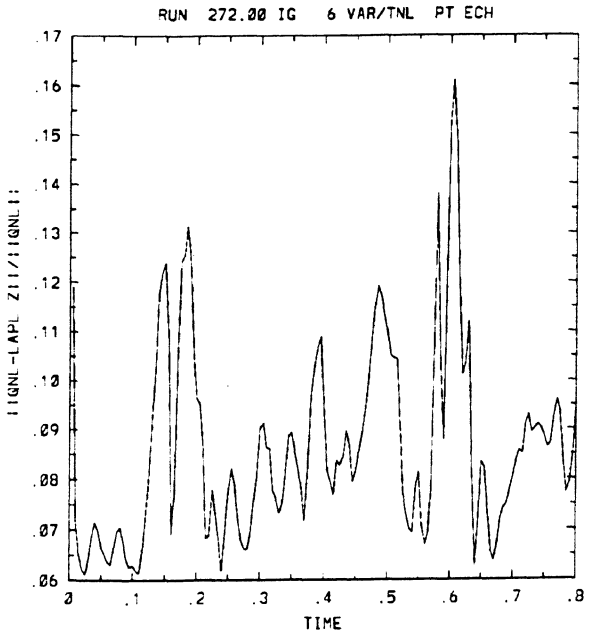
2)

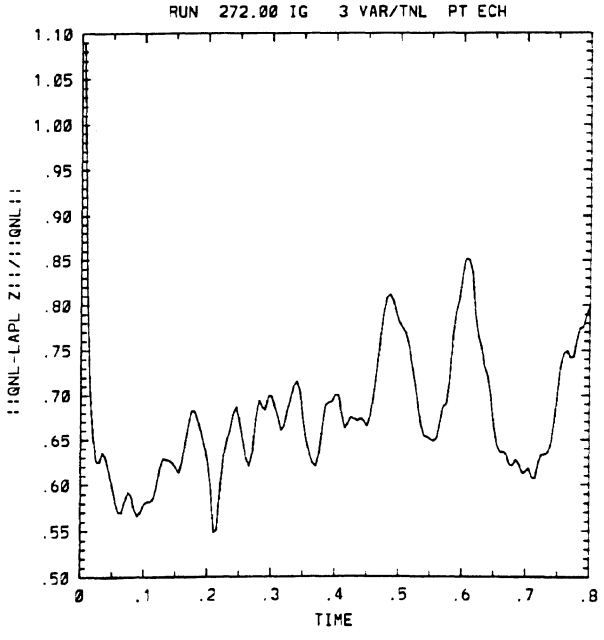
Figure I.22: *Turbulence libre. Galerkin classique* ($p = 4$). *1)* $\frac{||P_{n_i} J(\omega_m, \psi_m)||}{||\nu A \omega_{n_i}||}$; *2)* $\frac{||P_{m-n_i} J(\omega_m, \psi_m)||}{||\nu A \omega_{m-n_i}||}$. *A gauche* $n_i = 54$; *à droite* $n_i = 32$ ($m = 64$).





1)





2)

Figure I.23: *Turbulence libre. Galerkin classique* ($p = 4$). *1)* $\frac{||d\omega_{n_i}/dt||}{||P_{n_i}J(\omega_m,\psi_m)||}$; *2)* $\frac{||d\omega_{m-n_i}/dt||}{||P_{m-n_i}J(\omega_m,\psi_m)||}$. *A gauche* $n_i = 54$; *à droite* $n_i = 32$ ($m = 64$).

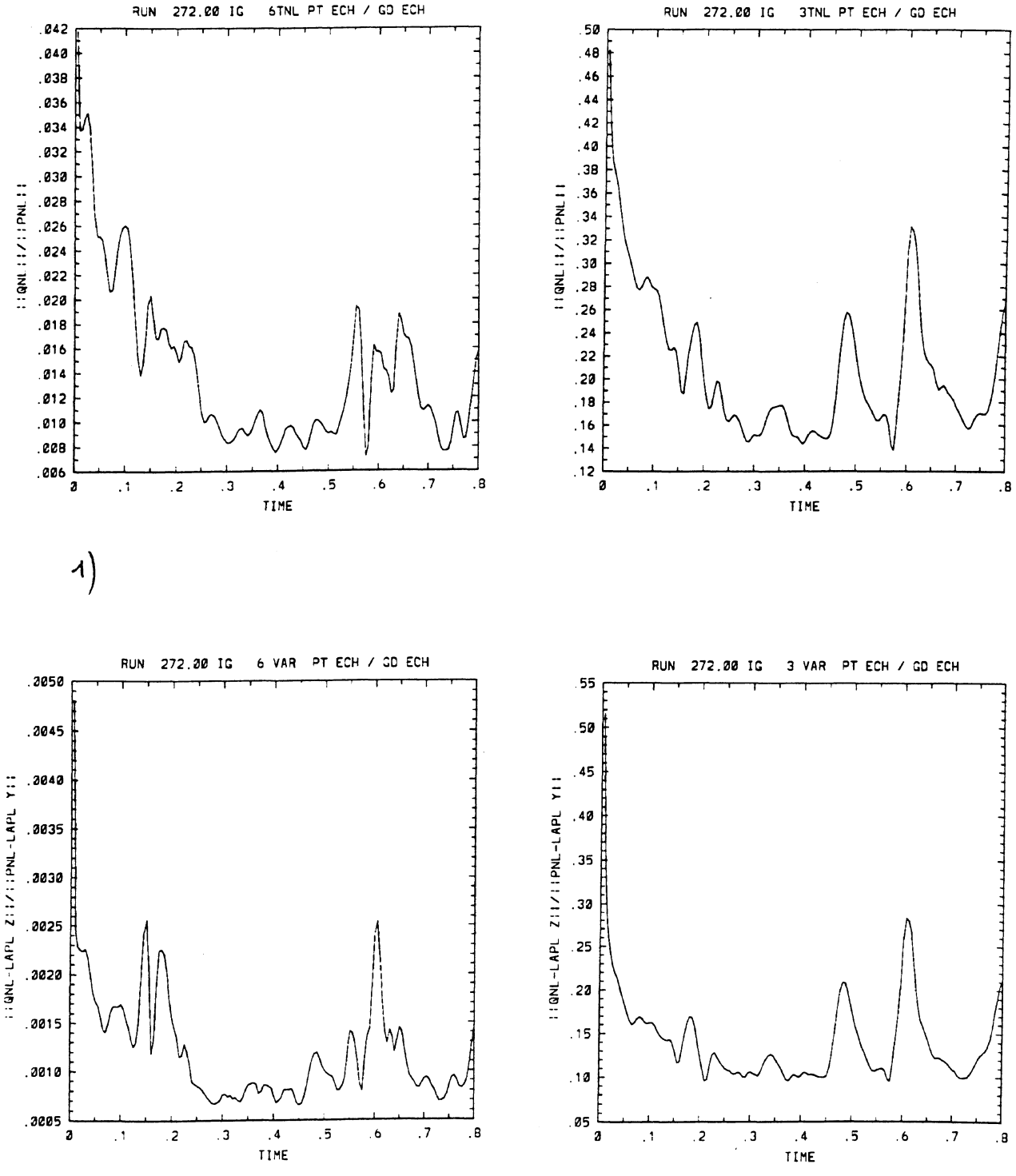


Figure I.24: *Turbulence libre. Galerkin classique* ($p = 4$). *1)* $\frac{||P_{m-n_i} J(\omega_m, \psi_m)||}{||P_{n_i} J(\omega_m, \psi_m)||}$; *2)* $\frac{||d\omega_{m-n_i}/dt||}{||d\omega_{n_i}/dt||}$. *A gauche* $n_i = 54$; *à droite* $n_i = 32$ ($m = 64$).

I.4 Les schémas numériques.

I.4.1 Première version.

La section 3 de l'annexe A donne une description détaillée de la méthode. La stratégie initialement proposée par Jauberteau [11] consiste à définir dynamiquement 3 zones dans l'ensemble des modes excités. Les frontières de ces zones n_{min} et n_{max} sont déterminées par les 2 rapports adimensionnés (9) et (10) de l'annexe A. On a :

1. une zone entièrement incluse dans la zone de dissipation ($n \geq n_{max}$) définissant les petites échelles figées puis relaxées (ces petites échelles et leurs interactions avec les grandes sont en effet négligeables localement en temps mais pas à long terme),

2. une zone tampon ($n_{min} \leq n \leq n_{max}$) (il est en effet difficile de préciser la limite zone de dissipation – cascade d'enstrophie) où les modes (et surtout leurs interactions avec les petits modes) ne sont pas suffisamment petits pour être négligés ou figés trop longtemps,

3. une zone définissant les grandes échelles ($n \leq n_{min}$) calculées à chaque pas de temps comme solution de l'équation (I.30) en tenant compte des modes de la première et deuxième zone dans le terme de couplage.

Les grandes structures w_{n_i} sont solutions de l'équation (I.30) tandis que les petites structures w_{m-n_i} sont figées pendant un temps égal au temps caractéristique des transferts évalué sur le niveau n_{max}

$$\tau = \frac{||\omega_{m-n_{max}}||_{L^2}}{||P_{m-n_{max}} J(\omega_m, \psi_m)||_{L^2}}.$$

et sont relaxées de temps en temps en intégrant :

$$\frac{\partial \omega_m}{\partial t} + \nu A \omega_m + J(\omega_m, \psi_m) = f \tag{I.32}$$

sur l'ensemble des modes pendant un temps suffisamment long.

Les tests effectués dans le cas de la turbulence forcée (cf. Annexe A) ont été repris avec $\theta_1 = 0.075$ et $\theta_2 = 0.085$. Le temps de calcul moyen par itération est égal à 1.4610^{-2} seconde ; ce qui représente un gain de temps de l'ordre de 36% par rapport à Galerkin classique (cf. table 2 de l'annexe A). Comme le montre la figure I.25, l'évolution de l'énergie au cours du temps est rigoureusement identique ainsi que le comportement statistique du champ tourbillonnaire.

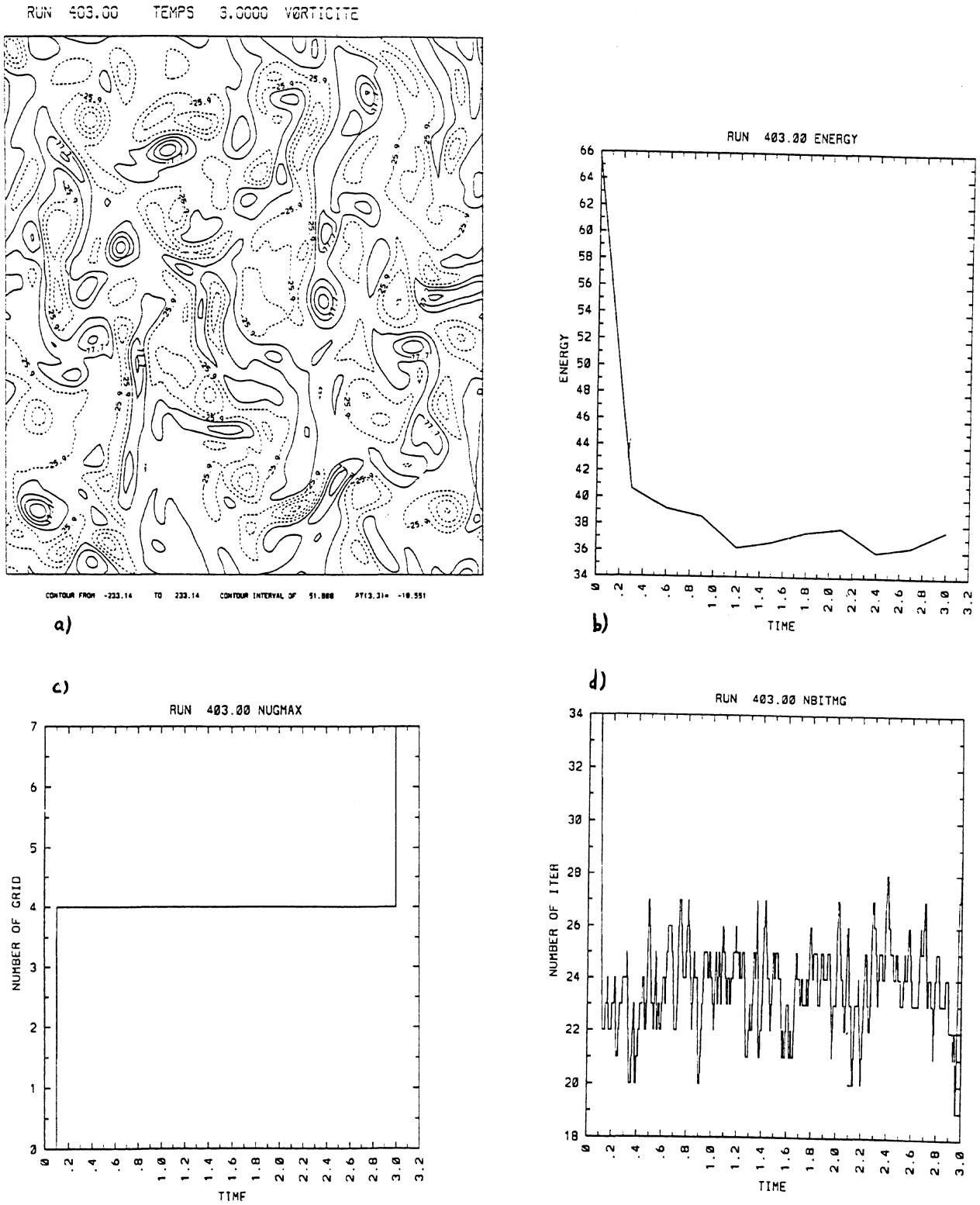


Figure I.25: *Turbulence forcée (GNL 1ere version). a) champ de vorticité à* $t = 3.0$ *b) énergie au cours du temps ; c)* n_{max} *au cours du temps ; d) évolution du temps caractéristique.*

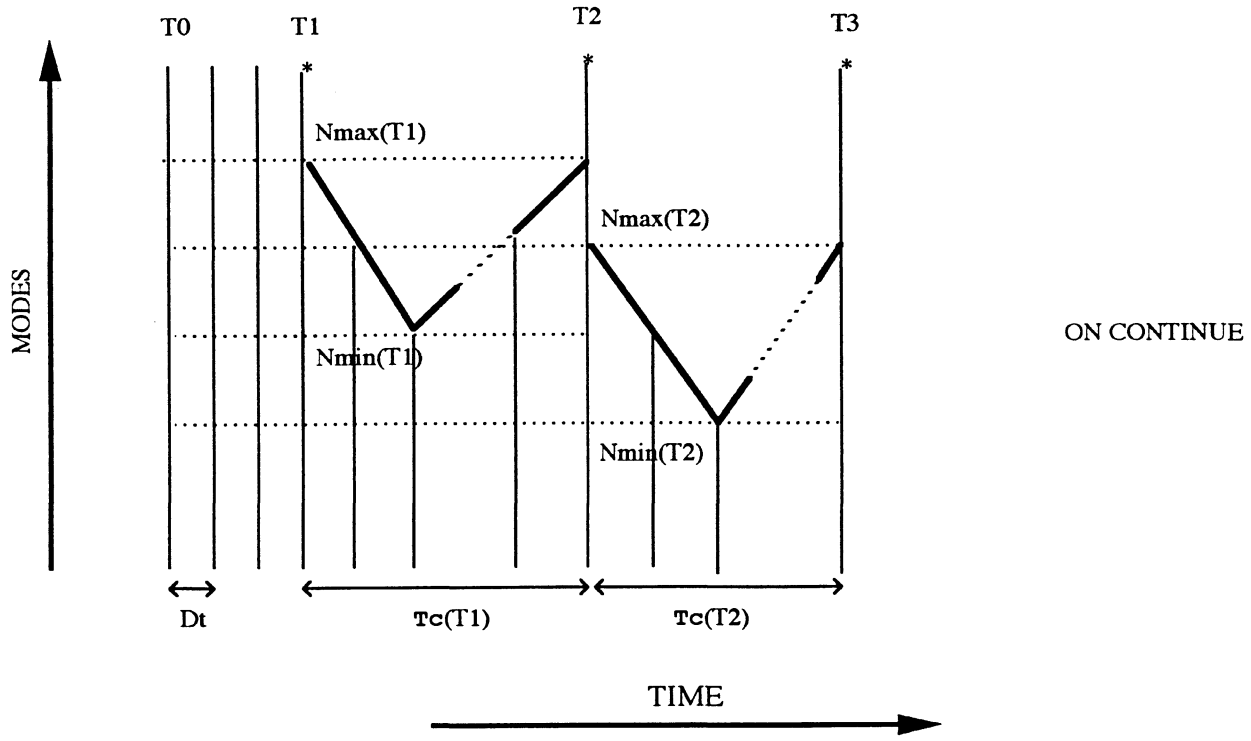


Figure I.26: *Organigramme de GNL2.* $\star$ *symbolise la détermination de* $T_c(t_i)$, $N_{max}(t_i)$, $N_{min}(t_i)$. *Les traits verticaux représentent les modes intégrés à l'instant considéré.*

I.4.2 Deuxième version.

En tenant compte des résultats théoriques et numériques, une variante de ce schéma consiste à remplacer l'étape de relaxation par la résolution de l'équation approchée de la variété inertielle approximative $\mathcal{M}_1$ ie par calculer $\omega_{m-n_{max}}$ solution de :

$$\nu A\omega_{m-n_{max}} + P_{m-n_{max}} J(\omega_{n_{max}}, \psi_{n_{max}}) = P_{m-n_{max}} f \tag{I.33}$$

Pratiquement le terme non linéaire est légèrement différent. Si la fonction de courant $\psi^k_{n_{max}}$ et la vorticité $\omega^k_{n_{max}}$ associées aux grandes échelles sont connues à l'instant $k\Delta t$, alors la vorticité $\omega^k_{m-n_{max}}$ associées aux petites échelles à l'instant $k\Delta t$, est solution de :

$$\nu A\omega^k_{m-n_{max}} + P_{m-n_{max}} J(\omega^k_{n_{max}} + \omega^{k-1}_{m-n_{max}}, \psi^k_{n_{max}} + \psi^{k-1}_{m-n_{max}}) = P_{m-n_{max}} f^k \tag{I.34}$$

La figure I.26 donne un organigramme de cet algorithme. Aux temps t_1, t_2,..., t_i,... les niveaux $n_{max}(t_i)$, $n_{min}(t_i)$ et le temps $\tilde{\tau}_{m-n_{max}}(t_i)$ sont déterminés après réactualisation de l'ensemble des modes. Il est à noter que généralement t_1 diffère de l'instant initial t_0 car pendant le temps de transition t_c (cf I.3.2 : les différents termes du système) les approximations effectuées en (I.33) ne sont pas justifiées.

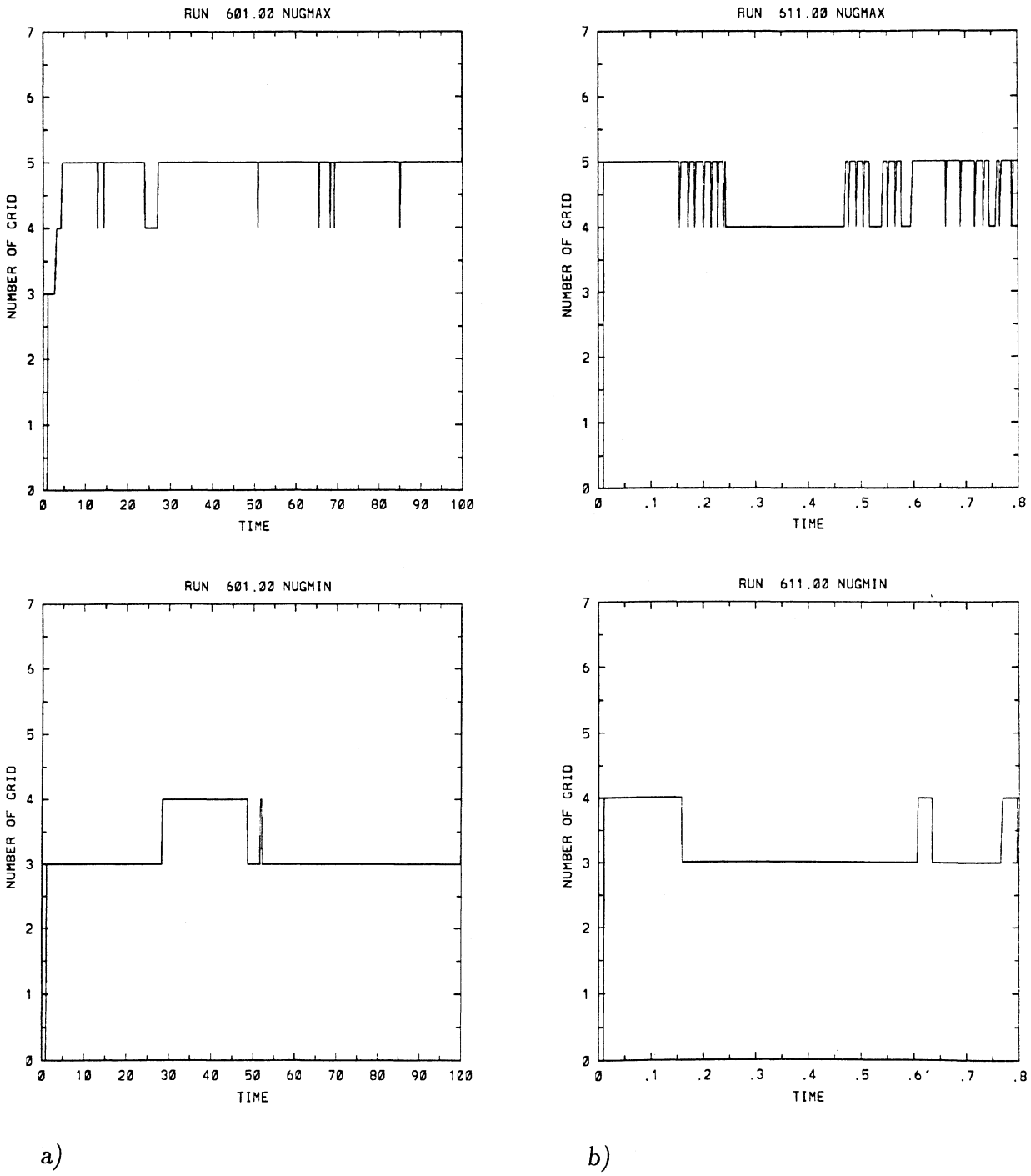
	p	Δt	θ_1	θ_2	Temps cpu GC par itération	Temps cpu GNL2 par itération	gain %
2 Vortex	4	0.01	0.01	0.02	$2.43\ 10^{-2}$	$1.99\ 10^{-2}$	18
Turbulence libre	4	0.0001	0.05	0.05	$2.40\ 10^{-2}$	$1.88\ 10^{-2}$	22
Turbulence forcée	4	0.0001	0.05	0.05	$2.29\ 10^{-2}$	$2.00\ 10^{-2}$	13

Tableau I.3: *Données et résultats expérimentaux.*
GC = Galerkin classique.
GNL2 = Galerkin non linéaire 2e version.

Le tableau (I.3) présente les performances en temps de calcul de cette version. Les temps caractéristiques et les rapports qui déterminent les niveaux sont identiques à ceux prévus par la méthode de Galerkin classique et les résultats numériques de cette version sont très proche de ceux obtenus par GC. Par rapport à la première version le temps d'exécution est légèrement supérieur car s'agissant d'une moyenne sur l'ensemble des itérations, la détermination coûteuse des niveaux est plus fréquente que dans la première version et les niveaux sont plus élevés comme le montre la figure I.27 en comparaison avec les figures (8) et (11) de l'annexe A.

Il est alors envisageable de choisir des critères de détermination des différents espaces de projection moins sévères et obtenir ainsi de meilleurs taux de réduction de temps CPU.

Par exemple dans le cas de la turbulence forcée avec $\theta_1 = 0.075$ et $\theta_2 = 0.085$, le gain de temps est de l'ordre de 43% ie un coût de 1.310^{-2} seconde par itération. La figure I.28 montre que n_{max} est légèrement supérieur à celui obtenu avec la première version. Les temps caractéristiques de transfert sont identiques à ceux prévus par Galerkin classique. Enfin les spectres d'énergie à $t = 3.0$ diffèrent aux modes 1 et 2 et sont rigoureusement identiques au-delà de k_I.



a) b)

Figure I.27: n_{max}, n_{min} *au cours du temps (GNL 2eme version) ; a) interaction de 2 vortex ; b) champ turbulent en régime libre* ($p = 4$).

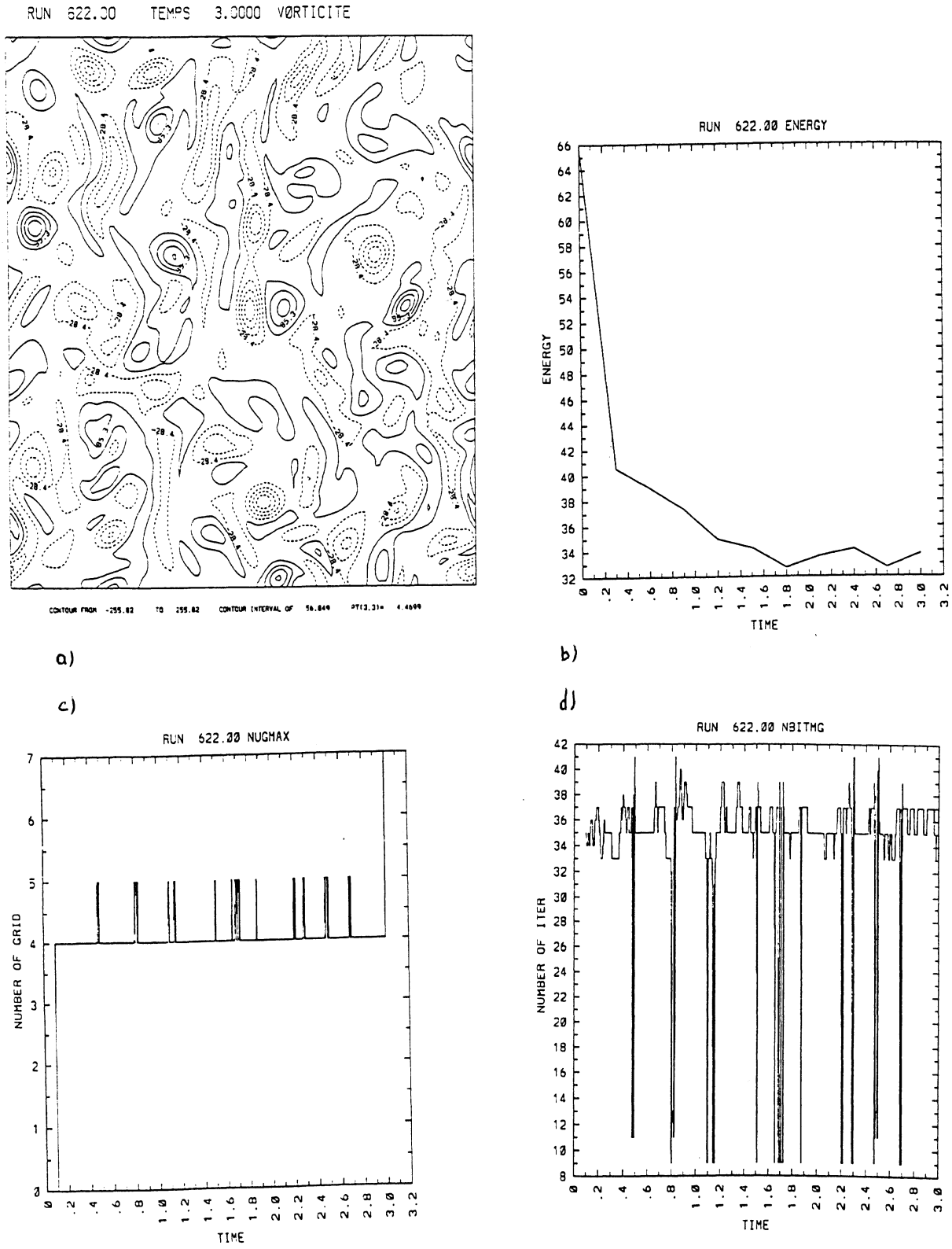


Figure I.28: *Turbulence forcée (GNL 2eme version); a) champ de vorticité à t = 3.0 ; b) énergie au cours du temps ; c) n_{max}, n_{min} au cours du temps ; d) évolution du temps caractéristique.*

I.4.3 Conclusion.

Si l'on se limite à une faible résolution, la méthode de Galerkin non linéaire (2ieme version) associée à un laplacien itéré modélisant les effets des échelles virtuelles sur les grandes échelles restitue une solution identique à celle issue de la méthode de Galerkin classique et apporte un gain de temps certes modéré (tout particulièrement dans le cas où $p \geq 8$ c'est-à-dire dans le cas où la zone de dissipation est extrêmement réduite). Dans le cas d'une résolution élévée, si l'on accepte qu'une partie du spectre soit affectée par la dissipation alors la méthode galerkin non linéaire combinée avec un laplacien itéré peut s'avérer très efficace en coût et permettre ainsi une résolution encore plus fine.

Lors des simulations d'écoulements turbulents, 3 cas sont possibles :

- dans le premier cas, une zone de dissipation est entièrement simulée (soit il s'agit de simulation directe et de faibles nombres de Reynolds, soit il s'agit d'un modèle turbulent simulant une zone de dissipation), la méthode GNL est alors applicable avec une efficacité fonction de la largeur de la zone de dissipation,
- dans le deuxième cas, la troncature se situe dans la zone inertielle loin de l'échelle de dissipation, GNL est alors inapplicable et un modèle turbulent doit être implémenté pour simuler les transferts,
- dans le dernier cas, la troncature se situe dans la zone inertielle mais proche de la zone de dissipation, il est alors envisageable de simuler cette derniere zone avec la méthode de Galerkin non linéaire pour un coût supplémentaire raisonnable.

Parmi les études à mener par la suite, figurent :

- l'amélioration des critères des niveaux ; il serait souhaitable de mettre au point des critères plus économiques,
- l'étude de la prédicibilité ou temps de bifurcation des solutions ; étude de l'influence de la méthode GNL sur cette notion,
- la recherche de la meilleure troncature possible et l'introduction d'autres bases telles que les ondelettes car les structures dynamiquement actives ne sont pas uniformément distribuées si l'on tient compte de l'intermittence spatio-temporelle.

Bibliographie

[I.1] C. Basdevant, *Le modèle de simulation numérique de turbulence bidimensionnelle du L.M.D.*, Note interne LMD 114, 1982.

[I.2] C. Basdevant and R. Sadourny, *Modélisation des échelles virtuelles dans la simulation numérique des écoulements turbulents bidimensionnels* J. de Mécanique théorique et appliquée, 1983, 243-269.

[I.3] C. Basdevant and R. Sadourny, *Ergodic properties of inviscid truncated models of 2 dimensioned incompressible flows*, J. of Fluid Mechanics, 69, 673-688, 1975.

[I.4] P. Constantin, C. Foias et R. Temam, *On the dimension of the attractors in two-dimensional turbulence*, Physica D 30, 284-296, 1988.

[I.5] P. Constantin et C. Foias, *Global Lyapunov exponents, Kaplan–Yorke formulas and dimension of attractors for two-dimensional Navier–Stokes equations* Comm. Pure App. Math. 38, 1-27, 1985.

[I.6] T. Dubois, F. Jauberteau and R. Temam, *Solution of the incompressible Navier-Stokes equations by the nonlinear Galerkin method*, to appear in J. of Comp. Phys.

[I.7] C. Foias et R. Temam, *Structure of the set of stationary solutions of the Navier–Stokes equations*, Comm. Pure Appl. Math. 30, 149-164, 1977.

[I.8] C. Foias et R. Temam, *Some analytic and geometric properties of the solutions of the Navier–Stokes equations*, J. Math. Pure Appl. 58, 339-368, 1979.

[I.9] C. Foias, G. Sell et R. Temam, *Variétés inertielles des équations dififferentielles dissipatives* C.R.A.S. I, 301, 139-142, 1985.

[I.10] C. Foias, O. Manley and R. Temam, *On the interaction of small and large eddies in two–dimensional turbulent flows*, Math. Mod. and Num. Anal. (M2AN), 22, 1988, 93-114.

[I.11] F. Jauberteau, *Résolution numérique des équations de Navier–Stokes instationnaires par méthodes spectrales. Méthode de Galerkin nonlinéaire*, Thèse Université de Paris–Sud, 1990.

[I.12] F. Jauberteau, C. Rosier and R. Temam, *The nonlinear Galerkin method in computational fluid dynamics*, App. Num. Math. 6, 1989-90, 361-370.

[I.13] F. Jauberteau, C. Rosier and R. Temam, *: A Nonlinear Galerkin Method for Navier Stokes Equations*, Comp. Meth. in App. Mec. and Eng. 80, 1990 , 245-260.

[I.14] P. Le Roy, *Cascade inverse et dispersion turbulente en turbulence bidimensionnelle*, Thèse de Doctorat de l'Ecole Nationale des Ponts et Chaussées, 1988.

[I.15] Lesieur, *Turbulence in fluids. Stochastic and numerical modelling*, Klurver Academic Publishers, 2nd revised edition.

[I.16] M. Marion and R. Temam, *Nonlinear Galerkin Methods*, SIAM J. Num. Anal., 26, 1989, 1139-1157.

[I.17] S.A. Orszag, *Numerical simulation of incompressible flow within simple boundaries in Galerkin spectral representation*, Studies in applied mathematics, 50, 1971, 239-327.

[I.18] K. Promislov, *Time analyticity and gevrey regularity for solutions of a class of dissipative partial differential equations*, Nonlinear Anal., Th., Met. et App.,Vol. 16, No 11, pp 959-980, 1991.

[I.19] D. Ruelle, *Chaotic evolution and strange attractors*, Academia nazionale dei Lincey, Cambridge University Press 89.

[I.20] R. Sadourny and C. Basdevant, *Une classe d'opérateurs adaptés à la modélisation de la diffusion turbulente en dimension deux*, C.R.A.S.,Série II, t. 292 (27 avril 1981).

[I.21] J.Shen, *Hopf bifurcation of the unsteady regularized driven cavity flow*, J. of Comp. Phys., 95, 228-245, 1991.

[I.22] J.Shen, *A nonlinear Galerkin method for the Navier–Stokes equations*, Comp. Meth. in Appl. Mech. and Eng., 80, 245-260, 1990.

[I.23] C. Staquet, *Etude numérique de la turbulence bidimensionnelle homogéne et cisaillée*, Thèse de Doctorat de l'Institut National Polytechnique de Grenoble,1985.

[I.24] R. Temam, *Navier–Stokes equations*, Amsterdam, North Holland, 1984.

[I.25] R. Temam, *Variétés inertielles approximatives pour les équations de Navier–Stokes bidimensionnelles*, C.R.A.S., Série II, t. 306, 1988, 399-402.

[I.26] R. Temam, *Inertial Manifolds* The mathematical intelligencer vol 12, no 4, 1990.

[I.27] R. Temam, *Approximation of attractors, large eddies simulation and multi–scale methods*, Proc. of the Royal Society A, special issue commemorating the work of A.N. Kolmogorov, 1991.

[I.28] E. S. Titi, *On approximate inertial manifolds to the Navier–Stokes equations*, J. of Math. Anal. and Appl., Vol. 149, No 2, 540-557, 1990.

I.5 Annexe A

NONLINEAR GALERKIN METHOD AND SUBGRID SCALE MODEL FOR 2D TURBULENT FLOWS

Frédéric PASCAL, Claude BASDEVANT
Laboratoire de Météorologie Dynamique
24 Rue Lhomond
75231 PARIS CEDEX, FRANCE

Abstract

The Nonlinear Galerkin method, derived by Marion and Temam from results in dynamical systems theory, is investigated in the framework of numerical simulation of 2-dimensional incompressible turbulent flows.

We use the Nonlinear Galerkin method together with a hyperviscosity subgrid scale parametrization. A first part states the theoretical background. In the second part we define the scheme and derive some technical improvements for the method. The last part reports numerical experiments conducted for freely decaying as well as forced turbulence. Our main conclusion is that the performance of the Nonlinear Galerkin method can be important only if the dissipation range is large.

Support grant

This work was supported by Contrat D.R.E.T. 90/1551/A000 : *Développement d'algorithmes de simulation numérique des équations de Navier–Stokes.*

To appear in *Theoretical and Computational Fluid Dynamics* .

1 Introduction.

From recent developments in dynamical systems theory, the nonlinear Galerkin method has been proposed by Marion and Temam [7] as a new numerical scheme for numerical simulation of Navier–Stokes equations. This algorithm currently developed in the framework of Fourier spectral method for the periodic case (Jauberteau *et al.* [5]), seems appropriate for long time integration of Navier-Stokes equations, numerical simulation of turbulent flows being performed at a small fraction of the computational effort usually required by traditional methods.

The first computational tests of this new method was conducted by Jauberteau [4], Jauberteau *et al.* [6] and by Dubois *et al.* [2], in cases where the exact stationary solution of the equations were known or discretization in space and time was choosen sufficiently small so that the unique deterministic flow associated to given initial values could be calculated. However tests with exact solutions do not take into account the main difficulties encountered in numerical simulations of turbulent flows of practical interest. For instance in Jauberteau [4] and Jauberteau *et al.* [6] the shape of the energy spectrum is constant with time. And present power and memory size of computers do not allow to handle all scales from large scales down to small ones in the dissipation range as soon as the Reynolds number is moderatly large. In practice for high Reynolds numbers and turbulence experiments, the numerical cutoff lies in the inertial range. The simulation is no longer a direct numerical simulation but a so called large eddy simulation ; a subgrid model has to be incorporated to avoid the drift of the solution toward statistical equilibrium (Basdevant and Sadourny [1]).

In this paper, we study the nonlinear Galerkin method in the framework of numerical simulation of two-dimensional turbulent flows, when the cutoff wavenumber lies inside the enstrophy inertial range. We will show that provided some modifications are applied to the original method, and provided a subgrid scale parametrization is included in the numerical model, the nonlinear Galerkin algorithm is well adapted for large scale simulation of turbulent flows at large Reynolds numbers.

In a first part we recall briefly the theoretical background and the main ideas of the schemes given in references [2], [4], and [6]. Then in a second part, we describe the practical algorithms and the improvements that seem to us necessary : particulary in the criterion for selection of grid levels and in the choice of integration time on these levels. In the last section, we present some numerical experiments performed on the Cray 2 of the Centre de Calcul Vectoriel pour la Recherche in Palaiseau, France. A first series of experiments deals with decaying turbulence (vortex pair merging or decaying homogenous turbulence). A second series studies the case of forced turbulence until stationary regime is reached.

The method applied on a 128 square grid together with an iterated laplacian $-(-\Delta)^4$, leads to a relative gain of computing time of approximatively 25% compared to the classical Galerkin method with the same subgrid model.

2 The nonlinear Galerkin method.

Let $\Omega = (0, L_1) \times (0, L_2)$ be a bounded domain of $\mathcal{R}^2$. The Navier–Stokes equations of periodic two–dimensional incompressible viscous flows in Ω are given by :

$$\begin{aligned} \frac{\partial u}{\partial t} - \nu \Delta u + (u.\nabla)u + \nabla p = f \quad & \text{in } Q = \Omega \times \mathcal{R}^+ \\ \nabla . u = 0 \quad & \text{in } Q \\ u(x,0) = u_0(x) \quad & \text{in } \Omega \end{aligned} \tag{1}$$

with periodic boundary conditions in both directions.

Let denote by A the basic dissipative operator and H an appropriate divergence free Hilbert space. The equations can be written as :

$$\frac{du}{dt} - \nu A u + B(u) = f \text{ in } H \tag{2}$$

plus boundary conditions and initial conditions.

There is a stream function ψ such that

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -\frac{\partial \psi}{\partial x_2} \\ +\frac{\partial \psi}{\partial x_1} \end{pmatrix}$$

and if $\omega = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} = \Delta \psi$ refers to vorticity and $J(\psi, \omega) = \frac{\partial \psi}{\partial x_1}\frac{\partial \omega}{\partial x_2} - \frac{\partial \psi}{\partial x_2}\frac{\partial \omega}{\partial x_1}$ to the jacobian operator, system (1) becomes :

$$\begin{aligned} \frac{\partial \omega}{\partial t} + J(\psi, \omega) = \nu \Delta \omega + \text{curl} f \\ \omega = \Delta \psi \\ \omega(x,0) = \omega_0(x) \end{aligned} \tag{3}$$

Let $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots$ denote the eigenvalues of the Stokes operator and $w_1, w_2, \ldots$ the corresponding eigenvectors. In the periodic case, these functions are combinations of cosines and sines functions, and we can write the Fourier series expansion of the vorticity, stream function and velocity ($L_1 = L_2 = 2\pi$) :

$$u(x,t) = \sum_{j \in \mathcal{Z}^2} \hat{u}_j(t) w_j(x) = \sum_{j \in \mathcal{Z}^2} \hat{u}_j(t) e^{ij.x}$$

The classical Galerkin method produces approximate orbits of the system lying on the linear space spanned by the first m eigenvectors $w_1, \ldots, w_m$. New schemes such as the nonlinear Galerkin methods proposed in [6] can be considered as algorithms for the computational estimation of approximate inertial manifolds : i.e. smooth finite dimensional manifolds which attract all orbits at an exponential rate into a thin neighborhood (see [3]). These schemes consist of introducing the following approximations u_m and u_n of the velocity u ($n < m$) :

$$u_m = u_n + z_{m-n}$$

$$\text{where} \quad u_n = \sum_{|j| \leq n} \hat{u}_j w_j \quad \text{and} \quad z_{m-n} = \sum_{n+1 \leq |j| \leq m} \hat{u}_j w_j$$

u_n corresponds to large wavelengths, and z_{m-n} appears as a correction of the approximation u_n and corresponds to small wavelengths.

Denoting by P_n the orthogonal projector onto the span of the first n eigenvectors of A and $Q_n = P_m - P_n$, the following pair of non linear coupled equations in u_n and z_{m-n} replaces the initial system :

$$\frac{du_n}{dt} + \nu A u_n + P_n B(u_n + z_{m-n}, u_n + z_{m-n}) = P_n f \quad (4)$$

$$\frac{dz_{m-n}}{dt} + \nu A z_{m-n} + Q_n B(u_n + z_{m-n}, u_n + z_{m-n}) = Q_n f \quad (5)$$

The small scale structures z_{m-n} relax on a time much smaller than that of the large scales u_n, while if n is large enough z_{m-n} is small compared to u_n (see references [3] for proof). Equation (5) can then be simplified to read :

$$\nu A z_{m-n} + Q_n B(u_n, u_n) = Q_n f \quad (6)$$

This determines an approximate inertial manifold of system (1). Equation (6) can be considered as an approximate interaction law between small and large eddies ; z_{m-n} can be viewed as a correction to u_n when it is expressed as a function of u_n and substituted in (4)

In Dubois *et al.* [2], it is suggested that the cutoff number n can be dynamically defined by tracking the ratio of $||z_{m-n}||_{L^2}$ to time discretization error and the ratio of $||z_{m-n}||_{L^2}$ to $||u_n||_{L^2}$. In practice, several intermediate cutoff numbers n corresponding to different intermediate grid levels are determined and adjusted from time to time. u_n is computed with a V–cycle multigrid procedure along a fixed number of time steps N_T and by solving (4).

As for small scales z_{m-n}, they are not evaluated at each time step, rather the simplified equation (6) is solved at every N_T iterations.

3 Description of the numerical scheme and improvements.

The nonlinear Galerkin method can be seen as an adaptive choice of the optimal number of collocation points during time integration. Given m the largest wavenumber, the N_m associated collocation points define the finest grid in the physical space on which the solution of the system can be approximated. This number of modes m depends on the size of the computer, on Reynolds number, the subgrid scale model, and on the flow scales of interested.

Then the classical Galerkin method consists (apart from time discretization) of finding functions ψ_m and ω_m such that :

$$\frac{\partial \omega_m}{\partial t} + \nu A \omega_m + P_m J(\psi_m, \omega_m) = P_m \mathrm{curl} f \quad (7)$$

with

$$\omega_m = \sum_{|j| \leq m} \hat{\omega}_j(t) w_j(x) \quad \text{and} \quad \psi_m = - \sum_{|j| \leq m} \frac{\hat{\omega}_j(t)}{\lambda_j^2} w_j(x)$$

The nonlinear Galerkin scheme proposed in Dubois *et al.* [2] consists of determining three adaptive zones evaluated from time to time. Being given several possible numbers of modes to

be used ($4 \leq n_1 < \cdots < n_i < \cdots \leq m$), we denote by P_{n_i} the orthogonal projection onto the space spanned by the n_i first eigenvectors and $Q_{n_i} = P_m - P_{n_i}$. We then obtain :

$$\frac{\partial \omega_{n_i}}{\partial t} + \nu A\omega_{n_i} + P_{n_i}J(\psi_{n_i},\omega_{n_i}) = \\ +P_{n_i}\mathrm{curl}f - P_{n_i}\{J(\psi_{m-n_i},\omega_{n_i}) + J(\psi_{n_i},\omega_{m-n_i}) + J(\psi_{m-n_i},\omega_{m-n_i})\} \tag{8}$$

where

$$\begin{array}{ll} \omega_m = \omega_{n_i} + \omega_{m-n_i} & \\ \psi_m = \psi_{n_i} + \psi_{m-n_i} & \text{and} \quad \begin{array}{l} \omega_{n_i} = P_{n_i}\omega_m \\ \omega_{m-n_i} = Q_{n_i}\omega_m \end{array} \end{array}$$

In [2] the first cutoff wavenumber called n_{max} is defined by :

$$\forall n_i > n_{max} \quad \frac{||\omega_{m-n_i}||_{L^2}}{(\Delta t)^r} \leq \theta_1$$

where $\theta_1 \ll 1$ is a constant and r represents the order of the time discretization method. It means that higher order modes are smaller than the method's accuracy and need not be computed. Let us stress that θ_1 is not dimensionless and that in practice it is not simple to compare spatial and temporal discretization errors. In our case, we prefer to take the following dimensionless criterion :

$$\forall n_i \geq n_{max} \quad \frac{||P_{n_i}\{J(\psi_{m-n_i},\omega_{n_i}) + J(\psi_{n_i},\omega_{m-n_i}) + J(\psi_{m-n_i},\omega_{m-n_i})\}||_{L^2}}{||P_{n_i}J(\psi_{n_i},\omega_{n_i})||_{L^2}} \leq \theta_1 \ll 1 \tag{9}$$

where $P_{n_i}\{J(\psi_{m-n_i},\omega_{n_i}) + J(\psi_{n_i},\omega_{m-n_i}) + J(\psi_{m-n_i},\omega_{m-n_i})\}$ is the coupling term between large and small scales and expresses the action of small eddies on large eddies. Criterion (9) means that for $n_i > n_{max}$ the interactions between structures are small enough to be neglected. Figure 1 diplays the time evolution of ratio (9) computed within a Galerkin model, for 2 cutoff wavenumbers. The appearance of strong bursts imposes the dynamical computation of the cutoff wavenumber.

The second cutoff wavenumber n_{min} proposed in [2] is the following one (see fig. 2) :

$$\forall n_i > n_{min} \quad \frac{||\omega_{m-n_i}||_{L^2}}{||\omega_{n_i}||_{L^2}} \leq \theta_2 \tag{10}$$

where $\theta_2 \ll 1$ is a dimensionless constant. (10) expresses that the contribution of small scales to total vorticity is negligible. During time integration, ω_{m-n_i} can be frozen because its amplitude and evolution are smaller than those of ω_{n_i}, then the evolution of ω_{m-n_i} is quasistatic. A consequence of (10) is that the cutoff wavenumber n_{min} must be within or close to the dissipation range ; for instance in the case of a k^{-3} energy spectrum extanding from large scales up to dissipation scale k_d, looking for n_i inside the inertial range and neglecting enstrophy contained in the dissipation range, condition (10) reads :

$$\frac{||\omega_{m-n_i}||_{L^2}}{||\omega_{n_i}||_{L^2}} \simeq \sqrt{\frac{\int_{n_i}^{k_d}\frac{dk}{k}}{\int_1^{n_i}\frac{dk}{k}}} = \sqrt{\frac{\ln\frac{k_d}{n_i}}{\ln n_i}} \leq \theta_2$$

and then :

$$n_i \geq {k_d}^{1/(1+{\theta_2}^2)}$$

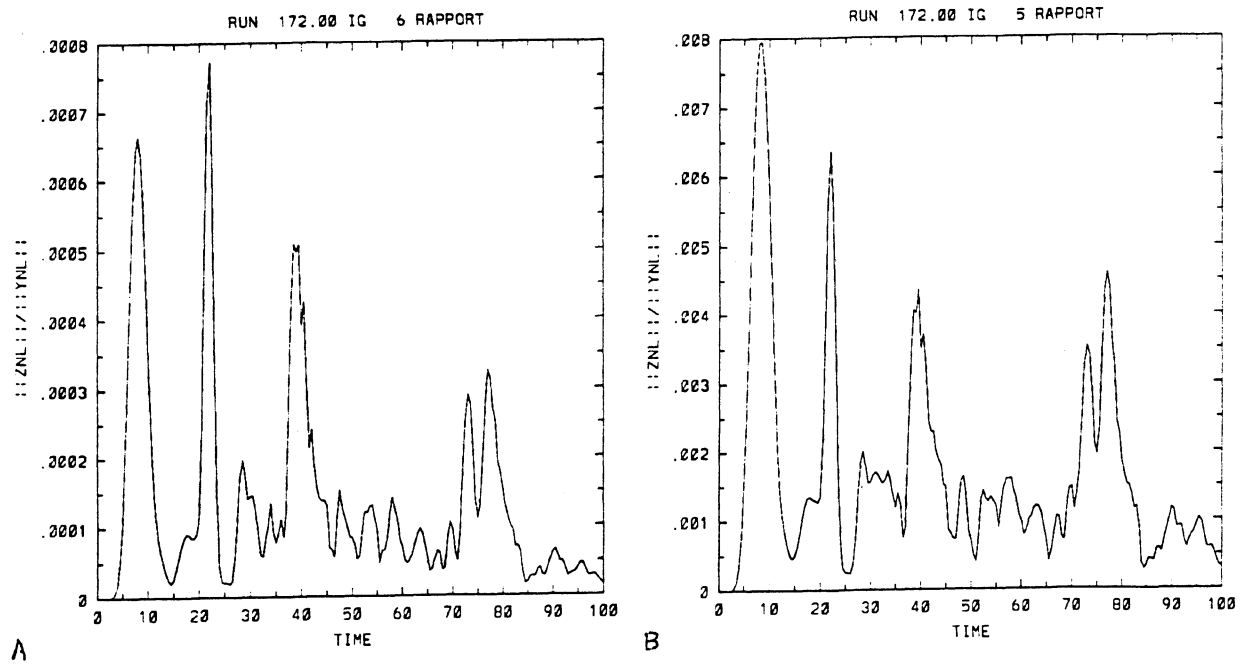


Figure 1: *Coupling term between small and large scales (9) computed within a classical Galerkin model.* $m = 128$. *a) On the grid* $n_i = 108$. *b) On the grid* $n_i = 96$.

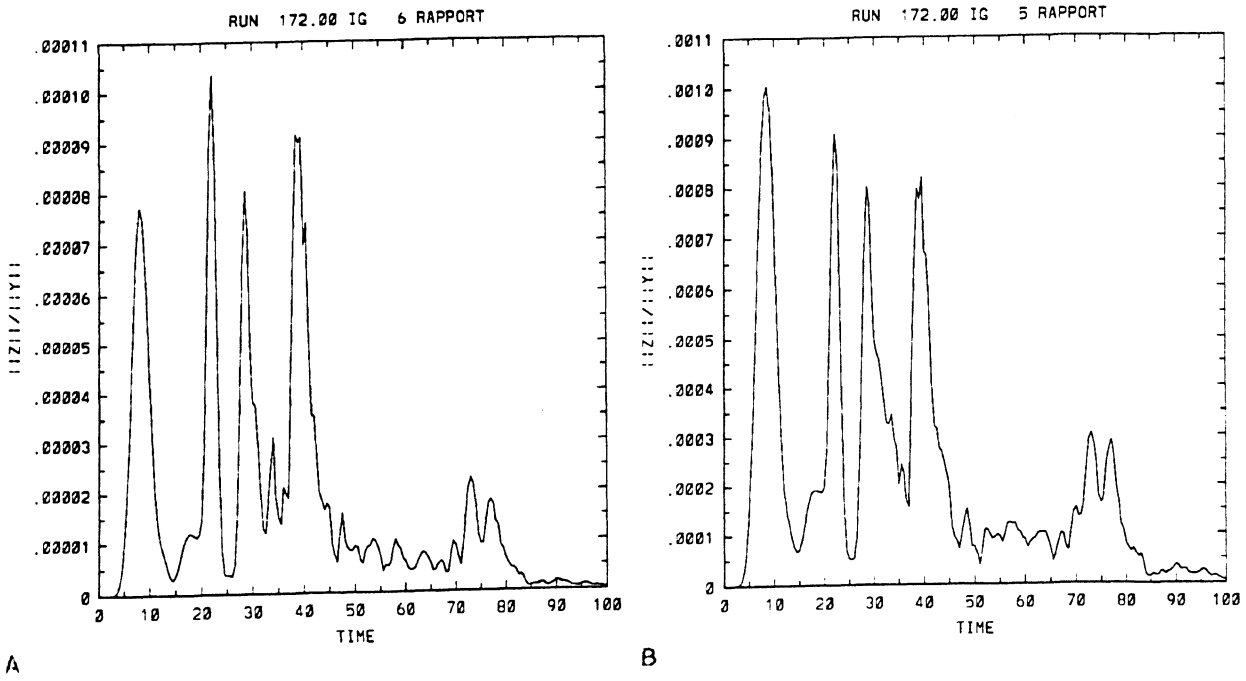


Figure 2: *Ratio of small to large scales (10) computed within a classical Galerkin model.* $m = 128$. *a) On the grid* $n_i = 108$. *b) On the grid* $n_i = 96$.

Let us remark that n_{min} is less than n_{max}. The choice of θ_1 and θ_2 will be discussed in the next section.

Modes corresponding to wavenumbers smaller than n_{min} belong to large scales and are integrated each time step with an exact integration scheme for the linear part and for the non linear terms with :

- an explicit Runge–Kutta scheme of the third order in references [6]
- an explicit Adams–Bashforth scheme of the second order in our case.

Modes between n_{min} and n_{max} are integrated with a V–cycle multigrid strategy and with a static approximation of the coupling terms. At each time step, (8) is solved on a level determined according the V–cycle. This zone permits a transition between stationary and dynamic approximation. These modes and their actions on the lower modes are not sufficiently small for a stationary approximation.

The multigrid strategy is performed during a time τ dynamically evaluated by :

$$\tau = \frac{||\omega_{m-n_{max}}||_{L^2}}{||Q_{n_{max}} J(\psi_m, \omega_m)||_{L^2}} \tag{11}$$

which is approximatively equal to the characteristic time of the small structures $\omega_{m-n_{max}}$, namely :

$$\frac{\omega_{m-n_{max}}}{\dot{\omega}_{m-n_{max}}}$$

This time is the smallest characteristic time of ω_{m-n_i} for $n_i < n_{max}$ (see fig. 3). So during this time, coupling terms can be fixed without introducing large errors. In test presented in [6], this time was fixed by hand.

Modes greater than n_{max} are integrated from time to time. Jauberteau *et al.* [6] propose to compute these modes as a function of lower modes according to the reduce equation (6) which can be rewritten as :

$$\nu A \omega_{m-n_i} + Q_{n_i} J(\psi_{n_i}, \omega_{n_i}) = Q_{n_i} \mathrm{curl} f$$

However this approximation supposes the complete relaxation of small scales. Instead, as Jauberteau in [4], we propose to compute the action of ω_{n_i} on ω_{m-n_i} by solving the original equation (7) on the finest grid. The computational time is almost the same and the integration is more accurate.

The highest modes which are frozen during characteristic time τ, must be computed during a time that we propose to call *relaxation time*, large enough to express as best as possible the interaction of low and high modes and also to take the effect of viscosity into account. The flow chart (see fig. 4) summarizes the algorithm.

4 Numerical experiments.

The first series of experiments performed to test the schemes described in previous sections concerns a vortex pair in interaction without forcing. Two vortices of the same sign orbit around each other and if their distance is not too large, they roll up and merge to form a unique vortex which dissipates. Figure 5 shows the initial field.

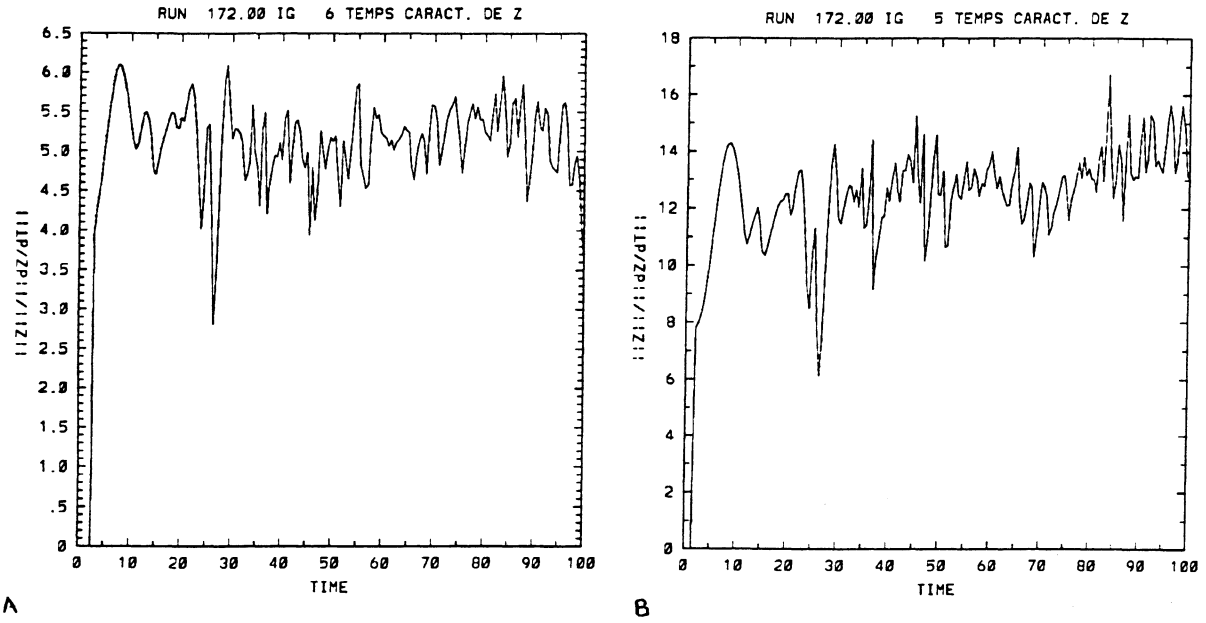


Figure 3: *Characteristic time (11) computed with a classical Galerkin method.* $m = 128$. *a) On the grid* $n_i = 108$. *b) On the grid* $n_i = 96$.

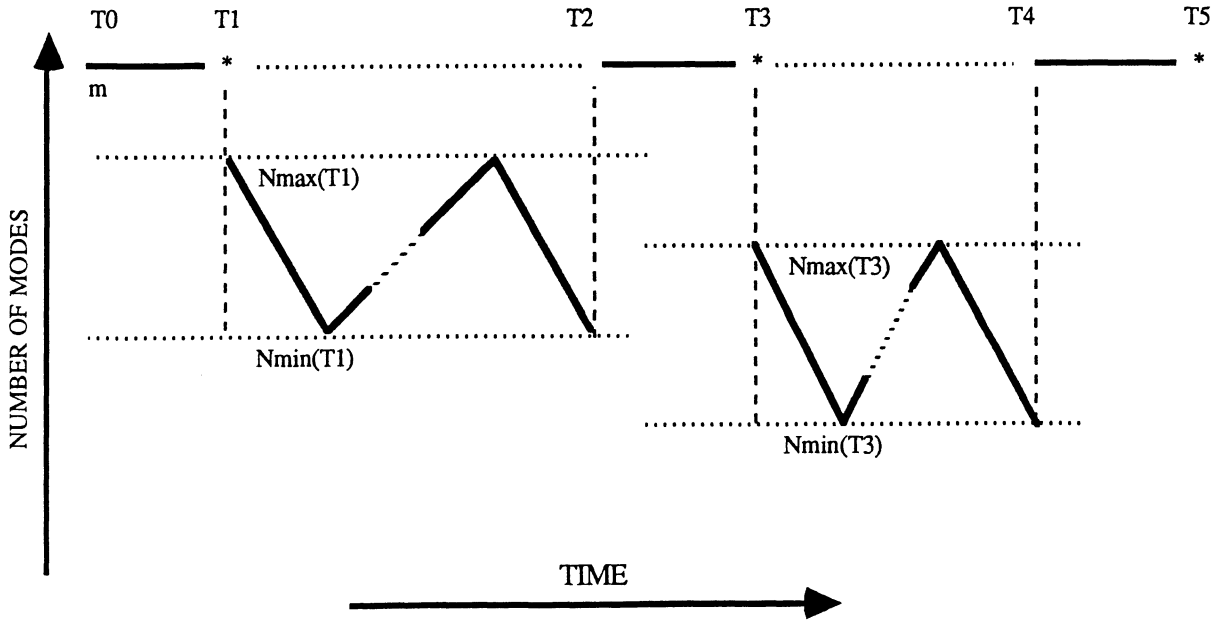


Figure 4: *The nonlinear Galerkin method. (*) marks the practical times for the determination of* n_{max}, n_{min} *and the number of time step for multigrid integration (see section 2).*

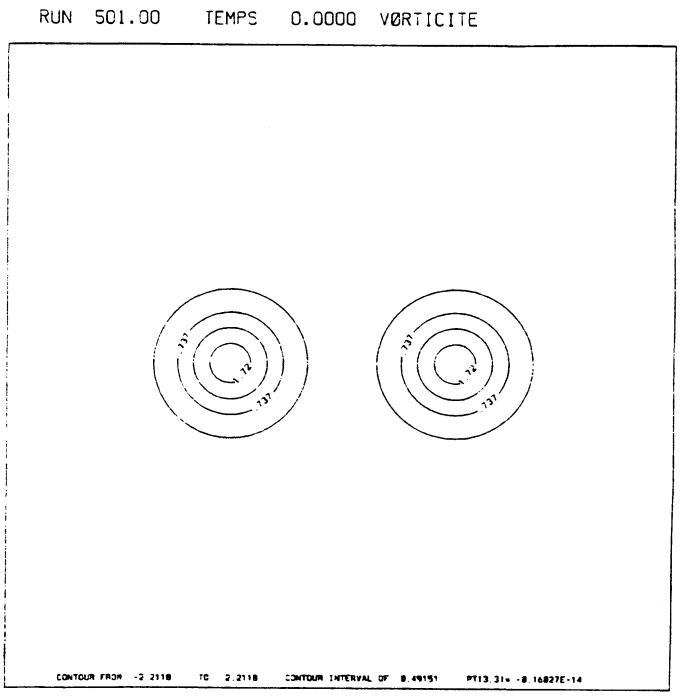


Figure 5: *2 vortices of same sign centered at* $(2\pi/3, \pi)$ *and* $(4\pi/3, \pi)$ *respectively. Vorticity intensity in core* $=$ 2.0. *Initial energy* $=$ $1.2\ 10^{-2}$. *Initial enstrophy* $=$ $3.7\ 10^{-2}$.

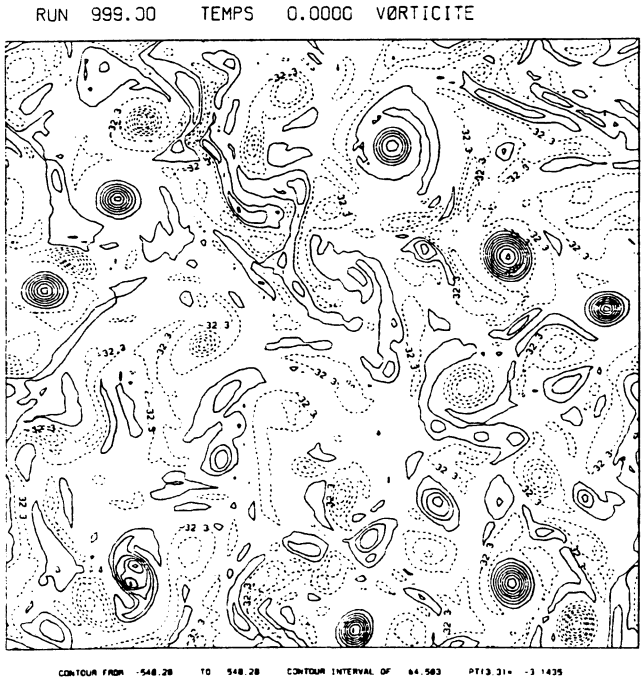


Figure 6: *Initial turbulent flow : we let turbulence operate. Initial energy* $=$ 49.5. *Initial enstrophy* $=$ 2522.9.

Next experiments still concern a decaying turbulent flow, but the initial field is a motion obtained from a fully developed turbulent flow and we let turbulence operate (see fig. 6).

We are interested in large eddy simulations i.e. with a Reynolds number so large that the numerical cutoff falls inside the inertial range. In that case the numerical model has to incorporate a subgrid scale parametrization. It models the effect of scales which are not explicitely described owing to insufficient resolution. In all experiments presented here, we used the hyperviscosity model consisting of a linear diffusion with a high dissipativity introduced in Basdevant and Sadourny [1] of the following form :

$$-\frac{\nu}{(k_T)^{2p}}(\Delta)^p = -\frac{\nu}{(k_T)^{2p}}(\text{curl curl})^p$$

with $p = 2, 4$ or 8 and where k_T is the cutoff mode and ν is choosen such that $\nu\Delta t \simeq 0.5$. This subgrid modelling assumes that the large scale eddies are independant of the detailed structure of small scales, then the Reynolds number is no longer a pertinent parameter provided it is large enough i.e. the dissipation scale is negligible compared to the cutoff one.

Experiments are realized in all cases with a spectral dealiased code with a square cutoff at 0.66 ratio. The classical Galerkin scheme is performed on a 128×128 grid, with time integration consisting of an exact integration for the linear part, with an explicit Adams–Bashforth scheme for the nonlinear terms.

The nonlinear Galerkin method scheme is computed with the finest mesh equal to the 128×128 grid ; the different permissible coarse grids are given in the table 1 where n represents the number of modes in each direction.

number of level	1	2	3	4	5	6	7
n	32	48	64	80	96	108	128
$n/2$	16	24	32	40	48	54	64
$n/2 \times 0.66$	10	16	21	26	32	36	42

Table 1: *Permissible grid levels for nonlinear Galerkin method*

The full nonlinear term $J(\psi_m, \omega_m)$ is evaluated once just after the determination of the cutoff numbers n_{max} and n_{min}, while the coupling term of the equation, $P_{n_i}\{J(\psi_{m-n_i}, \omega_{n_i}) + J(\psi_{n_i}, \omega_{m-n_i}) + J(\psi_{m-n_i}, \omega_{m-n_i})\}$, is obtained at each intermediate level n_i by the difference between $P_{n_i}J(\psi_m, \omega_m)$ and $P_{n_i}J(\psi_{n_i}, \omega_{n_i})$; it is fixed while multigrid integration is in progress so as to minimize computing time.

Figure 7 and 10 show readily the similarity of the energy spectra and vorticity contours for different decaying turbulences as computed by the two methods. Some small differences are observed between the two schemes (see fig. 10), especially the values of vorticity and the positions of some coherent structures. This is due to the small scales not being integrated at each time step. Figure 8 and 11 show how the levels n_{max} and n_{min} evolve. With energy spectra (figure 7 and 10), they permit to observe that n_{min} stays inside dissipation range. Figure 9 and 12 show that the number of time steps during V–cycle multigrid strategy is strongly dependent in time and in level n_{max} where this number is evaluated.

In practice, θ_1 and θ_2 have to be choosen so that cutoff n_{min} and n_{max} stay inside the dissipation range. Otherwise, the enstrophy cascade is not well simulated and the high dissipativity modelling subgrid scales is not really taken into account. The enstrophy not being dissipated,

accumulates and some small unphysical structures appear in the case of the interaction vortex pair (see fig. 13). In the case of homogeneous turbulence, some vortex pairs do not merge (compare fig. 14 and fig. 15). Accordingly, the larger is the dissipative zone, the lower are the cutoffs n_{min} and n_{max}, and the more important is the gain of computational time. That gain is due to the fact that the approximate solution is obtained on coarser grid. So it is not surprising that the gain decreases with p ; for $p = 8$, cpu time increases because the integration is carried out on the sixth grid, as determined by n_{max} and n_{min}. Figure 16 and 17 show how n_{min} and n_{max} behave depending on dissipativity p.

The last experiments relate to forced turbulence. The initial field is a zero mean random vorticity field : the forcing is obtain by keeping the amplitude of a mode $(0, k_F)$ constant during time integration and the model is integrated until statistical stationary regime is reached. Figure 18 and 19 illustrate results. Because small scales are differently integrated with the classical Galerkin method and the nonlinear Galerkin method, some small differences appear and they spread to all scales of the flow after a time called predictability time. Only large scale behavior of the flow are similar as well as energy vs time and enstrophy vs time.

The table 2 summarizes performances of the new method based on the different series of numerical experiments.

	p	Δt	θ_1	θ_2	cpu time CG per iteration	cpu time NLG per iteration	gain %
Vortex pair	2	0.01	0.01	0.02	$2.36\ 10^{-2}$	$1.26\ 10^{-2}$	47
	4	0.01	0.01	0.02	$2.43\ 10^{-2}$	$1.93\ 10^{-2}$	20
	8	0.01	0.01	0.02	$2.40\ 10^{-2}$	$2.30\ 10^{-2}$	04
Decaying turbulence	4	0.0001	0.05	0.05	$2.40\ 10^{-2}$	$1.77\ 10^{-2}$	26
	8	0.0001	0.05	0.05	$2.43\ 10^{-2}$	$2.43\ 10^{-2}$	0
Forced turbulence	4	0.0001	0.05	0.05	$2.29\ 10^{-2}$	$1.71\ 10^{-2}$	25

Table 2: *Data and results of experiments. CG = classical galerkin method. NLG = nonlinear galerkin method.*

In conclusion, the nonlinear Galerkin method constructed from recent developments in dynamical system theory has been adapted with success to numerical simulation of turbulence with a subgrid scale model. However the computational gain is important only when the dissipation range is large. We defined dimensionless criteria and characteristic times necessary for practical implementation. Further developments should define a dynamical time step, increasing for coarser grids.

The technique is applicable both to the decaying and forced turbulence and it generates the expected large–scale structures.

Acknowledgments

We are grateful to A. Babiano, T. Dubois, M. Farge, F. Jauberteau, D. Oueslati, T. Philipovitch and R. Temam for helpful discussions.

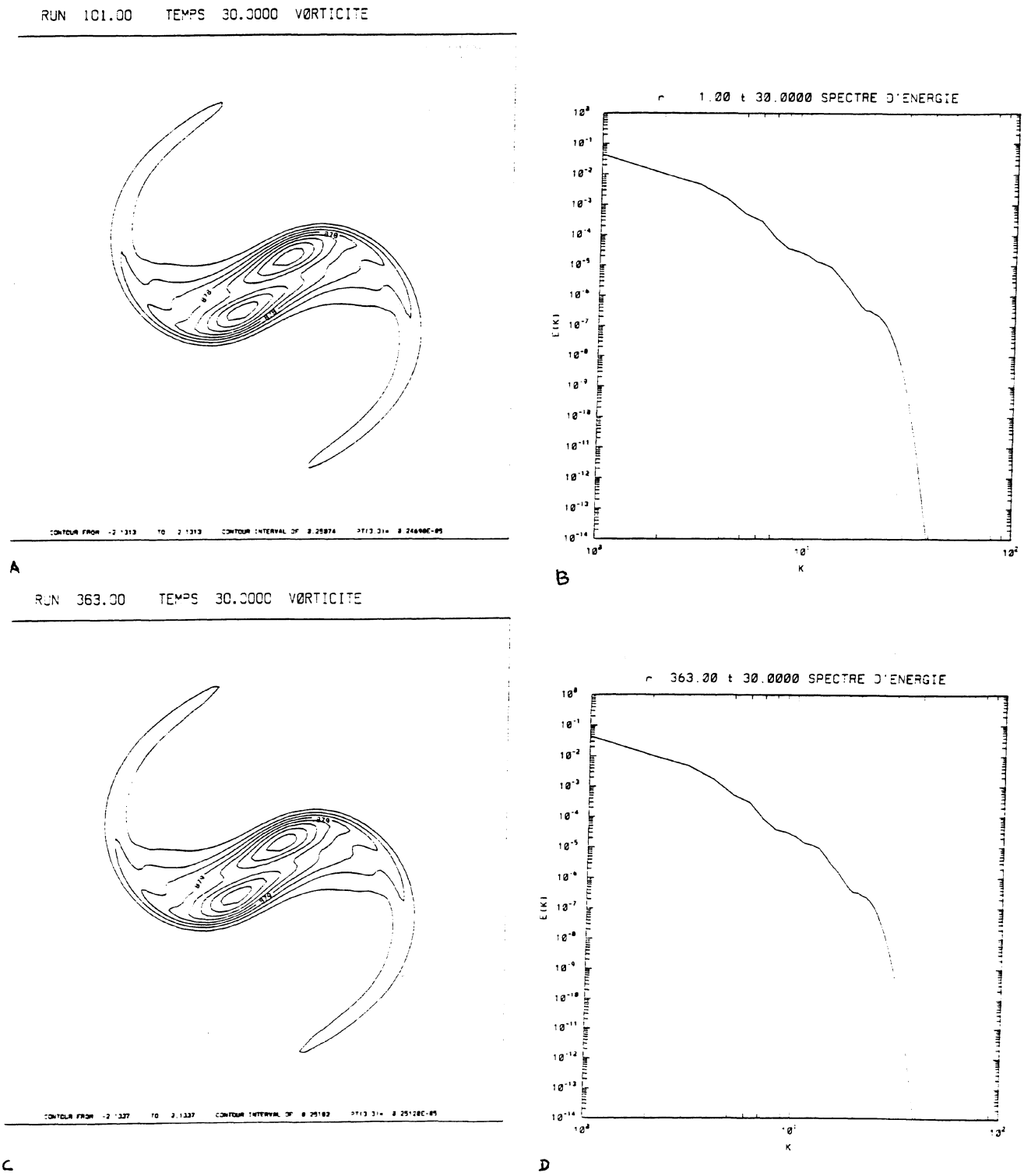


Figure 7: *Interacting vortex pair (no forcing) at $t = 30.0$. ($p = 4$). a) Vorticity contours computed with classical Galerkin scheme. b) Energy spectrum computed with classical Galerkin scheme. c) Vorticity contours computed with nonlinear Galerkin scheme. d) Energy spectrum computed with nonlinear Galerkin scheme.*

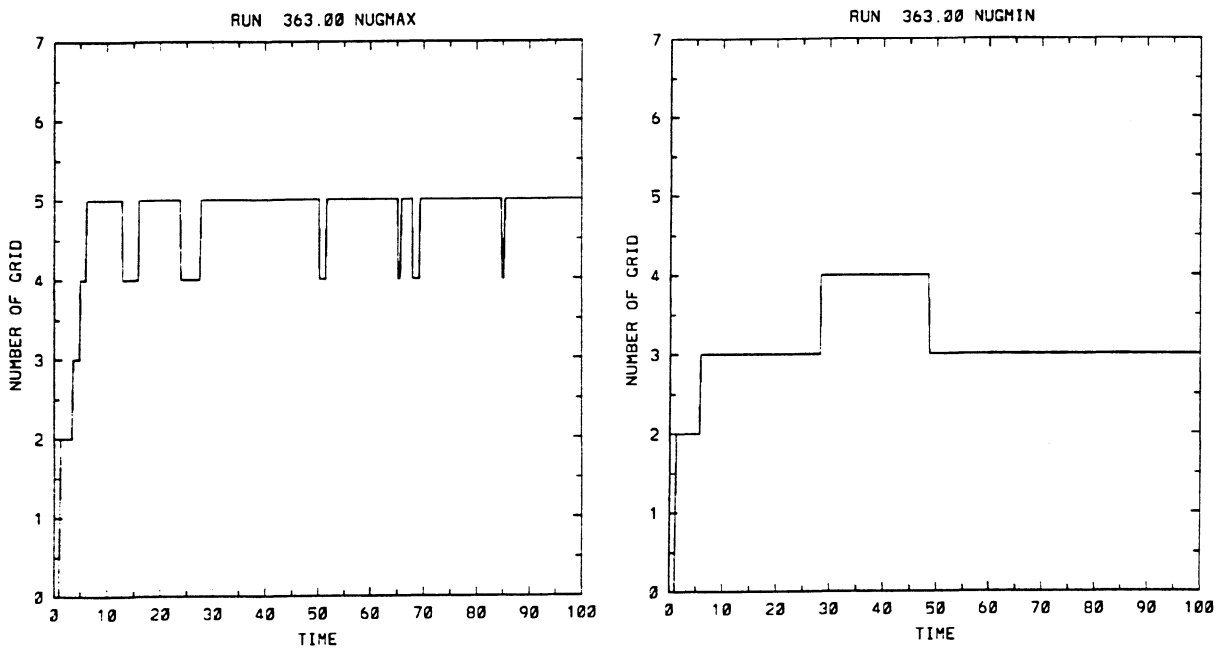


Figure 8: *Interacting vortex pair. Nonlinear Galerkin method.* n_{max}, n_{min} *vs time.*

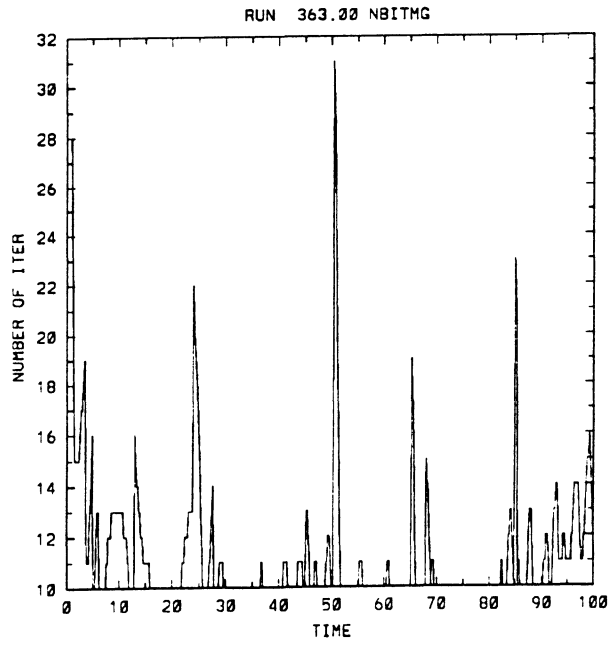


Figure 9: *Interacting vortex pair. Nonlinear Galerkin method. Caracteristic time vs time.*

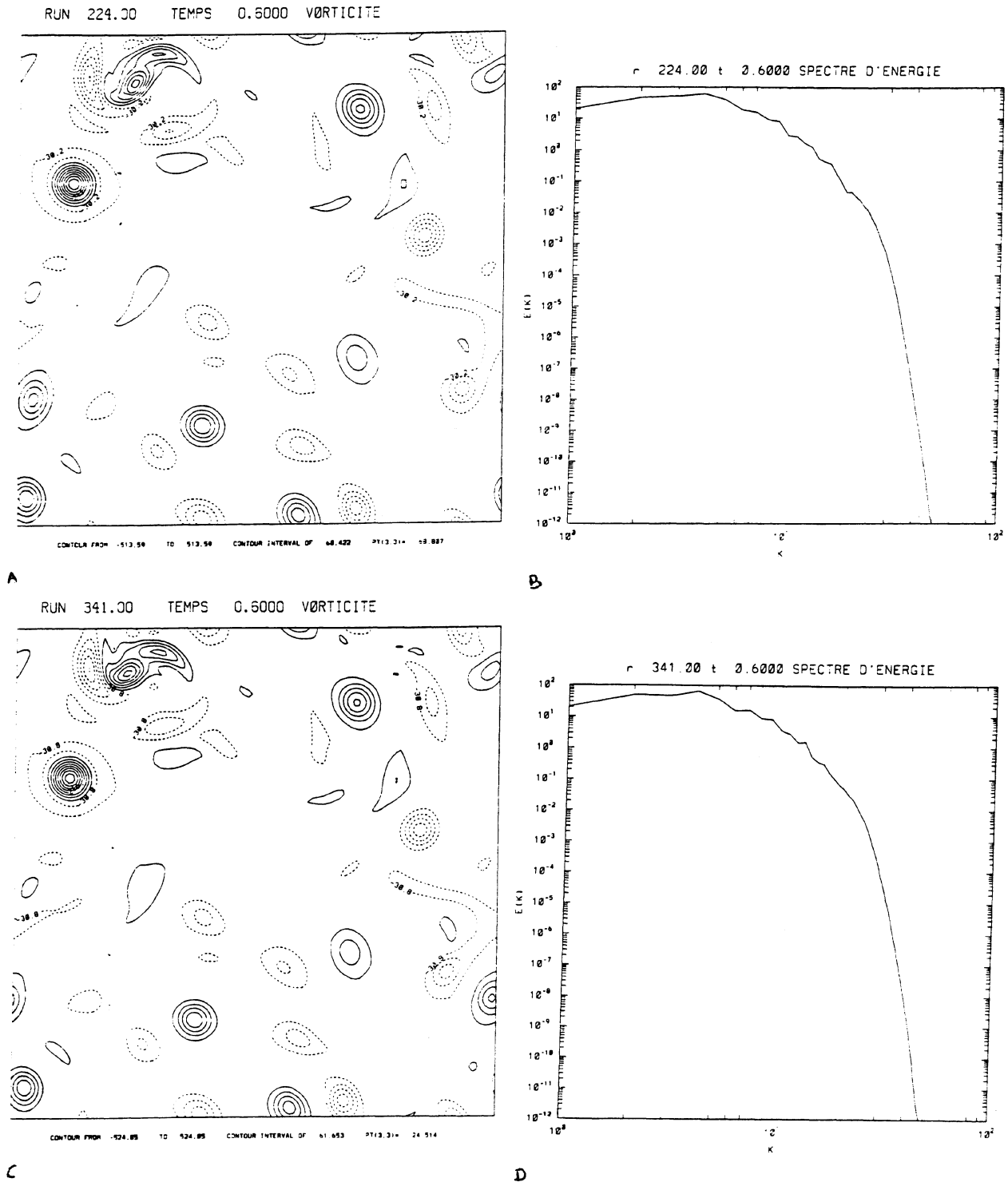


Figure 10: *Decaying turbulent flow at* $t = 0.60$. $(p = 4)$. a) *Vorticity contours computed with classical Galerkin scheme. b) Energy spectrum computed with classical Galerkin scheme. c) Vorticity contours computed with nonlinear Galerkin scheme. d) Energy spectrum computed with nonlinear Galerkin scheme.*

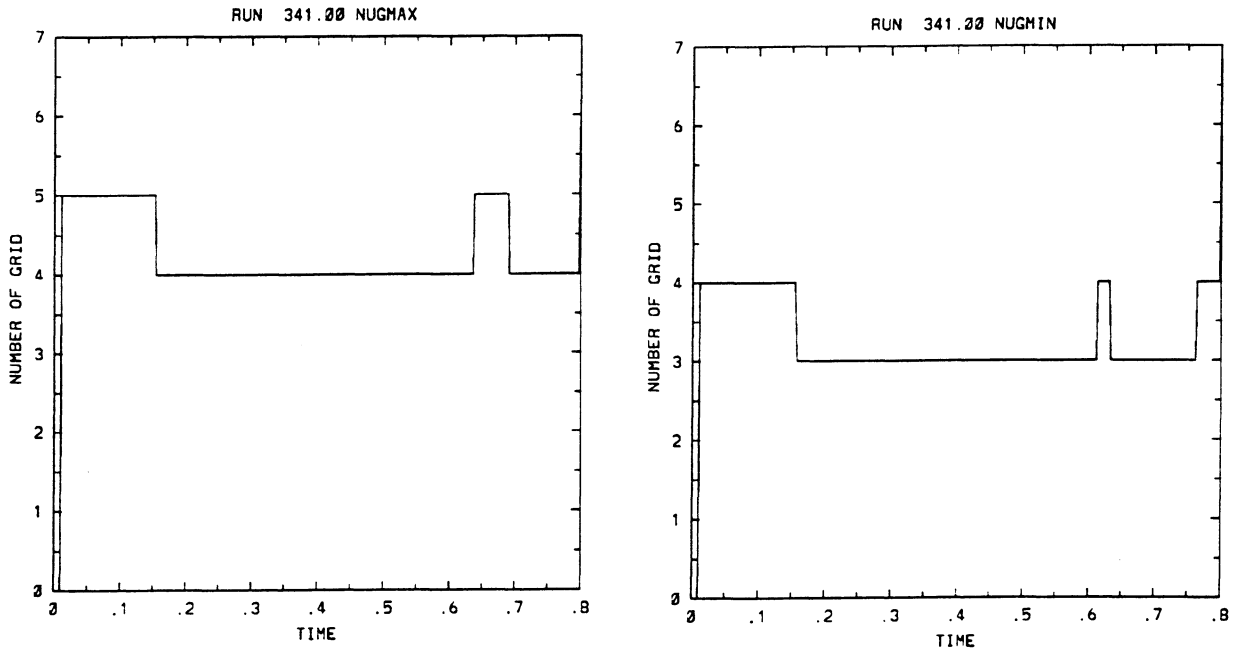


Figure 11: *Decaying turbulent flow. Nonlinear Galerkin method.* n_{max}, n_{min} *vs time.*

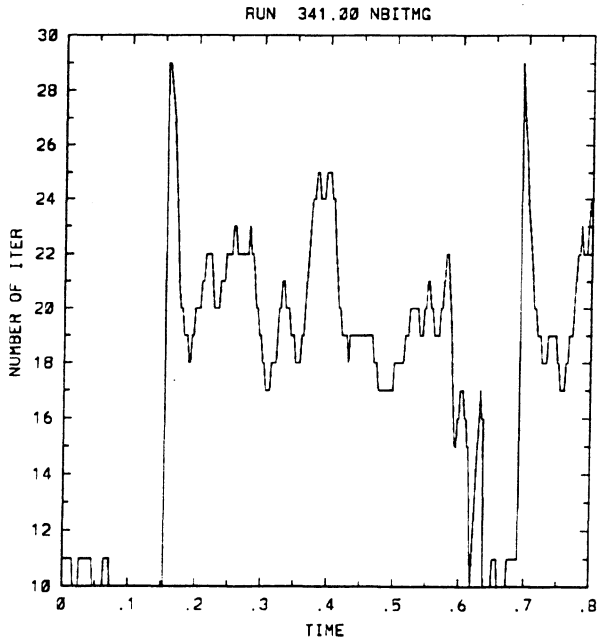


Figure 12: *Decaying turbulent flow. Nonlinear Galerkin method. Caracteristic time vs time.*

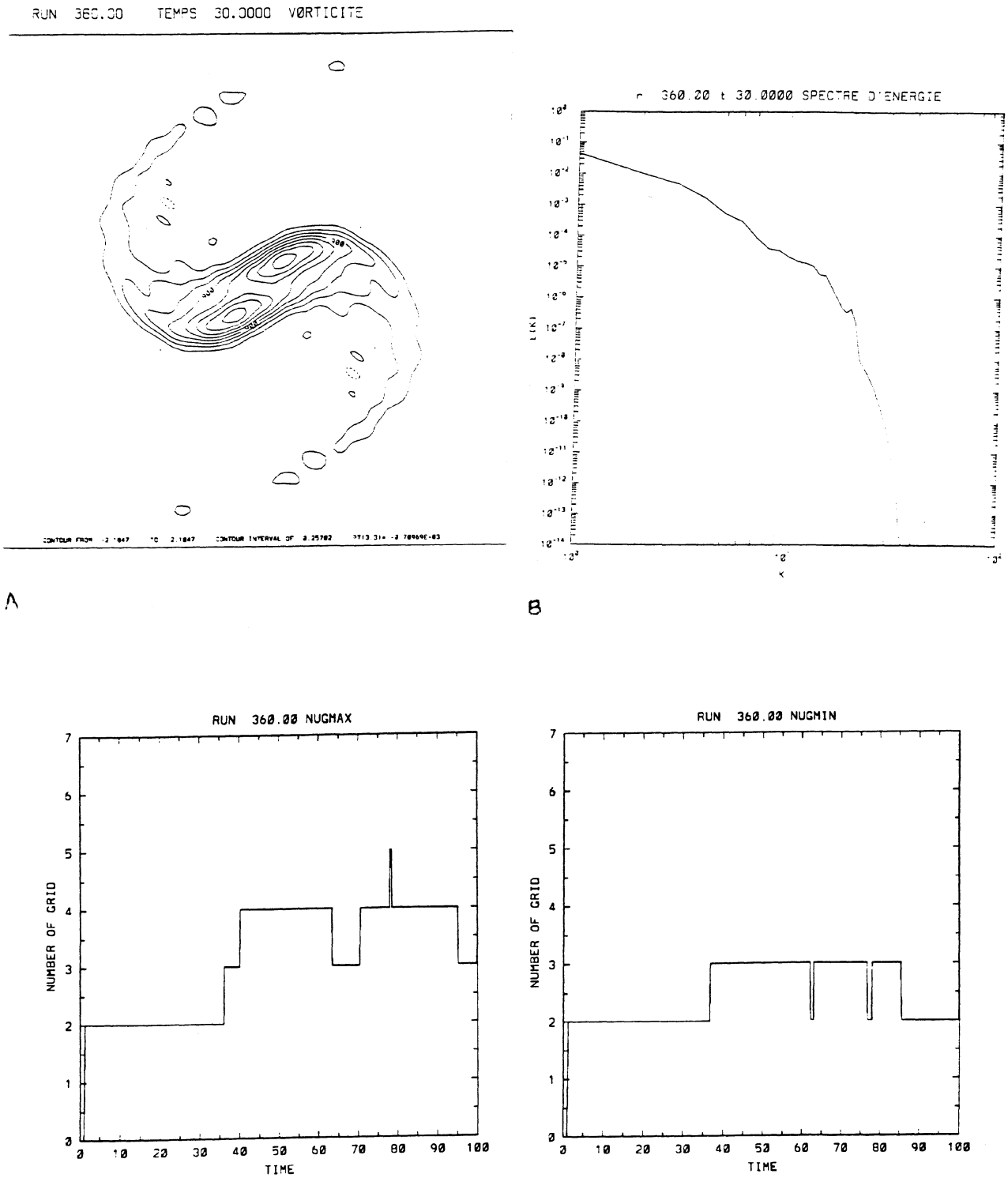


Figure 13: *Nonlinear Galerkin scheme with cutoff inside the inertial range. a) Vorticity contours at* $t = 30.0$. *b) Energy spectrum at* $t = 30.0$. *c)* n_{max}, n_{min} *vs time.*

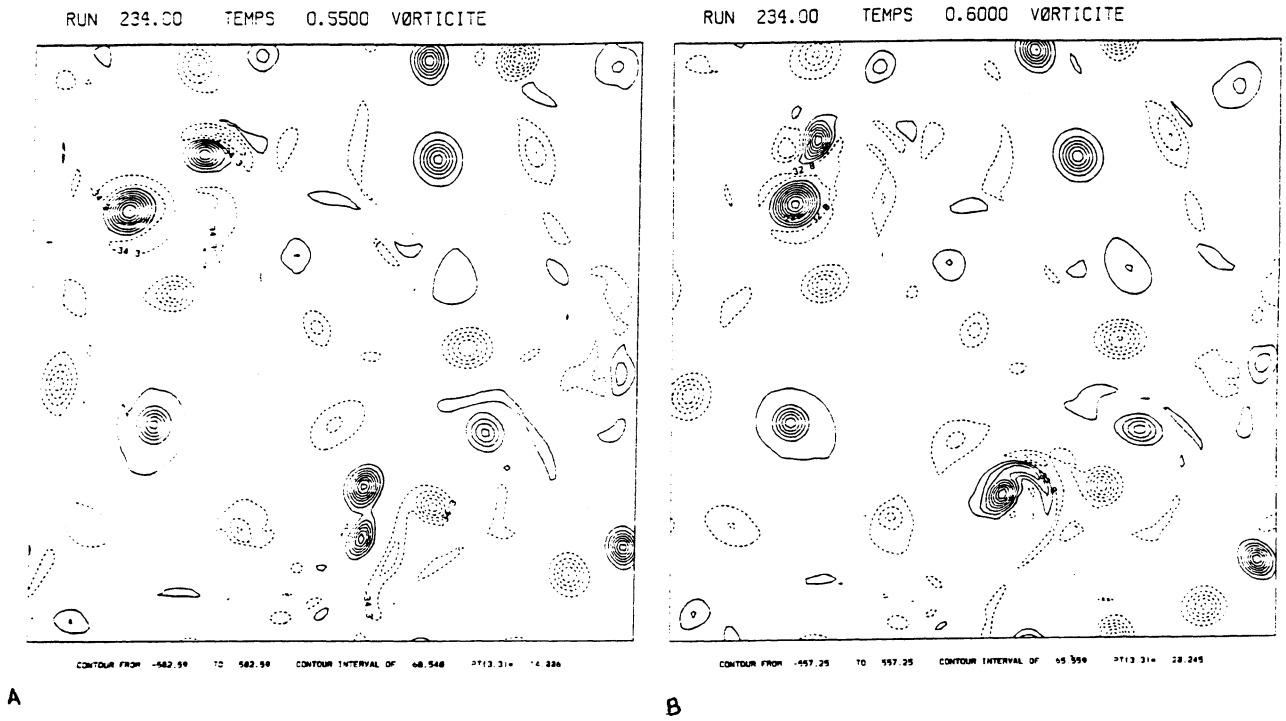


Figure 14: *Decaying turbulent flow ($p = 8$) computed with the classical Galerkin scheme. a) Vorticity contours at $t = 0.55$. b) Vorticity contours at $t = 0.60$.*

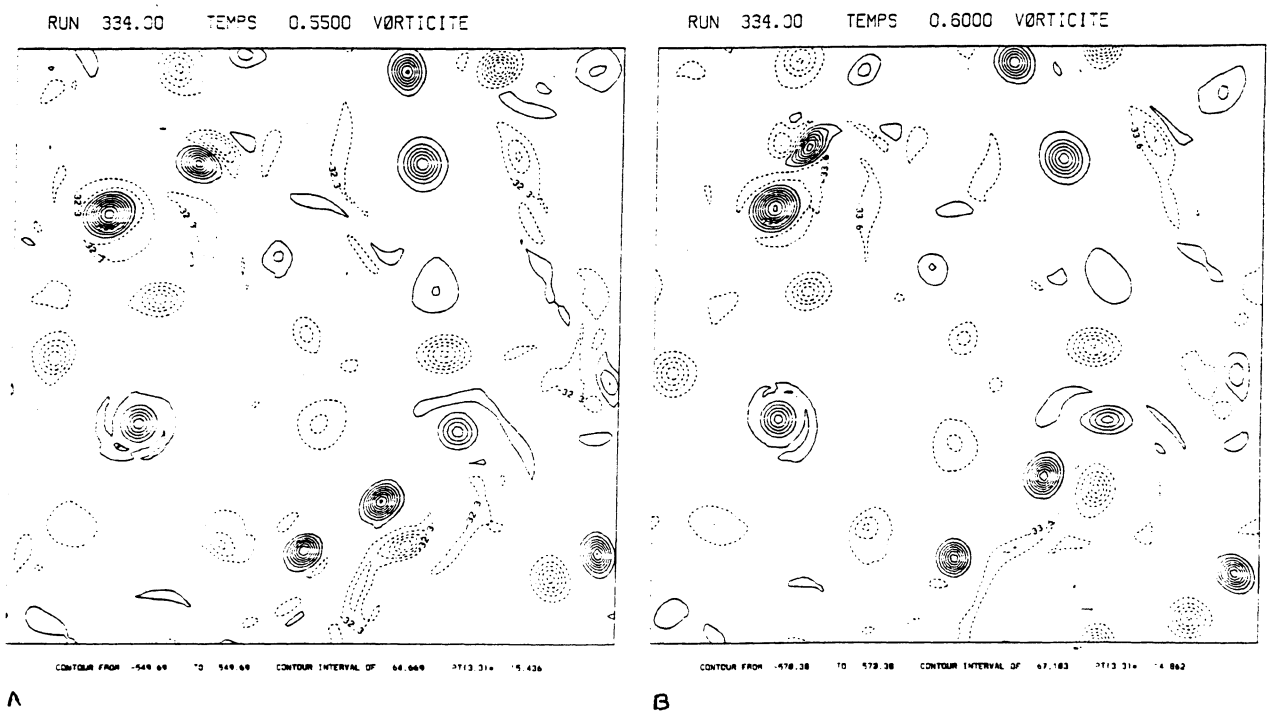


Figure 15: *Decaying turbulent flow ($p = 8$) computed with nonlinear Galerkin scheme with cutoff inside the inertial range. a) Vorticity contours at $t = 0.55$. b) Vorticity contours at $t = 0.60$.*

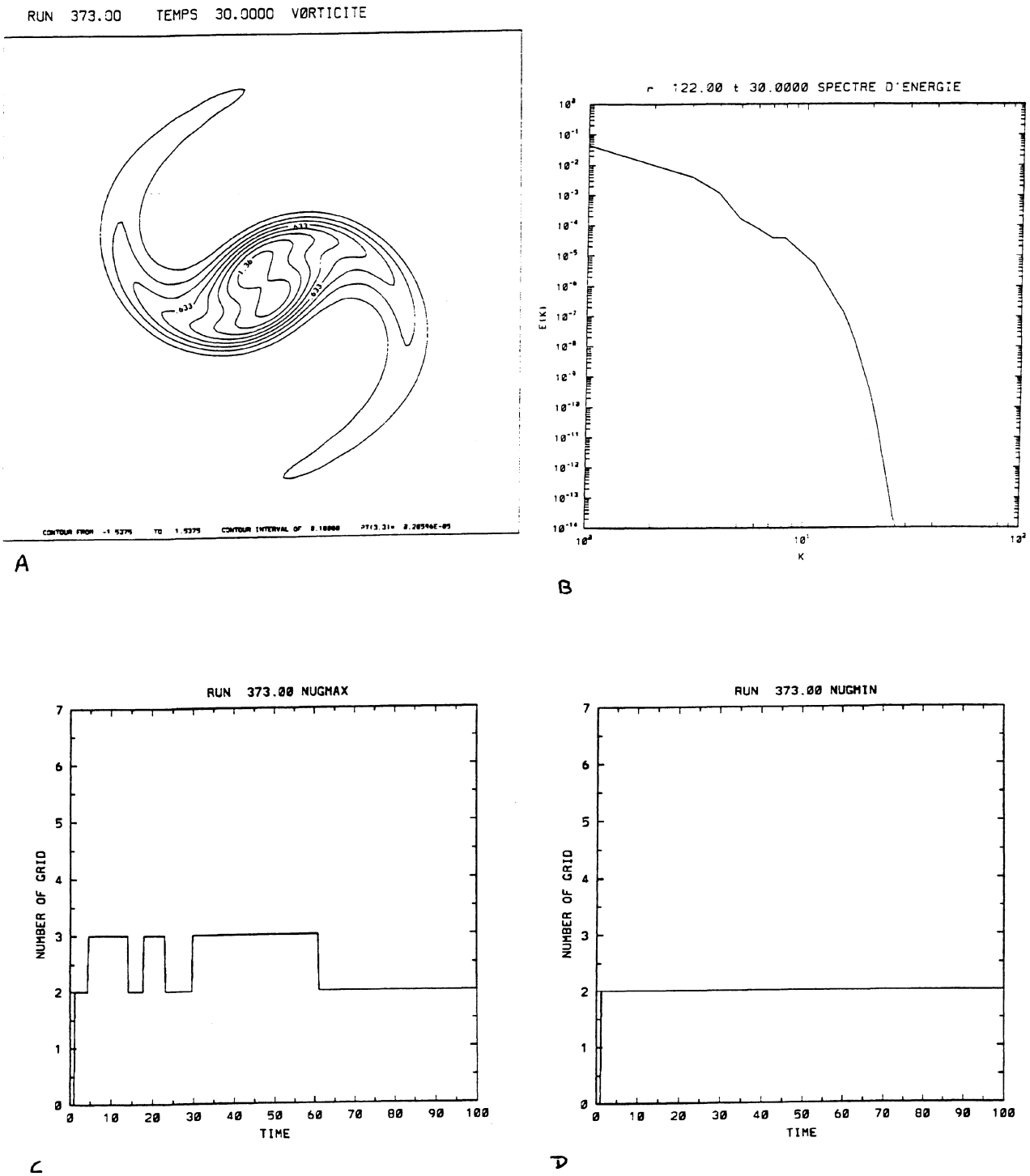


Figure 16: *Influence of dissipativity : $p = 2$.* a) *Vorticity contours at $t = 30.0$.* b) *Energy spectrum at $t = 30.0$.* c) *n_{max} vs time.* d) *n_{min} vs time.*

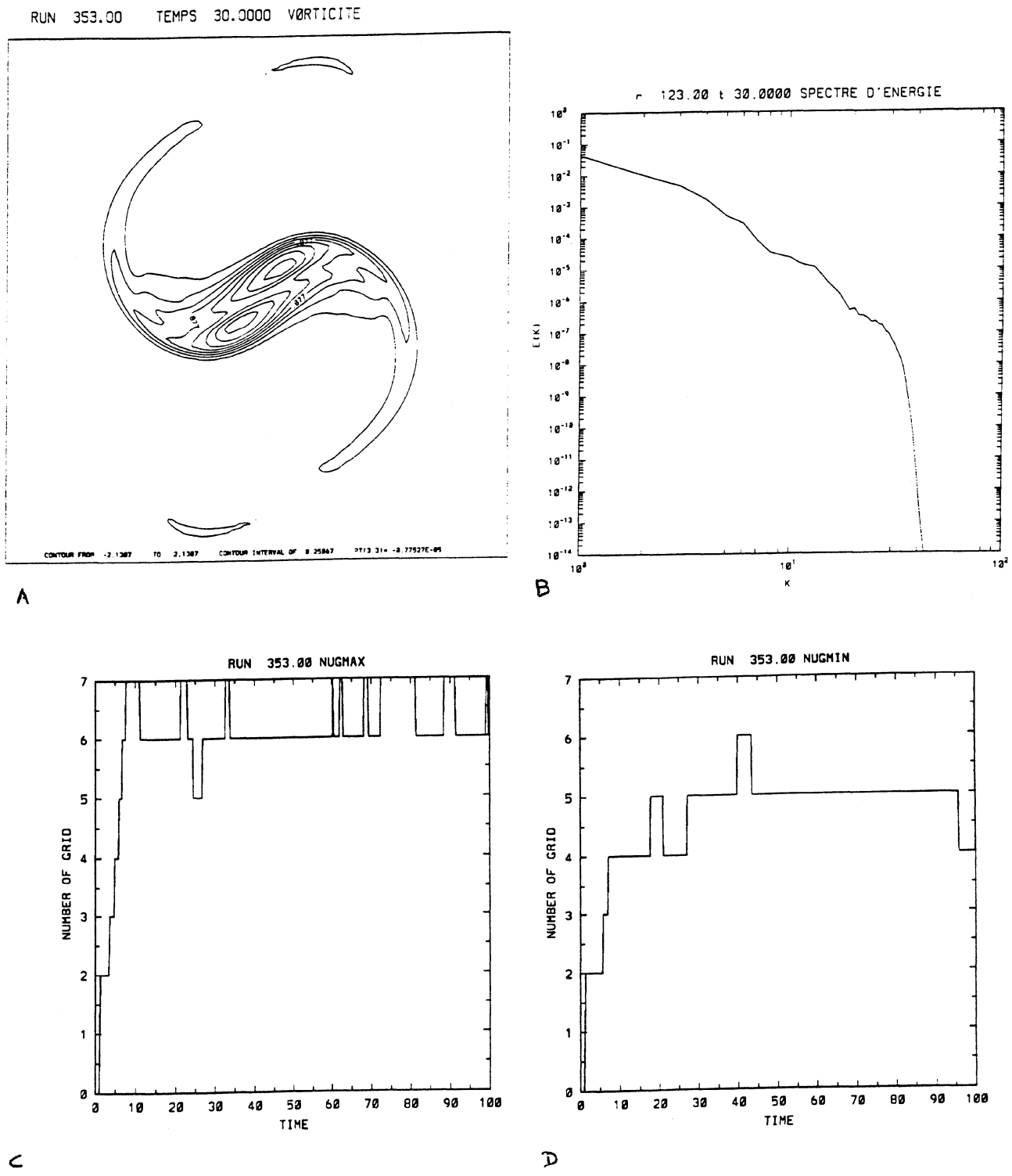


Figure 17: *Influence of dissipativity : $p = 8$. a) Vorticity contours at $t = 30.0$. b) Energy spectrum at $t = 30.0$. c) n_{max} vs time. d) n_{min} vs time.*

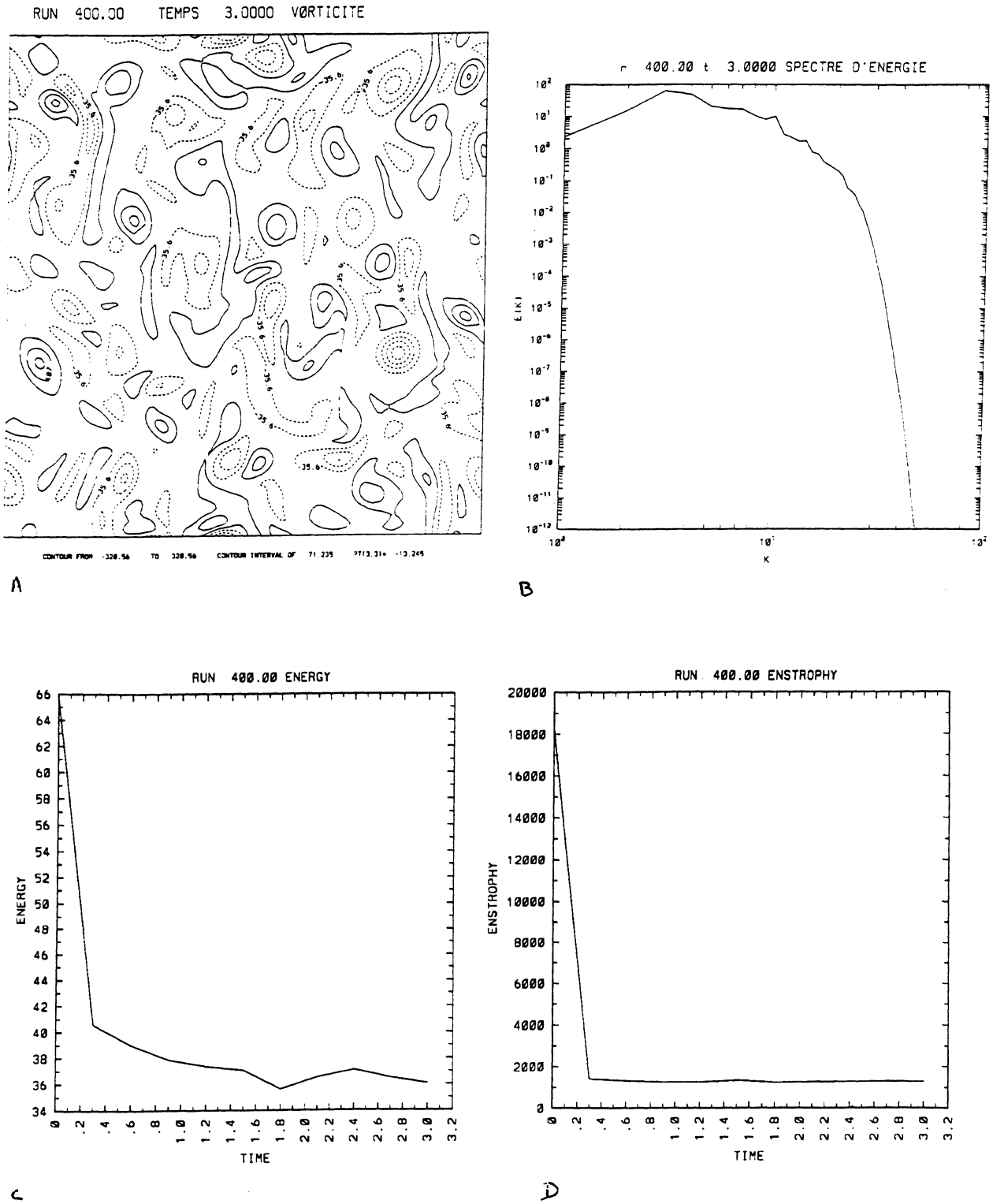


Figure 18: *Forced turbulence performed with classical Galerkin method ($p = 4$, $m = 128$ and $k_F = 10$). Initial energy* $= 65.8$. *Initial enstrophy* $= 18383.5$. *a) Vorticity contours at* $t = 30.0$. *b) Energy spectrum at* $t = 30.0$. *c) Energy vs time. d) Enstrophy vs time.*

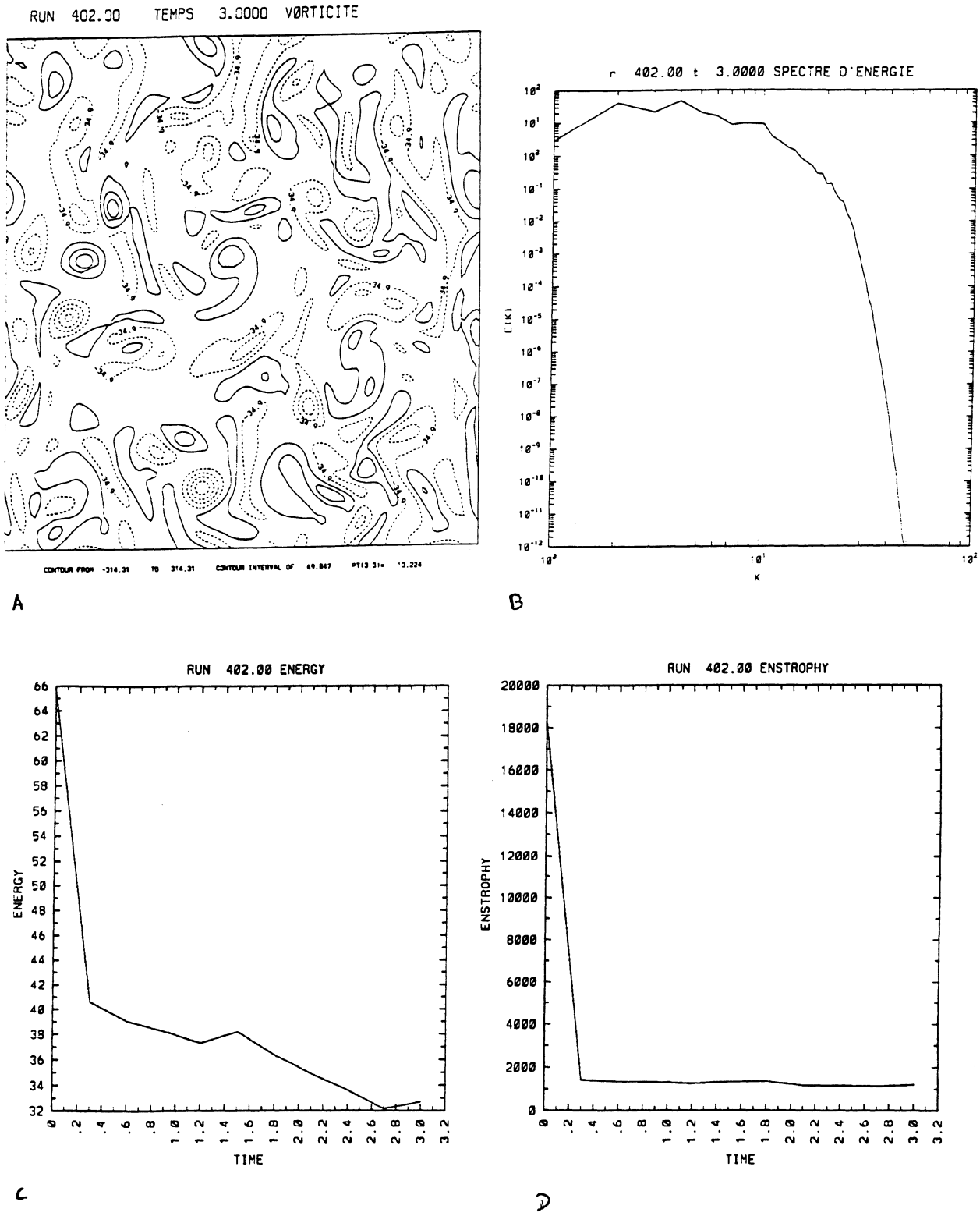


Figure 19: *Forced turbulence performed with nonlinear Galerkin method ($p = 4$, $m = 128$ and $k_F = 10$). Initial energy = 65.8. Initial enstrophy = 18383.5. a) Vorticity contours at $t = 30.0$. b) Energy spectrum at $t = 30.0$. c) Energy vs time. d) Enstrophy vs time.*

References

[1] C. Basdevant and R. Sadourny, *Modélisation des échelles virtuelles dans la simulation numérique des écoulements turbulents bidimensionnels* J. de Mécanique théorique et appliquée, 1983, 243-269.

[2] T. Dubois, F. Jauberteau and R. Temam, *Solution of the incompressible Navier-Stokes equations by the nonlinear Galerkin method*, to appear in J. of Comp. Phys.

[3] C. Foias, O. Manley and R. Temam, *On the interaction of small and large eddies in two–dimensional turbulent flows*, Math. Mod. and Num. Anal. (M2AN), 22, 1988, 93-114.

[4] F. Jauberteau, *Résolution numérique des équations de Navier–Stokes instationnaires par méthodes spectrales. Méthode de Galerkin nonlinéaire*, Thèse Université de Paris–Sud, 1990.

[5] F. Jauberteau, C. Rosier and R. Temam, *The nonlinear Galerkin method in computational fluid dynamics*, App. Num. Math. 6, 1989-90, 361-370.

[6] F. Jauberteau, C. Rosier and R. Temam, *A Nonlinear Galerkin Method for Navier Stokes Equations*, Comp. Meth. in App. Mec. and Eng. 80, 1990 , 245-260.

[7] M. Marion and R. Temam, *Nonlinear Galerkin Methods*, SIAM J. Num. Anal., 26, 1989, 1139-1157.

Chapitre II

RESOLUTION DES EQUATIONS DE NAVIER–STOKES PAR LA METHODE DES ELEMENTS FINIS 4P1-P1. (GALERKIN CLASSIQUE)

L'objet de ce chapitre est de présenter une méthode d'éléments finis pour résoudre les équations de Navier–Stokes bidimensionnelles. Il s'agit d'une approche dite de "Galerkin classique" utilisant l'élément fini, conforme, mixte 4P1-P1 (ou encore isoP2-P1). Le problème modèle de la cavité entrainée est présenté en section II.2, puis la section II.3 décrit la méthode de discrétisation en espace et la section II.4 présente le schéma de discrétisation en temps. Le problème discret de Stokes est résolu par un algorithme de type Uzawa et le problème non linéaire par un schéma du moindre carré.

L'aboutissement de ce travail est la réalisation d'un code numérique. Quelques résultats concernant la cavité entrainée sont proposés dans la section II.4.3 et prouvent le bon fonctionnement du code développé. Ce code que l'on appellera "Galerkin classique", d'une part sert de référence et permet d'autre part d'étudier et d'analyser a posteriori le comportement de nouvelles bases telles que la base hiérarchique qui introduit des structures de tailles différentes (voir l'étude au chapitre III) contrairement à la base canonique. C'est également une base de travail pour la réalisation d'un code résolvant les équations de Navier–Stokes avec une méthode de type Galerkin non linéaire.

II.1 Présentation des éléments finis.

Pourquoi les éléments finis?[1]

La méthode des éléments finis est une méthode qui concerne la résolution de problèmes mathématiques et physiques définis par des équations aux dérivées partielles où les degrés de libertés (en nombre infini) du système sont remplacés par un nombre fini de paramètres : c'est la discrétisation du problème. Bien que les méthodes dites spectrales, différences finies... procèdent aussi de cette façon, la technique des éléments finis est cependant la seule actuellement en mécanique des fluides qui puisse s'appliquer à toutes les équations, quels que soient le domaine d'application et la complexité de la géométrie, les conditions aux limites et les conditions initiales.

La méthode des éléments finis (approche de Galerkin) consiste à chercher une solution approchée u_h de l'inconnue u dans un espace V_h de dimension finie, u_h étant de la forme :

$$u_h(x,t) = \sum_{i=1}^{I} u_i(t)\phi_i(x)$$

où I est fini. u_i sont les degrés de liberté qu'il faut donc déterminer et ϕ_i les fonctions d'une base de V_h.

Supposons par exemple que notre problème ait une formulation variationnelle (ou formulation faible) qui s'écrive à l'instant t :

$$\begin{array}{l} \text{Trouver } u \in V \text{ tel que} \\ (\frac{\partial u}{\partial t}, v) + a(u,v) = (f,v) \quad \forall v \in V \end{array} \tag{II.1}$$

où $a(.,.)$ est une forme bilinéaire sur $V \times V$ et $(f,.)$ est une forme linéaire sur V. Soit V_h, l'espace de discrétisation des éléments finis ie un espace approchant V, de dimension finie engendré par les fonctions $\phi_1, \cdots, \phi_I$. Le problème discrétisé associé à (II.1) est alors le suivant :

$$\begin{array}{l} \text{Trouver } (u_i)_{i=1,I} \text{ tel que} \\ \sum_{i=1}^{I} \frac{\partial u_i}{\partial t}(t)(\phi_i, \phi_j) + \sum_{i=1}^{I} u_i(t) a(\phi_i, \phi_j) = (f, \phi_j) \quad \forall j \in [1, I] \end{array} \tag{II.2}$$

C'est un système d'équations différentielles ordinaires pour lequel existent de nombreuses méthodes numériques (par exemple les schémas aux différences finis).

La méthode des éléments finis est donc entièrement définie par le choix de V_h et de ses fonctions de base ϕ_i. Pour cela, le domaine d'application est divisé en petits sous domaines ou "éléments" (généralement des triangles ou des quadrangles en 2D) et les fonctions u_h de V_h sont définies élément par élément.

Par exemple les éléments P1 (si les sous-domaines sont des triangles) ou Q1 (si les sous-domaines sont des quadrangles) sont tels que la restriction de u_h à un

[1] D'après O.C. Zienkiewicz [11]

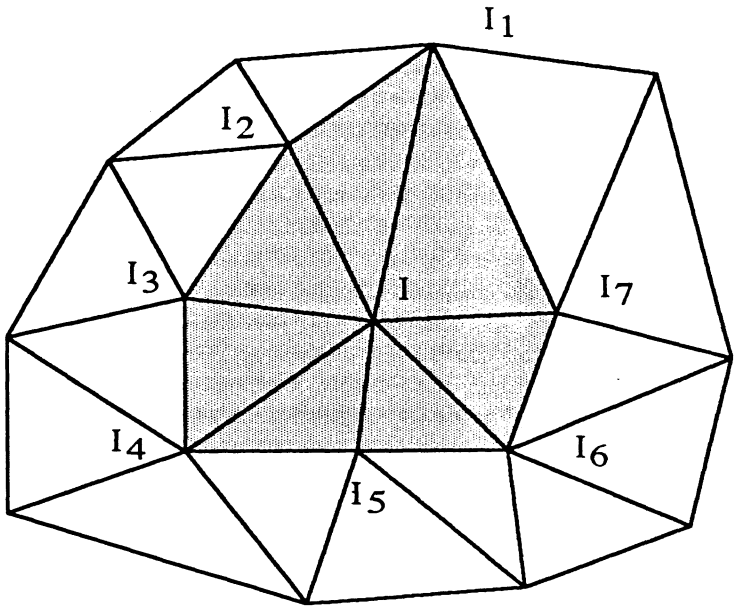


Figure II.1: *Support d'une fonction de base pour l'élément P1 : le support de ϕ_i est inclus dans la zone hachurée*

élément est un polynôme de degré (global si P1, par rapport à chaque variable si Q1) inférieur ou égal à 1. Dans ce cas la fonction ϕ_i de la base nodale est une fonction de V_h qui vaut 1 au "i^{eme}" noeud (un noeud étant un sommet de triangle ou de quadrangle) et 0 aux autres noeuds. L'existence d'une telle base est assurée par les propriétés telles que l'unisolvance des éléments P1 ou Q1.

Les contributions des éléments sont très localisées ; pour l'élément P1 par exemple le support de la fonction ϕ_i associée au "i^{eme}" noeud se réduit à quelques triangles (cf. figure II.1). Les matrices associées au système discret (II.2)

$$(a(\phi_i, \phi_j) + \frac{(\phi_i, \phi_j)}{\Delta t})_{(i,j)\in I^2}$$

sont alors creuses, réduisant considérablement la taille mémoire nécessaire a priori.

Pour les éléments finis de Lagrange, les degrés de liberté u_i sont les valeurs des fonctions inconnues aux noeuds. Pour les éléments finis de Hermite, les degrés de liberté sont, outre les valeurs des fonctions, les valeurs des dérivés. Il est à noter que pour les éléments mixtes (cf Thomasset [4]), le gradient de la vitesse $grad\, u$ est considéré comme une variable indépendante (on a 2 approximations ; une pour u et une pour $grad\, u$). Ces éléments sont fréquemment utilisés pour les problèmes de Stokes et Navier–Stokes.

Le choix des noeuds est lié au degré des polynomes et à la régularité de V_h. Si u_h est continue alors on parle d'éléments finis conformes. Dans le cas contraire il s'agit d'éléments non conformes et il faut alors tenir compte des contributions dues aux interfaces des éléments.

II.2 Un problème modèle : la cavité entrainée.

Ce paragraphe concerne la description d'un problème modèle qui permet d'étudier et de tester les codes numériques résolvant les équations de Navier–Stokes incompressibles : il s'agit de l'écoulement d'un fluide dans une cavité carrée $[0,1]^2$, les forces extérieures f étant nulles. Les conditions aux bords sont :

A)

$$u = \begin{pmatrix} u_1(x,t) \\ u_2(x,t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{sur les bords} \quad \begin{matrix} x_1 = 0, \\ x_1 = 1, \\ x_2 = 0, \end{matrix}$$

$$u = \begin{pmatrix} u_1(x,t) \\ u_2(x,t) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{sur le bord } x_2 = 1,$$

B)

$$u = \begin{pmatrix} u_1(x,t) \\ u_2(x,t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{sur les bords} \quad \begin{matrix} x_1 = 0, \\ x_1 = 1, \\ x_2 = 0, \end{matrix}$$

$$u = \begin{pmatrix} u_1(x,t) \\ u_2(x,t) \end{pmatrix} = \begin{pmatrix} (1-(2x_1-1)^2)^2 \\ 0 \end{pmatrix} \quad \text{sur le bord } x_2 = 1.$$

Le mouvement du fluide est dû à la distribution de la vitesse sur le bord supérieur d'où la dénomination "cavité entrainée" ou en anglais "wall driven forced cavity". La figure II.2 donne la configuration de cette cavité et quelques nomenclatures typiques du problème.

Le choix de B) au lieu de A) permet de régulariser l'écoulement dans la cavité en supprimant les singularités dans les 2 coins supérieurs dues à la discontinuité des conditions au bord. On peut observer la présence de ces singularités dans le cas A) avec la représentation du champ de gradient de la pression en figure II.3. Ce gradient est en effet très élevé dans les coins supérieurs pour la cavité A) à l'inverse de la cavité B). Il est à noter que les nombres de Reynolds (cf. Shen [28]) ne sont pas les mêmes (physiquement) dans les 2 cas puisque les distributions de la vitesse sur le bord supérieur sont différentes.

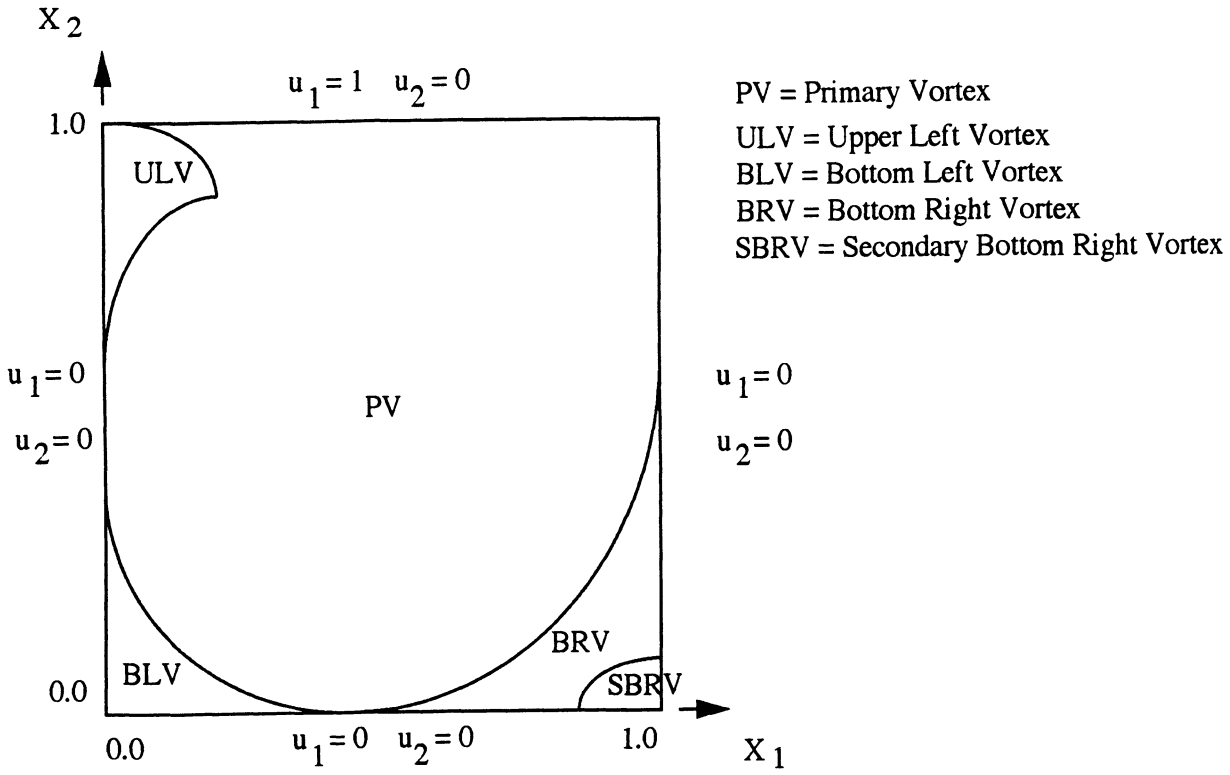


Figure II.2: *Configuration de la cavité entrainée, coordonnées, conditions au bord et nomenclature.*

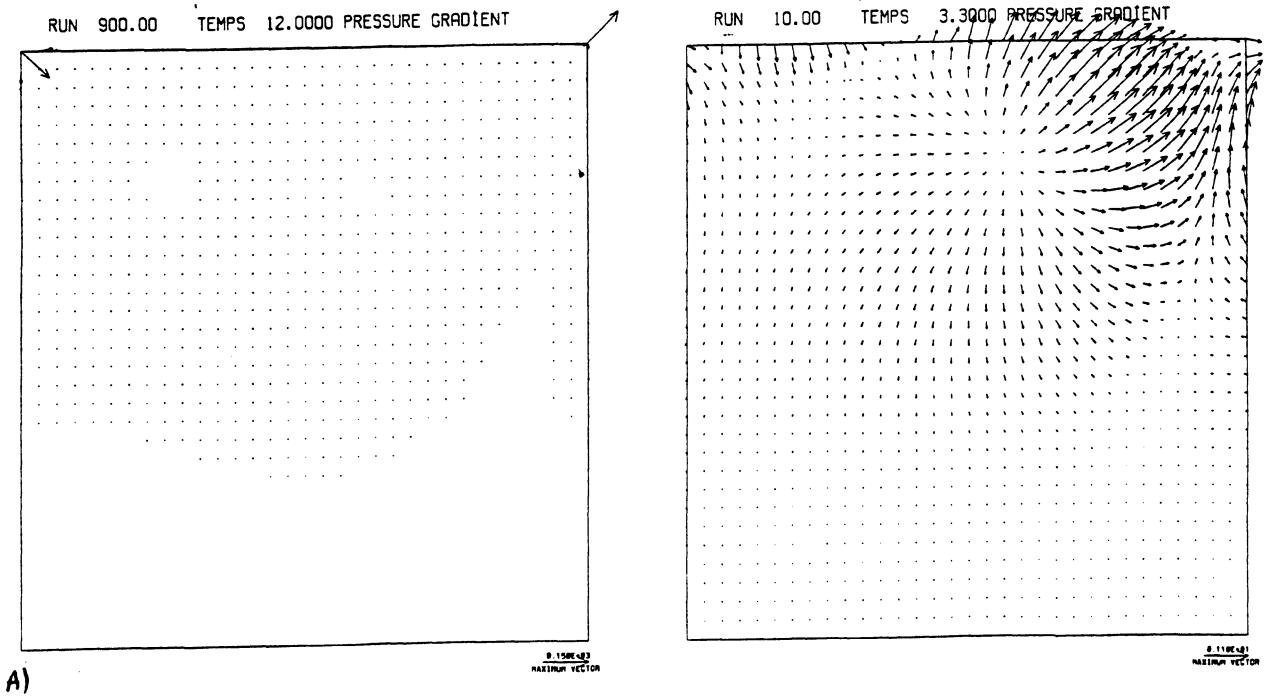


Figure II.3: *Gradient de la pression pour* $Re = 100$ *: a)* $u_1(x_1, 1) = 1$ *; b)* $u_1(x_1, 1) = (1 - (2x_1 - 1)^2)^2$.

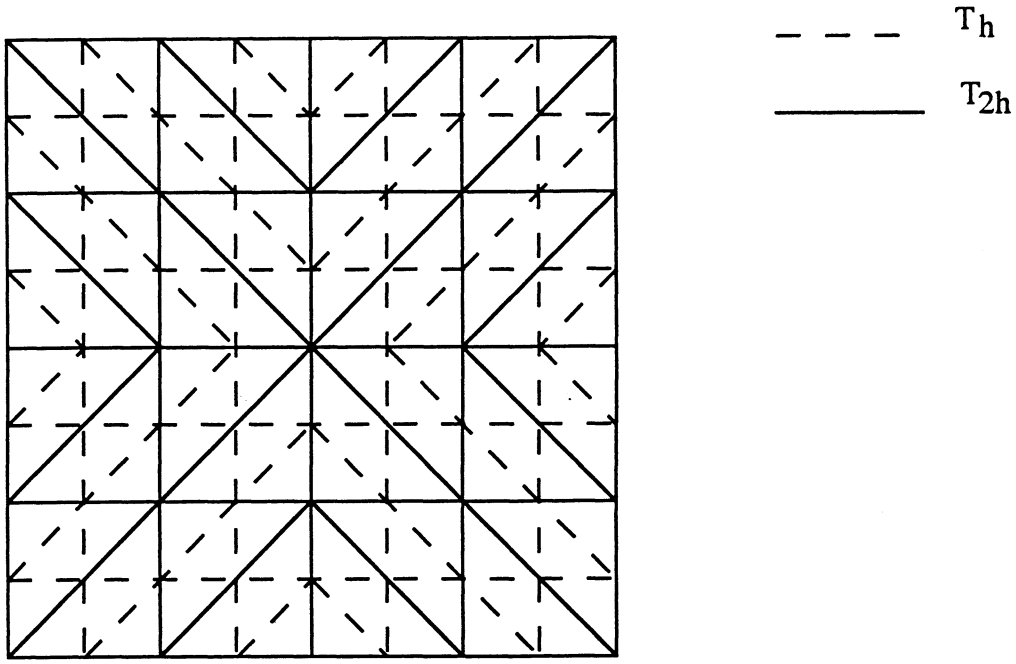


Figure II.4: *Triangulation de la cavité.*

Le maillage est une triangulation régulière (voir figure II.4) bien qu'un raffinement ou resserrement des mailles serait souhaitable au voisinage du bord supérieur pour mieux simuler les singularités. La forme en X permet à tout triangle τ du maillage d'avoir au moins un sommet à l'intérieur de Ω, condition nécessaire pour assurer la convergence de la discrétisation (voir la section II.3.1).

Afin de permettre une étude qualitative et quantitative du problème modèle, les résultats des simulations numériques seront présentés avec les outils suivants(cf. Thomasset [4], Goodrich *et al* [27]) :

- les lignes de courant (ensemble des points (x_1, x_2) du domaine Ω tels que $\psi(x_1, x_2)$ soit constant) : elles permettent de comparer qualitativement les résultats et de suivre l'évolution dynamique de l'écoulement, en particulier l'apparition et la disparition des tourbillons secondaires dits "recirculations" ou "vortex" ou "contre tourbillons". 2 types de représentation sont possibles
  - les lignes sont régulierement espacées, les valeurs variant du minimum au maximum de la fonction de courant.
  - les valeurs des lignes sont tabulés (cf tableau II.1),
- les lignes de vorticité (ensemble des points (x_1, x_2) tels que $\omega(x_1, x_2)$ soit constant) : elles sont extrèmement importantes dans la représentation des champs turbulents. Ce type de graphique permet de dépister d'éventuels défauts de résolution. Les lignes -6.0, $-5.0, \cdots, 0.0, \cdots, 5.0$, 6.0 sont représentées.

Valeurs de ψ							
a	-0.11	b	-0.1	c	-0.08	d	-0.06
e	-0.04	f	-0.02	g	$-1\ 10^{-2}$	h	$-3\ 10^{-3}$
i	$-1\ 10^{-3}$	j	$-3\ 10^{-4}$	k	$-1\ 10^{-4}$	l	$-3\ 10^{-5}$
m	$-1\ 10^{-5}$	n	$-1\ 10^{-6}$	o	$-1\ 10^{-7}$	p	$-1\ 10^{-8}$
q	$-1\ 10^{-9}$	r	$-1\ 10^{-10}$	s	0.		
α	$1\ 10^{-10}$	β	$1\ 10^{-9}$	γ	$1\ 10^{-8}$	δ	$1\ 10^{-7}$
ϵ	$1\ 10^{-6}$	ζ	$1\ 10^{-5}$	η	$1\ 10^{-4}$	θ	$3\ 10^{-4}$
ι	$1\ 10^{-3}$	κ	$3\ 10^{-3}$	λ	$1\ 10^{-2}$		

Tableau II.1: *Valeurs tabulées des lignes de courant*

- les contours 0.0025, 0.0050, 0.0075,... ,0.5 de l'énergie cinétique définie ponctuellement par

$$E(x_1, x_2) = \frac{1}{2}(u_1(x_1, x_2)^2 + u_2(x_1, x_2)^2).$$

- le champ des vitesses permet une excellente visualisation qualitative de l'écoulement.
- le champ des vitesses normalisées permet d'apprécier qualitativement les structures de petites tailles.
- les isobares (régulièrement distantes) révèlent les singularités.
- le gradient de la pression (normalisé et non normalisé) donne une idée de la précision de la résolution et révèle les tourbillons principaux comme "source" du champ.
- la localisation du tourbillon principal $(\bar{x}_1, \bar{x}_2)$, la valeur en ce point de la fonction de courant : $\psi(\bar{x}_1, \bar{x}_2)$ (extremum local de ψ) et du tourbillon $\omega(\bar{x}_1, \bar{x}_2)$; la localisation et les valeurs des tourbillons secondaires (extrmums locaux de ψ) permettent de comparer quantitativement les résultats numériques.
- le profil médian vertical de la vitesse horizontale ie la variation le long de l'axe vertical $x_1 = 0.5$ de la composante horizontale de la vitesse ie u_1.
- le profil médian horizontal de la vitesse verticale ie la variation le long de l'axe horizontal $x_2 = 0.5$ de la composante verticale de la vitesse ie u_2.
- les extremums de $u_1(0.5, .)$ (umin), $u_2(., 0.5)$ (vmin et vmax) et leurs positions (ymin, xmin et xmax).

II.3 Résolution du problème de Stokes.

L'objet de cette section est de présenter la méthode des éléments finis mixtes 4P1-P1 employée pour approcher le problème de Stokes et de décrire un solveur efficace du problème discrétisé ainsi obtenu.

II.3.1 Introduction des éléments 4P1-P1.

Formulation continue du problème.

Désormais on suppose que le domaine Ω est un domaine borné polygonal de $\mathbb{R}^2$ (hypothèse $H1$) et l'on suppose toujours que la condition au bord est :

$$u = g \text{ sur } \partial\Omega$$

On considère la formulation variationnelle (introduite en I.1.2) associée au problème de Stokes généralisé :

$$\begin{array}{l} \text{Trouver } (u,p) \in (H_g^1(\Omega))^2 \times L^2(\Omega) \text{ tel que} \\ \qquad \begin{array}{rcll} \alpha(u,v) + \frac{1}{Re}(\nabla u, \nabla v) - (p, \mathrm{div} v) & = & (f,v) & \forall v \in (H_0^1(\Omega))^2 \\ (q\, \mathrm{div} u) & = & 0 & \forall q \in L^2(\Omega) \end{array} \end{array} \tag{II.3}$$

que l'on peut réécrire :

$$\begin{array}{ll} a(u,v) - b(v,p) = (f,v) & \forall v \in (H_0^1(\Omega))^2 \\ b(u,q) = 0 & \forall q \in L^2(\Omega) \end{array} \tag{II.4}$$

avec

$$a(\phi,\psi) = \alpha(\phi,\psi) + \frac{1}{Re}(\nabla\phi, \nabla\psi) = \alpha(\phi,\psi) + \frac{1}{Re}((\phi,\psi))$$

$$b(\phi,q) = (q, \mathrm{div}\phi).$$

Il est à noter que le problème de Stokes (II.4) peut être mis sous la forme d'un problème de point selle :

$$\begin{array}{l} \text{trouver } (u,p) \in (H_g^1(\Omega))^2 \times L^2(\Omega) \text{ tel que} \\ \frac{1}{2}a(u,u) - b(u,p) - (f,u) = \min_{v \in (H_g^1(\Omega))^2} \max_{q \in L^2(\Omega)} \{\frac{1}{2}a(v,v) - b(v,q) - (f,v)\} \end{array} \tag{II.5}$$

ou encore d'un problème de minimisation avec contraintes :

$$\begin{array}{l} \text{trouver } u \in (H_g^1(\Omega))^2 \text{ tel que} \\ \frac{1}{2}a(u,u) - (f,u) = \min_{v \in \{v \in (H_g^1(\Omega))^2 \ / \ b(v,q)=0, \ \forall q \in L^2(\Omega)\}} \{\frac{1}{2}a(v,v) - (f,v)\} \end{array} \tag{II.6}$$

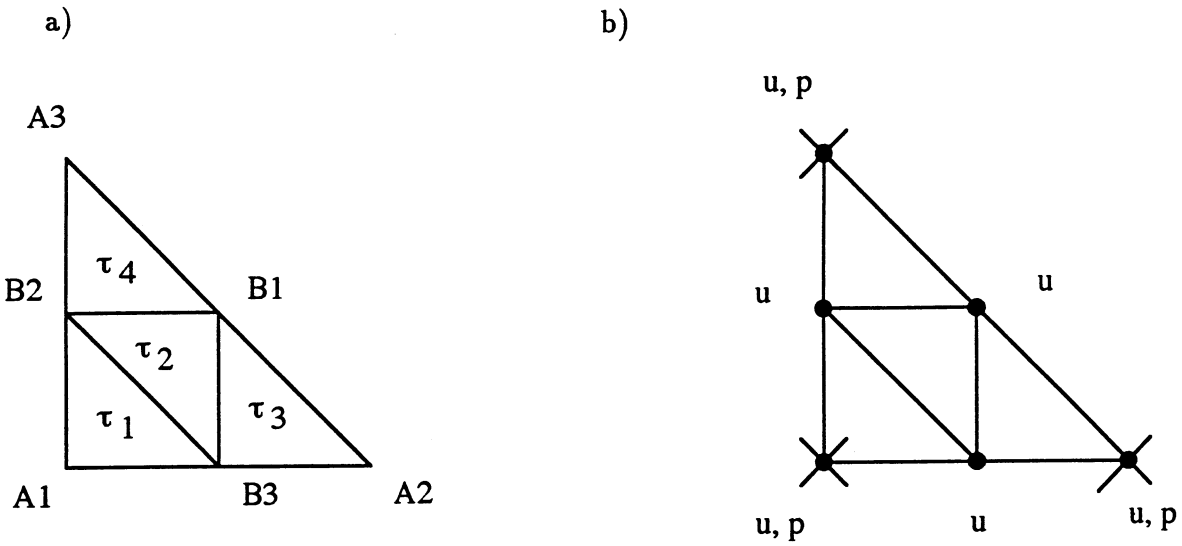


Figure II.5: a) *Subdivision d'un triangle τ en 4 sous triangles ; b) Position des degrés de liberté pour l'élément 4P1-P1.*

Triangulation

Pour définir l'élément fini lagrangien, conforme, mixte 4P1-P1, on considère une triangulation régulière T_{2h} (hypothèse $H2$) de Ω c'est-à-dire un ensemble de triangles ($2h$ étant le plus grand côté possible de ces triangles, paramètre qui tend vers 0) tels que :

- $\tau_i \cap \tau_j$ est un côté, un sommet ou l'ensemble vide pour tout triangle τ_i, τ_j de T_{2h} avec $i \neq j$,
- $\cup_i \tau_i = \Omega_h$ avec $\Omega_h = \Omega$,
- Pour tout τ_i, le diamètre de τ_i (ie le plus petit diamètre des cercles contenant τ_i) noté $\rho(\tau_i)$ et la rondeur de τ_i (ie le plus grand diamètre des cercles inscrits dans τ_i) noté $\rho'(\tau_i)$ vérifient :

$$\frac{\rho(\tau_i)}{\rho'(\tau_i)} \leq K,$$

K étant une constante indépendante de h.

On suppose de plus (hypothèse $H3$) que chaque triangle de T_{2h} a au moins un sommet n'appartenant pas à $\partial\Omega$. On définit T_h la triangulation obtenue à partir de la triangulation T_{2h} en divisant chaque triangle τ de T_{2h} en 4 sous triangles congruents (cf. figure II.5).

Espaces d'approximation

$P_1(\tau)$ étant l'ensemble des polynômes de degré inférieur ou égal à 1 sur τ, on note V_h et V_{2h} les espaces des fonctions continues linéaires par morceaux de Ω dans $\mathbb{R}$:

$$\begin{aligned} V_h &= \{v_h \in C^0(\Omega) \ / \ \forall \tau \in T_h, \ v_{h|\tau} \in P_1(\tau)\} \subset H^1(\Omega). \\ V_{2h} &= \{v_h \in C^0(\Omega) \ / \ \forall \tau \in T_{2h}, \ v_{h|\tau} \in P_1(\tau)\}. \end{aligned}$$

Soit $g_h \in V_h^2$ une approximation de g telle que

$$\int_{\partial\Omega} g_h.n \, d\sigma = 0,$$

les espaces d'éléments finis de dimension finie approchant $(H^1(\Omega))^2$, $(H^1_g(\Omega))^2$, $(H^1_0(\Omega))^2$ et $L^2(\Omega)$ sont alors respectivement :

$$\begin{aligned} \mathcal{V}_h &= V_h \times V_h \\ \mathcal{V}_{g,h} &= V_{g,h} \times V_{g,h} \\ &= \{v_h \in \mathcal{V}_h, \ v_h = g_h \text{ sur } \partial\Omega\} \times \{v_h \in \mathcal{V}_h, \ v_h = g_h \text{ sur } \partial\Omega\} \\ \mathcal{V}_{0,h} &= V_{0,h} \times V_{0,h} \\ &= (V_h \cap H^1_0(\Omega)) \times (V_h \cap H^1_0(\Omega)) \\ \mathcal{Q}_h &= V_{2h} \subset L^2(\Omega) \end{aligned}$$

Les degrés de liberté pour la vitesse sont donc les sommets A_i et les milieux des côtés B_i des triangles de T_{2h} (cf. figure II.5) et pour la pression les sommets des triangles de T_{2h}.

Formulation discrète du problème.

La version discrète de (II.3) s'écrit (S_h) :

$$\begin{aligned} &\text{Trouver } (u_h, p_h) \in \mathcal{V}_{g,h} \times \mathcal{Q}_h \text{ tel que} \\ &\alpha(u_h, v_h) + \frac{1}{Re}(\nabla u_h, \nabla v_h) - (p_h, \mathrm{div} v_h) = (f, v_h) \quad \forall v_h \in \mathcal{V}_{0,h} \\ &\qquad\qquad (q_h, \mathrm{div} u_h) = 0 \quad \forall q_h \in \mathcal{Q}_h, \end{aligned} \tag{II.7}$$

le problème de point selle étant discrétisé sous la forme :

$$\begin{aligned} &\text{trouver } u_h \in \mathcal{V}_{g,h} \text{ tel que} \\ &u_h = \min_{v_h \in \mathcal{V}_{g,h}} \max_{q_h \in \mathcal{Q}_h} \{\tfrac{1}{2} a(v_h, v_h) - b(v_h, q_h) - (f, v_h)\} \end{aligned} \tag{II.8}$$

Sous les hypothèses $H1$, $H2$, $H3$, les approximations $\mathcal{V}_{0,h}$ et $\mathcal{Q}_h$ de $(H^1_0(\Omega))^2$ et de $L^2(\Omega)$ vérifient la condition Inf-Sup, résultat technique démontré dans Bercovier et Pironneau [13] ie :

$$\begin{aligned} &\exists C \text{ indépendant de } h \text{ telle que} \\ &\sup_{v_h \in \mathcal{V}_{0,h} \setminus \{0\}} \frac{(v_h, \nabla q_h)}{(v_h, v_h)^{1/2}} \geq C(\nabla q_h, \nabla q_h)^{1/2} \quad \forall q_h \in \mathcal{Q}_h \end{aligned} \tag{II.9}$$

Cette relation assure l'existence et l'unicité de la solution de (S_h) dans $\mathcal{V}_{g,h} \times (\mathcal{Q}_h/\mathbb{R})$ ie la pression p_h est déterminée à une constante additive près.

Remarque II.1 *Il existe des méthodes directes utilisant par exemple des méthodes d'optimisation pour démontrer l'existence, en particulier à partir du problème de minimisation (II.6), mais l'unicité doit être traitée à part.*

La condition Inf-Sup permet également d'établir la convergence et l'ordre de la méthode. Pour la démonstration du théorème suivant, on se refèrera à l'article de Bercovier et Pironneau [13].

Théorème II.1 *Si les hypothèses $H1$, $H2$, $H3$ sont vérifiées et si $u \in (H^2(\Omega))^2$ et $p \in H^1(\Omega)/\mathbb{R}$ alors*

$$\begin{aligned} ||u_h - u||_{(H^1(\Omega))^2} &\leq Ch(||u||_{(H^2(\Omega))^2} + ||p||_{H^1(\Omega)/\mathbb{R}}) \\ ||p_h - p||_{L^2(\Omega)/\mathbb{R}} &\leq C'h(||u||_{(H^2(\Omega))^2} + ||p||_{H^1(\Omega)/\mathbb{R}}) \end{aligned}$$

où

$$||p||_{X/\mathbb{R}} = \inf_{\lambda \in \mathbb{R}} ||p + \lambda||_X.$$

Choix du 4P1-P1.

Les principales raisons pour lesquelles de nombreux auteurs ont choisi un tel élément sont résumées dans la suite :

1. L'élément 4P1-P1 vérifie la condition Inf-Sup (à l'inverse par exemple de l'élément P1-P0) qui traduit la compatibilité entre la divergence discrète et la pression discrète.

2. L'élément 4P1-P1 est conforme en vitesse (à l'inverse de l'élément P1 non conforme - P0 proposé par Crouzeix-Raviart [8]) et conforme en pression (ce qui n'est pas le cas pour P2-P0 proposé par Fortin [9]).

3. La construction de la base, ainsi que l'assemblage de la matrice du problème discret est facile (ce n'est plus le cas pour des bases à divergence nulle ou pour l'élément P2-P1 proposé par Taylor et Hood [10] ou pour l'élément P1 bulle - P1 d'après Arnold *et al* [6]).

4. Par rapport à P2-P1, la matrice est plus creuse, cela réduit l'encombrement mémoire et reste donc avantageux bien que l'ordre de convergence soit inférieur.

Remarque II.2 Dimension asymptotique des espaces de discrétisation

Nous allons vérifier formellement que la discrétisation proposée a un sens. Montrons que la dimension de l'espace d'approximation par éléments finis 4P1-P1 associé au problème de Stokes est strictement positive.

Si la frontière n'a qu'une seule composante connexe, l'identité d'Euler est pour les triangles :

$$S - C + T = 1$$

où S est le nombre de sommets, C le nombre de côtés et T le nombre de triangles dans la triangulation T_{2h}. Etant donné qu'il y a 3 côtés par triangle et que chaque côté appartient à 2 triangles, asymptotiquement on a la relation :

$$C \simeq \frac{3T}{2}$$

et par suite

$$S \simeq C - T = \frac{T}{2}.$$

On en déduit que

$$dim\ \mathcal{V}_h \simeq 2S + 2C \simeq T + 3T = 4T$$

$$dim\ \mathcal{Q}_h \simeq S = \frac{T}{2}.$$

Or les contraintes $(div\, v_h, q_h) = 0, \quad \forall q_h \in \mathcal{Q}_h$ sont indépendantes sauf une. En effet

$$(v_h, \nabla q_h) = 0 \quad \forall v_h \in \mathcal{V}_h \Longrightarrow q_h = \text{constante}.$$

On peut donc estimer la dimension de l'espace de discrétisation associé au problème de Stokes :

$$dim\ \{v \in \mathcal{V}_h \ /\ (div\, v_h, q_h) = 0 \quad \forall q_h \in \mathcal{Q}_h\} \simeq 4T - \frac{T}{2} \simeq \frac{7T}{2}. \qquad \text{(II.10)}$$

Remarque II.3 *Une formulation différente du problème de Stokes avec l'élément 4P1-P1 a été proposée et démontrée par Glowinski et Pironneau [14]. Cette formulation équivalente induit des schémas numériques différents.*

II.3.2 Méthode d'Uzawa et préconditionnement.

Forme matricielle du problème (S_h).

Notons Σ_h l'ensemble des sommets et des milieux des côtés des triangles de T_{2h}, Σ_{2h} l'ensemble des sommets des triangles de T_{2h} et posons :

$$\overset{\circ}{\Sigma}_h = \Sigma_h \cap \overset{\circ}{\Omega}$$

l'ensemble des sommets et milieux de côtés intérieurs à Ω,

$$\overset{\circ}{\Sigma}_{2h} = \Sigma_{2h} \cap \overset{\circ}{\Omega}$$

l'ensemble des sommets intérieurs à Ω.

On pose :

$$n_h = \text{Card}\,(\Sigma_h) \quad \text{et} \quad n_{0,h} = \text{Card}\,(\mathring{\Sigma}_h),$$
$$n_{2h} = \text{Card}\,(\Sigma_{2h})$$

On supposera que les noeuds intérieurs à Ω sont les points M_j pour j variant de 1 à $n_{0,h}$ et que les noeuds appartenant à $\partial\Omega$ sont les points M_j pour j variant de $n_{0,h}+1$ à n_h.

La base canonique dite nodale de V_h (respectivement V_{2h}) est notée $\{\phi_{h,i}\}_{i=1}^{n_h}$ (respectivement $\{\phi_{2h,i}\}_{i=1}^{n_{2h}}$). Elle est définie de la façon suivante :

$$\begin{array}{ll} \phi_{h,i} \in V_h & \forall i = 1, \cdots n_h \\ \phi_{h,i}(M_i) = 1 & \\ \phi_{h,i}(M_j) = 0 & \forall i,j = 1, \cdots, n_h \text{ avec } i \neq j \text{ et } \quad (M_i, M_j) \in \Sigma_h^2. \end{array}$$

La définition est rigoureusement identique pour $\phi_{2h,i}$.

Si l'on écrit le système de Stokes discrétisé pour les fonctions de base de $\mathcal{V}_{0,h}$ et $\mathcal{Q}_h$, alors (S_h) est équivalent au système linéaire suivant :

$$\begin{pmatrix} A_h & 0 & -B_{1,h}^t \\ 0 & A_h & -B_{2,h}^t \\ -B_{1,h} & -B_{2,h} & 0 \end{pmatrix} \begin{pmatrix} U_{1,h} \\ U_{2,h} \\ P_h \end{pmatrix} = \begin{pmatrix} F_{1,h} \\ F_{2,h} \\ 0 \end{pmatrix} \qquad \text{(II.11)}$$

où les matrices A_h, $B_{1,h}$ et $B_{2,h}$ sont définies par :

$$\begin{aligned} (A_h)_{ij} &= a(\phi_{h,i}, \phi_{h,j}) && 1 \leq i,j \leq n_{0,h} \\ (B_{1,h})_{ij} &= (\phi_{2h,i}, \frac{\partial \phi_{h,j}}{\partial x_1}) && 1 \leq i \leq n_{2h}, 1 \leq j \leq n_{0,h} \\ (B_{2,h})_{ij} &= (\phi_{2h,i}, \frac{\partial \phi_{h,j}}{\partial x_2}) && 1 \leq i \leq n_{2h}, 1 \leq j \leq n_{0,h} \end{aligned}$$

et les vecteurs colonnes $U_{1,h}$, $U_{2,h}$, $F_{1,h}$, $F_{2,h}$ et P_h sont les composantes de la vitesse, des forces extérieures et de la pression exprimées dans la base nodale de $V_{o,h}$ et V_{2h} :

$$U_{1,h} = \begin{pmatrix} u_{1,1} \\ . \\ . \\ . \\ u_{1,n_{0,h}} \end{pmatrix}, \quad U_{2,h} = \begin{pmatrix} u_{2,1} \\ . \\ . \\ . \\ u_{2,n_{0,h}} \end{pmatrix}$$

$$F_{1,h} = \begin{pmatrix} f_{1,1} \\ . \\ . \\ . \\ f_{1,n_{0,h}} \end{pmatrix}, \quad F_{2,h} = \begin{pmatrix} f_{2,1} \\ . \\ . \\ . \\ f_{2,n_{0,h}} \end{pmatrix}$$

$$P_h = \begin{pmatrix} p_1 \\ . \\ . \\ . \\ p_{n_{2h}} \end{pmatrix}$$

Comme la condition au bord du problème discret est $u_{n,h} = g_{n,h}$ pour $n = 1, 2$, il faut en tenir compte dans le système linéaire et tout particulièrement dans le second membre $f_{n,i}$ qui est déterminé par la relation :

$$f_{n,i} = \sum_{j=1}^{n_{0,h}} f_n(M_j)(\phi_{h,j}, \phi_{h,i}) - \sum_{j=n_{0,h}+1}^{n_h} g_{n,h}(M_j) a(\phi_{h,j}, \phi_{h,i}).$$

Enfin $u_{1,i}$, $u_{2,i}$, p_i sont tels que :

$$u_h = \begin{pmatrix} \sum_{i=1}^{n_{0,h}} u_{1,i}\phi_{h,i} + \sum_{i=n_{0,h}+1}^{n_h} g_{1,h}(M_i)\phi_{h,i} \\ \sum_{i=1}^{n_{0,h}} u_{2,i}\phi_{h,i} + \sum_{i=n_{0,h}+1}^{n_h} g_{2,h}(M_i)\phi_{h,i} \end{pmatrix}$$

$$p_h = \left(\sum_{i=1}^{n_{2h}} p_i \phi_{h,i} \right)$$

$u_{1,i}$, $u_{2,i}$ et p_j étant des valeurs approchées de $u_1(M_i)$, $u_2(M_i)$ et $p(M_j)$ pour $1 \leq i \leq n_{0,h}$ et $1 \leq j \leq n_{2h}$.

Le système (II.11) a une dimension $(2 \times n_{0,h} + n_{2h})^2$ bien trop grande pour envisager une méthode d'inversion directe. Seule une méthode itérative est appropriée, cependant la matrice n'est pas définie, car son rang est égal à $(2 \times n_{0,h} + n_{2h})^2 - 1$. A_h étant une matrice symétrique, définie, positive, et donc inversible, il semble donc plus adéquat de résoudre le système (II.11) réécrit sous la forme :

$$\begin{aligned} U_{1,h} &= A_h^{-1}(F_{1,h} + B_{1,h}^t P_h) \\ U_{2,h} &= A_h^{-1}(F_{2,h} + B_{2,h}^t P_h) \\ (B_{1,h}A_h^{-1}B_{1,h}^t + B_{2,h}A_h^{-1}B_{2,h}^t)P_h &= -B_{1,h}A_h^{-1}F_{1,h} - B_{2,h}A_h^{-1}F_{2,h} \end{aligned} \qquad \text{(II.12)}$$

La méthode d'Uzawa-Hurwicz décrite ci-dessous consiste à résoudre le système linéaire (II.12) par un gradient conjugué preconditionné. En effet la matrice $B_{1,h}A_h^{-1}B_{1,h}^t + B_{2,h}A_h^{-1}B_{2,h}^t$ est elle-même symétrique, définie positive.

Algorithme d'Uzawa.

Dans la suite, pour alléger les notations, l'indice h sera supprimé. On suppose que la matrice de préconditionnement M est une matrice facile à inverser. Les étapes de l'algorithme d'Uzawa sont les suivantes :

1. **Etape 1 : Initialisation**

   - P^0 est donné quelconque

- résoudre $AU_1^0 = (F_1 + B_1^t P^0)$
- résoudre $AU_2^0 = (F_2 + B_2^t P^0)$
- poser $g^0 = -B_1 U_1^0 - B_2 U_2^0$
- résoudre $Md^0 = g^0$

2. **Etape 2 :** P^k, U_1^k, U_2^k, g^k, d^k **étant connues**

- résoudre $AV_1 = B_1^t d^k$
- résoudre $AV_2 = B_2^t d^k$
- poser $\rho = \dfrac{(g^k, d^k)}{(d^k, B_1 V_1 + B_2 V_2)}$

3. **Etape 3 : Détermination de la nouvelle direction de descente**

- poser $P^{k+1} = P^k + \rho d^k$
- poser $U_1^{k+1} = U_1^k + \rho V_1 = U_1^k + \rho A^{-1} B_1^t d^k$
- poser $U_2^{k+1} = U_2^k + \rho V_2 = U_2^k + \rho A^{-1} B_2^t d^k$
- poser $g^{k+1} = g^k - \rho(B_1 V_1 + B_2 V_2) = g^k - \rho(B_1 A^{-1} B_1^t d^k + B_2 A^{-1} B_2^t d^k)$
- résoudre $Mz^{k+1} = g^{k+1}$
- poser $d^{k+1} = z^{k+1} + \dfrac{(g^{k+1}, z^{k+1})}{(g^k, z^k)} d^k$

4. **Etape 4 :** Retourner à l'étape 2 (ϵ est un paramètre donné)

- si $||B_1 U_1^{k+1} + B_2 U_2^{k+1}||_{2h} \geq \epsilon$ (test sur la divergence)
- ou si

$$\frac{||P^{k+1} - P^k||_{2h}}{||P^k||_{2h}} \geq \epsilon \quad \text{test sur la pression}$$

$$\frac{||U^{k+1} - U^k||_h}{||U^k||_h} \geq \epsilon \quad \text{test sur la vitesse}$$

où les normes $||.||_h$ et $||.||_{2h}$ sont données par :

$$||P||_{2h} = \begin{cases} \dfrac{\sum_{i=1}^{n_{2h}} |P_i|}{n_{2h}} \\ \text{ou} \\ \sqrt{\dfrac{\sum_{i=1}^{n_{2h}} P_i^2}{n_{2h}}} \end{cases} \qquad ||U||_h = \begin{cases} \dfrac{\sum_{i=1}^{n_h} (|U_{1,i}| + |U_{2,i}|)}{n_h} & (\text{norme } L_1) \\ \text{ou} \\ \sqrt{\dfrac{\sum_{i=1}^{n_h} (U_{1,i}^2 + U_{2,i}^2)}{n_{2h}}} & (\text{norme } L_2) \end{cases}$$

Il s'agit en fait de la résolution du problème du point selle (II.8) associé à Stokes par une méthode de gradient dont les étapes sont :

1. **Etape 1 : Initialisation**

   - $p_h^0 \in \mathcal{Q}_h$ est donné
   - on détermine u_h^0 comme étant solution du problème de Dirichlet :

$$u_h^0 \in \mathcal{V}_{g,h}$$

$$a(u_h^0, v_h) - (p_h^0, \operatorname{div} v_h) = (f, v_h) \quad \forall v_h \in \mathcal{V}_{0,h}$$

2. **Etape 2 :** Quand u_h^k et p_h^k sont connus, u_h^{k+1} et p_h^{k+1} sont définis par les relations :

$$\begin{aligned} u_h^{k+1} \in \mathcal{V}_{g,h}, \quad p_h^{k+1} \in \mathcal{Q}_h & \\ (p_h^{k+1} - p_h^k, q_h) = & \quad -\rho(\operatorname{div} u_h^k, q_h) \qquad \forall q_h \in \mathcal{Q}_h \\ a(u_h^{k+1}, v_h) - (p_h^{k+1}, \operatorname{div} v_h) = & \quad (f, v_h) \qquad \forall v_h \in \mathcal{V}_h \end{aligned}$$

   $-\operatorname{div} u_h^k$ est en effet le gradient au point p_h^k de la fonctionnelle :

$$p_h \mapsto \min_{v_h \in \mathcal{V}_h} \frac{1}{2} a(v_h, v_h) - (p_h, \operatorname{div} v_h) - (f, v_h).$$

On trouvera dans Temam [3] une démonstration directe de la convergence de u_h^k vers u_h et de p_h^k vers p_h. La méthode du gradient conjugué estime la valeur optimale de ρ.

Numériquement les étapes les plus coûteuses sont les étapes d'inversion des matrices et plus particuliérement de la matrice A effectuée lors de l'étape 2 pour chacune des composantes de la vitesse. En effet l'odre de A, $(n_{0,h})^2$, est nettement supérieur à l'ordre de M : $(n_{2h})^2$. L'inversion de A est réalisé par une méthode itérative du type gradient conjugué préconditionné par une décomposition de Choleski incomplet effectuée une fois pour toute au début du code numérique.

Un des avantages principaux de l'algorithme d'Uzawa est de découpler les 2 composantes de la vitesse, ce qui n'est pas le cas avec des méthodes du type lagrangien augmenté (cf Thomasset [4]).

Remarque II.4 *Dans le cas continue, la méthode d'Uzawa conserve la valeur moyenne de la pression initiale. En effet si* $\int_\Omega p^0 \, dx = 0$ *alors*

$$\int_\Omega p^{k+1} \, dx = \int_\Omega p^k \, dx - \rho \int_\Omega div \, u^k \, dx = \int_\Omega p^k \, dx = \int_\Omega p^0 \, dx$$

Dans le cas discret, si on suppose par exemple que $\sum_{i=1}^{n_{2h}} p_i^0 = 0$ *alors à l'étape 3, on pose :*

$$p_i^{k+1} = p_i^k + \rho(d_i^k - \frac{\sum_{j=1}^{n_{2h}} d_i^k}{n_{2h}}) \qquad 1 \leq i \leq n_{2h}.$$

Préconditionnement.

1. Le préconditionnement de Labadie
   Le préconditionnement de l'algorithme d'Uzawa a pour but d'accélérer la convergence. Or une matrice M de préconditionnement est d'une part une matrice facile à inverser et d'autre part la matrice inverse M^{-1} doit être proche de $B_{1,h}A_h^{-1}B_{1,h}^t + B_{2,h}A_h^{-1}B_{2,h}^t$ c'est-à-dire qu'elle doit vérifier :

   $$M^{-1}(B_{1,h}A_h^{-1}B_{1,h}^t + B_{2,h}A_h^{-1}B_{2,h}^t) \simeq Id$$

   où Id est la matrice identité.

   Formellement lorsque $\alpha \gg \frac{1}{Re}$, le terme de viscosité étant négligeable, on a $A \simeq \alpha Id$. On en déduit que

   $$B_{1,h}A_h^{-1}B_{1,h}^t + B_{2,h}A_h^{-1}B_{2,h}^t \simeq \frac{1}{\alpha}(B_{1,h}B_{1,h}^t + B_{2,h}B_{2,h}^t)$$

   et comme $B_{1,h}B_{1,h}^t + B_{2,h}B_{2,h}^t$ est une discrétisation élément fini du Laplacien, Labadie [7] propose de prendre

   $$(M)_{i,j} = (\nabla\phi_{2h,i}, \nabla\phi_{2h,j}) \quad 1 \leq i,j \leq n_{2h}.$$

   Dans ce cas, l'étape 2 de l'algorithme d'Uzawa consiste à résoudre le problème de Neumann avec des conditions au bord homogènes :

   Déterminer p_h^{k+1} tel que
   $(-\Delta(p_h^{k+1} - p_h^k), q_h) = -\rho(\operatorname{div} u_h^k, q_h) \quad \forall q_h \in \mathcal{Q}_h$

   Il est à noter que l'on trouve cette étape dans la méthode de projection proposée par Chorin [19] et Temam [3].

   Pour que la méthode soit efficace, il est cependant nécessaire que p soit suffisamment régulière (par exemple $p \in H^1(\Omega)$).

2. Le préconditionnement de Cahouet
   En s'appuyant sur un concept de multi-solveurs couplant

   - l'évaluation des gradients singuliers avec le schéma d'Uzawa sur le maillage fin c'est-à-dire sur le maillage associé à la vitesse,
   - l'évaluation rapide de la partie régulière avec un schéma de type Chorin-Temam sur un maillage grossier ie sur le maillage associé à la pression,

   Cahouet [7] propose de poser :

   $$M = (\frac{1}{Re}Id - \alpha\Delta^{-1})^{-1}.$$

Ainsi si (p_h^k, u_h^k) est la solution obtenue par Uzawa, une correction δp_h^k de la pression est calculée en utilisant un schéma de type Chorin-Temam ie :

$$\delta p_h^k = (\frac{1}{Re} Id - \alpha \Delta^{-1}) \mathrm{div}\, u_h^k.$$

Il est à noter que l'on retrouve le préconditionnement de Labadie lorsque $\alpha \gg \frac{1}{Re}$.

Ces préconditionnements accélèrent la convergence du schéma d'Uzawa, et leurs coûts restent raisonnables, puisque les problèmes supplémentaires à résoudre sont des problèmes de Neumann d'un ordre inférieur à ceux associés à A_h. La résolution des problèmes de Neumann est réalisé par un gradient conjugué préconditionné par une décomposition incomplète de Choleski calculée une fois pour toute en début de code, le préconditionnement permettant la convergence du gradient vers une solution particulière du problème (cf. non unicité de la solution des problèmes de Neumann).

II.3.3 Résultats numériques.

Problème de Stokes

Le processus itératif de l'algorithme d'Uzawa est initialisé par $P^0 = 0$. La figure II.6 représente les lignes de courant et de vorticité, les isobares et le contour de l'énergie cinétique : la solution est régulière et parfaitement symétrique par rapport à l'axe $x_1 = \frac{1}{2}$. La figure II.7 permet d'apprécier le nombre d'itérations nécessaires pour obtenir des erreurs relatives pour la pression et la vitesse en norme L_2 inférieures à $0.1\ 10^{-3}$ avec $Re = 1000$ et un pas d'espace

$$h = \frac{1}{129} = 0.0078 \quad \text{et} \quad h = \frac{1}{64} = 0.0156.$$

Afin de comparer les performances des schémas on a également représenté la norme L_2 de la divergence de la vitesse en fonction des itérations d'Uzawa. On remarque que les erreurs relatives et la norme de la divergence augmentent lorsque h diminue et lorsque le nombre de Reynolds augmente. (cf figure II.8)

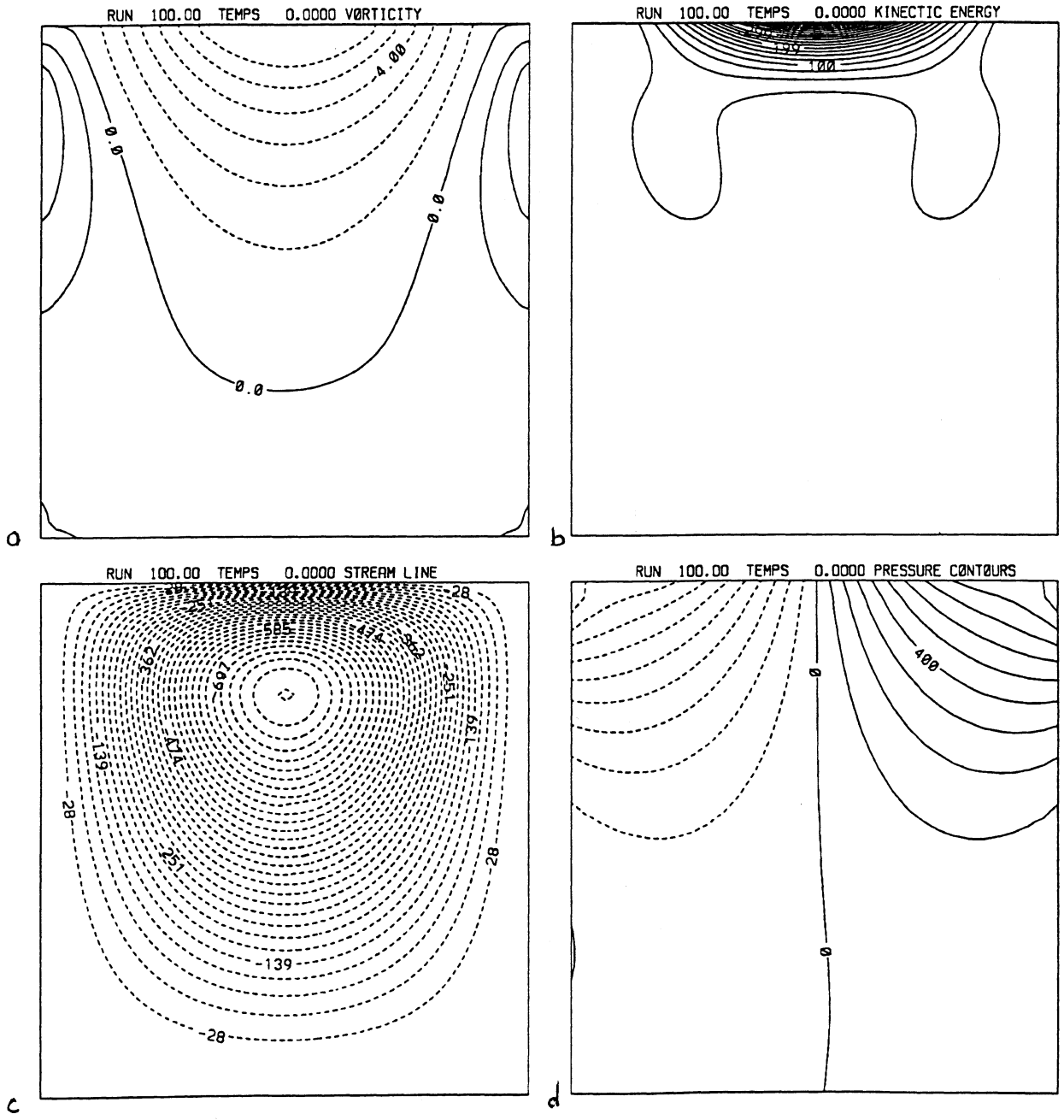


Figure II.6: *PROBLEME DE STOKES dans la cavité B) ; a) Isovorticité ; b) Energie ; c) Lignes de courant ; d) Isobares.*

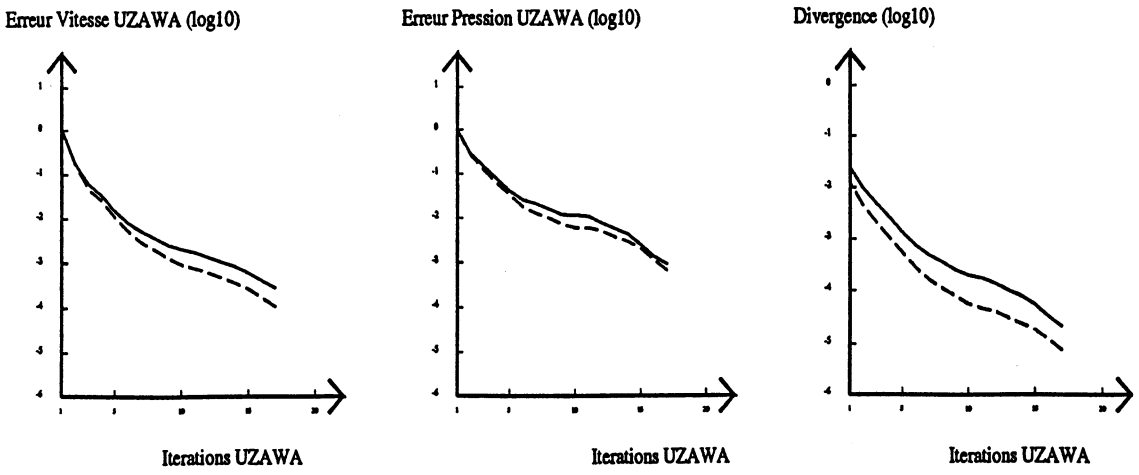


Figure II.7: *PROBLEME DE STOKES dans la cavité B) ; Courbes de convergence de la méthode d'Uzawa. En trait continu : maillage* 129×129 *; En trait pointillé : maillage* 65×65.

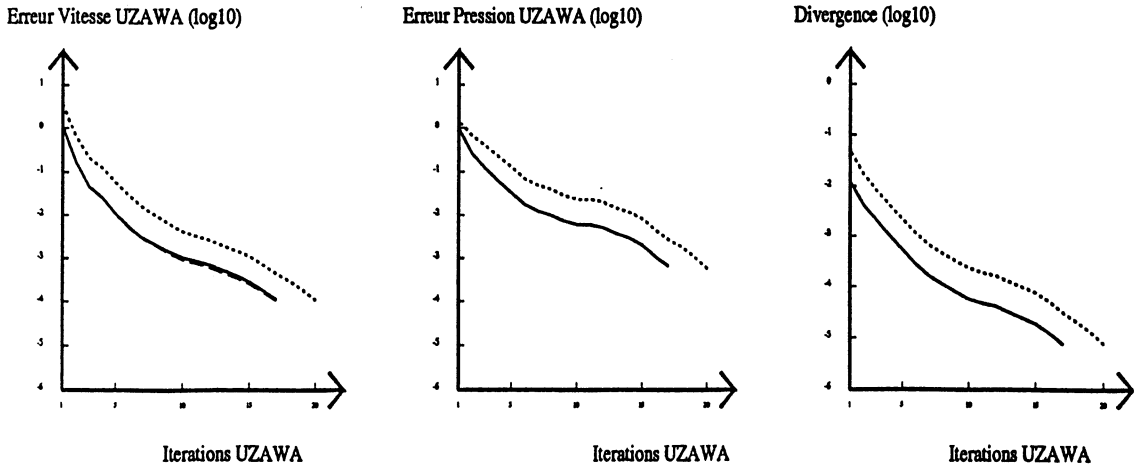


Figure II.8: *PROBLEME DE STOKES dans la cavité B) ; Courbes de convergence de la méthode d'Uzawa. En trait continu :* $Re = 100$ *; En trait discontinu :* $Re = 1000$ *; En trait pointillé :* $Re = 5000$.

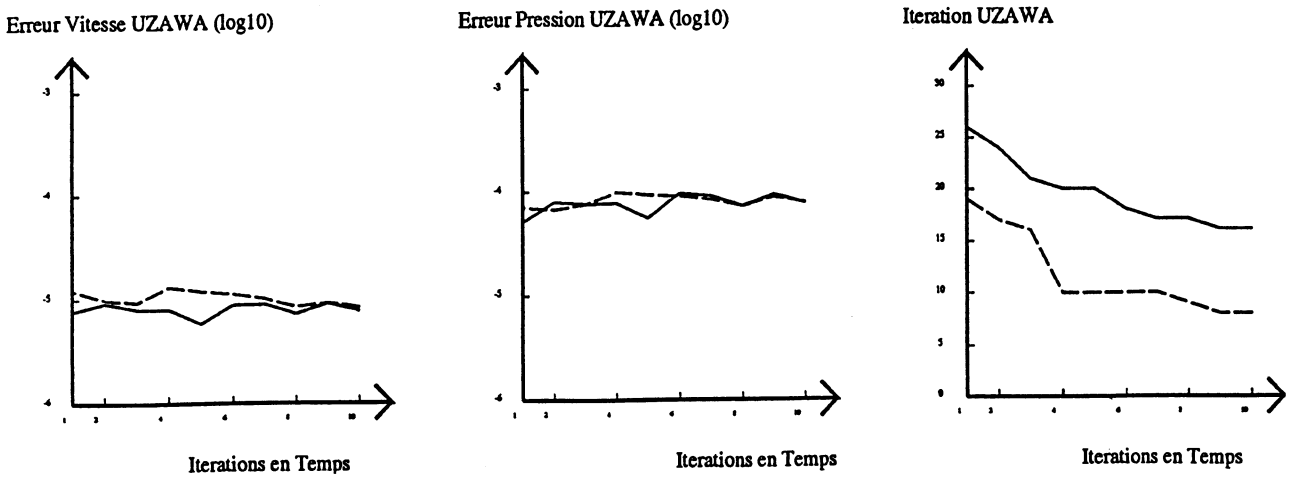


Figure II.9: *PROBLEME DE STOKES GENERALISE dans la cavité B) ; Erreurs et itérations de la méthode d'Uzawa pour 10 pas de temps de Navier-Stokes. En trait continue : sans préconditionnement ; En trait discontinue : préconditionnement de Labadie ; En trait pointillé : préconditionnement de Cahouet.*

Problème de Stokes généralisé

Les résultats sont comparables à ceux obtenus par Cahouet et Chabard. La figure II.9 représente, au cours des 10 premiers pas de temps du problème de Navier-Stokes, les erreurs relatives de la vitesse, de la pression et le nombre d'itérations pour la résolution du problème de Stokes généralisé par la méthode d'Uzawa. Comme

$$\alpha = 10 \gg \frac{1}{Re} = \frac{1}{100},$$

les préconditionnements de Labadie et Cahouet sont rigoureusement identiques et accélerent nettement la convergence d'Uzawa.

L'influence du paramètre α sur la solution (et particulièrement sur le champ de vitesse) est illustrée en figure II.10. Les conditions au bord sont celles de la cavité avec une vitesse d'entrainement égale à 1.0 et les forces extérieures sont nulles. Ce problème n'a pas de sens physique mais il est intéressant de constater que si α est grand, un très fort gradient de vitesse apparaît dans les couches supérieures du maillage.

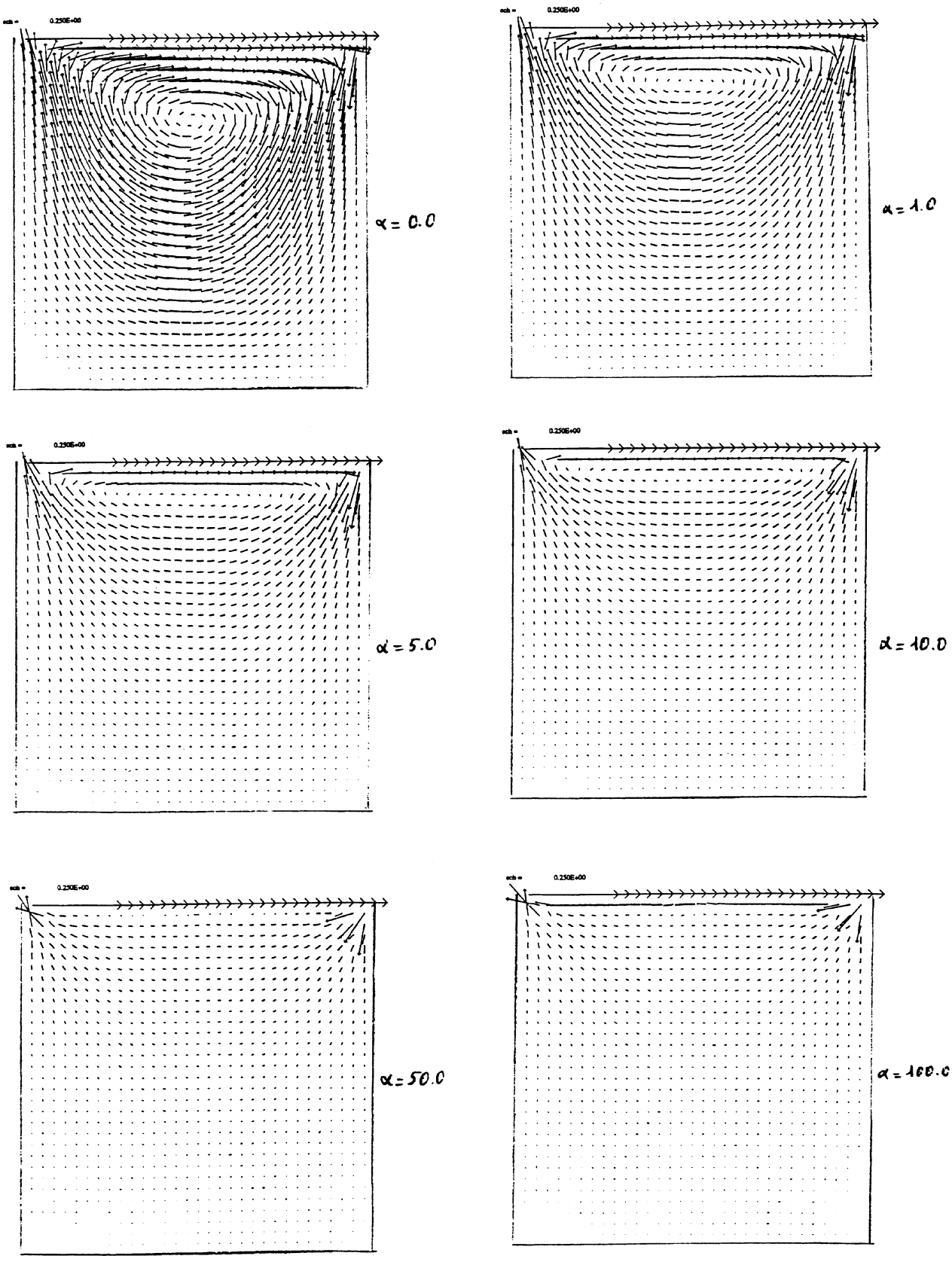


Figure II.10: *PROBLEME DE STOKES GENERALISE dans la cavité A) ; Champ de vitesse pour* $Re = 100$ *et* $\alpha = 0.0, 1.0, 5.0, 10.0, 50.0, 100.0$.

II.4 Résolution du problème de Navier–Stokes.

II.4.1 Discrétisation en espace et en temps.

Formulation discrète du problème

Les hypothèses $H1$, $H2$, $H3$ et les notations de la section II.3.1 sont conservées.

L'approximation par éléments finis mixtes 4P1-P1 du problème (N.S.) est la suivante (NS_h) (on se réfèrera à Le Tallec [15] pour une démonstration de la convergence de cette approximation) :

$$\begin{array}{l} \text{Trouver } (u_h(t), p_h(t)) \in \mathcal{V}_{g,h} \times \mathcal{Q}_h \text{ tel que} \\ \left(\frac{\partial u_h}{\partial t}, v_h\right) + \frac{1}{Re}(\nabla u_h, \nabla v_h) + b(u_h, u_h, v_h) - (p_h, \mathrm{div} v_h) = (f, v_h) \quad \forall v_h \in \mathcal{V}_{0,h} \\ (q_h, \mathrm{div} u_h) = 0 \quad \forall q_h \in \mathcal{Q}_h. \end{array} \tag{II.13}$$

Schéma en temps

Le problème étant discrétisé en espace, il s'agit de présenter maintenant une méthode de résolution adéquate en temps. 2 schémas par différences finis ont été envisagés :

1. Le premier schéma temporel envisagé est composé :
   - d'un schéma d'Adams-Bashforth explicite du second ordre pour les termes non linéaires,
   - d'un schéma de Crank-Nicholson semi-implicite du second ordre pour les termes linéaires.

   L'initialisation est effectuée par une méthode d'Euler :
   - Etape 1 : initialisation
     - $u_h^0 = u_{0,h}$ approximation de u_0 est donnée
     - u_h^1 et p_h^1 les approximations de u_h et p_h à l'instant Δt sont définies à partir de u_h^0 par :

$$\frac{(u_h^1 - u_h^0, v_h)}{\Delta t} + \frac{1}{2Re}(\nabla u_h^1, \nabla v_h) - (p_h^1, \mathrm{div} v_h) =$$

$$(f^1, v_h) - \frac{1}{2Re}(\nabla u_h^0, \nabla v_h) - ((u_h^0.\nabla)u_h^0, v_h)$$

$$(q_h, \mathrm{div} u_h^1) = 0 \qquad \forall v_h \in \mathcal{V}_{0,h} \quad \text{et} \quad \forall q_h \in \mathcal{Q}_h,$$

- Etape 2 : Pour $k \geq 1$, u_h^{k+1} et p_h^{k+1} sont définies à partir de u_h^k par:

$$\frac{(u_h^{k+1} - u_h^k, v_h)}{\Delta t} + \frac{1}{2Re}(\nabla u_h^{k+1}, \nabla v_h) - (p_h^{k+1}, \mathrm{div} v_h) =$$

$$(f^{k+1}, v_h) - \frac{1}{2Re}(\nabla u_h^k, \nabla v_h) - \left\{\frac{3}{2}((u_h^k.\nabla)u_h^k, v_h) - \frac{1}{2}((u_h^{k-1}.\nabla)u_h^{k-1}, v_h)\right\}$$

$$(q_h, \mathrm{div} u_h^{k+1}) = 0 \qquad \forall v_h \in \mathcal{V}_{0,h} \qquad \text{et} \qquad \forall q_h \in \mathcal{Q}_h,$$

Ce schéma d'ordre 2 a l'inconvénient d'être conditionnellement stable. Il est soumis à une condition CFL du type (le terme de transport, non linéaire est traité implicitement) :

$$||u_h|| \frac{\Delta t}{h} \leq cst,$$

ce qui rend ce shéma très couteux en temps de calcul dès que le nombre de Reynolds est grand et que h est par conséquent petit.

Il est à noter que chacune des étapes de cet algorithme consiste à résoudre un problème de Stokes généralisé, d'où la nécessité d'avoir un solveur performant de Stokes tel qu'un Uzawa préconditionné. A l'étape 2 le schéma d'Uzawa est initialisé par la pression obtenue à l'étape précédente : p_h^k.

2. Le deuxième schéma temporel envisagé est un schéma de splitting (appelé également dans la littérature schéma de Peaceman-Rachford ou θ schéma). Cette méthode de splitting utilisée par Glowinski et Periaux [20] est très proche des méthodes à pas fractionnaires de Temam [22] et [23] et des méthodes de projection proposées par Temam et Chorin, modifiées par Shen. En effet tous ces schémas ont pour avantage principal (et pour objectif) de découpler les 2 difficultés du problème de Navier–Stokes que sont la non-linéarité et l'incompressibilité. On trouvera une analyse mathématique de ces schémas dans Temam [3].

   Dans cette discrétisation, à chaque pas de temps le problème est fractionné en plusieurs étapes plus simples : tout d'abord un problème non linéaire formé de l'équation du moment sans le gradient de pression est résolu puis la vitesse et la pression sont corrigées avec un (schéma à 2 pas) ou deux (schéma à 3 pas) problèmes de Stokes généralisés, c'est à dire que la vitesse précédemment obtenue est projetée sur l'espace à divergence nulle.

   Regardons dans un cadre abstrait ce que cela donne. On considère un problème du type

$$\begin{aligned} \frac{du}{dt} + A(u) &= f \\ u(0) &= u_0 \end{aligned}$$

où A est un opérateur vérifiant $A = A_1 + A_2$. Si u^k représente la solution à l'instant $k\Delta t$, la valeur à l'instant $(k+1)\Delta t$, u^{k+1} est déterminée de la façon suivante :

Schéma 1 : Schéma à 2 pas

$$\begin{aligned} \frac{u^{k+1/2}-u^k}{\Delta t/2} + A_1(u^{k+1/2}) + A_2(u^k) &= f^{k+1/2} \\ \frac{u^{k+1}-u^{k+1/2}}{\Delta t/2} + A_1(u^{k+1/2}) + A_2(u^{k+1}) &= f^{k+1} \end{aligned}$$

Schéma 2 : Schéma à 3 pas

$$\begin{aligned} \frac{u^{k+\theta}-u^k}{\theta\Delta t} + A_1(u^{k+\theta}) + A_2(u^k) &= f^{k+\theta} \\ \frac{u^{k+1-\theta}-u^{k+\theta}}{(1-2\theta)\Delta t} + A_1(u^{k+\theta}) + A_2(u^{k+1-\theta}) &= f^{k+1-\theta} \\ \frac{u^{k+1}-u^{k+1-\theta}}{\theta\Delta t} + A_1(u^{k+1}) + A_2(u^{k+1-\theta}) &= f^{k+1} \end{aligned}$$

où f^k est une approximation de $f(k\Delta t)$ et $\theta \in]0, \frac{1}{2}[$. Ainsi la première étape traite implicitement A_1 et explicitement A_2 et les rôles de A_1 et A_2 sont inversés lors des étapes suivantes. On trouvera une démonstration de la convergence de tels schémas dans le cas d'opérateurs non linéaires dans Lions et Mercier [21].

Pour Navier-Stokes, ces schémas s'écrivent :

Schéma 1 : Schéma à 2 pas

$$u_h^{k+1/2} = g_h^{k+1/2} \quad \text{sur } \partial\Omega$$

$$\frac{(u_h^{k+1/2}-u_h^k, v_h)}{\Delta t/2} + \frac{1}{2Re}(\nabla u_h^{k+1/2}, \nabla v_h) - (p_h^{k+1/2}, \mathrm{div} v_h) =$$

$$(f^{k+1/2}, v_h) - \frac{1}{2Re}(\nabla u_h^k, \nabla v_h) - ((u_h^k.\nabla)u_h^k, v_h)$$

$$(q_h, \mathrm{div} u_h^{k+1/2}) = 0 \qquad \forall v_h \in \mathcal{V}_{0,h} \quad \text{et} \quad \forall q_h \in \mathcal{Q}_h,$$

$$u_h^{k+1} = g_h^{k+1} \quad \text{sur } \partial\Omega$$

$$\frac{(u_h^{k+1} - u_h^{k+1/2}, v_h)}{\Delta t/2} + \frac{1}{2Re}(\nabla u_h^{k+1}, \nabla v_h) + ((u_h^{k+1}.\nabla)u_h^{k+1}, v_h) =$$

$$(f^{k+1}, v_h) - \frac{1}{2Re}(\nabla u_h^{k+1/2}, \nabla v_h) + (p_h^{k+1/2}, \mathrm{div} v_h)$$

$$\forall v_h \in \mathcal{V}_{0,h} \quad \text{et} \quad \forall q_h \in \mathcal{Q}_h,$$

Schéma 2 : Schéma à 3 pas

$$u_h^{k+\theta} = g_h^{k+\theta} \quad \text{sur } \partial\Omega$$

$$\frac{(u_h^{k+\theta} - u_h^k, v_h)}{\theta\Delta t} + \frac{\alpha}{2Re}(\nabla u_h^{k+\theta}, \nabla v_h) - (p_h^{k+\theta}, \mathrm{div} v_h) =$$

$$(f^{k+\theta}, v_h) - \frac{\beta}{2Re}(\nabla u_h^k, \nabla v_h) - ((u_h^k.\nabla)u_h^k, v_h)$$

$$(q_h, \mathrm{div} u_h^{k+\theta}) = 0 \qquad \forall v_h \in \mathcal{V}_{0,h} \quad \text{et} \quad \forall q_h \in \mathcal{Q}_h,$$

$$u_h^{k+1-\theta} = g_h^{k+1-\theta} \quad \text{sur } \partial\Omega$$

$$\frac{(u_h^{k+1-\theta} - u_h^{k+\theta}, v_h)}{(1-2\theta)\Delta t} + \frac{\beta}{2Re}(\nabla u_h^{k+1-\theta}, \nabla v_h) + ((u_h^{k+1-\theta}.\nabla)u_h^{k+1-\theta}, v_h) =$$

$$(f^{k+1-\theta}, v_h) - \frac{\alpha}{2Re}(\nabla u_h^{k+\theta}, \nabla v_h) + (p_h^{k+\theta}, \mathrm{div} v_h)$$

$$\forall v_h \in \mathcal{V}_{0,h} \quad \text{et} \quad \forall q_h \in \mathcal{Q}_h,$$

$$u_h^{k+1} = g_h^{k+1} \quad \text{sur } \partial\Omega$$

$$\frac{(u_h^{k+1} - u_h^{k+1-\theta}, v_h)}{\theta\Delta t} + \frac{\alpha}{2Re}(\nabla u_h^{k+1}, \nabla v_h) - (p_h^{k+1}, \mathrm{div} v_h) =$$

$$(f^{k+1}, v_h) - \frac{\beta}{2Re}(\nabla u_h^{k+1-\theta}, \nabla v_h) - ((u_h^{k+1-\theta}.\nabla)u_h^{k+1-\theta}, v_h)$$

$$(q_h, \mathrm{div} u_h^{k+1}) = 0 \qquad \forall v_h \in \mathcal{V}_{0,h} \quad \text{et} \quad \forall q_h \in \mathcal{Q}_h,$$

L'efficacité de la méthode d'Uzawa décrite en section (II.3) permet d'utiliser le deuxième schéma dont le premier et troisième problème sont des problèmes de stokes généralisés et qui apparaît comme une sorte de symmétrisation du premier. La résolution du problème non linéaire est l'étape la plus coûteuse (jusqu'à 10 fois plus de temps CPU). Il est à remarquer que $u_h^{k+1-\theta}$ n'est pas nécessairement à divergence nulle.

Si les paramètres α, β et θ sont :

$$\alpha = \frac{2}{3}\ ,\ \beta = \frac{1}{3}\ ,\ \theta = \frac{1}{4}\ ,$$

le deuxième schéma est inconditionnellement stable. De plus les seuls opérateurs qui apparaissent sont $\frac{2I}{\Delta t} - \frac{\Delta}{3Re}$ et $2(\frac{2I}{\Delta t} - \frac{\Delta}{3Re})$. Une seule matrice est donc à inverser, ce qui réduit considérablement la taille mémoire nécessaire et le coût des factorisations.

II.4.2 Méthode des moindres carrés.

Ce paragraphe décrit la méthode des moindres carrés employée pour résoudre le problème non linéaire issu de la discrétisation en temps. Son application au problème de Stokes est développée dans Bristeau*et al* [17].

En dimension finie.

Dans le cas d'un système d'équations de dimension finie, la méthode des moindres carrés consiste à chercher la solution de

$$\begin{array}{l} F(x) = 0 \\ \text{avec } F : \mathbb{R}^N \longrightarrow \mathbb{R}^N \text{ et } F = \{f_1, \cdots, f_N\} \end{array} \tag{II.14}$$

comme étant la solution du problème de minimisation suivant :

$$\begin{array}{l} \text{Trouver } x \in \mathbb{R}^N \text{ tel que} \\ ||F(x)|| \leq ||F(y)|| \quad \forall y \in \mathbb{R}^N \end{array} \tag{II.15}$$

où $||.||$ est une norme euclidienne. Le problème (II.15) est résolu par une méthode de gradient conjugué préconditionné.

Supposons par exemple que

$$||F(x)|| = (S^{-1}F(x), F(x))$$

où S est une matrice symétrique définie positive. Si $J : \mathbb{R}^N \longrightarrow \mathbb{R}$ est défini par $J(x) = \frac{1}{2}||F(x)||$, on a l'équivalence :

$$(II.15) \Longleftrightarrow \left\{ \begin{array}{l} \text{Trouver } x \in \mathbb{R}^N \text{ tel que} \\ J(x) \leq J(y) \quad \forall y \in \mathbb{R}^N \end{array} \right.$$

Or si l'on note $J'(x)$ la différentielle de J (rappelons que J' est définie par

$$(J'(y), z) = (S^{-1}F(y), F'(y)z) \quad \forall z \in \mathbb{R}^N),$$

le gradient conjugué est composé des étapes suivantes :

Etape 1 : $x^0 \in \mathbb{R}^N$ étant donné

- calculer $J'(x^0)$
- résoudre $Sg^0 = J'(x^0)$ ie $(Sg^0, z) = (J'(x^0), z) \quad \forall z \in \mathbb{R}^N$
- poser $w^0 = g^0$

Etape 2 x^k, w^k et g^k étant connus

- déterminer λ_k tel que $J(x^k - \lambda_k w^k) \leq J(x^k - \lambda w^k) \quad \forall \lambda \in \mathbb{R}$
- poser $x^{k+1} = x^k - \lambda_k w^k$
- calculer $J'(x^{k+1})$
- résoudre $Sg^{k+1} = J'(x^{k+1})$ ie $(Sg^{k+1}, z) = (J'(x^{k+1}), z) \quad \forall z \in \mathbb{R}^N$
- poser $w^{k+1} = g^{k+1} + \gamma_{k+1} w^k$ où $\gamma_{k+1} = \dfrac{(Sg^{k+1}, g^{k+1})}{(Sg^k, g^k)} = \dfrac{J(g^{k+1})}{J(g^k)}$

Etape 3 Arrêt si $||g^{k+1}|| \leq \epsilon$.

Problème de Dirichlet non linéaire.

Considérons maintenant le problème modèle de Dirichlet non linéaire :

$$\begin{array}{rcll} \text{Trouver } \phi \text{ telle que} & & & \\ -\Delta\phi - T(\phi) & = & 0 & \text{dans } \Omega \\ \phi & = & 0 & \text{sur } \partial\Omega \end{array} \tag{II.16}$$

La formulation moindre carré de ce problème est la suivante :

$$\min_{v \in H_0^1(\Omega)} ||\Delta(v - \zeta(v))||_{H^{-1}(\Omega)} \tag{II.17}$$

que l'on peut réécrire

$$\min_{v \in H_0^1(\Omega)} ||v - \zeta(v)||_{H_0^1(\Omega)} = \min_{v \in H_0^1(\Omega)} \int_\Omega |\nabla(v - \zeta(v)|^2 \, dx \tag{II.18}$$

où $\zeta(v)$ est solution de

$$\begin{array}{rcll} -\Delta\zeta(v) & = & T(v) & \text{sur } \Omega \\ \zeta(v) & = & 0 & \text{sur } \partial\Omega. \end{array}$$

Rappelons que H^{-1} est le dual topologique de $H_0^1(\Omega)$, que Δ est un isomorphisme de $H_0^1(\Omega)$ dans H^{-1} et que la norme classique de H^{-1} est donnée par la relation :

$$||f||_{H^{-1}} = \sup_{v \in H_0^1(\Omega)-\{0\}} \frac{|(f,v)|}{||v||_{H_0^1(\Omega)}}.$$

La résolution du problème de contrôle optimal (II.18) est réalisée par une méthode de gradient conjugué préconditionné.

Et Stokes

Dans le cas présent la formulation variationnelle du problème non linéaire est du type :

$$\begin{array}{l} \text{Trouver } u \in V_g \text{ satisfaisant} \\ \alpha(u,v) + \frac{1}{Re}(\nabla u, \nabla v) + ((u.\nabla)u, v) \;=\; (f,v) \quad \forall v \in V_0 \end{array} \tag{II.19}$$

Pour tout $v \in V_0$, on définit $y(v) \in V_0$ comme étant la solution du problème linéaire :

$$\begin{array}{rcll} \alpha(y(v),z) + \frac{1}{Re}(\nabla y(v), \nabla z) & = & \alpha(v,z) + \frac{1}{Re}(\nabla v, \nabla z) & \\ & + & ((v.\nabla)v, z) - (f,z) & \forall z \in V_0 \\ y(v) & = & 0 \quad \text{sur } \partial\Omega & \end{array} \tag{II.20}$$

Comme (II.20) a une unique solution pour tout v donné et comme pour $v = u$, on a $y(u) = 0$, une formulation moindre carré du problème (II.19) est la suivante :

$$\begin{array}{l} \text{Trouver } u \in V_g \text{ satisfaisant} \\ J(u) \leq J(v) \quad \forall v \in V_0 \end{array} \tag{II.21}$$

où $J(v)$ est définie par :

$$J(v) = \frac{\alpha}{2}|y(v)|^2 + \frac{1}{2Re}|\nabla y(v)|^2.$$

Le gradient conjugué pour résoudre le problème (II.21) est composé des étapes suivantes :

Etape 1 Initialisation : $u^0 \in V_g$ est donnée.

- Déterminer $g^0 \in V_0$ tel que

$$\alpha(g^0, z) + \frac{1}{Re}(\nabla g^0, \nabla z) = (J'(u^0), z) \quad \forall z \in V_0$$

- $w^0 = g^0$

Etape 2 u^k, w^k et g^k étant connues,

- Déterminer λ_k tel que $J(u^k - \lambda_k w^k) \leq J(u^k - \lambda w^k) \ \ \forall \lambda \in \mathbb{R}$
- poser $u^{k+1} = u^k - \lambda_k w^k$
- déterminer $y^{k+1} = y(u^{k+1})$
- calculer $J'(u^{k+1})$
- déterminer $g^{k+1} \in V_0$ tel que

$$\alpha(g^{k+1}, z) + \frac{1}{Re}(\nabla g^{k+1}, \nabla z) = (J'(u^{k+1}), z) \ \ \forall z \in V_0$$

- $w^{k+1} = g^{k+1} + \gamma_{k+1} w^k$ où $\gamma_{k+1} = \dfrac{\alpha|g^{k+1}|^2 + \frac{1}{Re}|\nabla g^{k+1}|^2}{\alpha|g^k|^2 + \frac{1}{Re}|\nabla g^k|^2} = \dfrac{J(g^{k+1})}{J(g^k)}$

Etape 3 : Arrêt si $\dfrac{J(g^{k+1})}{J(g^0)}$

Le calcul de la différentielle de J ie $J'(u^{k+1})$ est effectué en détail dans Glowinski [2] : pour tout z dans V_0 $J'(u^{k+1})$ vérifie :

$$\begin{aligned}(J'(u^{k+1}), z) &= \alpha(y(u^{k+1}), z) + \frac{1}{Re}(\nabla y(u^{k+1}), \nabla z) \\ &+ (y(u^{k+1}), (z.\nabla)u^{k+1}) + (y(u^{k+1}), (u^{k+1}.\nabla)z).\end{aligned} \tag{II.22}$$

Quelques itérations (2 ou 3) de la méthode de Newton suffisent pour estimer la valeur de λ_k et y^{k+1} est déterminé en introduisant les vecteurs y_1^k et y_2^k :

$$y^{k+1} = y^k - \lambda_k y_1^k + \lambda_k^2 y_2^k$$

où y_1^k est solution de

$$\begin{aligned}\alpha(y_1^k, z) + \frac{1}{Re}(\nabla y_1^k, \nabla z) &= \alpha(w^k, z) + \frac{1}{Re}(\nabla w^k, \nabla z) \\ &+ ((u^k.\nabla)w^k, z) + ((w^k.\nabla)u^k, z) \\ y_1^k &= 0 \quad \text{sur } \partial\Omega\end{aligned}$$

et y_2^k est solution de

$$\begin{aligned}\alpha(y_2^k, z) + \frac{1}{Re}(\nabla y_2^k, \nabla z) &= ((w^k\nabla)w^k, z) \\ y_2^k &= 0 \quad \text{sur } \partial\Omega\end{aligned}$$

Pour la détermination de y_1^k, de y_2^k et g^{k+1}, seuls des problèmes de Dirichlet avec des conditions homogènes sont à résoudre. Il y en a 3 par composantes (ie 6) pour chaque itération du moindre carré. Les étapes les plus coûteuses sont constituées du calcul des différents termes non linéaires (le calcul du jacobien J' compris), le reste étant des produits matrices-vecteurs.

Remarque II.5 *Le paramètre γ_{k+1} dit de Fletcher-Reeves peut être remplacé par le paramètre de Polak-Ribière :*

$$\gamma_{k+1} = \frac{\alpha(g^{k+1} - g^k, g^{k+1}) + \frac{1}{Re}(\nabla(g^{k+1} - g^k), \nabla g^{k+1})}{\alpha|g^k|^2 + \frac{1}{Re}|\nabla g^k|^2}$$

Remarque II.6 *Dans le cas continu, l'existence et l'unicité de la solution du problème non linéaire ne sont démontrées que si le terme non linéaire est $(u.\nabla u)u + \frac{1}{2}u.\nabla u$ (cf. Temam [3]). Numériquement les résultats sont identiques.*

Remarque II.7 *On peut utiliser un schéma temporel implicite où u_h^{k+1} est déterminé par :*

$$\begin{aligned} &\frac{(u_h^{k+1} - u_h^k, v_h)}{\Delta t} + \frac{1}{Re}(\nabla u_h^{k+1}, \nabla v_h) \\ &-(p_h^{k+1}, div v_h) + ((u_h^{k+1}.\nabla)u_h^{k+1}, v_h) \quad = \quad (f^{k+1}, v_h) \end{aligned}$$

$$(q_h, div u_h^{k+1}) = 0 \qquad \forall v_h \in \mathcal{V}_{0,h} \qquad et \qquad \forall q_h \in \mathcal{Q}_h,$$

Dans ce cas chaque itération du moindre carré nécessite 6 résolutions du problème de Stokes généralisé.

II.4.3 Programmation et résultats numériques.

L'écoulement est initialisé par la solution du problème de Stokes, en particulier cette solution vérifie la condition :

$$(q_h, div u_h) = 0 \quad \forall q_h \in \mathcal{Q}_h$$

Les calculs ont été réalisés pour $Re = 100$ et $Re = 1000$ dans le cas de la cavité A) ie dans le cas où la vitesse d'entrainement est 1.0, et pour $Re = 100, 400, 1000, 5000$ dans le cas de la cavité B). Pour de tels nombres de Reynolds, la solution du problème de Navier-Stokes dans la cavité entrainée évolue vers une solution stationnaire ; le critère de convergence est le suivant :

$$\begin{aligned} \frac{||p_h^{k+1} - p_h^k||_{2h}}{||p_h^k||_{2h}} &\leq 0.001 \\ \frac{||u_h^{k+1} - u_h^k||_h}{||u_h^k||_h} &\leq 0.001 \end{aligned}$$

Les résultats numériques pour la solution asymptotique stationnaire sont comparés avec les résultats des publications suivantes :

- pour la cavité A)
  - Ghia *et al* [26] : publication concernant le problème stationnaire
  - Bruneau et Jouron [25] : publication concernant le problème stationnaire
  - Sonke [30] : thèse concernant le problème instationnaire
- pour la cavité B)
  - Shen [29] : thèse concernant le problème instationnaire
  - Shen [28] : publication concernant le problème instationnaire.

Les principaux résultats sur la localisation des tourbillons et les profils médians (cf section II.2) sont résumés dans les tableaux II.2 et II.3 pour la cavité A) et dans les tableaux II.4, II.5, II.6 et II.7 pour la cavité B). On peut constater que les résultats obtenus dans le cas présent s'accordent de façon satisfaisante avec les résultats précédemment publiés (tout particulièrement pour $Re = 100$ et $Re = 400$). Les quelques différences avec Ghia *et al* et Bruneau-Jouron sont essentiellement dues :

- aux différents maillages utilisés : le raffinement et la localisation de ce raffinement améliorent les résultats. La méthode des différences finies est pour des géométries simples plus économique et permet des résolutions plus fines,
- à un critère de convergence trop grand : les calculs sont prématurément arrêtés. Qualitativement la solution ne varie plus à partir d'un certain temps, mais la convergence n'est pas terminée et la solution diffère de la solution du problème stationnaire,
- à la variété des méthodes et des philosophies numériques employées par les auteurs cités.

Evolution au cours du temps

La figure II.11 donne un aperçu de l'évolution des lignes de courant au cours du temps pour $Re = 5000$. Elle permet d'apprécier le développement et la dynamique (naissance, vie et mort) des contre-tourbillons. On y constate essentiellement la présence d'un tourbillon central légèrement décentré dont le centre se déplace au cours du temps et autour duquel apparaissent plusieurs tourbillons. On remarque que le contre tourbillon en bas à gauche qui existe à $t = 5.0$ rejoint le contre tourbillon de droite à $t = 10.0$ pour former un large tourbillon qui s'étale sur toute la largeur de la cavité, puis disparaît à $t = 30.0$ pour réapparaître et se développer jusqu'à $t = 60.0$. On note que la "recirculation" en haut à gauche est absente à l'instant 10.0.

$Re = 100$					
		Pascal	Ghia [26]	Bruneau [25]	Sonke [30]
Formulation		U, P	ω, ψ	U, P	U, ω
Problème		Evolutif	Stationnaire	Stationnaire	Evolutif
Discrétisation x		Galerkin	Multigrille	Multigrille	Galerkin
		Eléments finis	Différences finis	Différences finis	Différences finis
Maillage vitesse		65×65	129×129	129×129	40×40
Δt		0.1			?
PV	x	0.6094	0.6172	0.6172	0.6129
	y	0.7344	0.7344	0.7344	0.7419
	$\psi(x,y)$	-0.1036	-0.1034	-0.1026	-0.1041
	$\omega(x,y)$	-3.113			
BLV	x	0.0313	0.0313	0.0313	0.0484
	y	0.0313	0.0391	0.0391	0.0385
	$\psi(x,y)$	$0.1204\ 10^{-5}$	$0.175\ 10^{-5}$	$0.163\ 10^{-5}$	$0.4\ 10^{-5}$
	$\omega(x,y)$	0.00978			
BRV	x	0.9531	0.9453	0.9453	0.9355
	y	0.0625	0.0625	0.0625	0.0484
	$\psi(x,y)$	$0.109\ 10^{-4}$	0.12510^{-4}	0.12510^{-4}	$0.12\ 10^{-4}$
	$\omega(x,y)$	0.02454			
	umin	-0.209	-0.210	-0.210	-0.210
	ymin	0.453	0.453	0.453	0.463
	vmin	-0.238	-0.245	-0.252	-0.247
	xmin	-0.8125	0.8047	0.8125	0.8064
	vmax	0.172	0.175	0.179	0.178
	xmax	0.2344	0.2344	0.2344	0.2364

Tableau II.2: *Résultats pour la cavité entrainé A) avec la méthode de splitting à 3 pas - Uzawa - Moindres carrés. PV = Tourbillon principal ; BLV = Contre tourbillon gauche ; BRV = Contre tourbillon droit.*

$Re = 1000$					
		Pascal	Ghia [26]	Bruneau [25]	Sonke [30]
Formulation		U, P	ω, ψ	U, P	U, ω
Problème		Evolutif	Stationnaire	Stationnaire	Evolutif
Discrétisation x		Galerkin	Multigrille	Multigrille	Galerkin
		Eléments finis	Différences finis	Différences finis	Différences finis
Maillage vitesse		65×65	129×129	129×129	120×120
Δt		0.1			?
PV	x	0.5313	0.5313	0.5313	0.5368
	y	0.5625	0.5625	0.5586	0.5693
	$\psi(x,y)$	-0.1231	-0.1179	-0.1163	-0.1210
	$\omega(x,y)$	-0.02185			
BLV	x	0.0781	0.0859	0.0859	0.0834
	y	0.0781	0.0781	0.0820	0.0637
	$\psi(x,y)$	$0.228\ 10^{-3}$	$0.231\ 10^{-3}$	$0.325\ 10^{-3}$	$0.274\ 10^{-3}$
	$\omega(x,y)$	0.3242			
BRV	x	0.8594	0.8594	0.8711	0.8513
	y	0.1094	0.1094	0.1094	0.1237
	$\psi(x,y)$	$0.169\ 10^{-2}$	$0.175\ 10^{-2}$	$0.191\ 10^{-2}$	$0.301\ 10^{-2}$
	$\omega(x,y)$	1.219			
	umin	-0.399	-0.383	-0.376	-0.405
	ymin	0.172	0.172	0.172	0.171
	vmin	-0.536	-0.516	-0.521	-0.653
	xmin	0.906	0.906	0.910	0.902
	vmax	0.386	0.371	0.366	0.351
	xmax	0.156	0.156	0.152	0.164

Tableau II.3: *Résultats pour la cavité entrainé A) avec la méthode de splitting à 3 pas - Uzawa - Moindres carrés. PV = Tourbillon principal ; BLV = Contre tourbillon gauche ; BRV = Contre tourbillon droit.*

$Re = 100$				
		Pascal	Shen [29]	Shen [28]
Formulation		U, P	U, P	U, P
Problème		Evolutif	stationnaire	Evolutif
Discrétisation en x		Galerkin classique Eléments finis	Galerkin classique Spectral-Chebyshev	Galerkin classique Spectral-Chebyshev
Maillage vitesse		65×65	64×64	17×17
Δt		0.1		4.0
PV	x	0.6094	0.63	0.609
	y	0.7500	0.77	0.750
	$\psi(x,y)$	-0.08374	-0.08366	-0.0836
	$\omega(x,y)$	2.933		
BLV	x	0.03125	0.05	0.031
	y	0.03125	0.05	0.031
	$\psi(x,y)$	$0.429\ 10^{-7}$	$0.127\ 10^{-5}$	$0.140\ 10^{-5}$
	$\omega(x,y)$	0.00749		
BRV	x	0.9531	0.97	0.953
	y	0.04687	0.06	0.047
	$\psi(x,y)$	$0.456\ 10^{-5}$	$0.490\ 10^{-5}$	$0.467\ 10^{-5}$
	$\omega(x,y)$	0.0179		
	umin	-0.159		
	ymin	+0.469		
	vmin	-0.179		
	xmin	+0.812		
	vmax	+0.128		
	xmax	+0.266		

Tableau II.4: *Résultats pour la cavité entrainé B) ; PV = Tourbillon principal ; BLV = Contre tourbillon gauche ; BRV = Contre tourbillon droit*

$Re = 400$				
		Pascal	Shen [29]	Shen [28]
Formulation		U, P	U, P	U, P
Problème		Evolutif	stationnaire	Evolutif
Discrétisation en x		Galerkin classique Eléments finis	Galerkin classique Spectral-Chebyshev	Galerkin classique Spectral-Chebyshev
Maillage vitesse		65×65	64×64	17×17
Δt		0.1		0.5
PV	x	0.578	0.59	0.578
	y	0.625	0.63	0.625
	$\psi(x,y)$	-0.0857	-0.0857	-0.0858
	$\omega(x,y)$			
BLV	x	0.03125	0.05	0.031
	y	0.03125	0.06	0.047
	$\psi(x,y)$	$0.795\ 10^{-6}$	$0.259\ 10^{-5}$	$0.631\ 10^{-5}$
	$\omega(x,y)$			
BRV	x	0.906	0.92	0.922
	y	0.125	0.13	0.094
	$\psi(x,y)$	0.21810^{-3}	$0.256\ 10^{-3}$	$0.198\ 10^{-3}$
	$\omega(x,y)$			
	umin	-0.227		
	ymin	+0.328		
	vmin	-0.313		
	xmin	+0.844		
	vmax	+0.200		
	xmax	+0.281		

Tableau II.5: *Résultats pour la cavité entrainé B) ; PV = Tourbillon principal ; BLV = Contre tourbillon gauche ; BRV = Contre tourbillon droit*

$Re = 1000$					
		Pascal	Pascal	Shen [29]	Shen [28]
Maillage vitesse		65×65	129	64×64	25×25
Δt		0.1	0.05		0.15
PV	x	0.5469	0.5547	0.5469	0.547
	y	0.5781	0.5859	0.5781	0.578
	$\psi(x,y)$	-0.09220	-0.09028	-0.0843	-0.08717
	$\omega(x,y)$	-1.944	-2.168		
BLV	x	0.0625	0.0547	0.08	0.078
	y	0.0625	0.0547	0.09	0.063
	$\psi(x,y)$	$0.315\ 10^{-4}$	$0.127\ 10^{-4}$	$0.515\ 10^{-4}$	$0.828\ 10^{-4}$
	$\omega(x,y)$	0.0761	0.0408		
BRV	x	0.875	0.867		0.922
	y	0.125	0.125		0.094
	$\psi(x,y)$	$0.922\ 10^{-3}$	$0.885\ 10^{-3}$	$0.882\ 10^{-3}$	$0.568\ 10^{-3}$
	$\omega(x,y)$	0.6069	0.5099		
	umin	-0.279	-0.275		
	ymin	+0.219	+0.242		
	vmin	-0.376	-0.366		
	xmin	+0.891	+0.891		
	vmax	+0.260	+0.252		
	xmax	+0.219	+0.242		

Tableau II.6: *Résultats pour la cavité entrainé B) ; PV = Tourbillon principal ; BLV = Contre tourbillon gauche ; BRV = Contre tourbillon droit*

$Re = 5000$					
		Pascal	Pascal	Shen [29]	Shen [28]
Maillage vitesse		65×65	129		33×33
Δt		0.1	0.05		0.03
PV	x	0.5156	0.5390		0.516
	y	0.5156	0.5313		0.531
	$\psi(x,y)$	-0.1004	-0.0975		-0.0880
	$\omega(x,y)$	-2.043	-2.169		
BLV	x	0.078	0.0859		0.094
	y	0.125	0.1172		0.094
	$\psi(x,y)$	$0.718\ 10^{-3}$	$0.6723\ 10^{-3}$		$0.753\ 10^{-3}$
	$\omega(x,y)$	0.8389	0.7310		
BRV	x	0.797	0.8047		0.922
	y	0.0781	0.0781		0.94
	$\psi(x,y)$	$0.203\ 10^{-2}$	$0.242\ 10^{-2}$		$0.775\ 10^{-3}$
	$\omega(x,y)$	2.016	2.009		
ULV	x	0.094	0.0781		0.078
	y	0.906	0.906		0.92
	$\psi(x,y)$	$0.866\ 10^{-3}$	$0.786\ 10^{-3}$		$0.678\ 10^{-3}$
	$\omega(x,y)$	1.034	1.159		
SBR	x	0.9844	0.9844		
	y	0.01565	0.01565		
	$\psi(x,y)$	$-4.5\ 10^{-7}$	$-2.391\ 10^{-7}$		
	$\omega(x,y)$	$-0.417\ 10^{-2}$	$-0.234\ 10^{-1}$		
	umin	-0.350	-0.318		
	ymin	+0.094	+0.094		
	vmin	-0.428	-0.414		
	xmin	+0.938	+0.945		
	vmax	+0.339	+0.310		
	xmax	+0.094	+0.102		

Tableau II.7: *Résultats pour la cavité entrainé B) ; PV = Tourbillon principal ; BLV = Contre tourbillon gauche ; BRV = Contre tourbillon droit ; ULV = Contre tourbillon droit en haut ; SBR = Contre contre tourbillon en bas à droite*

A partir de $t = 50.0$ la forme générale de l'écoulement est établie, les variations sont minimes et un contre contre tourbillon négatif de faible intensité apparaît dans le coin inférieur droit (cf figure II.14).

A $t = 60.0$, on considère la solution comme étant stationnaire. Les isobares (cf figure II.13) sont circulaires et la pression posséde un fort gradient dans le coin supérieur droit de la cavité. La figure II.14 montre les lignes tabulées au niveau des contres tourbillons qui sont bien "captés" par le schéma numérique.

Les lignes d'isovorticité (dont l'évolution est représentée en figure II.12) montrent à partir de $t = 50.0$ une région centrale assez large de vorticité constante entourée par une région quasi circulaire fine où la vorticité varie fortement (région formée de "plusieurs couches" de vorticité). Cette zone centrale s'élargit au cours du temps ; à $t = 5.0$ étroite, elle est entourée de 2 tourbillons de vorticité puis la zone de fort gradient augmente et s'avance en s'enroulant autour de la zone centrale.

Convergence

Les courbes de la figure II.15 montrent que la convergence vers la solution stationnaire est lente. On vérifie que le nombre d'itérations d'Uzawa est constant et faible (on retrouve l'efficacité du solveur de Stokes) devant le nombre d'itérations du moindre carré (c'est bien l'étape la plus coûteuse). Les petites oscillations sur la courbe de la divergence sont dues au fait que le critère d'arrêt pour Stokes porte sur la pression et la vitesse. On trouvera dans le tableau II.8 les différents temps de convergence.

Influence du nombre de Reynolds et de la précision du maillage

Les figures II.16, II.17 et II.18 montrent la géométrie de l'écoulement en fonction du nombre de Reynolds. Pour de faibles nombres de Reynolds (100, 400, 1000) le contre tourbillon en haut à gauche n'existe pas. Les recirculations deviennent de plus en plus importantes et l'écoulement stationnaire de plus en plus complexe au fur et à mesure que Re augmente.

Un maillage 65^2 est suffisant pour de petits nombres de Reynolds, mais pour $Re = 5000$ la figure II.19 illustre les effets d'un maillage insuffisant sur le gradient de pression. Celui ci est en effet discontinu dans le cas 65^2 et nettement plus régulier dans le cas 129^2. Pour $Re = 10000$ les oscillations de la vorticité et de l'énergie (figure II.20) induisent très clairement que le maillage 65×65 est trop grossier pour une résolution raisonnable avec un tel nombre de Reynolds. Le schéma diverge d'ailleurs quelques itérations plus tard.

	Itérations	Temps	Temps cpu / itérations
$Re = 100$ $h = 1/64$ $\Delta t = 0.1$ Cavité B)	33	3.3	Stardent
$Re = 400$ $h = 1/64$ $\Delta t = 0.1$ Cavité B)	119	11.9	Stardent
$Re = 1000$ $h = 1/64$ $\Delta t = 0.1$ Cavité B)	200	20.0	Cray 17s Stardent 75s
$Re = 5000$ $h = 1/64$ $\Delta t = 0.1$ Cavité B)	≥ 600	≥ 60.0	Cray 28s Stardent 206s
$Re = 5000$ $h = 1/128$ $\Delta t = 0.05$ Cavité B)	≥ 1200	≥ 60.0	Cray 60s Stardent 540s

Tableau II.8: *Temps de convergence et temps CPU pour Navier-Stokes. Précision demandée égale à 0.001*

Temps de calcul

Le tableau II.8 donne les temps de calcul moyens par itération en temps pour la résolution des équations de Navier-Stokes par la méthode de splitting-Uzawa-moindres carrés. Les calculs ont été réalisés sur les ordinateurs :

- Cray 2 au CCVR (Palaiseau)
- Stardent 3000 : bi-processeur de Stardent au Laboratoire d'Analyse numérique d'Orsay.

Il est à noter que le taux de vectorisation du programme ie le rapport

$$\frac{\text{Temps cpu non vectorisé}}{\text{Temps cpu vectorisé}}$$

est faible, de l'ordre de 1.75. En effet l'assemblage indirect des matrices creuses induit des dépendances vectorielles.

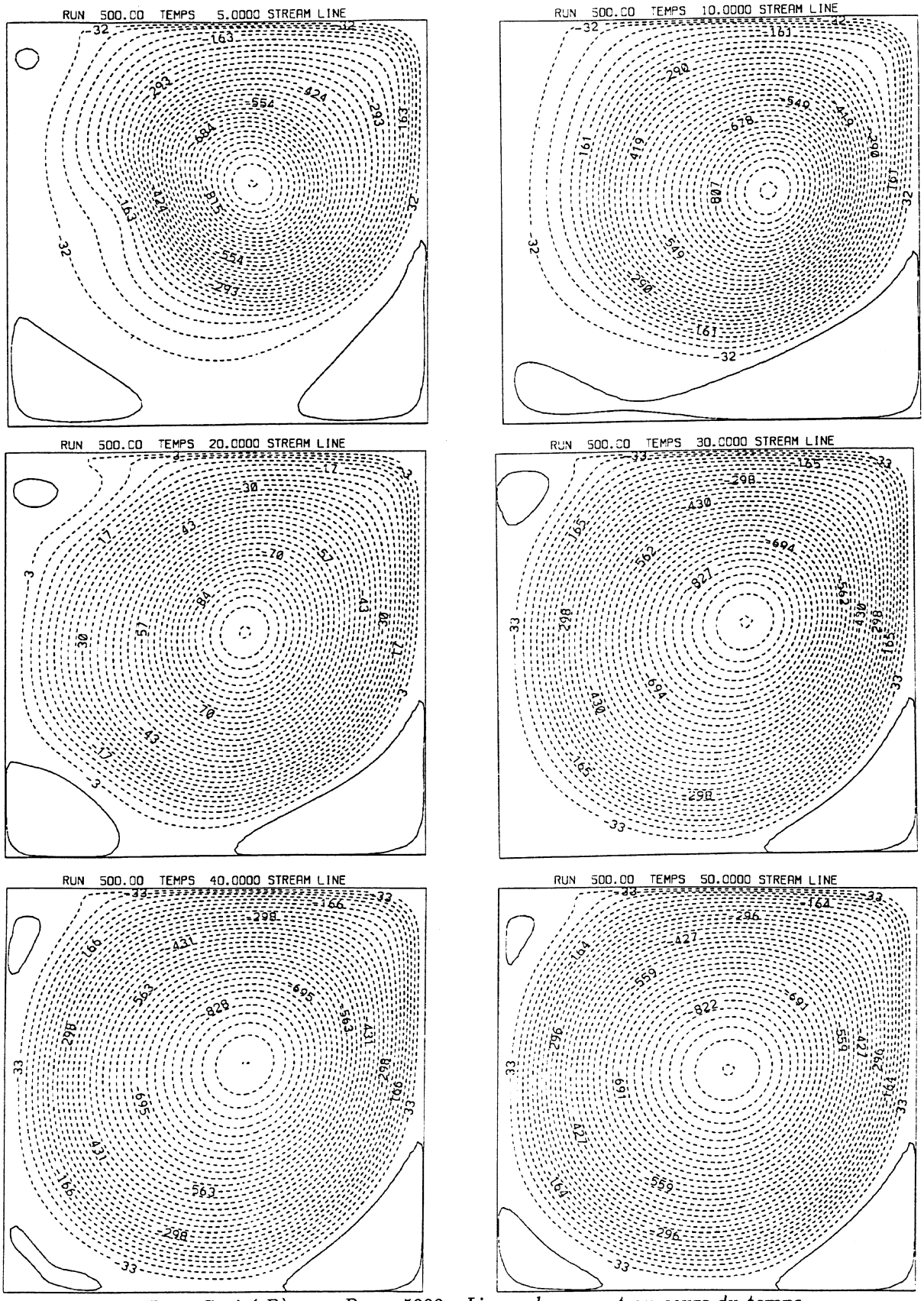


Figure II.11: *Cavité B) pour* $Re = 5000$ *: Lignes de courant au cours du temps*

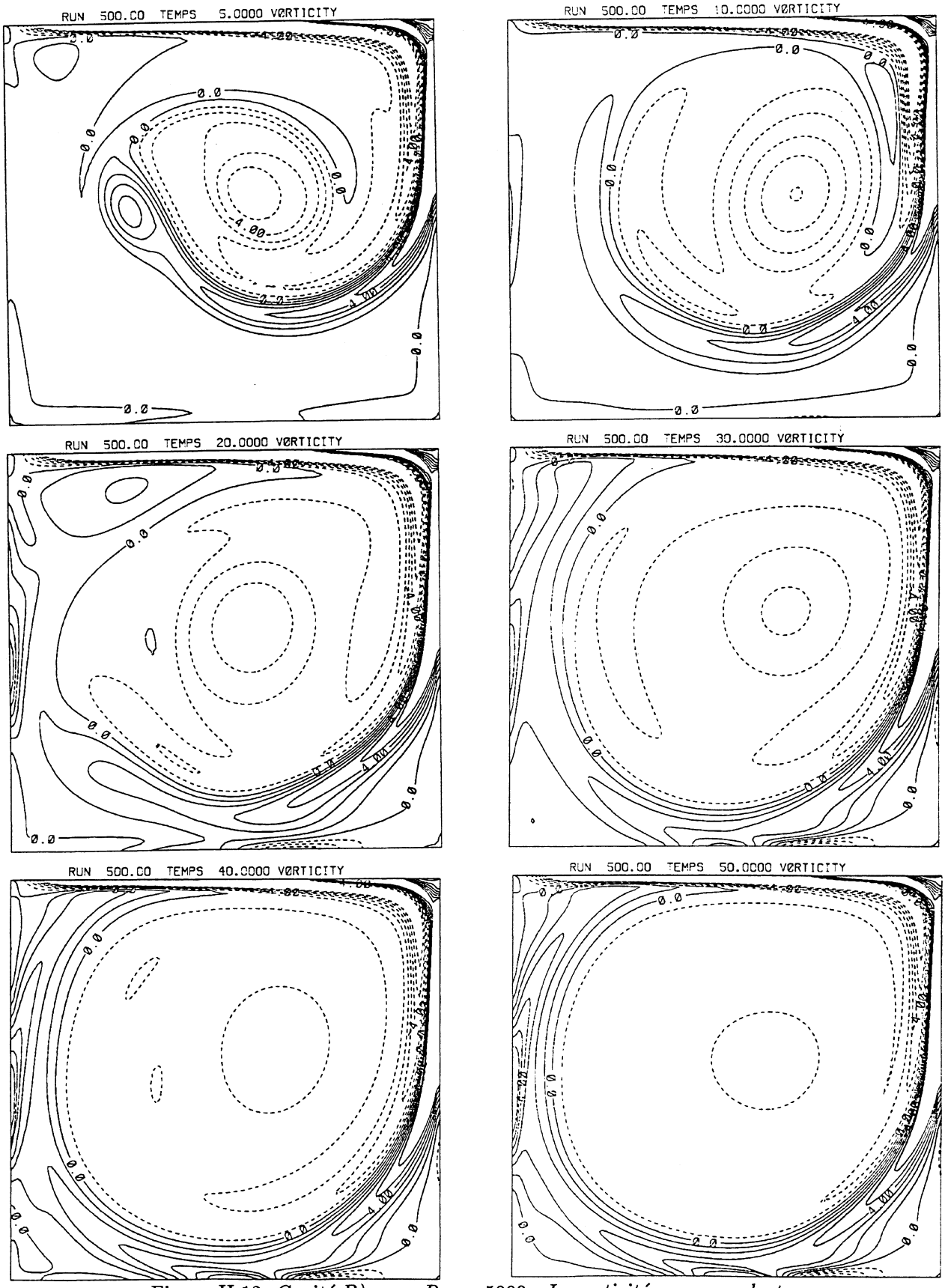


Figure II.12: *Cavité B) pour* $Re = 5000$ *: Isovorticité au cours du temps*

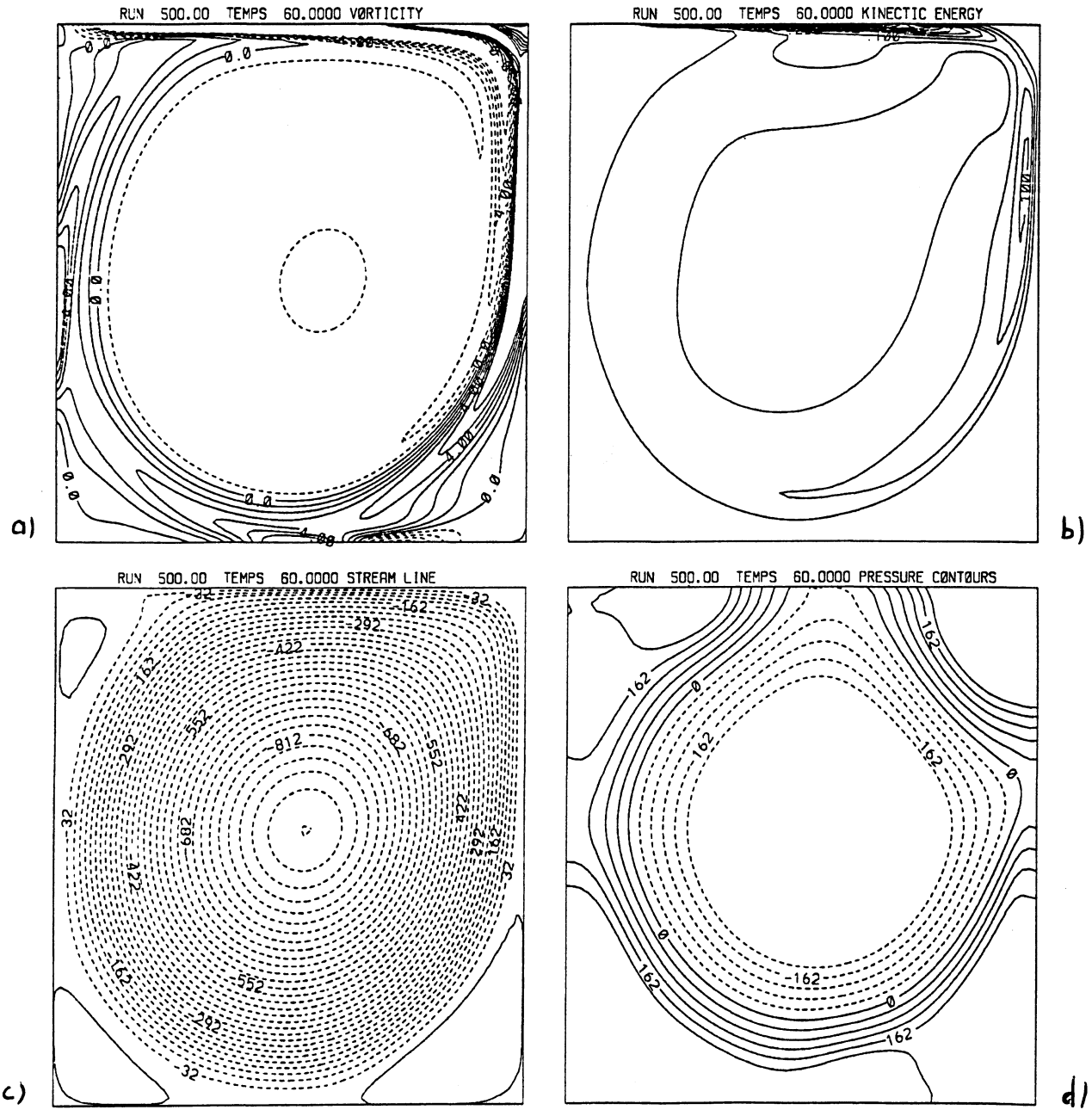


Figure II.13: *Cavité B) pour* $Re = 5000$ *à l'instant* $t = 60.0$.
a) *Isovorticité ; b) Energie ; c) Lignes de courant ; d) Isobares.*

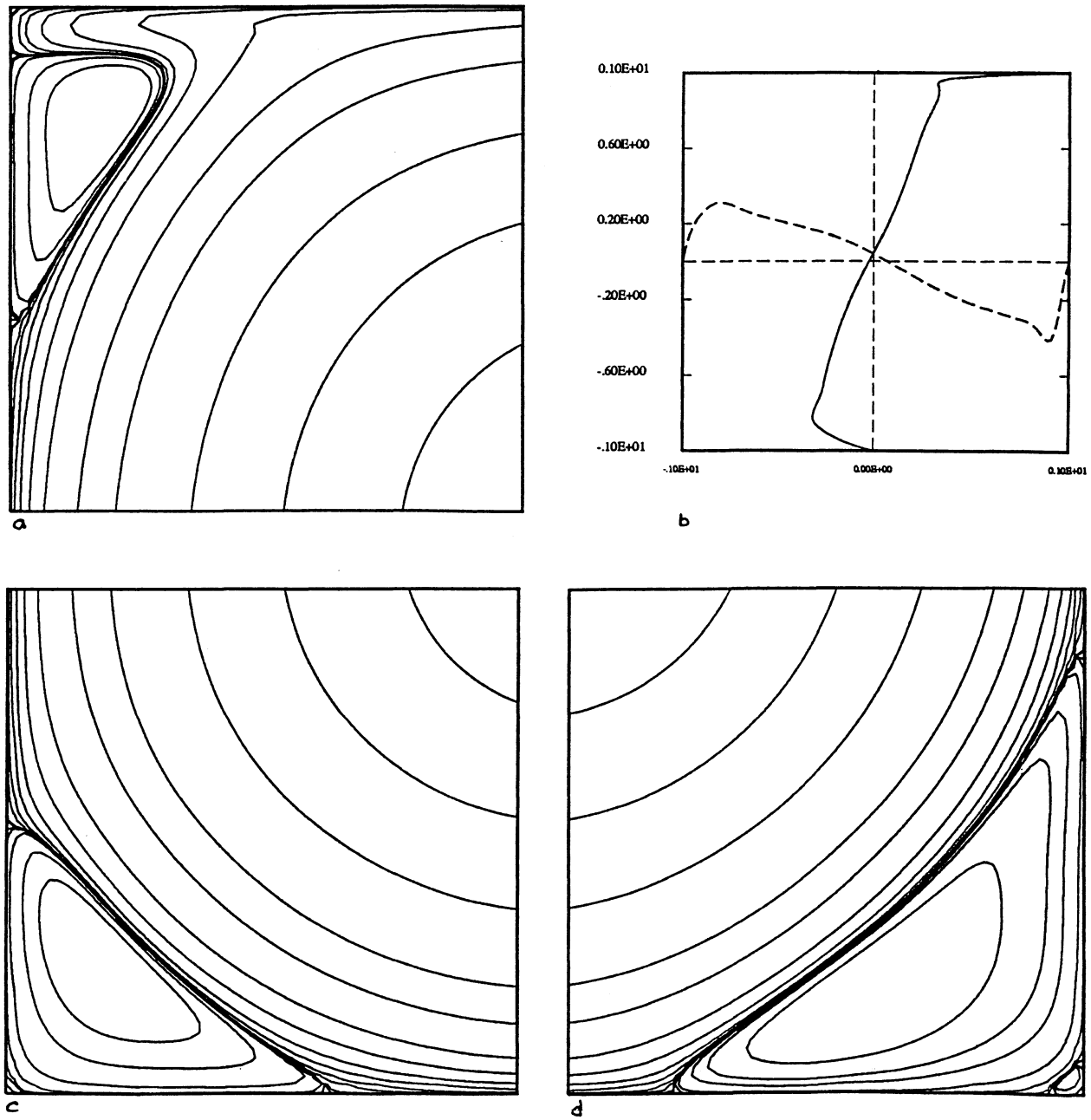


Figure II.14: *Cavité B) pour* $Re = 5000$ *à l'instant* $t = 60.0$.
a) Lignes de courant coin inférieur gauche ; b) Profils médians de la vitesse ; c) Lignes de courant coin inférieur droit ; d) Lignes de courant coin supérieur droit.

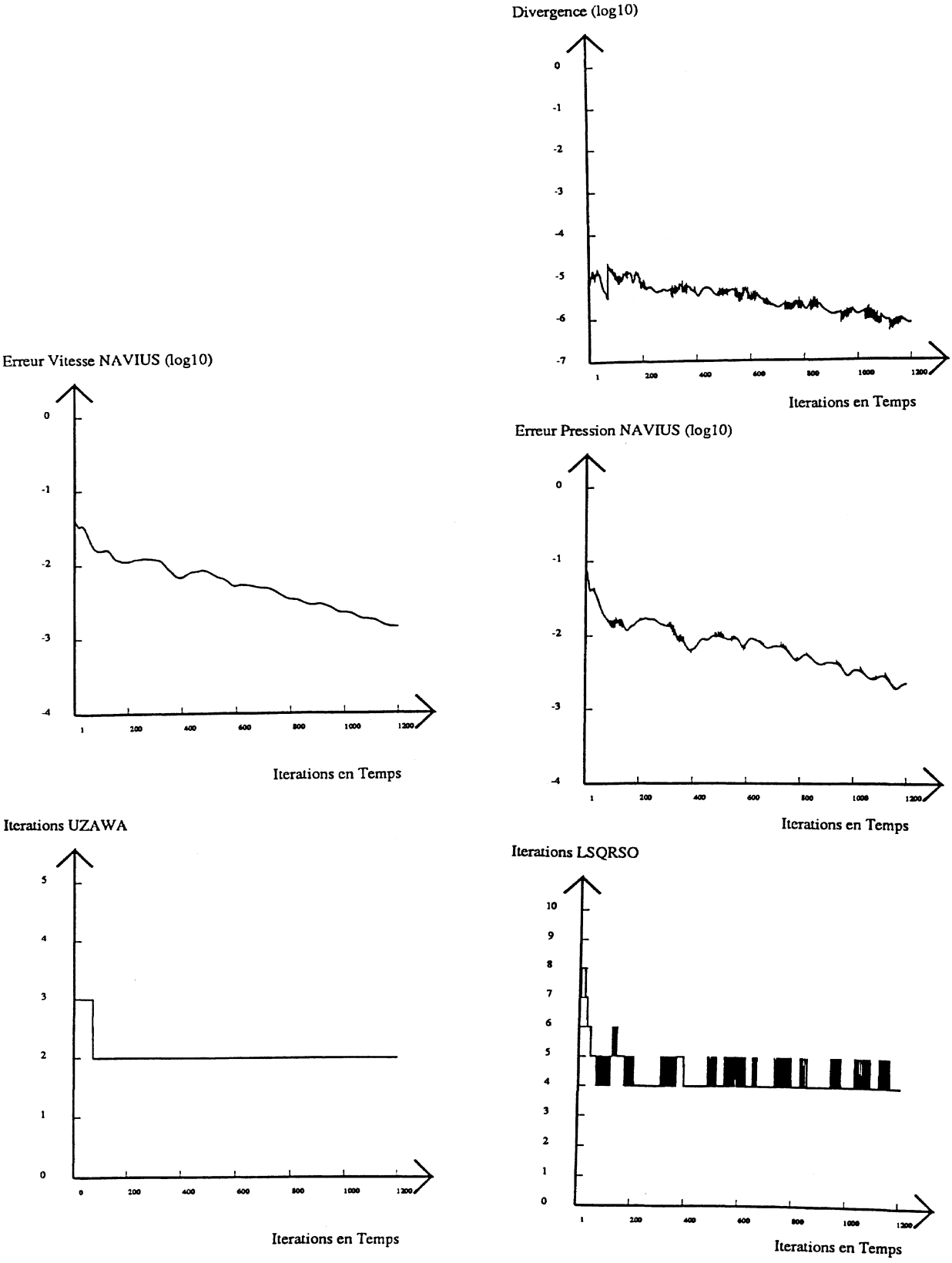


Figure II.15: *Cavité B) pour* $Re = 5000$ *: Courbes de convergence.*

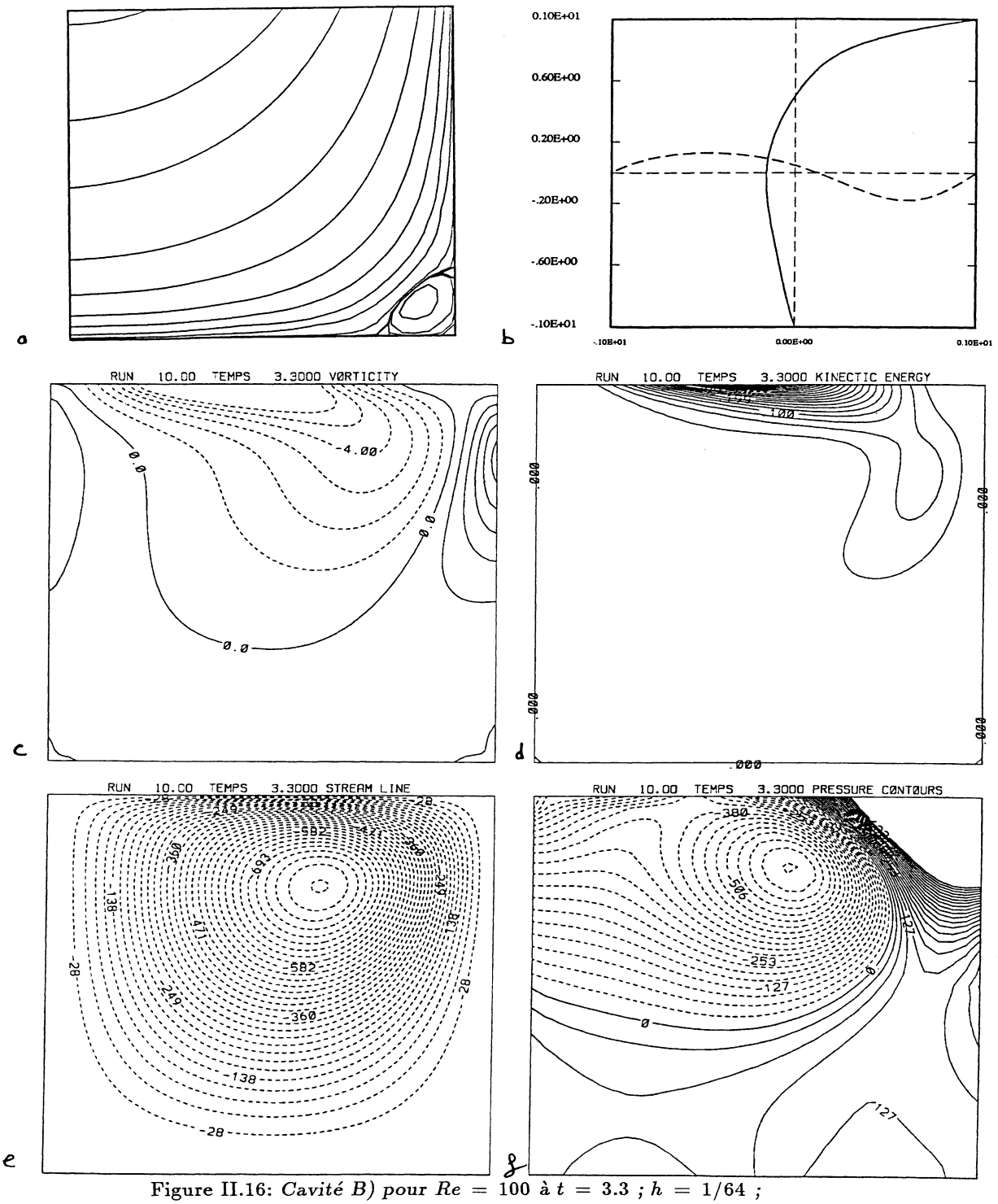


Figure II.16: *Cavité B) pour* $Re = 100$ *à* $t = 3.3$ *;* $h = 1/64$ *;*
a) Lignes de courant coin inférieur droit ; b) Profils médians de la vitesse ; c) Isovorticité ; d) Energie ; e) Lignes de courant ; f) Isobares.

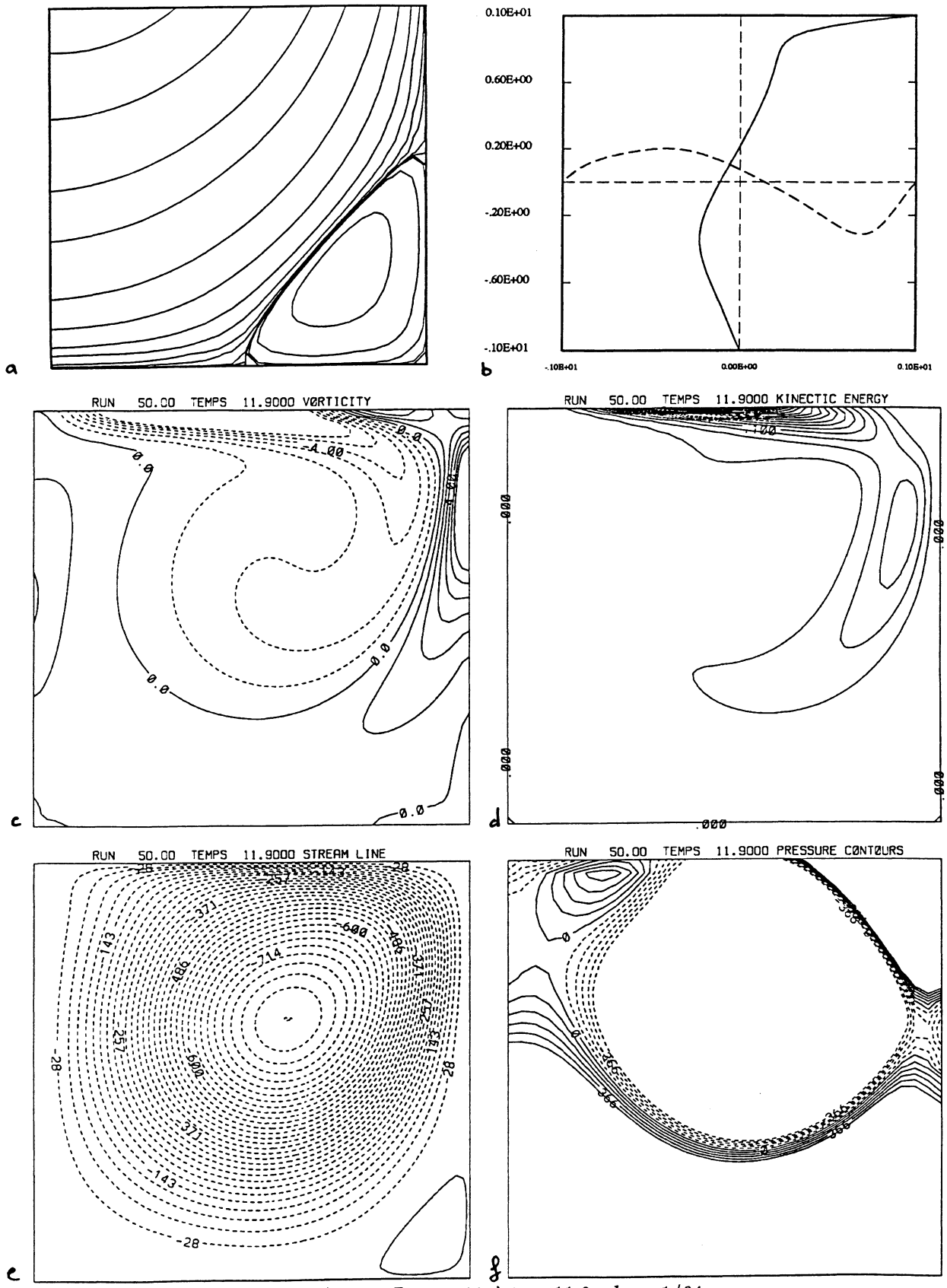


Figure II.17: *Cavité B) pour* $Re = 400$ *à* $t = 11.9$ *;* $h = 1/64$ *;*
a) Lignes de courant coin inférieur droit ; b) Profils médians de la vitesse ; c) Isovorticité ; d) Energie ; e) Lignes de courant ; f) Isobares.

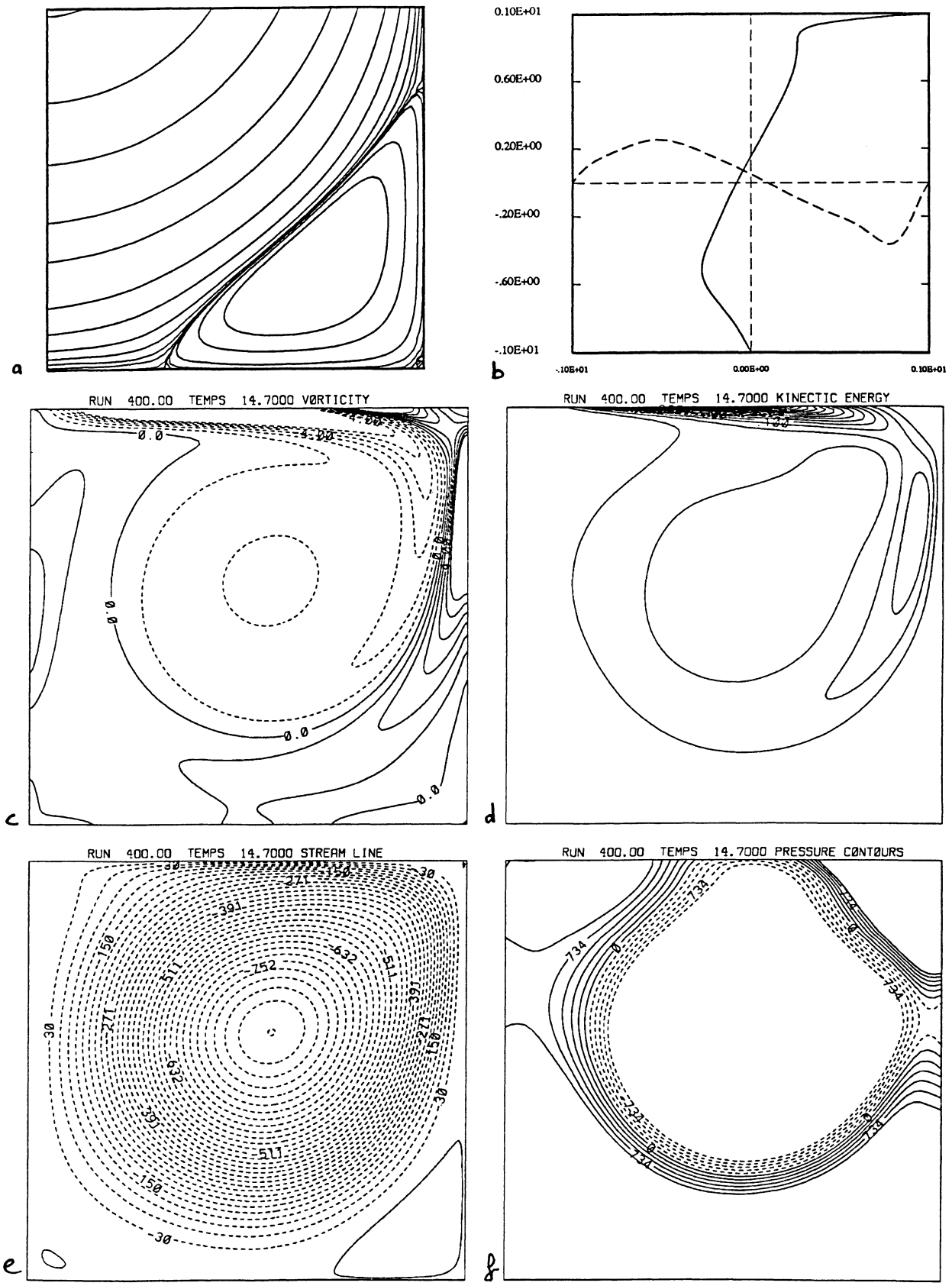


Figure II.18: *Cavité B) pour* $Re = 1000$ *à* $t = 14.7$ *;* $h = 1/128$ *;*
a) Lignes de courant coin inférieur droit ; b) Profils médians de la vitesse ; c) Isovorticité ; d) Energie ; e) Lignes de courant ; f) Isobares.

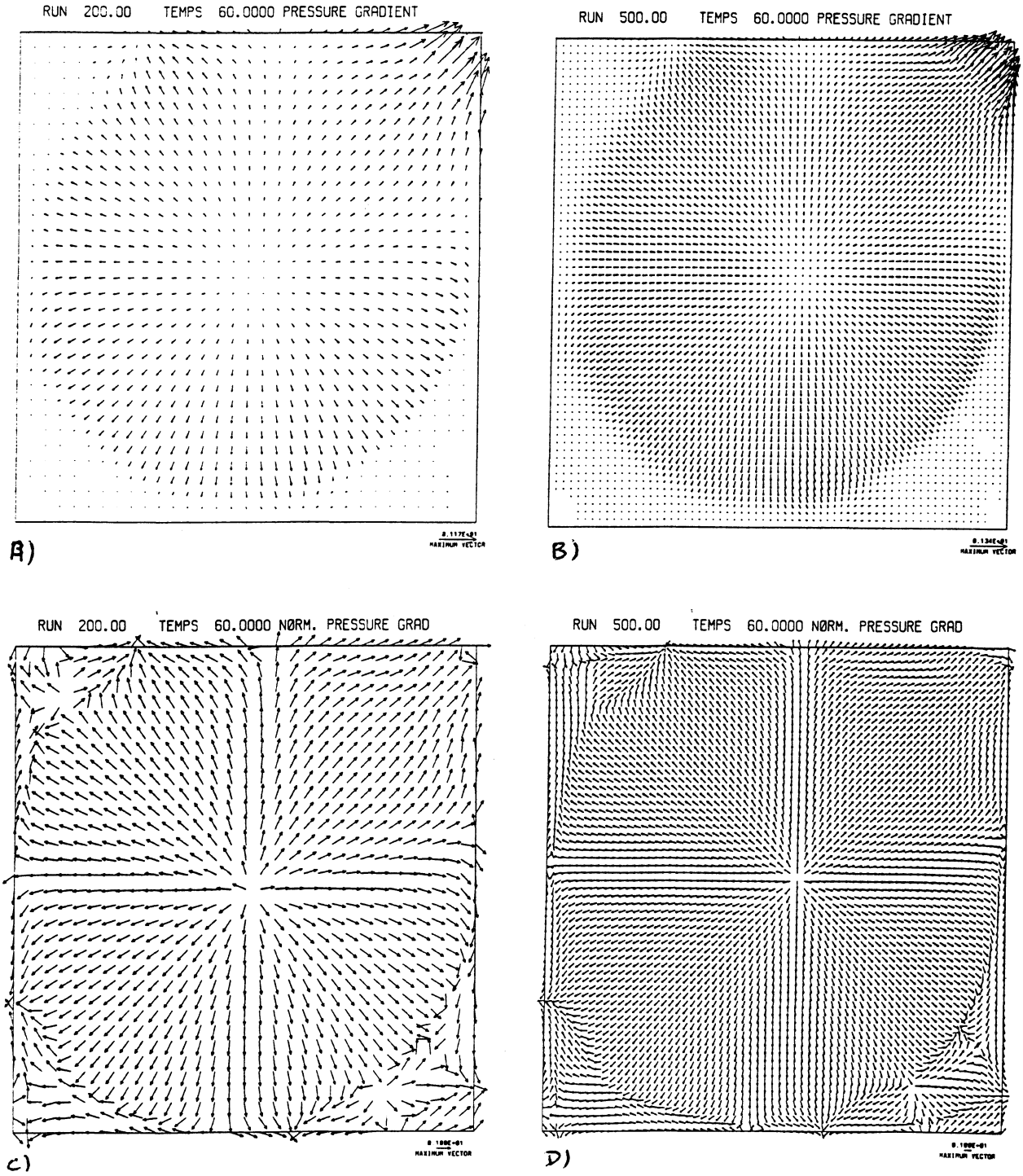


Figure II.19: *Cavité B) pour* $Re = 5000$ *:* $t = 60.0$ *; a) Champ de gradient de pression pour* 65^2 *; b) Champ de gradient de pression pour* 128^2 *; c) Champ de gradient de pression normalisé pour* 65^2 *; d) Champ de gradient de pression normalisé pour* 128^2.

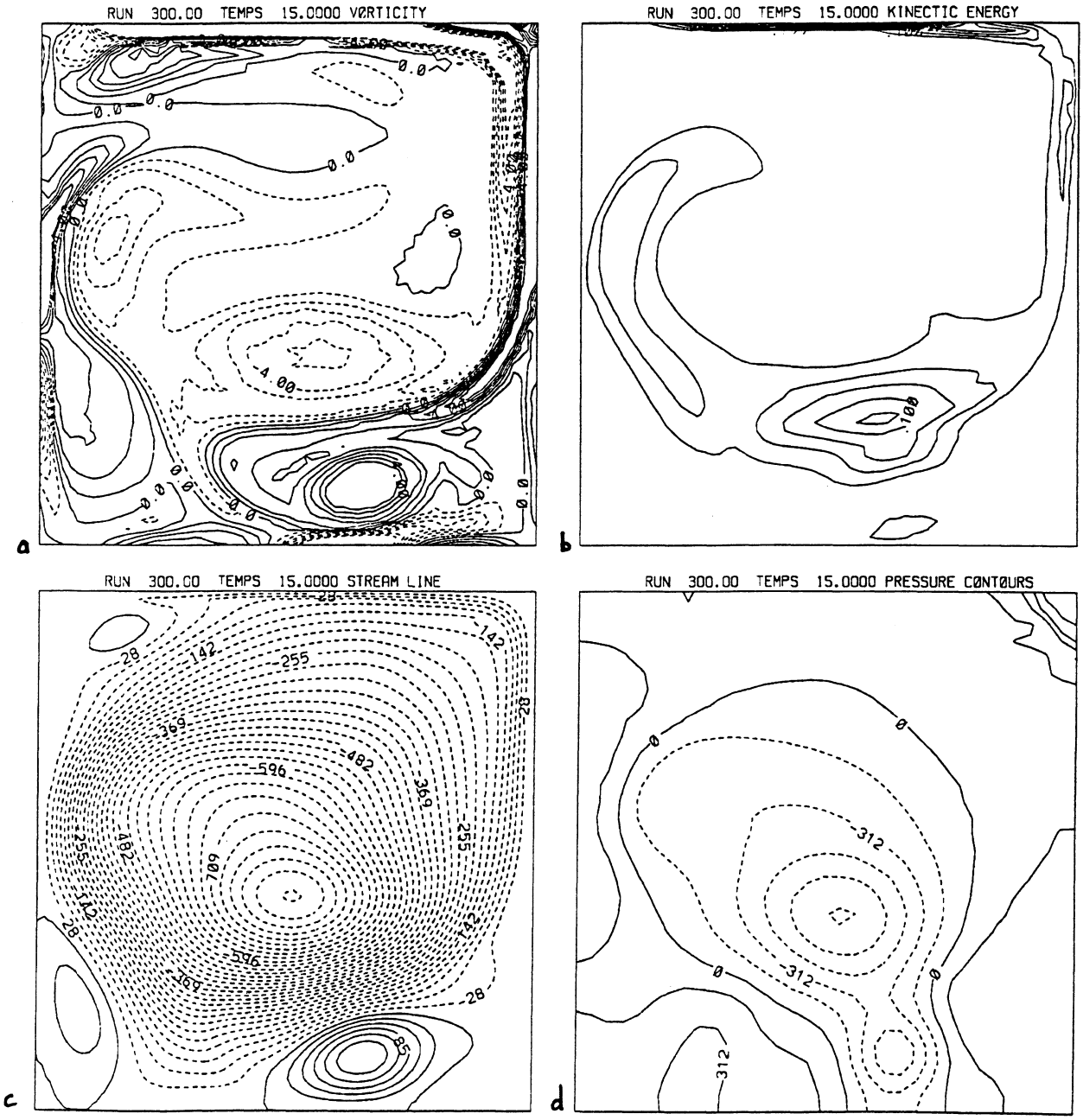


Figure II.20: *Cavité B) pour* $Re = 10000$ *à l'instant* $t = 15.0$ *;* $h = 1/64$ *; a) Isovorticité ; b) Energie ; c) Lignes de courant ; d) Isobares.*

Bibliographie

[II.1] OUVRAGES GENERAUX :

[II.2] R. Glowinski, *Numerical methods for non linear variational problems* Springer Verlag, 1984.

[II.3] R. Temam, *Navier–Stokes equations*, Amsterdam, North Holland, 1984.

[II.4] F. Thomasset, *Implementation of finite element methods for Navier–Stokes equations*, Springer Verlag, 1981.

[II.5] ELEMENTS FINIS :

[II.6] D. Arnold, F. Brezzi and M. Fortin, *A stable finite element for the Stokes equations*, Calcolo 21(4), 337-344, 1984.

[II.7] Cahouet et Chabard, *Some fast 3-D finite element solvers for Generalized Stokes problem*, Rapport EDF, HE/41/87.03, 1987.

[II.8] M. Crouzeix and P.A. Raviart, *Conforming and non conforming finite element methods for the stationnary Stokes equations*, RAIRO, R3, 33-76, 1973.

[II.9] M. Fortin, *Calcul numérique des écoulements des fluides de Bingham et des fluides newtoniens incompressibles par la méthode des éléments finis*, Thèse, Université Paris VI, 1972.

[II.10] P. Hood and G. Taylor, *Navier–Stokes equations using mixed interpolation*, in Finite element in flow problem, Oden ed., UAD Press, 1974.

[II.11] O.C. Zienkiewicz, *Why finite elements ?*, Finite Elements in Fluids, Vol 1, edited by Gallagher, Oden, Taylor, Zienkiewicz ; Wiley & Sons.

[II.12] ELEMENT P1-4P1 :

[II.13] M. bercovier and O. Pironneau, *Error estimates for finite element solution of the Stokes problem in the primitive variables*, Numer. Math. 33, 211-224, 1979.

[II.14] R. Glowinski and O. Pironneau, *On a mixed Finite element approximation of the Stokes problem*, Numer. Math., 33, 397-424, 1979.

[II.15] Le Tallec, *A mixed finite element approximation of the Navier-Stokes equation*, Numer. Math. 35, 381-404, 1980.

[II.16] MOINDRES CARRES :

[II.17] M.O. Bristeau, O. Pironneau, R. Glowinski, J. Periaux and P.Perrier, *On the numerical solution of non linear problems in fluid dynamics by least squares and finite element methods (1) least square formulations and conjugate gradient solution of the continuous problems*, Comp. Meth. in appl. Mech. and Eng. 17/18, 619-657, 1979.

[II.18] SCHEMAS EN TEMPS :

[II.19] A.J. Chorin, *Numerical solution of incompressible flow problems*, Numerical Analysis,2 , 64-71, 1968.

[II.20] R. Glowinski et J.F. Periaux, *Numerical methods for nonlinear problem in fluid dynamics*, Supercomputing, A. Lichnevsky, C. Saguez ed., North Holland, 1987.

[II.21] P.L. Lions, B. Mercier, *Splitting algorithms for the sum of two nonlinear operator*, SIAM J. Num. Anal. 16, 964-979, 1979.

[II.22] R. Temam, *Sur l' approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires(I)*, Arch. Rational Mech. Anal. 32, 1969, p135-153

[II.23] R. Temam, *Sur l' approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires(II)*, Arch. Rational Mech. Anal. 33, 1969, p377-385

[II.24] CAVITE ENTRAINEE :

[II.25] C.H. Bruneau, C. Jouron, *An efficient scheme for solving steady incompressible Navier-Stokes equations*, Jour. of Comp. Phys. 89, 389-413, 1990.

[II.26] Ghia, Ghia et Shin, *High-Re Solutions for incompressible Flow using the Navier-Stokes equations and a multigrid method*, J. of Comp. Phys. 48, 387-411, 1982.

[II.27] J.H. Goodrich, K. Gustafson, K. Halasi, *Hopf bifurcation in the driven cavity*, Jour. of Comp. Phys. 90, 219-261, 1990.

[II.28] J. shen, *Hopf bifurcation of the unsteady regularized driven cavity flow*, Jour. of Comp. Phys., 95, 228-245, 1991.

[II.29] J. shen, *Résolution numérique des équations de Stokes et Navier-Stokes par les méthodes spectrales*, thèse de 3e cycle, Université de Paris-Sud, 1987.

[II.30] L. Sonke Tabuguia, *Etude numérique des équations de Navier-Stokes en milieux multiplement connexes, en formulation vitesse-tourbillon, par une approche multidomaines*, thèse de 3e cycle, Université de Paris-Sud, 1989.

Chapitre III

LA BASE HIERARCHIQUE.

La cavité 2π périodique et la discrétisation spectrale ont des applications physiques et industrielles limitées. Dans le cadre d'une discrétisation spatiale par éléments finis, Marion et Temam [6] ont proposé de développer des idées similaires à celles exposées dans le chapitre I et introduites dans Marion et Temam [5] en s'appuyant sur des résultats de la théorie des systèmes dynamiques.

Une idée naturelle pour introduire une partition de l'espace de discrétisation par éléments finis en 2 sous-espaces supplémentaires consiste à utiliser le multigrille et la base hiérarchique, notion introduite par Zienkiewicz *et al* [11] et développée par Yserentant [9]. Cette partition induit une décomposition du champ calculé en petites et grandes structures interagissant non linéairement et pouvant être intégrées séparément. Cependant l'interprétation physique des petites structures liées au champ de vitesse et au champ de pression semble plus difficile que dans le cas spectral.

Dans un premier temps, nous décrivons et définissons la base hiérarchique dont on donne quelques propriétés essentielles accompagnées de leur démonstration. Puis dans une deuxième partie, est étudiée numériquement la décomposition sur la base hiérarchique de la solution du problème de Navier-Stokes a postériori i.e. à partir des résultats obtenus par la méthode de Galerkin classique décrite au chapitre II. L'évolution des petites structures et de l'ordre de grandeur des différents termes des équations projetées sur les 2 sous-espaces supplémentaires permettent d'établir quelques motivations numériques pour la méthode de Galerkin non linéaire.

Une méthode du type multigrille issue de la décomposition sur la base hiérarchique est ensuite développée et appliquée au problème de Dirichlet avec des conditions au bord homogènes. Des résultats similaires aux méthodes multigrilles classiques sont discutés.

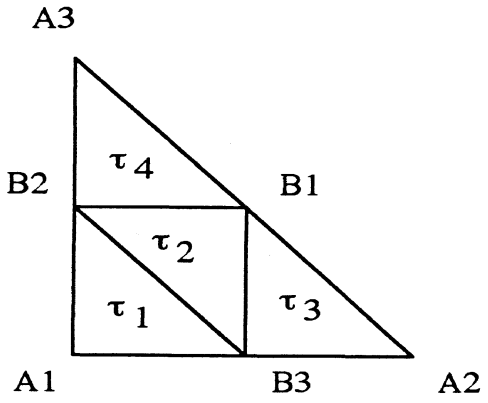


Figure III.1: *Division en 4 sous-triangles congruents d'un triangle τ de T_{2h}.*

III.1 Introduction de la base hiérarchique.

III.1.1 Définition.

Soit T_{2h} une triangulation régulière du domaine Ω (définie en II.3.1), on considère T_h la triangulation obtenue à partir de T_{2h} en divisant chaque triangle τ de T_{2h} en 4 sous-triangles congruents (cf. figure III.1).

Soit Σ_{2h} (respectivement Σ_h) l'ensemble des sommets des triangles du maillage T_{2h} (respectivement T_h). On pose :

$$n_{2h} = card(\Sigma_{2h}) \quad \text{et} \quad n_h = card(\Sigma_h).$$

Les noeuds i.e. les sommets des triangles de T_h créés lors du raffinement sont les éléments de $\Sigma_h \setminus \Sigma_{2h}$. La triangulation T_{2h} forme le maillage grossier et T_h le maillage fin. On suppose que les noeuds de Σ_{2h} sont numérotés de 1 à n_{2h} et que les noeuds de $\Sigma_h \setminus \Sigma_{2h}$ sont numérotés de $n_{2h}+1$ à n_h (voir l'exemple de numérotation de la cavité entrainée en figure III.2).

On associe à la triangulation T_h (respectivement T_{2h}) l'espace d'approximation par éléments finis V_h (respectivement V_{2h}) défini en II.3.1 :

$$\begin{aligned} V_h &= \{v_h \in C^0(\Omega) \ / \ \forall \tau \in T_h, \ v_{h|\tau} \in P_1(\tau)\} \\ V_{2h} &= \{v_h \in C^0(\Omega) \ / \ \forall \tau \in T_{2h}, \ v_{h|\tau} \in P_1(\tau)\}. \end{aligned}$$

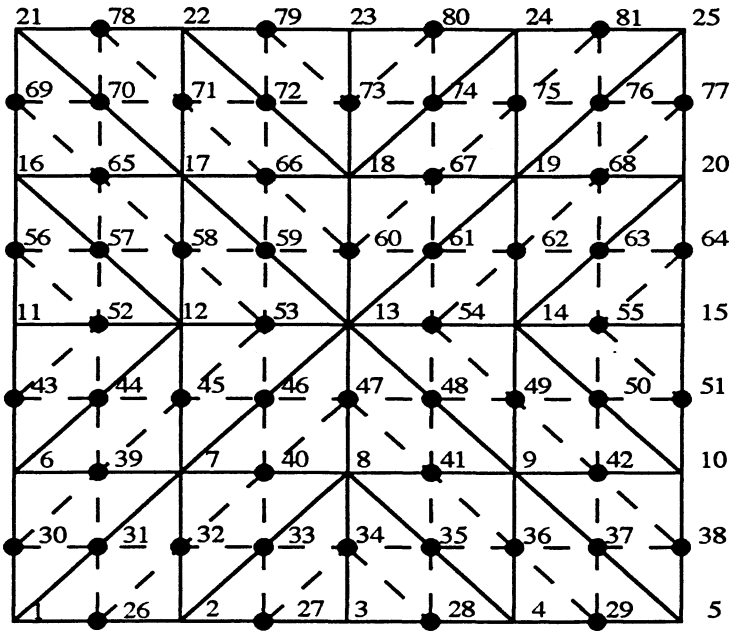


Figure III.2: *Numérotation de 1 à n_{2h} des noeuds de la triangulation T_{2h}. Numérotation de $n_{2h}+1$ à n_h des noeuds de $\Sigma_h \setminus \Sigma_{2h}$ ($n_h = 81$ et $n_{2h} = 25$).*

V_h est un espace de dimension finie n_h pour lequel on peut définir 2 types de bases :

1. **La base nodale** notée $\{\hat{\phi}_{h,i}\ ;\ i \in [1, n_h]\}$.
   Il s'agit de la base canonique où la fonction $\hat{\phi}_{h,i}$ associée au ieme sommet M_i est définie par :

$$\begin{array}{rl} \hat{\phi}_{h,i} \in V_h & \forall i = 1, \cdots, n_h \\ \hat{\phi}_{h,i}(M_i) = 1 & \\ \hat{\phi}_{h,i}(M_j) = 0 & \forall i, j = 1, \cdots, n_h \text{ avec } i \neq j \text{ et } (M_i, M_j) \in \Sigma_h^2. \end{array}$$

2. **La base hiérarchique** notée $\{\phi_{h,i}\ ;\ i \in [1, n_h]\}$ définie de la façon suivante :

   - pour $i = 1, \cdots, n_{2h}$: $\phi_{h,i} = \hat{\phi}_{2h,i}$ fonction de la base nodale de V_{2h} associée au ieme noeud de la grille grossière,
   - pour $i = n_{2h}+1, \cdots, n_h$: $\phi_{h,i} = \hat{\phi}_{h,i}$ fonction de la base nodale de V_h associée au ieme noeud de la grille fine.

   Il est à noter que les supports des fonctions de cette base sont de taille différente (cf. figure III.3). Une fonction peut en effet être différente de 0 sur un triangle ne contenant pas le noeud associé à cette fonction.

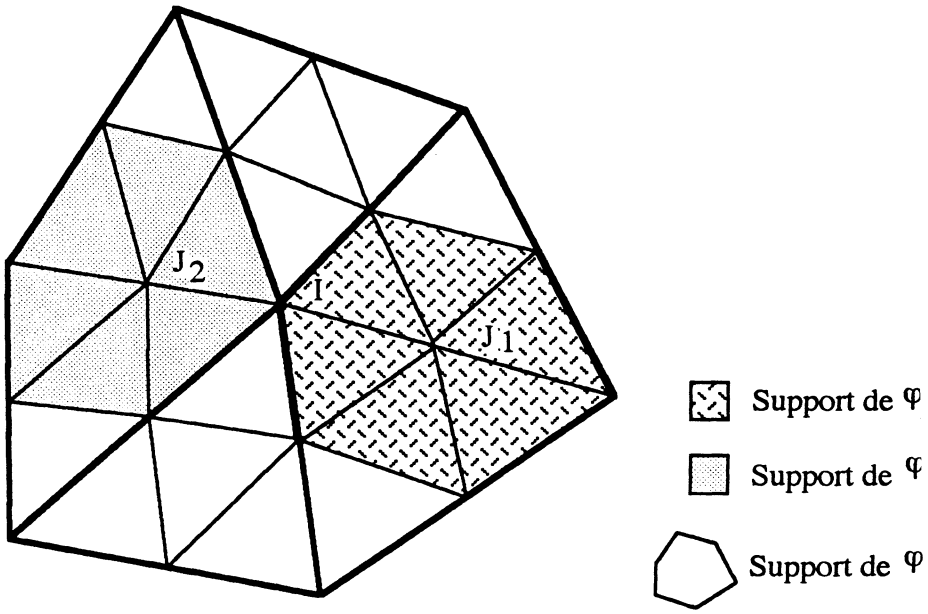


Figure III.3: *Support des fonctions de la base hiérarchique.*

Remarque III.1 *On peut prendre en toute généralité les triangulations T_{ph} et T_h ($p \geq 2$) où T_h est obtenue en divisant chaque triangle de T_{ph} en p^2 sous-triangles congruents. Dans ce cas la base hiérarchique $\{\phi_{h,i}\ ;\ i \in [1, n_h]\}$ est telle que :*

- *pour $i = 1, \cdots, n_{ph}$: $\phi_{h,i} = \hat{\phi}_{ph,i}$ fonction de la base nodale de V_{ph},*
- *pour $i = n_{ph}+1, \cdots, n_h$: $\phi_{h,i} = \hat{\phi}_{h,i}$ fonction de la base nodale de V_h.*

Par simplicité nous traiterons uniquement le cas $p = 2$ sauf pour étude numérique spécifique.

Remarque III.2 *Il est possible de définir récursivement la base hiérarchique sur V_h s'il y a plus de 2 triangulations i.e. niveaux ou grilles. Soient $T_{2^k h}, \cdots, T_{2h}, T_h$ les triangulations construites récursivement :*

- *$T_{2^k h}$ est une triangulation régulière de Ω,*
- *$T_{2^{j-1} h}$ est obtenue à partir de $T_{2^j h}$ en subdivisant chaque triangle de $T_{2^j h}$ en 4 sous-triangles congruents ($j \in [1, k]$).*

La base hiérarchique sur V_h est alors donnée récursivement par :

- *pour $i = 1, \cdots, n_{2h}$: $\phi_{h,i}$ sont les fonctions de la base hiérarchique de V_{2h},*
- *pour $i = n_{2h}+1, \cdots, n_h$: $\phi_{h,i} = \hat{\phi}_{h,i}$ fonction de la base nodale de V_h,*

- *la base hiérarchique du plus petit sous-espace $V_{2^k h}$ est égale à la base nodale de cet espace.*

La base hiérarchique induit ainsi un partitionnement de l'espace V_h :

$$V_h = V_{2h} \oplus W_h$$

où $W_h = Vect(\phi_{h,n_{2h}+1}, \cdots, \phi_{h,n_h})$. Soit $u_h \in V_h$, sa décomposition dans la base hiérarchique s'écrit alors :

$$u_h = y_h + z_h$$

c'est-à-dire :

$$u_h = \sum_{i=1}^{n_{2h}} y_i \hat{\phi}_{2h,i} + \sum_{i=n_{2h}+1}^{n_h} z_i \hat{\phi}_{h,i}$$

avec $y_h \in V_{2h}$ et $z_h \in W_h$ et où y_i et z_i sont définis de façon unique par :

$$\begin{aligned} \forall i \in [1, n_{2h}] \quad & y_i = u_h(A_i) \\ \forall i \in [n_{2h}+1, n_h] \quad & z_i = u_h(B_i) - \frac{u_h(A_{i_1}) + u_h(A_{i_2})}{2} \end{aligned}$$

Les sommets A_i, A_{i_1} et A_{i_2} sont des noeuds du maillage grossier (ils appartiennent donc à Σ_{2h}) et B_i, le milieu du côté $[A_{i_1}, Ai_2]$ est un noeud issu du raffinement (cf. figure III.1) i.e. un noeud de $\Sigma_h \setminus \Sigma_{2h}$.

Il est clair que y_h est du même ordre que u_h tandis que z_h est d'un ordre h^2 plus petit : par similitude avec le spectral, y_h est appelée grande structure et z_h petite structure. La figure III.4 représente le champ de vitesse pour la cavité entrainée B) avec $Re = 100$ projeté sur la base hiérarchique. On y remarque que z_i est généralement petit devant y_i.

Remarque III.3 *Plus précisément on a :*

$$z_i = -u''(B_i).(A_{i_1}\vec{A}_{i_2})^2 + O(||A_{i_1}\vec{A}_{i_2}||^3).$$

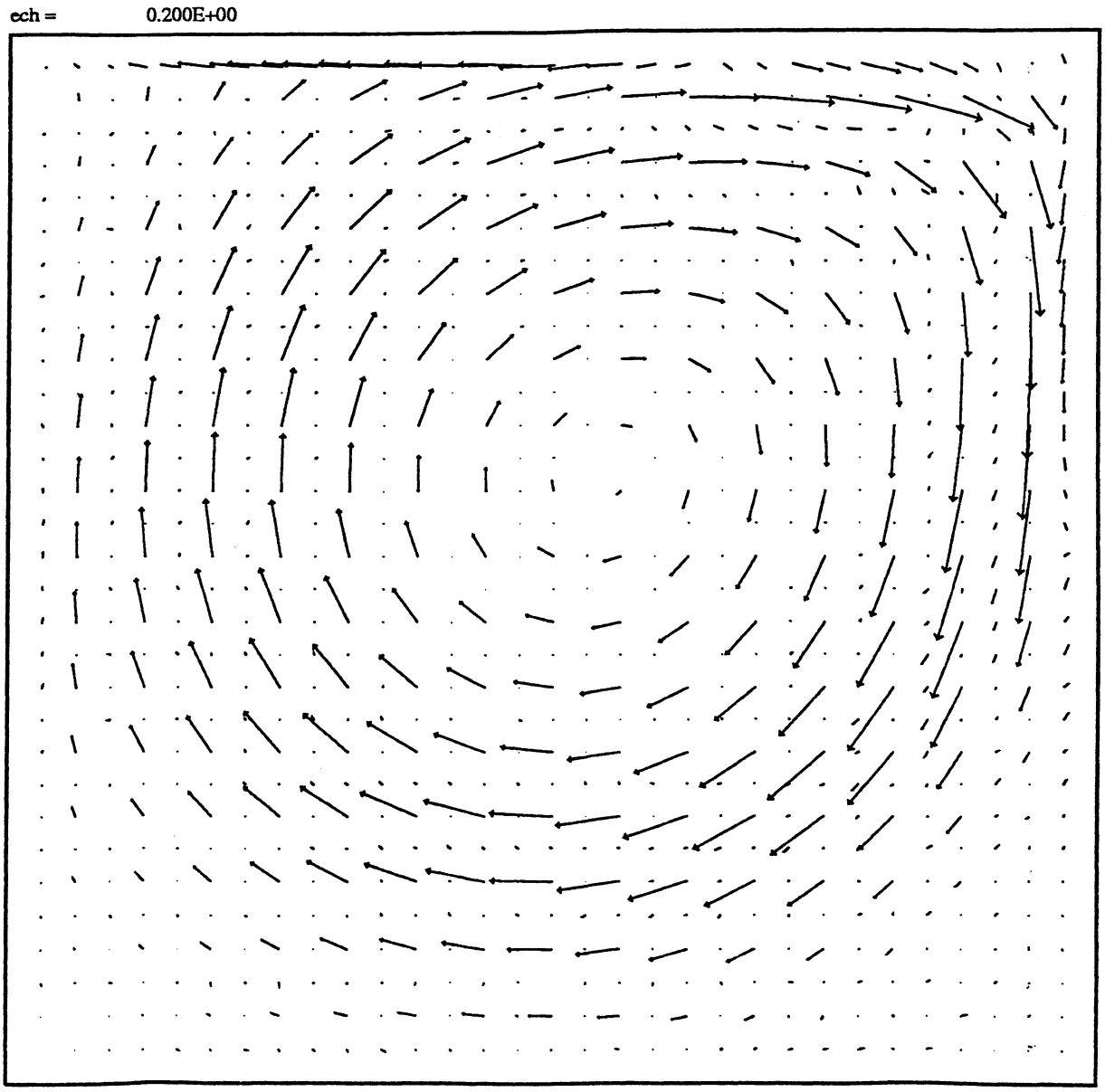


Figure III.4: *Champ de vitesse projeté sur la base hiérarchique pour la cavité entrainée B) avec* $Re = 100$, *le résultat étant obtenu par la méthode de Galerkin classique décrite au chapitre II.*

III.1.2 Propriétés de la base.

On note $(.,.)$ le produit scalaire de $L^2(\Omega)$ et $|.|$ la norme de $L^2(\Omega)$ et on pose

$$((u,v)) = (\nabla u, \nabla v) \text{ et } ||u|| = ((u,u)) \quad \forall u, v \in H^1(\Omega),$$

on a alors dans les espaces V_h, V_{2h} et W_h les relations suivantes :

$$\exists c_1 \in \mathbb{R} \text{ indépendant de } h \ / \ |u_h| \leq c_1 ||u_h|| \quad \forall u_h \in V_{0,h} \tag{III.1}$$

$$\exists S_1(h) \in \mathbb{R} \ / \ S_1(h)||u_h|| \leq |u_h| \quad \forall u_h \in V_h \tag{III.2}$$

$$\exists S_1(h) \longrightarrow 0 \text{ quand } h \to 0 \ / \ |z_h| \leq S_2(h)||z_h|| \quad \forall z_h \in W_h \tag{III.3}$$

$$\begin{array}{l} \exists \delta \text{ indépendant de } h \ / \ 0 < \delta < 1 \text{ et} \\ |((y_h, z_h))| \leq (1-\delta).||y_h||.||z_h|| \quad \forall y_h \in V_{2h} \quad \forall z_h \in W_h \end{array} \tag{III.4}$$

Si ρ_h est le maximun des diamètres et ρ'_h est le minimun des rondeurs des triangles de T_h, on a alors :

$$S_1(h) = c_2 \rho'_h \quad \text{et} \quad S_2(h) = c_3 \rho_h.$$

III.1.3 Démonstration des propriétés.

Inégalité (III.1) Il s'agit de l'inégalité de Poincaré : $c_1 = 2l$ où l est la plus grande largeur de Ω dans les directions x_1 et x_2.

Inégalité (III.2) Notons

$$\mu_h = \sup_{\{\tau \in T_h, \ \forall \phi \text{ linéaire sur } \tau\}} \frac{\int_\tau |\nabla \phi(x)|^2 \, dx}{\int_\tau |\phi(x)|^2 \, dx}$$

alors

$$\sup_{u_h \in V_h} \frac{||u_h||^2}{|u_h|^2} = \sup_{u_h \in V_h} \frac{\int_\Omega |\nabla u_h(x)|^2 \, dx}{\int_\Omega |u_h(x)|^2 \, dx} \leq \mu_h.$$

En effet pour tout u_h de V_h, on a

$$\int_\tau |\nabla u_{h|_\tau}(x)|^2 \, dx \leq \mu_h \int_\tau |u_{h|_\tau}(x)|^2 \, dx$$

d'où le résultat en sommant sur l'ensemble des triangles.

Afin d'estimer μ_h, fixons τ un triangle de T_h et ϕ une fonction linéaire sur τ. Soit F la fonction affine qui transforme le triangle de référence $\hat{\tau}$ en τ :

$$x = F(\hat{x}) = B\hat{x} + b.$$

Comme $\int_{\hat{\tau}} |\nabla \hat{\phi}(\hat{x})|^2 \, d\hat{x}$ et $\int_{\hat{\tau}} |\hat{\phi}(\hat{x})|^2 \, d\hat{x}$ sont une semi-norme et une norme sur l'espace de dimension finie des fonctions linéaires sur $\hat{\tau}$, on a pour toute fonction affine $\hat{\phi}$ sur $\hat{\tau}$:

$$\exists \hat{\mu} \in \mathbb{R} \ / \ \int_{\hat{\tau}} |\nabla \hat{\phi}(\hat{x})|^2 \, d\hat{x} \leq \hat{\mu} \int_{\hat{\tau}} |\hat{\phi}(\hat{x})|^2 \, d\hat{x}. \tag{III.5}$$

Or $\phi \circ F$ est linéaire sur $\hat{\tau}$, (III.5) implique :

$$\int_{\hat{\tau}} (|\frac{\partial \phi(F\hat{x})}{\partial \hat{x}_1}|^2 + |\frac{\partial \phi(F\hat{x})}{\partial \hat{x}_2}|^2) \, d\hat{x} \leq \hat{\mu} \int_{\hat{\tau}} |\phi(F\hat{x})|^2 \, d\hat{x}. \tag{III.6}$$

Comme

$$\frac{\partial \phi(F\hat{x})}{\partial x_i} = \sum_{j=1}^{2} \frac{\partial \phi(F\hat{x})}{\partial \hat{x}_j} . F^{-1}_{ji},$$

on en déduit que :

$$\sum_{j=1}^{2} |\frac{\partial \phi(F\hat{x})}{\partial x_i}|^2 \leq ||F^{-1}||^2 \sum_{j=1}^{2} |\frac{\partial \phi(F\hat{x})}{\partial \hat{x}_j}|^2.$$

Par suite en utilisant la relation (III.6),

$$\int_{\hat{\tau}} (|\frac{\partial \phi(F\hat{x})}{\partial x_1}|^2 + |\frac{\partial \phi(F\hat{x})}{\partial x_2}|^2) \, d\hat{x} \leq \hat{\mu} ||F^{-1}||^2 \int_{\hat{\tau}} |\phi(F\hat{x})|^2 \, d\hat{x}$$

mais $||F^{-1}|| \leq \frac{\rho_{\hat{\tau}}}{\rho'_{\tau}}$ (cf. Raviart et Thomas [7]), d'où par changement de variables dans les intégrales

$$\frac{\int_{\tau} |\nabla \phi(x)|^2 \, dx}{\int_{\tau} |\phi(x)|^2 \, dx} \leq \hat{\mu} \frac{\rho^2_{\hat{\tau}}}{\rho'^2_{\tau}} \leq \hat{\mu} \frac{\rho^2_{\hat{\tau}}}{\rho'^2_{h}}.$$

Par suite $\mu_h \leq c_2^{-2} \rho'^{-2}_h$, ce qui entraine l'inégalité (III.2).

Inégalité (III.3) Soit τ un triangle de T_{2h}, si on a

$$\int_{\tau} |z_h(x)|^2 \, dx \leq c \rho^2_h \int_{\tau} |\nabla z_h(x)|^2 \, dx \tag{III.7}$$

comme

$$\Omega = \bigcup_{\tau \in T_h} \tau \text{ et } \rho_h = \sup_{\tau \in T_h} \rho_\tau,$$

on en déduit :

$$\int_{\Omega} |z_h(x)|^2 \, dx \leq c \rho^2_h \int_{\Omega} |\nabla z_h(x)|^2 \, dx.$$

Montrons (III.7). Soit $\hat{\tau}$ le triangle de référence et F la fonction affine qui transforme $\hat{\tau}$ en τ. On note $\hat{z}_h = z_h \circ F$. Comme $z_h(\hat{A}_i) = 0$ pour tout i, il est clair que $\hat{z}_h(\hat{A}_i) = 0$.

Par suite l'inégalité de Poincaré appliquée à $\hat{z}_h$, fonction de W_h, permet d'écrire :

$$\int_{\hat{\tau}} |\hat{z}_h(\hat{x})|^2 \, d\hat{x} \leq c \int_{\hat{\tau}} \sum_{i=1}^{2} |\frac{\partial \hat{z}_h}{\partial \hat{x}_i}(\hat{x})|^2 \, d\hat{x}. \qquad \text{(III.8)}$$

Or

$$\frac{\partial \hat{z}_h}{\partial \hat{x}_i}(\hat{x}) = \sum_{j=1}^{2} \frac{\partial \hat{z}_h}{\partial x_j}(\hat{x}).F_{ji},$$

d'où

$$\sum_{i=1}^{2} |\frac{\partial \hat{z}_h}{\partial \hat{x}_i}(\hat{x})|^2 \leq ||F||^2 \sum_{j=1}^{2} |\frac{\partial \hat{z}_h}{\partial x_j}(\hat{x})|^2.$$

Or $||F|| \leq \frac{\rho_\tau}{\rho'_{\hat{\tau}}}$, d'où par changement de variables dans les intégrales :

$$\int_{\tau} |z_h(x)|^2 \, dx \leq c\rho_\tau^2 \int_{\tau} |\nabla z_h(x)|^2 \, dx.$$

Inégalité (III.4) Soit τ un triangle de T_{2h}, rappelons les principales étapes de la démonstration (effectuée dans Marion et Temam [6]) du résultat suivant :

$$\left| \int_{\tau} \nabla y . \nabla z \, dx \right| \leq (1 - \delta) . \left(\int_{\tau} |\nabla y|^2 \, dx \right)^{1/2} \left(\int_{\tau} |\nabla z|^2 \, dx \right)^{1/2}$$

où $1 - \delta = (1 - \frac{\eta}{2})^{1/2}$, η étant le réel $(0 < \eta < 1)$ caractérisant la minoration des angles de la triangulation régulière T_h (cf. II.3.1) ie :

$\forall \tau$ triangle de T_h, $\dfrac{\rho(\tau)}{\rho'(\tau)} \leq K$ où K est une constante indépendante de h.

Cette propriété est en effet équivalente à :

$$\exists \eta, \ 0 < \eta < 1 \ / \ \forall \theta \text{ angle de la triangulation } |\cos \theta| \leq 1 - \eta.$$

Soient $\lambda_1(x)$, $\lambda_2(x)$, $\lambda_3(x)$ les coordonnées barycentriques associées à B_1, B_2, B_3 milieux des côtés de τ (cf. figure III.1). Supposons que le triangle τ soit défini par $A_1(0,0)$, $A_2(b,0)$ et $A_3(c,d)$ alors :

$$\begin{aligned}
\lambda_1(x_1, x_2) &= \frac{2}{b} x_1 - \frac{2(c-b)}{db} x_2 - 1 \\
\lambda_2(x_1, x_2) &= 1 - \frac{2}{b} x_1 + \frac{2c}{db} x_2 \\
\lambda_3(x_1, x_2) &= 1 - \frac{2}{d} x_2
\end{aligned}$$

On a donc :

$$\nabla\lambda_1 = \begin{pmatrix} \frac{2}{b} \\ -\frac{2(c-b)}{db} \end{pmatrix} \text{ et } \nabla\lambda_2 = \begin{pmatrix} -\frac{2}{b} \\ \frac{2c}{db} \end{pmatrix} \tag{III.9}$$

Ce qui permet d'écrire :

$$\begin{aligned} |\nabla\lambda_1.\nabla\lambda_2| &= |4\frac{d^2 + c(c-b)}{b^2d^2}| \\ |\nabla\lambda_1|^2 &= |4\frac{d^2 + (c-b)^2}{b^2d^2}| \\ |\nabla\lambda_2|^2 &= |4\frac{d^2 + c^2}{b^2d^2}| \end{aligned}$$

d'où

$$\frac{|\nabla\lambda_1.\nabla\lambda_2|}{|\nabla\lambda_1|^2|\nabla\lambda_2|^2} = \frac{|c^2 + d^2 - cb|}{\sqrt{(d^2 + c^2)(d^2 + (c-b)^2)}}$$

C'est exactement $|\cos(A_1\widehat{A_3}A_2)|$ et par suite

$$\frac{|\nabla\lambda_1.\nabla\lambda_2|}{|\nabla\lambda_1|^2|\nabla\lambda_2|^2} \leq 1 - \eta. \tag{III.10}$$

Si l'on pose $y_h(A_i) = y_i$ et $z_h(B_i) = z_i$, y_h et z_h s'expriment alors en fonction des coordonnées barycentriques λ_1, λ_2, et λ_3 de la façon suivante :

- sur le triangle τ

$$\begin{aligned} y_h &= \frac{1}{2}(y_3 - y_1)\lambda_1 + \frac{1}{2}(y_3 - y_2)\lambda_2 + \frac{1}{2}(y_1 + y_2) \\ \nabla y_h &= \frac{1}{2}(y_3 - y_1)\nabla\lambda_1 + \frac{1}{2}(y_3 - y_2)\nabla\lambda_2 \end{aligned}$$

- sur le triangle $B_1B_2B_3$

$$z_h = z_1\lambda_1 + z_2\lambda_2 + z_3\lambda_3 \quad \text{et} \quad \nabla z_h = (z_1 - z_3)\nabla\lambda_1 + (z_2 - z_3)\nabla\lambda_2$$

- sur le triangle $A_1B_2B_3$

$$z_h = (z_2 + z_3)\lambda_1 + z_2\lambda_2 + z_3\lambda_3 \quad \text{et} \quad \nabla z_h = z_2\nabla\lambda_1 + (z_2 - z_3)\nabla\lambda_2$$

- sur le triangle $A_2B_1B_3$

$$z_h = z_1\lambda_1 + (z_1 + z_3)\lambda_2 + z_3\lambda_3 \quad \text{et} \quad \nabla z_h = (z_1 - z_3)\nabla\lambda_1 + z_1\nabla\lambda_2$$

- sur le triangle $A_3B_1B_2$

$$z_h = z_1\lambda_1 + z_2\lambda_2 + (z_1 + z_2)\lambda_3 \quad \text{et} \quad \nabla z_h = -z_2\nabla\lambda_1 + -z_1\nabla\lambda_2$$

Les intégrales $\int_\tau |\nabla y_h|^2\,dx$, $\int_\tau |\nabla z_h|^2\,dx$ et $\int_\tau \nabla y_h.\nabla z_h\,dx$ sont alors calculées en fonction de y_i et z_i ; l'inégalité de Schwarz et l'inégalité (III.10) permettent de conclure.

III.2 Equations de Navier-Stokes et base hiérarchique.

Dans ce paragraphe, nous étudions pour le problème de Navier-Stokes les grandes et petites structures (à un instant donné ou l'évolution au cours du temps) issues de la décomposition du champ de vitesse et de la pression sur la base hiérarchique. Ces analyses sont effectuées à partir de résultats numériques obtenues par la méthode de Galerkin classique développée au second chapitre pour la cavité entrainée régularisée.

On supposera pour simplifier les notations que les conditions au bord sont les conditions homogènes :

$$u = 0 \quad \text{sur } \partial\Omega.$$

III.2.1 Décomposition sur la base hierarchique.

Les notations sont identiques à celles décrites au chapitre II section 3. On considère que T_{2h} est une triangulation régulière et que la triangulation T_h est obtenue par subdivision en 4 sous-triangles congruents des triangles de T_{2h}. On suppose désormais que T_{2h} est issue de T_{4h} (triangulation régulière) par raffinement.

Rappelons que la discrétisation par élément fini 4P1-P1 du problème de Navier-Stokes consiste à chercher $u_h(t)$ dans $\mathcal{V}_{0,h} = V_{0,h} \times V_{0,h}$ et $p_h(t)$ dans $\mathcal{Q}_h = V_{2h}$ comme solutions du système :

$$\left\{ \begin{array}{rcl} (\frac{\partial u_h}{\partial t}, v_h) + \frac{1}{Re}(\nabla u_h, \nabla v_h) + ((u_h.\nabla)u_h, v_h) + (\nabla p_h, v_h) & = & (f, v_h) \\ (\nabla.u_h, q_h) & = & 0 \\ \forall v_h \in \mathcal{V}_{0,h} \quad \text{et} \quad \forall q_h \in \mathcal{Q}_h. & & \end{array} \right. \tag{III.11}$$

Munissons $V_{0,h}$ de la base hiérarchique $\{\phi_{h,i} \ ; \ i \in [1, n_{0,h}]\}$ (on supposera que les noeuds appartenant au bord sont numérotés de $n_{0,h}+1$ à n_h) définie par :

- pour $i = 1, \cdots, n_{0,2h} : \phi_{h,i} = \hat{\phi}_{2h,i}$
- pour $i = n_{0,2h}+1, \cdots, n_{0,h} : \phi_{h,i} = \hat{\phi}_{h,i}$.

De même on munit V_{2h} de la base hiérarchique $\{\phi_{2h,i} \ ; \ i \in [1, n_{2h}]\}$ où

- pour $i = 1, \cdots, n_{4h} : \phi_{2h,i} = \hat{\phi}_{4h,i}$
- pour $i = n_{4h}+1, \cdots, n_{2h} : \phi_{2h,i} = \hat{\phi}_{2h,i}$

On a donc le partitionnement des espaces d'approximation par éléments finis :

$$\mathcal{V}_{0,h} = \mathcal{V}_{0,2h} \oplus \mathcal{W}_{0,h} \quad \text{et} \quad \mathcal{Q}_h = V_{4h} \oplus W_{2h},$$

où $\mathcal{V}_{0,2h} = V_{0,2h} \times V_{0,2h}$ et $\mathcal{W}_{0,h} = W_{0,h} \times W_{0,h}$.

Soient $u_h \in \mathcal{V}_{0,h}$ et $p_h \in \mathcal{Q}_h$, leurs décompositions dans les bases hiérarchiques s'écrivent :

$$u_h = uy_h + uz_h \quad \text{avec} \quad \begin{cases} uy_h = \begin{pmatrix} uy_h^1 \\ uy_h^2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n_{2h}} uy_i^1 \, \hat{\phi}_{2h,i} \\ \sum_{i=1}^{n_{2h}} uy_i^2 \, \hat{\phi}_{2h,i} \end{pmatrix} & \in \mathcal{V}_{0,2h} \\ uz_h = \begin{pmatrix} uz_h^1 \\ uz_h^2 \end{pmatrix} = \begin{pmatrix} \sum_{i=n_{2h}+1}^{n_h} uz_i^1 \, \hat{\phi}_{h,i} \\ \sum_{i=n_{2h}+1}^{n_h} uz_i^2 \, \hat{\phi}_{h,i} \end{pmatrix} & \in \mathcal{W}_{0,h} \end{cases}$$

$$p_h = py_h + pz_h \quad \text{avec} \quad \begin{cases} py_h = \sum_{i=1}^{n_{4h}} py_i \, \hat{\phi}_{4h,i} & \in V_{4h} \\ pz_h = \sum_{i=n_{4h}+1}^{n_{2h}} pz_i \, \hat{\phi}_{h,i} & \in W_{2h}. \end{cases}$$

Par suite le système (III.11) se réécrit sous la forme proposée par Laminie *et al* [4] :

$$\begin{cases} (\frac{\partial(uy_h + uz_h)}{\partial t}, vy_h) + \\ \frac{1}{Re}(\nabla uy_h, \nabla vy_h) + \frac{1}{Re}(\nabla uz_h, \nabla vy_h) + \\ (\nabla py_h, vy_h) + (\nabla pz_h, vy_h) + \\ (u_h.\nabla u_h, vy_h) = (f, vy_h) \\ (\nabla.uy_h + \nabla.uz_h, qy_h) = 0 \\ \forall vy_h \in \mathcal{V}_{0,2h} \ \text{ et } \ \forall qy_h \in V_{4h} \end{cases} \tag{III.12}$$

$$\begin{cases} (\frac{\partial(uy_h + uz_h)}{\partial t}, vz_h) + \\ \frac{1}{Re}(\nabla uy_h, \nabla vz_h) + \frac{1}{Re}(\nabla uz_h, \nabla vz_h) + \\ (\nabla py_h, vz_h) + (\nabla pz_h, vz_h) + \\ (u_h.\nabla u_h, vz_h) = (f, vz_h) \\ (\nabla.uy_h + \nabla.uz_h, qz_h) = 0 \\ \forall vz_h \in \mathcal{W}_{0,h} \ \text{ et } \ \forall qz_h \in W_{2h} \end{cases} \tag{III.13}$$

Remarque III.4 *Si les conditions au bord sont du type $u = g$ alors l'espace $\mathcal{V}_{0,h}$ est remplacé par $\mathcal{V}_{g,h}$ vérifiant :*

$$\mathcal{V}_{g,h} = \{v_h \ / \ v_h \in \mathcal{V}_h, \ v_h = g_h \ sur \ \partial\Omega\},$$

g_h étant une approximation de g telle que (on note n la normale extérieure à Ω) :

$$\int_{\partial\Omega} g_h.n\, d\sigma = 0.$$

La décomposition sur la base hiérarchique est alors dans ce cas :

$$\mathcal{V}_{g,h} = \mathcal{V}_{g,2h} \oplus \mathcal{W}_{g,h}$$

où

$$\begin{aligned} \mathcal{V}_{g,2h} &= \{vy_h \;/\; vy_h \in \mathcal{V}_{2h},\ vy_h = gy_h \ sur\ \partial\Omega\} \\ \mathcal{W}_{g,h} &= \{vz_h \;/\; vz_h \in \mathcal{W}_h,\ vz_h = gz_h \ sur\ \partial\Omega\} \end{aligned}$$

les fonctions gy_h et gz_h satisfaisant :

$$\int_{\partial\Omega} (gy_h + gz_h).n\, d\sigma = 0.$$

Dans le cas de la cavité entrainée, l'approximation g_h est définie dans la base nodale par :

$$g_h = \sum_{A_i \in \partial\Omega} g(A_i)\hat{\phi}_{h,i}$$

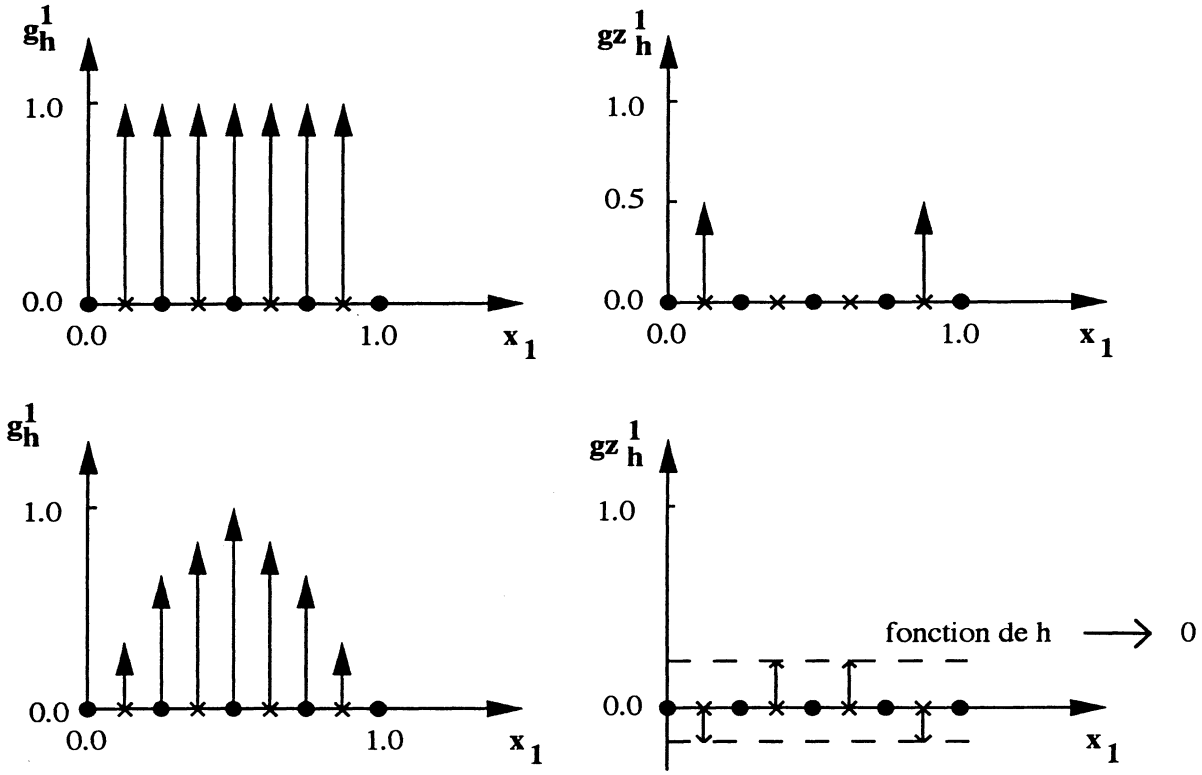
et dans la base hiérarchique par :

$$\begin{aligned} g_h &= gy_h + gz_h \\ g_h &= \sum_{\left\{\begin{array}{l} A_i \in \partial\Omega \\ A_i \in \Sigma_{2h} \end{array}\right.} g(A_i)\,\phi_{h,i} \\ &+ \sum_{\left\{\begin{array}{l} B_i \in \partial\Omega \\ B_i \in \Sigma_h \setminus \Sigma_{2h} \\ B_i = milieu[A_{i_1}, A_{i_2}] \end{array}\right.} \{g(B_i) - \frac{1}{2}g(A_{i_1}) - \frac{1}{2}g(A_{i_2})\}\,\phi_{h,i} \end{aligned}$$

Il est clair que non seulement $\int_{\partial\Omega} g_h.n\, d\sigma = 0$, mais on a aussi dans ce cas particulier :

$$\int_{\partial\Omega} gy_h.n\, d\sigma = 0 \quad et \quad \int_{\partial\Omega} gz_h.n\, d\sigma = 0$$

Il est à noter (cf. figure III.5) que pour la cavité entrainée A) où la vitesse d'entrainement est uniformément égale à 1.0, la première composante de gz_h est partout nulle sauf en 2 points de la paroi supérieure où elle vaut 0.5, quantité indépendante de h. Dans ce cas bien que gz_h tende vers 0 lorsque h tend vers 0 en norme L_2, gz_h ne tend pas vers 0 ponctuellement. C'est par contre le cas pour la cavité régularisée B). C'est une des raisons du choix de cette cavité.



A) Base nodale B) Base hiérarchique

Figure III.5: *Première composante de la vitesse sur le bord supérieur :*
a) $g_h^1 = gy_h^1 + gz_h^1$ *dans la base nodale ;* *b)* gz_h^1 *dans la base hiérarchique.*

III.2.2 Décomposition de la solution stationnaire.

Les figures III.6 à III.11 montrent le champ de vitesse de la solution stationnaire pour la cavité régularisée B) dans des parties du domaine qui présentent de forts gradients (telles le coin inférieur droit ou le coin supérieur gauche) pour des nombres de Reynolds faibles (100 et 400) et plus forts (1000 et 5000) avec des maillages plus ou moins grossiers. Ce champ de vitesse est tantôt représenté dans la base nodale (le mouvement du fluide y est alors clairement visible), tantôt représenté dans la base hiérarchique (des petites et grandes "vitesses" apparaissent correspondant généralement aux petites et grandes structures uz_h et uy_h).

Pour $Re = 100$ (figure III.6), bien que le maillage soit grossier : $h = \frac{1}{64}$ (il est suffisant pour la méthode de Galerkin classique) l'ordre de grandeur de uy_h (vitesse dont les flèches ont pour origine une croix) est nettement supérieur à l'ordre de grandeur de uz_h (y compris près du bord supérieur). Il est à noter que uz_h ne présente pas une très grande "continuité" ; uz_h semble "partir" dans toutes les directions suivant les noeuds ce qui peut s'expliquer par le fait que uz_h n'est pas une vitesse mais la projection d'une dérivée seconde sur des directions variables en fonction du point.

Lorsque l'on compare le champ de vitesse (exprimé dans la base hiérarchique) de la solution stationnaire pour $Re = 400$ et pour $Re = 100$, on constate que l'amplitude de uy_h pour $Re = 400$ diminue légèrement par rapport à $Re = 100$ (en ce qui concerne du moins le coin supérieur gauche du domaine) mais surtout

l'amplitude de uz_h augmente jusqu'à doubler en particulier près du bord supérieur de la cavité (cf. figure III.6 et III.7), le phénomène de discontinuité en amplitude et en direction s'amplifiant.

La différence d'ordre entre uy_h et uz_h étant localement claire pour une telle précision du maillage et de tels Reynolds (elle sera confirmée avec l'évolution au cours du temps des normes L_2 à la section III.2.3), il est raisonnable d'envisager d'intégrer ces 2 termes séparément ie de considérer le système (III.12) comme étant un système en (uy_h, py_h) avec (uz_h, pz_h) fixé et de considérer le système (III.13) comme étant un système en (uz_h, pz_h) avec (uy_h, py_h) fixé. Certaines approximations s'avérant vraies, il est alors possible d'obtenir l'équation d'une variété inertielle approximative approchant l'attracteur.

Les figures III.8 et III.9 pour $Re = 1000$ et III.10 et III.11 pour $Re = 5000$ montrent qu'il en est autrement pour des nombres de Reynolds plus importants et pour des maillages alors trop grossiers.

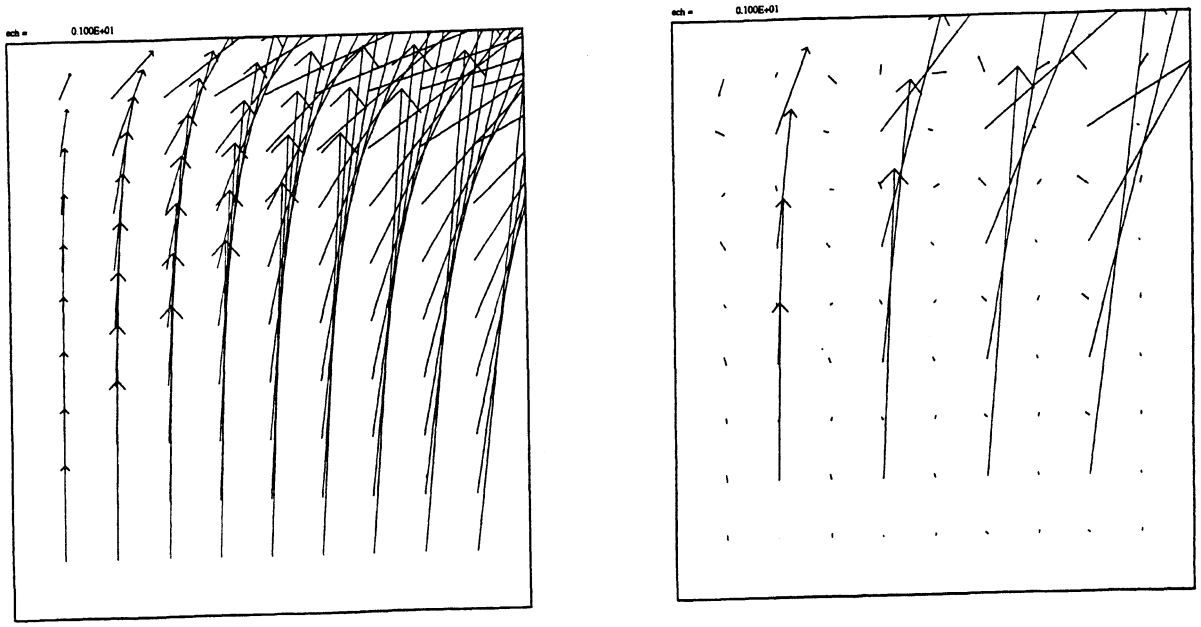
Remarquons tout d'abord en figure III.8 que $|uz_h(A)| \gg |uz_h(B)|$. Ceci illustre la remarque (III.3) sur uz_h : A se situe au milieu de l'hypoténuse d'un triangle du maillage tandis que B est le milieu d'un côté de ce triangle. D'autre part $uy_h(C)$ est en module plus petit que $uz_h(C_1)$, $uz_h(C_2)$, $uz_h(C_3)$ et $uz_h(C_4)$, où C_1, C_2, C_3 et C_4 sont des points voisins de C détérminés sur la figure III.8. La présence du contre-tourbillon et de forts gradients de la vitesse (assez mal simulés à cause de la faible précision en espace) au voisinage de C imposent de grandeurs valeurs locales à uz_h qui reste néanmoins petit en norme L_2 (i.e. globalement). Cela ne contredit pas l'estimation (III.3) :

$$|uz_h| \leq S_2(h)|\nabla uz_h|.$$

En effet le gradient de uz_h bien que borné peut être très grand.

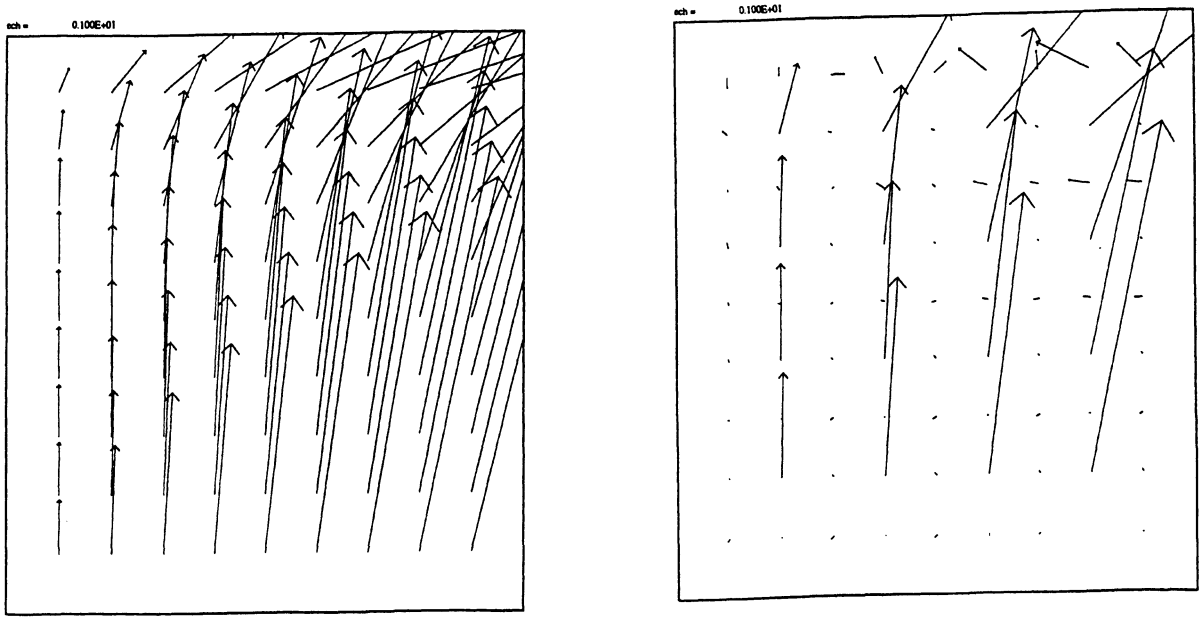
Comme uy_h est indépendant de h tandis que uz_h en dépend quadratiquement (cf. remarque III.3) et comme près du bord supérieur le gradient de vitesse est important, il est nécessaire de réduire le pas d'espace pour simuler correctement les structures de l'écoulement et pour que uz_h diminue de façon significative. C'est ce que l'on constate sur la figure III.8 où l'échelle de représentation est identique pour les 4 illustrations et où le pas a été divisé par 2. On notera que uz_h est particulièrement petit dans les zones régulières de l'écoulement.

Lorsque le nombre de Reynolds est égal à 5000 (et de manière générale lorsqu'il augmente) la composante uz_h de la solution stationnaire peut être importante et fortement discontinue (il suffit de regarder ces structures sur les figures III.10 et III.11) et ce malgré un maillage fin (par exemple $h = 0.0078$) dès que l'on se situe dans une zone de fort gradient. Il est certainement impossible de négliger certains termes des systèmes en uy_h (III.12) et uz_h (III.13) ; il faut en tenir compte lors de la mise en place des schémas numériques tels que la méthode de Galerkin non linéaire ou bien envisager un maillage plus fin (localement et adaptatif par exemple).



a) b)

Figure III.6: *Re* = 100 *: coin supérieur gauche de la cavité entrainée B). Champ de vitesse a) en base nodale ; b) en base hiérarchique.*



a) b)

Figure III.7: *Re* = 400 *: coin supérieur gauche de la cavité entrainée B). Champ de vitesse a) en base nodale ; b) en base hiérarchique.*

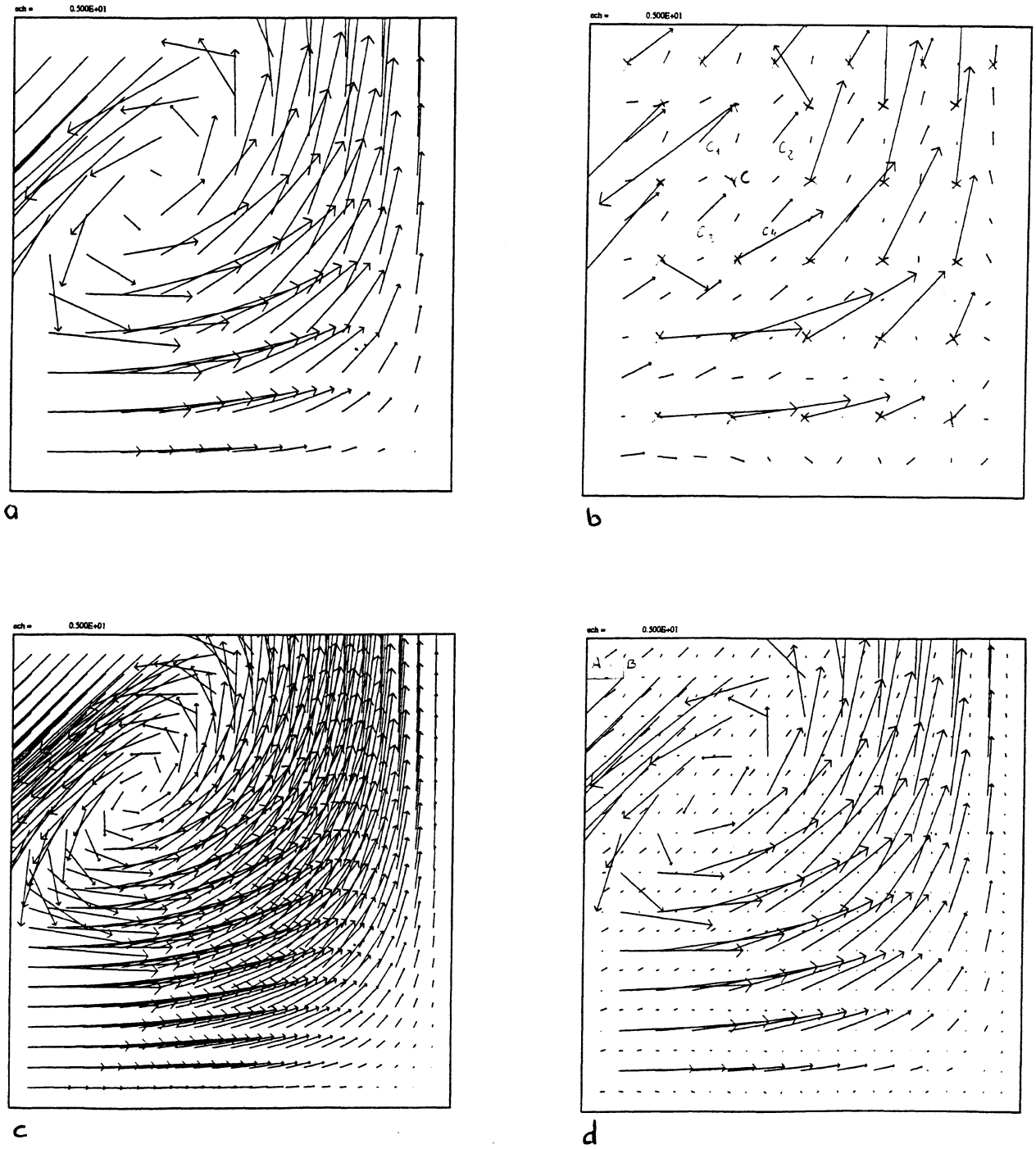


Figure III.8: *Re* = 1000 : *coin inférieur droit de la cavité entrainée B). Champ de vitesse décomposé sur la base nodale et la base hiérarchique pour* $h = \frac{1}{64}$ *(a, b) et* $h = \frac{1}{128}$ *(c, d).*

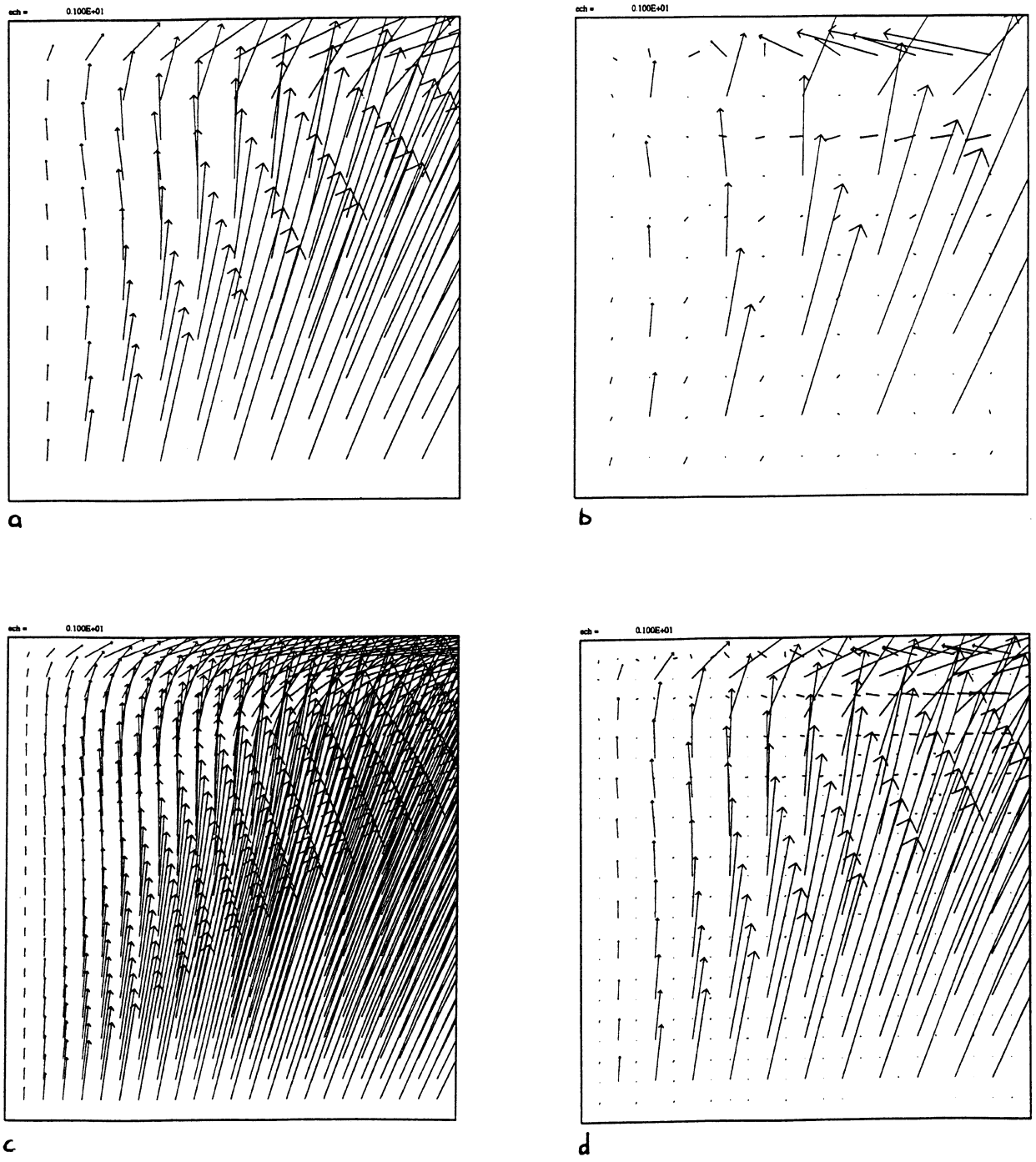


Figure III.9: *Re* = 1000 *: coin supérieur gauche de la cavité entrainée B). Champ de vitesse décomposé sur la base nodale et la base hiérarchique pour* $h = \frac{1}{64}$ *(a, b) et* $h = \frac{1}{128}$ *(c, d).*

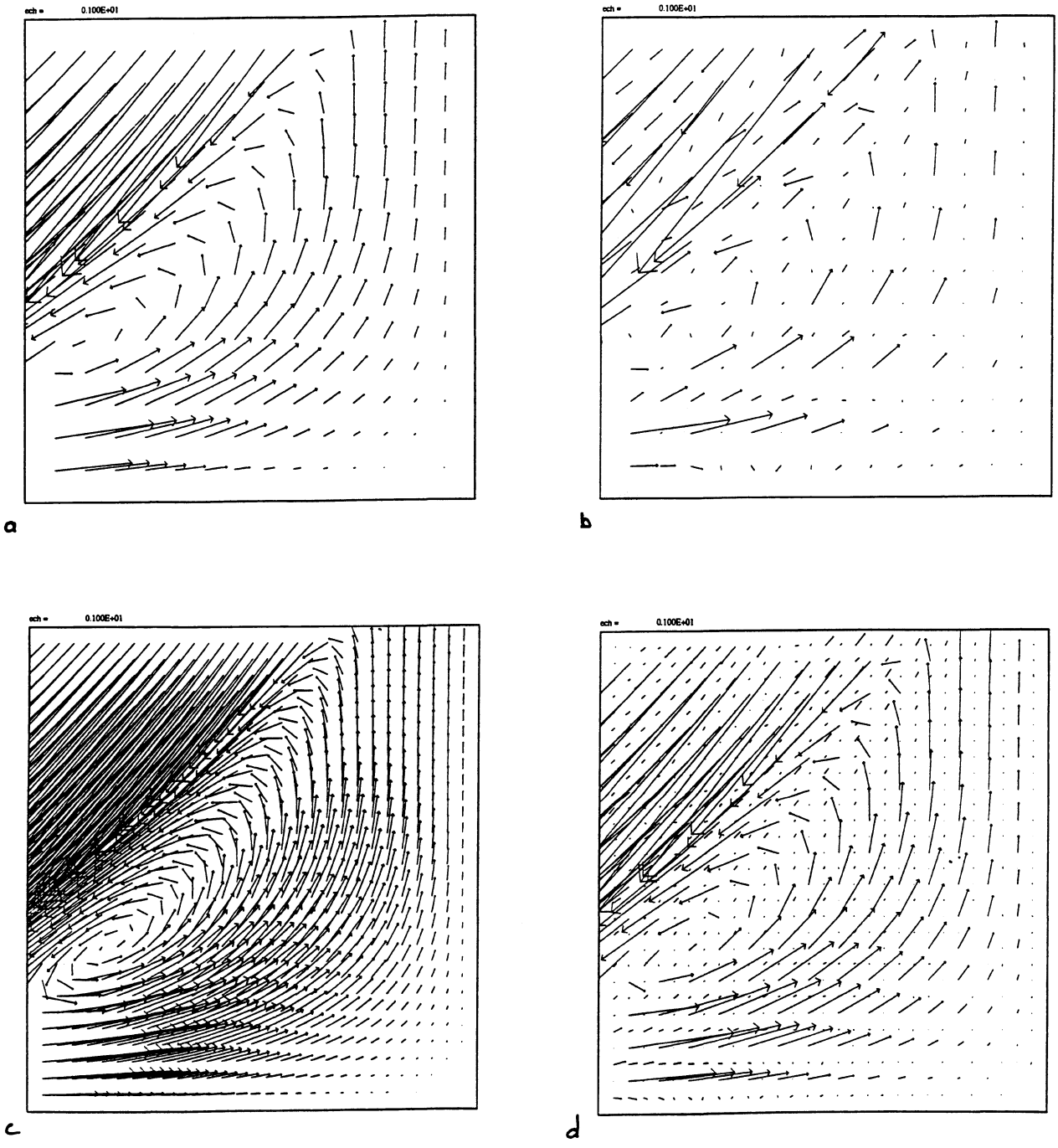


Figure III.10: *Re* = 5000 *: coin inférieur droit de la cavité entrainée B). Champ de vitesse décomposé sur la base nodale et la base hiérarchique pour* $h = \frac{1}{64}$ *(a, b) et* $h = \frac{1}{128}$ *(c, d).*

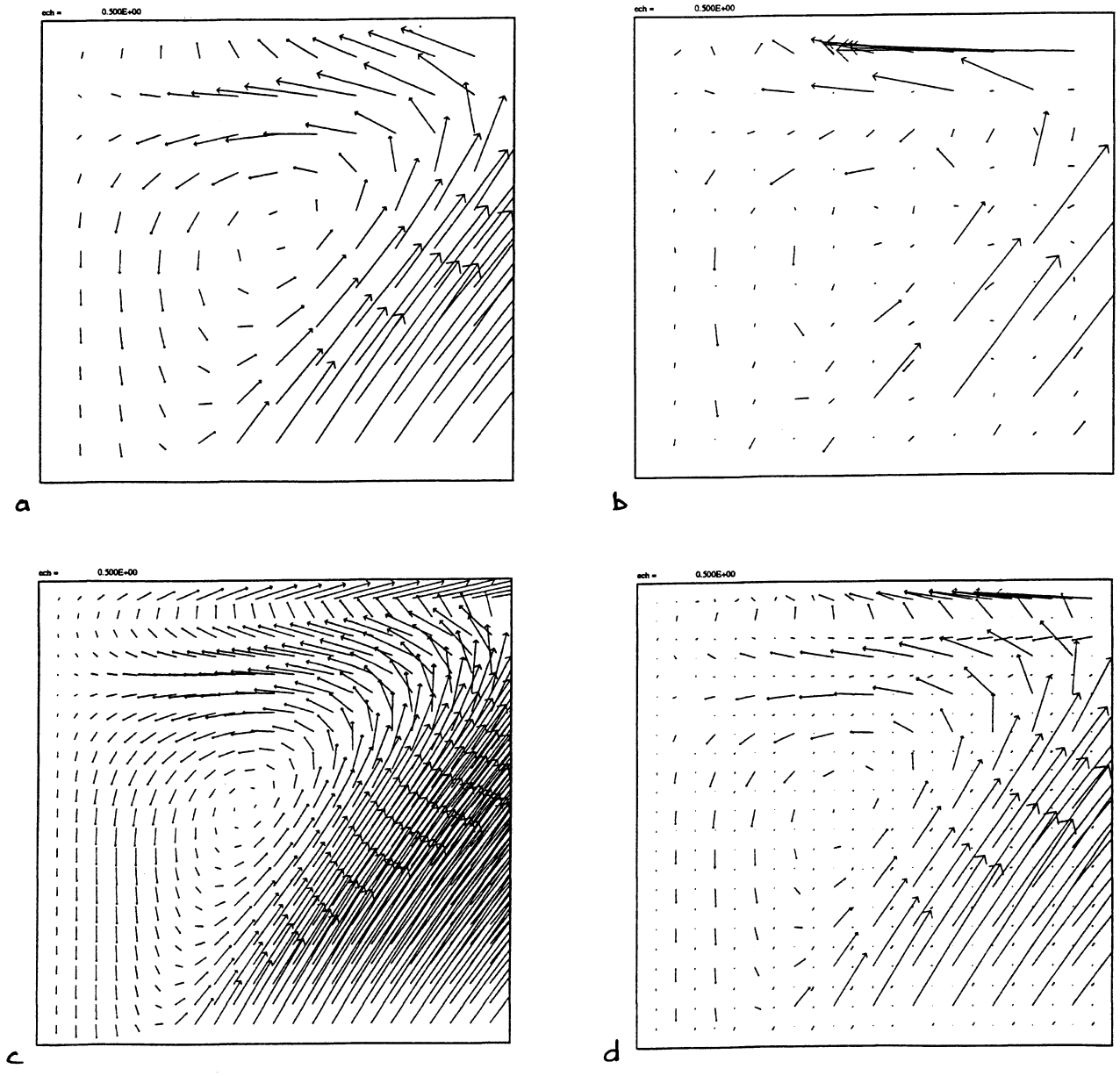


Figure III.11: *Re* = 5000 *: coin supérieur gauche de la cavité entrainée B). Champ de vitesse décomposé sur la base nodale et la base hiérarchique pour* $h = \frac{1}{64}$ *(a, b) et* $h = \frac{1}{128}$ *(c, d).*

III.2.3 Evolution des petites et grandes structures.

L'objet de ce paragraphe est d'analyser le comportement des différents termes des systèmes (III.12) et (III.13) issus de la projection des équations de Navier-Stokes sur la base hiérarchique, lorsque la solution évolue vers la solution stationnaire, la condition initiale étant la solution du problème de Stokes i.e. à divergence nulle. La décomposition sur la base hiérarchique est réalisée a postériori.

Norme l_2 des petites et grandes structures.

L'évolution de la norme l_2 des composantes de uy_h et uz_h pour $Re = 100$ avec un pas de discrétisation $h = \frac{1}{64}$, pour $Re = 5000$ avec $h = \frac{1}{64}$ et pour $Re = 5000$ avec $h = \frac{1}{128}$ est proposée dans les figures III.12 à III.14. Il s'agit des quantités :

$$|uy_h^n|_{l_2} = \left(\frac{\sum_{i=1}^{n_{2h}}(uy_i^n)^2}{n_{2h}}\right)^{1/2} \qquad |uz_h^n|_{l_2} = \left(\frac{\sum_{i=n_{2h}+1}^{n_h}(uz_i^n)^2}{n_h - n_{2h}}\right)^{1/2}$$

ainsi que du rapport $\frac{|uz_h^n|_{l_2}}{|uy_h^n|_{l_2}}$ pour $n = 1,\ 2$.

Remarque III.5 *La norme l_2 est différente de la norme L_2.*

$$|v_h|_{l_2} \neq |v_h| = \left(\int_\Omega v_h^2(x)\,dx\right)^{1/2}$$

L'ordre de grandeur de la norme l_2 des 2 composantes du champ uy_h est indépendante du maillage et du nombre de Reynolds (pour les nombres considérés). Les valeurs sont de l'ordre de 0.170-0.173 (1e composante) et 0.107-0.112 (2e composante) pour $Re = 100$, 0.165-0.190 (1e composante) et 0.105-0.165 (2e composante) pour $Re = 5000$ avec un pas en espace $h = \frac{1}{64}$, 0.155-0.175 (1e composante) et 0.110-0.165 (2e composante) pour $Re = 5000$ avec un pas de discrétisation $h = \frac{1}{128}$. On remarque que ces normes restent constantes au cours du temps pour de faibles nombres de Reynolds et ne présentent que de très légères variations pour des valeurs plus élevées.

Il en est tout autrement pour les composantes de uz_h. Leur norme est d'une part divisée environ par 3.0-3.5 (on attend 4 d'après la remarque III.3) lorsque h le pas de discrétisation en espace est réduit d'un facteur 2 (voir figure III.13 et III.14). D'autre part, elle est d'autant plus faible que Re est lui-même faible i.e. que le champ de vitesse (et par conséquent l'écoulement du fluide) est plus régulier. Il est à noter qu'initialement c'est-à-dire pour le problème de Stokes associé, les normes $|uz_h^1|_{l_2}$ et $|uz_h^2|_{l_2}$ sont très petites, négligeables devant h : dans ce cas le gradient de vitesse est en effet très faible. Pour $Re = 100$, ces normes tendent ensuite rapidement vers une limite (celle de la solution stationnaire).

Pour $Re = 5000$, la valeur initiale est multipliée par 10, puis les normes oscillent et convergent vers une limite à partir de l'instant $t = 40$.

L'évolution de la norme des composantes de uz_h, comme le champ de vitesse de la solution stationnaire dans la section précédente, montre l'importance du nombre de Reynolds et donc de la régularité ou de la non régularité de l'écoulement dans la relation (III.3) et dans son interprétation en terme d'ordre de grandeur des petites structures. Elle confirme cependant la dépendance quadratique en h.

La méthode de Galerkin non linéaire (voir l'introduction théorique au chapitre I) consiste à chercher u_h sur une variété inertielle approximative construite en tenant compte de ces petites structures et de la possibilité de négliger certains termes dans les équations (III.12) et (III.13). Il semble donc naturel de regarder :

- $(uy_h.\nabla)uy_h$: le terme non linéaire associé à uy_h
- les termes d'interaction entre uy_h et uz_h.

Les termes non linéaires.

En figure III.15, est représentée en échelle logarithmique l'évolution au cours du temps des normes suivantes :

$$||(f_h.\nabla)g_h^n||_c = \left(\frac{\sum_{i=1}^{n_{2h}}((f_h.\nabla)g_h^n, \hat{\phi}_{2h,i})^2}{n_{2h}}\right)^{1/2} \quad \text{pour} \quad n = 1,\ 2$$

avec

$$\begin{aligned} f_h &= uy_h,\ g_h = uy_h \\ f_h &= uy_h,\ g_h = uz_h \\ f_h &= uz_h,\ g_h = uy_h \\ f_h &= uz_h,\ g_h = uz_h. \end{aligned}$$

Il s'agit de la norme L_2 de la projection de $(f_h.\nabla)g_h^n$ sur $\mathcal{V}_{2h}$ notée *projection sur la grille grossière.* La norme L_2 de la projection de $(f_h.\nabla)g_h^n$ sur $\mathcal{W}_h$ appelée *projection sur la grille fine* et donnée par :

$$||(f_h.\nabla)g_h^n||_f = \left(\frac{\sum_{i=n_{2h}+1}^{n_h}((f_h.\nabla)g_h^n, \hat{\phi}_{h,i})^2}{n_h - n_{2h}}\right)^{1/2} \quad \text{pour} \quad n = 1,\ 2$$

est légèrement inférieure à la norme de la projection sur $\mathcal{V}_{2h}$ mais son comportement est identique.

La figure III.15 permet d'établir les relations :

$$||(uy_h.\nabla)uy_h^n|| \geq ||(uy_h.\nabla)uz_h^n|| \geq ||(uz_h.\nabla)uy_h^n|| \geq ||(uz_h.\nabla)uz_h^n||.$$

En particulier le dernier terme non linéaire par rapport aux petites structures est négligeable devant tous les autres pour $Re = 100$ et $Re = 1000$. Pour

$Re = 100$ les termes $(uy_h.\nabla)uz_h^n$ et $(uz_h.\nabla)uy_h^n$ sont également négligeables. Pour des nombres de Reynolds plus élevés, $(uy_h.\nabla)uy_h$ est de moins en moins prépondérant. Dans le cas de la cavité entrainée avec $Re = 5000$, il semble difficile de négliger telle ou telle partie du terme non linéaire étant donné leur valeur a postériori si la discrétisation en espace est trop grossière. La figure III.16 pour $h = \frac{1}{128}$ et $Re = 5000$ permet cependant d'envisager une telle possibilité si h est suffisamment petit. Il est à noter que des oscillations de ces termes apparaissent jusqu'à $t = 35$ pour $Re = 5000$ le régime stationnaire s'établissant par la suite.

Les termes de divergence.

Comme $\operatorname{div} u_h = 0$ i.e. $\operatorname{div} uy_h + \operatorname{div} uz_h = 0$, on déduit que :

$$(\operatorname{div} uy_h, \hat{\phi}_{4h,i}) = -(\operatorname{div} uz_h, \hat{\phi}_{4h,i}) \quad \forall i \in [1, n_{4h}]$$

$$(\operatorname{div} uz_h, \hat{\phi}_{2h,i}) = -(\operatorname{div} uy_h, \hat{\phi}_{2h,i}) \quad \forall i \in [n_{4h}+1, n_{2h}]$$

et par suite

$$\left(\frac{\sum_{i=1}^{n_{4h}}(\operatorname{div} uy_h, \hat{\phi}_{4h,i})^2}{n_{4h}}\right)^{1/2} = \left(\frac{\sum_{i=1}^{n_{4h}}(\operatorname{div} uz_h, \hat{\phi}_{4h,i})^2}{n_{4h}}\right)^{1/2}$$

$$\left(\frac{\sum_{i=n_{4h}+1}^{n_{2h}}(\operatorname{div} uy_h, \hat{\phi}_{2h,i})^2}{n_{2h}-n_{4h}}\right)^{1/2} = \left(\frac{\sum_{i=n_{4h}+1}^{n_{2h}}(\operatorname{div} uz_h, \hat{\phi}_{2h,i})^2}{n_{2h}-n_{4h}}\right)^{1/2}$$

C'est exactement ce que prouve la figure III.17 qui donne le rapport de ces quantités au cours du temps. De plus l'ordre de grandeur de chacun ce ces termes est très faible (de l'ordre de 10^{-5}).

La méthode de Galerkin non linéaire a pour objectif d'intégrer uy_h en tenant compte de la correction que constitue la composante uz_h de la vitesse. Un problème encore ouvert se pose : peut-on simplifier le sytème (III.12) en posant

$$(\operatorname{div} uy_h, qy_h) = 0 \quad \forall qy_h \in V_{4h} \ ?$$

Les termes de diffusion et du gradient de pression.

Comme le montre la figure III.18 avec $Re = 5000$ les rapports

$$\frac{(\sum_{i=1}^{n_{2h}}(\nabla uz_h, \nabla\hat{\phi}_{2h,i})^2)^{1/2}}{(\sum_{i=1}^{n_{2h}}(\nabla uy_h, \nabla\hat{\phi}_{2h,i})^2)^{1/2}} \quad \text{et} \quad \frac{(\sum_{i=n_{2h}+1}^{n_h}(\nabla uz_h, \nabla\hat{\phi}_{h,i})^2)^{1/2}}{(\sum_{i=n_{2h}+1}^{n_h}(\nabla uy_h, \nabla\hat{\phi}_{h,i})^2)^{1/2}}$$

i.e. les rapports des normes l_2 des projections des laplaciens sur $\mathcal{V}_{2h}$ et $\mathcal{W}_h$ sont élevés, dépassent (c'est le cas du second) la valeur 1 et diminuent avec h prouvant ainsi que les petites structures ont un rôle non négligeable en particulier dans le phénomène de diffusion. Ce rapport certe plus faible avec $Re = 100$ reste important, de l'ordre de 0.6.

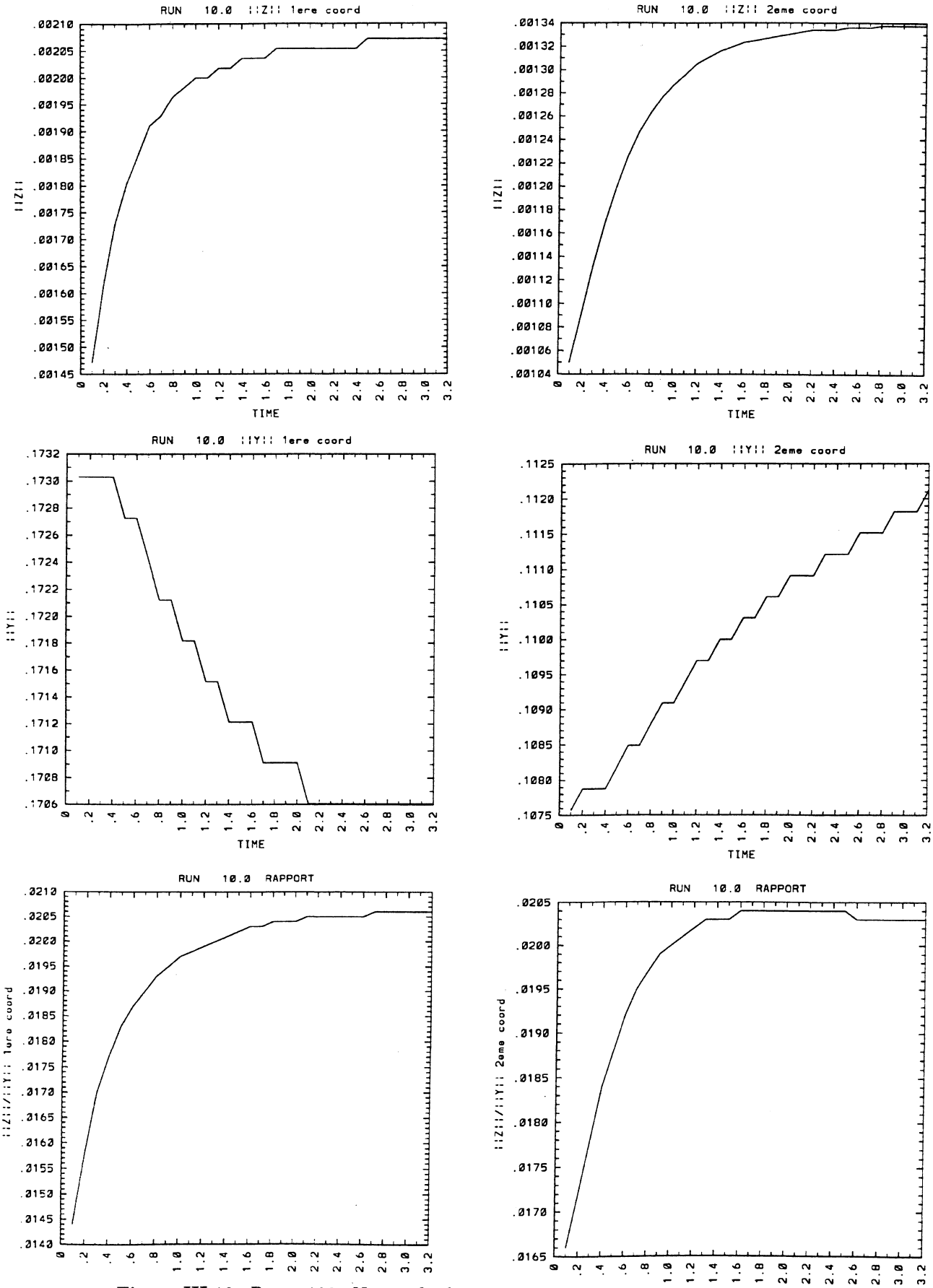


Figure III.12: *Re* = 100. *Norme* l_2 *des composantes de* uy_h *et* uz_h *et rapport de ces normes au cours du temps.*

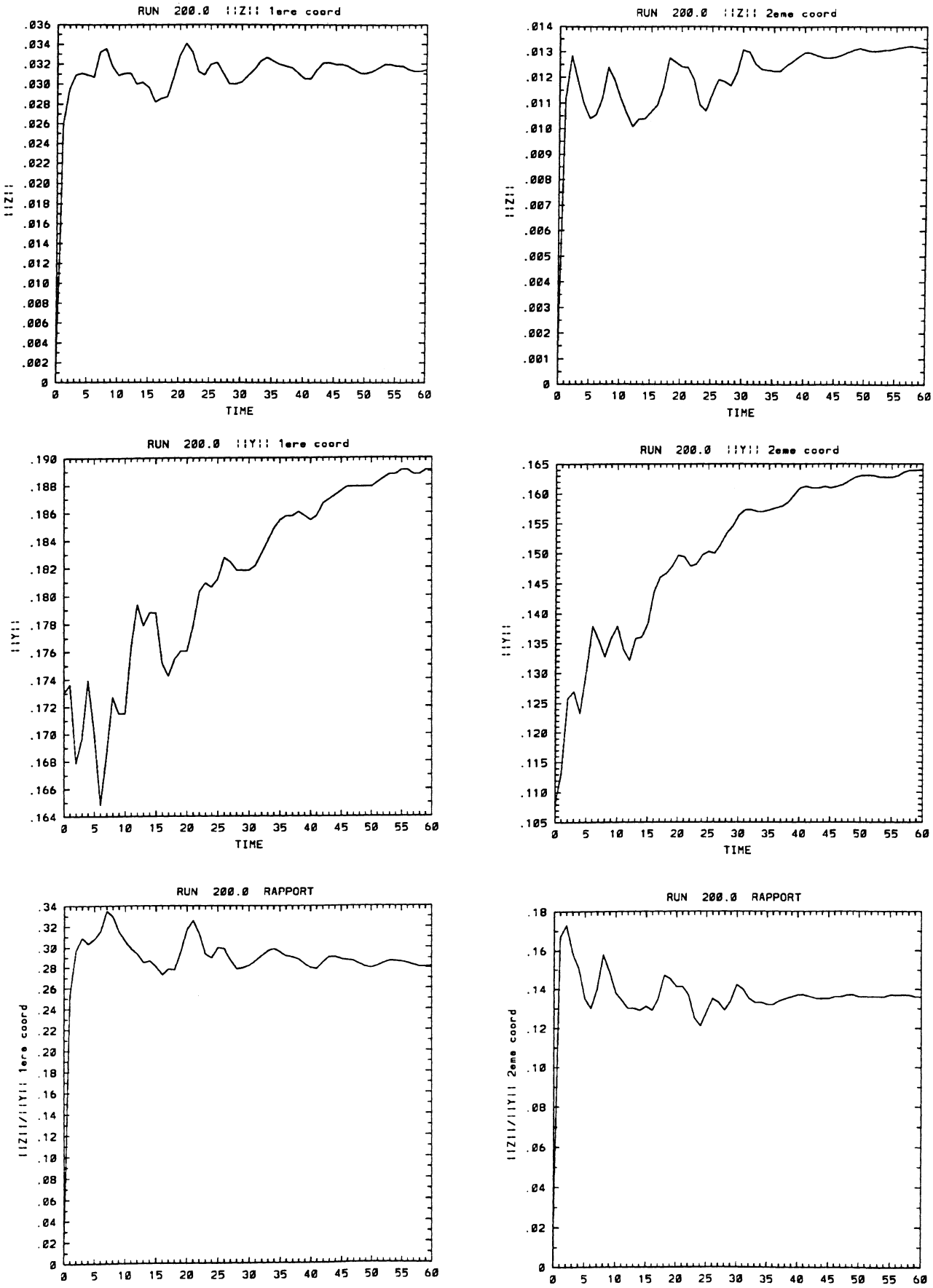
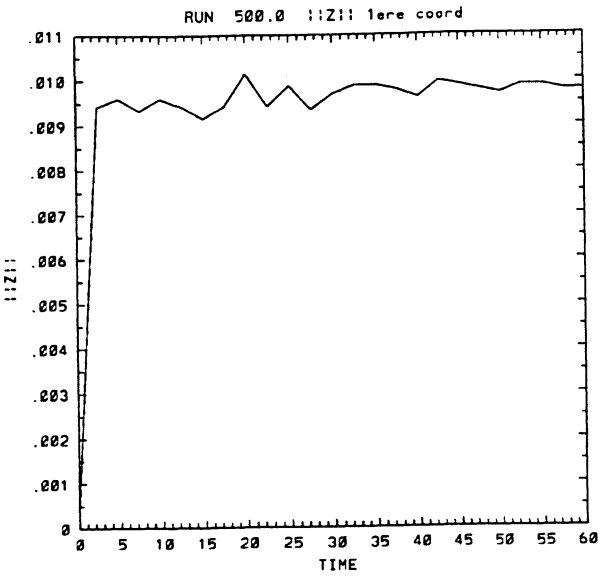
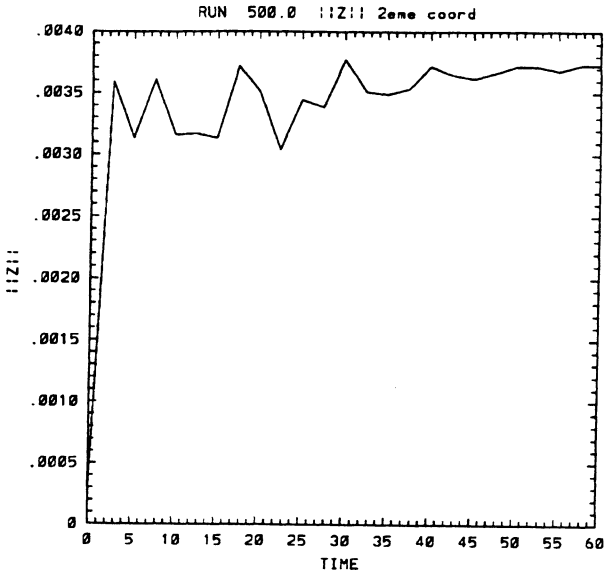
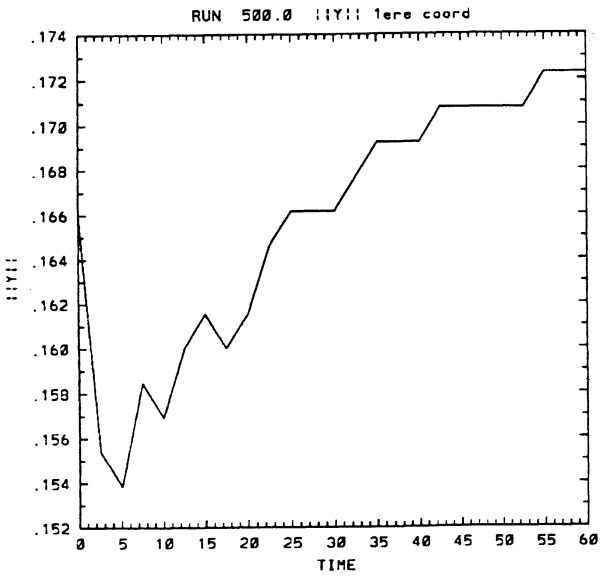
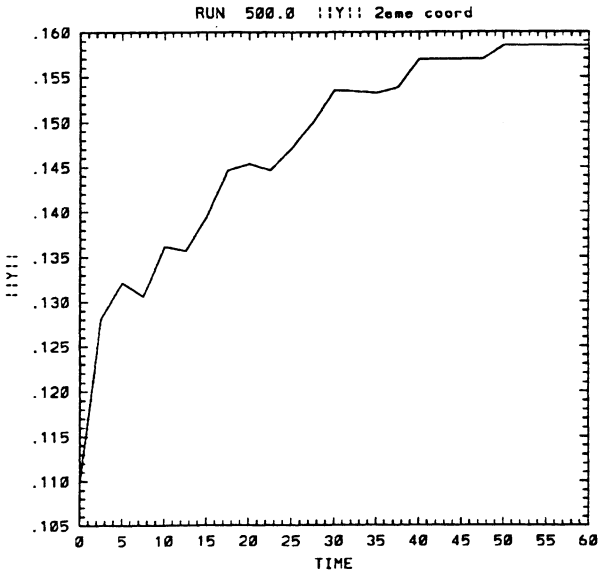


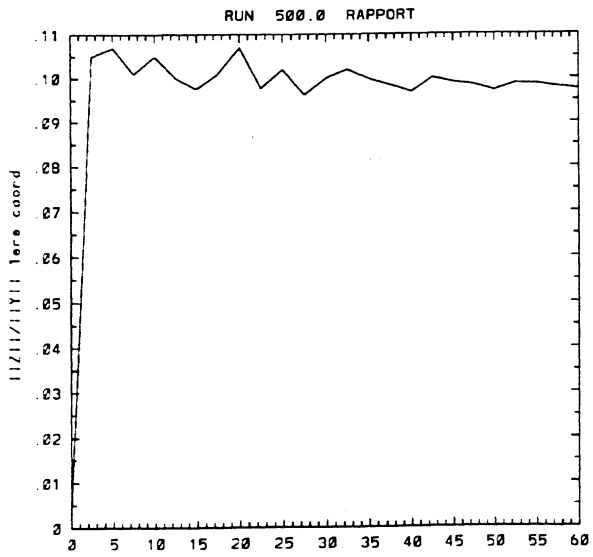
Figure III.13: *$Re = 5000$; $h = \frac{1}{64}$. Norme l_2 des composantes de uy_h et uz_h et rapport de ces normes au cours du temps.*











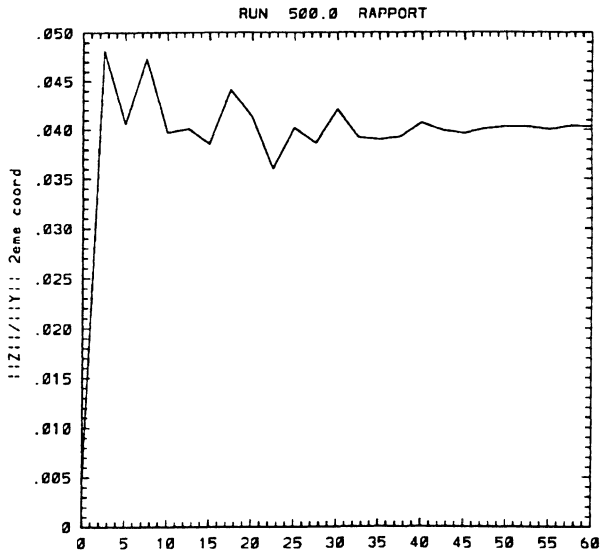


Figure III.14: $Re = 5000$; $h = \frac{1}{128}$; *cavité B). Norme* l_2 *des composantes de* uy_h *et* uz_h *et de leur rapport.*

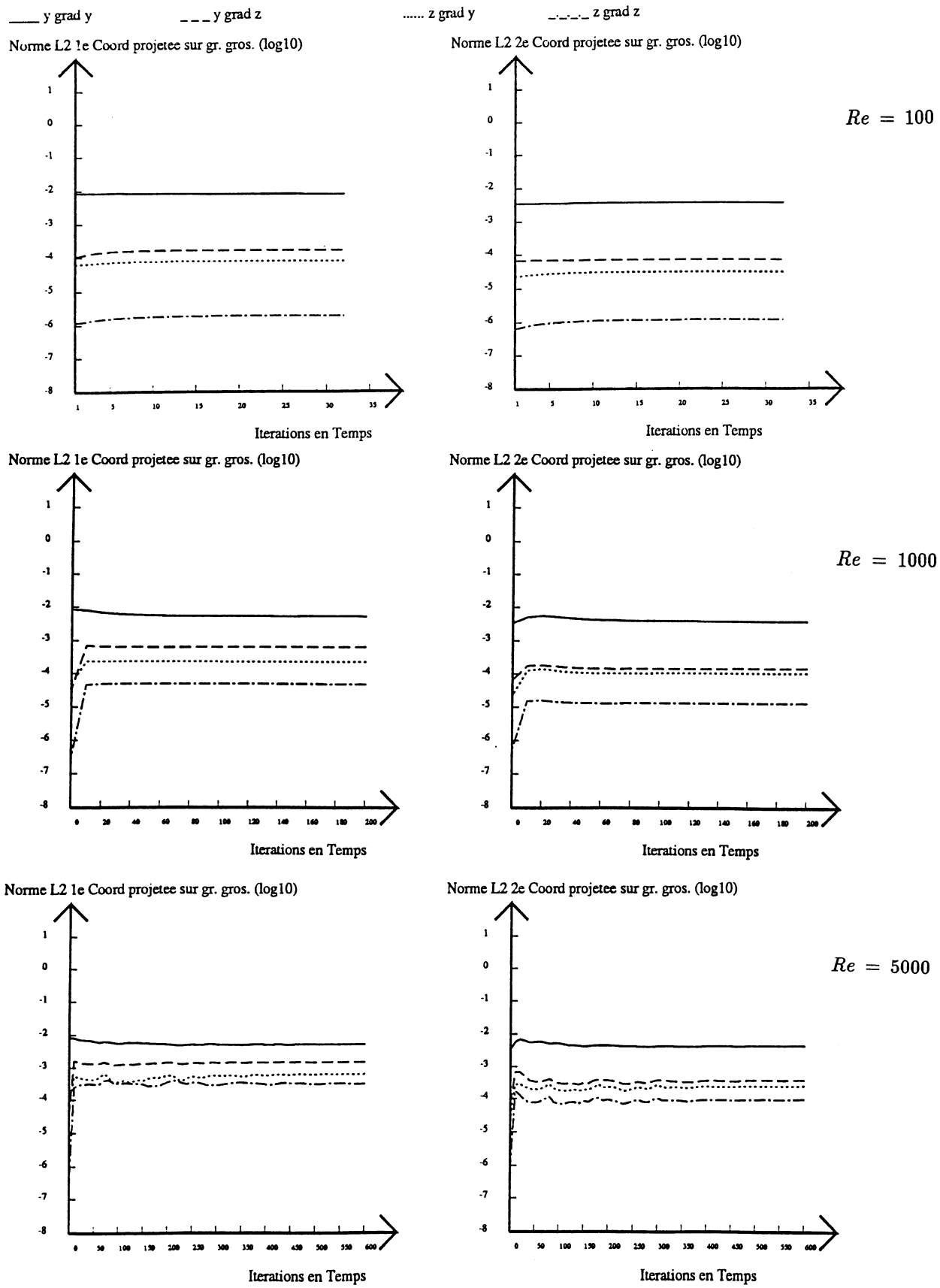


Figure III.15: *Cavité B) ; $h = \frac{1}{64}$. Norme L_2 des differentes composantes du terme non linéaire dans V_{2h}.*

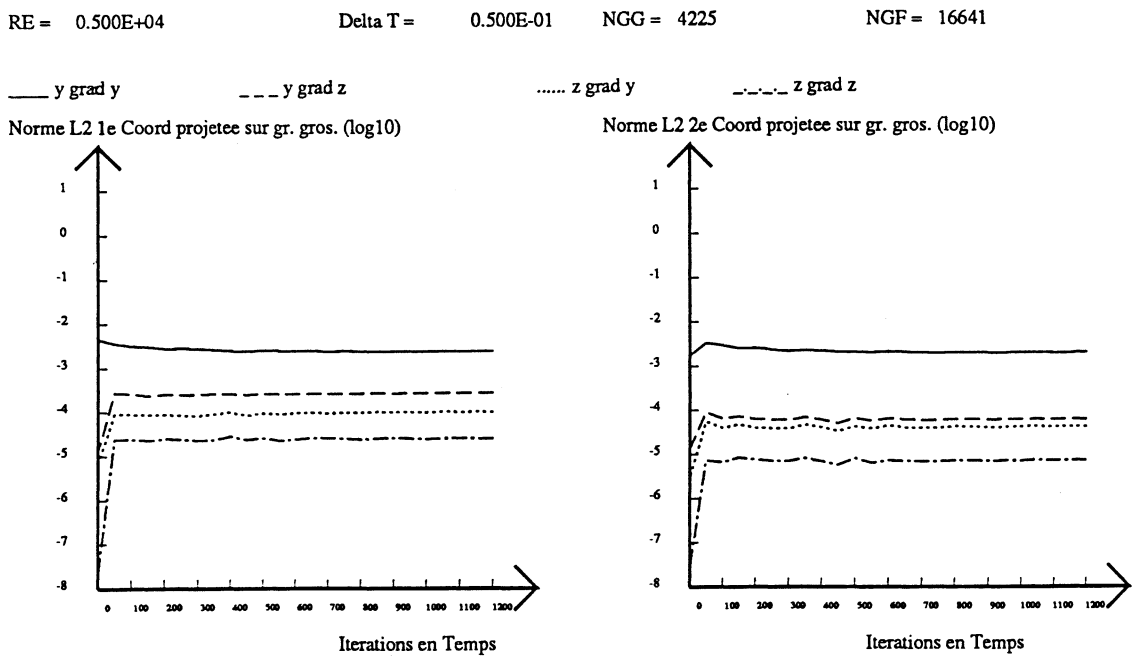


Figure III.16: *Re* = 5000 ; $h = \frac{1}{128}$. *Norme* L_2 *des différentes composantes du terme non linéaire dans* $\mathcal{V}_{2h}$.

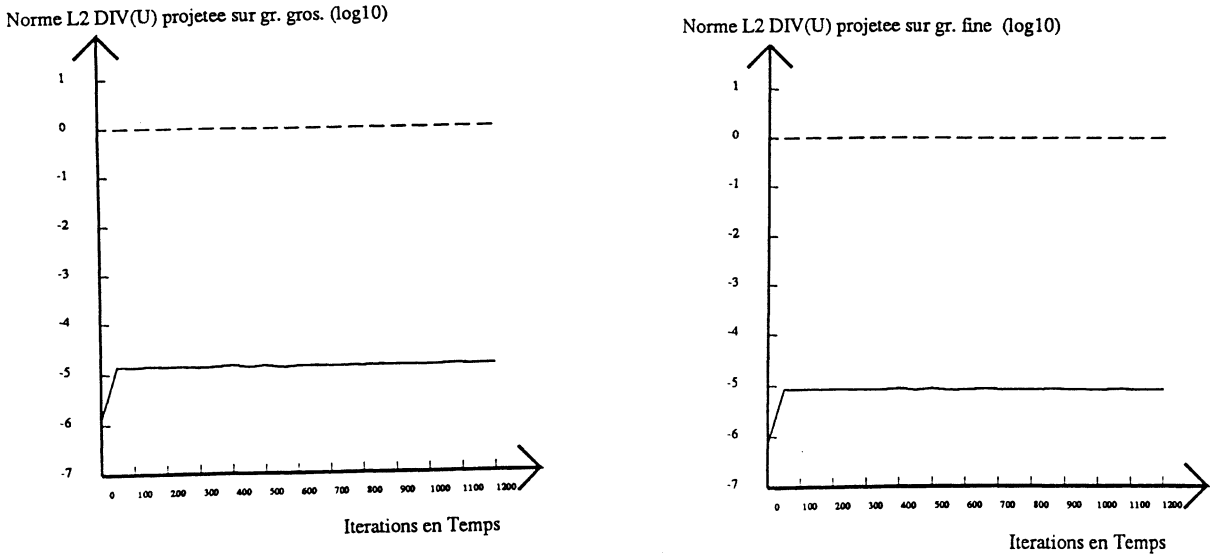
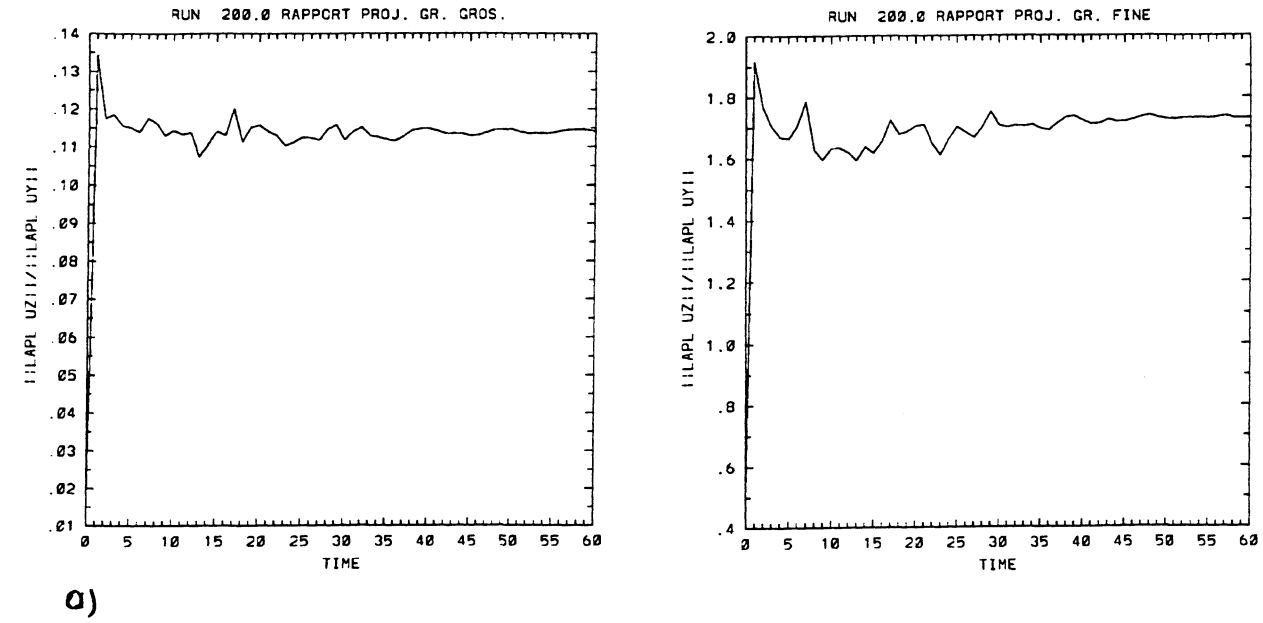
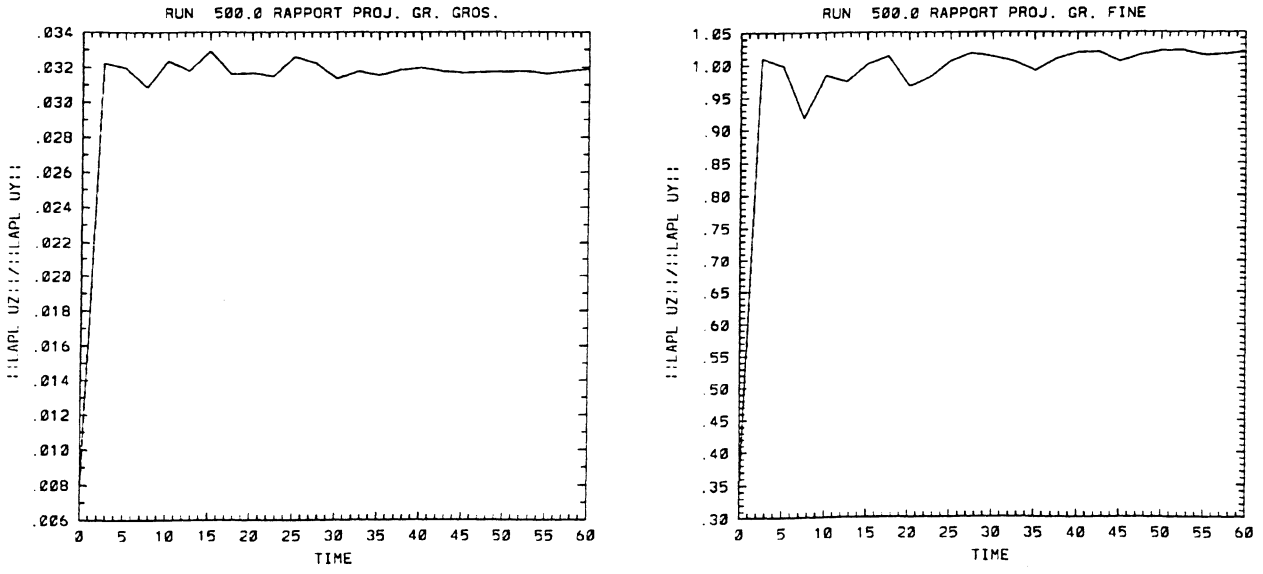


Figure III.17: *Re* = 5000 ; $h = \frac{1}{128}$.
(——) : Norme L_2 *de* $divuy_h$. *(- - -) : Rapport des normes* L_2 *de* $divuz_h$ *et de* $divuy_h$. *A gauche : projection sur* $\mathcal{V}_{2h}$. *A droite : projection sur* $\mathcal{W}_{2h}$.



a)



b)

Figure III.18: $Re = 5000$. *Rapport des normes* L_2 *des termes de diffusion i.e.* $\frac{(\sum_i(\nabla uz_h, \nabla\phi_i)^2)^{1/2}}{(\sum_i(\nabla uy_h, \nabla\phi_i)^2)^{1/2}}$. *a)* : $h = \frac{1}{64}$. *b)* : $h = \frac{1}{128}$.

III.3 Etude préliminaire : Base hiérarchique et problème de Poisson.

Ce paragraphe présente une étude préliminaire concernant la discrétisation multigrille utilisant la base hiérarchique appliquée au problème de Poisson. Dans un premier temps, la formulation variationnelle est écrite dans la base nodale et dans la base hiérarchique. Différents solveurs tenant compte de la structure des systèmes sont ensuite exposés et des études numériques discutées.

III.3.1 Le problème.

On suppose que le domaine Ω est $[-1,1]\times[-1,1]$. On se propose de résoudre le problème de Poisson dont on sait l'existence et l'unicité de la solution (cf. Raviart et Thomas [7]) :
Trouver $u \in H^1(\Omega)$ solution de

$$\begin{aligned} -\frac{\Delta u}{Re} + \alpha u &= f \text{ dans } \Omega \\ u &= g \text{ sur } \partial\Omega \end{aligned} \qquad \text{(III.14)}$$

où $f \in L^2(\Omega)$, $g \in H^{-1/2}(\partial\Omega)$, $Re > 0$ et $\alpha \geq 0$. Pour simplifier les notations, on supposera dans la suite que $g = 0$. Numériquement, on tient compte des conditions au bord dans le second membre du système linéaire associé à (III.14).

III.3.2 Formulation variationnelle et discrétisation.

Une formulation équivalente à (III.14) est la formulation variationnelle suivante $(\mathcal{P})$:
Trouver $u \in H_0^1(\Omega)$ solution de

$$a(u,v) = (f,v) \quad \forall v \in H_0^1(\Omega) \qquad \text{(III.15)}$$

où $a(u,v) = \frac{((u,v))}{Re} + \alpha\,(u,v)$.

On considère la triangulation T_h et l'espace d'approximation par éléments finis V_h définis au chapitre II. Le problème (III.15) est approché par le problème discret suivant $(\mathcal{P}_h)$:
Trouver $u_h \in V_{0h} = \{u_h \in V_h \ / \ u_h = 0 \text{ sur } \partial\Omega\}$ solution de

$$a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h \qquad \text{(III.16)}$$

On connaît l'existence, l'unicité et la convergence de u_h vers u lorsque h tend vers 0. On suppose que les noeuds intérieurs au domaine sont numérotés de 1 à $n_{0,h}$ quelque soit le choix de la base de V_h.

III.3.3 Systèmes linéaires associés.

Dans la base nodale.

Si V_h est muni de la base nodale $\{\hat{\phi}_{h,i}\}_{i=1,n_h}$, la formulation $(\mathcal{P}_h)$ exprimée dans cette base se met sous la forme matricielle suivante :

$$\hat{A}\hat{u} = \hat{b} \tag{III.17}$$

où

- $\hat{A}$ est la matrice symétrique, définie positive telle que :

$$(\hat{A})_{ij} = a(\hat{\phi}_{h,i}, \hat{\phi}_{h,j}) \quad 1 \leq i,j \leq n_{0,h}$$

- $\hat{b}$ est le vecteur colonne : $\hat{b}_i = (f, \hat{\phi}_{h,i}) \quad 1 \leq i \leq n_{0,h}$

- $\hat{u}$ est le vecteur colonne des composantes de u dans la base nodale :

$$u_h = \sum_{i=1}^{n_{0,h}} \hat{u}_i \hat{\phi}_{h,i}$$

où $\hat{u}_i = u(M_i)$, M_i étant le ieme noeud intérieur.

Remarque III.6 *Si la numérotation des noeuds est la numérotation usuelle dans le domaine rectangulaire i.e. croissante de gauche vers la droite, puis de bas en haut alors la matrice $\hat{A}$ a une structure bande.*

Remarque III.7 *Si la numérotation des noeuds est la numérotation induite par la base hiérarchique (cf section III.1) sur 2 niveaux, alors $\hat{A}$ se met sous la forme d'une matrice par blocs :*

$$\hat{A} = \begin{pmatrix} \hat{A}_{cc} & \hat{A}_{cf} \\ \hat{A}_{fc} & \hat{A}_{ff} \end{pmatrix}$$

où

- $\hat{A}_{fc} = \hat{A}_{cf}^t$,
- $\mathcal{A}_{ff}$ *est une matrice carrée symétrique, définie, positive d'ordre* $(n_{0,h} - n_{0,2h})^2$,
- $\hat{A}_{cc}$ *est une matrice diagonale d'ordre* $(n_{0,2h})^2$, *en effet les* $n_{0,2h}$ *premières fonctions de la base nodale (avec la numérotation issue de la base hiérarchique) ont des supports 2 à 2 disjoints.*

Il a été constaté qu'une méthode itérative par blocs du type Gauss-Seidel appliquée à une telle matrice converge très lentement. En effet $\hat{A}_{cc}^{-1}\hat{b}_c$ est une très mauvaise approximation de $\hat{u}_c$.

Dans la base hiérarchique.

Si V_h est muni de la base hiérarchique (on se restreint à 2 niveaux), la formulation $(\mathcal{P}_h)$ exprimée dans cette base conduit au système matriciel suivant :

$$A \begin{pmatrix} u_c \\ u_f \end{pmatrix} = \begin{pmatrix} A_{cc} & A_{cf} \\ A_{fc} & A_{ff} \end{pmatrix} \begin{pmatrix} u_c \\ u_f \end{pmatrix} = \begin{pmatrix} b_c \\ c_f \end{pmatrix} \qquad \text{(III.18)}$$

où

$$\begin{array}{rcll} (A_{cc})_{ij} & = & a(\hat{\phi}_{2h,i}, \hat{\phi}_{2h,j}) & 1 \leq i,j \leq n_{0,2h} \\ (A_{ff})_{ij} & = & a(\hat{\phi}_{h,i}, \hat{\phi}_{h,j}) & n_{0,2h}+1 \leq i,j \leq n_{0,h} \\ (A_{cf})_{ij} & = & a(\hat{\phi}_{2h,i}, \hat{\phi}_{h,j}) & 1 \leq i \leq n_{0,2h} \\ & & & n_{0,2h}+1 \leq j \leq n_{0,h} \\ A_{fc} & = & A_{cf}^t & \\ (b_c)_i & = & (f, \hat{\phi}_{2h,i}) & 1 \leq i \leq n_{0,2h} \\ (b_f)_i & = & (f, \hat{\phi}_{h,i}) & n_{0,2h}+1 \leq i \leq n_{0,h} \end{array}$$

avec

$$u_h = \sum_{i=1}^{n_{0,2h}} (u_c)_i \, \hat{\phi}_{2h,i} + \sum_{i=n_{0,2h}+1}^{n_{0,h}} (u_f)_i \, \hat{\phi}_{h,i}$$

u_c correspond aux grosses structures et u_f aux faibles structures.

Remarque III.8 *Si le système (III.18) est mis sous la forme*

$$A_{cc}u_c + A_{cf}u_f = b_c \qquad \text{(III.19)}$$

$$A_{fc}u_c + A_{ff}u_f = b_f \qquad \text{(III.20)}$$

alors (III.20) apparaît comme une correction du problème (III.15) discrétisé sur la grille grossière avec la base nodale de V_{2h} :

$$A_{cc}u_c = b_c.$$

Ainsi le raffinement du maillage permet de corriger la valeur approchée u_c de la solution évaluée sur la grille grossière T_{2h}.

Lien entre les systèmes $\hat{A}$ et A.

Nous allons décrire dans un premier temps la matrice de passage de la base nodale à la base hiérarchique : S est une matrice carrée triangulaire inférieure définie par :

$$S = \begin{pmatrix} I_c & 0 \\ R & I_f \end{pmatrix}$$

I_c et I_f sont les matrices identités de $\mathbb{R}^{n_{0,2h}^2}$ et de $\mathbb{R}^{(n_{0,h}-n_{0,2h})^2}$. R est la matrice de "raffinement" ou matrice de "passage du maillage T_{2h} au maillage T_h". R est telle que :

- seuls 2 termes par lignes sont au plus non nuls,
- si le sommet B_i est créé lors du raffinement comme étant le milieu de $[A_{j_1}, A_{j_2}]$ alors $S_{ij_1} = S_{ij_2} = \frac{1}{2}$.

Le systéme linéaire associé à la base nodale est lié au système linéaire associé à la base hiérarchique par les 3 relations suivantes :

$$\begin{array}{lcccl} \text{pour les matrices} & : & A & = & S^t \hat{A} S \\ \text{pour les seconds membres} & : & b & = & S^t \hat{b} \\ \text{pour les vecteurs solutions} & : & \hat{u} & = & Su \end{array}$$

En explicitant blocs par blocs ces 3 égalités, on obtient :

$$\left\{ \begin{array}{lcl} A_{cc} & = & \hat{A}_{cc} + R^t \hat{A}_{fc} + \hat{A}_{cf} R + R^t \hat{A}_{ff} R \\ A_{cf} & = & \hat{A}_{cf} + R^t \hat{A}_{ff} \\ A_{fc} & = & \hat{A}_{fc} + \hat{A}_{ff} R \\ A_{ff} & = & \hat{A}_{ff} \end{array} \right. \tag{III.21}$$

$$\left\{ \begin{array}{lcl} b_c & = & \hat{b}_c + R^t \hat{b}_f \\ b_f & = & \hat{b}_f \\ \hat{u}_c & = & u_c \\ \hat{u}_f & = & R u_c + u_f \end{array} \right. \tag{III.22}$$

III.3.4 Description des solveurs.

Le soveur de référence : ICCG en base nodale.

Le solveur de référence consiste en une méthode itérative (ponctuelle) appliquée au système écrit dans la base nodale de V_h : il s'agit du gradient conjugué préconditionné par une factorisation incomplète de Choleski (noté dans toute la suite ICCG).

Rappelons que pour le système $\hat{A}\hat{u} = \hat{b}$, le gradient conjugué préconditionné avec M pour matrice de préconditionnement consiste à :

Figure III.19: *Profil de $\hat{A}$ et de la factorisé incomplète de Choleski L.*

Etape 1: u^0 étant donné quelconque

- calculer $r^0 = b - \hat{A}u^0$
- résoudre $Mz^0 = r^0$
- poser $d^0 = z^0$

Etape 2: d^k, u^k, r^k, z^k étant connues

- poser $u^{k+1} = u^k + \lambda_k d^k$ avec $\lambda_k = \dfrac{(z^k, Mz^k)}{(d^k, \hat{A}d^k)}$
- poser $r^{k+1} = r^k - \lambda_k \hat{A}d^k$
- résoudre $Mz^{k+1} = r^{k+1}$
- poser $d^{k+1} = z^{k+1} + s_{k+1}d^k$ avec $s_{k+1} = \dfrac{(z^{k+1}, Mz^{k+1})}{(z^k, Mz^k)}$

Etape 3: arrêter si $||r^{k+1}|| \leq \epsilon$

M doit être proche de $\hat{A}$ et facile à inverser. M est pris égal à LL^t où L est la factorisée incomplète de Choleski i.e. une matrice triangulaire inférieure dont chaque ligne a le même nombre de termes non nuls que la partie correspondante de $\hat{A}$ plus 4 autres termes issus de la factorisation de choleski. On trouve dans Golub et Meurant [3] une étude mathématique de cette factorisation. Dans le cas de matrices diagonales, les profils sont représentés en figure III.19.

Le solveur du type multigrille en base hiérarchique.

Le solveur utilisant la structure par blocs de la matrice A est une méthode itérative de Gauss-Seidel, symétrique par blocs (SSOR). u_c et u_f étant initialisées à 0, la procédure est la suivante (une schématisation de ce solveur est présentée en figure III.20) :

- **Initialisation :**
  - résoudre $A_{cc}u_c = b_c$
  - calculer $r_f = b_f - A_{fc}u_c$
- **Itération :**
  - résoudre $A_{ff}u_f = r_f$
  - calculer $r_c = b_c - A_{cf}u_f$
  - résoudre $A_{cc}u_c = r_c$
  - calculer $r_f = b_f - A_{fc}u_c$

La procédure se termine lorsque la précision demandée est atteinte (le critère porte sur l'erreur relative). La solution dans la base nodale sur la grille fine T_h est donnée par :

$$\begin{aligned} \hat{u}_c &= u_c \\ \hat{u}_f &= Ru_c + u_f \end{aligned}$$

La procédure commence par évaluer sur la grille grossière T_{2h} une approximation de la solution. Puis une "valeur de correction" (d'un ordre inférieur) est calculée sur le maillage fin constitué des noeud créés lors du raffinement. L'étape suivante consiste à réévaluer la valeur approchée de la solution sur la grille grossière en tenant compte de ces corrections.

Le système en A_{cc} peut être résolu :

- soit de façon récursive (en utilisant les triangulations T_{4h},...,T_{2^kh}),
- soit par une méthode directe (si elle correspond au maillage le plus grossier et si son ordre est suffisamment petit)
- soit par un ICCG.

u_f étant une correction d'un ordre inférieur à u_c, il est raisonnable de penser que seules quelques itérations d'une méthode itérative sont suffisantes pour relaxer $A_{ff}u_f = r_f$. Les tests ont été effectués avec 1 itération de ICCG ou 10 itérations de relaxation i.e. de SSOR.

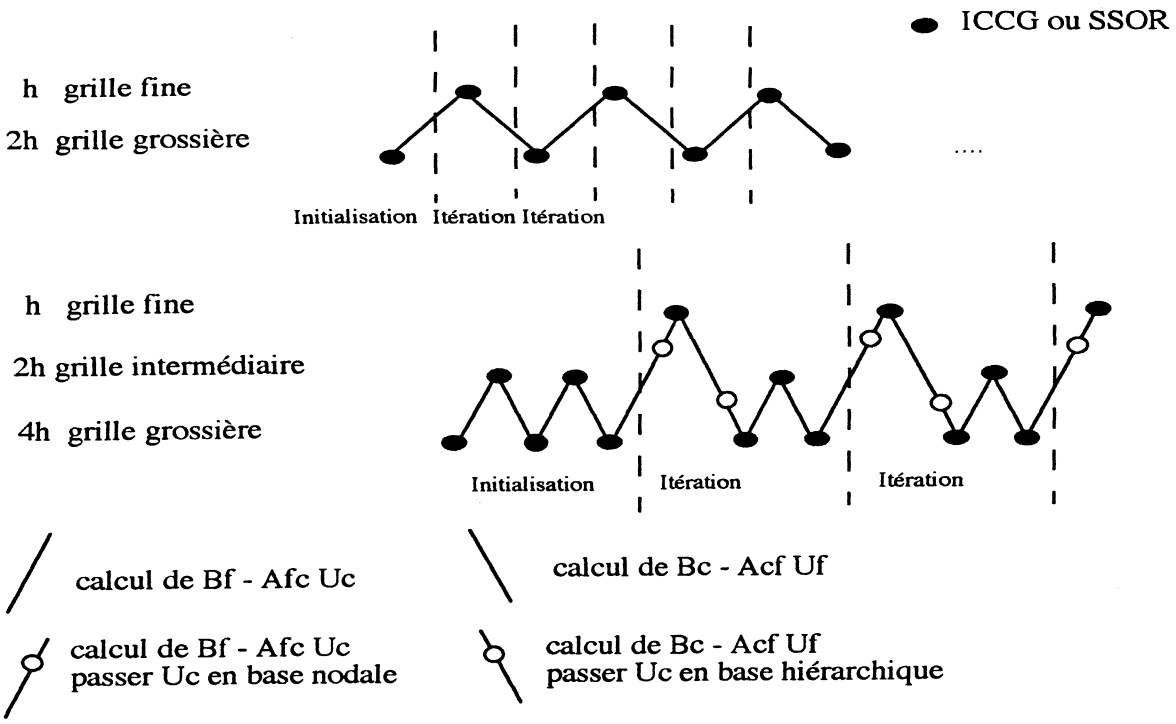


Figure III.20: *Schématisation des solveurs.*

Les points positifs de cette méthode sont résumés ci-dessous :

- l'initialisation fournit une bonne approximation de u_c (on peut donc espérer que la convergence sera rapide),
- u_f est petit (d'un ordre h^2 plus petit sous réserve que la dérivée seconde de u ne soit pas trop grande), ce qui implique un nombre réduit d'itérations pour l'estimer car son influence sur u_c est modérée,
- les 2 systèmes non linéaires sont de taille plus petite (leurs inversions coûtent donc moins cher) en particulier A_{cc}.

Cependant cette méthode présente un certain nombre de points négatifs :

- la solution est obtenue dans la base hiérarchique ; il faut donc procéder à la transformation base hiérarchique-base nodale si l'on procède récursivement i.e. si l'on considère plus de 2 niveaux,
- les différentes matrices sont assemblées dans la base hiérarchique,
- des calculs intermédiaires sont nécessaires pour estimer r_c et r_f,
- u_f est de l'ordre de h^2 fois la dérivée seconde : un maillage très fin est donc nécessaire pour appliquer cette méthode si u présente de forts gradients

Remarque III.9 *Une variante équivalente consiste à non pas relaxer u_c et u_f à chaque itération mais à chercher une correction y_c et y_f de ces quantités. La procédure est alors la suivante :*

- **Initialisation :**
  - *résoudre* $A_{cc}u_c = b_c$
  - *calculer*
$$\begin{pmatrix} r_c \\ r_f \end{pmatrix} = \begin{pmatrix} b_c - A_{cc}u_c \\ b_f - A_{fc}u_c \end{pmatrix}$$
- **Itération :**
  - *résoudre* $A_{ff}y_f = r_f$
  - *calculer*
$$\begin{pmatrix} r_c \\ r_f \end{pmatrix} = \begin{pmatrix} r_c - A_{cf}y_f \\ r_f - A_{ff}y_f \end{pmatrix}$$
  - *poser* $u_f = u_f + y_f$
  - *résoudre* $A_{cc}y_c = r_c$
  - *calculer*
$$\begin{pmatrix} r_c \\ r_f \end{pmatrix} = \begin{pmatrix} r_c - A_{cc}y_c \\ r_f - A_{fc}y_c \end{pmatrix}$$
  - *poser* $u_c = u_c + y_c$

Remarque III.10 *Il est possible d'éviter l'assemblage des matrices dans la base hiérarchique. Il suffit pour cela d'assembler les matrices $\hat{A}_{ff}$, $\hat{A}_{fc} = \hat{A}^t_{fc}$ avec la base nodale sur la grille fine et d'assembler A_{cc} avec la base nodale sur la grille grossière. En tenant compte des relations (III.21) et (III.22) liant A et $\hat{A}$, la procédure itérative s'écrit :*

- **Initialisation :**
  - *calculer* $b_c = \hat{b}_c + R^t\hat{b}_f$
  - *résoudre* $A_{cc}\hat{u}_c = b_c$ *soit directement, soit itérativement, soit récursivement*
  - *calculer* $y_c = R\hat{u}_c$
  - *calculer* $\hat{r}_f = \hat{b}_f - \hat{A}_{fc}\hat{u}_c - \hat{A}_{ff}y_c$

- **Itération :**
  - *résoudre* $\hat{A}_{ff} y_f = \hat{r}_f$
  - *poser* $\hat{u}_f = \hat{u}_f + y_f$
  - *calculer* $r_c = b_c - \hat{A}_{cf} y_f - R^t A_{ff} y_f$
  - *résoudre* $A_{cc} \hat{u}_c = r_c$
  - *poser* $y_c = R\hat{u}_c$
  - *calculer* $\hat{r}_f = \hat{b}_f - \hat{A}_{fc}\hat{u}_c - \hat{A}_{ff} y_c$

Il est clair que le nombre d'opérations par itération est nettement plus grand que lorsque les matrices sont assemblées dans la base hiérarchique. L'assemblage des matrices est cependant plus simple.

Cette méthode (très voisine de celle proposée par Bank *et al* [1]) diffère en plusieurs points des méthodes multigrilles classiques exposées par W.L. Briggs [2]. La méthode V-cycle par exemple consiste à relaxer la solution u sur la grille la plus fine avec une méthode de Gauss-Seidel (la grille fine étant constituée de l'ensemble des noeuds de la triangulation T_h :Σ_h). Dans le cas présent, les points de la grille fine sont constitués des noeuds crées lors du raffinement i.e. $\Sigma_h \setminus \Sigma_{2h}$. Puis l'équation résiduelle i.e. $Ae = r$ avec $r = f - Au$ est résolue sur la grille grossière afin de relaxer les basses fréquences qui n'ont pu l'être sur la grille fine. Un opérateur de restriction pour passer de la grille fine à la grille grossière I_c^f et un opérateur d'interpolation pour passer de la grille grossière I_f^c sont donc nécessaires. La solution sur la grille fine est corrigée en posant $u = u + I_f^c e$ puis relaxée. La méthode FMV (full multigrid V-cycle) consiste en un V-cycle pour lequel la solution sur la grille fine est initialement évaluée sur la grille grossière puis interpolée.

Remarque III.11 *L'efficacité du solveur décrit ci-dessus nous permet également de l'envisager comme préconditionnement d'un gradient conjugué appliqué à A car il fournit dès les premières itérations une excellente approximation de l'inverse de A. Pratiquement le préconditionnement est constitué de la première itération de cette procédure itérative par bloc. Ce préconditionnement est étudié numériquement dans la suite.*

III.3.5 Résultats numériques.

Les test numériques ont été réalisés en imposant la solution exacte :
$u^{ex} = \sin[(x-1)(x+1)(y-1)(y+1)]$. On note u_h^{ex} la discrétisation de u^{ex} dans $V_{0,h}$ et on associe à u_h^{ex} les vecteurs colonnes

$$U_h^{ex} = \begin{pmatrix} U_{hc}^{ex} \\ U_{hf}^{ex} \end{pmatrix} \quad \text{et } \hat{U}_h^{ex}$$

des coordonnées de u_h^{ex} dans la base hiérarchique et dans la base nodale.

On se propose d'appliquer les différents schémas développés dans la section précédente au problème :

$$\text{Trouver } \hat{U}_h \ / \ \hat{A}\hat{U}_h = \hat{A}\hat{U}_h^{ex}.$$

On note :

ALGO 1 : V_h étant muni de la base hiérarchique sur 2 grilles, l'algorithme consiste en un gradient conjugué préconditionné par la première itération du processus itératif par blocs décrit ci-dessus.

ALGO 2 : V_h étant muni de la base hiérarchique sur 2 grilles, l'algorithme est le processus itératif par blocs où chacun des systèmes est résolu par la méthode ICCG.

ALGO 3 : V_h étant muni de la base hiérarchique sur 2 grilles, l'algorithme est le processus itératif par blocs où la matrice A_{cc} est inversée par la méthode ICCG et la matrice A_{ff} par une méthode de relaxation du type SSOR.

ALGO 4 : V_h étant muni de la base nodale, la matrice $\hat{A}$ est inversée par un gradient conjugué préconditionné par le factorisé incomplet de Choleski.

Les itérés étant notés U_h^k,

$$\frac{||U_h^k - U_h^{ex}||_{l_2}}{||U_h^{ex}||_{l_2}} \quad \text{et} \quad \frac{||U_h^k - U_h^{k-1}||_{l_2}}{||U_h^k||_{l_2}}$$

représente respectivement l'erreur absolue et l'erreur relative. La convergence est atteinte lorsque l'erreur relative est inférieur à une valeur ϵ donnée.

On suppose dans un premier temps que $\alpha = 0$.

Influence de la subdivision

Les premières expériences numériques testent les 3 premiers schémas avec différentes subdivisions des triangles. Pour cela le maillage T_h qui représente la grille fine est obtenu à partir du maillage grossier T_{ph} en subdivisant les triangles en p^2 sous-triangles congruents avec successivement $p = 2, 4, 8$.

Si la frontière n'a qu'une seule composante connexe, l'identité d'Euler pour les triangles $S - C + T = 1$ où S est le nombre de sommets, C le nombre de côtés et T le nombre de triangles a permis d'établir (cf. chapitre II) qu'asymptotiquement :

$$S \simeq \frac{T}{2} \quad \text{et} \quad C \simeq \frac{3T}{2}.$$

Si l'on note n_1 (respectivement n_2) le nombre de divisions du maillage grossier dans la direction 1 (respectivement dans la direction 2), on a alors :

- si $p = 2$ i.e. si le nombre de sous-triangles est 4

$$\text{Ordre de } A = (4n_1n_2)^2 \quad \text{et} \quad \frac{\text{Ordre de } A_{cc}}{\text{Ordre de } A_{ff}} = \frac{(n_1n_2)^2}{(3n_1n_2)^2} = \frac{1}{9}$$

- si $p = 4$ i.e. si le nombre de sous-triangles est 16

$$\text{Ordre de } A = (9n_1n_2)^2 \quad \text{et} \quad \frac{\text{Ordre de } A_{cc}}{\text{Ordre de } A_{ff}} = \frac{(n_1n_2)^2}{(8n_1n_2)^2} = \frac{1}{72}$$

- si $p = 8$ i.e. si le nombre de sous-triangles est 64

$$\text{Ordre de } A = (64n_1n_2)^2 \quad \text{et} \quad \frac{\text{Ordre de } A_{cc}}{\text{Ordre de } A_{ff}} = \frac{(n_1n_2)^2}{(63n_1n_2)^2} = \frac{1}{3969}.$$

Il est donc clair que plus p est grand, plus la matrice A_{cc} est petite et plus A_{ff} tend vers la matrice A.

La figure III.21 représente l'erreur absolue en fonction du nombre de degrés de liberté pour $p = 2$, $p = 4$ et $p = 8$ après 1 itération, 5 itérations et 15 itérations des 3 premiers schémas : Algo 1, Algo 2 et Algo 3.

Après une itération, les schémas qui donnent la meilleure approximation de la solution sont les processus itératifs par blocs du type multigrille ; ils sont d'autant meilleurs qu'il y a de noeuds bien qu'une limite de l'ordre de 10^{-5} (la précision des méthodes d'inversion des sous-systèmes) semble être atteinte au delà de 10000 points pour $p = 2$. Pour $p = 4$ et $p = 8$, les schémas itératifs par blocs (Algo 2 et 3) sont moins performants : en effet l'erreur absolue est inférieure à celle obtenue pour $p = 2$, la différence diminuant avec le nombre de noeuds. En ce qui concerne le gradient conjugué, quelque soient le nombre de sous-triangles (seul le préconditionnement change lorsque p change) et le nombre de noeuds, le premier itéré est très éloigné de la solution exacte.

Il semble donc que d'une part si le rapport du nombre de noeuds de la grille grossière sur le nombre de noeuds de la grille fine est faible (par exemple pour $p = 8$), il est nécessaire que le nombre total de noeuds soit important pour que la première approximation soit très proche de la solution exacte et que d'autre part il est inutile d'augmenter démesurément le nombre total de noeuds pour améliorer la première approximation (car il existe une limite inférieure dépendant de la précision de l'inversion de la matrice A_{cc}).

Question temps CPU le schéma 2 avec $p = 4$ est plus rapide qu'avec $p = 2$ ou $p = 8$. Pour 66000 degrés de liberté, c'est le moins cher et le plus efficace des 3 schémas après une itération.

On constate après 5 itérations, que le gradient conjugué converge très vite (résultat classique) et c'est d'autant plus vrai qu'il y a peu de points. Par contre les schémas itératifs par blocs convergent lentement (résultat classique pour des méthodes de Gauss-Seidel), l'approximation essentielle ayant eu lieu lors de la première itération. Cela confirme que ces schémas sont avant tout d'excellents préconditionneurs de méthodes itératives telles que le gradient conjugué. Le schéma 3 converge moins vite que le schéma 2 et c'est d'autant plus clair que p augmente : l'approximation de u_f par quelques itérations d'une méthode de relaxation est de plus en plus grossière au fur et à mesure que l'ordre de la matrice A_{ff} augmente.

Les courbes de l'erreur absolue à la 15e itération confirment que pour $p = 8$ les schémas itératifs par blocs (Algo 2 et 3) convergent lentement après le premier itéré si le nombre de points de discrétisation est faible. Cependant à la vue de la figure III.21, on peut déduire que si ce nombre est suffisamment important, les schémas du type multigrille à la 15e itération sont meilleurs que les gradients conjugués. C'est en effet ce qu'il se produit avec $p = 2$ pour plus de 10000 points et avec $p = 4$ pour plus de 40000 points. Pour de tels nombres de points et de tels itérés, la différence entre $p = 2$, 4 et 8 se situe au niveau du temps d'exécution : le schéma 3 (avec SSOR) coûte très cher quelque soit p et le temps de calcul pour le schéma 2 et 3 est minimal avec $p = 4$.

En conclusion, si la précision exigée est très élevée, si le nombre d'itérations est important et si le maillage est trop grossier, il faut alors priviligier des algorithmes tels que le gradient conjugué préconditionné par le processus itératif par blocs. Par contre dans le cas contraire, quelques itérations du processus itératif basé sur la décomposition sur la base hiérarchique suffisent avec une subdivision $p = 2$ et ce pour un coût extrêmement compétitif. Ce processus itératif avec une subdivision $p = 4$ ou 8 n'est envisageable vis à vis de $p = 2$ et du gradient conjugué que si le nombre de points de discrétisation est très élevé, ce qui dans notre exemple bien trop régulier n'est absolument pas nécessaire.

Comparaison des schémas avec $p = 2$.

Désormais les tests numériques sont réalisés avec les triangulations T_h et T_{2h}. Les figures III.22, III.23 et III.24 donnent l'évolution du temps de calcul CPU, de l'erreur relative et de l'erreur absolue au cours des itérations de chacun des schémas dans les cas où le nombre de degrés de liberté vaut 4225 (i.e. 65^2), 40401 (i.e. 129^2) , 66049 (i.e. 257^2). On y remarque que les schémas itératifs par blocs atteignent très rapidement la précision exigée (10^{-5}) alors que les gradients convergent plus lentement. Chaque itération des schémas "multigrille-base hiérarchique" exige plus d'opérations qu'une itération du gradient conjugué mais comme seul un petit nombre est nécessaire, ces schémas sont les moins coûteux.

Ces résultats sont également observables sur les courbes en fonction du nombre de noeuds (cf. figure III.25) où l'on peut observer que le nombre d'itérations nécessaires aux schémas multigrilles est constant (de l'ordre de 3) quelque soit la

précision du maillage à l'inverse des gradients conjugués dont le nombre d'itérations nécessaires augmente avec la précision. L'efficacité du préconditionnement par la méthode du type multigrille y est également visible puisque le nombre d'itérations dans ce cas là est inférieur au nombre d'itérations d'un ICCG ponctuel.

Multigrille sur plus de 2 niveaux.

Les figures III.26 et III.27 comparent le gradient conjugué préconditionné par choleski incomplet appliqué au système issu de la base nodale et le schéma du type multigrille appliqué au système issu de la base hiérarchique (l'ensemble des inversions étant réalisé par ICCG) avec 2 puis 3 et 4 niveaux distincts de plus en plus fins.

On y constate à nouveau que la méthode du type multigrille est performante et que pour les schémas avec 3 et 4 niveaux l'efficacité est également fonction du nombre de points du maillage. Par exemple il faut 16000 points pour que la précision soit atteinte en 3 itérations (cf. définition d'une itération s'il y a plus de 2 niveaux en section III.3.4) avec un schémas à 3 niveaux et plus de 26000 pour un schéma à 4 niveaux (la grille la plus grossière étant dans ce cas la grille 21×21). Plus il y a de niveaux, plus il faut de noeuds pour que les schémas itératifs par blocs soient performants.

En effet le passage d'une grille à une autre est une opération coûteuse que l'on compense si les grandeurs u_c sont de bonnes approximations de la solution, i.e. si le maillage est suffisamment fin.

Application au problème de Stokes.

Il a été montré en II.3 que la méthode d'Uzawa utilisée pour résoudre le problème de Stokes nécessite par itération la résolution de plusieurs problèmes de Dirichlet. Les différentes méthodes étudiées dans les paragraphes précédents (avec 2 niveaux) sont donc appliquées à ces problèmes. Les résultats sont résumés en figure III.28.

Comme cela était attendu, ces schémas n'influent pas sur la norme L_2 de la divergence de la vitesse u, par contre le temps CPU est nettement réduit avec les méthodes du type multigrille dès que le pas de discrétisation est faible. On retrouve que le nombre d'itérations nécessaires pour ces schémas est inférieur au nombre d 'itérations des gradients conjugués. On remarque cependant que l'inversion de la matrice A_{ff} par quelques itérations de SSOR est trop grossière par rapport à l'inversion constituée d'une itération de ICCG d'où la convergence plus lente du schéma "Algo 3".

Cette efficacité des méthodes multigrille-base hiérarchique pour résoudre le problème de Stokes est tout à fait similaire à celle obtenue par Verfurth [8] avec des méthodes multigrilles classiques telles que le V-cycle, le W-cycle ou le FMV-cycle.

Cas où $\alpha \neq 0$.

La dernière expérience numérique concerne le problème de Dirichlet suivant :

$$-\frac{\Delta u}{100} + 100\, u = f \text{ dans } \Omega \qquad \text{(III.23)}$$

$$u = g \text{ sur } \partial\Omega \qquad \text{(III.24)}$$

Comme $\frac{1}{100} \ll 100$, la matrice $\hat{A}$ est très proche de

$$(\hat{A})_{ij} \simeq \frac{1}{100}((\hat{\phi}_{h,i}, \hat{\phi}_{h,j}))$$

matrice très bien conditionnée, proche de l'identité et facilement inversible avec un gradient conjugué préconditionnné par Choleski incomplet. Par exemple jusqu'à $n_h = 37249$ i.e. un maillage 193×193, le nombre d'itérations nécessaires pour obtenir une erreur relative inférieure à 10^{-4} est de l'ordre de 3, ce qui est très faible.

Or les méthodes multigrilles atteignent de telles précisions en 2 ou 3 itérations. De plus dans le cas des méthodes issues de la décomposition sur la base hiérarchique, chaque itération coûte plus cher en temps d'exécution (cf. la figure III.29 qui compare ICCG et les méthodes issues de la décomposition sur la base hiérarchique) et ce pour les raisons suivantes :

- le passage d'une grille à l'autre implique un certain nombre d'opérations
- la matrice A_{ff} est mal conditionnée d'où une convergence plus lente.

Dans ce cas, les méthodes proposées et développées dans les paragraphes précédents sont d'un intérêt moindre par rapport aux méthodes plus classiques telles que le gradient conjugué préconditionné.

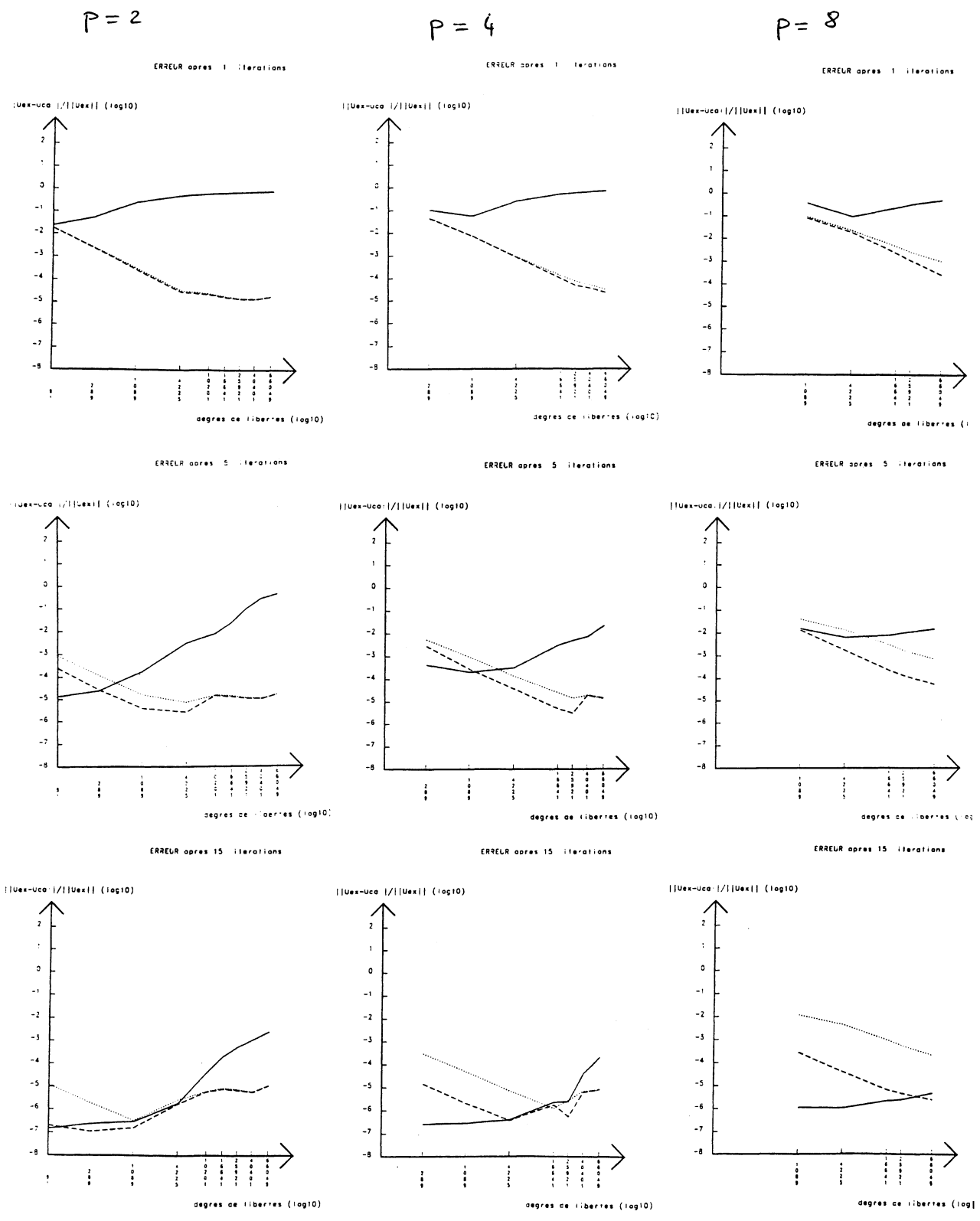


Figure III.21: *Problème de Dirichlet. Erreur absolue pour les algorithmes 1 (——), 2 (- - - -), 3 (.....) après 1, 5 et 15 itérations.*

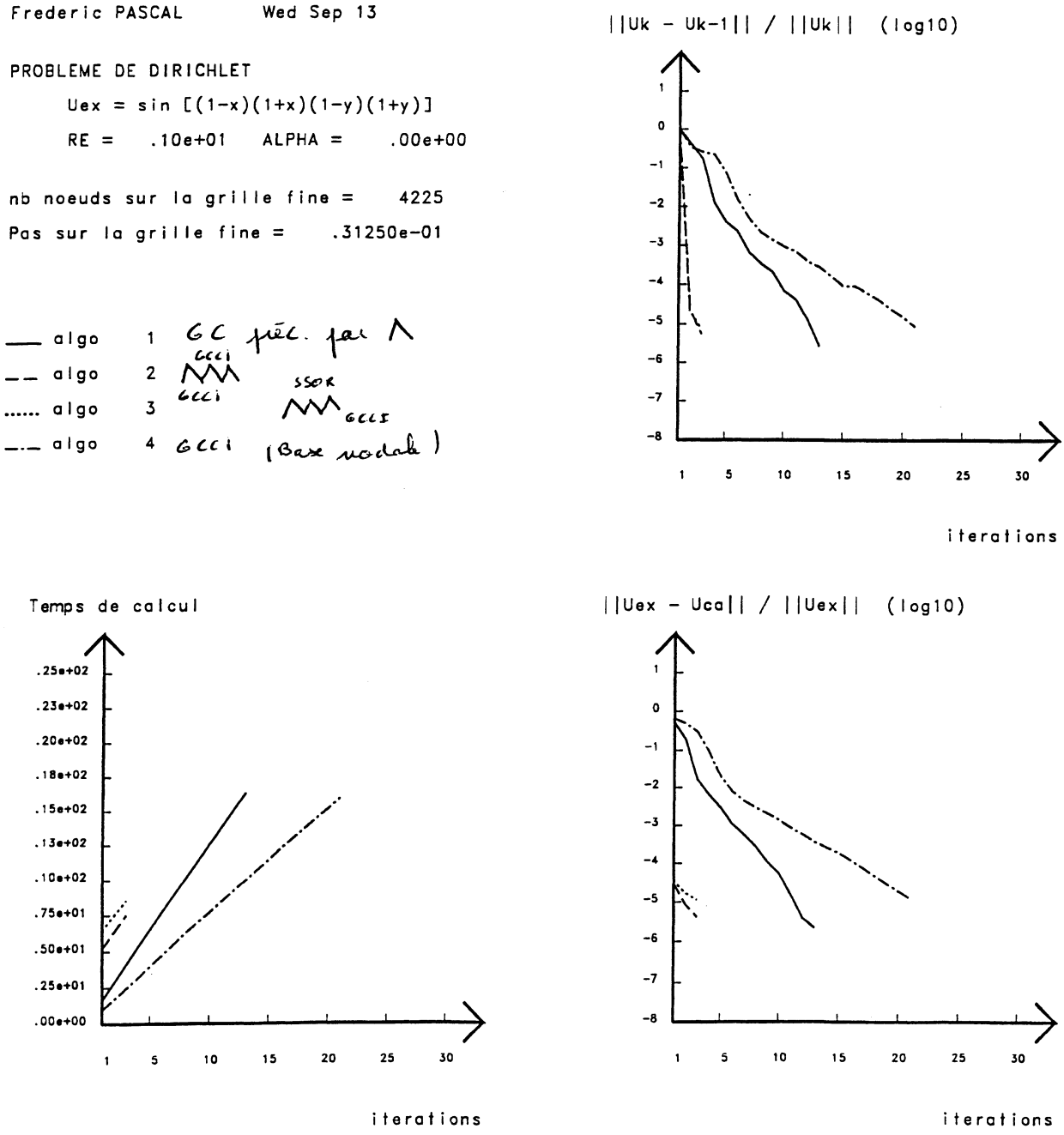


Figure III.22: *Problème de Dirichlet. Comparaison de la méthode ICCG-base nodale et des méthodes multigrille-base hiérarchique.* $h = 0.03125$.

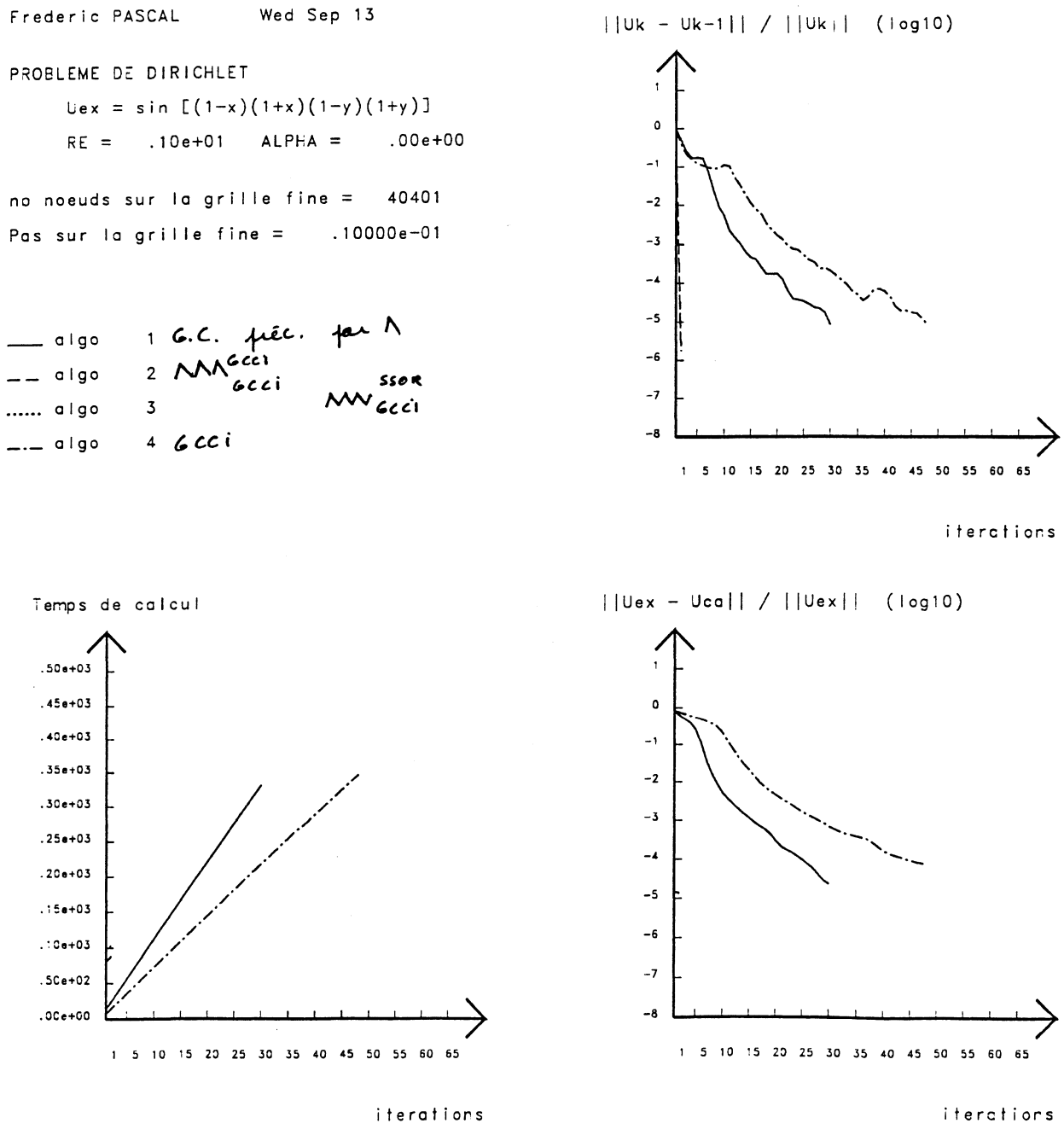


Figure III.23: *Problème de Dirichlet. Comparaison de la méthode ICCG-base nodale et des méthodes multigrille-base hiérarchique.* $h = 0.0156$.

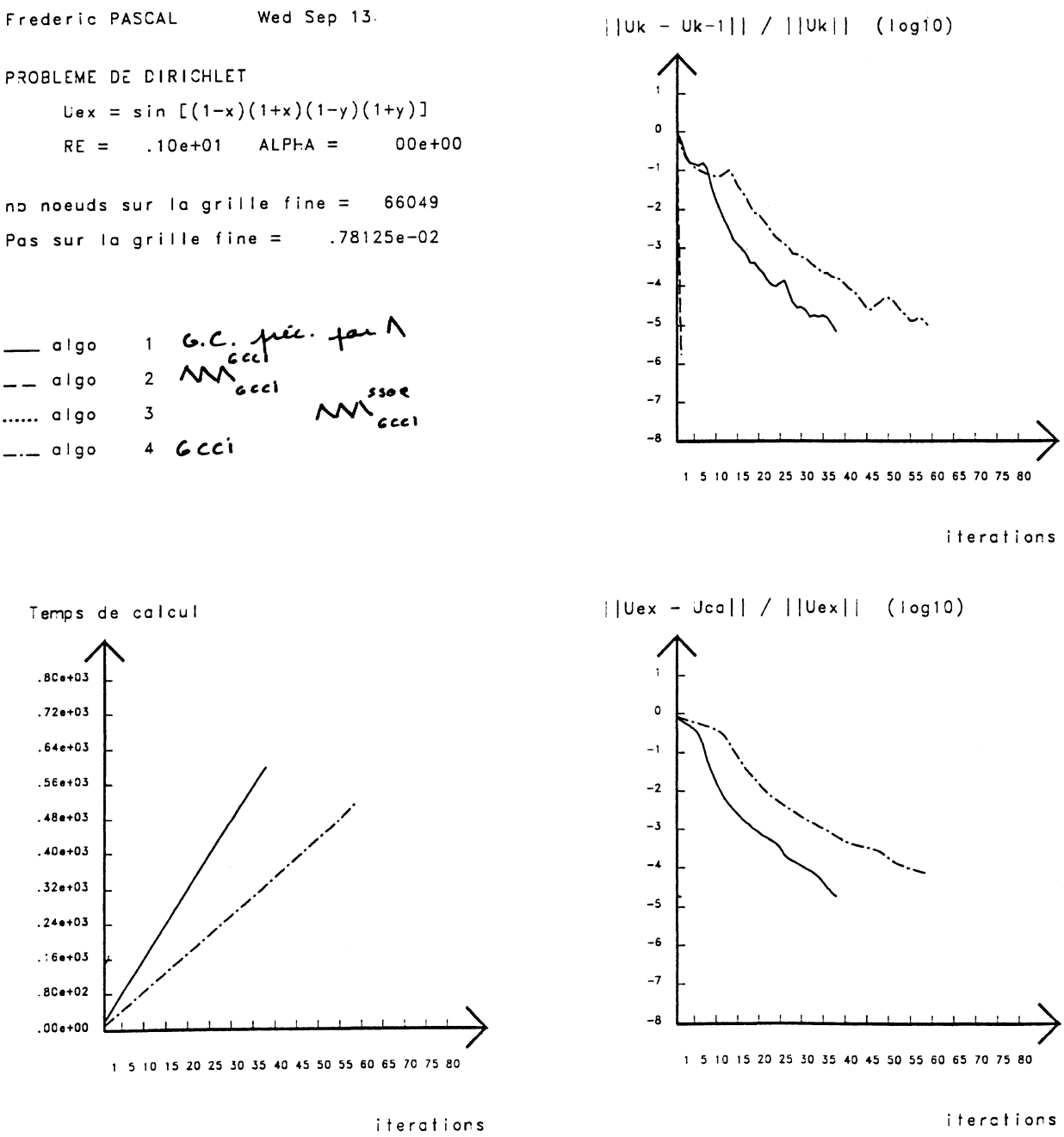


Figure III.24: *Problème de Dirichlet. Comparaison de la méthode ICCG-base nodale et des méthodes multigrille-base hiérarchique.* $h = 0.0078$.

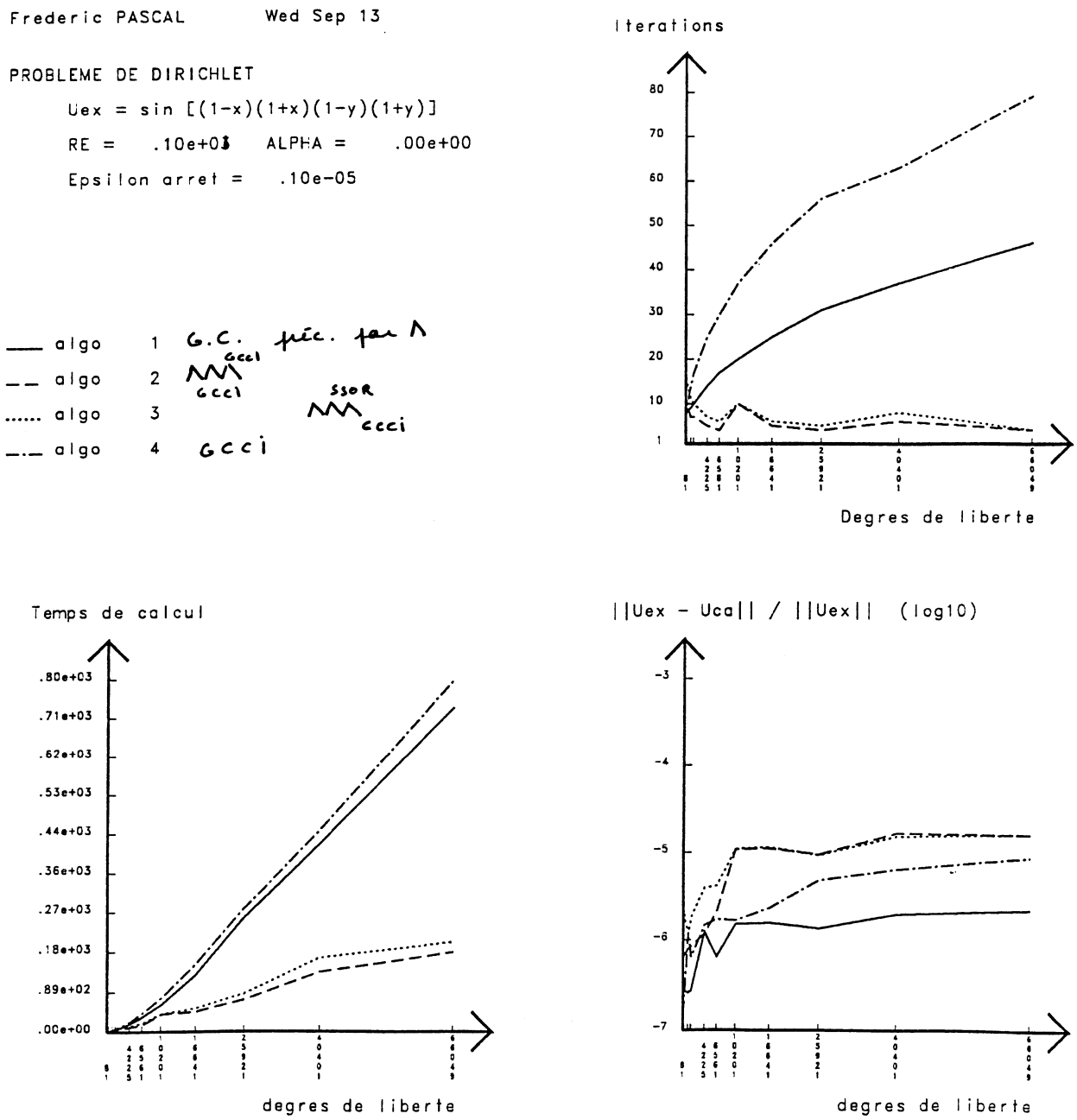


Figure III.25: *Problème de Dirichlet. Comparaison de la méthode ICCG-base nodale et des méthodes multigrille-base hiérarchique. Etude de la convergence en fonction du nombre de degrés de liberté.*

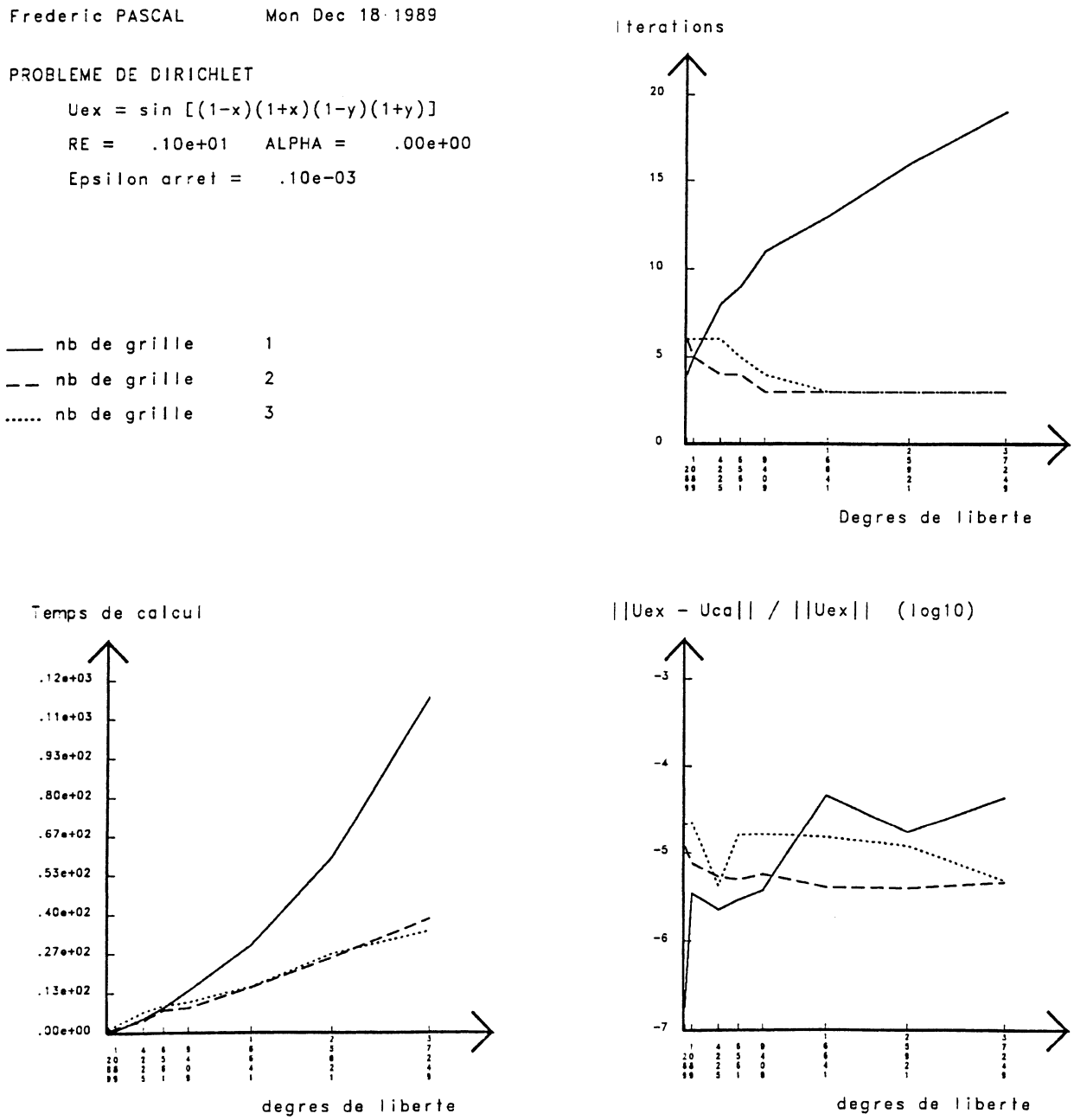


Figure III.26: *Problème de Dirichlet. Comparaison de la méthode ICCG-base nodale et des méthodes multigrille-base hiérarchique sur 2 et 3 niveaux.*

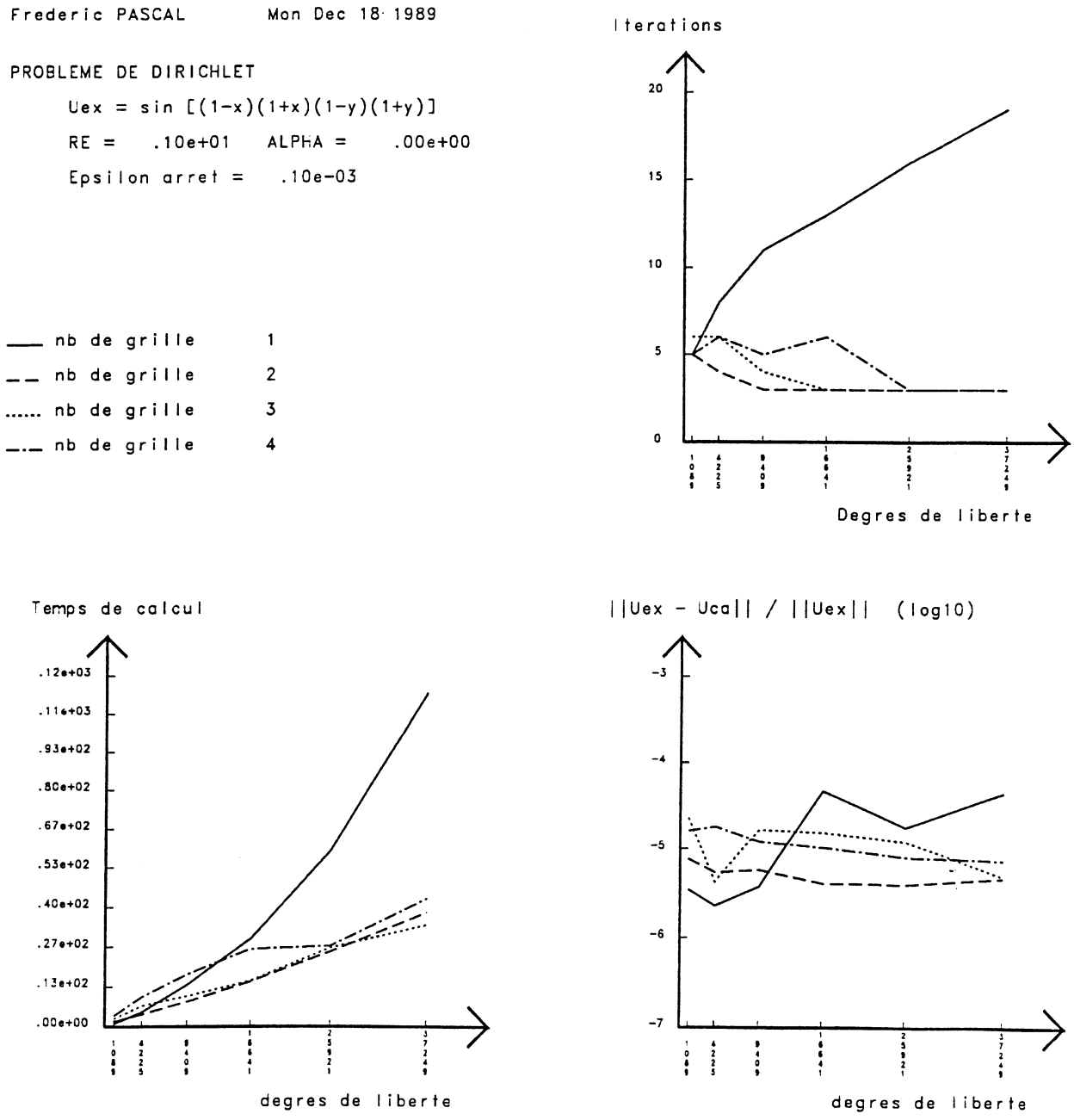


Figure III.27: *Problème de Dirichlet. Comparaison de la méthode ICCG-base nodale et des méthodes multigrille-base hiérarchique sur 2, 3 et 4 niveaux.*

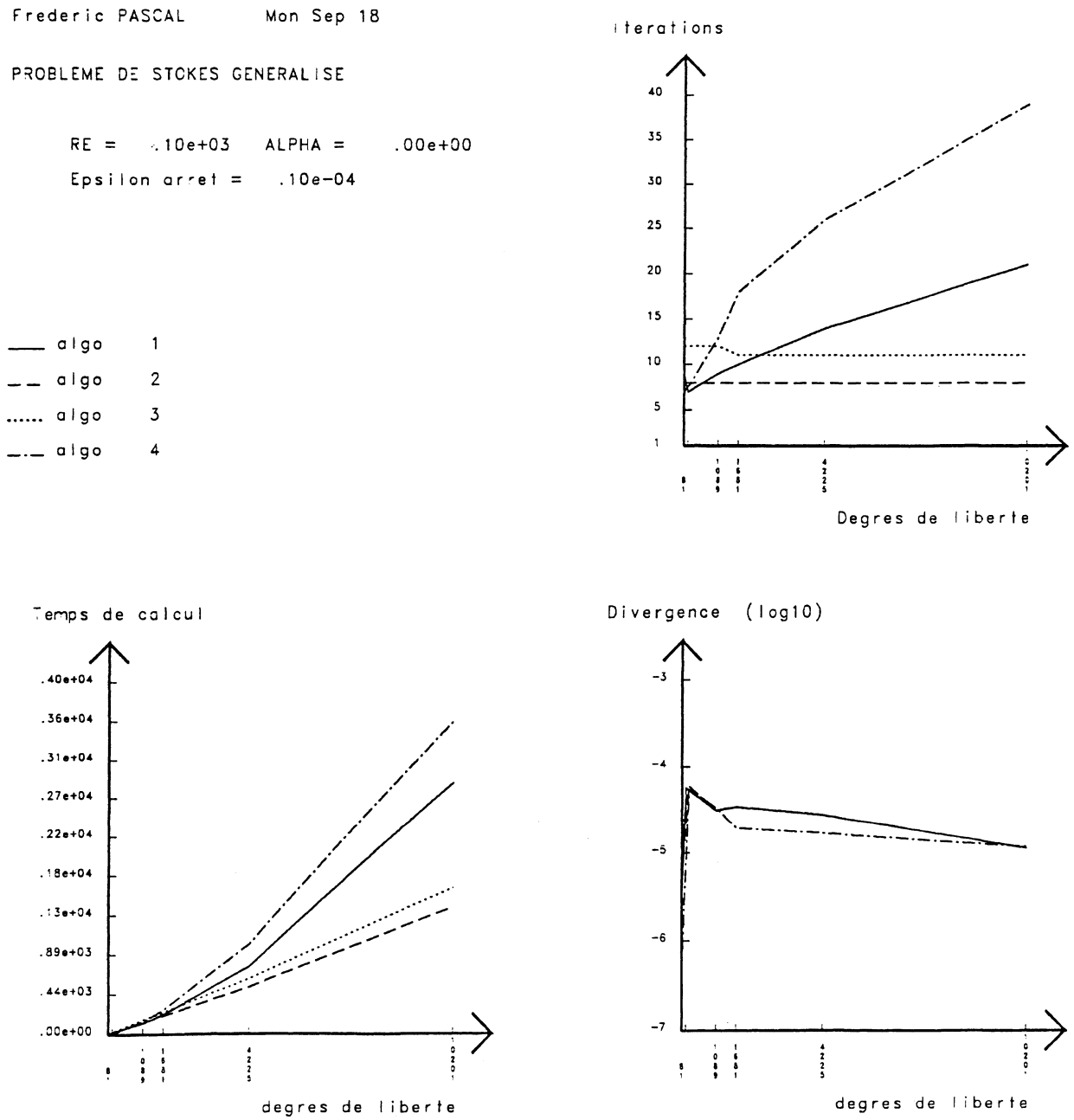


Figure III.28: *Application au problème de Stokes. Etude de la convergence en fonction du nombre de degrés de liberté.*

Frederic PASCAL Fri Dec 15 1989

PROBLEME DE DIRICHLET

Uex = sin [(1-x)(1+x)(1-y)(1+y)]

RE = .10e+03 ALPHA = .10e+03

Epsilon arret = .10e-03

___ nb de grille 1

_ _ nb de grille 2

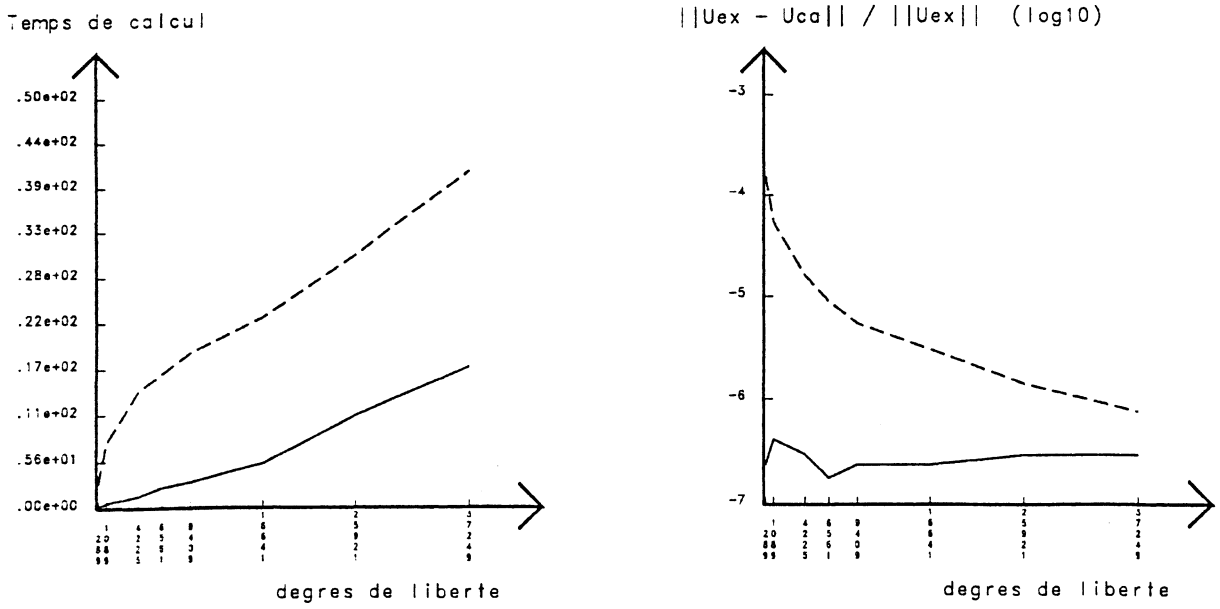


Figure III.29: *Problème de Dirichlet. Cas où* $\alpha \neq 0$.
Etude du problème $-\frac{\Delta u}{100} + 100\ u = f$.

Bibliographie

[III.1] R.E. Bank, T.E. Dupont, H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math. 52, 427-458, 1988.

[III.2] W.L. Briggs, *A multigrid tutorial*, S.I.A.M., Philadelphia, Pennsylvania, 1987.

[III.3] G. Golub, G. Meurant, *Résolution numérique des grands systèmes linéaires*, CEA-EDF-INRIA, Ecole d'été d'analyse numérique, 49, Eyrolles.

[III.4] J.L. Laminie, F. Pascal, R. Temam, *Nonlinear Galerkin method with finite element approximation*, XII ICNMFD, Oxford, 1989.

[III.5] M. Marion and R. Temam, *Nonlinear Galerkin Methods*, SIAM J. Num. Anal., 26, 1139-1157, 1989.

[III.6] M. Marion and R. Temam, *Nonlinear Galerkin Methods : the finite element case*, Numer. Math. 57, 205-226, 1990.

[III.7] P.A. Raviart, J.M. Thomas *Introduction à l'analyse numérique des équations aux dérivées partielles*, Masson.

[III.8] R. Verfurth, *Iterative methods for the numerical solution of mixed finite element approximations of the Stokes problem*, Rapport INRIA, 379, 1985.

[III.9] H. Yserentant, *Hierarchical bases give conjugate gradient type methods, a multigrid speed of convergence*, Appl. Math. and Comp. 19, 347-358, 1986.

[III.10] H. Yserentant, *On the multilevel splitting of finite element spaces*, Numer. Math. 49, 379-412, 1986.

[III.11] O.C. Zienkiewicz, D.W. Kelly, J. Gago, I. Babuska, *Hierarchical finite element approaches, error estimates and adaptive refinement*, The mathematics of finite elements and applications IV (J.R. Whiteman, ed.), Mafelap 1981, London, 1982.

Chapitre IV

BASE HIERARCHIQUE ET PROBLEMES D'EVOLUTION : IMPLEMENTATION DE LA METHODE DE GALERKIN NON LINEAIRE 2D

En tenant compte des motivations théoriques exposées au chapitre I, ainsi que des conclusions numériques discutées au chapitre III et des résultats théoriques développés par Marion et Temam [5], on se propose d'implémenter les méthodes de Galerkin non linéaires en discrétisation par éléments finis.

Dans un premier temps, les idées qui s'appuient sur la décomposition de l'espace de discrétisation induite de la base hiérarchique sont appliquées à un problème linéaire du type équation de la chaleur. Les schémas explicites voient leur stabilité améliorée et les schémas implicites leur temps d'exécution réduit.

Dans une deuxième partie, les méthodes sont implémentées pour les équations de Burgers généralisées, évolutives qui sont une forme simplifiée des équations de Navier-Stokes mais qui ont un comportement qualitatif similaire. Dans un premier temps la solution exacte est connue grâce à la transformation de Cole-Hopf, puis dans un deuxième temps seules les conditions aux limites sont connues. Les méthodes de Galerkin non linéaires qui diffèrent par le traitement des termes d'évolution convergent vers la solution et peuvent réaliser un gain de temps de calcul jusqu'à 30%. Mais dans certains cas, ces méthodes sont moins stables que des méthodes classiques.

Enfin la dernière partie concerne une première application des nouveaux schémas au problème modèle de la cavité entrainée régularisée avec un nombre de Reynolds égal à 100.

IV.1 Equation de la chaleur.

IV.1.1 Le problème et méthode d'intégration classique.

Le premier problème évolutif sur lequel ont été testés les schémas numériques basés sur la décomposition en grandes et petites structures de la solution, est un problème linéaire du type équation de la chaleur consistant à trouver u satisfaisant :

$$\frac{\partial u}{\partial t} - \frac{\Delta u}{Re} = f \quad \text{dans } \Omega \tag{IV.1}$$

où $u(0,x)$ vaut $u_0(x)$ et Re est un paramètre strictement positif, les conditions au bord étant :

- de Dirichlet : $u = g$ sur $\partial\Omega$
- mixtes i.e. de Dirichlet sur $\partial\Omega_1$ et de Neumann sur $\partial\Omega_2$ avec $\partial\Omega = \partial\Omega_1 \cup \partial\Omega_2$
  - $u = g$ sur $\partial\Omega_1$,
  - $\frac{\partial u}{\partial n} = 0$ sur $\partial\Omega_2$.

(IV.1) étant équivalent au problème variationnel suivant :
Trouver $u \in \{w \in H^1(\Omega) \ / \ w = g \text{ sur } \partial\Omega_1\}$ tel que

$$\begin{aligned} (\frac{\partial u}{\partial t}, \hat{u}) + \frac{1}{Re}((u,\hat{u})) &= (f,\hat{u}) \quad \forall \hat{u} \in \{w \in H^1(\Omega) \ / \ w = 0 \text{ sur } \partial\Omega_1\} \\ u(0,x) &= u_0(x) \end{aligned}$$

si l'on considère la triangulation régulière T_h (définie en I.3) et V_h l'espace de discrétisation par élément fini P1, le problème discret correspondant à (IV.1) est le suivant :
Trouver $u_h \in \{w_h \in V_h \ / \ w_h = g_h \text{ sur } \partial\Omega_1\}$ satisfaisant

$$\begin{aligned} (\frac{\partial u_h}{\partial t}, \hat{u}_h) + \frac{1}{Re}((u_h,\hat{u}_h)) &= (f_h,\hat{u}_h) \quad \forall \hat{u}_h \in \{w_h \in V_h \ / \ w_h = 0 \text{ sur } \partial\Omega_1\} \\ u_h(0,x) &= u_{0,h}(x) \end{aligned}$$

Par simplicité on considèrera que $\partial\Omega_1 = \partial\Omega$.

Les schémas classiques de discrétisation en temps sont :

1. **Schéma 1 :** schéma d'Euler explicite. Il s'agit d'un schéma conditionnellement stable dont la condition de stabilité est du type

$$||u_h||\frac{\Delta t}{h^2} < 1.$$

u_h^k, une approximation de u_h à l'instant $k\Delta t$ est définie récursivement par :

$$u_h^0 = u_{0,h} \tag{IV.2}$$

$$\frac{1}{\Delta t}(u_h^{k+1} - u_h^k, \hat{u}_h) = -\frac{1}{Re}((u_h^k, \hat{u}_h)) + (f_h^{k+1}, \hat{u}_h) \quad \forall \hat{u}_h \in V_{0,h} \tag{IV.3}$$

où f_h^{k+1} est une approximation de f_h à l'instant $(k+1)\Delta t$.

Lorsque V_h est muni de sa base nodale, (IV.3) est équivalent au système matriciel suivant :

$$\frac{\hat{M}}{\Delta t}U^{k+1} = (\frac{\hat{M}}{\Delta t} - \frac{\hat{A}}{Re})U^k + B^{k+1}$$

où U^k est le vecteur colonne des coordonnées de u_h^k dans la base nodale et

$$(\hat{M})_{ij} = (\hat{\phi}_{h,i}, \hat{\phi}_{h,j}) \quad 1 \leq i,j \leq n_{0,h}$$
$$(\hat{A})_{ij} = ((\hat{\phi}_{h,i}, \hat{\phi}_{h,j})) \quad 1 \leq i,j \leq n_{0,h}$$

2. **Schéma 2 :** schéma d'Euler implicite. Ce schéma est inconditionnellement stable et u_h^{k+1} est déterminée à l'instant $(k+1)\Delta t$ par :

$$\frac{1}{\Delta t}(u_h^{k+1} - u_h^k, \hat{u}_h) + \frac{1}{Re}((u_h^{k+1}, \hat{u}_h)) = (f_h^{k+1}, \hat{u}_h) \quad \forall \hat{u}_h \in V_{0,h} \tag{IV.4}$$

Le système matriciel est alors dans ce cas :

$$(\frac{\hat{M}}{\Delta t} + \frac{\hat{A}}{Re})U^{k+1} = \frac{\hat{M}}{\Delta t}U^k + B^{k+1}$$

IV.1.2 Schémas induits de la base hiérarchique.

On a vu au chapitre III que la base hiérarchique conduit à une décomposition de l'espace V_h :

$$V_h = V_{2h} \oplus W_h$$

en introduisant des petites (z_h) et grandes structures (y_h) telles que :

$$u_h = y_h + z_h.$$

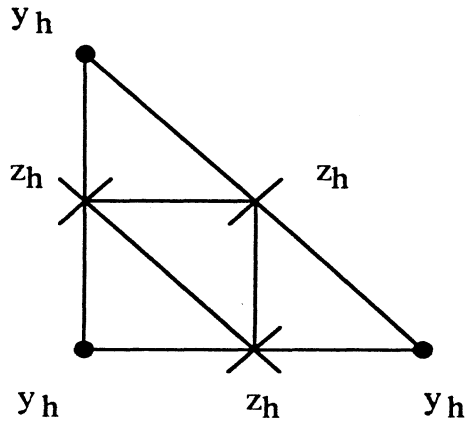


Figure IV.1: *Position des degrés de liberté pour les problèmes de la chaleur et de Burgers résolus par les méthodes issues de la base hiérarchique.*

Le système discret se met alors sous la forme :

$$\left\{\begin{array}{rcl} (\frac{\partial y_h}{\partial t},\hat{y}_h)+(\frac{\partial z_h}{\partial t},\hat{y}_h)+\frac{1}{Re}((y_h,\hat{y}_h))+\frac{1}{Re}((z_h,\hat{y}_h)) &=& (f_h,\hat{y}_h) \\ y_h(0,x) &=& y_{0,h}(x) \\ \forall \hat{y}_h \in \{w_h \in V_{2h} \ / \ w_h = 0 \text{ sur } \partial\Omega_1\} & & \end{array}\right. \quad \text{(IV.5)}$$

$$\left\{\begin{array}{rcl} (\frac{\partial y_h}{\partial t},\hat{z}_h)+(\frac{\partial z_h}{\partial t},\hat{z}_h)+\frac{1}{Re}((y_h,\hat{z}_h))+\frac{1}{Re}((z_h,\hat{z}_h)) &=& (f_h,\hat{z}_h) \\ z_h(0,x) &=& z_{0,h}(x) \\ \forall \hat{z}_h \in \{w_h \in W_h \ / \ w_h = 0 \text{ sur } \partial\Omega_1\} & & \end{array}\right. \quad \text{(IV.6)}$$

les conditions au bord étant induites de celles portant sur u_h.

On considère désormais que l'équation (IV.5) est un problème en y_h ; z_h étant fixée. Il s'agit de la discrétisation P1 du problème variationnel sur le maillage grossier T_{2h} corrigée par les valeurs "incrémentales" z_h. Ces dernières sont déterminées par l'équation (IV.6) où y_h est fixée. Il est à noter que les degrés de liberté associés à y_h sont les sommets des triangles de T_{2h} et ceux associés à z_h sont les milieux des côtés des triangles de T_{2h} (cf. figure IV.1).

Nous proposons par la suite plusieurs schémas en temps et l'étude numérique de leur convergence, exactitude, rapidité et stabilité. Dans l'ensemble de ces schémas, les termes d'évolution de couplage ie $(\frac{\partial z_h}{\partial t},\hat{y}_h)$ dans (IV.5) et $(\frac{\partial y_h}{\partial t},\hat{z}_h)$ dans (IV.6) sont négligés.

Les 4 premiers schémas traitent successivement les problèmes (IV.5) et (IV.6) explicitement et implicitement. Les 2 derniers ont une philosophie différente puisque le terme d'évolution $(\frac{\partial z_h}{\partial t}, \hat{z}_h)$ est négligé dans (IV.6). Dans ce dernier cas, (IV.5) est discrétisée implicitement puis explicitement tandis que (IV.6) est considéré comme un problème stationnaire.

1. **Schéma 3 :** (Explicite-Explicite)
   Les discrétisations en temps des équations (IV.5) et (IV.6) sont des méthodes d'Euler explicites ; à chaque pas de temps le couple (y_h^{k+1}, z_h^{k+1}) est déterminé de la façon suivante :

$$\begin{array}{lcll} \frac{1}{\Delta t}(y_h^{k+1} - y_h^k, \hat{y}_h) + \frac{1}{Re}((y_h^k + z_h^k, \hat{y}_h)) & = & (f_h^{k+1}, \hat{y}_h) & \forall \hat{y}_h \in V_{0,2h} \\ \frac{1}{\Delta t}(z_h^{k+1} - z_h^k, \hat{z}_h) + \frac{1}{Re}((y_h^{k+1} + z_h^k, \hat{z}_h)) & = & (f_h^{k+1}, \hat{z}_h) & \forall \hat{z}_h \in W_{0,h} \end{array}$$

   Ce système s'écrit matriciellement dans la base hiérarchique de V_h :

$$\begin{array}{rcl} \frac{M_{cc}}{\Delta t} Y^{k+1} & = & (\frac{M_{cc}}{\Delta t} - \frac{A_{cc}}{Re}) Y^k - \frac{A_{cf}}{Re} Z^k + B_c^{k+1} \\ \text{puis } \frac{M_{ff}}{\Delta t} Z^{k+1} & = & (\frac{M_{ff}}{\Delta t} - \frac{A_{ff}}{Re}) Z^k - \frac{A_{cf}^t}{Re} Y^{k+1} + B_f^{k+1} \end{array}$$

   Y^k et Z^k sont les vecteurs colonnes des coordonnées de y_h^k dans la base de V_{2h} et de z_h^k dans la base de W_h et les matrices M_{cc}, A_{cc}, M_{ff}, A_{ff}, A_{cf} sont données par les relations :

$$\begin{array}{rcll} (M_{cc})_{ij} & = & (\hat{\phi}_{2h,i}, \hat{\phi}_{2h,j}) & 1 \leq i,j \leq n_{0,2h} \\ (A_{cc})_{ij} & = & ((\hat{\phi}_{2h,i}, \hat{\phi}_{2h,j})) & 1 \leq i,j \leq n_{0,2h} \\ (M_{ff})_{ij} & = & (\hat{\phi}_{h,i}, \hat{\phi}_{h,j}) & n_{0,2h}+1 \leq i,j \leq n_{0,h} \\ (A_{ff})_{ij} & = & ((\hat{\phi}_{h,i}, \hat{\phi}_{h,j})) & n_{0,2h}+1 \leq i,j \leq n_{0,h} \\ (A_{cf})_{ij} & = & ((\hat{\phi}_{2h,i}, \hat{\phi}_{h,j})) & 1 \leq i \leq n_{0,2h} \\ & & & n_{0,2h}+1 \leq j \leq n_{0,h} \end{array}$$

   A chaque pas de temps, la valeur y_h^{k+1} est d'abord déterminée ; c'est une bonne approximation à cet instant donné de la solution u. Elle est ensuite corrigée en determinant la correction z_h^{k+1}.

2. **Schéma 4 :** (Explicite-Implicite)
   La méthode de discrétisation en temps de (IV.5) est la méthode d'euler explicite et celle de (IV.6) la méthode d'euler implicite. Ainsi à chaque pas de temps, les systèmes on se propose de résoudre :

$$\begin{array}{lcll} \frac{1}{\Delta t}(y_h^{k+1} - y_h^k, \hat{y}_h) + \frac{1}{Re}((y_h^k + z_h^k, \hat{y}_h)) & = & (f_h^{k+1}, \hat{y}_h) & \forall \hat{y}_h \in V_{0,2h} \\ \frac{1}{\Delta t}(z_h^{k+1} - z_h^k, \hat{z}_h) + \frac{1}{Re}((y_h^{k+1} + z_h^{k+1}, \hat{z}_h)) & = & (f_h^{k+1}, \hat{z}_h) & \forall \hat{z}_h \in W_{0,h} \end{array}$$

soit matriciellement

$$\begin{aligned} \frac{M_{cc}}{\Delta t}Y^{k+1} &= (\frac{M_{cc}}{\Delta t} - \frac{A_{cc}}{Re})Y^k - \frac{A_{cf}}{Re}Z^k + B_c^{k+1} \\ \text{puis } (\frac{M_{ff}}{\Delta t} + \frac{A_{ff}}{Re})Z^{k+1} &= \frac{M_{ff}}{\Delta t}Z^k - \frac{A_{cf}^t}{Re}Y^{k+1} + B_f^{k+1} \end{aligned}$$

3. **Schéma 5 :** (Implicite-Explicite)
L'équation (IV.5) est intégrée implicitement et l'équation (IV.6) explicitement i.e. :

$$\begin{aligned} \frac{1}{\Delta t}(y_h^{k+1} - y_h^k, \hat{y}_h) + \frac{1}{Re}((y_h^{k+1} + z_h^k, \hat{y}_h)) &= (f_h^{k+1}, \hat{y}_h) \quad \forall \hat{y}_h \in V_{0,2h} \\ \frac{1}{\Delta t}(z_h^{k+1} - z_h^k, \hat{z}_h) + \frac{1}{Re}((y_h^{k+1} + z_h^k, \hat{z}_h)) &= (f_h^{k+1}, \hat{z}_h) \quad \forall \hat{z}_h \in W_{0,h} \end{aligned}$$

soit matriciellement

$$\begin{aligned} (\frac{M_{cc}}{\Delta t} + \frac{A_{cc}}{Re})Y^{k+1} &= \frac{M_{cc}}{\Delta t}Y^k - \frac{A_{cf}}{Re}Z^k + B_c^{k+1} \\ \text{puis } \frac{M_{ff}}{\Delta t}Z^{k+1} &= (\frac{M_{ff}}{\Delta t} - \frac{A_{ff}}{Re})Z^k - \frac{A_{cf}^t}{Re}Y^{k+1} + B_f^{k+1} \end{aligned}$$

4. **Schéma 6 :** (Implicite-Implicite)
Les 2 équations sont traitées implicitement, le système s'écrit donc :

$$\begin{aligned} \frac{1}{\Delta t}(y_h^{k+1} - y_h^k, \hat{y}_h) + \frac{1}{Re}((y_h^{k+1} + z_h^k, \hat{y}_h)) &= (f_h^{k+1}, \hat{y}_h) \quad \forall \hat{y}_h \in V_{0,2h} \\ \frac{1}{\Delta t}(z_h^{k+1} - z_h^k, \hat{z}_h) + \frac{1}{Re}((y_h^{k+1} + z_h^{k+1}, \hat{z}_h)) &= (f_h^{k+1}, \hat{z}_h) \quad \forall \hat{z}_h \in W_{0,h} \end{aligned}$$

c'est-à-dire matriciellement

$$\begin{aligned} (\frac{M_{cc}}{\Delta t} + \frac{A_{cc}}{Re})Y^{k+1} &= \frac{M_{cc}}{\Delta t}Y^k - \frac{A_{cf}}{Re}Z^k + B_c^{k+1} \\ \text{puis } (\frac{M_{ff}}{\Delta t} + \frac{A_{ff}}{Re})Z^{k+1} &= \frac{M_{ff}}{\Delta t}Z^k - \frac{A_{cf}^t}{Re}Y^{k+1} + B_f^{k+1} \end{aligned}$$

5. **Schéma 7 :** (Explicite-Stationnaire)
A chaque pas de temps, y_h^{k+1} est déterminée de façon explicite :

$$\frac{1}{\Delta t}(y_h^{k+1} - y_h^k, \hat{y}_h) + \frac{1}{Re}((y_h^k + z_h^k, \hat{y}_h)) = (f_h^{k+1}, \hat{y}_h) \quad \forall \hat{y}_h \in V_{0,2h}$$

puis y_h^{k+1} est corrigée en calculant z_h^{k+1} comme solution stationnaire de (IV.6) :

$$\frac{1}{Re}((z_h^{k+1}, \hat{z}_h)) = -\frac{1}{Re}((y_h^{k+1}, \hat{z}_h)) + (f_h^{k+1}, \hat{z}_h) \quad \forall \hat{z}_h \in W_{0,h}$$

i.e. matriciellement

$$\frac{A_{ff}}{Re}Z^{k+1} = -\frac{A_{cf}^t}{Re}Y^{k+1} + B_f^{k+1}$$

6. **Schéma 8 :** (Implicite-Stationnaire)
   Ce schéma est identique au schéma 7 mais y_h^{k+1} est déterminée de façon implicite.

L'inversion des matrices $\frac{M_{cc}}{\Delta t}$ et $\frac{M_{cc}}{\Delta t} + \frac{A_{cc}}{Re}$ est réalisée par un ICCG et celle de $\frac{M_{ff}}{\Delta t}$ et $\frac{M_{ff}}{\Delta t} + \frac{A_{ff}}{Re}$ par une méthode de relaxation du type SSOR.

IV.1.3 Résultats numériques.

3 types d'expériences numériques ont été menés à terme :

- le premier concerne des conditions au bord du type Dirichlet homogène dans la cavité $[-1,1] \times [-1,1]$ et une solution exacte connue :
  $u_{ex} = (x-1)(x+1)(y-1)(y+1)$,
- le deuxième porte sur des conditions au bord du type Dirichlet non homogènes sur le domaine $[-1,1] \times [-1,1]$:
  - $u = 0$ sur les parois verticales et sur la paroi horizontale inférieure,
  - $u(x_1, 1) = 10(1 - x_1^2)^2$ sur la paroi horizontale supérieure

  avec un forçage f constant de densité 1.
- les dernières expériences ont été effectuées avec un forçage f constant 1 et avec des conditions au bord mixtes de Dirichlet et Neumann :
  - $u(x_1, x_2) = 0$ sur
    $\{(x_1, x_2)/x_2 = -1; -1 \leq x_1 \leq 1\} \cup \{(x_1, x_2)/x_1 = +1; -1 \leq x_2 \leq 1\}$,
  - $u(x_1, x_2) = 10(1 - x_1^2)^2$ sur $\{(x_1, x_2)/x_2 = +1; -1 \leq x_1 \leq 1\}$,
  - $\frac{\partial u}{\partial n}(x_1, x_2) = 0$ sur $\{(x_1, x_2)/x_1 = -1; -1 \leq x_2 \leq 1\}$

Les paramètres de discrétisation valent

$$h = \frac{1}{8}, \frac{1}{16}, \frac{1}{32} \text{ et } \frac{\Delta t}{h^2} = 6.4, 0.512, 0.320, 0.128, 0.064.$$

Les tableaux IV.1 et IV.2 donnent pour la première expérience l'erreur relative à l'instant $t = 1.0$ et le temps d'exécution moyen par itération lorsque l'on fait varier le pas de temps. Quand il y a convergence, ils permettent de vérifier l'exactitude de l'ensemble des schémas.

Les résultats concernant le temps de calcul sont similaires pour les 2 autres tests numériques : le gain de temps varie de 10% à 20% pour les schémas 7 et 8 dès que le maillage est suffisamment fin ($h = 0.0625$ et $h = 0.03125$). Ce gain n'est plus que de l'ordre de 5% pour les schémas qui ne négligent pas l'évolution de z_h.

Mais les résultats les plus intéressants portent sur la stabilité relative des schémas :

- le schéma 6 qui traite implicitement les 2 équations est, tout comme le schéma classique implicite, inconditionnellement stable.
- bien que l'on ait approché le problème (IV.6) par un problème stationnaire, le schéma 8 qui traite y_h implicitement est inconditionnellement stable.
- les schémas 4 et 7 ont pour point commun d'être explicite en y_h. Le schéma 4 intègre implicitement l'équation (IV.6) et le schéma 7 l'approche statiquement. Ces opérations ne destabilisent pas le problème en y_h qui impose donc sa condition de stabilité à l'ensemble du schéma : une condition 4 fois plus faible. Il s'agit en effet de la condition d'Euler explicite sur le maillage T_{2h}.
- la condition de stabilité des schémas 3 et 5 est celle de la méthode d'Euler explicite pour le problème z_h dont le pas de discrétisation en espace est celui du maillage fin.

Les figures IV.2 et IV.3 qui représentent la solution à l'instant $t = 1.0$ ainsi que l'erreur relative et son évolution au cours du temps pour le schéma classique d'Euler implicite et pour le schéma 8 permettent d'apprécier l'exactitude du schéma 8 et ce malgré les approximations effectuées dans l'équation (IV.6).

Ainsi pour le problème de la chaleur, la décompostion due à l'utilisation de la base hiérarchique a permis de mettre au point :

- des schémas qui améliorent la stabilité de la méthode d'Euler explicite (environ d'un facteur 4);
- des schémas aussi stables qu'Euler implicite et d'un coût inférieur.

	$\Delta t/h^2$ $h = 16$	Schéma 1	Schéma 2	Schéma 3	Schéma 4
Err. relative	6.4	DV	$0.12\ 10^{-2}$	DV	DV
Temps CPU			$0.88\ 10^{-1}$s		
Err. relative	0.512	DV	$0.73\ 10^{-4}$	DV	DV
Temps CPU			$0.55\ 10^{-1}$s		
Err. relative	0.320	DV	$0.45\ 10^{-4}$	DV	$0.45\ 10^{-4}$
Temps CPU			$0.53\ 10^{-1}$s		$0.46\ 10^{-1}$s
Err. relative	0.128	DV		DV	
Temps CPU					
Err. relative	0.064	$0.89\ 10^{-5}$	$0.89\ 10^{-5}$	$0.90\ 10^{-5}$	$0.90\ 10^{-5}$
Temps CPU		$0.45\ 10^{-1}$s	$0.45\ 10^{-1}$s	$0.43\ 10^{-1}$s	$0.43\ 10^{-1}$s

Tableau IV.1: *Erreur relative et temps de calcul pour les différents schémas avec* $h = 0.0625$. *Schémas 1 et 2 : schémas classiques. Schémas 3 à 8 : base hiérarchique.*

	$\Delta t/h^2$ $h = 16$	Schéma 5	Schéma 6	Schéma 7	Schéma 8
Err. relative	6.4	DV	$0.12\ 10^{-2}$	DV	$0.12\ 10^{-2}$
Temps CPU			$0.50\ 10^{-1}$s		$0.46\ 10^{-1}$s
Err. relative	0.512	DV	$0.74\ 10^{-4}$	DV	$0.73\ 10^{-4}$
Temps CPU			$0.46\ 10^{-1}$s		$0.42\ 10^{-1}$s
Err. relative	0.320	DV	$0.45\ 10^{-4}$	$0.45\ 10^{-4}$	$0.46\ 10^{-4}$
Temps CPU			$0.46\ 10^{-1}$s	$0.42\ 10^{-1}$s	$0.42\ 10^{-1}$s
Err. relative	0.128	DV			
Temps CPU					
Err. relative	0.064	$0.90\ 10^{-5}$	$0.90\ 10^{-5}$	$0.90\ 10^{-5}$	$0.90\ 10^{-5}$
Temps CPU		$0.43\ 10^{-1}$s	$0.43\ 10^{-1}$s	$0.39\ 10^{-1}$s	$0.39\ 10^{-1}$s

Tableau IV.2: *Erreur relative et temps de calcul pour les différents schémas avec* $h = 0.0625$. *Schémas 1 et 2 : schémas classiques. Schémas 3 à 8 : base hiérarchique. DV = divergence.*

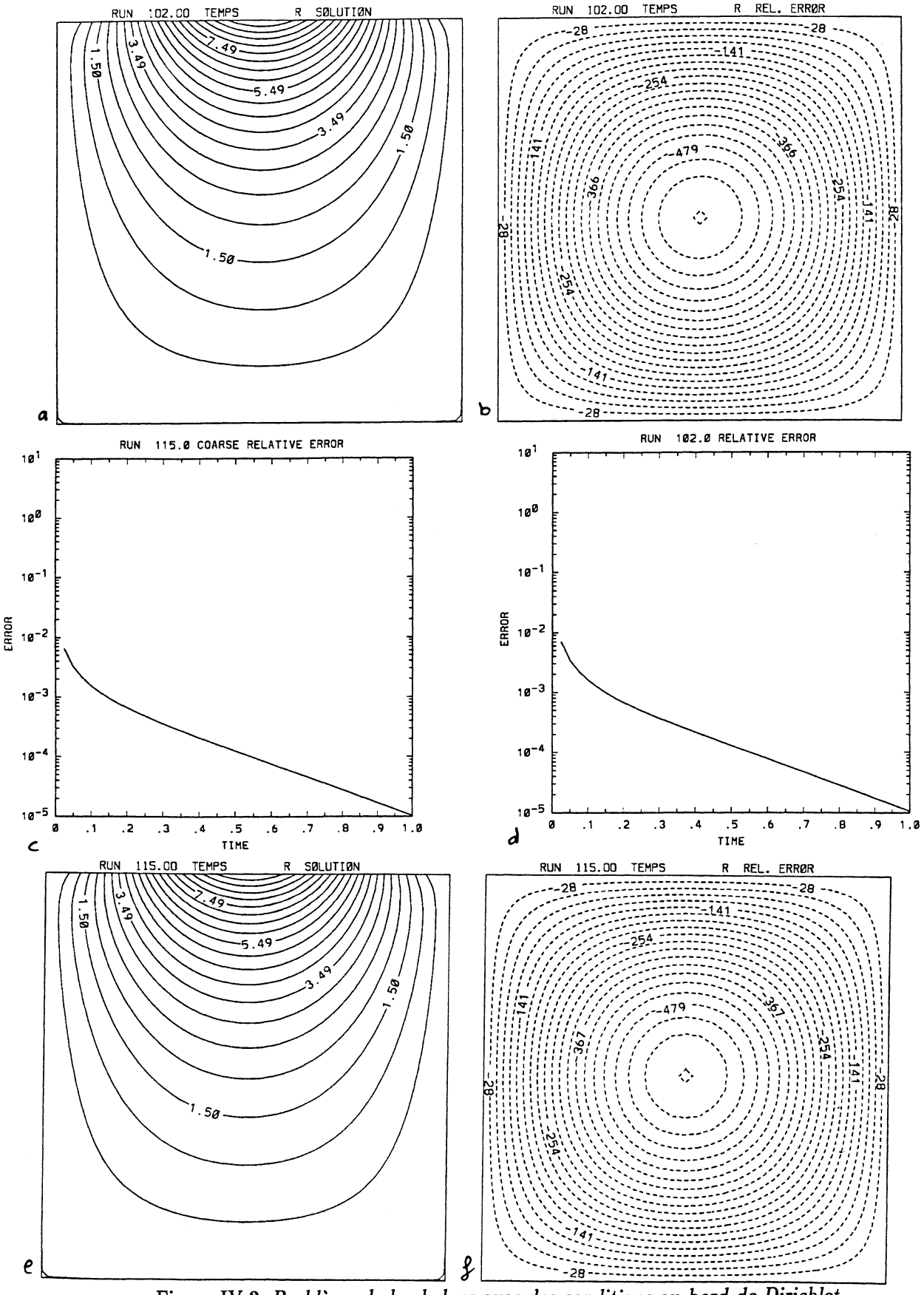


Figure IV.2: *Problème de la chaleur avec des conditions au bord de Dirichlet.* a), b), d) *Run 102 : Schéma 2.* c), e), f) *Run 115 : Schéma 8.* a), e) *: solution.* b), f) *: erreur relative.* c), d) *: norme* L_2 *de l'erreur relative.*

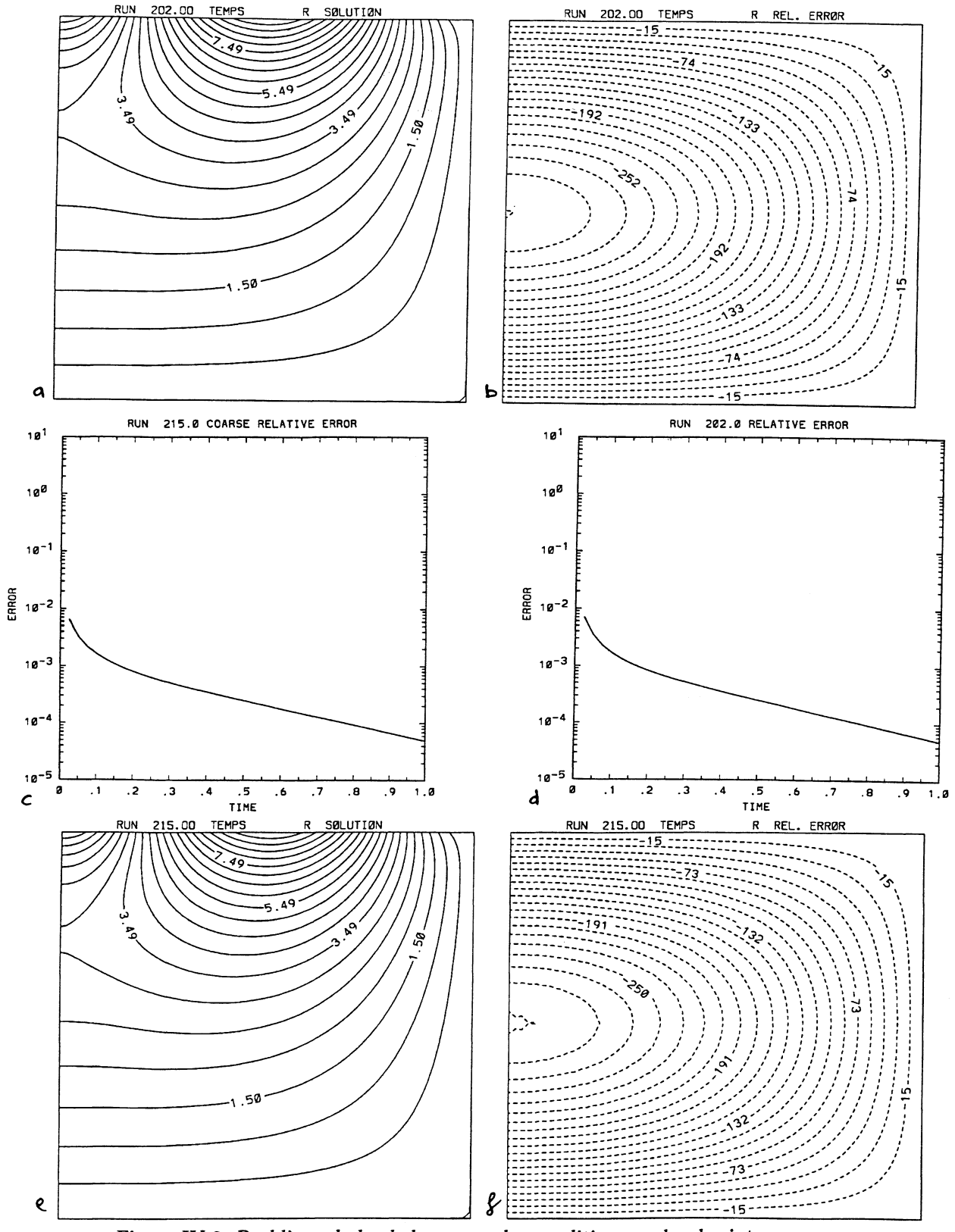


Figure IV.3: *Problème de la chaleur avec des conditions au bord mixtes.*
a), b), d) Run 102 : Schéma 2. c), e), f) Run 115 : Schéma 8. a), e) : solution.
b), f) : erreur relative. c), d) : norme L_2 de l'erreur relative.

IV.2 Un problème non linéaire : équations de Burgers 2D.

Les équations de Burgers généralisées sont une forme simplifiée des équations de Navier-Stokes qu'elles approchent qualitativement. Elles conservent les parties convectives et dissipatives mais les termes du gradient de pression et de la divergence de la vitesse ne sont pas retenus. Ainsi la solution du problème de Burgers peut présenter de forts gradients (un des comportements essentiels de la solution de Navier-Stokes) dus à l'interaction des termes non linéaires et des termes dissipatifs mais peut ne pas vérifier l'équation de continuité.

Ces équations sont données par les relations :

$$\begin{aligned} \frac{\partial u}{\partial t} - \frac{\Delta u}{Re} + (u.\nabla)u &= 0 \quad \text{dans } \Omega \\ u &= g \quad \text{sur } \partial\Omega \end{aligned}$$

c'est-à-dire si $u = (u_1, u_2)$:

$$\begin{aligned} &\frac{\partial u_1}{\partial t} - \frac{1}{Re}\left(\frac{\partial^2 u_1}{\partial x_1^2} + \frac{\partial^2 u_1}{\partial x_2^2}\right) + \left(u_1\frac{\partial}{\partial x_1} + u_2\frac{\partial}{\partial x_2}\right)u_1 = 0 \\ &\frac{\partial u_2}{\partial t} - \frac{1}{Re}\left(\frac{\partial^2 u_2}{\partial x_1^2} + \frac{\partial^2 u_2}{\partial x_2^2}\right) + \left(u_1\frac{\partial}{\partial x_1} + u_2\frac{\partial}{\partial x_2}\right)u_2 = 0 \end{aligned} \qquad \text{(IV.7)}$$

avec des conditions initiales.

Dans le cadre d'un domaine rectangulaire dont les dimensions seront précisées ultérieurement, il est possible de construire une solution exacte en utilisant la transformation de Cole-Hopf. Fletcher [1] donne les étapes de cette construction.

Cette solution qui peut présenter de faibles ou forts gradients internes ou des gradients près du bord permet de mesurer le temps d'exécution, l'exactitude et l'efficacité des schémas numériques (voir par exemple l'étude de Fletcher [2] sur des méthodes d'éléments finis et différences finis).

Dans un premier temps la méthode de Galerkin non linéaire a été appliquée à ces équations connaissant la solution exacte, puis pour étudier la stabilité et l'influence du nombre de Reynolds dans un cas où la solution exacte est inconnue mais converge vers une solution stationnaire.

IV.2.1 Tests avec la solution exacte connue.

On suppose $\Omega = [-1, 1] \times [0, \frac{\pi}{6\alpha}]$. L'expression de la solution exacte est la suivante :

$$u_1 = -\frac{2}{Re}\left[\frac{a_1 + a_3x_2 + \alpha a_4\{e^{\alpha(x_1-x_0)} - e^{-\alpha(x_1-x_0)}\}\cos\alpha x_2}{a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4\{e^{\alpha(x_1-x_0)} + e^{-\alpha(x_1-x_0)}\}\sin\alpha x_2}\right]$$

$$u_2 = -\frac{2}{Re}\left[\frac{a_2 + a_3x_2 - \alpha a_4\{e^{\alpha(x_1-x_0)} + e^{-\alpha(x_1-x_0)}\}\sin\alpha x_2}{a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4\{e^{\alpha(x_1-x_0)} + e^{-\alpha(x_1-x_0)}\}\cos\alpha x_2}\right]$$

La solution présente un faible gradient interne (au domaine) pour le choix des paramètres suivants :

$$a_0 = a_1 = 110.13 \ ; \ a_2 = a_3 = 0 \ ; \ a_4 = 1.0$$

$$\alpha = 5 \ ; \ x_0 = 1; Re = 100.$$

Dans le cas où ces paramètres sont :

$$a_0 = a_1 = 1.2962 \ 10^{13} \ ; \ a_2 = a_3 = 0 \ ; \ a_4 = 1.0$$

$$\alpha = 25 \ ; \ x_0 = 1; Re = 100$$

la composante u_1 de la solution possède un fort gradient interne dans la première direction. Les profils médiants des composantes de u i.e. $u_1(., \frac{\pi}{3\alpha})$, $u_2(., \frac{\pi}{3\alpha})$ sur la médiane horizontale, $u_1(0,.)$, $u_2(0,.)$ sur la médiane verticale, représentés sur les tracés IV.5 à IV.8 permettent d'apprécier ces gradients.

La méthode de Galerkin classique en éléments P1 : GUS.

Elle consiste à chercher $u_h \in \mathcal{V}_h = V_h \times V_h$ muni de la base nodale telle que :

$$\begin{array}{rcll} (\frac{\partial u_h}{\partial t}, \hat{u}_h) + \frac{1}{Re}((u_h, \hat{u}_h)) + ((u_h.\nabla)u_h, \hat{u}_h) & = & 0 & \forall \hat{u}_h \in \mathcal{V}_{0,h} \\ u_h & = & g_h & \text{sur } \partial\Omega \\ u_h(0,x) & = & u_{0,h}(x) & \text{dans } \Omega \end{array}$$

La discrétisation en temps initialisée par une méthode d'Euler est composée :

- d'un schéma d'Adams-Bashforth explicite pour les termes non linéaires
- d'un schéma d'Euler implicite pour les termes linéaires.

u_h^{k-1}, u_h^k étant connues, u_h^{k+1} est solution de l'équation :

$$\frac{1}{\Delta t}(u_h^{k+1} - u_h^k, \hat{u}_h) + \frac{1}{Re}((u_h^{k+1}, \hat{u}_h)) = -\frac{3}{2}((u_h^k.\nabla)u_h^k, \hat{u}_h) + \frac{1}{2}((u_h^{k-1}.\nabla)u_h^{k-1}, \hat{u}_h)$$

$\forall \hat{u}_h \in \mathcal{V}_{0,h}$

qui s'écrit matriciellement dans la base nodale de V_h de la façon suivante (le solveur utilisé est un ICCG) :

$$(\frac{\hat{M}}{\Delta t} + \frac{\hat{A}}{Re})U_l^{k+1} = \frac{\hat{M}}{\Delta t}U_l^k + B_l^{k+1} \quad \text{pour } l = 1,2$$

où

$$\begin{array}{rcll} (\hat{M})_{ij} & = & (\hat{\phi}_{h,i}, \hat{\phi}_{h,j}) & 1 \leq i,j \leq n_{0,h} \\ (\hat{A})_{ij} & = & ((\hat{\phi}_{h,i}, \hat{\phi}_{h,j})) & 1 \leq i,j \leq n_{0,h} \end{array}$$

Les méthodes de Galerkin non linéaires

Elles consistent à chercher $u_h \in \mathcal{V}_h$ sous la forme $y_h + z_h$ avec $y_h \in \mathcal{V}_{2h}$ et $z_h \in \mathcal{W}_h$, décomposition issue de la base hiérarchique, y_h et z_h étant solutions de :

$$\begin{cases} (\frac{\partial y_h}{\partial t}, \hat{y}_h) + (\frac{\partial z_h}{\partial t}, \hat{y}_h) + \frac{1}{Re}((y_h + z_h, \hat{y}_h)) + ((u_h.\nabla)u_h, \hat{y}_h)) = 0 \\ \forall \hat{y}_h \in \mathcal{V}_{0,2h} \end{cases} \quad \text{(IV.8)}$$

$$\begin{cases} (\frac{\partial y_h}{\partial t}, \hat{z}_h) + (\frac{\partial z_h}{\partial t}, \hat{z}_h) + \frac{1}{Re}((y_h + z_h, \hat{z}_h)) + ((u_h.\nabla)u_h, \hat{z}_h)) = 0 \\ \forall \hat{z}_h \in \mathcal{W}_{0,h} \end{cases} \quad \text{(IV.9)}$$

avec des conditions initiales

$$\begin{array}{rcll} y_h(0,x) & = & y_{0,h}(x) & \text{dans } \Omega \\ z_h(0,x) & = & z_{0,h}(x) & \text{dans } \Omega \end{array} \quad \text{(IV.10)}$$

et des conditions au bord

$$\begin{array}{rcll} y_h & = & gy_h & \text{sur } \partial\Omega \\ z_h & = & gz_h & \text{sur } \partial\Omega \end{array} \quad \text{(IV.11)}$$

où $gy_h \in \mathcal{V}_{2h}$ et $gz_h \in \mathcal{W}_h$ vérifient $g_h = gy_h + gz_h$.

Les propriétés de la base hiérarchique (établies en section III.1.2) permettent d'envisager des approximations de certains termes des équations (IV.8) et (IV.9). 3 schémas numériques dont le dernier a été étudié théoriquement par Marion et Temam [5] sont considérés et appliqués au problème de Burgers généralisé. En voici une description :

1. **Schéma GNL1 :**

   Par analogie avec la discrétisation spectrale, les termes variationnels couplant les équations (IV.8) et (IV.9) à savoir

   $$(\frac{\partial z_h}{\partial t}, \hat{y}_h) \text{ dans } (IV.8) \quad \text{et} \quad (\frac{\partial y_h}{\partial t}, \hat{z}_h) \text{ dans } (IV.9)$$

   sont négligés. La convergence théorique de cet algorithme est encore un problème ouvert : aucune estimation a priori n'a pu être établie étant donné qu'aucune approximation n'est effectuée dans les termes non linéaires. La discrétisation en temps est constituée d'un schéma d'Euler implicite pour les termes linéaires et d'un schéma d'Adams-Bashforth explicite pour les termes non linéaires.

Si y_h^k et z_h^k sont connues, alors $y_h^{k+1} \in \mathcal{V}_{2h}$ à l'instant $(k+1)\Delta t$ est définie par :

$$\left\{ \begin{array}{lll} \frac{1}{\Delta t}(y_h^{k+1} - y_h^k, \hat{y}_h) & + & \\ \frac{1}{Re}((y_h^{k+1} + z_h^k, \hat{y}_h)) & = & -\frac{3}{2}((u_h^k \ .\nabla)u_h^k \ , \hat{y}_h) \\ & & +\frac{1}{2}((u_h^{k-1}.\nabla)u_h^{k-1}, \hat{y}_h) \\ \forall \hat{y}_h \in \mathcal{V}_{0,2h} & & \end{array} \right.$$

et $z_h^{k+1} \in \mathcal{W}_h$ à l'instant $(k+1)\Delta t$ est donnée par :

$$\left\{ \begin{array}{lll} \frac{1}{\Delta t}(z_h^{k+1} - z_h^k, \hat{z}_h) & + & \\ \frac{1}{Re}((y_h^{k+1} + z_h^{k+1}, \hat{z}_h)) & = & -\frac{3}{2}((u_h^k \ .\nabla)u_h^k \ , \hat{z}_h) \\ & & +\frac{1}{2}((u_h^{k-1}.\nabla)u_h^{k-1}, \hat{z}_h) \\ \forall \hat{z}_h \in \mathcal{W}_{0,h} & & \end{array} \right.$$

avec l'équivalent discret des conditions au bord (IV.11).

La première équation est un problème en y_h dont la discrétisation en éléments finis est conforme ; les degrés de libertés associés à y_h sont les sommets des triangles de la triangulation grossière T_{2h} (cf. figure IV.1). La deuxième équation est un problème en z_h dont la discrétisation en éléments finis est non conforme ; les degrés de liberté sont les milieux des côtés des triangles de T_{2h} (cf. figure IV.1). A chaque pas de temps, les matrices

$$\frac{M_{cc}}{\Delta t} + \frac{A_{cc}}{Re} \quad \text{et} \quad \frac{M_{ff}}{\Delta t} + \frac{A_{ff}}{Re}$$

sont à inverser pour chacune des composantes ; le solveur est un ICCG.

2. **Schéma GNL2 :**

z_h a une petite amplitude et étant donné que l'on a

$$(u_h.\nabla)u_h = (y_h.\nabla)y_h + (y_h.\nabla)z_h + (z_h.\nabla)y_h + (z_h.\nabla)z_h,$$

$((z_h.\nabla)z_h, \hat{y}_h)$ est négligé dans la première équation et $((y_h.\nabla)z_h + (z_h.\nabla)y_h + (z_h.\nabla)z_h, \hat{z}_h)$ est négligé dans la deuxième équation. La discrétisation en temps étant identique à celle de GNL1, le schéma GNL2 consiste à chaque pas de temps à déterminer les grandes structures y_h et les petites structures z_h comme solutions des systèmes suivants avec l'équivalent discret

des conditions au bord (IV.11) :

$$\left\{\begin{array}{lll} \frac{1}{\Delta t}(y_h^{k+1}-y_h^k,\hat{y}_h) & + & \\ \frac{1}{Re}((y_h^{k+1}+z_h^k,\hat{y}_h)) & = & -\frac{3}{2}((y_h^k\ .\nabla)y_h^k\ +(y_h^k\ .\nabla)z_h^k\ +(z_h^k\ .\nabla)y_h^k\ ,\hat{y}_h) \\ & & +\frac{1}{2}((y_h^{k-1}.\nabla)y_h^{k-1}+(y_h^{k-1}.\nabla)z_h^{k-1}+(z_h^{k-1}.\nabla)y_h^{k-1},\hat{y}_h) \\ \forall \hat{y}_h \in \mathcal{V}_{0,2h} & & \end{array}\right.$$

$(y_h^k,\ z_h^k,\ y_h^{k-1},\ z_h^{k-1},$ étant connues)

$$\left\{\begin{array}{lll} \frac{1}{\Delta t}(z_h^{k+1}-z_h^k,\hat{z}_h) & + & \\ \frac{1}{Re}((y_h^{k+1}+z_h^{k+1},\hat{z}_h)) & = & -\frac{3}{2}((y_h^k\ .\nabla)y_h^k\ ,\hat{z}_h) \\ & & +\frac{1}{2}((y_h^{k-1}.\nabla)y_h^{k-1},\hat{z}_h) \\ \forall \hat{z}_h \in \mathcal{W}_{0,h} & & \end{array}\right.$$

$(z_h^k,\ y_h^{k+1},\ y_h^k,\ y_h^{k-1},$ étant connues).

3. Schéma GNL3 :

Ce schéma diffère des précédents dans le sens où l'on néglige en plus le terme d'évolution en z_h dans l'équation (IV.9). Ainsi $y_h(t)$ et $z_h(t)$ sont déterminées dans $\mathcal{V}_{2h}$ et $\mathcal{W}_h$ par :

$$\left\{\begin{array}{lll} (\frac{\partial y_h}{\partial t},\hat{y}_h)+\frac{1}{Re}((y_h+z_h,\hat{y}_h)) & + & \\ ((y_h.\nabla)y_h+(y_h.\nabla)z_h+(z_h.\nabla)y_h,\hat{y}_h) & = & 0 \\ \forall \hat{y}_h \in \mathcal{V}_{0,2h} & & \end{array}\right. \qquad \text{(IV.12)}$$

$$\left\{\begin{array}{lll} \frac{1}{Re}((y_h+z_h,\hat{z}_h))+((y_h.\nabla)y_h,\hat{z}_h) & = & 0 \\ \forall \hat{z}_h \in \mathcal{W}_{0,h} & & \end{array}\right. \qquad \text{(IV.13)}$$

(IV.13) est un problème linéaire en $z_h(t)$ qui s'exprime comme fonction quadratique de $y_h(t)$. (IV.12) est l'équation (modifiée non linéairement par z_h) que vérifie $u_{2h} \in \mathcal{V}_{2h}$ (solution approchée sur le maillage grossier T_{2h}).

Ainsi $u_h = y_h + z_h$ est une solution approchée appartenant à une variété $\mathcal{M}_{1h}$ de $\mathcal{V}_h$ de dimension égale à la dimension de $\mathcal{V}_{2h}$ et d'équation :

$$z_h = -y_h + A_h^{-1}(-B_h(y_h,y_h))$$

où A_h et B_h sont les opérateurs vérifiant :

$$\begin{array}{llll} A_h \in \mathrm{L}(\mathcal{V}_h) & / & (A_h\phi_h,\psi_h)=\frac{1}{Re}((\phi_h,\psi_h)) & \forall \phi_h,\ \psi_h \in \mathcal{V}_h \\ B_h(y_h,y_h) \in \mathcal{V}_h & / & (B_h(y_h,y_h),\psi_h)=((y_h.\nabla)y_h,\psi_h) & \forall \psi_h \in \mathcal{V}_h \end{array}$$

Marion et Temam [5], après avoir donné des estimations a priori pour y_h et z_h, ont montré l'existence et l'unicité pour tout temps de y_h et z_h et établi les convergences au sens suivant (lorsque $h \longrightarrow 0$) :

- $u_h \longrightarrow u$ dans $L^p(0,T;(H^1(\Omega))^2) \quad \forall T, \quad \forall p \in [1,+\infty]$
- $y_h \longrightarrow u$ dans $L^2(0,T;(L^2(\Omega))^2)$
- $z_h \longrightarrow 0$ dans $L^p(0,T;(L^2(\Omega))^2)$

La discrétisation en temps consiste à déterminer y_h^{k+1} et z_h^{k+1} de la manière suivante :

$$\left\{\begin{array}{lll} \frac{1}{\Delta t}(y_h^{k+1}-y_h^k,\hat{y}_h) & + & \\ \frac{1}{Re}((y_h^{k+1}+z_h^k,\hat{y}_h)) & = & -\frac{3}{2}((y_h^k\ .\nabla)y_h^k\ +(y_h^k\ .\nabla)z_h^k\ +(z_h^k\ .\nabla)y_h^k\ ,\hat{y}_h) \\ & & +\frac{1}{2}((y_h^{k-1}.\nabla)y_h^{k-1}+(y_h^{k-1}.\nabla)z_h^{k-1}+(z_h^{k-1}.\nabla)y_h^{k-1},\hat{y}_h) \\ \forall \hat{y}_h \in \mathcal{V}_{0,2h} & & \end{array}\right.$$

puis

$$\left\{\begin{array}{lll} \frac{1}{Re}((y_h^{k+1}+z_h^{k+1},\hat{z}_h)) & = & -\frac{3}{2}((y_h^k\ .\nabla)y_h^k\ ,\hat{z}_h) \\ & & +\frac{1}{2}((y_h^{k-1}.\nabla)y_h^{k-1},\hat{z}_h) \\ \forall \hat{z}_h \in \mathcal{W}_{0,h} & & \end{array}\right.$$

Résultats numériques.

La précision exigée pour la convergence des gradients conjugués et des schémas GUS, GNL1, GNL2, GNL3 est égale à 0.00001. Rappelons que pour les nombres de Reynolds considérés la solution converge vers une solution stationnaire. Le tableau IV.3 (respectivement IV.4) donne le nombre d'itérations et le temps CPU moyen par itération pour obtenir la convergence dans le cas où la solution exacte présente un faible gradient (respectivement un fort gradient) et compare ces temps avec le temps nécessaire à la méthode de Galerkin classique.

Dès que le nombre de degrés de liberté est suffisant, les méthodes de Galerkin non linéaires sont extrèmement performantes question temps de calcul et nécessitent autant d'itérations pour atteindre la solution stationnaire que la méthode de Galerkin classique. Cependant la présence d'un fort gradient atténue cette efficacité et impose un maillage très fin : seulement 7% de gain de temps avec $h = \frac{1}{128}$.

Il est à noter que la suppression d'une partie des termes non linéaires réduit considérablement le temps d'exécution (en effet chaque partie est calculée séparément). Ces approximations, dans le cas présent ne modifient pas la précision des schémas GNL et parfois l'améliorent : la norme L_2 de l'erreur absolue est inférieure à celle obtenue par la méthode de Galerkin classique (voir les courbes

de l'erreur pour la première composante en figure IV.5 et pour la deuxième composante en figure IV.6).

Les profils sur les médianes verticales et horizontales de la différence entre la solution exacte et la solution calculée sont de faible amplitude et montrent l'exactitude des nouveaux schémas, y compris GNL3 qui fige l'évolution des petites structures (voir figure IV.5 à IV.8). Il est à remarquer que la présence d'une zone de fort gradient dans la première direction engendre de fortes oscillations de l'erreur dans cette direction et au voisinage de cette zone que n'atténue pas ou n'amplifie pas les méthodes de Galerkin non linéaires.

Enfin le gain de temps CPU s'explique par le fait qu'à chaque pas de temps 2 systèmes linéaires de dimension réduite sont inversés.

En conclusion, aux vues de ces premières expériences numériques, les méthodes de Galerkin non linéaires convergent vers la bonne solution et plus rapidement.

IV.2.2 Tests avec la solution exacte inconnue.

On suppose $\Omega = [0, 0.5] \times [0, 0.5]$. La condition initiale est :

$$\begin{aligned} u_1(x_1, x_2, 0) &= \sin(\pi x_1) + \sin(\pi x_2) \\ u_2(x_1, x_2, 0) &= x_1 + x_2 \end{aligned}$$

Les conditions au bord sont celles de Jain et Holla [4] à savoir :

$$\begin{aligned} u_1(0, x_2, t) &= \cos(\pi x_2) & \\ u_2(0, x_2, t) &= x_2 & 0 \leq x_2 \leq 0.5 \\ u_1(0.5, x_2, t) &= 1 + \cos(\pi x_2) & \\ u_2(0.5, x_2, t) &= 0.5 + x_2 & 0 \leq x_2 \leq 0.5 \\ u_1(x_1, 0, t) &= \sin(\pi x_1) + 1 & \\ u_2(x_1, 0, t) &= x_1 & 0 \leq x_1 \leq 0.5 \\ u_1(x_1, 0.5, t) &= \sin(\pi x_1) & \\ u_2(x_1, 0.5, t) &= x_1 + 0.5 & 0 \leq x_1 \leq 0.5 \end{aligned}$$

Reynolds 100.

Les figures IV.9 et IV.10 donnent les profils sur la médiane verticale et la médiane horizontale des composantes de la solution calculée avec la méthode de Galerkin classique et la méthode GNL2 aux temps 0.02, 0.2, 0.5. Les profils de la différence des 2 solutions en figure IV.11 permettent de noter que les 2 solutions sont très voisines, la différence ne dépassant par $1.0\ 10^{-2}$ à $t = 0.5$.

La solution initialement régulière évolue rapidement jusqu'à $t = 0.5$, au delà de ce temps les variations sont en effet très faibles. On constate qu'au cours du temps, de forts gradients s'établissent en particulier dans la 1e direction pour les 2 composantes de la vitesse et dans la 2e direction pour la 2e composante.

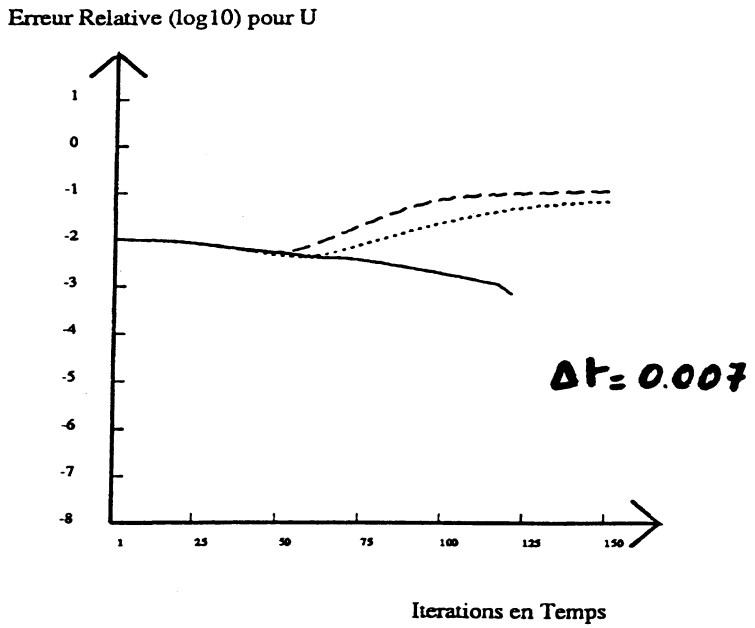


Figure IV.4: *Problème de Burgers. Courbes de convergence pour Galerkin classique, GNL1, GNL2.* $Re = 100$, $h = \frac{1}{128}$.

C'est dans ces zones de forts gradients que se situe la plus grande différence entre les 2 solutions. On retrouve donc que la présence de forts gradients limite la convergence et certainement la stabilité des nouveaux schémas tant que le maillage est trop grossier.

Pour $h = \frac{1}{256}$ et $\Delta t = 0.005$, l'erreur relative est inférieure à 10^{-3} à $t = 0.58$. Pour un pas de temps plus élevé, $\Delta t = 0.007$, la méthode de Galerkin classique converge mais la précision 10^{-3} n'est atteinte qu'à $t = 0.85$. En ce qui concerne les schémas GNL1 et GNL2 avec un tel pas de temps, l'erreur relative décroit pendant une soixantaine d'itérations puis augmente. Les schémas de Galerkin non linéaires sont donc moins stables que les schémas classiques (cf. figure IV.4) (En deçà de $\Delta t = 0.007$, la méthode classique ne converge pas.) Pour GNL3, des résultats similaires à GNL2 sont observés.

Reynolds 500.

Pour un tel nombre de Reynolds, la composante u_1 de la solution présente au voisinage de $x_1 = 0.5$ un très fort gradient qui n'est correctement simulé avec des méthodes de Galerkin classiques que si le pas de discrétisation est suffisamment faible (cf. figure IV.12). En effet pour $h = \frac{1}{128}$, u_1 oscille dans cette zone et ces oscillations disparaissent presque complètement pour $h = \frac{1}{256}$.

Cette régularisation de u_1 se retrouve avec GNL2, cependant des oscillations apparaissent sur l'ensemble du domaine.

maillage	méthode	nombre itérations	temps cpu moyen / itérations	pourcentage / GUS (itération)	pourcentage / GUS (total)
65 × 65	GUS	7	1.91		
	GNL1	9	2.67	+40 %	+80 %
	GNL2	9	1.93	+01 %	+30 %
	GNL3	9	2.00	+05 %	+34 %
129 × 129	GUS	6	12.72		
	GNL1	7	11.46	-09 %	-05 %
	GNL2	7	08.50	-33 %	-22 %
	GNL3	7	08.55	-32 %	-22 %
257 × 257	GUS	6	70.76		
	GNL1	6	50.57	-29 %	-29 %
	GNL2	6	39.11	-45 %	-45 %
	GNL3	6	40.08	-43 %	-43 %

Tableau IV.3: *Problème de Burgers : performances des schémas lorsque la solution présente un faible gradient interne.* $\Delta t = 1.0$.

maillage	méthode	nombre itérations	temps cpu moyen / itérations	pourcentage / GUS (itération)	pourcentage / GUS (total)
65 × 65	GUS	8	1.47		
	GNL1	10	2.58	+136%	+165%
	GNL2	10	1.82	+22 %	+55 %
	GNL3	10	1.80	+21 %	+53 %
129 × 129	GUS	7	09.56		
	GNL1	8	10.87	+14 %	+30 %
	GNL2	8	07.77	-19 %	-07 %
	GNL3	8	07.77	-19 %	-07 %
257 × 257	GUS	8	50.76		
	GNL1	8	45.23	-11 %	-11 %
	GNL2	7	35.02	-31 %	-40 %
	GNL3	7	36.00	-30 %	-40 %

Tableau IV.4: *Problème de Burgers : performances des schémas lorsque la solution présente un fort gradient interne.* $\Delta t = 1.0$.

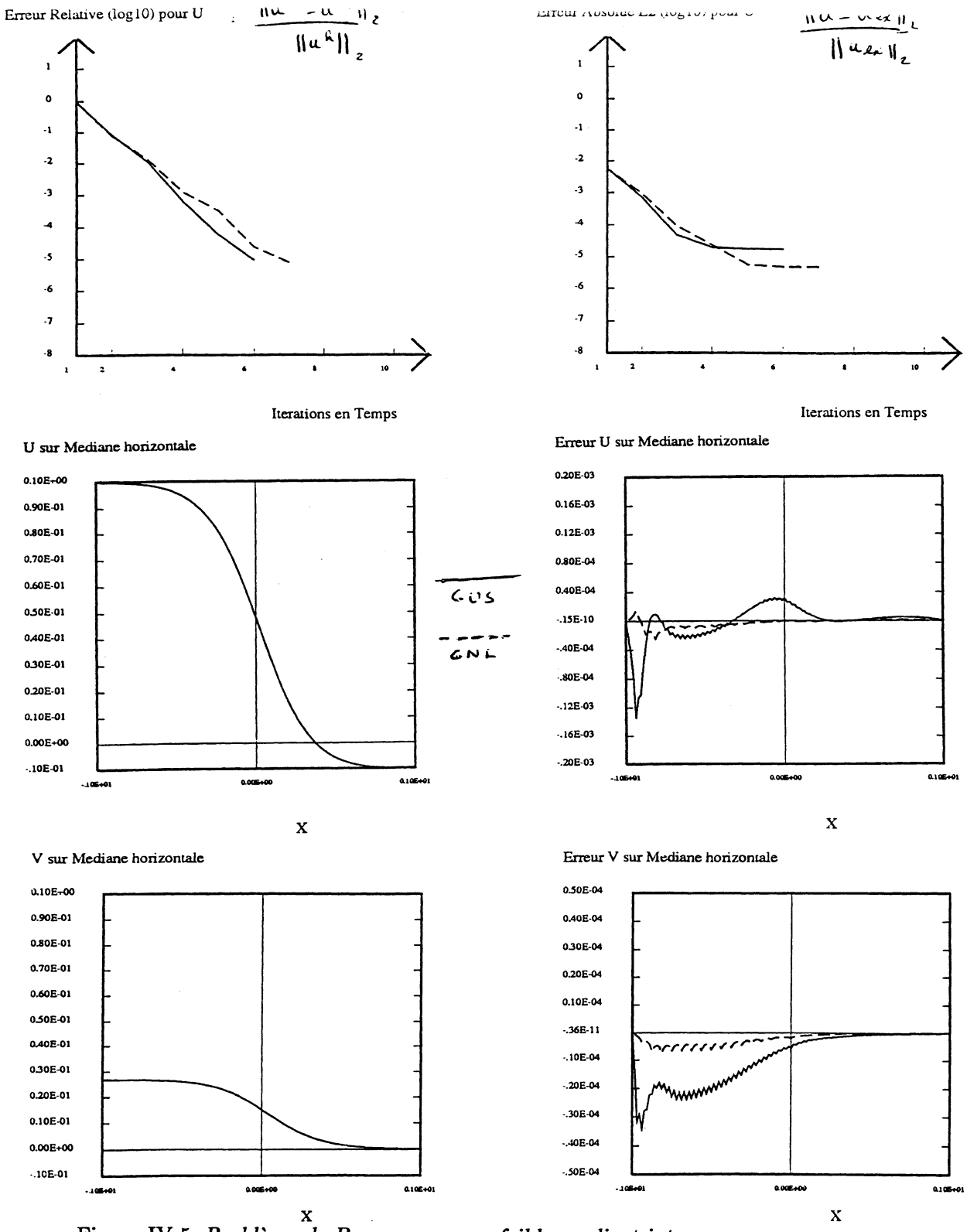


Figure IV.5: *Problème de Burgers avec un faible gradient interne.*
On note $U = u_1$; $V = u_2$; $X = x_1$; $Y = x_2$.
En haut : erreurs. A gauche : profils des composantes de la solution calculée. A droite : profils des différences avec la solution exacte.

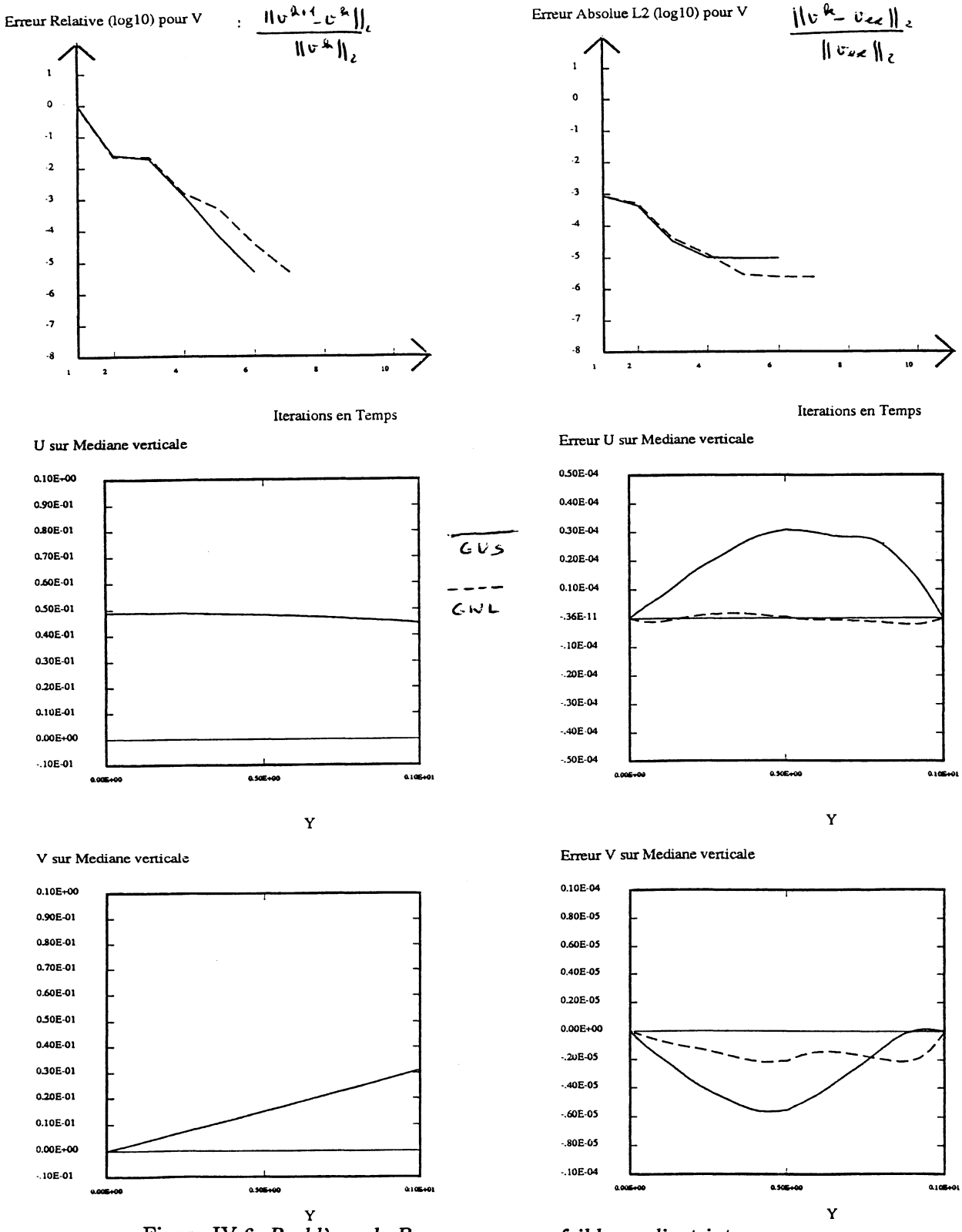
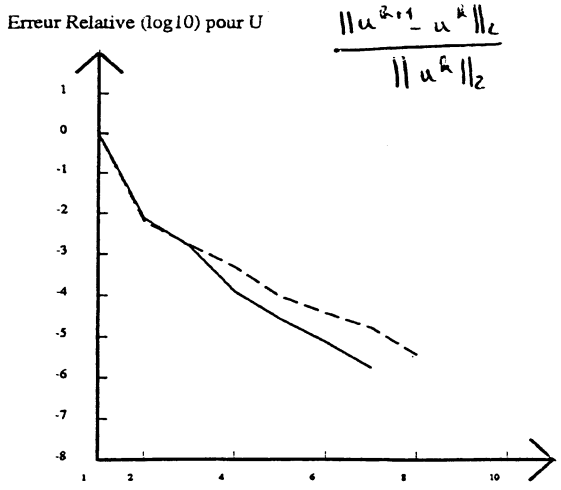
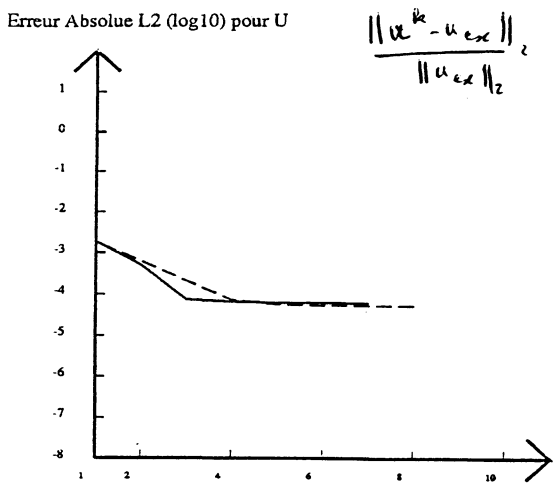
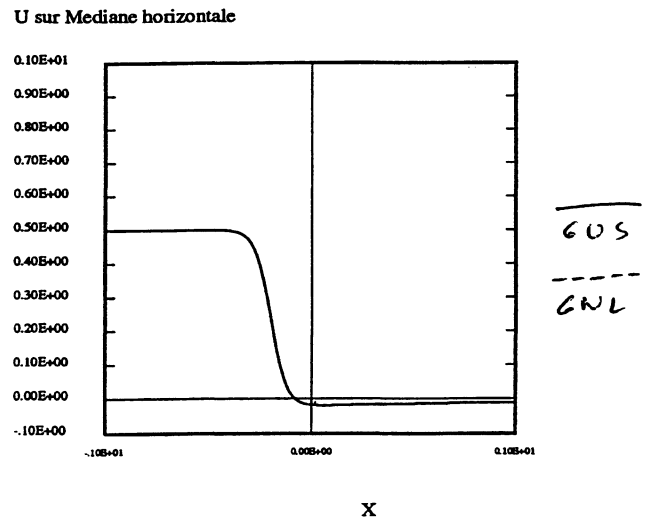
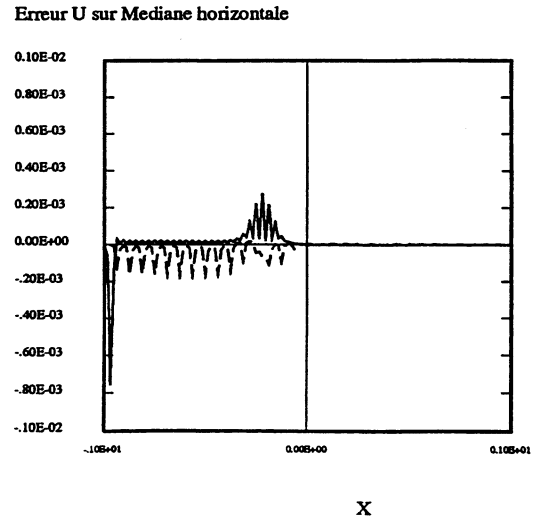


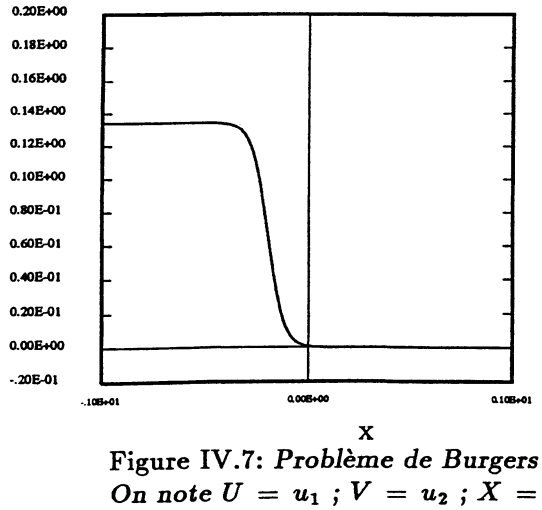
Figure IV.6: *Problème de Burgers avec un faible gradient interne.*
On note $U = u_1$; $V = u_2$; $X = x_1$; $Y = x_2$.
En haut : erreurs. A gauche : profils des composantes de la solution calculée. A droite : profils des différences avec la solution exacte.











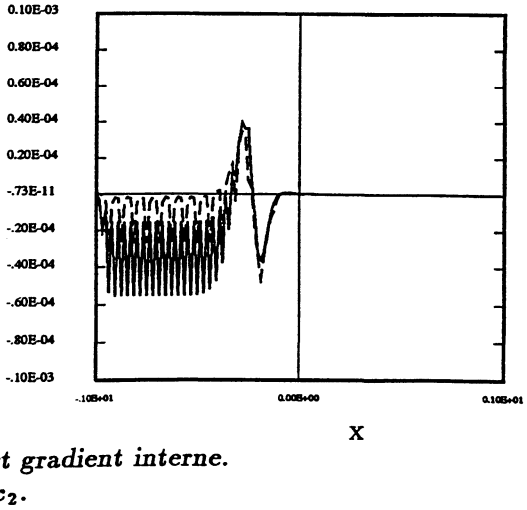


Figure IV.7: *Problème de Burgers avec un fort gradient interne.*
On note $U = u_1$; $V = u_2$; $X = x_1$; $Y = x_2$.
En haut : erreurs. A gauche : profils des composantes de la solution calculée. A droite : profils des différences avec la solution exacte.

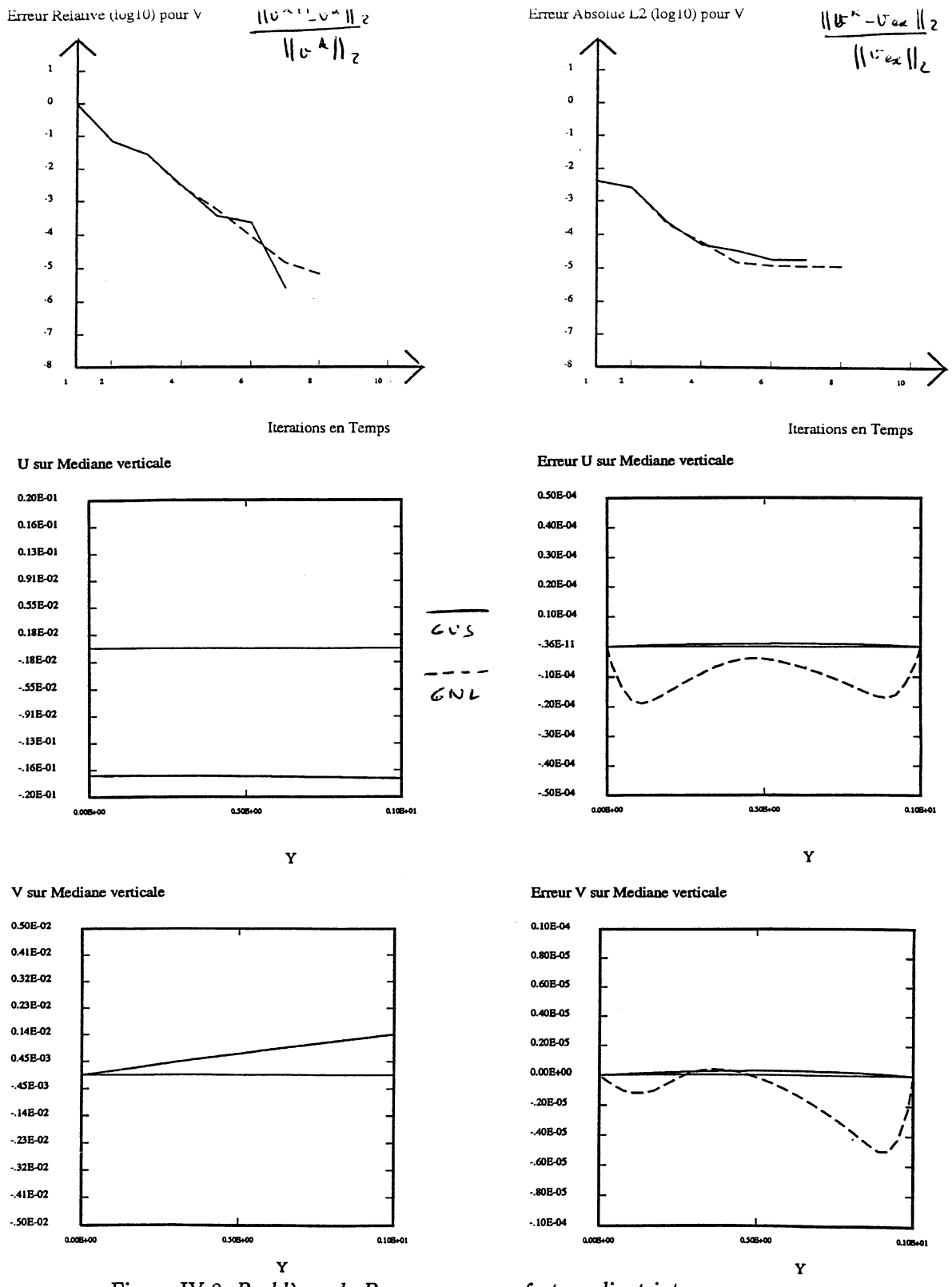


Figure IV.8: *Problème de Burgers avec un fort gradient interne.*
On note $U = u_1$; $V = u_2$; $X = x_1$; $Y = x_2$.
En haut : erreurs. A gauche : profils des composantes de la solution calculée. A droite : profils des différences avec la solution exacte.

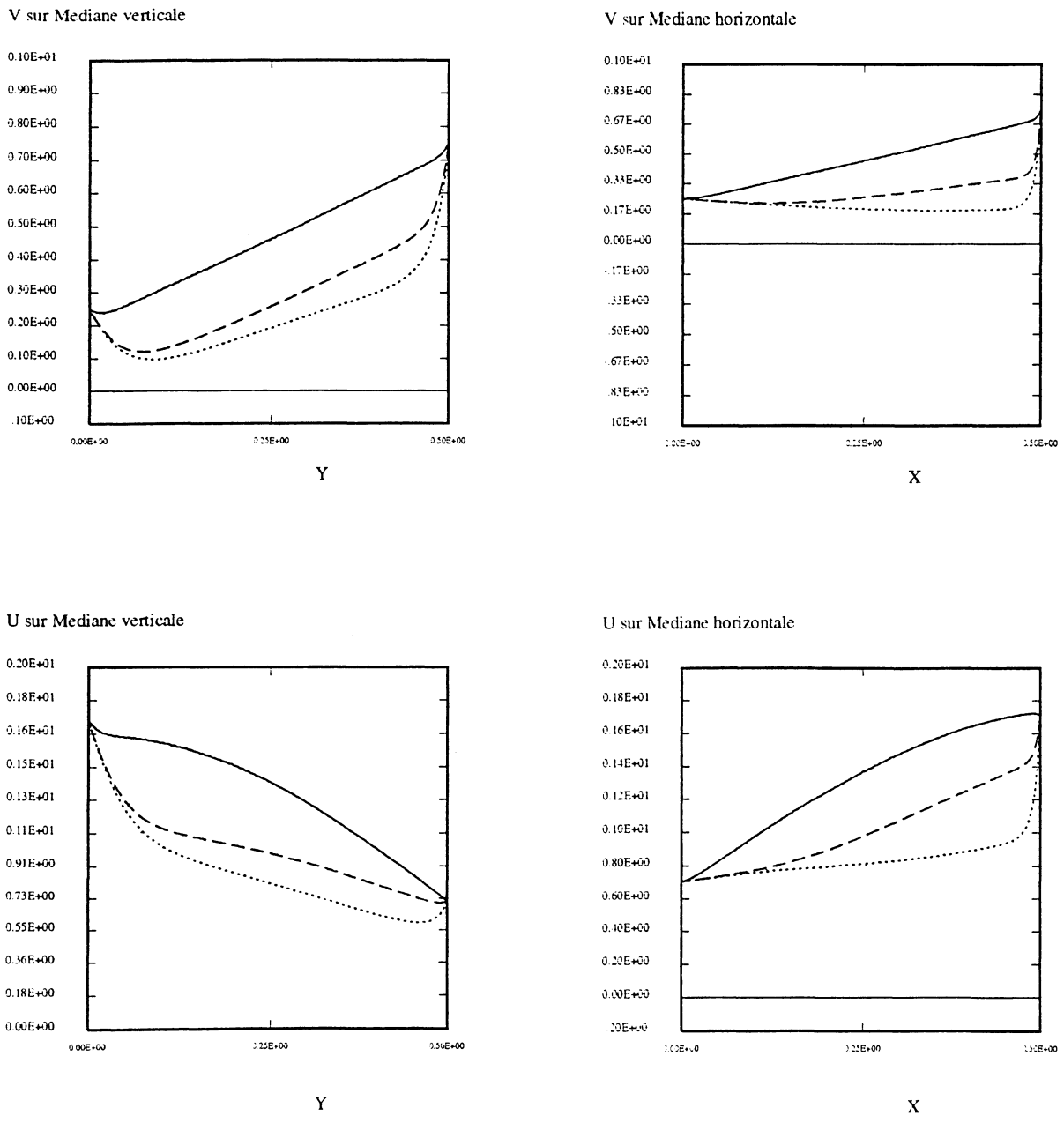
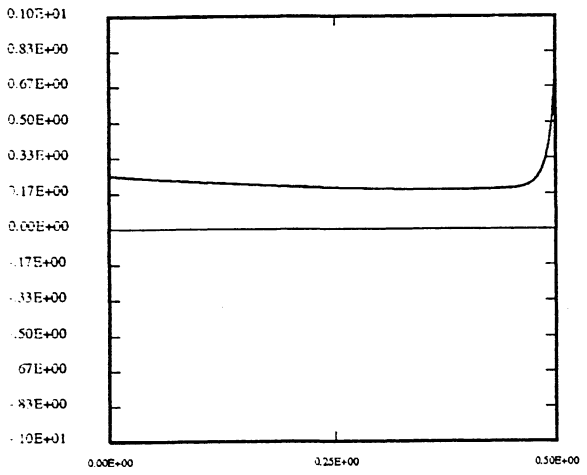
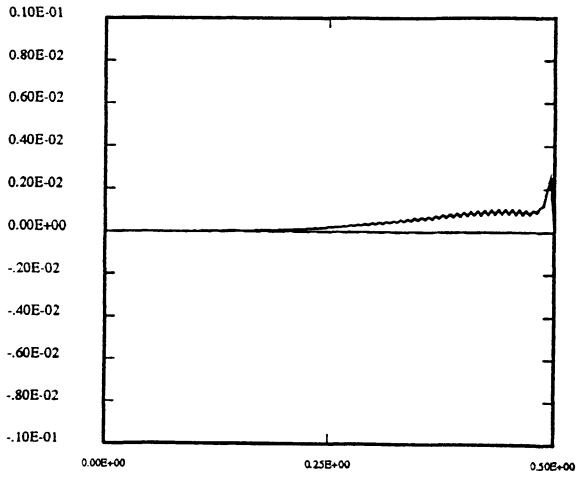


Figure IV.9: *Problème de Burgers par la méthode de Galerkin non linéaire GNL2.*
$Re = 100.$ $h = \frac{1}{256}$; $\Delta t = 5.10^{-3}$.
Profils sur la médiane verticale et la médiane horizontale de la solution.
$(\text{——}) : t = 0.02$; $(\text{- - - -}) : t = 0.2$; $(\text{.....}) : t = 0.5$.

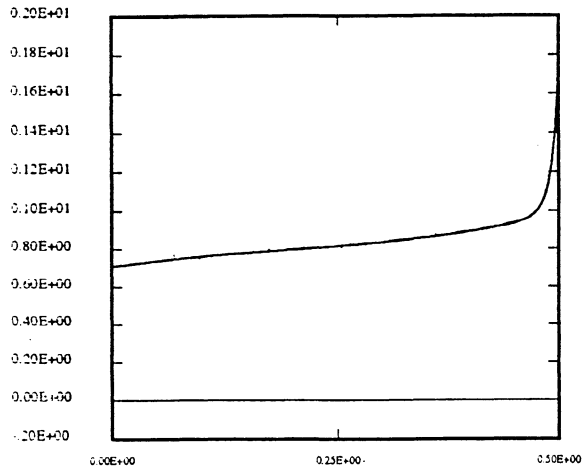
V sur Mediane horizontale



Erreur V sur Mediane horizontale



U sur Mediane horizontale



Erreur U sur Mediane horizontale

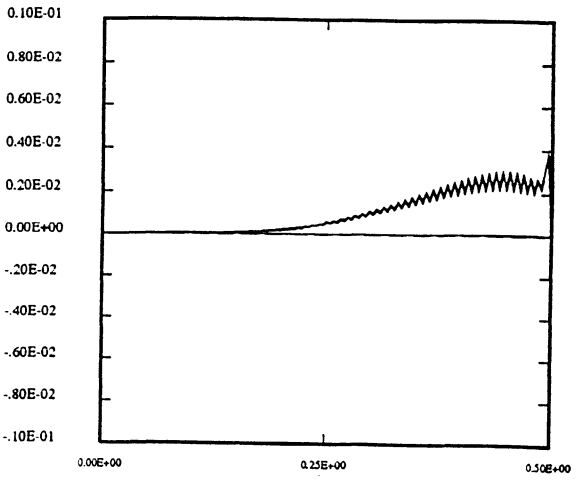
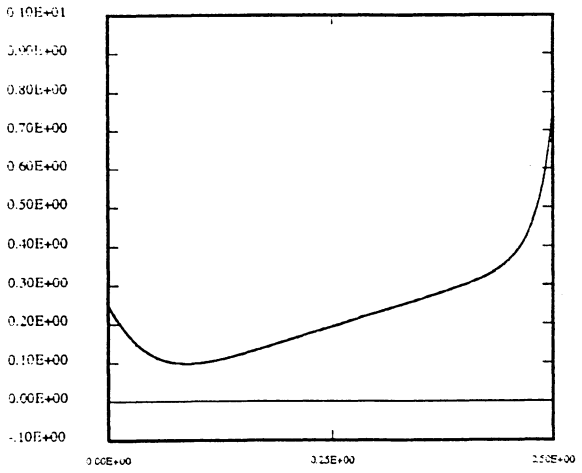
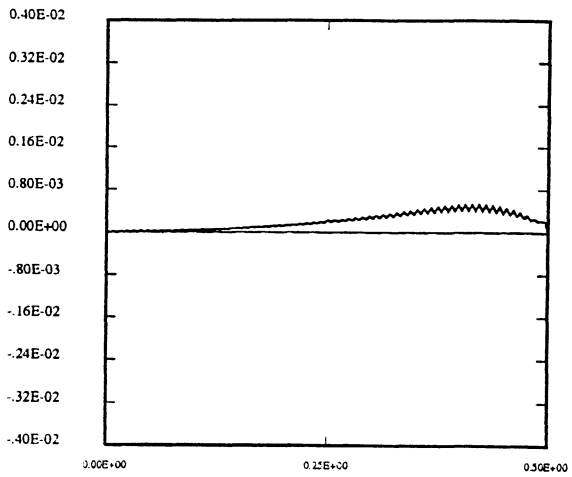


Figure IV.10: *Problème de Burgers.* $Re = 100.$ $h = \frac{1}{256}$; $\Delta t = 5.10^{-3}$.
Profils sur la médiane horizontale de la solution par GUS et GNL2 à $t = 0.5$.
Profils sur la médiane horizontale de la différence des 2 solutions à $t = 0.5$.

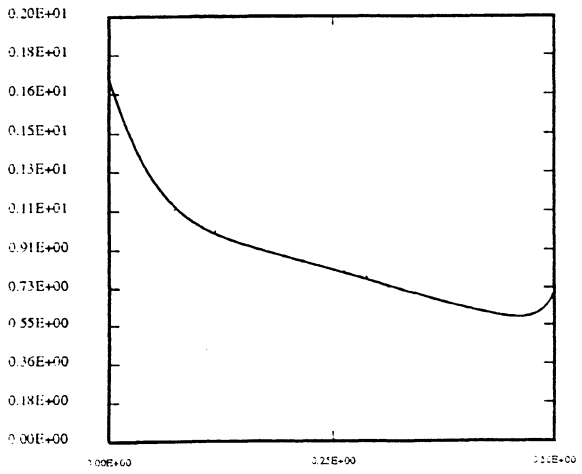
V sur Mediane verticale



Erreur V sur Mediane verticale



U sur Mediane verticale



Erreur U sur Mediane verticale

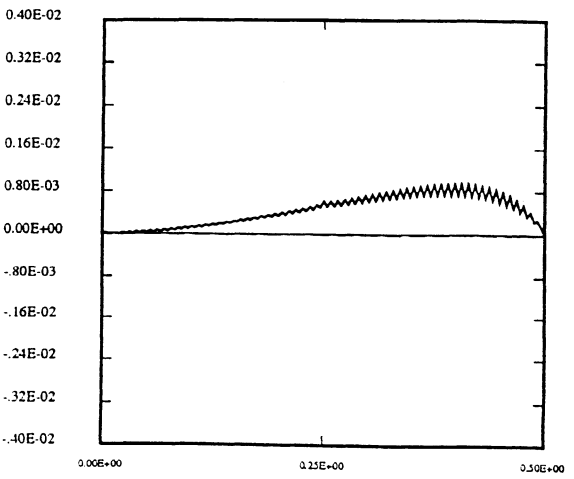


Figure IV.11: *Problème de Burgers.* $Re = 100.$ $h = \frac{1}{256}$; $\Delta t = 5.10^{-3}$.
Profils sur la médiane verticale de la solution par GUS et GNL2 à $t = 0.5$.
Profils sur la médiane verticale de la différence des 2 solutions à $t = 0.5$.

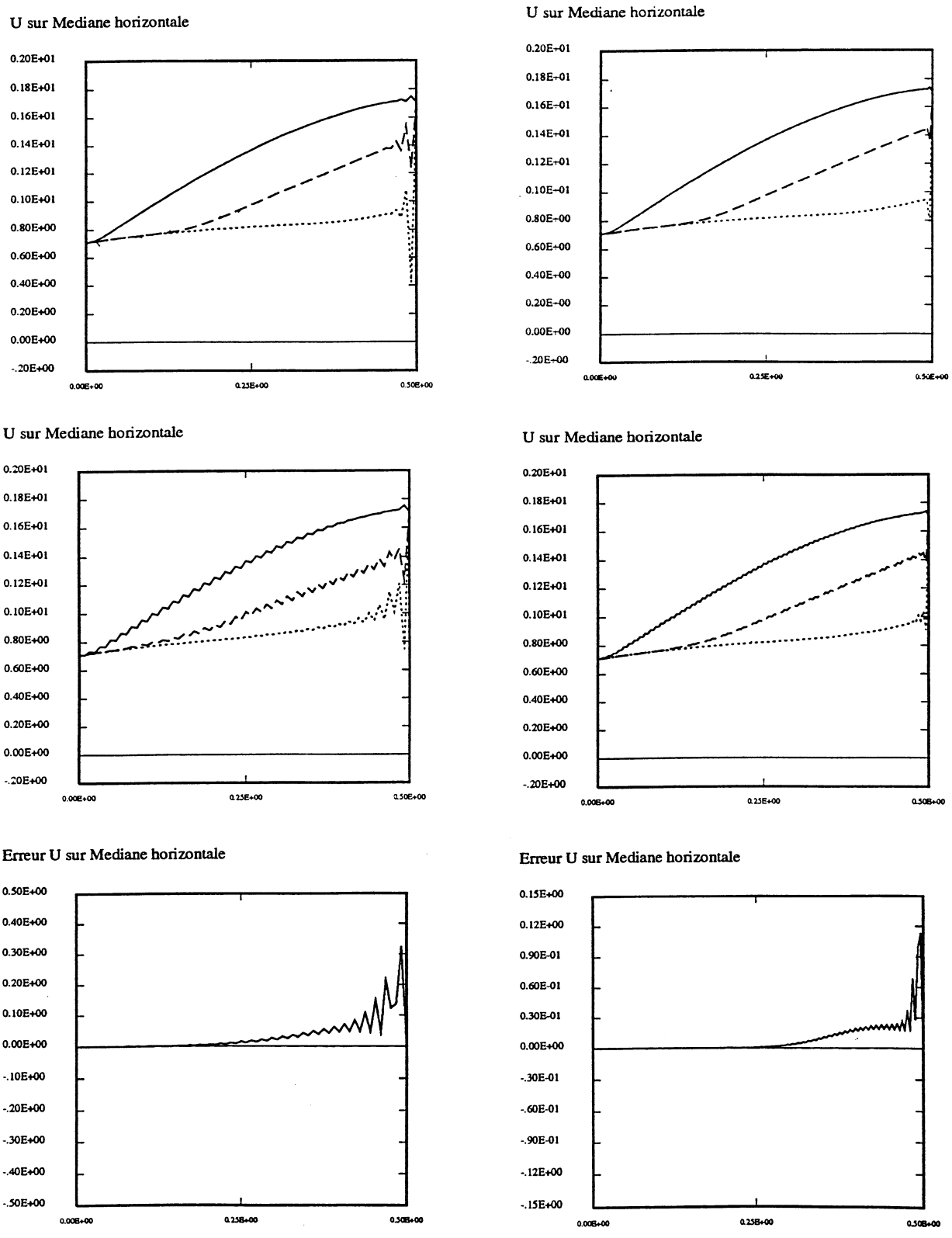


Figure IV.12: *Problème de Burgers.* $Re = 500$.
A gauche $h = \frac{1}{128}$ *et* $\Delta t = 1.10^{-2}$ *; A droite* $h = \frac{1}{256}$ *et* $\Delta t = 5.10^{-3}$.
En haut : méthode de Galerkin classique ;
Au milieu : méthode GNL2 ; En bas : différence à l'instant $t = 0.5$.

IV.3 Problème de Navier-Stokes.

IV.3.1 Description des algorithmes.

Les résultats concernant les méthodes de Galerkin non linéaires appliquées au probléme de Burgers permettent d'envisager l'implémentation de ces méthodes pour le problème de Navier-Stokes évolutif. On propose de décomposer le champ de vitesse et la pression en petites et grandes structures.

De façon identique à la section III.2 considérons la décomposition des espaces de discrétisation liée à la base hiérarchique :

$$\mathcal{V}_{0,h} = \mathcal{V}_{0,2h} \oplus \mathcal{W}_{0,h} \quad \text{et} \quad \mathcal{Q}_h = V_{4h} \oplus \mathcal{W}_{2h}$$

La méthode de Galerkin non linéaire consiste à chercher la solution approchée $(u_h, p_h) \in \mathcal{V}_{0,h} \times \mathcal{Q}_h$ du problème discrétisé $(N.S.)_h$ sous la forme :

$$\begin{array}{rcl} u_h & = & uy_h + uz_h \quad \text{avec } uy_h \in \mathcal{V}_{0,2h} \quad uz_h \in \mathcal{W}_{0,h} \\ p_h & = & py_h + pz_h \quad \text{avec } py_h \in V_{4h} \quad pz_h \in W_{2h} \end{array}$$

où uy_h, uz_h, py_h, pz_h sont définis par :

$$\left\{ \begin{array}{rcl} \dfrac{\partial(uy_h, vy_h)}{\partial t} & + & \\ \dfrac{1}{Re}(\nabla uy_h, \nabla vy_h) + (\nabla py_h, vy_h) & + & \\ ((uy_h.\nabla)uy_h + (uy_h.\nabla)uz_h + (uz_h.\nabla)uy_h, vy_h) & = & (f_h, vy_h) \\ & - & \dfrac{1}{Re}(\nabla uz_h, \nabla vy_h) - (\nabla pz_h, vy_h) \\ (\nabla.uy_h, qy_h) & = & -(\nabla.uz_h, qy_h) \\ \forall vy_h \in \mathcal{V}_{0,2h} \ \text{ et } \ \forall qy_h \in V_{4h} & & \end{array} \right. \tag{IV.14}$$

$$\left\{ \begin{array}{rcl} \dfrac{\partial(uz_h, vz_h)}{\partial t} & + & \\ + \dfrac{1}{Re}(\nabla uz_h, \nabla vz_h) + (\nabla pz_h, vz_h) & + & \\ ((uy_h.\nabla)uy_h, vz_h) & = & (f_h, vz_h) \\ & - & \dfrac{1}{Re}(\nabla uy_h, \nabla vz_h) - (\nabla py_h, vz_h) \\ (\nabla.uz_h, qz_h) & = & -(\nabla.uy_h, qz_h) \\ \forall vz_h \in \mathcal{W}_{0,h} \ \text{ et } \ \forall qz_h \in W_{2h} & & \end{array} \right. \tag{IV.15}$$

Le système (IV.14) où le terme d'évolution de uz_h est négligé, est interprété comme un problème en (uy_h, py_h) correspondant à la discrétisation de Galerkin classique sur la triangulation grossière T_{2h}.

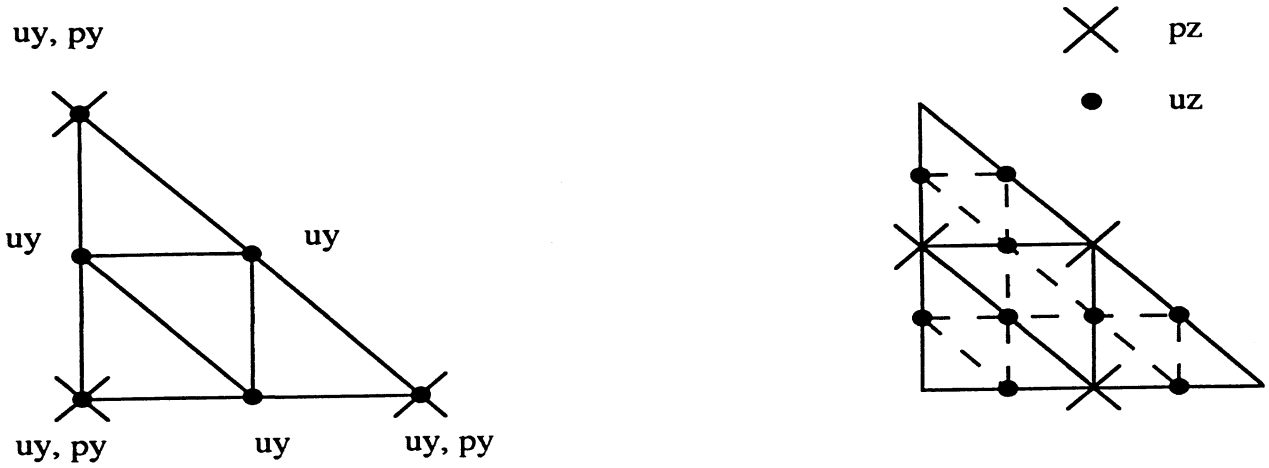


Figure IV.13: *Position des degrés de liberté pour les problèmes en* (uy_h, py_h), (uz_h, pz_h).

Le système (IV.15) est considéré comme un problème en (uz_h, pz_h), c'est-à-dire un problème de Stokes dont la discrétisation par éléments finis est non conforme et qui vérifie la condition inf-sup (communication d'Olivier Goubet et de Martine Marion). Les positions relatives des degrés de liberté sont représentées en figure IV.13 sur le triangle de référence.

Remarque IV.1 *L'étude des différentes structures au chapitre III pour Navier-Stokes montre que les termes* $\nabla.uy_h$ *et* $\nabla.uz_h$ *sont des termes très petits en norme* L_2. *Il semblait donc naturel d'écrire la condition de continuité pour le système (IV.14) sous la forme :*

$$(\nabla.uy_h, qy_h) = 0 \quad \forall qy_h \in \mathcal{V}_{4h}$$

condition vérifiée par uy_h *dans la discrétisation de Galerkin sur* T_{2h}.

Etant donné qu'aucun résultat théorique n'a été prouvé avec une telle approximation, sur suggestion de M. Marion il a été décidé de garder :

$$(\nabla.uy_h, qy_h) = -(\nabla.uz_h, qz_h) \quad \forall qy_h \in \mathcal{V}_{4h}$$

La discrétisation en temps associée (et notée **schéma GNL2**) est composée d'une méthode implicite pour le terme linéaire et d'une méthode explicite d'Adams-Bashforth pour le terme non linéaire.

(uy_h, py_h) est solution du problème de Stokes suivant :

$$\left\{\begin{array}{lll}
\frac{1}{\Delta t}(uy_h^{k+1} - uy_h^k, vy_h) & + & \\
\frac{1}{Re}((uy_h^{k+1} + uz_h^k, vy_h)) & + & \\
(\nabla(py_h^{k+1} + pz_h^k), vy_h) & = & (f_h^{k+1}, vy_h) \\
 & - & \frac{1}{Re}(\nabla uz_h^k, \nabla vy_h) - (\nabla pz_h^k, vy_h) \\
 & - & \frac{3}{2}((uy_h^k \ .\nabla)uy_h^k \ + (uy_h^k \ .\nabla)uz_h^k \ + (uz_h^k \ .\nabla)uy_h^k \ , vy_h) \\
 & + & \frac{1}{2}((uy_h^{k-1}.\nabla)uy_h^{k-1} + (uy_h^{k-1}.\nabla)uz_h^{k-1} + (uz_h^{k-1}.\nabla)uy_h^{k-1}, vy_h) \\
\forall vy_h \in \mathcal{V}_{0,2h} & & \\
 & & \\
(\nabla.uy_h^{k+1}, qy_h) & = & -(\nabla.uz_h^k, qy_h) \\
\forall qy_h \in V_{4h} & &
\end{array}\right.$$

(uy_h^{k+1}, py_h^{k+1}) sont une approximation de la solution sur T_{2h} que l'on corrige sur la grille fine T_h en déterminant (uz_h^{k+1}, pz_h^{k+1}) :

$$\left\{\begin{array}{lll}
\frac{1}{\Delta t}(uz_h^{k+1} - uz_h^k, vz_h) & + & \\
\frac{1}{Re}((uy_h^{k+1} + uz_h^{k+1}, vz_h)) & + & \\
(\nabla(py_h^{k+1} + pz_h^{k+1}), vz_h) & = & (f_h^{k+1}, vz_h) \\
 & - & \frac{1}{Re}(\nabla uy_h^{k+1}, \nabla vz_h) - (\nabla py_h^{k+1}, vz_h) \\
 & - & \frac{3}{2}((uy_h^k \ .\nabla)uy_h^k \ , vz_h) \\
 & + & \frac{1}{2}((uy_h^{k-1}.\nabla)uy_h^{k-1}, vz_h) \\
\forall vz_h \in \mathcal{W}_{0,h} & & \\
 & & \\
(\nabla.uz_h^{k+1}, qz_h) & = & -(\nabla.uy_h^{k+1}, qz_h) \\
\forall qy_h \in W_{2h} & &
\end{array}\right.$$

A chaque pas de temps, chacun de ces problèmes de Stokes est résolu par la méthode d'Uzawa décrite au chapitre II.

Remarque IV.2 *On note :*

- **GNL1** : *le schéma identique au schéma GNL2 où aucune partie du terme non linéaire n'est négligée,*
- **GNL3** : *le schéma qui néglige en plus l'évolution de uz_h dans l'équation (IV.15).*

Remarque IV.3 *Une des différences essentielles (aussi bien d'un point de vue théorique que d'un point de vue numérique) avec la discrétisation pseudo-spectrale*

est la non orthogonalité de uy_h et uz_h en norme H^1 en dimension 2. Il est à noter qu'en dimension 1, on a $((uy_h, uz_h)) = 0$. Goubet [3] propose la construction (à partir des ondelettes) d'un nouveau type d'éléments finis dont la décomposition en structures uy_h et uz_h est orthogonale.

IV.3.2 Les résultats numériques.

Les premières expériences numériques ont été réalisées dans la cavité entrainée avec un nombre de Reynolds égal à 100 et un pas de discrétisation h de l'ordre de 0.0156 correspondant à un maillage 65×65. On constate que

1. (cf. figure IV.14) si les schémas GNL3 et GUS convergent pour δt égal à 0.1, le schéma GNL2 diverge. Des instabilités se produisent donc si le maillage est trop grossier.

2. (cf. tableau IV.5) si les schémas convergent (GNL3 avec $\Delta t = 0.1$ et GNL2 avec $\Delta t = 0.01$) alors l'écoulement du fluide est correctement simulé. Le tableau IV.5 permet d'apprécier l'exactitude de la convergence et la position des tourbillons et des contres-tourbillons comparées avec la méthode de Galerkin classique (tableau II.4).

La figure IV.15 représente le champ de différence, multiplié par un facteur 100, entre les 2 champs de vitesse, le premier obtenue par la méthode classique et le second par GNL3. Cette figure révèle l'origine des instabilités des nouveaux schémas. La différence est en effet maximale au voisinage du bord supérieur i.e. dans la zone où les petites structures sont les plus grandes (cf. figure III.6), lieu où le gradient de vitesse est le plus élevé. Il est donc nécessaire pour réduire ces instabilités d'augmenter la précision du maillage (éventuellement localement).

Enfin lors du calcul de la solution de Stokes, les pressions py_h^{k+1} et pz_h^{k+1} sont de moyennes nulles : ce qui permet d'évaluer la pression sur le maillage T_{2h} exprimée dans la base nodale bien que py_h et pz_h soient déterminées en des points distincts du maillage.

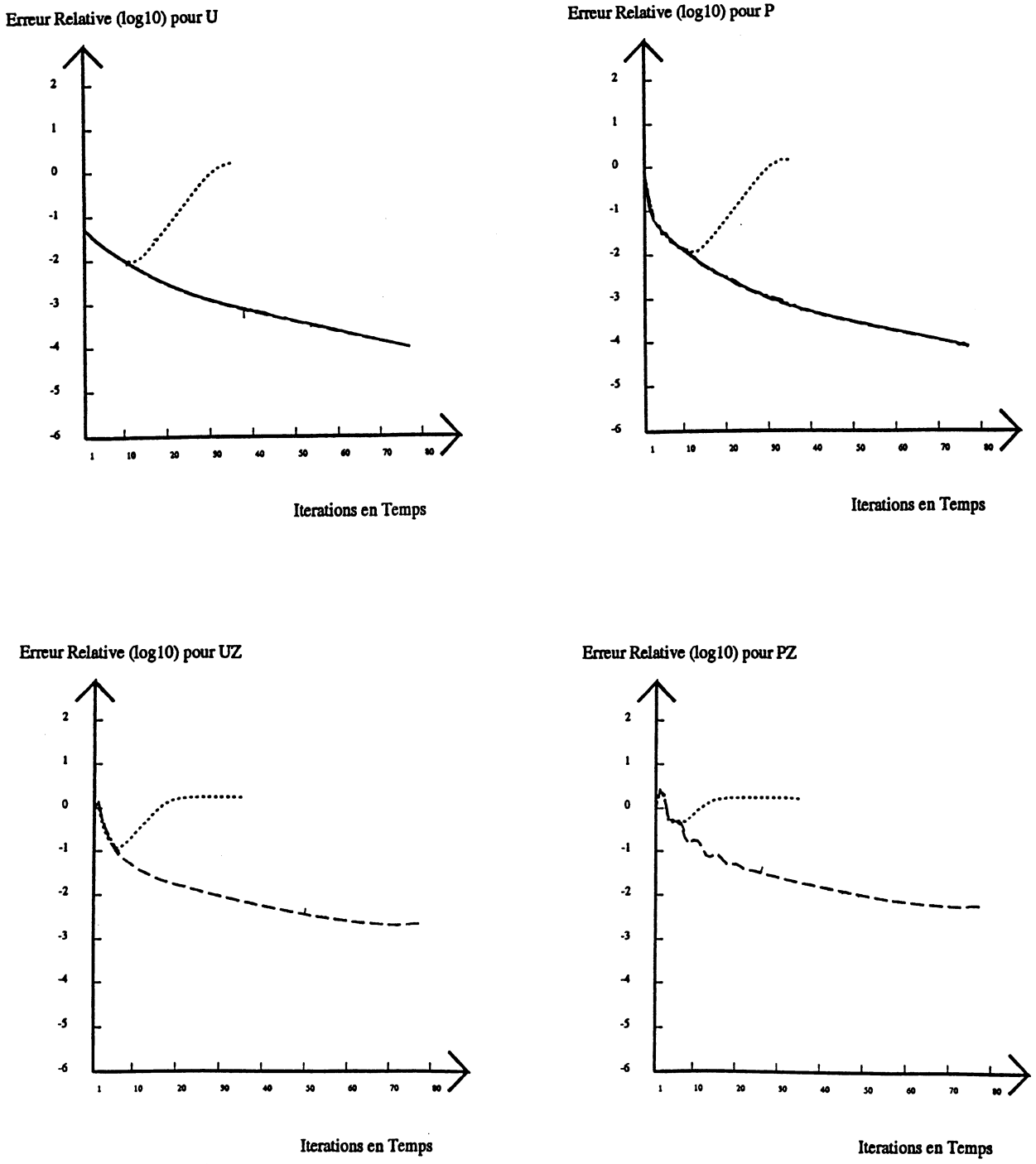


Figure IV.14: *Problème de Navier-Stokes. Re = 100 et* $h = \frac{1}{64}$. *Courbes de convergence de* u_h, p_h, uz_h, pz_h *pour les méthodes de Galerkin classique (——), les méthodes GNL3 (- - -), GNL2 (.....).*

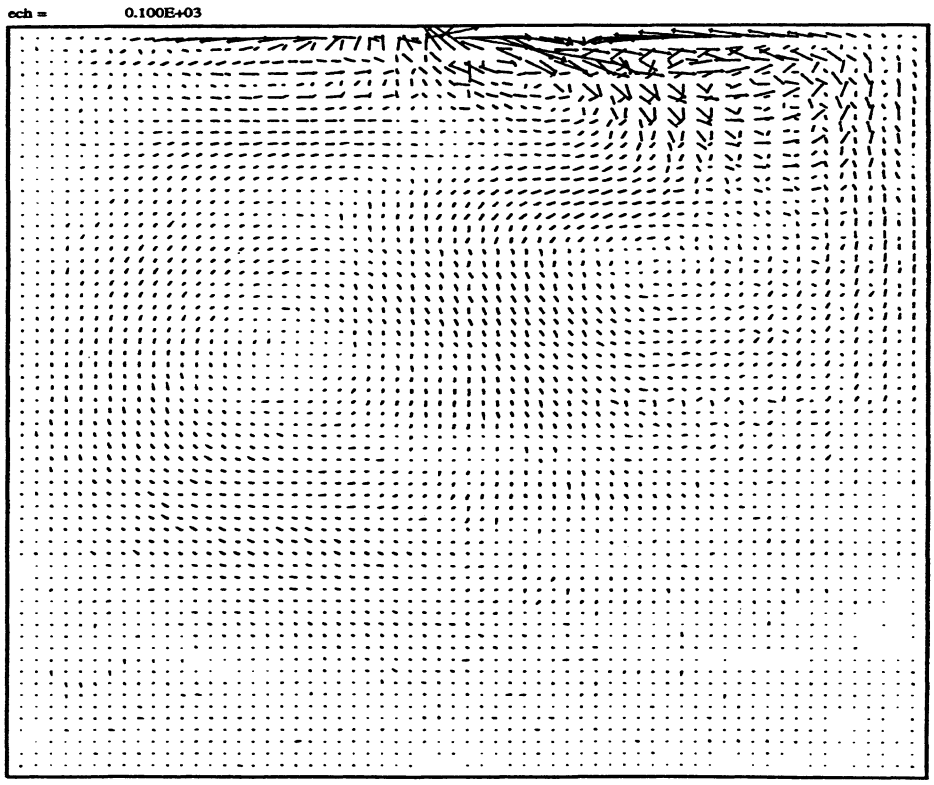


Figure IV.15: *Cavité entrainée B) avec $Re = 100$. Champ de différence ($\times 100$) entre la vitesse obtenue par la méthode de Galerkin classique et celle obtenue par GNL3.*

$Re = 100$		GNL3
PV	x	0.6094
	y	0.7500
	$\psi(x,y)$	-0.08374
	$\omega(x,y)$	2.932
BLV	x	0.03125
	y	0.03125
	$\psi(x,y)$	$0.909\ 10^{-6}$
	$\omega(x,y)$	0.00746
BRV	x	0.9531
	y	0.04687
	$\psi(x,y)$	$0.385\ 10^{-5}$
	$\omega(x,y)$	0.0179

Tableau IV.5: *Méthode de Galerkin non linéaire. Cavité entrainé B).*
PV = Tourbillon principal ;
BLV = Contre tourbillon gauche ;
BRV = Contre tourbillon droit

Bibliographie

[IV.1] C.A.J. Fletcher, *Generating exact solutions of the two-dimensional Burgers' equations*, Inter. Jour. for Num. Met. in fluids, 3, 213-216, 1983.

[IV.2] C.A.J. Fletcher, *A comparison of finite element and finite difference solutions of the one and two dimensional Burgers' equations.* J.C.P. 51, 159-188, 1983.

[IV.3] 0. Goubet, *Nonlinear Galerkin methods using hierarchical almost-orthogonal finite element bases*, soumis à J. of nonlinear Analysis, theory, methods and applications.

[IV.4] P.C. Jain and D.N. Holla, *Numerical solutions of coupled Burgers' equation* Int. J. Nonlinear Mechanics, 13, 213-222, 1978.

[IV.5] M. Marion and R. Temam, *Nonlinear Galerkin Methods : the finite element case*, Numer. Math. 57, 205-226, 1990.

CONCLUSION.

Au cours de ce travail, j'ai dans un premier temps regardé les difficultés dues à l'incompressibilité, la non linéarité et l'évolution en temps qui apparaissent dans le traitement numérique des équations régissant l'écoulement des fluides newtoniens, visqueux, incompressibles. Ainsi un code numérique qui permet d'obtenir des bases de donnés concernant la cavité entrainée, a été mis au point pour résoudre les équations de Navier-Stokes en dicrétisation par éléments finis mixtes, conformes P1-4P1.

Parallèlement, j'ai mis en oeuvre de nouveaux schémas numériques s'appuyant sur la théorie des systèmes dynamiques. Basés sur l'introduction de petites et grandes structures du champ de vitesse, ces schémas ont été développés dans le cadre d'une part d'une discrétisation en espace pseudo-spectrale et d'autre part d'une discrétisation par éléments finis.

Dans le premier cas, ces méthodes dites de Galerkin non linéaires pour lesquelles j'ai du introduire de nouveaux critères dynamiques de selection des modes actifs et passifs basés sur l'enstrophie et les termes non linéaires de couplage, étaient couplées avec un modèle de turbulence. Ces algorithmes conservent le comportement statistique turbulent de l'écoulement (en particulier la pente du spectre d'énergie) et ont permis des gains de temps non négligeables.

Cependant des limites apparaissent. Je montre que ces gains sont fonctions de la largeur de la zone de dissipation qui peut s'avèrer très etroite dans certains modèles de turbulence. En effet, la troncature déterminant les grandes et petites échelles doit se situer proche de cette zone et les critères doivent être en conséquence adaptés. Les méthodes de Galerkin non linéaires qui cherchent la solution sur des variétés inertielles approximatives sont donc très efficaces avec des simulations directes.

D'autres critères plus économiques, d'autres troncatures tenant compte du comportement intermittent de l'écoulement pour calculer les modes actifs et passifs, d'autres variétés inertielles approximatives ou d'autres schémas temporels de ces modes sont à envisager.

Dans le cadre d'une discrétisation par éléments finis, l'introduction de la base hiérarchique fait apparaître à l'inverse de la base nodale, de "petites" et "grandes" structures. Ces idées étant nouvelles, j'ai dans une première étape, étudié numériquement cette base à partir des donnés obtenus de façon classique pour la cavité entrainée régularisée. La présence de forts gradients (une caractéristique des écoulements du type Navier-Stokes) s'avère extrèmement contraignante. Les structures ne sont petites devant les grandes (condition importante pour éventuellement figer leur évolution) que si le pas de discrétisation est très faible. Il est donc nécessaire pour mettre en oeuvre les méthodes de Galerkin non linéaires de travailler avec un maillage très précis, éventuellement adaptatif en temps et en espace, et localement très fin.

La structure multigrille de la base hiérarchique a ensuite permis de mettre au point un préconditionneur du gradient conjugué pour résoudre les problèmes elliptiques.

Cette décomposition de la solution discrète appliquée à un problème linéaire a permis d'améliorer la condition de stabilité du schéma d'Euler explicite. Il serait souhaitable maintenant d'envisager un forçage dépendant du temps.

Pour les problèmes de Burgers généralisés et de Navier-Stokes, les méthodes de Galerkin non linéaires ont été développées avec des nombres de Reynolds de l'ordre de 100. Des résultats similaires à ceux des méthodes classiques ont été obtenus et des gains de temps d'exécution sont observés bien que le calcul des termes non linéaires ne soit pas optimisé. Cependant aux vues des résultats numériques, certaines questions restent sans réponse :

- ces nouveaux schémas sont-ils plus stables que les schémas de Galerkin classiques ? On observe que cela n'est pas le cas, mais Re n'est-il pas trop faible ou le maillage trop grossier ?

- des oscillations du champ de vitesse pour Burgers et de la vorticité pour Navier-Stokes apparaissent avec les méthodes de Galerkin non linéaires pour des nombres de Reynolds plus élevés. Ces oscillations sont-elles irréductibles aux schémas ou dépendent-elles de la précision du maillage dans les zones de forts gradients ? L'écoulement discrétisé sur la grille grossière est-il alors satisfaisant ?

- quel est le comportement de la pression ? A quoi correspondent les petites structures de la pression ? Ne peut-on pas simplifier la condition de continuité ?