

# Schéma de haute précision en maillage quadrangulaire structuré pour des écoulements compressibles non visqueux

Romuald Carpentier

► **To cite this version:**

Romuald Carpentier. Schéma de haute précision en maillage quadrangulaire structuré pour des écoulements compressibles non visqueux. [Rapport de recherche] RR-3909, INRIA. 2000, pp.63. <inria-00072744>

**HAL Id: inria-00072744**

**<https://hal.inria.fr/inria-00072744>**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*Schéma de haute précision en maillage  
quadrangulaire structuré pour des écoulements  
compressibles non visqueux.*

Romuald CARPENTIER

**N° 3909**

avril 2000

THÈME 4



*Rapport  
de recherche*



# Schéma de haute précision en maillage quadrangulaire structuré pour des écoulements compressibles non visqueux.

Romuald CARPENTIER\*

Thème 4 — Simulation et optimisation  
de systèmes complexes  
Projets SINUS

Rapport de recherche n° 3909 — avril 2000 — 63 pages

**Résumé :** Une famille de schémas décentrés aux volumes finis de haute précision est proposée pour les équations bidimensionnelles d'advection, Burgers et Euler. Le schéma ENO simplifié a son coût calcul optimisé par l'utilisation d'une seule molécule par volume. Après avoir établi le principe du calcul, un algorithme est proposé et une étude en coût est effectuée afin de définir l'ordre de précision optimal. Enfin, des vérifications sont réalisées sur différents modèles avec conditions de bords.

**Mots-clés :** Mécanique des fluides numérique, Euler, schémas ENO.

\*Email: [Romuald.Carpentier@sophia.inria.fr](mailto:Romuald.Carpentier@sophia.inria.fr)

# High accuracy inviscid scheme on quadrangular structured meshes for compressible flow without viscosity.

**Abstract:** A family of finite volume upwind schemes with high accuracy is developed for advection, Burgers, Euler bidimensional equations. By taking just one stencil per volume, the ENO scheme is simplified. After establishing the basis of the calculus, an precise algorithm and a study of its complexity are proposed. By this way, an optimal choice for the order of accuracy is identified. At last, some verifications on different test cases with boundary conditions are carried out.

**Key-words:** Numerical fluid mecanic, Euler, ENO schemes.

---

# Table des matières

<b>1</b>	<b>INTRODUCTION</b>	<b>2</b>
<b>2</b>	<b>METHODOLOGIE</b>	<b>3</b>
2.1	La construction polynômiale . . . . .	4
2.2	Calcul du flux convectif . . . . .	11
2.3	Le calcul du terme source . . . . .	14
2.4	L'avancée en temps . . . . .	19
<b>3</b>	<b>FAISABILITE</b>	<b>21</b>
3.1	Algorithme . . . . .	21
3.2	Estimation de l'espace mémoire . . . . .	25
3.3	Estimation du temps de calcul . . . . .	27
3.4	Vérifications . . . . .	27
3.5	Choix de précision . . . . .	40
<b>4</b>	<b>TRAITEMENT DES BORDS</b>	<b>47</b>
4.1	Dégénérescence des molécules . . . . .	47
4.2	Implémentations des conditions . . . . .	53
4.3	Vérifications . . . . .	53
<b>5</b>	<b>CAS TEST EULER</b>	<b>55</b>
5.1	Onde de densité . . . . .	55
5.2	Tube à choc . . . . .	57
5.3	Ecoulement stationnaire . . . . .	57
<b>6</b>	<b>CONCLUSIONS ET PERSPECTIVES</b>	<b>61</b>

# 1 INTRODUCTION

L'approximation numérique des équations de Navier Stokes en fluide compressible pour les applications industrielles comportant des couplages aéroacoustiques nécessite l'usage de schémas ayant les propriétés suivantes:

1. la conservation
2. la capture avec dissipation des singularités (ou petits détails)
3. un ordre de précision élevé
4. se prêter aux maillages arbitrairement non structurés.

Godunov a donné une assez bonne réponse aux points 1 et 2. Ceci est résumé par deux idées: les solveurs de Riemann et le décentrage par discontinuité. Pour le point 3, l'interpolation d'ordre élevé est le fondement de deux types de familles: Galerkin discontinu [1] ou les méthodes de reconstruction [2] – [3]. Pour un ordre de précision élevé, les méthodes de reconstruction ENO sont plus attractives du fait:

- d'un moindre coût du traitement des termes de dissipation,
- d'une capture plus fine des discontinuités entre deux cellules voisines,
- d'une extension élégante de la méthode M.U.S.C.L. permettant de réutiliser des logiciels existants.

De plus et pour répondre au point 4, la méthode de reconstruction de type ENO a été étendue aux maillages non structurés [4].

Cependant, ENO utilise pour chaque volume un nombre important de molécules; WENO [5] en est une simplification qui est encore coûteuse. De ce point de vu, la présente étude a pour but de proposer une simplification de la méthode de reconstruction applicable à des écoulements ne présentant pas de chocs (sinon des limiteurs sont introduits). Cette simplification doit être capable de donner une bonne prédiction des ondes acoustiques pour le modèle des équations de Navier Stokes. La simplification principale du modèle repose sur la construction d'une seule molécule par volume.

• **Objectifs du travail:** Nous souhaitons réaliser, à moindre coût, un schéma explicite bidimensionnel d'ordre  $D + 1$  ( $D \in \{1,2,3,4...\}$ ) en volumes finis préservant cet ordre:

- pour un modèle linéaire ou non linéaire s'écrivant:

$$\frac{\partial \mathcal{U}}{\partial t} + \frac{\partial F(\mathcal{U})}{\partial x} + \frac{\partial G(\mathcal{U})}{\partial y} = S(t,x,y) \quad (1)$$

où  $U(t,x,y)$  est une composante du vecteur inconnu  $\mathcal{U}$  de dimension  $u$ .

$F$  et  $G$  sont les termes convectifs

$S(t,x,y)$  est un terme source connu.

+ conditions de bords pour fermer le problème.

- sur un maillage structuré de quadrangles  $Q_0$  pas forcément réguliers ou/et orthogonaux.

• **Hypothèses:** Dans le cahier des charges de ce schéma, une hypothèse importante est faite afin d'amoindrir le coût de calcul: nous ne chercherons pas à capturer les chocs parce que l'application visée (Cas test de l'ONERA) n'en présente pas! En d'autres termes:

- Le maillage est fixe. Il n'évolue pas au cours du temps.
- Les molécules sont fixes. Elles n'évoluent pas au cours du temps.
- Chaque composante  $U(t,x,y)$  de  $\mathcal{U}$  est une fonction  $C^D$ , c'est à dire, pouvant se laisser approcher localement par un polynôme de degré  $D$  (formule de Taylor).
- $F(\mathcal{U})$  et  $G(\mathcal{U})$  sont des fonctions  $C^D$  par rapport à chaque composante  $U(t,x,y)$  de  $\mathcal{U}$ .

• **Plan du rapport:** Le premier paragraphe est consacré à la description de la méthode utilisée qui se déroule en 4 phases:

- la (re)construction polynômiale.
- le calcul du flux convectif.
- le calcul du terme source.
- l'avancée en temps.

Le second paragraphe traite de la mise en oeuvre informatique. Il se compose:

- d'un algorithme visant un moindre temps de calcul.
- d'une estimation de l'espace mémoire requis par ordre.
- d'une estimation du temps de calcul par ordre.
- d'une vérification des ordres atteints.
- d'un choix d'un ordre pour continuer l'étude.

Le troisième paragraphe est orienté sur le traitement des conditions de bords qui amène une prise en compte de:

- la dégénérescence des molécules.
- l'implémentation des conditions de bords.
- la préservation de l'ordre.

Le quatrième paragraphe est un déballage de cas tests sur le système des équations d'Euler.

Le premier est effectué dans le contexte des hypothèses choisies, les autres non...

- onde de densité.
- tube à choc.
- écoulement stationnaire.

Le dernier paragraphe traite des conclusions et des perspectives que ce type de schéma apporte ainsi que des remarques observées lors d'essais numériques et qui demanderaient d'être étudiées plus en détails.

## 2 METHODOLOGIE

Pour décrire la méthode, nous écrivons (1) pour  $u = 1$ , c'est à dire, dans le cas où  $\mathcal{U}$  n'a qu'une seule composante  $U(t,x,y)$ .

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} = S(t,x,y) \quad (2)$$



Nous écrivons (2) en intégrant l'équation en volumes finis sur chaque quadrangle  $Q_0$  constituant le maillage et nous utilisons la formule de Green pour transformer l'intégration des termes convectifs en un flux sur le bord  $\partial Q_0$  du quadrangle  $Q_0$ :

$$\int \int_{Q_0} \frac{\partial U}{\partial t} dx dy + \int_{\partial Q_0} \mathcal{F}(U, \vec{n}) ds = \int \int_{Q_0} S(t, x, y) dx dy \quad (3)$$

où  $\mathcal{F}(U, \vec{n}) = F(U) n_x + G(U) n_y$

avec  $\vec{n} = (n_x, n_y)$  est la normale locale unitaire extérieure au quadrangle  $Q_0$ .

Nous nous proposons de résoudre (3) non pas en recherchant la fonction  $U(t, x, y)$  au cours du temps  $t = n\Delta t$  noté alors  $U^n(x, y)$ , mais en évaluant à un ordre  $D + 1$  sa valeur moyenne sur chaque quadrangle  $Q_0$  que nous noterons et définirons comme suit:

$$\bar{U}_0^n = \frac{1}{\text{Aire}(Q_0)} \int \int_{Q_0} U^n(x, y) dx dy \quad (4)$$

Pour obtenir une telle précision sur  $\bar{U}_0^n$ , il est nécessaire, d'après (4), d'avoir sur  $U^n(x, y)$  et sur chacun des termes de (3) une évaluation à l'ordre  $D + 1$ . La démarche proposée est décrite dans les quatres étapes suivantes.

## 2.1 La construction polynômiale

Il est supposé dans les hypothèses qu'à tout temps  $t = n\Delta t$ ,  $U^n(x, y)$  peut être estimé sur chaque quadrangle  $Q_0$  comme un polynôme local  $P_0^n(x, y)$  de degré  $D$ .

$$\forall (x, y) \in Q_0, \quad U^n(x, y) = P_0^n(x, y) + O(\Delta^{D+1}) \quad (5)$$

où  $\Delta$  est le pas d'espace moyen du maillage.

Le polynôme  $P_0^n(x, y)$  est caractérisé par  $NINC + 1$  coefficients à déterminer; où:

$$NINC + 1 = \frac{(D + 1)(D + 2)}{2}. \quad (6)$$

Ne connaissant que les valeurs moyennes de  $U^n(x, y)$ , il est légitime de demander que le polynôme  $P_0^n(x, y)$  défini sur chaque quadrangle  $Q_0$  ait sa valeur moyenne  $\bar{P}_{0/0}^n$  égale à la valeur moyenne  $\bar{U}_0^n$  sur le quadrangle  $Q_0$ :

$$\bar{P}_{0/0}^n = \frac{1}{\text{Aire}(Q_0)} \int \int_{Q_0} P_0^n(x, y) dx dy = \bar{U}_0^n \quad (7)$$

L'équation (7) ne permet cependant que de déterminer un coefficient du polynôme  $P_0^n(x, y)$  sur les  $NINC + 1$  recherchés. Il faut donc trouver d'autres équations permettant de définir pleinement chaque polynôme  $P_0^n(x, y)$ . Pour ce faire, nous définissons autour de chaque quadrangle  $Q_0$ , une molécule ou stencil, constituée de  $NST_0 \geq NINC$  quadrangles proches de

$Q_0$  (FIG.1).

**Remarque:** Compte tenu du bord du domaine, le nombre de quadrangles constituant le stencil n'est pas le même pour tous les quadrangles  $Q_0$  d'où la nécessité de faire porter à  $NST$  l'indice 0 pour se rappeler ultérieurement que  $NST_0$  peut changer de valeur selon la situation géométrique du quadrangle  $Q_0$  considéré. Pour l'heure, nous considérons,  $Q_0$  comme un quadrangle loin des bords du domaine et nous rouvrirons cette parenthèse au paragraphe 4 lors du traitement spécifique des bords.

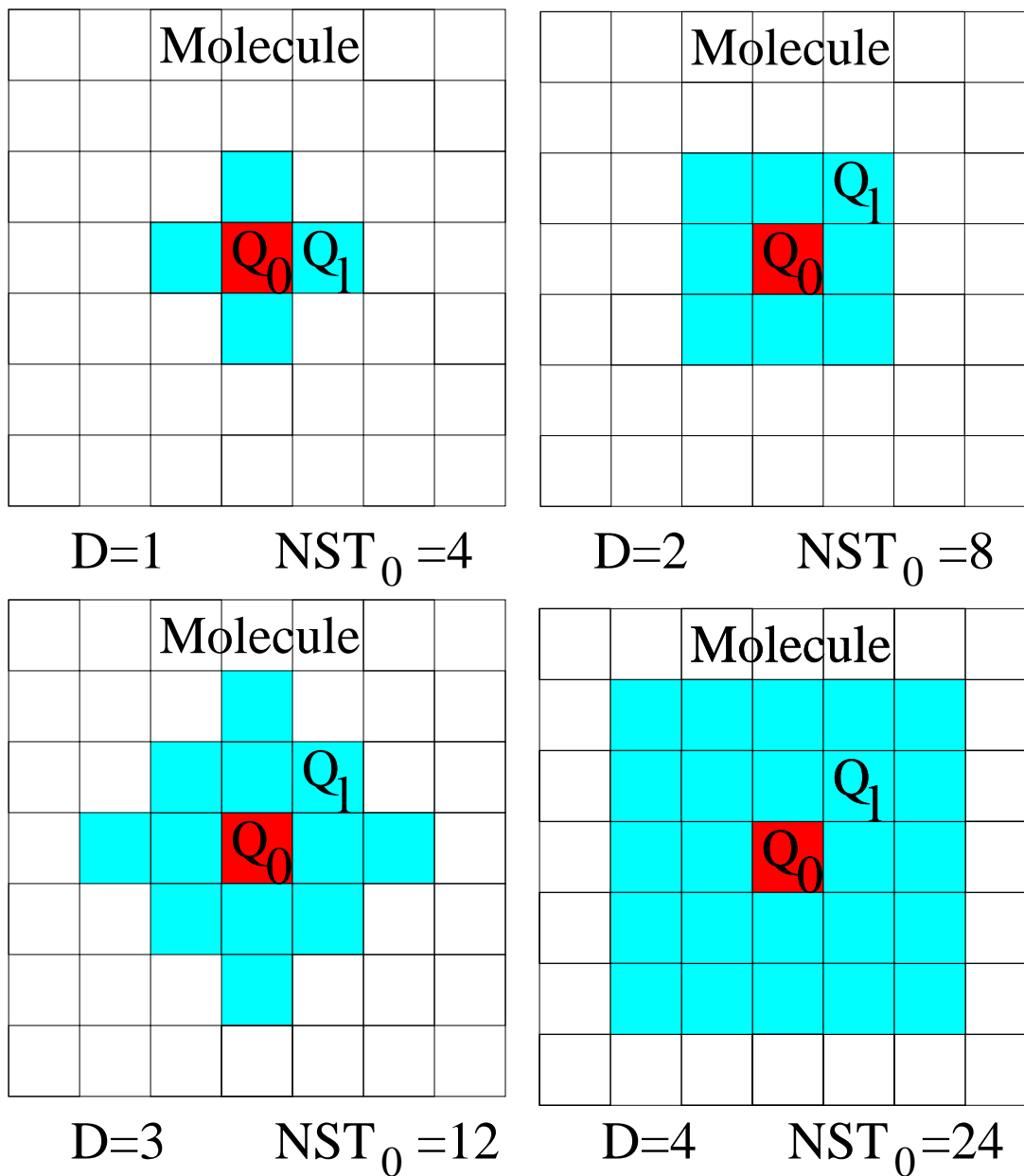


FIG.1: molécule de  $l=1$  à  $NST_0$  quadrangles  $Q_l$  pour une cellule centrale.

A l'instar de (7), nous demandons que la valeur moyenne  $\bar{P}_{0/l}^n$  du polynôme  $P_0^n(x,y)$  sur chaque quadrangle  $Q_l$  du stencil associé au quadrangle  $Q_0$  vérifie les relations:

$$\forall l \in [1, NST_0], \bar{P}_{0/l}^n = \frac{1}{\text{Aire}(Q_l)} \int \int_{Q_l} P_0^n(x,y) dx dy = \bar{U}_l \quad (8)$$

Avec les équations contenues dans (8) et parce que  $NST_0$  est dans la majorité des cas bien supérieur à  $NINC$ , nous avons trop d'équations pour déterminer les  $NINC + 1$  coefficients du polynôme  $P_0^n(x,y)$ . C'est pourquoi, plutôt que (8), est choisie une minimisation par les moindres carrés de la fonctionnelle  $H$  suivante:

$$H = \sum_{l=1}^{NST_0} \mu_{0/l} (\bar{P}_{0/l}^n - \bar{U}_l^n)^2 \quad (9)$$

où  $\mu_{0/l}$  est un poids positif pouvant permettre de tenir compte, par exemple:

- de la proximité du quadrangle  $Q_l$  par rapport au quadrangle  $Q_0$  en définissant  $\mu_{0/l}$  comme étant l'inverse de la distance entre le barycentre  $G_0$  du quadrangle  $Q_0$  au barycentre  $G_l$  du quadrangle  $Q_l$  (MU1);
- de l'importance des aires des cellules appartenant au stencil en définissant  $\mu_{0/l}$  comme étant l'aire du quadrangle  $Q_l$  (MU2);
- des deux items précédents en définissant  $\mu_{0/l}$  comme étant l'aire de  $Q_l$  sur la distance entre les barycentres de  $Q_0$  et  $Q_l$  (MU3);
- ou bien enfin, d'une même importance accordée à chaque quadrangle du stencil:  $\mu_{0/l} = 1$  (MU4).

Pour aller plus loin dans la résolution de (9), il est temps de définir plus explicitement le polynôme  $P_0^n(x,y)$  de degré  $D$ . Une expression, à première vue alambiquée, est la suivante:

$$P_0^n(x,y) = \bar{U}_0^n + \sum_{m=1}^{NINC} a_{0,m}^n \{b_{0,m}(x,y) - \bar{b}_{0/0,m}\} \quad (10)$$

avec:

$a_{0,m}$  les  $NINC$  inconnues réelles du polynôme  $P_0^n(x,y)$

$b_{0,m}(x,y)$  les  $m$  monômes de base constituant le polynôme:  $P_0^n(x,y)$

$$b_{0,m}(x,y) = (x - x_0)^{im(m)} (y - y_0)^{jm(m)}$$

$\bar{b}_{0/0,m}$  la valeur moyenne de  $b_{0,m}(x,y)$  sur le quadrangle  $Q_0$ :

$$\bar{b}_{0/0,m} = \frac{1}{\text{Aire}(Q_0)} \int \int_{Q_0} b_{0,m}(x,y) dx dy$$

où

$(x_0, y_0)$  sont les coordonnées du barycentre  $G_0$  du quadrangle  $Q_0$ ,

$im(m)$  et  $jm(m)$  sont respectivement les tableaux de correspondance

des puissances en  $x$  et en  $y$  donnés dans TAB.1 :

<i>ORDRE</i>	<i>DEGRE(D)</i>	$m$	$i = im(m)$	$j = jm(m)$
2	$D = 1$	1	1	0
		2	0	1
3	$D = 2$	3	2	0
		4	1	1
		5	0	2
4	$D = 3$	6	3	0
		7	2	1
		8	1	2
		9	0	3
5	$D = 4$	10	4	0
		11	3	1
		12	2	2
		13	1	3
		14	0	4
...	...	...	...	...
$D_0 + 1$	$D = D_0$	$\frac{D_0(D_0 + 1)}{2}$	$D_0$	0
		...	...	...
		<i>NINC</i>	0	$D_0$

TAB.1: tableau de correspondance des puissances  $im(m)$  et  $jm(m)$  repéré par le seul indice  $m$  (11)

En fait, l'écriture (10) de  $P_0(x,y)$  est astucieuse pour trois raisons:

- $P_0(x,y)$  est construit de manière à vérifier automatiquement (7), c'est-à-dire, de manière à préserver très fortement la conservation de la valeur moyenne  $\bar{U}_0^n$ .
- Le terme constant étant choisi, seules *NINC* inconnues sont à déterminer au moyen de (9). Il est clair que cette économie peut paraître dérisoire surtout lorsque  $NINC \gg 1$ , néanmoins, elle prend toute son importance tant au point de vue du temps de calcul que celui du stockage des matrices issues de (9) comme nous allons le voir.
- En nous plaçant dans le référentiel barycentrique de chaque quadrangle  $Q_0$ , les différences de monômes de puissances élevées en  $x$  et en  $y$  que nous serons amenés à faire, restent d'un même ordre de grandeur. Les calculs que nous ferons avec seront alors mieux évalués.

Maintenant que le polynôme  $P_0^n(x,y)$  est clairement défini dans (10), revenons à (9) en déterminant tout d'abord, la valeur moyenne  $\bar{P}_{0/l}^n$  sur un quadrangle  $Q_l$  appartenant au

stencil du quadrangle  $Q_0$ .

$$\bar{P}_{0/l}^n = \frac{1}{\text{Aire}(Q_l)} \int \int_{Q_l} P_0^n(x,y) dx dy \quad (12)$$

$$= \bar{U}_0^n + \sum_{m=1}^{NINC} a_{0,m}^n \{ \bar{b}_{0/l,m} - \bar{b}_{0/0,m} \}$$

$$\text{où } \bar{b}_{0/l,m} = \frac{1}{\text{Aire}(Q_l)} \int \int_{Q_l} (x - x_0)^{im(m)} (y - y_0)^{jm(m)} dx dy$$

En utilisant (12), (9) devient:

$$H = \sum_{l=1}^{NST_0} \mu_{0/l} \left( \bar{U}_0^n - \bar{U}_l^n + \sum_{m=1}^{NINC} a_{0,m}^n c_{0/l,m} \right)^2 \quad (13)$$

$$\text{où } c_{0/l,m} = \{ \bar{b}_{0/l,m} - \bar{b}_{0/0,m} \}$$

Une minimisation par les moindres carrés de  $H$  par rapport aux inconnues  $a_{m_0,0}^n$  s'écrit avec les  $NINC$  équations suivantes:

$$\forall m_0 \in \{1, 2, \dots, NINC\}, \frac{\partial H}{\partial a_{0,m_0}^n} = 0 \quad (14)$$

En utilisant l'expression (13) de  $H$ , introduite dans (14), il vient le système:

$$\forall m_0 \in \{1, 2, \dots, NINC\}, \sum_{m=1}^{NINC} a_{0,m}^n \left( \sum_{l=1}^{NST_0} \mu_{0/l} c_{0/l,m} c_{0/l,m_0} \right) = \sum_{l=1}^{NST_0} \mu_{0/l} c_{0/l,m_0} (\bar{U}_l^n - \bar{U}_0^n) \quad (15)$$

par lequel les coefficients  $a_{0,m}^n$  du polynôme  $P_0^n(x,y)$  associé au quadrangle  $Q_0$  peuvent être déterminés. Le système (15) peut s'écrire sous la forme matricielle suivante:

$$\mathcal{M}_0 \alpha_0^n = \beta_0^n \quad (16)$$

avec:

$$\mathcal{M}_0 = (M_{0,m,m_0}) = \left( \sum_{l=1}^{NST_0} \mu_{0/l} c_{0/l,m} c_{0/l,m_0} \right) \text{ une matrice } NINC \times NINC \text{ symétrique,}$$

$$\alpha_0^n = (a_{0,1}^n, \dots, a_{0,NINC}^n) \text{ le vecteur des coefficients inconnus du polynôme } P_0^n(x,y),$$

$$\beta_0^n = \left( \sum_{l=1}^{NST_0} \mu_{0/l} c_{0/l,1} (\bar{U}_l^n - \bar{U}_0^n), \dots, \sum_{l=1}^{NST_0} \mu_{0/l} c_{0/l,NINC} (\bar{U}_l^n - \bar{U}_0^n) \right)$$

le vecteur du second membre dépendant des valeurs moyennes.

Avant d'aller plus loin dans la détermination du vecteur  $\alpha_0^n$ , on se propose de donner ici la méthode utilisée pour déterminer les coefficients  $M_{0,m,m_0}$  de la matrice  $\mathcal{M}_0$  associée à chaque quadrangle  $Q_0$ . Ces coefficients dépendent des  $c_{0/l,m}$ , eux-mêmes dépendants des  $\bar{b}_{0/l,m}$  définis

dans (12).

Soient 4 points  $A1(x_1, y_1), A2(x_2, y_2), A3(x_3, y_3)$  et  $A4(x_4, y_4)$  définissant un quadrangle  $Q_i = (A1, A2, A3, A4)$  quelconque. Soit  $G_0(x_0, y_0)$  un point quelconque. On peut toujours décomposer le quadrangle  $Q_i$  en deux triangles (FIG.2):  $T1 = (A1, A2, A3)$  et  $T2 = (A1, A3, A4)$ .

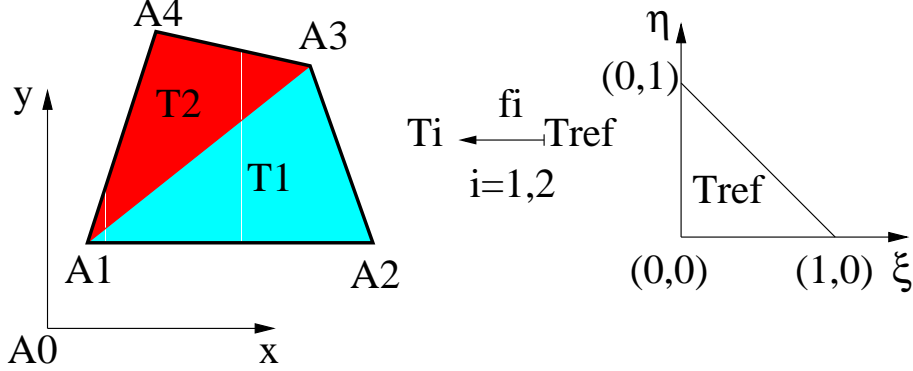


FIG.2: décomposition de  $Q_i$  en 2 triangles  $T_i$ . L'application  $f_i$  transforme  $T_{ref}$  en  $T_i$ .

Le calcul de la valeur moyenne  $\bar{b}_{0/l,m}$  du monôme  $b_{0,m}(x,y) = (x - x_0)^{im(m)}(y - y_0^{jm(m)})$  sur  $Q_i$  peut être décomposé de la manière suivante:

$$\begin{aligned} \bar{b}_{0/l,m} &= \frac{1}{Aire(Q_i)} \int \int_{Q_i} b_{0,m}(x,y) \, dx dy \\ &= \frac{1}{Aire(Q_i)} \sum_{i=1}^2 \left\{ \int \int_{T_i} b_{0,m}(x,y) \, dx dy \right\} \end{aligned} \quad (17)$$

Le triangle de référence  $T_{ref}$  peut être transformé en triangle  $T_i$  par l'application linéaire  $f_i : (\xi, \eta) \mapsto (x_i(\xi, \eta), y_i(\xi, \eta))$  suivante:

$$f1 : \begin{vmatrix} \xi \\ \eta \end{vmatrix} \mapsto \begin{vmatrix} x1(\xi, \eta) \\ y1(\xi, \eta) \end{vmatrix} = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix} \begin{vmatrix} \xi \\ \eta \end{vmatrix} + \begin{vmatrix} x_1 - x_0 \\ y_1 - y_0 \end{vmatrix} \quad (18)$$

$$f2 : \begin{vmatrix} \xi \\ \eta \end{vmatrix} \mapsto \begin{vmatrix} x2(\xi, \eta) \\ y2(\xi, \eta) \end{vmatrix} = \begin{pmatrix} x_3 - x_1 & x_4 - x_1 \\ y_3 - y_1 & y_4 - y_1 \end{pmatrix} \begin{vmatrix} \xi \\ \eta \end{vmatrix} + \begin{vmatrix} x_1 - x_0 \\ y_1 - y_0 \end{vmatrix} \quad (19)$$

En effectuant un tel changement de variables dans (17) et en remarquant que:

$$\forall i \in \{1,2\}, \det(f_i) = 2 \, Aire(T_i)$$

alors, est obtenu:

$$\bar{b}_{0/l,m} = \frac{2}{Aire(Q_i)} \sum_{i=1}^2 Aire(T_i) \int \int_{T_{ref}} b_{0,m}(x_i(\xi, \eta), y_i(\xi, \eta)) \, d\xi d\eta \quad (20)$$

Quel que soit le monôme  $b_{0,m}(x,y)$ , il est d'un degré en  $x$  ou en  $y$  inférieur ou égal à  $D$ . Le changement de variable s'étant effectué de manière linéaire, le monôme  $b_{0,m}(x,y)$  exprimé en  $\xi$  et  $\eta$  est d'un même degré. Pour calculer (20) d'une manière générale et automatique,

l'intégration est réalisée au moyen de *NGS* points de Gauss de coordonnées  $(\xi_g, \eta_g)$  accompagnés de leur poids réel  $p_g$  respectif, permettant l'intégration exacte du monôme considéré, c'est-à-dire de degré  $D$ , comme suit:

$$\bar{b}_{0/l,m} = \frac{1}{\text{Aire}(Q_l)} \sum_{i=1}^2 \text{Aire}(T_i) \left\{ \sum_{g=1}^{NGS} p_g b_{0,m}(x_i(\xi_g, \eta_g), y_i(\xi_g, \eta_g)) d\xi d\eta \right\} \quad (21)$$

où:

ORDRE	DEGRE(D)	NGS	$\xi_g$	$\eta_g$	$p_g$																		
2	1	1	1/3	1/3	1																		
3	2	3	1/2 0 1/2	1/2 1/2 0	1/3 1/3 1/3																		
4	3	4	1/3 1/5 3/5 1/5	1/3 1/5 1/5 3/5	-27/48 25/48 25/48 25/48																		
5	4	6	$a = 0.445948490915965$ $pa = 0.111690794839005$ <table style="display: inline-table; border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 0 5px;"><math>a</math></td> <td style="padding: 0 5px;"><math>a</math></td> <td style="padding: 0 5px;"><math>2pa</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 0 5px;"><math>1 - 2a</math></td> <td style="padding: 0 5px;"><math>a</math></td> <td style="padding: 0 5px;"><math>2pa</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 0 5px;"><math>a</math></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 0 5px;"><math>1 - 2a</math></td> <td style="padding: 0 5px;"><math>2pa</math></td> </tr> </table> $b = 0.091576213509771$ $pb = 0.054975871827661$ <table style="display: inline-table; border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 0 5px;"><math>b</math></td> <td style="padding: 0 5px;"><math>b</math></td> <td style="padding: 0 5px;"><math>2pb</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 0 5px;"><math>1 - 2b</math></td> <td style="padding: 0 5px;"><math>b</math></td> <td style="padding: 0 5px;"><math>2pb</math></td> </tr> <tr> <td style="border-right: 1px solid black; padding: 0 5px;"><math>b</math></td> <td style="border-left: 1px solid black; border-right: 1px solid black; padding: 0 5px;"><math>1 - 2b</math></td> <td style="padding: 0 5px;"><math>2pb</math></td> </tr> </table>			$a$	$a$	$2pa$	$1 - 2a$	$a$	$2pa$	$a$	$1 - 2a$	$2pa$	$b$	$b$	$2pb$	$1 - 2b$	$b$	$2pb$	$b$	$1 - 2b$	$2pb$
$a$	$a$	$2pa$																					
$1 - 2a$	$a$	$2pa$																					
$a$	$1 - 2a$	$2pa$																					
$b$	$b$	$2pb$																					
$1 - 2b$	$b$	$2pb$																					
$b$	$1 - 2b$	$2pb$																					

TAB.2: coordonnées et poids des points de Gauss dans le triangle de référence. (22)

Dans (21), il faut bien comprendre que  $(x_i(\xi_g, \eta_g), y_i(\xi_g, \eta_g))$  ne sont rien d'autre que les coordonnées d'un point de Gauss dans le triangle  $T_i$  du quadrangle  $Q_l$  que nous écrivons

$(x_{l,i,g}, y_{l,i,g})$  afin de soulager l'écriture des formules à venir. Avec cette notation, explicitons (21):

$$\bar{b}_{0/l,m} = \frac{1}{\text{Aire}(Q_l)} \sum_{i=1}^2 \text{Aire}(Ti) \left\{ \sum_{g=1}^{NGS} p_g (x_{l,i,g} - x_0)^{im(m)} (y_{l,i,g} - y_0)^{jm(m)} \right\} \quad (23)$$

L'expression (23) est valable pour tout  $m$  et pour tout quadrangle  $Q_l$  du stencil y compris l'extension  $l = 0$  c'est à dire pour le quadrangle  $Q_0$  lui même donnant alors:

$$\bar{b}_{0/0,m} = \frac{1}{\text{Aire}(Q_0)} \sum_{i=1}^2 \text{Aire}(Ti) \left\{ \sum_{g=1}^{NGS} p_g (x_{0,i,g} - x_0)^{im(m)} (y_{0,i,g} - y_0)^{jm(m)} \right\} \quad (24)$$

On constate alors que  $\bar{b}_{0/l,m}$  dépend uniquement de la géométrie du maillage. Si le maillage n'évolue pas au cours du temps, il suffit d'en faire son évaluation une seule et unique fois. De fait,  $c_{0/l,m}$  donné par (13) est également indépendant du temps et donc, si le stencil associé à  $Q_0$  n'évolue pas au cours du temps,  $M_{0,m,m_0}$  donné par (16) aussi.

Dès lors, il devient très intéressant de ne pas résoudre le système (16) à chaque itération, mais de stocker la matrice inverse symétrique  $\mathcal{M}_0^{-1}$  pour chaque quadrangle  $Q_0$  afin d'obtenir plus rapidement le vecteur  $\alpha_0^n$  des coefficients du polynôme  $P_0^n(x,y)$  car, seul, le vecteur  $\beta_0^n$ , défini dans (16), devra être réactualisé à chaque pas de temps.

$$\alpha_0^n = \mathcal{M}_0^{-1} \beta_0^n \quad (25)$$

En conséquence, initialement, la construction de la matrice  $\mathcal{M}_0$  demande beaucoup de calculs en pré-traitement mais une fois  $\mathcal{M}_0^{-1}$  déterminée et stockée [6], nous appliquons (25), ce qui représente une énorme économie en temps de calcul dans la phase itérative du schéma explicite. Nous appellerons l'application de (25), la phase de **reconstruction polynômiale**.

**Remarque:** Sur le plan du stockage informatique, ce n'est pas la matrice inverse  $\mathcal{M}_0^{-1}$  qui est la plus coûteuse en espace mémoire puisqu'elle est symétrique et que, par conséquent, elle est peut se mémoriser en seulement:

$$\frac{NINC(NINC + 1)}{2} \text{ réels plutôt que } NINC^2 \text{ réels par quadrangle;}$$

mais le stockage des  $c_{0/l,m}$ , utiles à la réactualisation de  $\beta_0^n$ , qui sont au nombre de:

$$NST_0 \times NINC \geq NINC^2 \text{ réels par quadrangle!}$$

## 2.2 Calcul du flux convectif

La circulation de  $\mathcal{F}(U, \vec{n})$  autour du quadrangle  $Q_0$  donnée dans (3) est décomposée en la sommation de 4 intégrales portant sur les 4 interfaces droites que  $Q_0$  réalise avec les 4



quadrangles  $Q_I$  voisins "forts".

$$\int_{\partial Q_0} \mathcal{F}(U, \vec{n}) ds = \sum_{I=1}^4 \int_{Q_0 \cap Q_I} \mathcal{F}(U, \vec{n}) ds \quad (26)$$

Pour aller plus loin dans l'expression (26), nous considérons que  $\mathcal{F}(U, \vec{n})$  s'exprime sur chaque quadrangle  $Q_0$  comme un polynôme  $\mathcal{P}_0(x, y)$  de degré quelconque mais dont l'erreur est toujours en  $O(\Delta^{D+1})$ .

$$\mathcal{F}(U^n(x, y), \vec{n}) = \mathcal{F}(P_0^n(x, y), \vec{n}) = \mathcal{P}_0^n(x, y) + O(\Delta^{D+1}) \quad (27)$$

L'intégration sur une interface commune à deux quadrangles  $Q_0$  et  $Q_I$  peut alors encore s'effectuer au moyen de:

$$NGL \geq \frac{D+1}{2} \quad (28)$$

points de Gauss préservant cet ordre mais cette fois-ci sur un segment  $[A1, A2]$  et non plus une surface (FIG.3).

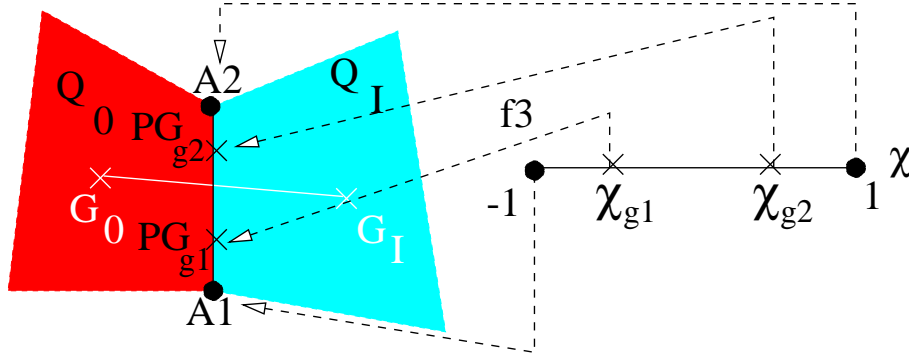


FIG.3: L'application  $f3$  transforme le segment  $[-1, 1]$  de référence en l'interface  $[A1, A2]$ .

Le segment de référence  $[-1, 1]$  peut être transformé en segment  $[A1, A2]$  au moyen de l'application linéaire  $f3: (\chi) \mapsto (x3(\chi), y3(\chi))$  suivante:

$$f3 : \chi \mapsto \begin{cases} x3(\chi) \\ y3(\chi) \end{cases} = \frac{1}{2} \begin{cases} (x_2 - x_1)\chi + (x_2 + x_1) \\ (y_2 - y_1)\chi + (y_2 + y_1) \end{cases} \quad (29)$$

Ainsi, le flux convectif peut s'exprimer sur l'interface  $[A1, A2]$  comme:

$$\int_{[A1, A2]} \mathcal{P}_0^n(x, y) ds = \frac{\|\vec{N}\|}{2} \int_{-1}^1 \mathcal{P}_0^n(x3(\chi), y3(\chi)) d\chi \quad (30)$$

où:  $\|\vec{N}\|$  représente la norme du vecteur normale intégré correspondant à la longueur de l'interface  $[A1, A2]$

On utilise les points de Gauss (TAB.3) pour exprimer (30):

$$\int_{[A1,A2]} \mathcal{P}_0^n(x,y) ds = \|\vec{N}\| \sum_{g=1}^{NGL} q_g \mathcal{P}_0^n(x3(\chi_g), y3(\chi_g)) \quad (31)$$

où  $\chi_g$  est la coordonnée d'un point de Gauss et  $q_g$  son poids

ORDRE	DEGRE(D)	NGL	$\chi_g$	$q_g$
1-2	0-1	1	0	1
3-4	2-3	2	$\pm 1/\sqrt{3}$	1/2
5-6	4-5	3	0 $\pm \sqrt{3/5}$	4/9 5/18

TAB.3: coordonnées et poids des points de Gauss sur le segment de référence. (32)

En fait, comme dans (21), il faut comprendre que le couple  $(x3(\chi_g), y3(\chi_g))$  correspond aux coordonnées du point de Gauss considéré sur le segment  $[A1,A2]$  que l'on nommera  $PG_g(x_g, y_g)$ . L'expression (31) devient alors:

$$\int_{[A1,A2]} \mathcal{P}_0^n(x,y) ds = \sum_{g=1}^{NGL} q_g \mathcal{F}(\mathcal{P}_0^n(x_g, y_g), \vec{N}) \quad (33)$$

Cependant, sur l'interface commune entre  $Q_0$  et  $Q_I$ , le flux peut aussi s'écrire au moyen du polynôme  $\mathcal{P}_I^n(x,y)$ :

$$\int_{[A1,A2]} \mathcal{P}_I^n(x,y) ds = \sum_{g=1}^{NGL} q_g \mathcal{F}(\mathcal{P}_I^n(x_g, y_g), \vec{N}) \quad (34)$$

On constate alors que résoudre ce problème de Riemann en utilisant le solveur de Roe [7] revient à poser  $NGL$  fois ce problème pour des valeurs interpolées de part et d'autre de l'interface en de mêmes points de Gauss. Pour alléger l'écriture, nous poserons:

$$U_{0_g}^n = \mathcal{P}_0^n(x_g, y_g) \text{ et } U_{I_g}^n = \mathcal{P}_I^n(x_g, y_g)$$

La différence essentielle avec le schéma de Roe classique est que maintenant il faut pondérer chaque contribution obtenue par  $q_g$ . Pour le reste, rien ne change:

$$\int_{Q_0 \cap Q_I} \mathcal{F}(U^n(x,y), \vec{n}) ds = \Phi_{0I}(\bar{U}^n) = \sum_{g=1}^{NGL} q_g \varphi(U_{0_g}^n, U_{I_g}^n, \vec{N}) \quad (35)$$

avec:

$$\varphi(U_{0_g}^n, U_{I_g}^n, \vec{N}) = \frac{1}{2} \left\{ \mathcal{F}(U_{0_g}^n, \vec{N}) + \mathcal{F}(U_{I_g}^n, \vec{N}) + \gamma \left| \frac{\partial \mathcal{F}(\tilde{U}_g)}{\partial U} \right| (U_{0_g}^n - U_{I_g}^n) \right\}$$

où:

$\tilde{U}_g$  est la moyenne de Roe définie à partir de  $U_{0_g}^n$  et  $U_{I_g}^n$   
 $\gamma$  est un coef. de pondération de la viscosité numérique [8].

En conséquence, plutôt que de redéterminer à chaque calcul du flux convectif les points de Gauss  $PG_g(x_g, y_g)$  sur chaque interface et les normales  $\vec{N}(N_x, N_y)$ , nous les stockons, tandis que les points des sommets des quadrangles peuvent être oubliés.

**Remarque:** dans (TAB.3), il est notable qu'il n'y a pas plus de flux à calculer, du fait de (28), pour un ordre 2 que 1, un ordre 4 que 3,...

### 2.3 Le calcul du terme source

La fonction  $S(t, x, y)$  est connue. Son intégration dans (3) sur chaque quadrangle  $Q_0$  peut donc être déterminée de façon très précise voire exactement. Nous poserons au temps  $t$ :

$$\bar{S}_0(t) = \frac{1}{\text{Aire}(Q_0)} \int \int_{Q_0} S(t, x, y) dx dy \quad (36)$$

Le calcul de ce terme n'apporte pas de difficulté mathématique mais plutôt un souci de coût informatique. Dans le cadre des applications que nous avons menées, pour calculer  $\bar{S}_0(t)$ , nous exploitons la décomposition de la fonction  $S(t, x, y)$  comme:

$$S(t, x, y) = \sum_{i=1}^{NS} s_i(x, y) \tau_i(t) \quad \text{avec } NS \in \mathbb{N} \text{ fini.} \quad (37)$$

Si bien que de (36), nous avons obtenu:

$$\bar{S}_0(t) = \sum_{i=1}^{NS} \tau_i(t) \bar{s}_{0,i} \quad \text{où } \bar{s}_{0,i} = \frac{1}{\text{Aire}(Q_0)} \int \int_{Q_0} s_i(x, y) dx dy \quad (38)$$

Compte tenu de (38), il est intéressant de calculer une seule fois les  $NS$  termes sources spatiaux:  $\bar{s}_{0,i}$  et de réactualiser la valeur  $\bar{S}_0(t)$  en remettant seulement à jour les valeurs  $\tau_i(t)$  au cours du temps.

L'idée générale étant donnée, nous détaillons ici, le calcul du terme source dans le cadre de l'équation de Burgers 2D:

$$\frac{\partial U}{\partial t} + c \cos \theta U \frac{\partial U}{\partial x} + c \sin \theta U \frac{\partial U}{\partial y} = S(t, x, y) \quad (39)$$

où l'on souhaite que (40) soit solution de (39):

$$U(t,x,y) = \cos\{2\pi(mx + ny - \alpha t)\} + \sin\{2\pi(mx + ny - \alpha t)\} \quad (40)$$

avec:  $(m,n) \in \mathbb{N}^2$ ,  $(c,\theta) \in \mathbb{R}^2$ ,  $\alpha = c(m \cos \theta + n \sin \theta)$

En injectant (40) dans (39), il vient:

$$S(t,x,y) = 2\pi\alpha \sum_{i=1}^4 \tau_i(t) s_i(x,y) \quad (41)$$

où:

$$\begin{aligned} \tau_1(t) &= \cos(2\pi\alpha t) - \sin(2\pi\alpha t), & s_1(x,y) &= \sin\{2\pi(mx + ny)\} \\ \tau_2(t) &= 2 \cos(2\pi\alpha t) \sin(2\pi\alpha t), & s_2(x,y) &= \sin\{4\pi(mx + ny)\} \\ \tau_3(t) &= -\cos(2\pi\alpha t) - \sin(2\pi\alpha t), & s_3(x,y) &= \cos\{2\pi(mx + ny)\} \\ \tau_4(t) &= \cos^2(2\pi\alpha t) - \sin^2(2\pi\alpha t), & s_4(x,y) &= \cos\{4\pi(mx + ny)\} \end{aligned}$$

Nous nous efforçons alors d'écrire l'ensemble des fonctions  $s_i$  sous une forme commune en posant:

$$s_i(x,y) = \sin(ax + by + c) \quad (42)$$

où

$i$	$a$	$b$	$c$
1	$2\pi m$	$2\pi n$	0
2	$4\pi m$	$4\pi n$	0
3	$2\pi m$	$2\pi n$	$\frac{\pi}{2}$
4	$4\pi m$	$4\pi n$	$\frac{\pi}{2}$

On se propose d'écrire  $s_i$  comme la divergence d'un vecteur  $\vec{V}_i$  afin de pouvoir transformer par la formule de Green l'intégration de  $s_i$  sur chaque quadrangle  $Q$  en la circulation de  $\vec{V}_i \cdot \vec{n}$  sur le contour  $\partial Q$  du quadrangle:

$$\int \int_Q s_i(x,y) \, dx dy = \int \int_Q \operatorname{div}(\vec{V}_i) \, dx dy = \int_{\partial Q} \vec{V}_i \cdot \vec{n} \, ds \quad (43)$$

$\vec{n}$  est la normale extérieure locale et unitaire du pourtour du quadrangle  $Q$  considéré. Ce pourtour peut être décomposé en la somme des 4 segments  $[A_j, A_{j+1}]$  délimités par les coins  $A_j$ , ( $j \in [1,4]$  et  $A_5 = A_1$ ) du quadrangle numérotés dans le sens trigonométrique (FIG.4).

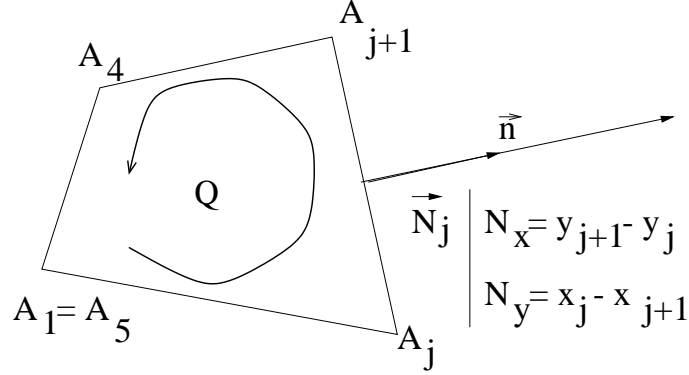


FIG.4: Décomposition du pourtour du quadrangle en 4 segments  $[A_j, A_{j+1}]$ .

où  $\vec{N}_j$  est la normale extérieure intégrée au segment.

En utilisant une paramétrisation selon  $e \in [0,1]$  de chaque segment  $[A_j, A_{j+1}]$  par la transformation linéaire  $f_4$  de type (29) définie par:

$$f_4 : e \mapsto \begin{pmatrix} x(e) \\ y(e) \end{pmatrix} = \begin{pmatrix} -N_y e + x_j \\ N_x e + y_j \end{pmatrix} \quad (44)$$

nous pouvons écrire (43) comme:

$$\int \int_Q s_i(x,y) dx dy = \sum_{j=1}^4 \int_{A_j}^{A_{j+1}} \vec{V}_i \cdot \vec{n} ds = \sum_{j=1}^4 \int_0^1 \vec{V}_j \cdot \vec{N}_j de \quad (45)$$

Ainsi, pour calculer exactement l'intégrale de chaque terme source  $s_i(x,y)$ , nous allons utiliser (45). Pour ce faire, il est temps d'écrire le vecteur  $\vec{V}_i$  selon la configuration  $(m,n)$  du mode de Fourier choisi en prenant en compte d'éventuelles particularités du maillage.

- **Cas  $m = 0$  et  $n = 0$  ( $\Rightarrow a = 0$  et  $b = 0$ )**

Dans ce cas particulier  $\alpha = 0$  et donc dans (41) c'est  $S(t,x,y) = 0$ . Il est alors inutile de calculer les  $s_i(x,y)$ .

- **Cas  $m = 0$  et  $n \neq 0$  ( $\Rightarrow a = 0$  et  $b \neq 0$ )**

$$(42) \Rightarrow s_i(x,y) = \sin(by + c) = -\frac{1}{b} \frac{\partial \cos(by + c)}{\partial y}$$

De fait, on définit le vecteur  $\vec{V}_i$  comme:

$$\vec{V}_i = \begin{pmatrix} 0 \\ -\frac{1}{b} \cos(by + c) \end{pmatrix}$$

d'où dans (45) et en utilisant (44):

$$\int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds = -\frac{N_y}{b} \int_0^1 \cos\{b(N_x e + y_j) + c\} \, de$$

– Sous cas  $N_x = 0$  alors:

$$\int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds = -\frac{N_y}{b} \cos(by_j + c)$$

– Sous cas  $N_x \neq 0$  alors:

$$\int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds = \frac{N_y}{b^2 N_x} \{\sin(by_j + c) - \sin(by_{j+1} + c)\}$$

– **Cas**  $m \neq 0$  et  $n = 0$  ( $\Rightarrow a \neq 0$  et  $b = 0$ )

$$(42) \Rightarrow s_i(x, y) = \sin(ax + c) = -\frac{1}{a} \frac{\partial \cos(ax + c)}{\partial x}$$

De fait, on définit  $\vec{V}_i$  comme:

$$\vec{V}_i = \begin{pmatrix} -\frac{1}{a} \cos(ax + c) \\ 0 \end{pmatrix}$$

d'où dans (45) et en utilisant (44):

$$\int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds = -\frac{N_x}{a} \int_0^1 \cos\{a(-N_y e + x_j) + c\} \, de$$

– Sous cas  $N_y = 0$  alors:

$$\int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds = -\frac{N_x}{a} \cos(ax_j + c)$$

– Sous cas  $N_y \neq 0$  alors:

$$\int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds = \frac{N_x}{a^2 N_y} \{\sin(ax_{j+1} + c) - \sin(ax_j + c)\}$$

– **Cas**  $m \neq 0$  et  $n \neq 0$  ( $\Rightarrow a \neq 0$  et  $b \neq 0$ )

$$(42) \Rightarrow s_i(x,y) = \sin(ax + by + c) = \sin(ax) \cos(by + c) + \sin(by + c) \cos(ax)$$

De fait, on peut définir  $\vec{V}_i$  comme:

$$\vec{V}_i = \begin{pmatrix} -\frac{1}{a} \cos(ax) \cos(by + c) \\ -\frac{1}{b} \cos(by + c) \cos(ax) \end{pmatrix}$$

d'où dans (45) et en utilisant (44):

$$\begin{aligned} \int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds &= -\frac{1}{2ab} (bN_x + aN_y) \\ &\left\{ \int_0^1 \cos[(bN_x - aN_y)e + ax_j + by_j + c] \, de \right. \\ &\left. + \int_0^1 \cos[(bN_x + aN_y)e - ax_j + by_j + c] \, de \right\} \end{aligned}$$

– Sous cas:  $bN_x + aN_y = 0$  alors:

$$\int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds = 0$$

– Sous cas:  $bN_x - aN_y = 0$  alors:

$$\begin{aligned} \int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds &= -\frac{N_x}{a} \cos(ax_j + by_j + c) \\ &+ \frac{1}{2ab} \{ \sin(ax_{j+1} - by_{j+1} - c) - \sin(ax_j - by_j - c) \} \end{aligned}$$

– Sous cas:  $bN_x + aN_y \neq 0$  et  $bN_x - aN_y \neq 0$  alors:

$$\begin{aligned} \int_{A_j}^{A_{j+1}} \vec{V}_j \cdot \vec{n} \, ds &= -\frac{1}{2ab} \left\{ \right. \\ &\frac{bN_x + aN_y}{bN_x - aN_y} \{ \sin(ax_{j+1} + by_{j+1} + c) - \sin(ax_j + by_j + c) \} \\ &\left. + \{ \sin(ax_{j+1} - by_{j+1} - c) - \sin(ax_j - by_j - c) \} \right\} \end{aligned}$$

### Remarques:

- En fait, chaque estimation de l'intégrale sur un segment  $[A_j, A_{j+1}]$  est effectuée une seule fois pour le calcul de l'intégrale de  $s_i(x, y)$  sur les deux quadrangles ayant ce segment en commun. La contribution change seulement de signe.
- Par ailleurs, il faut noter que l'obtention des valeurs moyennes  $\bar{s}_{0,i}$  sur n'importe quel quadrangle  $Q_0$  du maillage est utile non seulement à l'obtention du terme source (41) dans le cadre du modèle Burger 2D, mais aussi à la connaissance de la valeur moyenne exacte du mode de Fourier considéré. Dans un cadre beaucoup plus général, cette valeur nous sert alors de mesure étalon. En effet, de (40), nous pouvons tirer:

$$U(t, x, y) = -\tau_3(t)s_1(x, y) + \tau_1(t)s_3(x, y) \Rightarrow \bar{U}_0^n = -\tau_3(n\Delta t)\bar{s}_{0,1} + \tau_1(n\Delta t)\bar{s}_{0,3} \quad (46)$$

## 2.4 L'avancée en temps

L'intégration des flux convectifs étant effectuée avec une précision à l'ordre  $D + 1$  et celle des termes sources avec une précision machine, il convient de respecter au moins l'ordre  $D + 1$  lors de l'intégration temporelle afin de ne pas introduire d'erreurs supplémentaires et significatives dans la résolution de (3) devenant:

$$\int \int_{Q_0} \frac{\partial U}{\partial t} dx dy = Aire(Q_0)\Psi_0(t, \bar{U}^n) + O(D + 1) \quad (47)$$

où  $\Psi_0$  est la fonction regroupant les flux spatiaux  $\Phi_{0I}(\bar{U}^n)$  et le terme source  $\bar{S}_0(t)$  calculés précédemment pour chaque quadrangle  $Q_0$  à partir:

- des valeurs moyennes  $\bar{U}^n$  connues à l'instant initial  $n\Delta t$  (terme spatial)
- du temps  $t$  de son évaluation (terme source).

Nous l'écrivons:

$$\Psi_0(t, \bar{U}^n) = \bar{S}_0(t) - \frac{1}{Aire(Q_0)} \sum_{I=1}^4 \Phi_{0I}(\bar{U}^n) \quad (48)$$

Pour écrire une intégration temporelle de (47), nous utilisons une méthode de Runge-Kutta à  $NRK \geq D + 1$  pas. Cette méthode explicite permet de connaître à l'instant  $(n + 1)\Delta t$  la valeur moyenne de  $\bar{U}_0^{n+1}$  sur un quadrangle  $Q_0$  si l'on connaît à l'instant  $n\Delta t$  les valeurs moyennes de  $\bar{U}^n$  sur le domaine de calcul. D'une façon générale, cette méthode s'écrit:

$$\bar{U}_0^{n+1} = \bar{U}_0^n + \Delta t \sum_{i=1}^{NRK} r_i K_i \quad \text{où } r_i \in \mathbb{R} \quad (49)$$

avec:

$$K_1 = \Psi_0(\Delta t(n + \eta_1), \bar{U}^n), \quad \eta_1 \in \mathbb{R}$$

$$K_i = \Psi_0\left(\Delta t(n + \eta_i), \bar{U}^n + \Delta t \sum_{j=1}^{i-1} k_{ij} K_j\right), \quad (\eta_i, k_{ij}) \in \mathbb{R}^2, \quad i = 2, \dots, NRK$$



Pour une meilleure lisibilité de l'expression (49), nous nous proposons d'écrire explicitement les différents pas de la méthode. D'abord dans un cadre linéaire puis non linéaire pour  $D=1$  à 5.

- Pour un modèle linéaire, le nombre  $NRK$  de pas à effectuer est égal à l'ordre de précision voulue:

$$NRK = D + 1$$

Le choix des paramètres n'est pas unique et celui que nous donnons semble être le plus automatique et le moins coûteux en stockage [9]. Il s'écrit:

$$\bar{U}_0^{n+1} = \bar{U}_0^n + \Delta t K_{NRK} \quad (50)$$

avec:

$$K_1 = \Psi_0 \left( \Delta t (n + \eta_1), \bar{U}^n \right), \quad \eta_1 = \frac{1}{NRK}$$

$$K_i = \Psi_0 \left( \Delta t (n + \eta_i), \bar{U}^n + \eta_{i-1} \Delta t K_{i-1} \right), \quad \eta_i = \frac{1}{NRK + 1 - i}, \quad i = 2, \dots, NRK$$

On remarque alors que, pour passer à l'étape  $i$ , seule la mémorisation temporaire de l'étape précédente ( $i - 1$ ) est utile.

- Pour un modèle non linéaire, l'écriture est plus lourde et dépend de l'ordre recherché:
  - Ordre 2:  $NRK=2$

$$\bar{U}_0^{n+1} = \bar{U}_0^n + \Delta t K_2 \quad (51)$$

$$K_1 = \Psi_0 \left( n \Delta t, \bar{U}^n \right)$$

$$K_2 = \Psi_0 \left( \left( n + \frac{1}{2} \right) \Delta t, \bar{U}^n + \frac{\Delta t}{2} K_1 \right)$$

- Ordre 3:  $NRK=3$

$$\bar{U}_0^{n+1} = \bar{U}_0^n + \frac{\Delta t}{4} \{ K_1 + 3K_3 \} \quad (52)$$

$$K_1 = \Psi_0 \left( n \Delta t, \bar{U}^n \right)$$

$$K_2 = \Psi_0 \left( \left( n + \frac{1}{3} \right) \Delta t, \bar{U}^n + \frac{\Delta t}{3} K_1 \right)$$

$$K_3 = \Psi_0 \left( \left( n + \frac{2}{3} \right) \Delta t, \bar{U}^n + \frac{2\Delta t}{3} K_2 \right)$$

- Ordre 4:  $NRK=4$

$$\bar{U}_0^{n+1} = \bar{U}_0^n + \frac{\Delta t}{6} \{ K_1 + 2K_2 + 2K_3 + K_4 \} \quad (53)$$

$$K_1 = \Psi_0 \left( n \Delta t, \bar{U}^n \right)$$

$$K_2 = \Psi_0 \left( \left( n + \frac{1}{2} \right) \Delta t, \bar{U}^n + \frac{\Delta t}{2} K_1 \right)$$

$$\begin{aligned} K_3 &= \Psi_0 \left( (n + \frac{1}{2})\Delta t, \bar{U}^n + \frac{\Delta t}{2}K_2 \right) \\ K_4 &= \Psi_0 \left( (n + 1)\Delta t, \bar{U}^n + \Delta t K_3 \right) \end{aligned}$$

– Ordre 5: NRK=6

$$\begin{aligned} \bar{U}_0^{n+1} &= \bar{U}_0^n + \Delta t \left\{ \frac{16}{135}K_1 + \frac{6656}{12825}K_3 + \frac{28561}{56430}K_4 - \frac{9}{50}K_5 + \frac{2}{55}K_6 \right\} \\ K_1 &= \Psi_0 \left( n\Delta t, \bar{U}^n \right) \\ K_2 &= \Psi_0 \left( (n + \frac{1}{4})\Delta t, \bar{U}^n + \frac{\Delta t}{4}K_1 \right) \\ K_3 &= \Psi_0 \left( (n + \frac{3}{8})\Delta t, \bar{U}^n + \frac{3\Delta t}{32}(K_1 + 3K_2) \right) \\ K_4 &= \Psi_0 \left( (n + \frac{12}{13})\Delta t, \bar{U}^n + \frac{\Delta t}{2197}(1932K_1 - 7200K_2 + 7296K_3) \right) \\ K_5 &= \Psi_0 \left( (n + 1)\Delta t, \bar{U}^n + \Delta t \left( \frac{439}{216}K_1 - 8K_2 + \frac{3680}{513}K_3 - \frac{845}{4104}K_4 \right) \right) \\ K_6 &= \Psi_0 \left( (n + \frac{1}{2})\Delta t, \bar{U}^n + \Delta t \left( -\frac{8}{27}K_1 + 2K_2 - \frac{3544}{2565}K_3 + \frac{1859}{4104}K_4 - \frac{11}{40}K_5 \right) \right) \end{aligned} \quad (54)$$

– Ordre 6: NRK=8 ...

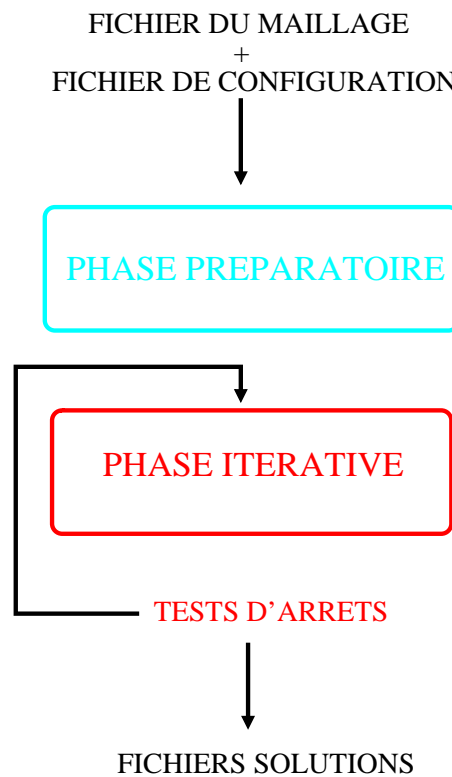
### 3 FAISABILITE

L'aspect mathématique ayant été abordé dans les grandes lignes lors du chapitre précédent, nous abordons ce chapitre d'abord sur le plan de la réalisation informatique puis sur une estimation grossière de l'espace mémoire requis mais aussi sur le temps de calcul que consomme cette nouvelle méthode. Ces points étant traités, nous vérifierons que les objectifs de précision sont atteints. En guise de conclusion, nous choisirons un ordre de précision sur les critères précédents.

#### 3.1 Algorithme

Nous donnons un aperçu du code numérique pour le cas d'une équation scalaire. Néanmoins, dans un cadre non scalaire, nous ferons porter aux tableaux nécessitant plusieurs champs, l'indice  $u$ , mesurant le nombre d'inconnues ( $u \in \mathbb{N}$ ).

Le schéma proposé se décompose en deux grandes phases: l'une préparatoire et l'autre itérative. Le programme principal ([SENO.f](#)) est donc bâti sous cette forme:



- La phase préparatoire (**MAT.f**):

Cette phase n'étant pas récursive, son coût en temps n'est pas important. Par contre et dans la mesure où beaucoup de mémoire est sollicité par la suite, durant cette phase, les déclarations seront locales aux sous-routines et l'on s'autorisera de recommencer de mêmes calculs dès lors qu'ils peuvent donner source à une forte diminution de l'espace mémoire requis. Dans cette phase préparatoire, nous n'avons qu'un but: initialiser les valeurs intemporelles (en rouge) nécessaires à l'autre phase de calcul.

- Appel à la sous-routine **DON.f**

Lecture du fichier de configuration du calcul

**NORDRE**: ordre du polynôme recherché ( $D + 1$ )

**NRK**: ordre de l'intégration temporelle.

**ADVEC, TETA**: paramètres  $c$  et  $\theta$  du modèle mathématique (39)

**CFL**: facteur propre au calcul du pas de temps

**ITEMAX**: nombre d'itérations à réaliser

**ROP**: le paramètre  $\gamma$  de la matrice de Roe

**NCHOIX**: choix du modèle mathématique (advection, Burgers)

**NVX, NVY**: paramètres  $(m, n)$  du mode de Fourier (39)

Déterminations et définitions de:

**NINC**: nombre de coefficients inconnus par polynôme (6)

**IM(NINC), JM(NINC)**: tableaux des puissances en  $x$  et en  $y$  (11)

**NGS**: nombre de points de Gauss pour l'intégration du polynôme sur un triangle de référence (22).

**NGL**: nombre de points de Gauss pour le calcul du flux sur un segment de référence (30).

**GS(NGS, 1 =  $\xi_g/2 = \eta_g/3 = p_g$ )**: mémorisation des points de Gauss sur le triangle de référence (22)

**GL(NGL, 1 =  $\xi_g/2 = q_g$ )**: mémorisation des points de Gauss sur le segment de référence (30).

**NST**: nombre de quadrangles constituant le stencil (FIG.1 - Cas périodique)

- Appel à la subroutine **MAIL.f**

Lecture du maillage structuré des sommets des quadrangles.

**COORD**( $j, k, 1 = x / 2 = y$ ): coordonnées des sommets du maillage

Déterminations et définitions de:

**NQ**: nombre de quadrangles constituant le maillage

**ITE**: mise à zéro du nombre d'itérations effectuées

**T**: mise à zéro de la variable du temps

**DT**: calcul du pas de temps  $\Delta t$

**FS(NQ, 4)**: mise à zéro des 4 termes sources  $\bar{s}_{0,i}$

- Appel à la subroutine **SEG.f**

Déterminations et définitions de:

**NSEG**: nombre total de segments (interfaces) constituant le maillage ( $2NQ$ )

**NUBO(NSEG, 2)**: tableau donnant à un segment orienté le numéro des sommets constituant ses extrémités

**VNOR(NSEG, 3)**: tableau associant à un segment les composantes de la normale extérieure et son module

**PG(NSEG, NGL, 2)**: coordonnées des points de Gauss sur chaque segment.

Cumul des termes sources au moyen de la subroutine **BUS.f**

Boucle sur les quadrangles

Appel à la subroutine **STEN.f**

**LST(NQ,NST)**: donne à un quadrangle les numéros des quadrangles constituant son stencil  
**JST(NST)**, **KST(NST)**: rappel momentané des coordonnées des quadrangles constituant le stencil.

Appel à la subroutine **CENT.f**

**AIRINV(NQ)**: aire inverse de chaque quadrangle  
**COORB(NQ,2)**: coordonnées barycentrique de chaque quadrangle  
**B(NQ,NINC)**: valeur moyenne  $\bar{b}_{0/l,m}$  (10).

Appel à la subroutine **BST.f**

**BNL(NINC,NST)**: valeur moyenne  $\bar{b}_{0/l,m}$  (12)  
**PONDE(NQ,NST)**: pondération sur le voisinage  $\mu_{0/l}$  (9)

**XM(NQ,NINC,NST)**: Stockage des coefficients  $c_{0/l,m}$  (13), utiles à la reconstruction de  $\beta_0^n$  et à la matrice  $\mathcal{M}_0$  (16).  
**COEF(NINC(NINC+1)/2)**: les coefficients de la matrice symétrique  $\mathcal{M}_0$  écrit en 1D.

Appel aux routines de la librairie NAG [6]:

**F06RDF**: donne la norme maximale  
**DSPTRF**: factorise la matrice  
**DSPCON**: évalue la possibilité d'inverser  
**DSPTRI**: inverse la matrice

**SP(NQ,NINC(NINC+1)/2)**: les coefficients de la matrice inverse  $\mathcal{M}_0^{-1}$  écrit en 1D.

- **UAS<sub>u</sub>(NQ)**: initialisation des  $u$  valeurs moyennes sur tous les quadrangles (46).
- Appel à la subroutine **COEFP**

Produit matrice vecteur (25)

$W_u(NINC)$ : mise à jour des  $u$  vecteurs  $\beta_0^n$  (16).

$A_u(NQ, 0 : NINC)$ : coefficients des  $u$  polynômes  $\alpha_0^n$  (16).

Tous les tableaux en rouge étant mis à jour, nous pouvons aborder la phase itérative du calcul.

- La phase itérative:

Cette phase est la plus importante dans la mesure où la majeure partie de l'exécution va se dérouler dans une boucle itérative autour de la méthode de Runge-Kutta à  $NRK$  pas ( $RK*.f$ ). La procédure d'arrêt se faisant sur le nombre maximal d'itérations ( $ITMAX$ ), il convient donc d'optimiser au mieux cette seconde phase.

- Sauvegarde des  $u$  valeurs moyennes de l'instant  $n$

$$SAU_u(NQ)$$

- Mise à zéro du tableau de cumul du flux

$$SAF_u(NQ)$$

- Puis et par itération, nous avons  $NRK$ :

- Mises à zéro des  $u$  flux convectifs puis leur calcul:

$$CE_u(NQ)$$

- Calculs et ajout aux flux convectifs des  $u \times NS$  termes sources en réactualisant seulement les facteurs temporels  $\tau_i^n$  ( $FLUR.f$ ).

- accumulation du flux dans:

$$SAF_u(NQ)$$

- Estimation des  $u$  valeurs moyennes intermédiaires:

$$A_u(NQ,0)$$

- Reconstruction des  $u$  polynômes intermédiaires ( $COEFP.f$ ):

$$A_u(NQ, 0 : NINC)$$

- Incrémentation du pas de temps jusqu'à  $n + 1$ .

### 3.2 Estimation de l'espace mémoire

C'est un point essentiel qui pèse lourd dans la faisabilité de ce type de schéma. Nous ne prétendons pas obtenir une mesure précise de l'espace mémoire total occupé par le code issu de ce schéma mais juste un ordre de grandeur en prenant en compte les principaux tableaux (fonction de  $NQ$  ou  $NSEG$ ) détaillés précédemment et dont la liste constitue vraiment un

minimum. Il faut aussi savoir qu'en FORTRAN 77, un entier est mémorisé au moyen de 4 octets alors qu'un réel pris en double précision en occupe 8.

Nous allons déterminer la fonction MEM donnant le nombre d'octets utilisés par quadrangle. MEM est dépendante de NINC (nombre d'inconnues pour définir un polynôme), NST (nombre de quadrangles constituant le stencil pour une cellule au centre du maillage), NGL (nombre de points de Gauss par interface) et  $u$  (dimension du vecteur inconnu  $\mathcal{U}$ ). En fait, MEM n'est fonction que de 2 variables:  $u$  et  $D$  (le degré du polynôme recherché) grâce au tableau récapitulatif suivant:

$$MEM(D,u) = 4NINC(NINC + 2NST + 2u) + 4NST + 12NINC + 32NGL + 32u + 92 \quad (55)$$

ORDRE	DEGRE (D)	NINC	NST	NGL
2	1	2	4	1
3	2	5	8	2
4	3	9	12	2
5	4	14	24	3

La représentation de la mémoire requise par quadrangle montre une augmentation exponentielle avec le degré  $D$  du polynôme d'approximation de  $U(t,x,y)$ (FIG.5) alors que cette augmentation est seulement linéaire avec la dimension  $u$  du vecteur inconnu.

En effet, ce qui occupe le plus de mémoire est le stockage de chaque matrice inverse et les coefficients de reconstruction du vecteur  $\beta_0^n$ . Or, par quadrangle, cette matrice ou ces coefficients sont les mêmes quelle que soit la dimension  $u$  du vecteur  $\mathcal{U}$  recherché. En d'autres termes, ce qui est important sur le plan de l'espace mémoire, c'est le degré  $D$  du polynôme d'interpolation et non la dimension  $u$  du vecteur  $\mathcal{U}$ .

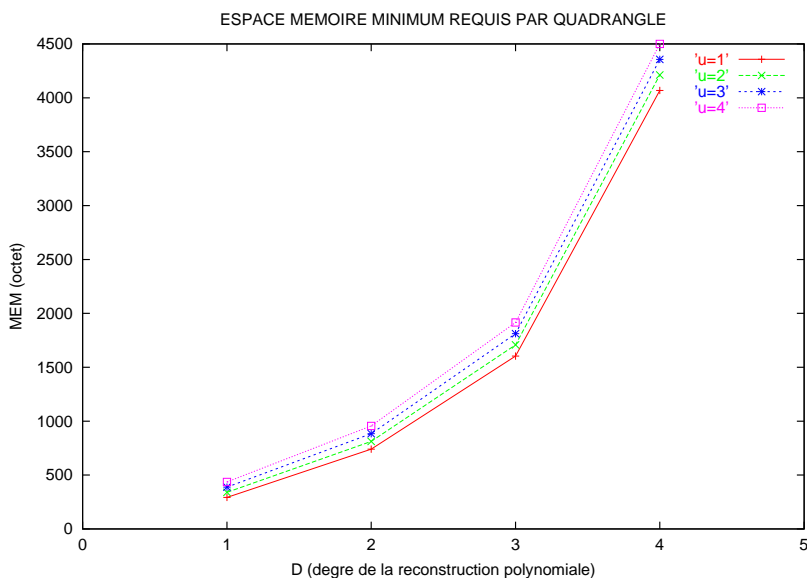


FIG.5: évolution exponentielle de la mémoire avec le degré d'interpolation.

### 3.3 Estimation du temps de calcul

Nous proposons de donner le rapport de temps entre un degré  $D \in [1,4]$  d'interpolation et le degré  $D = 1$  servant alors de référence. Nous procédons à 2 mesures du temps CPU, l'une à 0 itérations et l'autre à 100 itérations sur un même maillage. Nous effectuons par la suite la différence pour ne prendre en compte que le temps CPU utilisé durant la phase itérative. Puis, nous ramenons ce temps CPU en un temps CPU par quadrangle et par itération que nous notons  $TQI(D)$ . Par ailleurs, nous utilisons pour ces mesures de temps CPU, les modèles linéaires comme non linéaires de l'intégration temporelle de manière à prendre en compte la progression non linéaire du nombre d'étapes par la méthode de Runge-Kutta en fonction du degré ( $D = 4 \Rightarrow NRK = 6$  en non linéaire!).

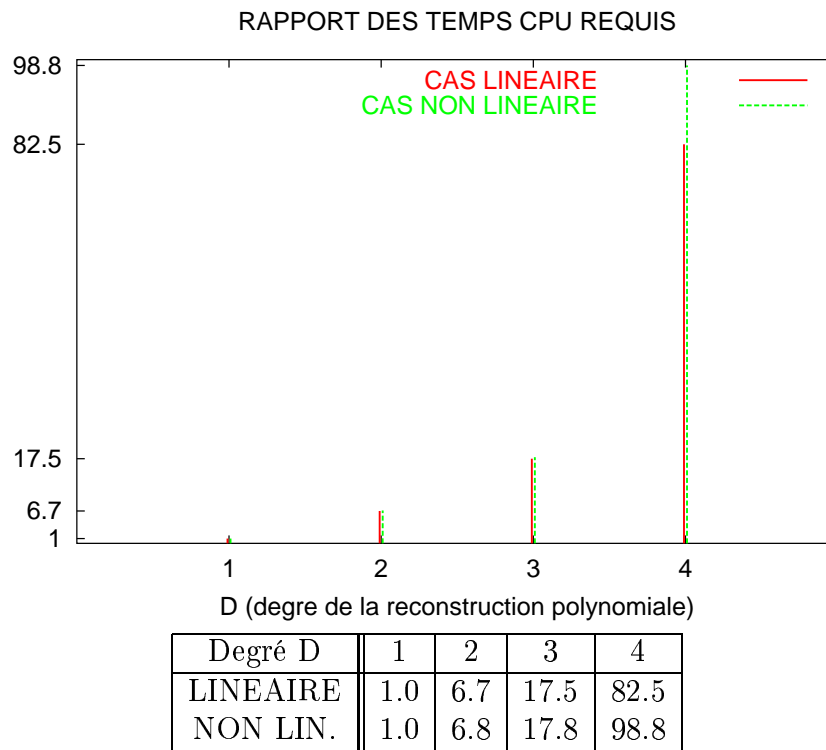


FIG.6: évolution du rapport des temps CPU  $\frac{TQI(D)}{TQI(1)}$  avec le degré  $D$  d'interpolation.

Pour être complet, il manque ici une évaluation des CFL admissibles en fonction de l'intégration temporelle choisie.

### 3.4 Vérifications

Nous souhaitons simplement vérifier ici l'ordre du schéma. Cependant, il faut bien avoir conscience que sous le terme simplement, il y a énormément d'ironie. En effet, vérifier une précision sous entend des types de mesures d'erreurs sur des valeurs à définir pour différents raffinements de maillages et dans des conditions spécifiques de calcul.



L'énorme difficulté est que les conditions spécifiques sont légions et d'en citer quelques unes, soulignons l'essentiel de ce que nous allons représenter:

- en abscisse et en échelle log., nous allons porter 4 raffinements d'un carré  $[0,1] \times [0,1]$

CODE	R16	R32	R64	R128
POINTS	$16 \times 16$	$32 \times 32$	$64 \times 64$	$128 \times 128$

- en ordonnée et en échelle log., nous allons porter une erreur.  
Sous le terme "erreur" nous envisageons 6 choix possibles car une erreur est afférente à:

- une valeur

Or deux valeurs attirent notre attention: la valeur moyenne (M)  $\bar{U}_0^n$  qui est la base de la reconstruction polynômiale et la valeur reconstruite au barycentre (B)  $U^n(x_0, y_0)$  propre à asseoir le bien fondé de la méthodologie pour les numériciens habitués aux schémas centré cellule.

- une mesure

Sur l'ensemble des valeurs  $V_i$  du quadrangle  $Q_i$ , on peut procéder à différentes mesures d'erreurs par rapport à la valeur exacte  $V_{e_i}$  en utilisant différentes normes comme:

$$\text{Norme L1: } \frac{1}{NQ} \sum_{i=1}^{NQ} |V_i - V_{e_i}|$$

$$\text{Norme L2: } \sqrt{\frac{1}{NQ} \sum_{i=1}^{NQ} (V_i - V_{e_i})^2}$$

...

$$\text{Norme LI } \max_{i=1}^{NQ} |V_i - V_{e_i}|$$

Ce qui conduit à la représentation de 6 erreurs possibles:

CODE	ML1	ML2	MLI	BL1	BL2	BLI
------	-----	-----	-----	-----	-----	-----

Abordons maintenant les conditions spécifiques des calculs en commençant par celles qui seront fixés une fois pour toutes sous ce chapitre:

- Le temps  $T = 6.25 \cdot 10^{-2}$
- Le nombre de Courant:  $CFL = 0.25$  (valeur non optimisée)
- L'état initial: mode (1,1) de Fourier avec  $c = 1$  et  $\theta = \frac{\pi}{4}$  (40)

$$U(x, y, t = 0) = \cos\{2\pi(x + y)\} + \sin\{2\pi(x + y)\}$$

- Condition de bord périodique: 

CODE	PER
------	-----

Citons maintenant, toutes les conditions spécifiques qu'il serait intéressant de faire varier:

- Le modèle numérique avec deux équations scalaires ( $u = 1$ ) envisagées:  
l'une linéaire (Advection 2D):

$$\frac{\partial U}{\partial t} + \frac{\sqrt{2}}{2} \left\{ \frac{\partial U}{\partial x} + \frac{\partial U}{\partial y} \right\} = 0 \quad (56)$$

l'autre non linéaire (Burgers 2D):

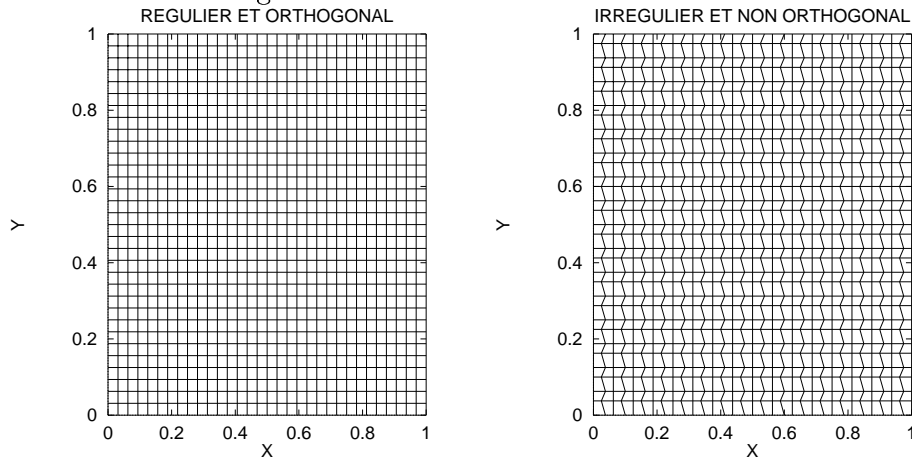
$$\frac{\partial U}{\partial t} + \frac{\sqrt{2}}{2} U \left\{ \frac{\partial U}{\partial x} + \frac{\partial U}{\partial y} \right\} = S \quad (57)$$

TYPE	LINEAIRE	NON LINEAIRE
CODE	ADV	BUR

- Le degré  $D$  d'interpolation polynômial devant donner un ordre  $D + 1$  au schéma numérique.

D	1	2	3	4
CODE	O2	O3	O4	O5

- La nature du maillage:



NATURE	REGULIER	IRREGULIER
CODE	REG	IRE

- La viscosité numérique du schéma de Roe piloté par le paramètre  $\gamma$  (35)

$\gamma$	0	1
CODE	G0	G1

- La pondération  $\mu$  de l'ajustement par les moindres carrés dans (9)

CODE	MU1	MU2	MU3	MU4
------	-----	-----	-----	-----

En conséquence, un calcul est maintenant bien définie en lui donnant l'ensemble des codes l'identifiant comme par exemple:

R16-MLI-BUR-O5-IRE-PER-G0-MU3

Ce calcul est un point de la représentation graphique:

MLI-BUR-O5-IRE-PER-G0-MU3

Si on voulait étudier toutes les combinaisons possibles, il faudrait mener 3072 calculs et établir 768 représentations graphiques qu'il faudrait comparer entre elles! Un tel travail n'étant guère imaginable, nous allons faire des choix en espérant ne pas passer à côté d'une phénoménologie importante...

- Choix d'une erreur:

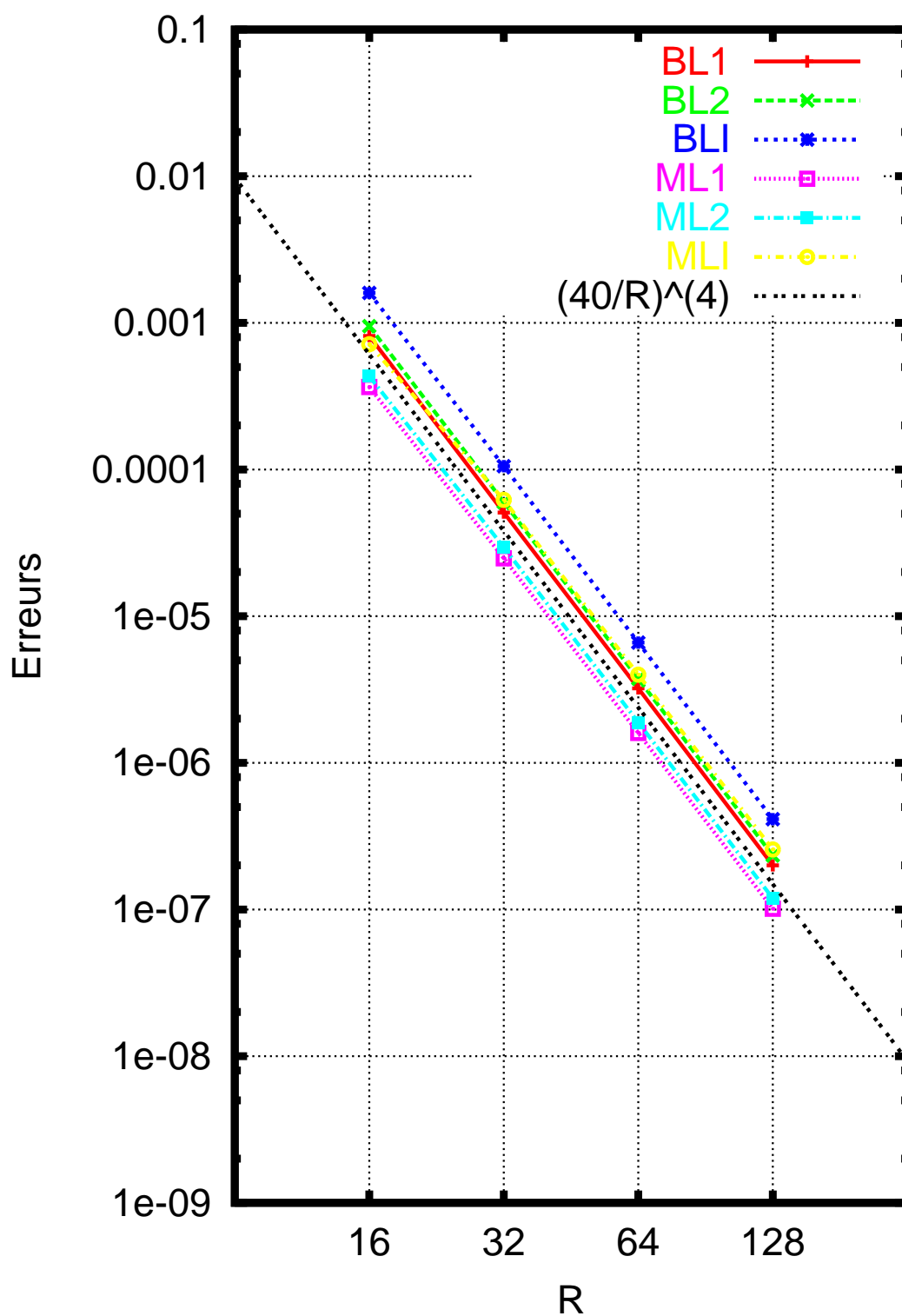
Nous procédons à toutes les mesures d'erreurs (\*) pour une configuration que l'on espère représentative de l'ensemble des configurations possibles. Afin de ne pas être accusé de choisir un cas trop trivial, nous prenons le modèle non linéaire, un ordre élevé ( $D = 3$ ), un maillage irrégulier, le schéma de Roe classique ( $\gamma = 1$ ) et une reconstruction par la méthode des moindres carrés sans pondération ( $\mu_{0/l} = 1$ ): \*-BUR-O4-IRE-PER-G1-MU4

Sur la représentation (FIG.7) en échelle log., nous constatons que toutes les erreurs se comportent comme des droites de pente  $-4$  en fonction du raffinement ( $R$ ) du maillage. Ce point nous confirme l'ordre 4 du schéma. Par ailleurs, on remarque que:

- pour une norme donnée ( $L$ ), les erreurs sur les valeurs moyennes ( $M$ ) sont toujours inférieures aux erreurs sur les valeurs reconstruites aux barycentres ( $B$ ).
- pour une valeur choisie ( $M$  ou  $B$ ), plus la norme est petite ( $L1$ ) et plus l'erreur est faible.

A partir de ces faits et avec l'idée de se placer dans les conditions les plus défavorables possibles nous choisirons par la suite l'erreur  $BL2$ . La norme infinie étant trop ponctuelle est écartée.

EXPERIENCE: \*-BUR-O4-IRE-PER-G1-MU4

FIG.7: Choix d'une erreur  $\Rightarrow$  BL2

- Choix de la pondération dans la méthode de moindres carrés.

Nous avons envisagé en (9) différentes façons de pondérer l'importance des quadrangles  $Q_l$  appartenant à la molécule du quadrangle  $Q_0$ .

- MU1: par l'inverse de la distance séparant le barycentre de  $Q_0$  au barycentre de chaque quadrangle  $Q_l$  ( $\mu_{0/l} = \frac{1}{\sqrt{(x_0-x_l)^2+(y_0-y_l)^2}}$ ).
- MU2: par l'aire de chaque quadrangle  $Q_l$  ( $\frac{1}{\text{AIRINV}(Q_l)}$ ).
- MU3: par le produit de MU1 et de MU2.
- MU4: par le choix de la facilité ( $\mu_{0/l} = 1$ )!

Nous souhaitons maintenant nous positionner envers ces quatre pondérations. Pour ce faire, nous reprenons l'expérience précédente en mesurant uniquement l'erreur BL2. La considération d'un maillage irrégulier s'impose pour pouvoir faire intervenir efficacement les différentes pondérations. Nous envisageons deux rapports de taille entre 2 segments voisins: 1.5 (FIG.8) et 5 (FIG.9) de manière à souligner l'effet de la pondération en fonction de l'irrégularité du maillage.

Nous vérifions sur les figures 8 et 9 que toutes les pondérations envisagées préservent l'ordre 4 souhaité pour cette expérience. Il est clair que l'on n'escompte pas obtenir un ordre supérieur mais une erreur moindre selon la pondération choisie. En fait, sur le maillage faiblement irrégulier (1.5), les erreurs BL2 sont vraiment semblables d'une pondération à une autre; ce n'est qu'avec un maillage plus irrégulier (5) que l'on observe finalement et à condition de s'armer d'une loupe (FIG.9), que:

$$MU4 \sim MU2 > MU1 \sim MU3$$

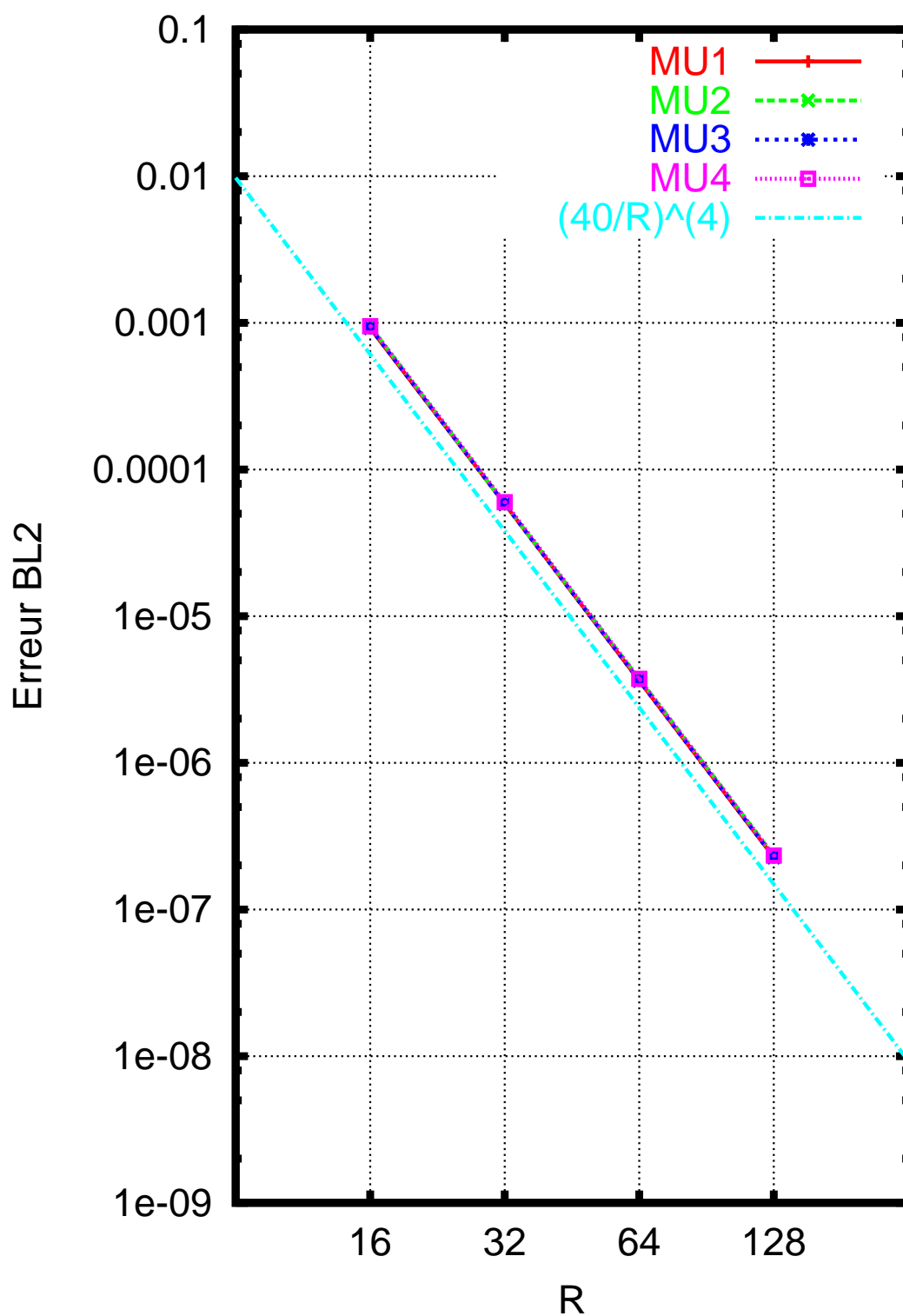
comme le confirme le tableau ci-dessous donnant les ajustements linéaires d'après (58) en fonction de la pondération.

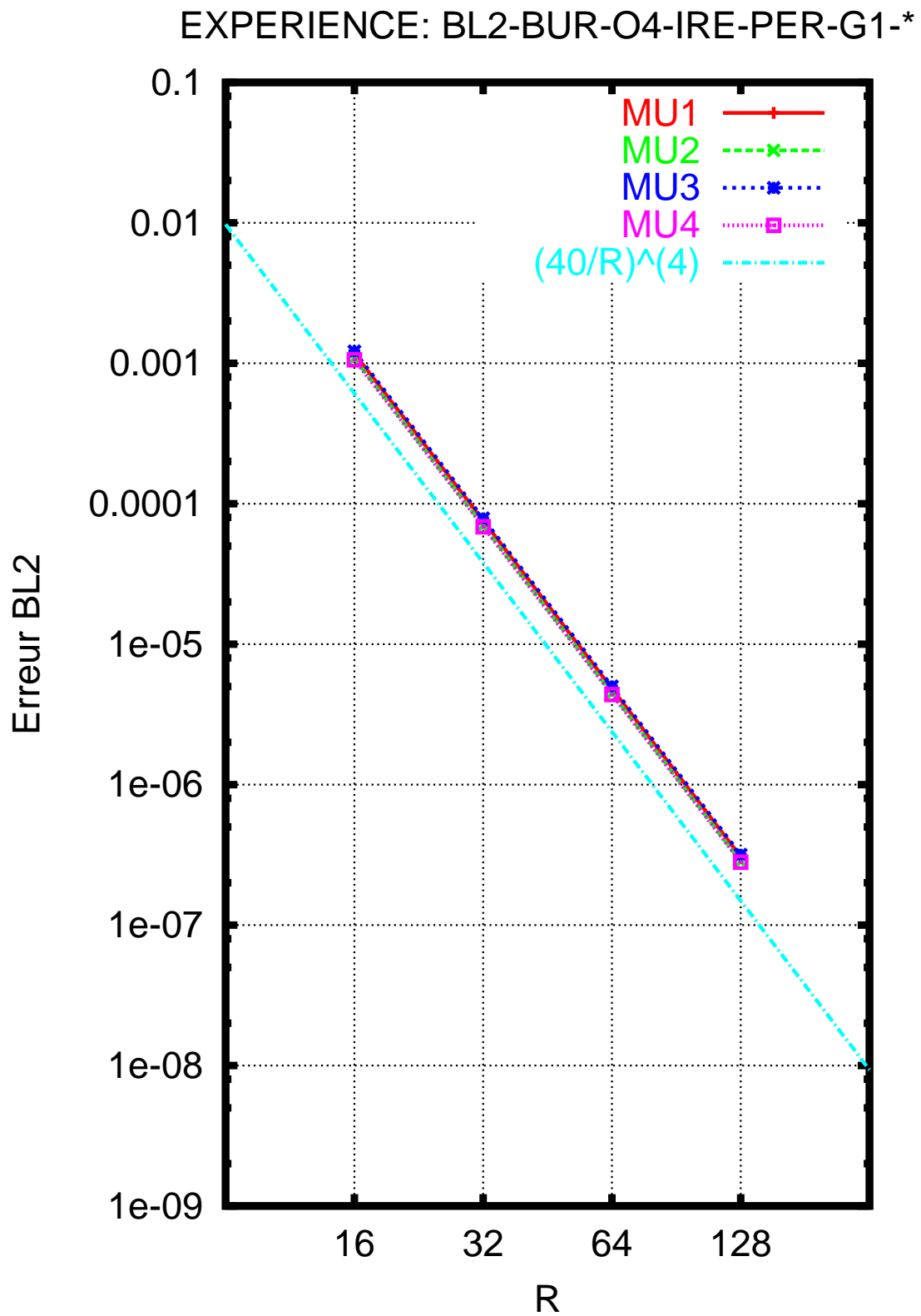
Pondération	Facteur $\Lambda$	Puissance p
MU1	70.34	3.97
MU2	64.89	3.97
MU3	73.64	3.97
MU4	62.80	3.97

Il ressort de ce tableau que les différences sont vraiment faibles et ne portent, comme prévu, que sur les facteurs  $\Lambda$  avec une variation maximale de 15%. Dès lors et de façon surprenante, on retiendra que de ne rien faire du tout est optimal (MU4) tant sur le plan de la précision que sur le plan du coût en espace mémoire. En effet, le stockage de  $NST$  réels supplémentaires par quadrangle (PONDE(NQ,NST)) est alors évité.

Nous prendrons donc pour la suite la pondération MU4 ( $\mu_{0/l} = 1$ ).

## EXPERIENCE: BL2-BUR-O4-IRE-PER-G1-\*

FIG.8: rapport 1.5 entre deux segments voisins  $\Rightarrow$  MU4

FIG.9: rapport 5 entre deux segments voisins  $\Rightarrow$  MU4

- Vérification de l'ordre

Pour vérifier l'ordre des schémas, nous vérifions par une méthode de moindres carrés le comportement de l'erreur BL2 en fonction du raffinement R du maillage en considérant toujours la pondération MU4 lors de l'approximation polynômiale.

$$ErreurBL2 = \Lambda \left(\frac{1}{R}\right)^p \quad (58)$$

Nous procédons dans le cadre du schéma de Roe classique ( $\gamma = 1$ ) aux 4 expérimentations suivantes:

BL2-ADV-* <b>-REG</b> -PER-G1-MU4			BL2-ADV-* <b>-IRE</b> -PER-G1-MU4		
Degré D	Facteur $\Lambda$	Puissance p	Degré D	Facteur $\Lambda$	Puissance p
1	5.69	2.07	1	5.80	2.06
2	34.66	3.01	2	35.14	3.01
3	67.92	4.05	3	68.73	4.05
4	1299.94	4.98	4	1315.09	4.98

BL2-BUR-* <b>-REG</b> -PER-G1-MU4			BL2-BUR-* <b>-IRE</b> -PER-G1-MU4		
Degré D	Facteur $\Lambda$	Puissance p	Degré D	Facteur $\Lambda$	Puissance p
1	5.75	2.00	1	5.26	1.96
2	26.68	2.93	2	27.30	2.93
3	59.70	4.00	3	60.00	4.00
4	1066.08	4.91	4	1093.65	4.91

TAB.4: ajustement par les moindres carrés des expériences

Nous observons dans le tableau 4 et sur les figures 10 que l'erreur BL2 représentée en échelle logarithmique en fonction du raffinement de maillage réalise effectivement et à quelques pourcents près l'ordre voulu (la puissance p) en fonction du degré D de la reconstruction polynômiale. Nous constatons aussi l'évolution de  $\Lambda$  selon le degré D voulu et c'est avec déception que l'on s'aperçoit que le degré D=4 occasionne pour de petits maillages ( $R < 27$ ) une erreur plus grande que le degré D=3!

Par ailleurs et d'une façon assez logique, l'erreur BL2 devient sensiblement plus importante en passant du modèle linéaire sur maillage régulier au maillage irrégulier puis du modèle non linéaire sur maillage régulier au maillage irrégulier. Cette propriété est d'autant plus sensible pour les degrés de reconstruction petits (D=1 et 2). Ainsi, pour une erreur BL2 de  $10^{-4}$ , il faut, pour D=1, passer d'un raffinement de  $198 \times 198$  à  $256 \times 256$  alors que pour D=4, une grille de  $27 \times 27$  suffit quelle que soit l'expérience.

Par ces expériences, nous avons donc vérifié que l'ordre du schéma numérique correspond à l'ordre attendu pour un modèle linéaire ou non et pour un maillage quelconque. Ceci dans le cadre du schéma de Roe classique.



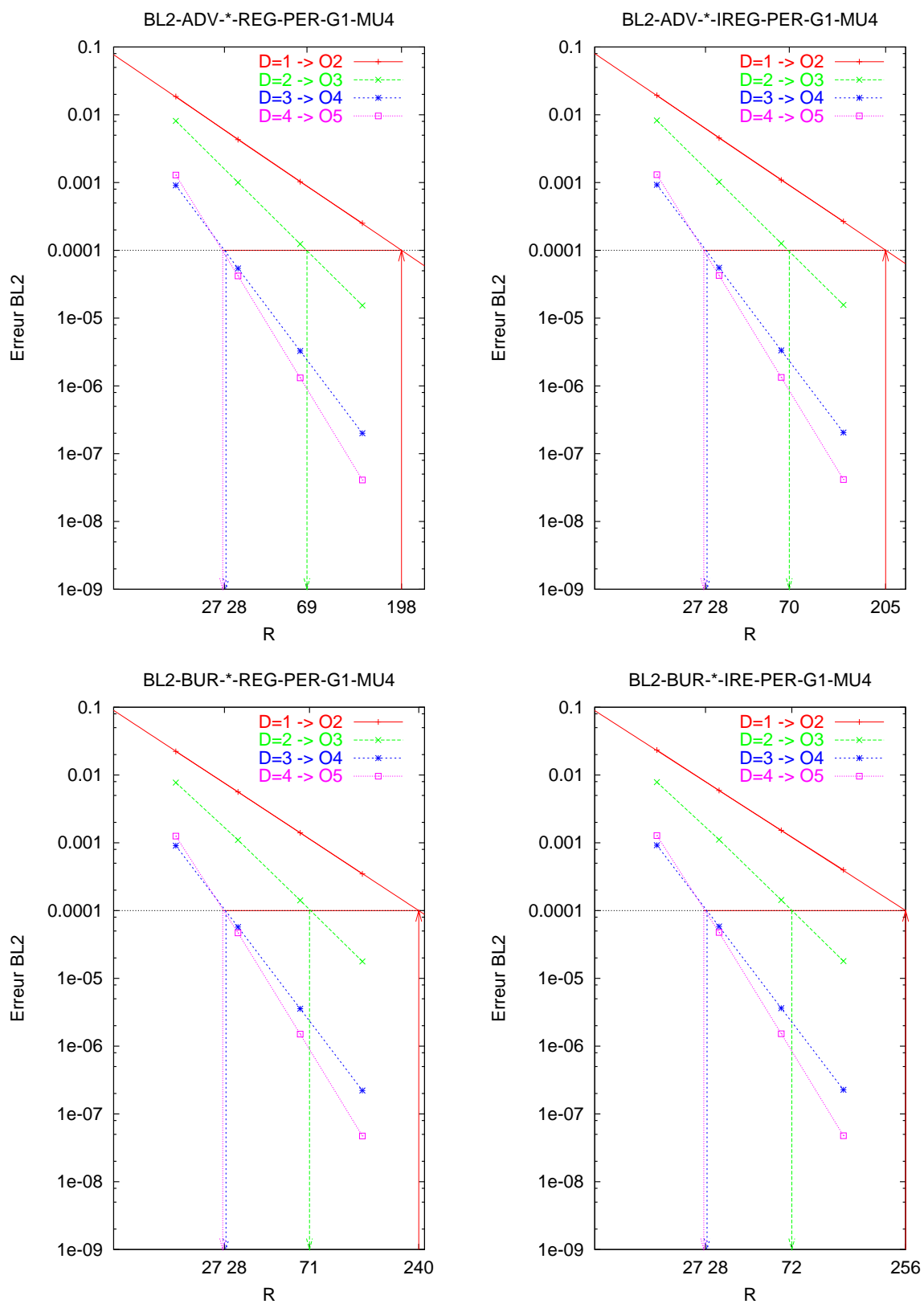


FIG.10: Validation des ordres pour le schéma classique de Roe (G1)

Maintenant, nous souhaitons ouvrir une parenthèse assez succincte sur le paramètre  $\gamma$  coefficient devant la matrice de Roe et pondérant la viscosité numérique du schéma. Jusqu'à présent, nous sommes restés dans le cadre du schéma de Roe classique avec  $\gamma = 1$  (G1) et nous nous proposons de supprimer la viscosité artificielle en posant  $\gamma = 0$  (G0). D'ores et déjà, nous savons que cette action rend le schéma instable dans le cadre non linéaire quelle que soit la nature du maillage. De fait, nous allons nous limiter au cadre linéaire et ne présenterons que des résultats obtenus sur maillage régulier où une super-convergence peut apparaître...

BL2-ADV-*-REG-PER-G0-MU4		
Degré D	Facteur $\Lambda$	Puissance p
1	3.69	1.98
2-CFL=0.25	65.82	3.90
2-CFL=0.5	20.41	3.45
2-CFL=1	18.08	3.02
2-RK4	34.03	3.02
3	48.32	3.98
4-CFL=0.25	3308.09	5.96
4-CFL=0.5	3352.87	5.96
4-CFL=1	2057.77	5.80
4-CFL=2	1461.94	5.02
4-RK6	1299.98	4.98

TAB.5: ajustement par les moindres carrés

Nous retrouvons pour les degrés de reconstruction impairs ( $D=1$  -  $D=3$ ), les ordres pairs attendus ( $O2$  -  $O4$ ). Pour les degrés de reconstruction pairs ( $D=2$  -  $D=4$ ), nous obtenons pratiquement un ordre supérieur à celui attendu ( $\sim O4$  -  $\sim O6$ ). Ces derniers résultats paraissent fort intéressants et si l'ordre supérieur n'est pas vraiment bien réalisé (3.90), la faute pouvait être rejetée sur l'erreur temporelle. En effet, si l'on considère que l'erreur au ordre impaire ( $2k+1$ ) s'écrit comme:

$$Erreur_{BL2} = f(\gamma)\Delta^{2k+1} + g(CFL)\Delta t^{2k+1}$$

et que  $f(0) = 0$  alors, l'ordre supérieur pourrait être obtenu en supprimant l'erreur temporelle grâce à une intégration de Runge-Kuta d'ordre supérieur. Nous avons donc procédé à des calcul utilisant un RK4 plutôt qu'un RK3 et un RK6 plutôt qu'un RK5. Les résultats sont alors bien pitoyables: plutôt que d'affiner la super-convergence précédente, les ordres 3 et 5 apparaissent très nettement. La conclusion que l'on peut en tirer, est qu'il y a eu compensation entre l'erreur temporelle avec un CFL=0.25 (pris arbitrairement) et l'erreur spatiale qui ne s'annule pas lorsque  $\gamma = 0$ . Nous nous trouvons alors dans des conditions de calculs dangereuses tant sur la restriction du modèle linéaire, du maillage régulier mais aussi du degré de reconstruction et du CFL choisi! Sans parler de la stabilité d'un tel schéma.

En conclusion, prendre  $\gamma = 0$  est une option qui, dans ce contexte de schéma de haute précision, semble bien inutile.

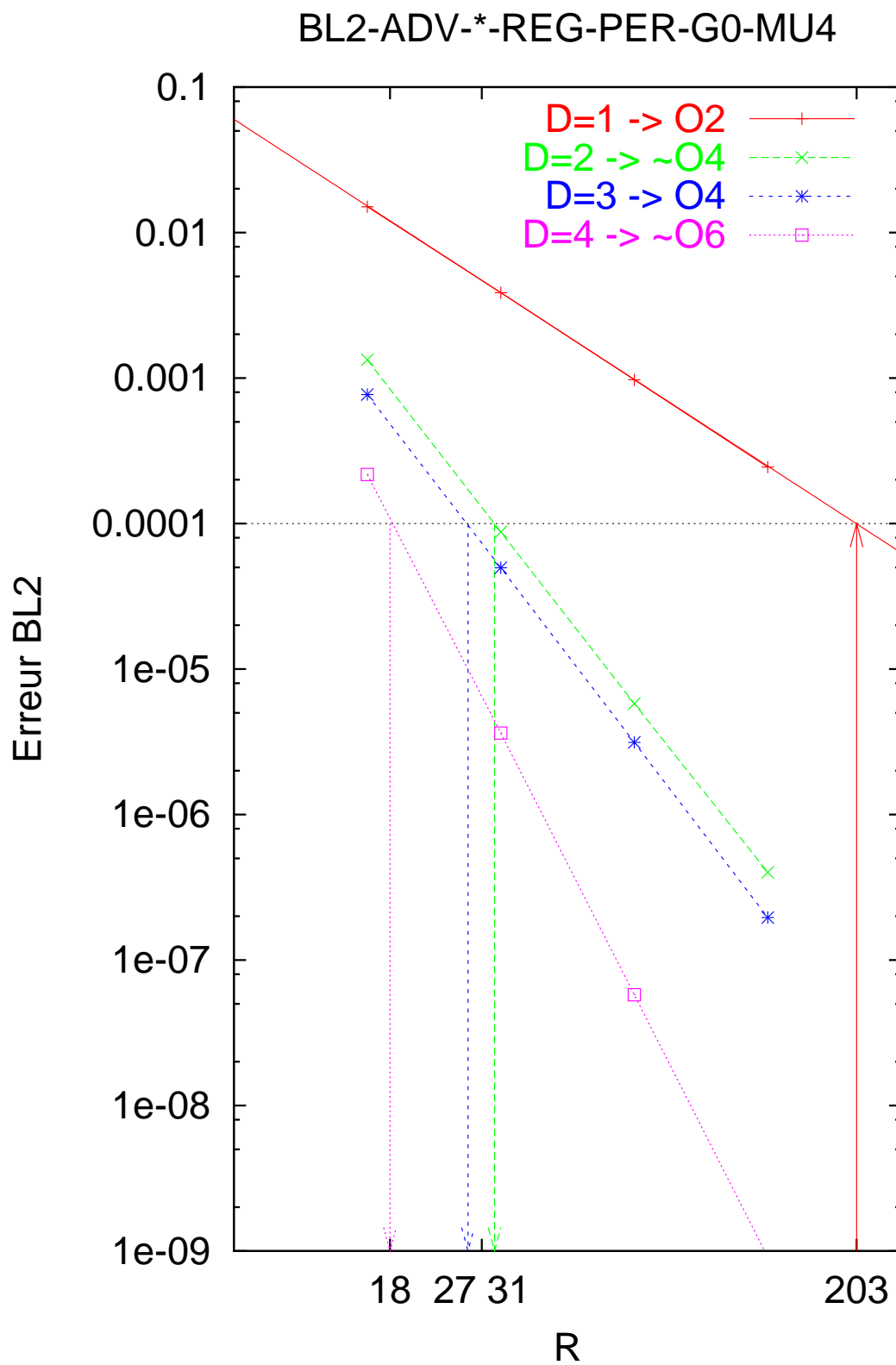


FIG.11: avec  $\gamma = 0$ , une super convergence pour  $D=2$  et  $D=4$  apparaît...

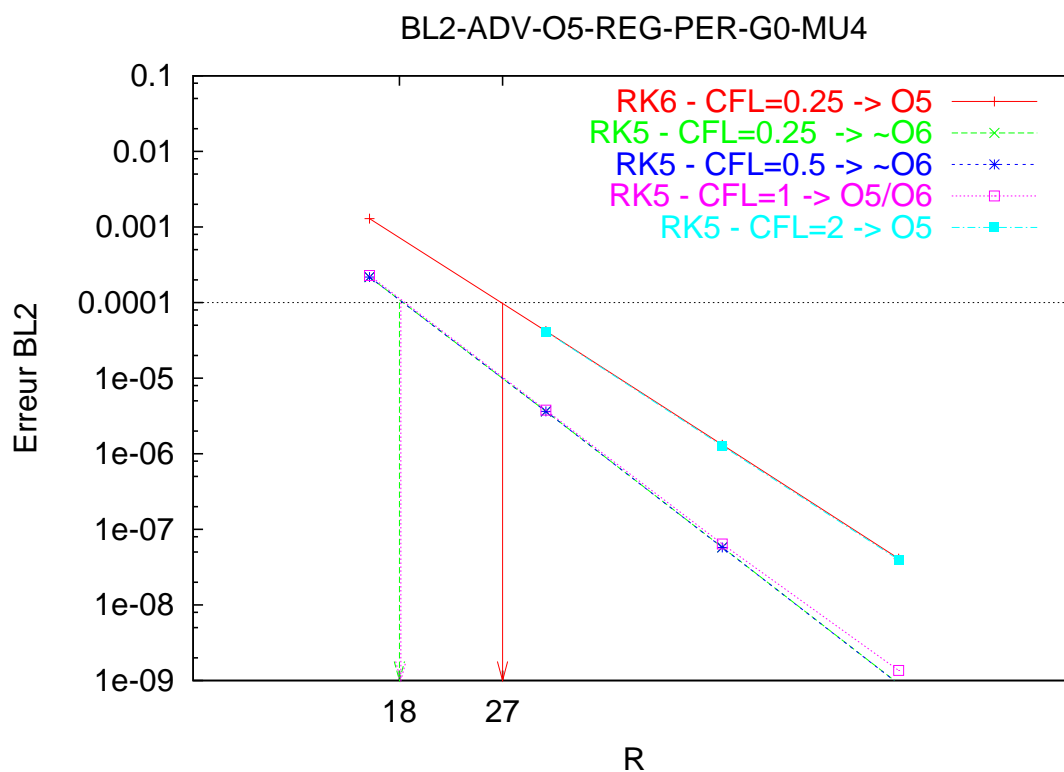
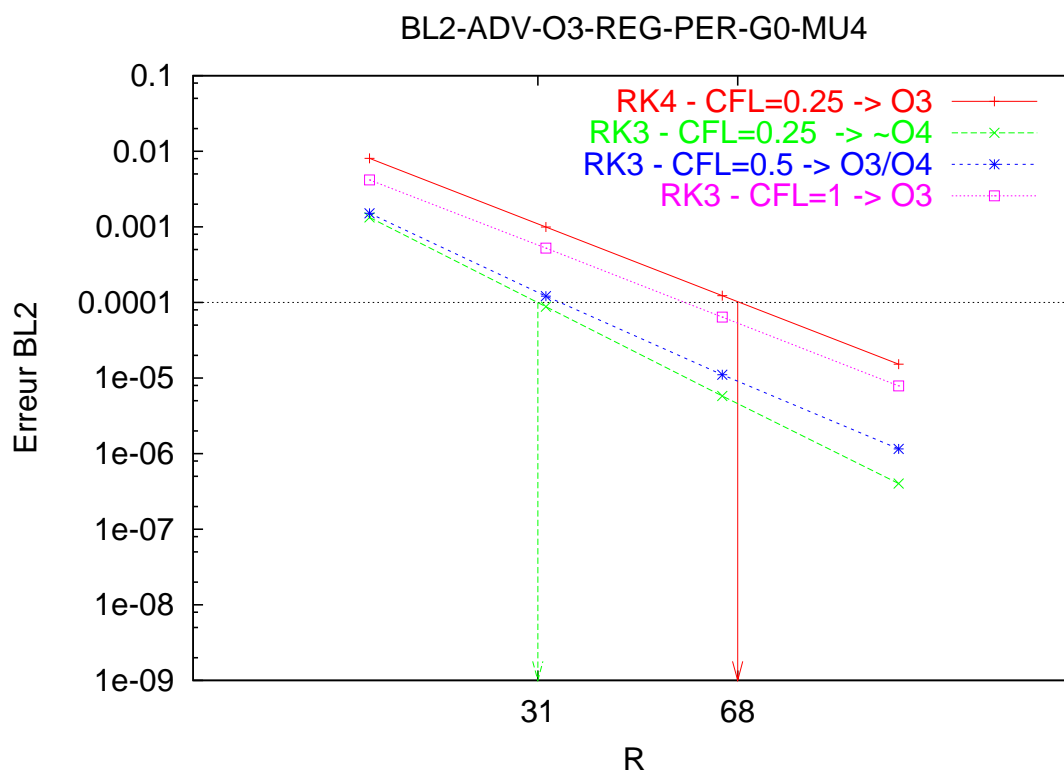


FIG.12: ... mais est fonction du CFL choisi!

### 3.5 Choix de précision

Pour continuer l'étude, nous allons maintenant privilégier un certain degré  $D$  de reconstruction polynômiale. Pour faire ce choix, nous allons d'abord estimer les coûts CPU pour obtenir une erreur BL2 écrite en  $10^{-pr}$  pour chaque degré  $D \in [1,4]$  notés  $C(pr,D)$ . Ensuite, nous formulerons le rapport du coût CPU lorsque  $D \geq 1$  par référence au coût CPU lorsque  $D = 1$ . Nous appellerons ce rapport la fonction gain  $G(pr,D)$  qui s'écrit:

$$G(pr,D) = \frac{C(pr,1)}{C(pr,D)} \quad (59)$$

Le coût CPU pour un degré  $D$  à une erreur BL2 donnés se calcule par le produit du temps CPU par quadrangle du maillage et par itération  $TQI(D)$  multiplié par le nombre de quadrangles du maillage  $R^2(pr,D)$  et par le nombre d'itérations  $I(pr,D)$  nécessaires pour atteindre le temps  $T$  du calcul où la mesure l'erreur BL2 est effectuée.

$$C(pr,D) = TQI(D)R^2(pr,D)I(pr,D) \quad (60)$$

En effet, le coût CPU, ne se résume pas seulement à un temps CPU ( $TQI(D)$ ) plus important avec  $D$  croissant jouant en la défaveur d'un degré  $D$  élevé. Il faut prendre en considération d'une part, que pour obtenir une précision donnée  $pr$ , les maillages deviennent de plus en plus grossiers avec  $D$  croissant. Ces maillages comptabilisent  $R^2(pr,D)$  quadrangles où  $R(pr,D)$  se déduit de la relation (58) et s'écrit:

$$R(pr,D) = 10^{\left(\frac{pr + \text{Log}[\Lambda(D)]}{p(D)}\right)} \quad (61)$$

où  $\Lambda(D)$  et  $p(D)$  sont donnés dans les tableaux 4 et 5.

D'autre part, que les maillages devenant plus grossiers, les pas de temps  $\Delta t$  peuvent être pris plus grands et donc permettre à un calcul de s'effectuer en moins d'itérations  $I(pr,D)$ . Pour un même CFL, ce nombre d'itérations est proportionnel à  $R(pr,D)$ .

L'ensemble est donc bien traduit par la relation (60) et comme l'expression du gain  $G(pr,D)$  est un rapport, le facteur de proportionnalité se simplifie laissant (59) s'écrire:

$$G(pr,D) = \frac{TQI(1)}{TQI(D)} \left( \frac{R(pr,1)}{R(pr,D)} \right)^3 \quad (62)$$

Par la figure 6, nous avons connaissance du rapport  $\frac{TQI(1)}{TQI(D)}$ . Par les tableaux TAB.4, TAB.5 et l'expression (61) à  $R(pr,D)$ . Nous avons donc accès à  $G(pr,D)$  pour tous les cas de figures que nous avons envisagés précédemment. Pour choisir un degré  $D$  de reconstruction polynômiale, il est intéressant pour chaque précision  $pr_0$  donnée de mettre en évidence le degré  $D_0$  de reconstruction donnant le gain  $G(pr_0,D_0)$  maximum. C'est ce qui a été réalisé sur les figures 13 à 17 où l'on observe des points anguleux dès lors que l'on change de degré. Dans la partie supérieure de chaque figure et pour une précision  $pr_0$ , est donnée la valeur  $R(pr_0,D_0)$ . Il faut élever au carré cette valeur pour obtenir le nombre de quadrangles nécessaires à la constitution du maillage permettant d'obtenir la précision voulue avec ce degré.

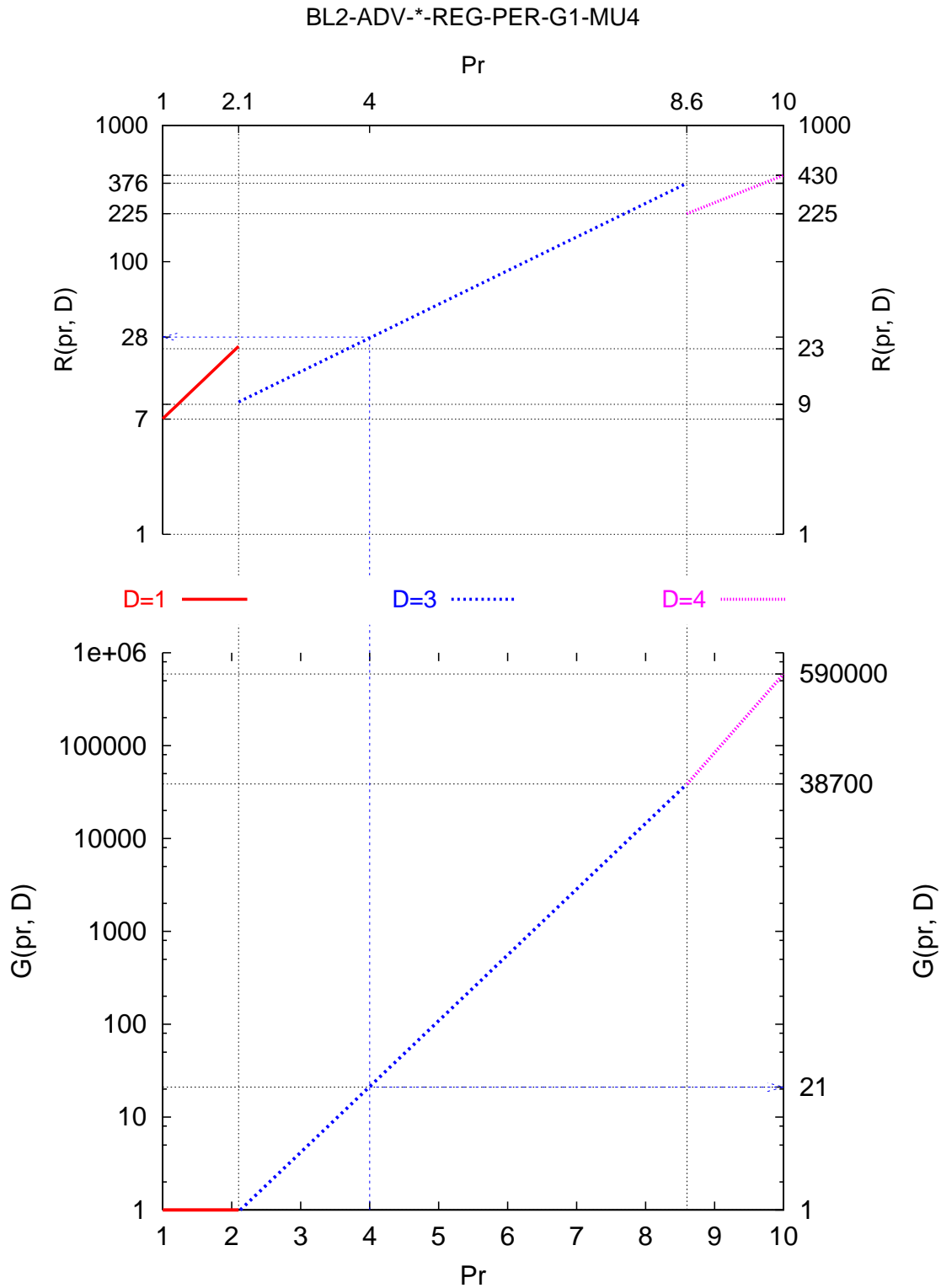


FIG.13: cas de l'advection, maillage régulier, conditions périodiques,  $\gamma = 1$ :  
 abscisse: erreur  $BL2=10^{-Pr}$ , ordonnée: fonction gain  $G$  et nombre de quadrangles  $R^2$ .  
 $10^{-2.1} \geq BL2 \geq 10^{-8.6}$ ,  $D = 3$  est préférable.  $D = 2$  n'est jamais optimal.  
 Ex:  $BL2 = 10^{-4} \Rightarrow D = 3$ : gain de 21 et maillage de  $28^2$  quadrangles!

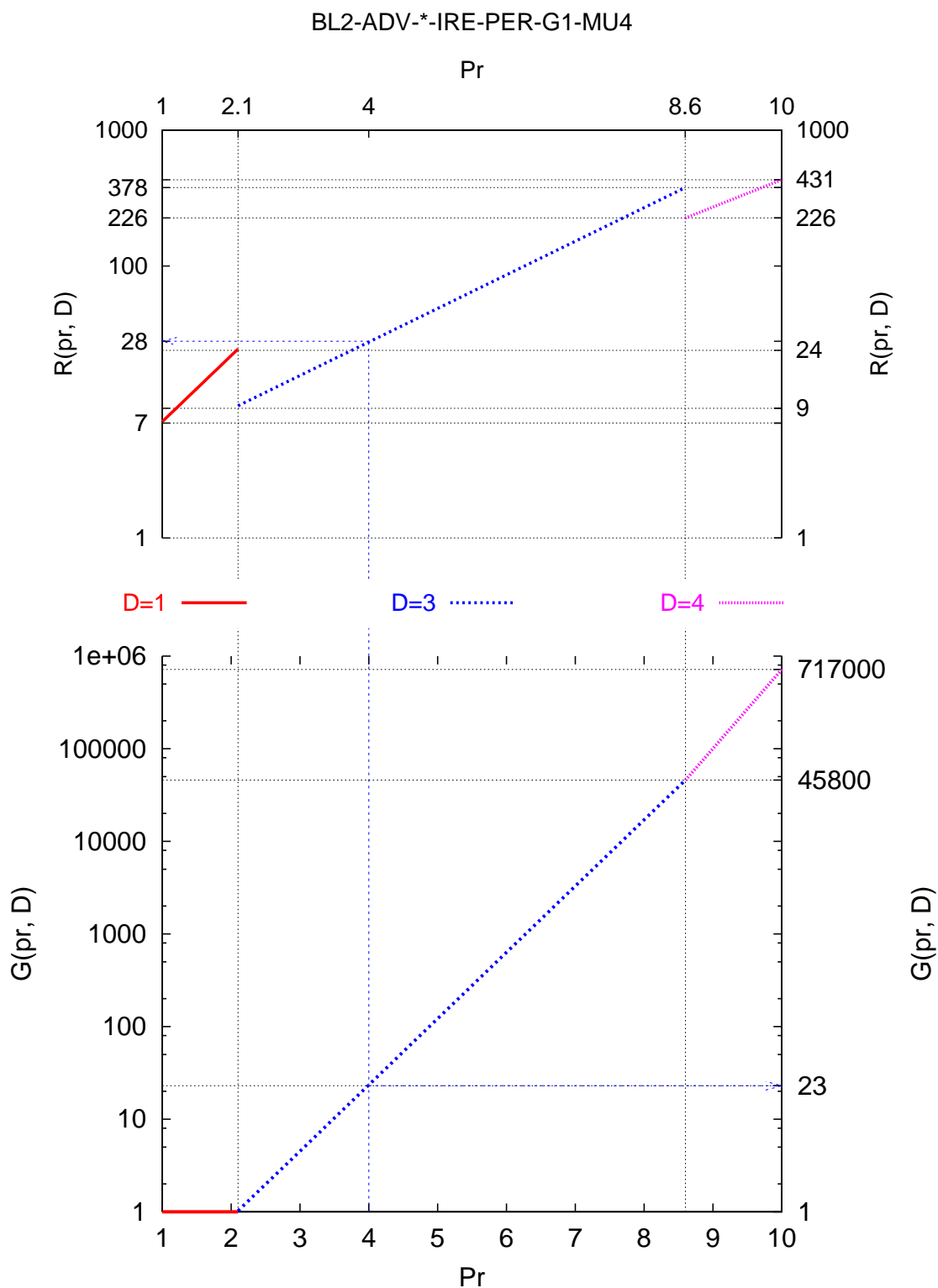


FIG.14: cas de l'advection, maillage irrégulier, conditions périodiques,  $\gamma = 1$ :  
 abscisse: erreur  $BL2=10^{-Pr}$ , ordonnée: fonction gain  $G$  et nombre de quadrangles  $R^2$ .  
 $10^{-2.1} \geq BL2 \geq 10^{-8.6}$ ,  $D = 3$  est préférable.  $D = 2$  n'est jamais optimal.  
 Ex:  $BL2 = 10^{-4} \Rightarrow D = 3$ : gain de 23 et maillage de  $28^2$  quadrangles!

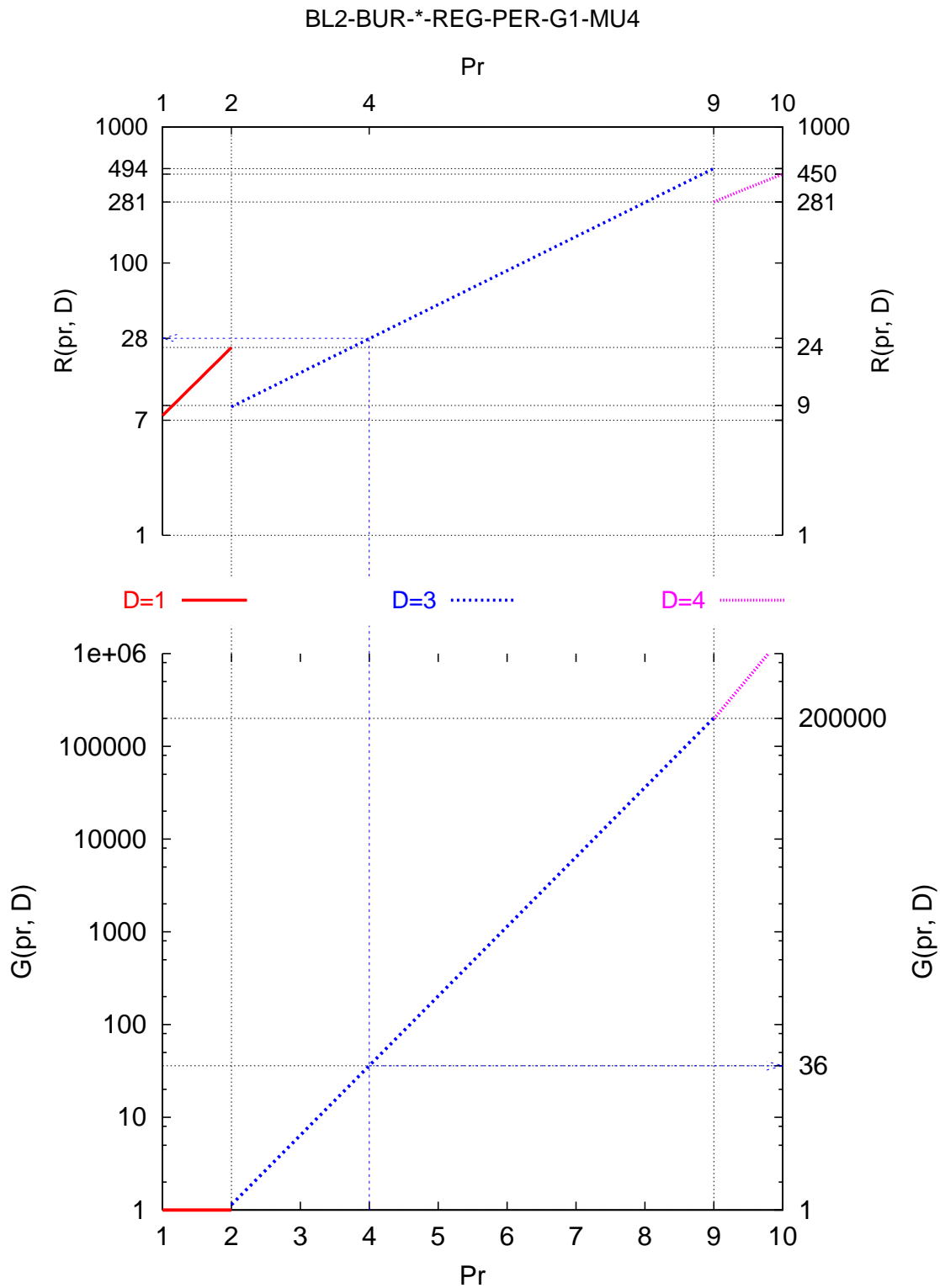


FIG.15: cas de Burgers, maillage régulier, conditions périodiques,  $\gamma = 1$ :  
 abscisse: erreur  $BL2=10^{-Pr}$ , ordonnée: fonction gain  $G$  et nombre de quadrangles  $R^2$ .  
 $10^{-2.0} \geq BL2 \geq 10^{-9.0}$ ,  $D = 3$  est préférable.  $D = 2$  n'est jamais optimal.  
 Ex:  $BL2 = 10^{-4} \Rightarrow D = 3$ : gain de 36 et maillage de  $28^2$  quadrangles!



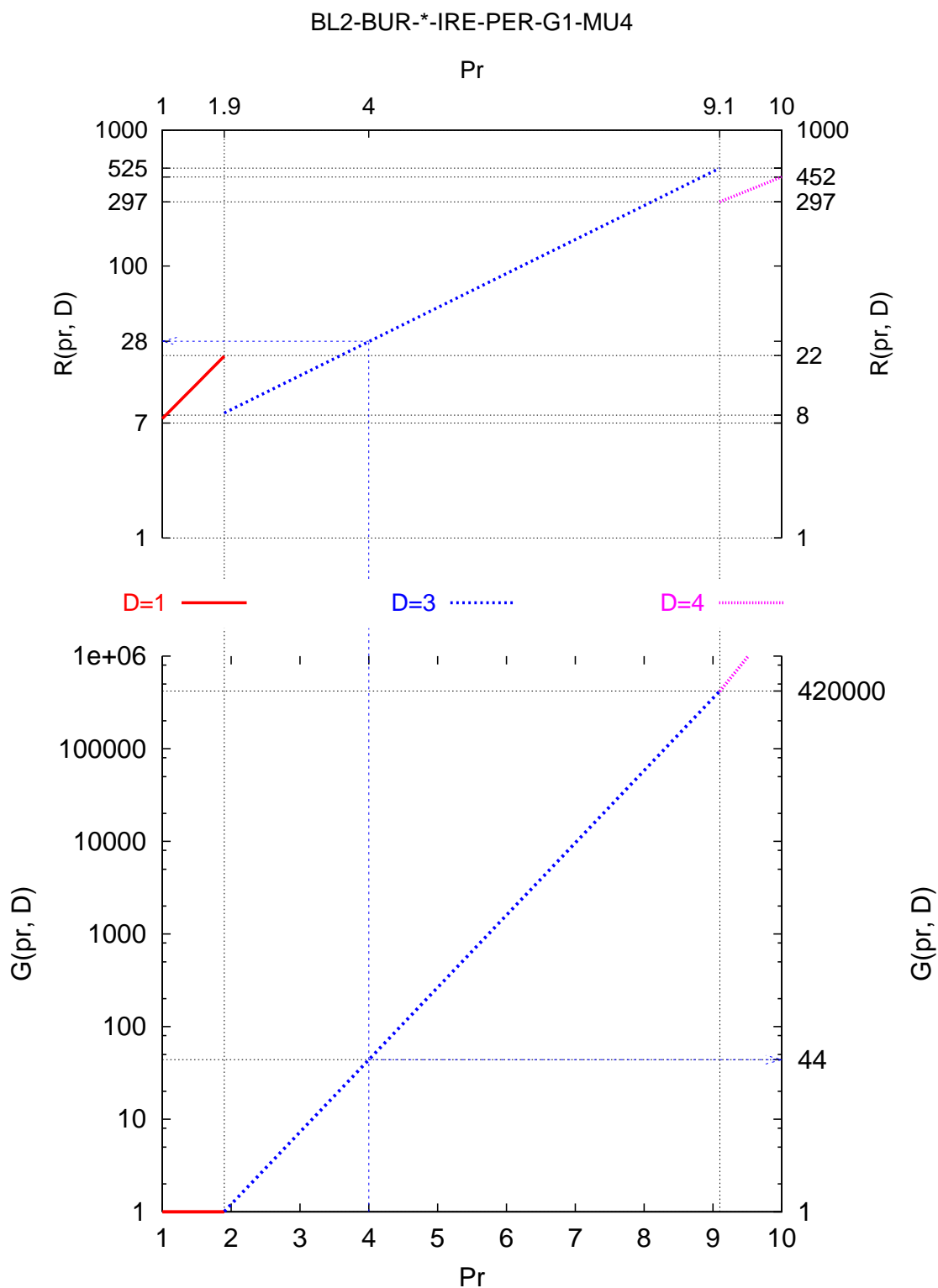


FIG.16: cas de Burgers, maillage irrégulier, conditions périodiques,  $\gamma = 1$ :  
 abscisse: erreur  $BL2=10^{-Pr}$ , ordonnée: fonction gain  $G$  et nombre de quadrangles  $R^2$ .  
 $10^{-1.9} \geq BL2 \geq 10^{-9.1}$ ,  $D = 3$  est préférable.  $D = 2$  n'est jamais optimal.  
 Ex:  $BL2 = 10^{-4} \Rightarrow D = 3$ : gain de 44 et maillage de  $28^2$  quadrangles!

## BL2-ADV-\*-REG-PER-G0-MU4

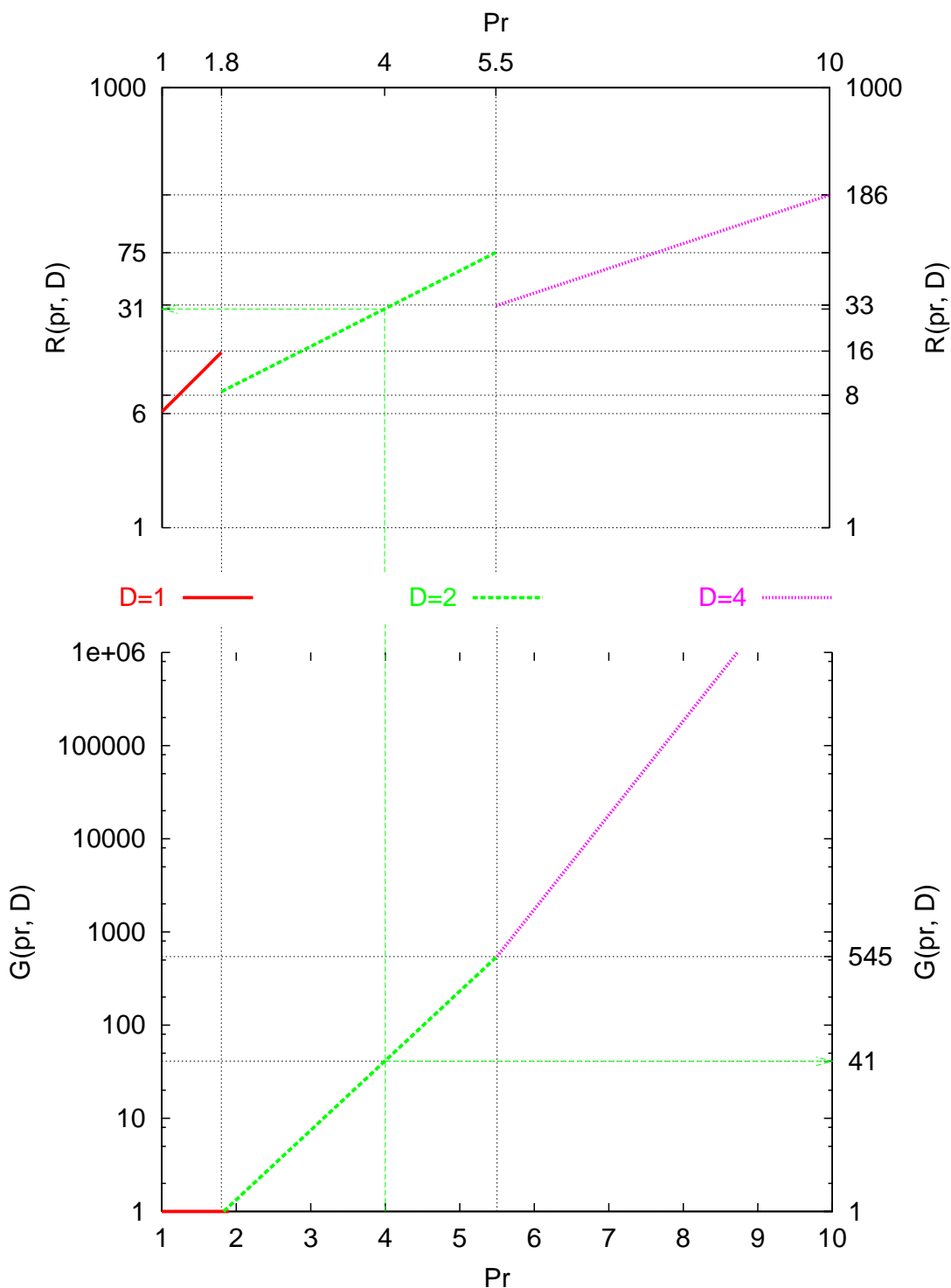


FIG.17: cas de l'advection, maillage régulier, conditions périodiques,  $\gamma = 0$ :  
 abscisse: erreur  $BL2=10^{-Pr}$ , ordonnée: fonction gain  $G$  et nombre de quadrangles  $R^2$ .  
 $10^{-1.8} \geq BL2 \geq 10^{-5.5} \Rightarrow D = 2$ ;  $10^{-5.5} \geq BL2 \Rightarrow D = 4$ ;  $D = 3$  n'est jamais optimal.  
 Ex:  $BL2 = 10^{-4} \Rightarrow D = 2$ : gain de 41 et maillage de  $31^2$  quadrangles!

En règle générale et si l'on se situe dans le cadre du schéma classique de Roe ( $\gamma = 1$ ), quel que soit le modèle choisi, linéaire ou non linéaire, sur maillage régulier ou non régulier, les figures 13 à 16 indiquent que:

- pour une précision jusqu'à la deuxième décimale, un degré  $D = 1$  suffit. Il est alors inutile d'utiliser des gants de boxe en choisissant un degré supérieur!
- pour une précision supérieure et jusqu'à la neuvième décimale alors le degré  $D = 3$  est privilégié.
- si une précision à la dixième décimale ou plus est souhaitée, alors l'on se confronte d'abord à des tailles de maillages avoisinant le million de quadrangles, ensuite à la précision machine...

Au bilan, le degré  $D = 2$  n'est jamais optimal (hormis le cas  $\gamma = 0$  FIG.17). Le degré  $D = 3$  est celui qui couvre la plage de haute précision réaliste  $[10^{-2}, 10^{-9}]$ . En pratique, pour le cas test C3 de l'ONERA, les maillages bidimensionnels envisagés allaient de 5000 à 30000 quadrangles ( $R < 200$ ). Avec les schémas d'ordre 2 des différents partenaires (PATRIC, THESEE, SIERRA...) et en utilisant la figure 10, celà signifie que l'on avait une précision au moins inférieure à  $10^{-2}$  mais que l'on ne pouvait pas atteindre une précision de  $10^{-4}$ ! Avec une reconstruction de degré  $D = 3$ , une telle précision est tout à fait envisageable et conduit aux gains et maillages ( $R^2$ ) suivants (en rouge):

Degré D	1	2	<b>3</b>	4
<b>BL2-ADV-*-REG-PER-G1-MU4</b>				
$G(4,D)$	1	4	<b>21</b>	5
$R(4,D)$	198	69	<b>28</b>	27
<b>BL2-ADV-*-IRE-PER-G1-MU4</b>				
$G(4,D)$	1	4	<b>23</b>	5
$R(4,D)$	205	70	<b>28</b>	27
<b>BL2-BUR-*-REG-PER-G1-MU4</b>				
$G(4,D)$	1	6	<b>36</b>	7
$R(4,D)$	240	71	<b>28</b>	27
<b>BL2-BUR-*-IRE-PER-G1-MU4</b>				
$G(4,D)$	1	7	<b>44</b>	9
$R(4,D)$	256	72	<b>28</b>	27
<b>BL2-ADV-*-REG-PER-G0-MU4</b>				
$G(4,D)$	1	<b>41</b>	<b>24</b>	17
$R(4,D)$	203	31	<b>27</b>	18

**TAB.6: Exemple réaliste avec  $pr = 4$  ( $\Leftrightarrow BL2 = 10^{-4}$ )!!!**

En dehors du cas pathologique de la pseudo super convergence (G0), le schéma utilisant une reconstruction polynômiale de degré  $D=3$ , c'est-à-dire, le schéma d'ordre 4, réalise au moins un gain de 20 par rapport à un schéma d'ordre 2. (Ce gain est vraiment minimal car nous n'avons pas pris en compte qu'il est usuel de doubler le CFL (le pas de temps) lorsque l'on

passé d'un RK2 pour  $D=1$  à un RK4 pour  $D=3$ )... En conséquence, nous prendrons pour la suite de notre étude:

$$D = 3 \text{ (O4), G1 } (\gamma = 1) \text{ et MU4 } (\mu_{0/l} = 1)$$

## 4 TRAITEMENT DES BORDS

Nous considérons maintenant que notre domaine de calcul n'est plus infini mais qu'il est donc délimité par des bords. La présence de ces bords implique deux modifications. D'une part, une molécule au voisinage d'un bord se trouve différente à une molécule en plein centre du maillage (*FIG.1*). D'autre part, à un bord correspond des conditions particulières qu'il faut introduire dans le calcul sans perdre de précision et le plus simplement possible...

### 4.1 Dégénérescence des molécules

La construction des molécules ou stencils au voisinage d'un bord doit répondre au cahier des charges suivant:

- une construction automatique. Nous ne souhaitons pas devoir définir à la main, selon la situation géographique d'un quadrangle  $Q_0$ , tous les  $NST_0$  quadrangles du stencil. Il nous faut une méthode générique de construction pour n'importe quel quadrangle  $Q_0$  sans avoir besoin d'étudier son positionnement par rapport aux bords du maillage. Par ailleurs, l'idée est de pouvoir étendre cet automatisme à des maillages non structurés.
- le stencil construit doit pouvoir permettre la définition du polynôme de degré 3. En d'autres termes, ce stencil doit être constitué d'un minimum de  $NST_0 = 9$  quadrangles autre que  $Q_0$  et laisser le système (16) inversible. Pour illustrer ce propos, nous donnons figure 18 un exemple trivial à ne pas suivre. En effet, si l'on a un grand nombre de valeurs pour une direction, on ne peut pas définir les coefficients du polynôme pour la direction orthogonale:

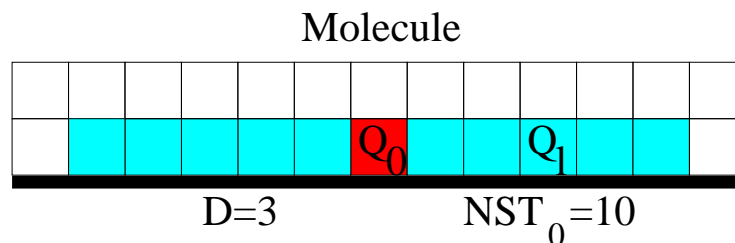


FIG.18: un exemple de mauvais choix de stencil.

Le cahier des charges étant posé, nous allons reprendre chaque item. Auparavant, une remarque s'impose: avec des conditions de bords, le nombre de quadrangles  $NST_0$  n'est plus constant comme cela l'était en conditions périodiques ( $NST = 12$ ). L'indice 0 prend donc tout son sens en indiquant par là qu'il réfère au quadrangle  $Q_0$  considéré. C'est pour cette raison qu'un tableau de  $NQ$  entiers à été alloué:  $NST(NQ)$ .

- La construction automatique

Définissons tout d'abord ce qu'est un voisin fort. Un voisin fort d'un quadrangle  $Q_0$  est un quadrangle qui partage un de ces côtés avec  $Q_0$ . Chaque quadrangle  $Q_0$  dispose donc d'un maximum de 4 voisins forts  $Q_{F1}, \dots, Q_{F4}$  et d'un minimum de 1 sinon  $Q_0$  serait le seul élément du maillage (FIG.19).

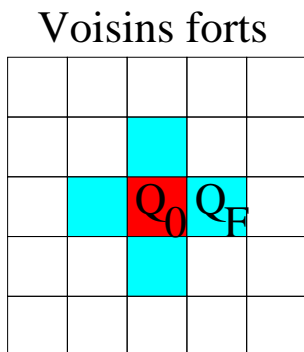


FIG.19: les voisins forts comme base de construction du stencil.

L'idée est alors de considérer les voisins forts des voisins forts de  $Q_0$  n'ayant pas encore été cités et ainsi de suite. Ceci, jusqu'à l'obtention d'un stencil de taille suffisante, c'est-à-dire, d'au moins 9 quadrangles dans le proche voisinage de  $Q_0$ . On s'aperçoit sur la figure 20 qu'en 2 passes, nous retrouvons le stencil considéré sur la figure 1 ( $D = 3$ ) lorsque l'on ne considère pas les bords.

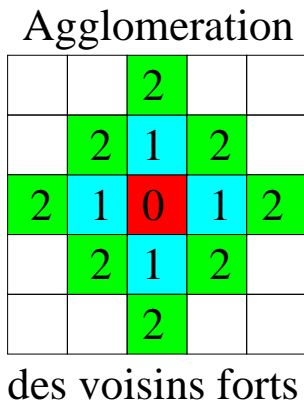


FIG.20: en 2 passes, le stencil standard ( $NST = 12$ ) est retrouvé.

Pour obtenir un stencil d'aspect le plus "centré" possible par rapport à  $Q_0$ , l'agglomération ne s'arrête pas au nombre de 9 quadrangles (nécessaires à la définition du polynôme) mais va jusqu'à la fin de la passe qui permet d'agglomérer au moins 9 quadrangles.

Appliquons maintenant cet automatisme à un quadrangle  $Q_0$  au voisinage d'un bord. Plusieurs cas de figures apparaissent selon la situation géographique de  $Q_0$  par rapport aux bords. Les figures (FIG.21) sur la page suivante représentent toutes les molécules observables, à une symétrie près, lorsque  $Q_0$  approche un angle du maillage.

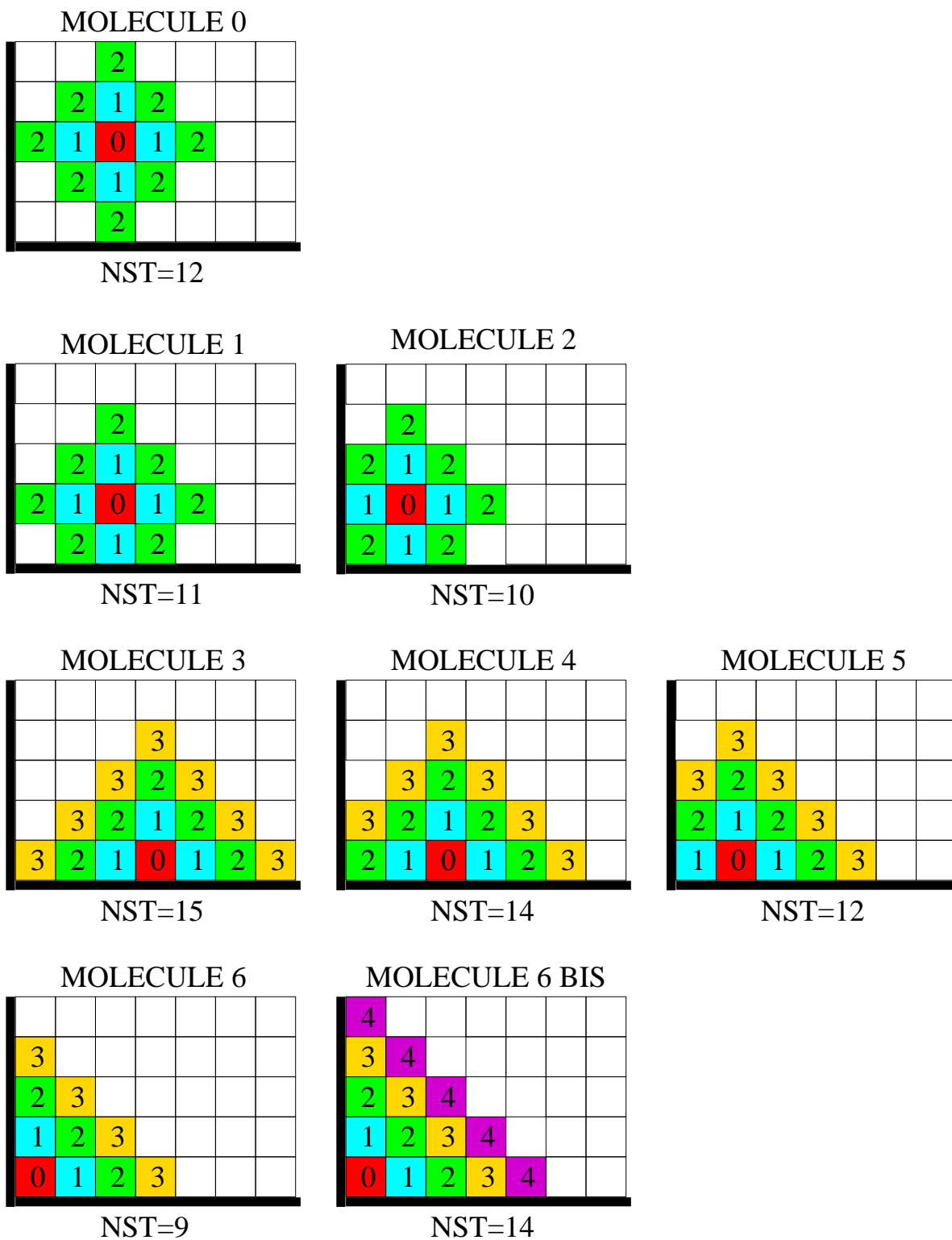


FIG.21: Les différentes molécules obtenues par agglomération.

Par cette approche, nous constatons que:

- si le quadrangle  $Q_0$  dispose d'une zone tampon de deux quadrangles par rapport aux bords, le stencil généré est la molécule 0. Ce stencil est donc la molécule la plus courante du maillage et correspond bien au stencil standard pour ( $D = 3$ ) composé de 12 quadrangles autres que  $Q_0$ .
- si le quadrangle  $Q_0$  dispose d'une zone tampon d'un seul quadrangle par rapport à l'un des bords, la molécule 0 dégénère en la molécule 1 voire 2 et passe d'un nombre de quadrangles autres que  $Q_0$  allant respectivement de 11 à 10. Ces valeurs étant supérieures à 9, une troisième passe n'est pas nécessaire.
- si le quadrangle  $Q_0$  est au contact d'un bord, deux passes ne suffisent plus. Une troisième est alors nécessaire augmentant brutalement le nombre de quadrangles autres que  $Q_0$ . Ce nombre va de 15 (molécule 3) pour ce réduire de nouveau jusqu'à 9 (molécule 6). Les molécules sont de plus en plus décentrées par rapport à  $Q_0$ .
- si le quadrangle  $Q_0$  est le quadrangle constituant le coin (molécule 6), le nombre de quadrangles autres que  $Q_0$  constituant le stencil est juste le nombre de quadrangles nécessaires à la définition du polynôme de degré 3. Il faut alors se rappeler que la construction polynômiale s'effectue par une méthode de moindres carrés. Ici, rien n'empêche d'utiliser cette méthode générale à toutes les molécules, cependant, vu le cas particulier, une résolution directe peut paraître préférable. Par ailleurs, il est très facile d'arrêter l'agglomération lorsque le nombre de quadrangles autres que  $Q_0$  est strictement supérieur à 9. Nous obtenons alors la molécule 6 bis constituée de 14 quadrangles.

En conclusion, cette méthode d'agglomération permet de générer des molécules automatiquement avec un nombre de quadrangles *a priori* suffisant à la définition d'un polynôme de degré 3. Il reste maintenant à vérifier que toutes les molécules envisagées permettent réellement d'inverser le système (16) et ne sont pas des molécules du genre de la figure 18...

- L'inversibilité de la matrice  $\mathcal{M}_0$

Avant d'inverser la matrice  $\mathcal{M}_0$ , il est fait usage de la routine DSPCON de la librairie NAG (voir section: Algorithme). Cette routine évalue l'indice de conditionnement de la matrice  $\mathcal{M}_0$  noté  $COND$  et le compare à la précision machine noté  $COMP$  d'une valeur ici de  $10^{-16}$ . Si  $COND > COMP$  alors la matrice est inversible mais elle le sera d'autant mieux que le rapport  $RCOND = \frac{COND}{COMP} \gg 1$ . En fait, l'idéal est d'avoir  $COND$  voisin de 1, c'est-à-dire, d'avoir  $RCOND$  proche de  $10^{16}$ .

Nous l'avons déjà dit mais le répétons: la matrice  $\mathcal{M}_0$  ne dépend que du maillage (raffinement, irrégularité) et du choix des molécules (de 0 à 6bis). Elle est indépendante du modèle mathématique (linéaire ou non linéaire) que l'on veut simuler.

Nous représentons le coin d'un maillage carré  $[0,1] \times [0,1]$  divisé en  $16 \times 16$  intervalles de manière régulière (FIG.22) ou irrégulière (FIG.23). Nous voulons mettre en évidence le comportement du rapport de conditionnement  $RCOND$  sur les molécules décrites sur la

figure 21. Auparavant, l'irrégularité du maillage apporte une fluctuation de  $RCOND$  qui donne à la figure 23 un aspect de "tôle ondulée".

MOLECULE	VALEURS DE $RCOND$ ( $\times 10^{10}$ )	
	REGULIER	IRREGULIER
0	1.0	$1.0 \pm 0.1$
1	1.0	$1.0 \pm 0.1$
2	1.4	$1.4 \pm 0.1$
3	1.3	$1.3 \pm 0.1$
4	1.4	$1.4 \pm 0.1$
5	0.4	0.4
6	0.07	0.07
6bis	0.25	0.23
6 (res. direct)	912.1	912.1

Sur un maillage  $16 \times 16$ , le niveau moyen de  $RCOND$  est de  $10^{10}$  (molécule 0). Lorsque l'on commence à décentrer le quadrangle  $Q_0$  au sein de la molécule, cela se passe plutôt bien, voire mieux que la molécule 0 pour la plus part d'entre elles (1 à 4). A l'approche du coin, la molécule 5 commence à poser problème ( $RCOND=0.4 \cdot 10^{10}$ ) mais ce problème est moindre par rapport à la molécule 6 ( $RCOND = 0.07 \cdot 10^{10}$ ) où nous avons tenté quelques améliorations. D'abord la molécule 6 est très particulière parce qu'elle est constituée du nombre de quadrangles ( $NST=9$ ) juste suffisant pour définir le polynôme. Le fait est que si l'on utilise la méthode des moindres carrés (mc) pour le définir, nous observons une très forte chute du rapport de conditionnement de la matrice ( $RCOND = 0.07 \cdot 10^{10}$ ). A partir de là, deux autres voies ont été tentées:

- résoudre le système directement (8), sans passer par la méthode des moindres carrés. Une très nette amélioration du rapport de conditionnement  $RCOND \sim 10^{13}$  est alors observée. Cependant, le fait d'utiliser une méthode directe oblige une résolution très spécifique et un stockage plus important ( $\mathcal{M}_0$  n'est plus symétrique). Par ailleurs, il n'y a plus d'effet de lissage et le polynôme peut donner lieu à de fortes oscillations.
- une extension de la molécule 6 à la molécule 6 bis obtenue par l'apport d'un quatrième niveau de quadrangles. La molécule passe alors de 9 à 14 quadrangles autres que  $Q_0$  et le rapport de conditionnement s'améliore un peu ( $RCOND = 0.25 \cdot 10^{10}$ ) par rapport au résultat obtenu sur la molécule 6 avec la méthode des moindres carrés.



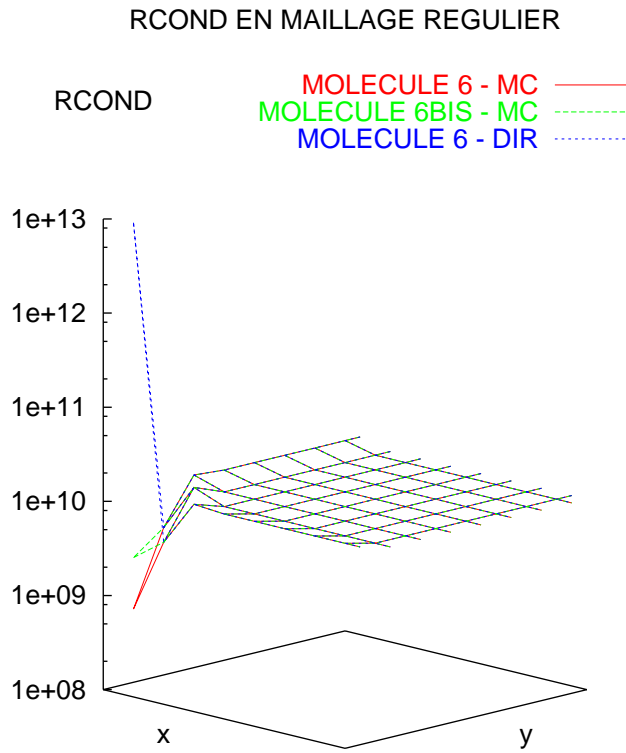


FIG.22: vue sur un coin du maillage régulier ( $16 \times 16$ ). Rapport du conditionnement de la matrice  $\mathcal{M}_0$

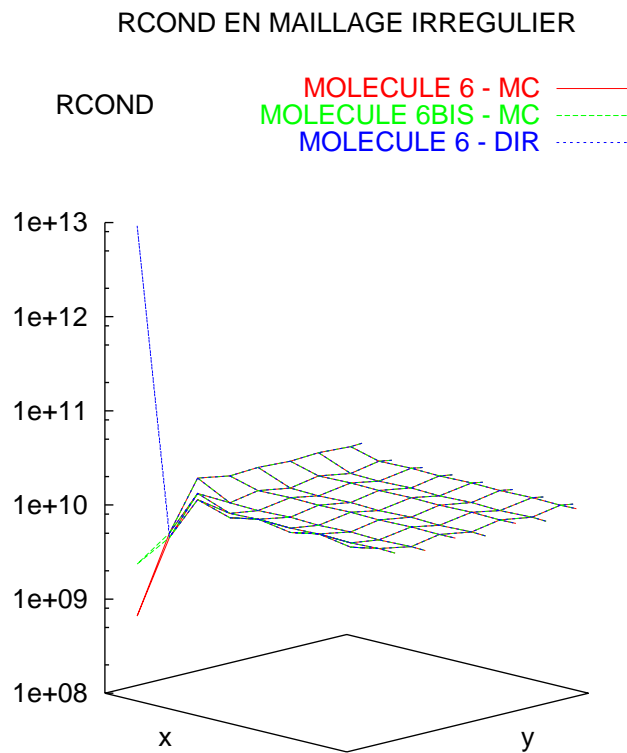


FIG.23: vue sur un coin du maillage irrégulier ( $16 \times 16$ ). Rapport du conditionnement de la matrice  $\mathcal{M}_0$

## 4.2 Implémentations des conditions

Pour les premiers cas tests scalaires, nous utilisons une condition de Dirichlet. En pratique, nous introduisons cette condition de bord de manière faible. En d'autres termes, nous n'imposons pas la valeur de bord comme devant être respectée par le polynôme associé au quadrangle de bord en tous les points ou en valeur moyenne; nous utilisons la valeur de bord par voie de flux. Pour cela, nous procédons comme pour le calcul d'un flux de Roe classique ( $\gamma = 1$ ) entre 2 quadrangles voisins (35) où  $U_{I_g}^n$  est remplacé par la valeur exacte et connue au point de Gauss  $(x_g, y_g)$  du bord notée  $U_{b_g}^n$ .

$$\varphi(U_{0_g}^n, U_{b_g}^n, \vec{N}) = \frac{1}{2} \left\{ \mathcal{F}(U_{0_g}^n, \vec{N}) + \mathcal{F}(U_{b_g}^n, \vec{N}) + \left| \frac{\partial \mathcal{F}(\tilde{U}_g)}{\partial U} \right| (U_{0_g}^n - U_{b_g}^n) \right\}$$

Il ne s'agit donc pas de véritable conditions de bord au sens où l'on va modifier l'expression du flux (voir condition de glissement ou de débit pour Euler) mais d'une extension du solveur ayant pour propriété de sélectionner lui même les caractéristiques entrantes ou sortantes. Ainsi et pour ce type de conditions entrantes ou sortantes, nous allons vérifier l'ordre du schéma en prenant en compte les molécules dégénérées.

## 4.3 Vérifications

Nous reprenons les expériences précédentes sur un mode (1,1) de Fourier avec  $c = 1$  et  $\theta = \frac{\pi}{4}$  (40) en considérant toujours l'erreur BL2, l'ordre 4 (O4), le schéma de Roe classique (G1) et la pondération MU4. Les calculs effectués portent:

- sur le modèle linéaire (advection) ou non (Burgers)
- sur l'irrégularité du maillage
- sur le choix de la molécule de coin allant de 9 (mol. 6) à 14 (mol. 6bis) quadrangles.

Nous constatons sur la figure 24 qu'indépendamment du modèle, de l'irrégularité et du choix de la molécule de coin, l'erreur BL2 commise en condition non périodique par rapport aux conditions périodiques est plus importante. Néanmoins, nous constatons que l'ordre est toujours préservé. En toutes circonstances, et sur les petits maillages, c'est-à-dire, là où l'effet de bord se fait le plus sentir, nous notons que le choix de la molécule 6 est très sensiblement préférable à la molécule 6bis contenant plus de quadrangles. Ce dernier point étant contraire à ce que nous espérons au regard de l'indice de conditionnement, nous garderons la molécule 6bis en justifiant ce choix (malheureux puisque plus coûteux), par le fait qu'il autorise un meilleur lissage dans l'interpolation lorsque la solution est oscillante.

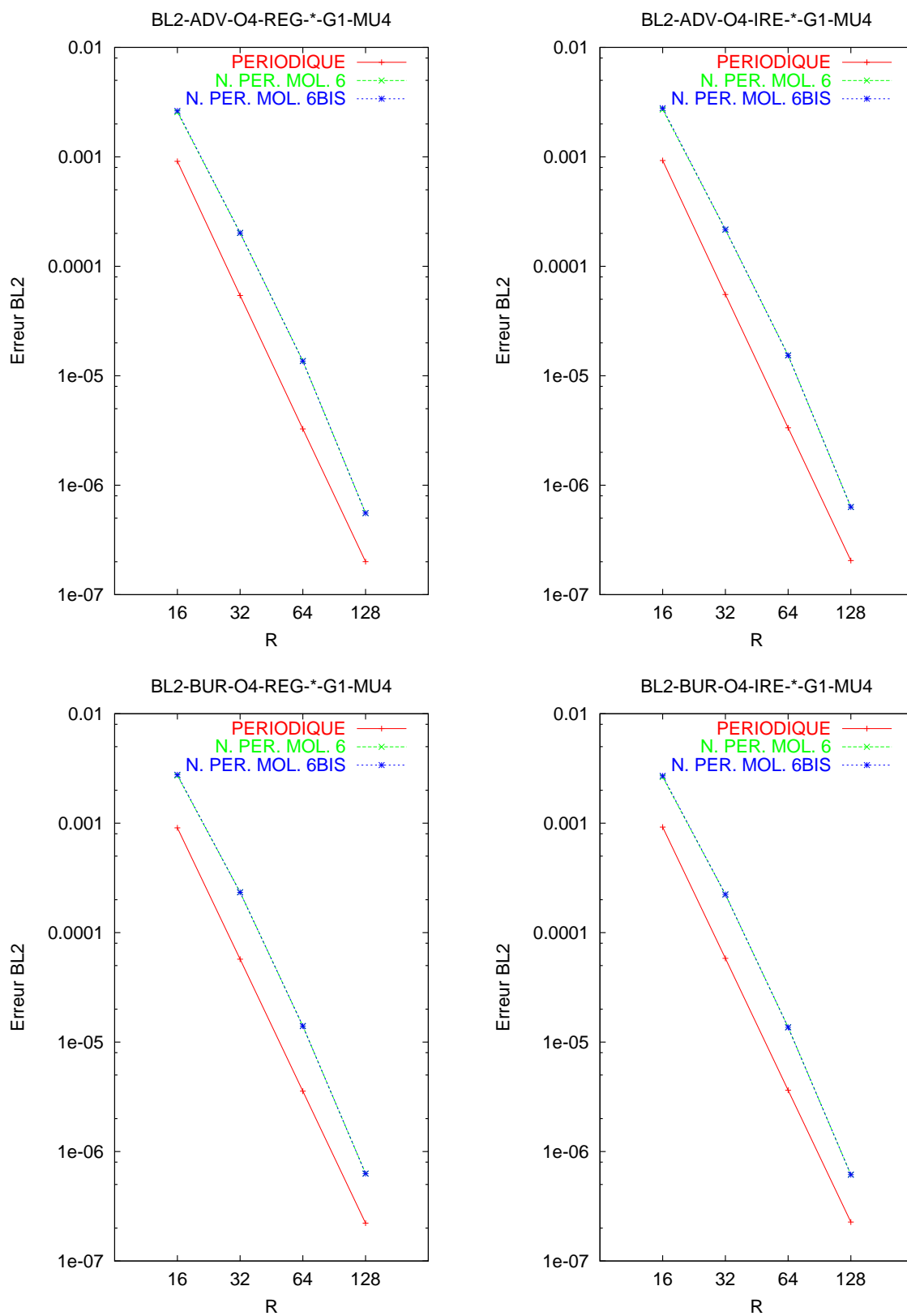


FIG.24: Conditions de bord = erreurs + grande mais ordre préservé.

## 5 CAS TEST EULER

Il est temps maintenant de passer au système que représente les équations Euler bidimensionnelles. Il convient donc d'établir non plus un polynôme mais quatre afin de reconstruire sur chaque quadrangle les valeurs conservatives:

$$\rho \text{ densité, } (\rho u, \rho v) \text{ quantités de mouvements et } E \text{ l'énergie.}$$

Cependant rappelons que ces reconstructions s'effectuent, pour chaque quadrangle, à partir de la même matrice  $\mathcal{M}_0$  (16).

Nous poursuivons la validation du schéma à travers de multiples exemples où nous prenons en compte toutes les remarques faites précédemment, à savoir:

- une reconstruction de degré 3 ( $D = 3$ ) devant donner au schéma une précision d'ordre 4 ( $O4$ ),
- pas de pondération lors de l'ajustement par la méthode des moindres carrés ( $MU4$ ),
- pas de pondération de la matrice de Roe ( $G1$ ).
- le stencil de coin contient 14 quadrangles autres que  $Q_0$  (MOLECULE 6 BIS).

### 5.1 Onde de densité

En tout premier lieu et sur un exemple très simple, nous souhaitons vérifier que l'ordre 4 du schéma est maintenu en utilisant les variables conservatives. Cette vérification s'effectue en réalisant une mesure d'erreur BL2 sur la densité pour un problème dont nous connaissons la solution exacte sur un maillage carré  $[0,1] \times [0,1]$ .

$$\begin{aligned} \rho &= 1 + \frac{1}{10} \left\{ \cos\{2\pi(x + y - \sqrt{2}t)\} + \sin\{2\pi(x + y - \sqrt{2}t)\} \right\} \\ u &= \frac{\sqrt{2}}{2} \\ v &= \frac{\sqrt{2}}{2} \\ P &= (\gamma - 1) \left\{ E - \frac{\rho}{2}(u^2 + v^2) \right\} = \frac{100}{\gamma} \quad \text{où } \gamma = 1.4 \end{aligned}$$

Nous utilisons les conditions de bord toujours par extension du solveur de Roe. Nous réalisons des calculs sur divers raffinements ( $R$ ) de maillages réguliers et irréguliers.

Sur la figure 25, nous constatons que l'ordre 4 est validé sur le système des équations d'Euler indépendamment de la nature du maillage et de la prise en compte des bords pour cet exemple très simple.

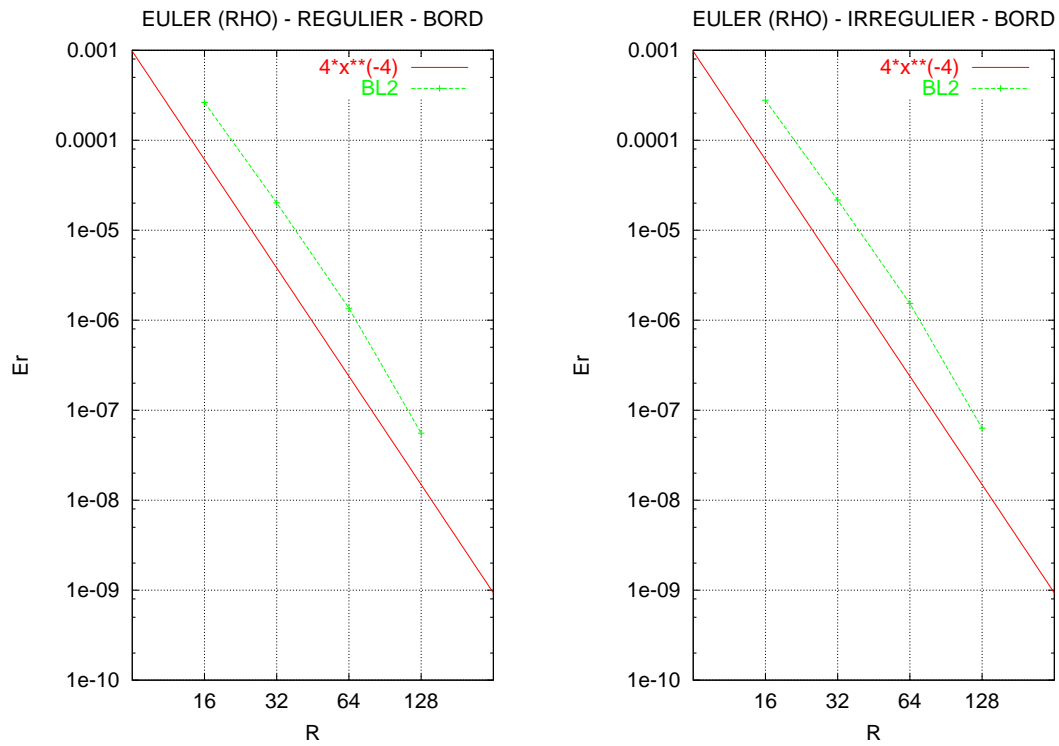


FIG.25: ordre 4 validé pour le système Euler.

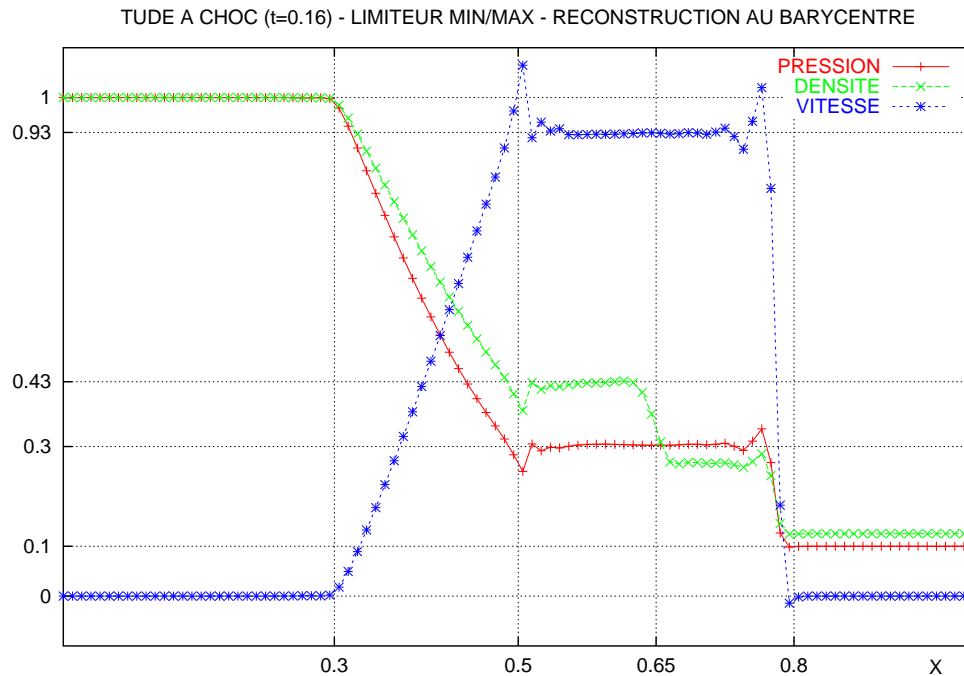


FIG.26: beaucoup d'oscillations, perte de précision mais niveaux usuels retrouvés.

## 5.2 Tube à choc

A des fins de validation, nous réalisons sur la figure 26, le cas du tube à choc de Sod 1D pour lequel nous prenons un maillage  $100 \times 10$  pour pouvoir définir les molécules bidimensionnelles. L'initialisation est donnée par:

$$\begin{cases} \rho = 1 & \rho u = 0 & \rho v = 0 & E = 2.5 & 0 \leq x \leq 0.5 \\ \rho = 0.125 & \rho u = 0 & \rho v = 0 & E = 0.25 & 0.5 \leq x \leq 1 \end{cases}$$

D'ores et déjà, nous savons que ce cas test va engendrer de très fortes oscillations du fait qu'il présente de fortes pentes (choc, discontinuité de contact, détente). Dans la mesure où les stencils sont immobiles et à l'inverse des schémas de ENO où un choix du stencil est effectué, nous ne pouvons atténuer ces oscillations que par l'introduction de limiteurs. D'ailleurs, le calcul ne passe pas si l'on n'introduit pas de limiteurs adhoc. Ceux que nous avons utilisés ne permettent pas d'éliminer néanmoins les overshoots. La mise au point de limiteurs est un sujet en soi, assez éloigné du contexte du couplage aéroacoustique, but que nous visons à court terme. Au bilan de la figure 26, nous pouvons considérer que ce type de schéma ne convient pas à des calculs présentant des chocs (ce qui avait été annoncé dès l'introduction) mais que la programmation du système des équations d'Euler est correct dans la mesure où les niveaux usuels des variables sont retrouvés.

## 5.3 Ecoulement stationnaire

On se propose de simuler l'écoulement d'un fluide passant d'un état sonique à supersonique à travers une tuyère. Le maillage structuré que nous créons pour cette expérience est réalisé de  $27 \times 8$  quadrangles et constitue donc un maillage très grossier de la demi tuyère (*FIG.27*). Dans la mesure où nous ne considérons à l'heure actuelle que le système des équations d'Euler, le bord pariétal et l'axe de symétrie portent une condition de glissement. En amont de l'écoulement, c'est à dire sur le côté gauche de la figure 27, le bord est caractérisé par des conditions entrantes tandis qu'à droite les conditions sont sortantes. Comme précédemment, nous n'imposons pas de valeurs sur les bords mais agissons par le calcul de flux de Roe estimés aux moyens des valeurs internes et des valeurs externes. Ces valeurs externes sont pour les bords:

- glissants: condition miroir sur la vitesse,
- entrants: nombre de Mach à 0.5, pression atmosphérique,
- sortants: sortie supersonique libre.

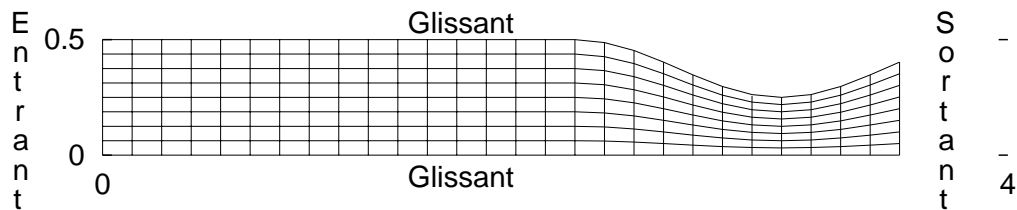
Nous prenons pour condition initiale: une densité constante, la pression atmosphérique, une vitesse de composante verticale nulle et horizontale respectant le nombre de Mach de 0,5.

Avant de présenter les résultats, il faut souligner les points suivants:

- les bords étant définis de manière très grossière, nous avons été obligés de maintenir un ordre 1 pour le calcul des flux de Roe sur ces bords. Si nous ne réalisons pas cette correction, nous observons une instabilité verticale qui s'amplifie au cours du temps et rend le calcul stationnaire impossible.

- nous avons dû modifier le solveur de Roe pour faire disparaître un choc entropique qui n'avait rien de physique.
- enfin, le logiciel graphique utilisé réalise une interpolation linéaire avec le faible nombre de valeurs données. En conséquence, dans les représentations graphiques ci-après, la connaissance de la solution sous forme polynômiale haute précision n'est pas exploitée.

**MAILLAGE DE LA DEMI TUYERE  
27 X 8 QUADRANGLES**



**FIG.27: maillage et conditions de bords**

Nous présentons 3 calculs sur ce maillage en utilisant toujours les stencils utiles à l'ordre 4 et donc les matrices associées pour définir les polynômes d'ordre 4. Ces trois calculs sont différents en ordre de précision parce que nous limitons la reconstruction polynômiale à l'ordre 1 (valeur moyenne) pour le premier, à l'ordre 2 (NINC=2) pour le second et à l'ordre 4 (NINC=9) pour le troisième calcul.

Sur la FIG.28, nous constatons d'abord une stationnarité de la solution en moins de 5000 itérations et une amplitude du nombre de Mach plus importante avec l'ordre du schéma.

Sur la FIG.29, où nous représentons de mêmes isovaleurs sur un zoom du col de la tuyère des solutions obtenues pour les différents ordres, nous constatons:

- une cassure des isovaleurs pour les ordres supérieurs à 1 qui s'explique par les conditions de bords bridées à l'ordre 1 pour des raisons d'instabilités verticales comme nous l'avons déjà mentionné.
- un resserrement des isovaleurs avec un ordre croissant de précision. Nous soulignons encore que l'interpolation de ces isovaleurs sont linéaires et que ce resserrement pourrait être encore plus important si le logiciel graphique utilisait pleinement les solutions trouvées sous leur forme polynômiale et non quelques valeurs ponctuelles du maillage grossier.

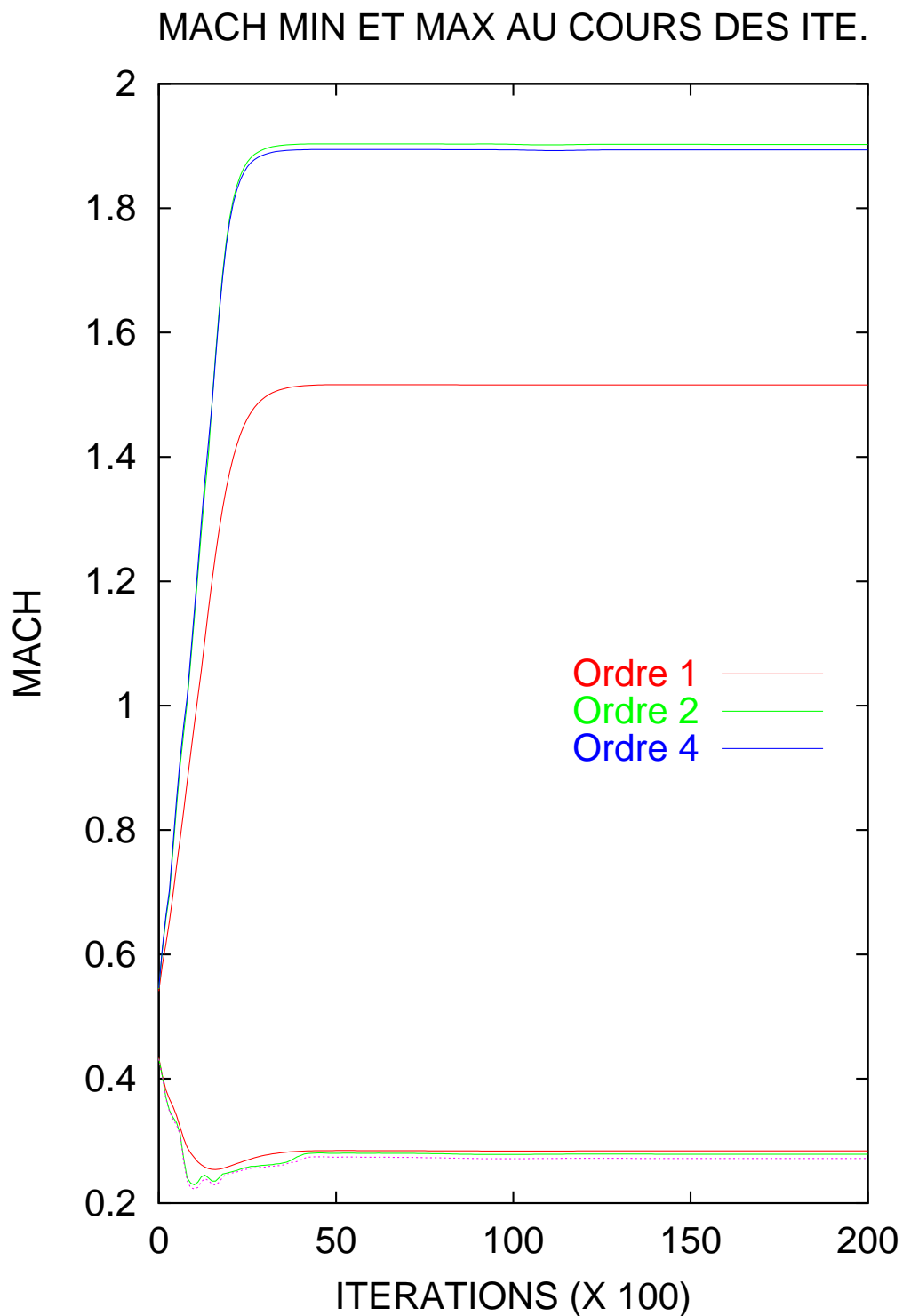
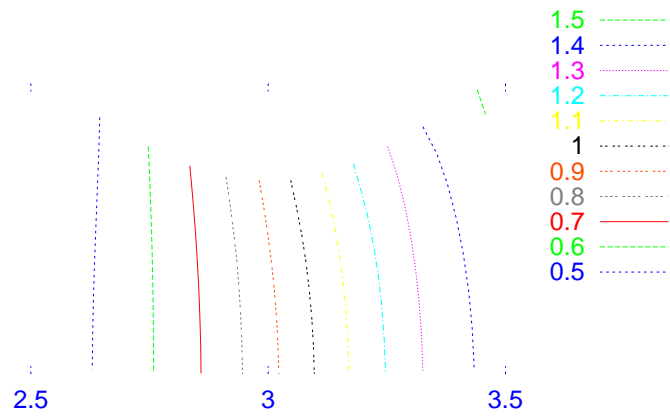


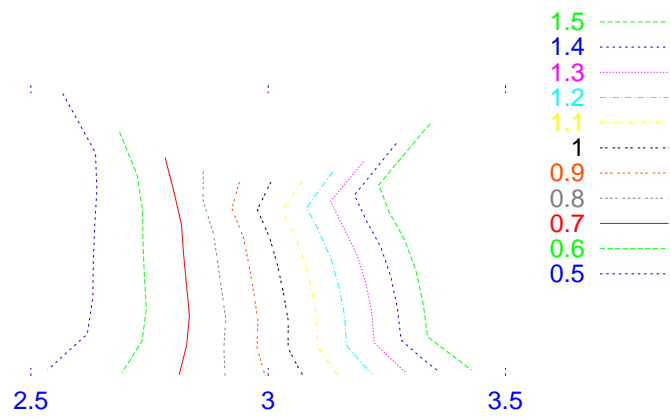
FIG.28: évolution du nombre de Mach au cours des itérations.



SCHEMA BRIDE A L ORDRE 1 / ISO-MACH



SCHEMA BRIDE A L ORDRE 2 / ISO-MACH



SCHEMA A L ORDRE 4 / ISO-MACH

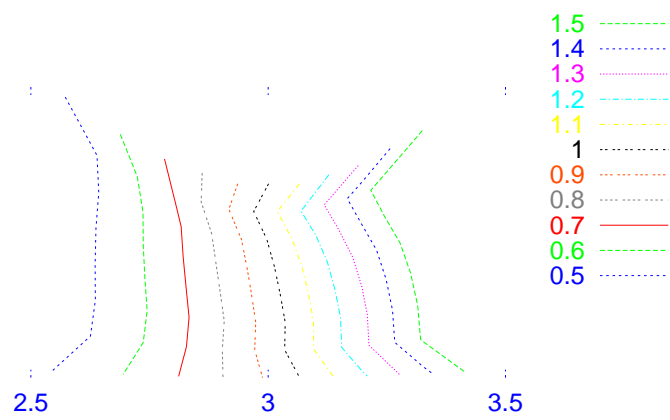


FIG.29: lignes iso-mach sur le col de la tuyère

## 6 CONCLUSIONS ET PERSPECTIVES

Il est possible d'obtenir un schéma de très haute précision applicable à la recherche de solutions régulières (pas de chocs) pour modèle linéaire ou non et sur un maillage quelconque pour un coût nettement inférieur à ENO (ou WENO). Cela tient essentiellement au fait de considérer pour chaque quadrangle une molécule fixe au cours du temps. Par ce biais, un prétraitement d'inversion des matrices dépendantes uniquement du maillage peut être effectué autorisant une reconstruction polynômiale rapide tandis que l'usage de méthodes par les moindres carrés génère des matrices symétriques permettant un moindre coût de stockage informatique.

Tout au long de cet écrit, nous avons pris garde de préserver au maximum des options de calculs sur lesquels nous pouvons maintenant et après étude nous positionner:

- L'ordre:  
l'ordre 4, c'est-à-dire, une reconstruction polynômiale de degré 3 est celui qui assure une bonne précision pour tous les modèles numériques envisagés avec un moindre coût. Un gain de 20 à 40 est obtenu par rapport à l'ordre 2 (FIG.13 à 16).
- La pondération de la viscosité numérique:  
le paramètre  $\gamma$  qui pondère la matrice de Roe a été envisagé de manière booléenne (0 ou 1). En supprimant la matrice de Roe ( $\gamma = 0$ ), les calculs ne peuvent être effectués que pour un modèle linéaire (instabilités sinon) et un semblant de super convergence apparaît uniquement sur maillage régulier mais pas pour l'ordre 4. De fait, nous conseillons ici, l'usage du schéma de Roe classique ( $\gamma = 1$ ).
- La pondération de la méthode des moindres carrés:  
dans le cadre d'un maillage irrégulier, il pouvait paraître raisonnable de prendre en compte les différences d'aspect de chaque quadrangle constituant le stencil. Ceci a été traduit par le paramètre  $\mu$  sous 4 déclinaisons. Au bilan, ne pas faire de pondération ( $\mu = 1$ ) semble être l'option optimale ne serait ce que sur l'économie de l'espace mémoire qu'utilise une telle pondération.

Par ailleurs, nous avons vérifié que l'ordre 4 était toujours préservé pour n'importe quel maillage (régulier ou irrégulier) et pour n'importe quel modèle (linéaire ou non linéaire). Cependant, nous notons une dégradation de la précision dès lors que l'on ne considère plus des bords périodiques. Ce point peut s'expliquer d'une part par la dégénérescence des stencils sur les bords et particulièrement sur les coins où le conditionnement de la matrice devient plus mauvais qu'ailleurs, d'autre part dans l'expression des conditions de bords que nous essayons de faire intervenir de manière faible, c'est-à-dire par calcul de flux de Roe. Il reste donc encore à cet égard quelques innovations à étudier de manière à ne pas être obligé de restreindre le calcul de l'écoulement stationnaire à un ordre 1 sur les bords... Pour cela, j'imagine trois pistes à suivre:

- Lorsque le maillage est grossier et qu'il remplace des bords courbes par des segments de droites (col de la tuyère), il y a une perte de précision lors de l'intégration par les points de Gauss. Il faudrait pouvoir donner la courbure du bord et avoir une intégration spécifique.

- Dans la mesure où les conditions de bords sont introduites par des valeurs extérieures au domaine, il est envisageable d'inclure dans les stencils de bords des quadrangles virtuels contenant ces valeurs extérieures et laissant ainsi une molécule où le quadrangle considéré sera moins décentré.
- Nous pouvons essayer d'obtenir un meilleur conditionnement des matrices en changeant de base polynômiale. Par exemple, les polynômes de Tchebychev tout en continuant à garder des expressions locales (référenciel barycentrique), en préservant la valeur moyenne comme terme constant et en utilisant toujours une méthode de moindres carrés.

Enfin, il ne faut pas oublier que les solutions obtenues par ce type de schémas s'expriment sous forme polynômiale et non en valeurs ponctuelles. Même si le maillage est pauvre en nombre de points, il est possible de connaître en n'importe quel lieu du domaine la valeur de(s) inconnue(s) à l'ordre de précision du schéma. C'est la "k-exactitude". Pour en profiter, il est nécessaire de mettre au point un logiciel graphique pouvant grâce à la connaissance des solutions polynômiales projeter aussi finement que souhaité les valeurs obtenues. Ce n'est que grâce à ce type de logiciel que ce nouveau schéma sera en mesure de montrer son avancée.

Dans les perspectives de plus en plus hypothétiques pour moi lorsque j'écris ces lignes, j'ajouterais que:

- l'approche non structuré du maillage n'est pas un réel problème dans la mesure où le code informatique actuel considère déjà des quadrangles numérotés sans avoir besoin de faire référence à ses coordonnées structurées.
- l'approche multi-éléments (triangles et quadrangles) est susceptible de poser des problèmes pour la construction des stencils compacts.
- inclure au schéma Euler actuel les termes de viscosité adéquates au système Navier Stokes pose deux problèmes:
  - estimer les dérivées secondes des valeurs par la dérivée seconde des polynômes préserve-t-il l'ordre 4 du schéma?
  - la mise au point des conditions de bords.

Après quoi, la programmation axisymétrique ayant été déjà introduite dans [10], ce type de schéma pourrait venir apporter une très haute précision à des codes numériques comme CPS.

---

## Références

- [1] B. COCKBURN, W. SHU, "*The Runge-Kutta discontinuous Galerkin method for convection laws V: multidimensional systems*", J. Comp. Phys, v141 (1998), pp.199-224.
- [2] T. BARTH, P FREDERICKSON, "*High order solution of the Euler equations on unstructured grids using quadratic reconstruction*", AIAA Paper No 90-0013.
- [3] B. VAN LEER, "*Computational methods for ideal compressible flow*", Cours van Karmann Institute, lecture series 1983-04, Comp. fluid dynamic (1983).
- [4] R. ABGRALL, "*On essentially non-oscillatory schemes on unstructured meshes: analysis and implementation*", J. Comp. Phys., v114 (1994), pp.45-58.
- [5] X-D. LIU, S. OSHER, T. CHAN, "*Weighted essentially non-oscillatory schemes*", J. Comp. Phys., v115 (1994), pp.200-212.
- [6] NAG Fortran Library Manual.
- [7] P.L. ROE, "*Approximate Riemann Solvers, Parameters vectors and Difference Schemes*", J. Comp. Phys., v43 (1981), p.357-371.
- [8] R. CARPENTIER, "*Comparaison entre des schémas 2D de type Roe sur maillage régulier triangle ou quadrangle. II: calcul au sommet - Le  $\beta\gamma$  schéma*", Rapport INRIA No 3360 (1998).
- [9] A. JAMESON, "*Numerical solution of Euler equations for compressible inviscid fluids in numerical methods for the Euler equations of fluid dynamics.*", SIAM Philadelphia (1985), pp.199-231.
- [10] R. CARPENTIER, "*Le  $\beta\gamma$  schéma en axisymétrique*", Rapport INRIA No 3676 (1999).



---

Unité de recherche INRIA Sophia Antipolis  
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique  
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

---

Éditeur  
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)  
<http://www.inria.fr>  
ISSN 0249-6399