

Partie II

Méthodes numériques

3	Les approches numériques en mécanique des fluides	29
4	Méthodes aux différences finies	37
5	Méthodes spectrales	59

3

Les approches numériques en mécanique des fluides

« Informatique : Alliance d'une science
inexacte et d'une activité humaine faillible »

— L. Fayard

Dictionnaire impertinent des branchés

Plan du chapitre

1. Bases de l'intégration numérique	29
§ 11. Introduction	29
§ 12. Équation modèle	30
2. Méthodes à discrétisation spatiale	31
§ 13. Approche des différences finies	31
§ 14. Approche des volumes finis	32
§ 15. Approche particulière	33
3. Approche spectrale	34
4. Quelles méthodes pour quelles simulations?	35

1. Bases de l'intégration numérique

§ 11. Introduction

LES MÉTHODES NUMÉRIQUES ONT ÉTÉ UTILISÉES DEPUIS L'ANTIQUITÉ pour l'évaluation de quantités utilisées dans la vie courante (dans l'architecture notamment). Citons par exemple l'évaluation du rapport de la diagonale du carré par rapport à son côté par les babyloniens 2000 ans avant notre ère. On le voit, la mesure de certaines quantités mathématiques *a priori* abstraites a été un des premiers problèmes de l'Homme. Ces méthodes ont par la suite été formalisées pour traiter certains problèmes précis (Par exemple les méthodes de Newton ou d'Euler). L'avènement de l'informatique à partir de 1950 a donné une nouvelle dimension à cette approche et a permis de traiter des problèmes de plus en plus complexes grâce à l'accroissement rapide de la puissance de calcul disponible.

En termes modernes, le but de ces méthodes est d'évaluer numériquement une quantité qui est solution d'une équation plus ou moins complexe. Les domaines d'applications sont

extrêmement larges et vont de la dynamique des populations humaines aux marchés financiers en passant par la physique subatomique. De manière générale, les approches numériques permettent de résoudre *de manière approchée* des problèmes où l'approche analytique échoue. On notera cependant que, vu le nombre de problèmes auxquels peuvent s'appliquer ces méthodes, il est important de cibler précisément le problème posé, et de connaître en partie la solution que l'on pense obtenir. En effet, une méthode peut exceller dans un domaine donné et être de bien piètre qualité dans le domaine dans lequel on veut trouver une solution.

Dans notre cas, le problème posé est essentiellement un problème de dynamique des fluides, éventuellement magnétisés. Dans la suite de ce chapitre je tenterai de décrire différentes méthodes numériques utilisées en physique pour calculer la dynamique de ce type d'écoulements. Je mettrai en particulier l'accent sur les possibilités et les limites de ces méthodes et je montrerai en quoi les choix effectués pour ce travail sont pertinents.

§ 12. Équation modèle

Notons $Q(x_i, t)$ le vecteur à n composantes décrivant le fluide aux coordonnées x_i et à l'instant t . On peut alors décrire l'évolution du fluide par un ensemble d'équations tels que :

$$\partial_t Q(x_i, t) = G(Q(x_i, t)) \quad (12.62)$$

Ce problème n'est complet que lorsque l'on a défini un état initial $Q(x_i, t_0) = Q_0(x_i)$ et des conditions aux limites à tout instant t . La fonction G introduite dans (12.62) ne dépend que de $Q(x_i, t)$ et de ses dérivées spatiales, mais peut faire intervenir des termes non linéaires. Remarquons de plus que l'écriture de l'équation (12.62) est donnée en variables eulériennes, ce qui suppose que l'on ne suivra pas numériquement les particules fluides. En fait, certaines formulations peuvent faire appel à des variables lagrangiennes (voir par exemple § 15). Cependant, ce type d'approche ne modifie pas fondamentalement le principe de l'intégration numérique que je présente ici.

On commence tout d'abord par discrétiser l'opérateur d'évolution temporelle. En pratique, on fera évoluer d'un pas de temps dt la solution $Q(x_i, t)$ vers $Q(x_i, t + dt)$ en évaluant le terme $G(Q(x_i, t))$:

$$Q(x_i, t + dt) = Q(x_i, t) + \mathcal{H}(G(Q), dt) \quad (12.63)$$

\mathcal{H} dépend de la méthode d'intégration temporelle¹⁰ utilisée, de la précision recherchée et de considérations de stabilités. Ces points seront discutés pour les méthodes utilisées dans cette thèse par la suite.

La deuxième étape dans l'établissement d'un schéma d'intégration est la discrétisation de l'opérateur de différenciation spatiale G . Comme je le montrerai, la différenciation spatiale peut revêtir différentes formes. Cette discrétisation dépend de la géométrie du problème, de la précision voulue, des conditions aux limites et de la physique mise en jeu dans les solutions recherchées.

Pour finir, notons que la séparation temporelle/spatiale présentée ici par soucis de simplification n'est pas forcément évidente en pratique. Certains schémas, tels que les volumes

¹⁰Nous n'utiliserons ici que des méthodes d'intégration temporelles explicites, ce qui nous permettra d'écrire (12.63). Il existe néanmoins des méthodes implicites, que l'on écrira sous la forme générale $\mathcal{H}(Q_{t+dt}, Q_t) = 0$

finis, font ainsi appel à une intégration mixte des termes temporels et spatiaux. Cependant, pour mon étude, cette séparation me sera très utile et me permettra de mettre en évidence les origines de problèmes rencontrés dans plusieurs schémas d'intégration.

Poursuivant la logique présentée ici, je vais séparer les méthodes de discrétisation de l'opérateur G en deux grandes classes : les méthodes spatiales et les méthodes spectrales.

2. Méthodes à discrétisation spatiale

L'approche la plus simple pour discrétiser un problème est de décomposer la fonction $Q(x_i, t)$ sur un ensemble de points de contrôle répartis dans l'espace. On cherche alors à décrire l'évolution de la solution en chacun des points.

§ 13. Approche des différences finies

L'approche aux différences finies est de loin la plus intuitive : on utilise une grille fixe pour discrétiser l'espace, et la solution est évaluée en chacun des points de grille. On calcule alors l'évolution de la solution en utilisant une forme discrète de l'équation (12.62) que l'on écrira sous la forme :

$$\partial_t Q_i(t) = \tilde{G}(Q_{i-k}(t), \dots, Q_{i+l}(t)) \quad (13.64)$$

où \tilde{G} est l'opérateur différences finies équivalent à G . \tilde{G} dépend de la solution Q en chacun des point x_j , notée Q_j . Plus concrètement, le passage en différences finies se fait en remplaçant les opérateurs de différentiation spatiale par leurs équivalents en différences finies, qui sont des fonctions linéaires des solutions Q_j au voisinage du point que l'on considère. On pourra alors écrire de manière générale :

$$\frac{\partial^{(p)} Q(x_i, t)}{\partial x^p} = \sum_{j=i-k}^{i+l} \alpha_j^p Q_j(t) \quad (13.65)$$

Le choix des α_j^p dépend du problème physique considéré et en particulier de la nature précise de l'équation à résoudre comme nous le verrons dans la suite. Ils dépendent aussi de la précision voulue dans l'évaluation de l'opérateur.

Un des avantages majeurs de la méthode aux différences finies est sa simplicité, ce qui permet de traiter relativement facilement une grand variété de problèmes. De plus, on peut assez simplement obtenir une précision accrue sur les opérateurs de différentiation spatiale, sous certaines contraintes de régularité de la solution. Cependant, il existe des cas pour lesquels les opérateurs différences finies ne donnent pas les résultats physiques auxquels on pourrait s'attendre. Par exemple, l'existence de discontinuités (chocs, changement de milieu) ou de forts gradients peut engendrer de fortes erreurs dans les formules de différences finies et peuvent notamment donner lieu à la formation d'oscillations en aval des chocs (voir Fig. 10). D'autre part, ces méthodes ne tiennent pas compte *a priori* de l'existence de lois de conservations. Ainsi, la masse ou l'énergie totale ne sont en principe pas conservées dans les schémas aux différences finies en raison des erreurs numériques d'intégration. On notera cependant qu'il est toujours possible de développer des méthodes aux différences finies conservatives, sous certaines contraintes.

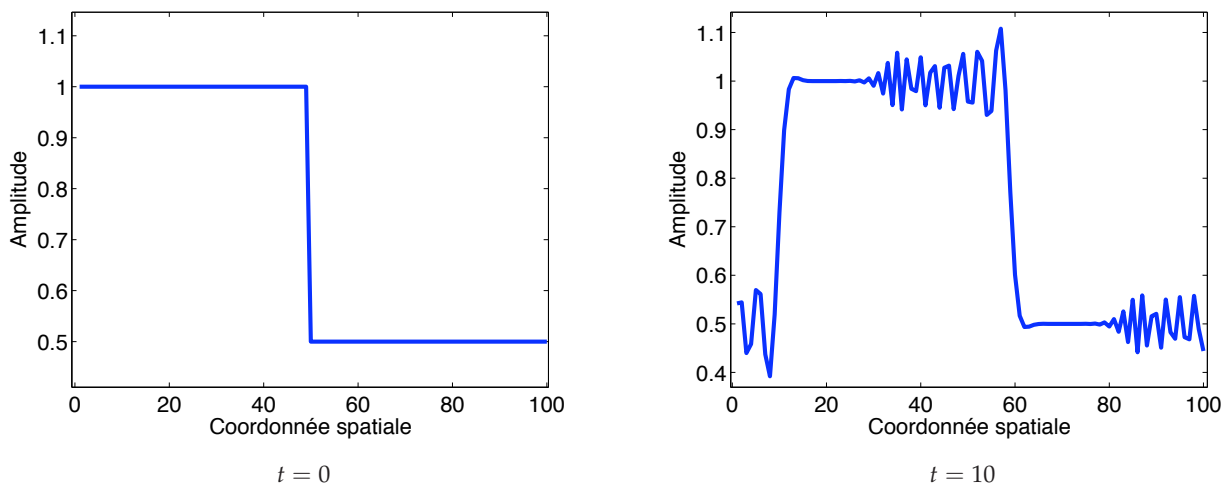


FIG. 10. Exemple d'oscillation due à la présence d'un choc dans une méthode aux différences finies. Résolution d'une équation d'advection $\partial_t u + \partial_x u = 0$ avec une intégration temporelle de type Runge-Kutta d'ordre 4 et une dérivation spatiale centrée d'ordre 5.

§ 14. Approche des volumes finis

Dans l'approche des volumes finis, on considère, de la même manière qu'en différences finies, une grille fixe dans l'espace. Cependant, au contraire des différences finies, les valeurs des variables sur les points de grille sont les *valeurs moyennes locales* des solutions du problème physique. Dans cette approche, on doit donc définir un ensemble de volumes de contrôle (rattachés aux points de contrôle) à l'intérieur desquels la moyenne est calculée.

Cette approche a certains avantages en mécanique des fluides, notamment en raison de l'existence de quantités conservées par les équations. En effet, dans la plupart des cas, on pourra écrire les équations fluides sous la forme :

$$\partial_t P = \nabla \cdot F(P) + S(P) \quad (14.66)$$

dite « forme conservative » de l'équation (12.62). P est une fonction arbitraire de la variable Q définie précédemment, F est la fonction de flux, et S est le terme source du système. Dans le cas unidimensionnel, en intégrant spatialement et temporellement l'équation (14.66), il vient :

$$\int_{x_{j-1/2}}^{x_{j+1/2}} P(t_{n+1}, x) dx - \int_{x_{j-1/2}}^{x_{j+1/2}} P(t_n, x) dx = \int_{t_n}^{t_{n+1}} F(x_{j+1/2}, t) dt - \int_{t_n}^{t_{n+1}} F(x_{j-1/2}, t) dt + \int_{t_n}^{t_{n+1}} \int_{x_{j-1/2}}^{x_{j+1/2}} S(x, t) dx dt. \quad (14.67)$$

Dans cette expression, on reconnaîtra dans le membre de droite les variables volumes finis $P_j(t_{n+1})$ et $P_j(t_n)$. Cette équation *exacte* décrit donc l'évolution temporelle des variables volumes finis en fonction des fonctions de flux intégrées temporellement et des termes sources. Toute la difficulté réside alors dans l'évaluation de ces termes, et en particulier du terme de flux. Notons de plus que le problème de l'intégration temporelle est caché dans l'évaluation des fonctions de flux et des termes sources.

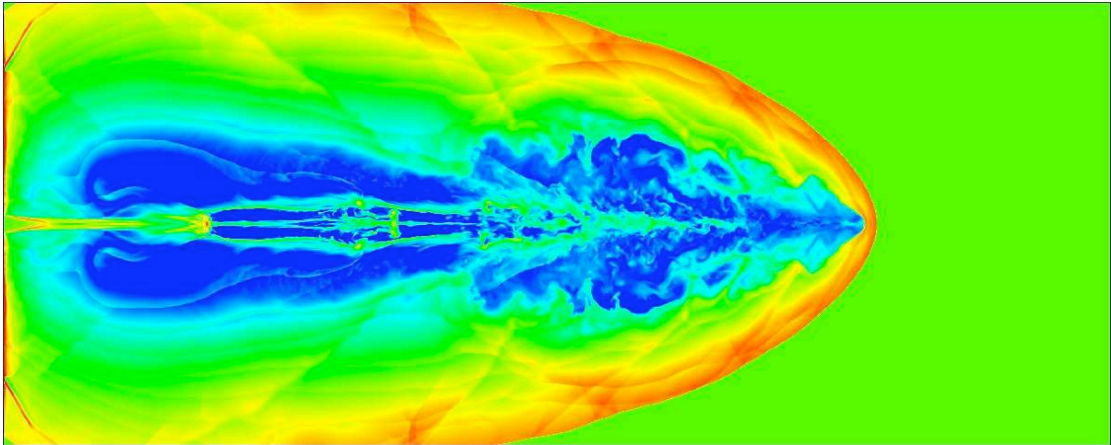


FIG. 11. Exemple de simulation aux volumes finis. Représentation de la densité lors de la propagation d'un jet astrophysique supersonique dans le milieu interstellaire. (Crédit G. Murphy)

De part sa forme intégrale, une méthode aux volumes finis peut facilement traiter les chocs dans le système physique sans engendrer d'oscillations parasites. De plus, de part leur forme conservative, ces méthodes permettent la conservation de certaines quantités (masse, énergie), à la précision de la machine, ce qui les rend très utiles dans les systèmes à géométrie complexe. Ces deux caractéristiques font que les méthodes des volumes finis sont très couramment utilisées en astrophysique, notamment parce que les vitesses y sont souvent supersoniques (d'où la formation des chocs) et que certaines quantités telles que la masse ou le moment cinétique doivent être conservées précisément (voir par exemple la formation d'un jet astrophysique, Fig. 11). Cependant, ces méthodes ne peuvent être que difficilement étendues à des ordres élevés, ce qui les rend moins adaptées aux écoulements sans discontinuités.

§ 15. Approche particulaire

L'approche particulaire fait appel à une discrétisation spatiale *mobile*. Cette méthode suppose que le fluide est constitué d'un nombre fini de particules que l'on suit dans leur mouvement. Ceci simplifie le problème de l'advection du fluide observé dans les méthodes précédentes. Cependant, l'évaluation des termes sources, et en particulier l'interaction entre les particules est assez difficile à décrire, ce qui peut engendrer certains problèmes (notamment au niveau des effets dissipatifs).

Par exemple, la méthode SPH (*Smoothed Particles Hydrodynamics*) est couramment employée en astrophysique pour des problèmes mettant en jeu une grande dynamique d'échelle et en particulier pour les milieux autogravitants (voir par exemple Fig. 12). Dans cette méthode, on suppose que les propriétés du fluide (température, pression...) en un point donné correspondent à une moyenne plus ou moins complexe des propriétés des particules fluides au voisinage de ce point. On remarquera toutefois que les écoulements magnétisés restent difficiles à simuler avec les méthodes particulières, et que les méthodes SPH-MHD sont encore en phase de développement (Price & Monaghan 2004).

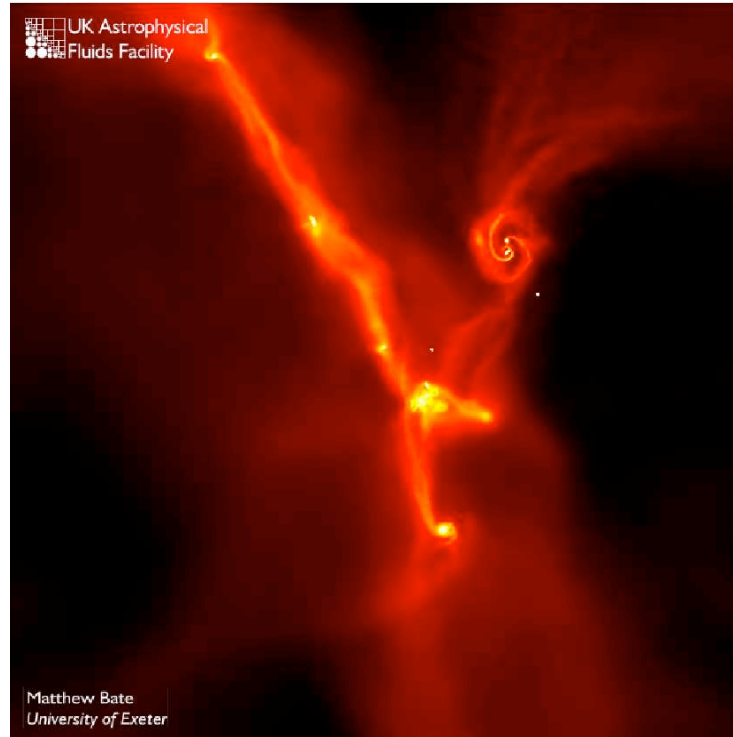


FIG. 12. Exemple de simulation particulière de type SPH. Représentation de la densité lors de l'effondrement et la fragmentation d'un nuage interstellaire. On peut voir apparaître des cœurs denses qui donneront naissance aux étoiles ainsi que le début de la formation de structures d'accrétion. D'après Bate *et al.* (2002).

3. Approche spectrale

Dans les cas où la forme de la solution attendue est connue et avec des géométries simples (sphérique, cartésienne), on peut projeter la solution sur une base de fonctions orthogonales respectant la symétrie et les conditions aux limites. On discrétise alors le problème en choisissant un sous ensemble fini de la base de projection et on réécrit les équations du problème sur cette base.

Un des avantages de cette méthode est que l'on peut dans la plupart des cas avoir des expressions simples et exactes des opérateurs de dérivation sur les éléments de la base. Par exemple, si l'on utilise une base de Fourier

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) \quad (15.68)$$

l'opérateur de dérivation spatiale s'écrit simplement :

$$\partial_{x_j} \varphi_{\mathbf{k}} = ik_j \varphi_{\mathbf{k}} \quad (15.69)$$

On comprend alors que la précision spatiale de ces méthodes est extrêmement bonne, à condition que la base utilisée soit cohérente avec la solution recherchée. D'autre part, certaines conditions telles que les conditions de non divergence du champ magnétique ou d'incompressibilité peuvent être appliquées de manière simple dans l'espace spectral. Enfin, la précision élevée des opérateurs spatiaux rendent ces méthodes très peu dissipatives et permet de diminuer de manière sensible la dissipation numérique dans le schéma d'intégration. Pour ces raisons, les méthodes spectrales ont été utilisées en mécanique des fluides depuis les années 1970 pour

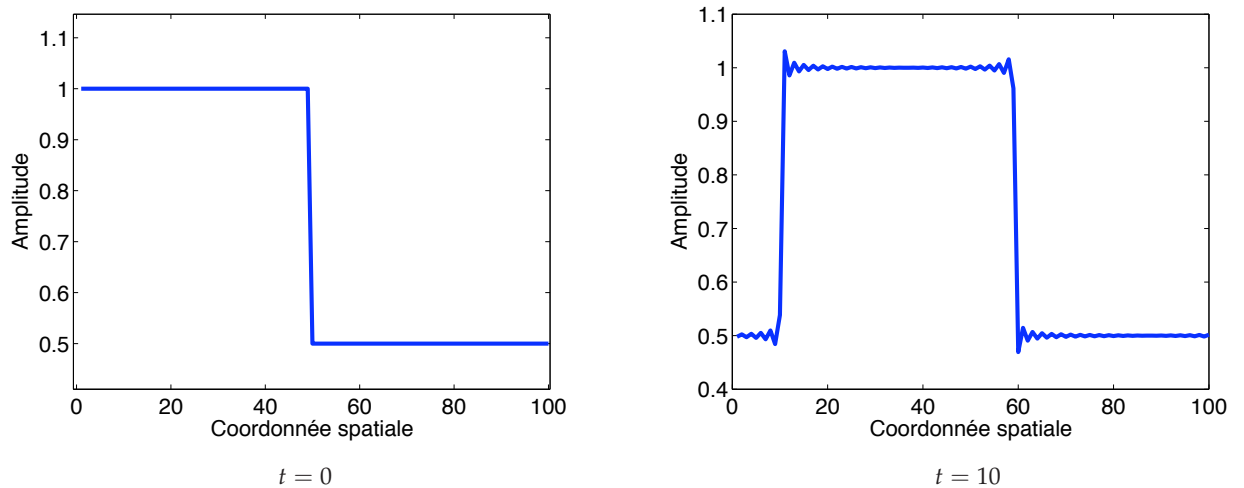


FIG. 13. Exemple d'oscillation due à la présence d'une discontinuité dans une méthode spectrale. Résolution d'une équation d'advection $\partial_t u + \partial_x u = 0$ avec une intégration temporelle de type Runge-Kutta d'ordre 4 et une dérivation spatiale spectrale dans l'espace de Fourier.

l'étude de la turbulence 3D et en particulier de la turbulence homogène (Orszag & Patterson 1972).

Cependant, les chocs sont ici aussi un exemple flagrant de solution ne pouvant être représentées par des bases continues tel que la base de Fourier. Ainsi, un test d'advection pour une méthode Fourier montre l'apparition d'oscillations de part et d'autre de la discontinuité (voir Fig. 13). Notons cependant que les oscillations sont beaucoup plus discrètes dans le cas des méthodes spectrales que pour les méthodes aux différences finies. D'autre part, l'utilisation des méthodes spectrales est plus coûteuse en temps de calcul que les méthodes précédentes. En effet, les calculs de dérivées nécessitent la plupart du temps des passages de l'espace réel à l'espace des fonctions de la base utilisée. Ces passages, ou *transformées*, se font par un produit de convolution, dont le temps de calcul évolue comme n^2 , où n est le nombre de points de grilles (ou le nombre d'éléments de la base utilisée). Comparativement aux méthodes différences finies qui évoluent en n , les méthodes spectrales deviennent donc rapidement très coûteuses en temps de calcul. Il existe cependant pour certaines bases des transformées « rapides » permettant de réduire de manière significative le temps de calcul. Les cas des transformées de Fourier rapides, ou *FFT*, en sont un exemple typique, pour lequel le temps de calcul évolue en $n \log n$. En pratique, l'utilisation d'une méthode spectrale nécessitera donc l'existence d'un algorithme de transformée rapide, ce qui limitera énormément le choix des bases possibles (essentiellement les bases de Fourier et de Tchebychev).

4. Quelles méthodes pour quelles simulations?

On l'a vu précédemment, les méthodes utilisables pour la mécanique des fluides sont nombreuses et ont chacune un certain nombre d'avantages et d'inconvénients. Dans le cadre de ce travail de thèse, l'étude numérique est concentrée sur les phénomènes de turbulence tridimensionnelle dans un disque d'accrétion potentiellement magnétisé.

Dans ce type de problème, la maîtrise des phénomènes de dissipation numérique est extrêmement importante, comme je le montrerai par la suite. Ainsi, il faut réduire au strict minimum ces phénomènes, c'est-à-dire utiliser des méthodes dont la précision sur le calcul des dérivées est grande. D'autre part, comme je l'ai montré dans l'introduction, on s'attend à obtenir une turbulence subsonique, ce qui limite la formation de chocs dans les simulations. Ces remarques tendent à exclure naturellement les méthodes aux volumes finis, à moins d'envisager une méthode de type Godunov d'ordre élevé, complexe à implémenter (Toro *et al.* 2001). De la même manière, les méthodes *SPH* ne seront pas utilisées, notamment en raison de leur forte dissipation.

Ainsi, seules des méthodes spectrales et de différences finies d'ordre élevées sont utilisées dans ce travail de thèse, compte tenu des hypothèses initiales. Dans la suite, je développerai quelques aspects de ces méthodes, leur implémentation, et leurs limites dans notre cas.

4

Méthodes aux différences finies

Plan du chapitre

1. Fondements.....	38
§ 16. Introduction.....	38
§ 17. Développement de Taylor.....	38
2. Un cas d'école (ou presque?) : L'équation linéaire d'advection.....	39
§ 18. Schéma temporel d'Euler.....	39
§ 18.1. La stabilité des schémas centrés en question.....	39
§ 18.2. Stabilité des schémas temporels.....	41
§ 19. Schémas de Runge-Kutta.....	42
§ 20. Intérêt des formules d'ordre élevés.....	43
§ 21. Advection d'une discontinuité.....	45
§ 22. Conclusions.....	46
3. Transport non linéaire.....	48
§ 23. L'équation de Burgers.....	48
4. Implémentation d'un code hydrodynamique aux différences finies.....	49
§ 24. Équations.....	49
§ 25. Conditions aux limites.....	49
§ 25.3. Conditions aux limites rigides.....	50
§ 25.4. Conditions aux limites shearing sheet.....	51
§ 26. Tests.....	52
5. Magnétohydrodynamique.....	53
§ 27. Le problème de la divergence de B	53
§ 28. Choix de Jauge.....	54
6. Parallélisation.....	55
§ 29. Choix d'une méthode de parallélisation.....	55
§ 30. Décomposition de domaine.....	56

1. Fondements

§ 16. Introduction

LES MÉTHODES AUX DIFFÉRENCES FINIES sont une classe de méthodes numériques très utilisée encore aujourd'hui, malgré leurs nombreux défauts. Dans ce chapitre, je ne présenterai pas de résultat fondamentalement nouveau, et on pourra consulter Forsythe & Wasow (1964) ou encore Press *et al.* (2002) pour avoir des détails supplémentaires sur cette approche, ainsi qu'une série d'algorithmes écrits dans différents langages de programmation. L'objectif de ce chapitre est donc de présenter de manière semi-quantitative quelques éléments sur la méthode aux différences finies, appliquée à la problématique de cette thèse. De plus, quelques unes des approches numériques présentées ici seront utilisées pour décrire les méthodes spectrales du prochain chapitre.

§ 17. Développement de Taylor

Considérons une fonction ψ sur une grille unidimensionnelle dont les coordonnées des points sont notées x_i et de pas δx supposé constant. La méthode des différences finies consiste à trouver une expression pour les dérivées à l'ordre q de ψ en fonction des valeurs $\psi_i = \psi(x_i)$, ce que l'on écrira sous la forme générale :

$$(\delta x)^q \psi_i^{(q)} = \sum_{j=i-n}^{i+m} \alpha_j^{(q)} \psi_j \quad (17.70)$$

où les $\alpha_j^{(q)}$ sont les inconnues. En utilisant les formules de Taylor jusqu'à l'ordre p , on montre facilement que :

$$\psi_j = \sum_{q=0}^p \frac{(j-i)^q \delta x^q}{q!} \psi_i^{(q)} + o(\delta x^p) \quad (17.71)$$

On pourra remarquer que les équations (17.70) et (17.71) correspondent l'une et l'autre à des formulations matricielles inverses si $p = m + n$. En effet, on peut écrire matriciellement l'équation (17.71) sous la forme :

$$\begin{pmatrix} \psi_{i-n} \\ \psi_{i-n+1} \\ \vdots \\ \psi_{i+m} \end{pmatrix} = \begin{pmatrix} 1 & -n & n^2/2 & \dots \\ 1 & -n+1 & (n-1)^2/2 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & m & m^2/2 & \dots \end{pmatrix} \begin{pmatrix} \psi_i \\ \psi_i' \delta x \\ \vdots \\ \psi_i^{(p)} \delta x^p \end{pmatrix} \quad (17.72)$$

En inversant la matrice intervenant dans (17.72), on obtiendra la matrice des $\alpha_j^{(n)}$ qui sera une représentation différences finies d'ordre p des dérivées de ψ . On pourra remarquer qu'il existe *a priori* une infinité de représentations d'une dérivée à un ordre donné. Cependant, on choisira de préférence les formules faisant intervenir les valeurs de ψ au voisinage du point où l'on cherche la dérivée. Le choix précis de la formule à utiliser sera dicté par des considérations de stabilité et de convergence que l'on verra par la suite. On trouvera en annexe quelques formules différences finies dans les cas les plus courants.

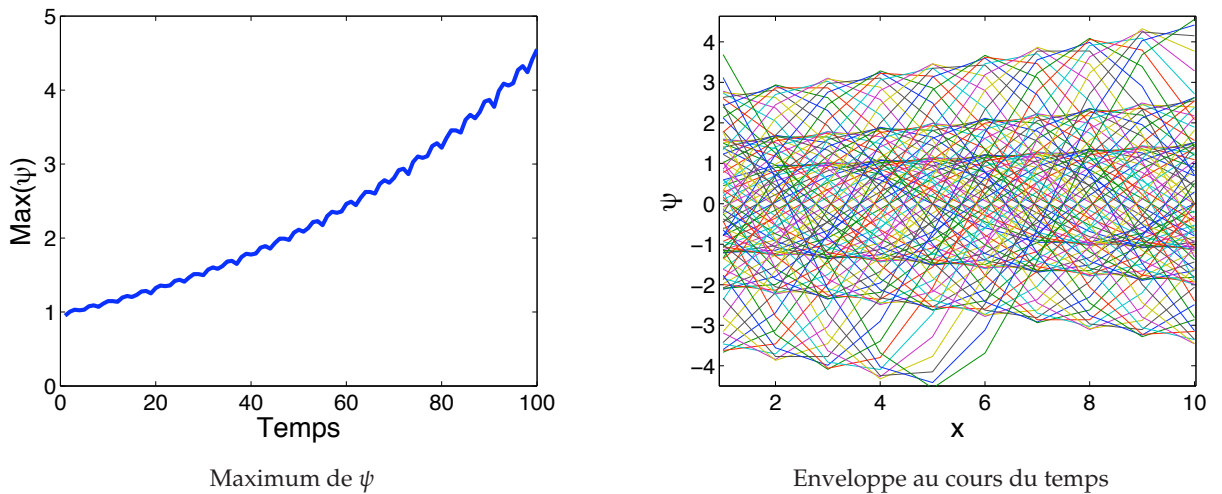


FIG. 14. Évolution d'une fonction sinus advectée par un schéma d'Euler et une formule différences finies centrée d'ordre 2. Simulation sur 100 pas de temps et 10 points de grille. On met en évidence l'accroissement de l'amplitude de la solution : cet algorithme est instable.

2. Un cas d'école (ou presque?) : L'équation linéaire d'advection

Le cas de l'équation d'advection est un problème classique d'étude en mécanique des fluides. En effet, ces équations Eulériennes font intervenir systématiquement un terme de « transport » où les champs sont simplement déplacés dans l'espace. Une telle équation d'advection s'écrit :

$$\partial_t \psi + c \partial_x \psi = 0 \quad (17.73)$$

où c est la vitesse d'advection. En posant $y = x - ct$, on remarque que (17.73) est équivalente à $d\psi/dy = 0$. On montre alors facilement que la solution de (17.73) s'écrit :

$$\psi(x, t) = \psi(x_0 - c(t_0 - t), t_0) \quad (17.74)$$

ce qui est conforme à l'idée du transport de ψ à une vitesse c sur l'axe x .

§ 18. Schéma temporel d'Euler

§ 18.1. La stabilité des schémas centrés en question

La formulation la plus simple que l'on puisse imaginer pour l'équation d'advection est d'utiliser un schéma d'intégration temporel d'Euler. Ce dernier s'écrit sous la forme intuitive :

$$\psi^{n+1} = \psi^n - c \delta t \partial_x \psi^n \quad (18.75)$$

où l'on a noté ψ^n la solution au pas de temps n . On peut alors exprimer la dérivée spatiale avec une formule de différences finies centrée du deuxième ordre. On écrira alors l'algorithme sous la forme :

$$\psi_i^{n+1} = \psi_i^n - c \delta t \frac{\psi_{i+1}^n - \psi_{i-1}^n}{2 \delta x} \quad (18.76)$$

Malheureusement, la programmation de cet algorithme test pour l'advection d'une fonction sinus montre rapidement que la solution est instable (voir Fig. 14).

Ce type de problème peut être étudié de manière systématique en utilisant l'analyse de Von Neumann. Cette analyse suppose que la fonction ψ est une fonction sinusoïdale sous la forme :

$$\psi_j^n = \Gamma^n \exp(ikj\delta x) \quad (18.77)$$

Ainsi, en injectant (18.77) dans (18.76), on trouve :

$$\Gamma^{n+1} = \Gamma^n \exp(ikj\delta x) \left(1 - \frac{ic\delta t}{\delta x} \sin(kj\delta x)\right) \quad (18.78)$$

Cette relation montre clairement que le taux de croissance $\tau = |\Gamma^{n+1}/\Gamma^n|$ est plus grand que 1, c'est-à-dire que la solution croît au cours du temps, et ce pour toutes les valeurs de $c\delta t/\delta x$. On dira alors que l'algorithme (18.76) est inconditionnellement instable.

Il est alors couramment admis que ce schéma doit son instabilité à la représentation différences finies centrée. En effet, on pourra remarquer que physiquement, on advecte en un point ce qu'il y avait en amont de ce point l'instant auparavant. Ainsi, on peut remplacer l'algorithme (18.76) par un schéma *upwind*, c'est-à-dire prenant en compte la physique de la propagation, sous la forme :

$$\psi_i^{n+1} = \psi_i^n - c\delta t \frac{\psi_i^n - \psi_{i-1}^n}{\delta x} \quad (18.79)$$

On vérifie alors facilement que :

$$\tau^2 = 1 - \frac{2c\delta t}{\delta x} [1 - \cos(k\delta x)] \left(1 - \frac{c\delta t}{\delta x}\right) \quad (18.80)$$

ce qui montre que ce schéma est stable si $c\delta t/\delta x < 1$. Cette condition, appelée condition CFL (Courant, Friedrichs, Lewy), donne un maximum pour le pas de temps utilisé dans le schéma d'intégration. On la retrouvera dans tous les schémas explicites en temps¹¹, mais son expression précise peut varier comme nous le verrons par la suite.

En comparant le schéma (18.79) au schéma centré (18.76), on pourra remarquer qu'ils diffèrent par un terme de dissipation faisant intervenir une dérivée seconde. En effet, on peut réécrire (18.79) sous la forme :

$$\psi_i^{n+1} = \psi_i^n - c\delta t \frac{\psi_{i+1}^n - \psi_{i-1}^n}{2\delta x} + c\delta t \frac{\psi_{i+1}^n - 2\psi_i^n + \psi_{i-1}^n}{2\delta x} \quad (18.81)$$

où l'on reconnaîtra dans le membre de droite un terme de la forme $c\delta x \partial_x^2 \psi$. Ainsi, le schéma (18.79) est beaucoup plus dissipatif que le schéma centré d'origine, ce qui peut poser problème au niveau de la solution trouvée numériquement, qui se trouve être diffusée très rapidement. Remarquons que de manière générale, les schémas *upwind* diffèrent des schémas centrés par un terme de dissipation qui peut être d'ordre plus élevé et qui accroît naturellement la stabilité de ces schémas. Ainsi, cet exemple est souvent utilisé dans la littérature pour justifier l'utilisation de schémas *upwind* dans les méthodes numériques (voir par exemple Toro 1999). Cependant, nous allons voir que ces instabilités numériques ont en fait une origine différente.

¹¹Un schéma explicite est un schéma dont la solution au temps $n + 1$ ne dépend que de la solution aux temps inférieures ou égaux à n .

§ 18.2. Stabilité des schémas temporels

Tout d'abord, imaginons une méthode de différenciation spatiale d'ordre infini (on peut par exemple imaginer une méthode spectrale dont l'ordre est très élevé). Cette méthode permet de calculer *exactement* la dérivée d'une fonction, pourvue que ses propriétés de continuités soient suffisantes (par exemple une fonction trigonométrique). Considérons alors le schéma d'intégration spatiale d'Euler décrit au paragraphe précédent avec cette méthode de différenciation spatiale. Une analyse de Von Neumann d'un tel schéma montre alors immédiatement que le taux de croissance :

$$\tau^2 = 1 + c^2 \delta t^2 k^2 \quad (18.82)$$

est toujours supérieur à 1, c'est-à-dire que cette méthode est à nouveau instable. Ceci montre que même avec une méthode de différenciation spatiale exacte, notre méthode d'intégration reste instable. Ceci pointe vers le fait que le problème de stabilité vient plus du schéma d'intégration temporel que du schéma spatial...

Pour étudier les schémas d'intégration temporels, et en particulier leur stabilité, je propose d'utiliser une approche graphique, originellement développée par Canuto *et al.* (1988) pour les méthodes spectrales, qui permet de comprendre les choix effectués dans cette thèse. Cette approche est basée sur une analyse de Von Neumann, que l'on va étendre pour l'intégration temporelle. Tout d'abord, en utilisant les fonctions de Von Neumann (18.77), ainsi qu'une formulation aux différences finies du type (17.70), on peut écrire :

$$\partial_x \psi_k^n = \frac{\psi_k^n}{\delta x} \left(\sum_{j=k-n}^{k+m} \alpha_j^{(1)} \exp [ik(j-k)\delta x] \right) \quad (18.83)$$

Ce que l'on réduira à l'expression générique :

$$\partial_x \psi_i^n = \psi_i^n \frac{f(k\delta x)}{\delta x} \quad (18.84)$$

où f est une fonction à valeurs dans \mathbb{C} , caractéristique de la formule de différences finies utilisée. En utilisant cette formulation, on remarque que l'équation différentielle d'advection (17.73) se réduit à :

$$\partial_t \psi = -c \psi_i^n \frac{f(k\delta x)}{\delta x} \quad (18.85)$$

qui est donc une équation différentielle ordinaire. Aussi, pour la plupart des schémas d'intégration temporelles explicites (et en particulier ceux que je vais étudier dans la suite), on pourra écrire :

$$\psi_i^{n+1} = \psi_i^n \left[1 - \mathcal{T} \left(\frac{c \delta t f(k\delta x)}{\delta x} \right) \right] \quad (18.86)$$

Cette forme de l'équation d'advection présente l'avantage de pouvoir séparer assez simplement le rôle de l'intégration temporel du schéma spatial. En effet, la stabilité du schéma sera vérifiée si $|1 - \mathcal{T}(\dots)| < 1$. Ainsi, on pourra supposer dans un premier temps que l'argument de \mathcal{T} prend toutes les valeurs possibles dans \mathbb{C} . On peut alors connaître les régions pour lesquelles le schéma temporel est stable. Il est ensuite facile de spécifier, pour un schéma différences finies, si la stabilité est possible ou pas.

Dans le cas du schéma d'intégration temporel d'Euler, la fonction \mathcal{T} est la fonction identité. On peut alors tracer un graphique représentant $|1 - \mathcal{T}(q)|$ en fonction des parties réelles et imaginaires de l'argument q (Fig. 15). Ce graphique montre le comportement général du

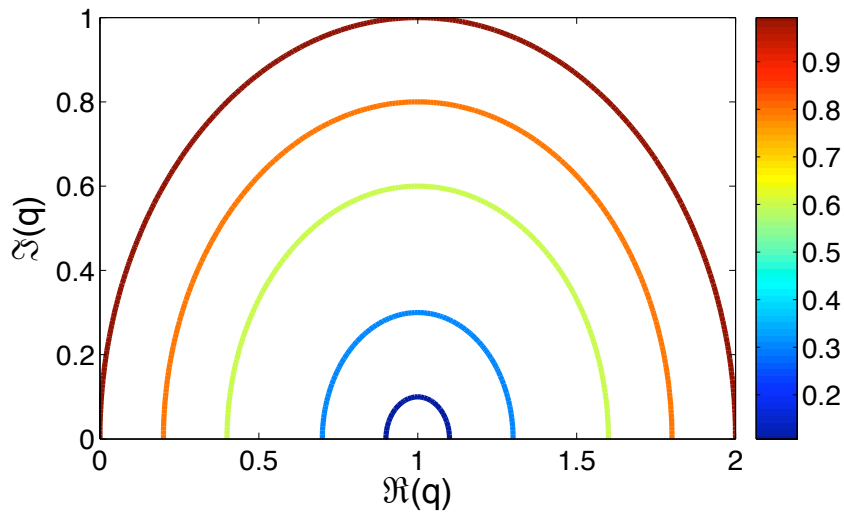


FIG. 15. Tracé de $|1 - \mathcal{T}(q)|$ pour un schéma temporel d'Euler en fonction des valeurs réelles et imaginaires de q . On voit clairement qu'un schéma différences finies symétrique [$\Re(q) = 0$] ne peut pas être stable avec un tel schéma temporel.

schéma temporel et permet de tirer quelques conclusions globales. On remarquera par exemple qu'un schéma différences finies centré ne pourra donner que des valeurs de f imaginaires pures [voir par exemple (18.78)], ce qui rendra nécessairement instable le schéma d'Euler. La même remarque s'applique lorsqu'on utilise une méthode exacte (ou spectrale) pour lesquelles $f = ik$. Ainsi, la nécessité d'utiliser un schéma upwind pour une équation d'advection découle seulement de l'utilisation d'un schéma d'Euler pour l'intégration temporelle, ce que nous allons pouvoir vérifier par la suite.

§ 19. Schémas de Runge-Kutta

On l'a vu précédemment, l'utilisation d'un schéma d'Euler impose l'utilisation de formules différences finies upwind d'ordre faible qui sont beaucoup plus dissipatives que des formules centrées ou des formules d'ordre élevé. Ainsi, il est souhaitable de trouver une méthode d'intégration, de préférence explicite en temps, qui soit stable, y compris pour les formules centrées. Un tel schéma peut être recherché parmi les schémas de Runge-Kutta, déjà testés pour le calcul de solutions d'équations différentielles ordinaires.

Nous nous proposons donc ici de tester les schémas de Runge-Kutta d'ordre 2 et 4, qui sont en pratique les plus utilisés. En utilisant les expressions algorithmiques de ces schémas (voir Annexe B), on peut se ramener à des expressions formelles, permettant d'obtenir un critère de stabilité générique. Pour se faire, on injecte l'équation différentielle ordinaire (18.85) dans un des algorithmes de Runge-Kutta. Cette équation étant très simple, on la résout formellement d'un pas d'intégration, ce qui nous permet finalement d'écrire la solution au pas de temps $n + 1$ sous la forme (18.86). En pratique, on trouve :

$$\psi_i^{n+1} = \psi_i^n \left(1 - q + \frac{1}{2}q^2 \right) \quad (19.87)$$

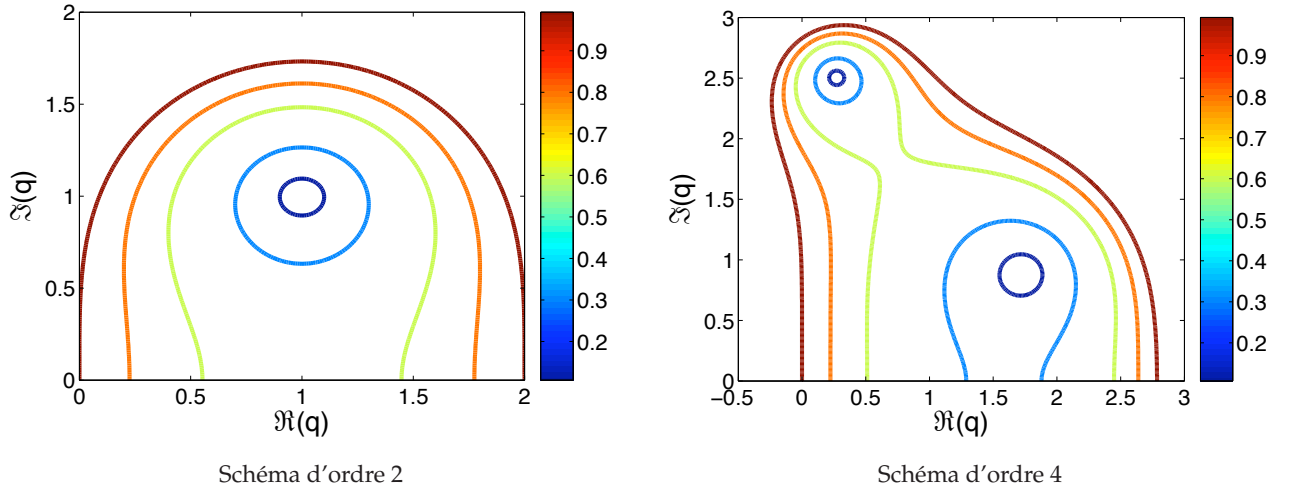


FIG. 16. Tracé de $|1 - \mathcal{T}(q)|$ pour 2 schémas temporels de Runge-Kutta en fonction des valeurs réelles et imaginaires de q . On montre ici qu'un schéma d'ordre 4 est stable pour des valeurs imaginaires pures de q , contrairement au schéma d'ordre 2.

pour le schéma de Runge-Kutta d'ordre 2, en posant $q = \frac{c\delta t}{\delta x} f(k\delta x)$. De même, le schéma de Runge-Kutta donne naturellement :

$$\psi_i^{n+1} = \psi_i^n \left(1 - q + \frac{1}{2}q^2 - \frac{1}{6}q^3 + \frac{1}{24}q^4 \right) \quad (19.88)$$

A partir de ces expressions, il suffit de faire une analyse similaire à celle faite pour le schéma d'Euler pour connaître les propriétés de stabilité du schéma considéré (voir Fig. 16). Cette analyse montre de nouveau une instabilité systématique pour les schémas symétriques de l'algorithme de Runge-Kutta d'ordre 2. Cependant, l'algorithme d'ordre 4 ne présente pas ce défaut. Dans le cas simple où l'on considère un schéma symétrique, c'est à dire q imaginaire pur, on montre facilement que ce schéma est stable si $q < \sqrt{8}$, ce qui constitue une condition CFL généralisée.

Pour résumer les résultats de cette partie, j'ai effectué des tests d'advection d'une fonction sinus sur 1000 pas de temps en utilisant les différents schémas temporels et spatiaux étudiés ici (voir Fig. 17). On y retrouve la très forte dissipation du schéma upwind (cette dissipation est toujours présente, quel que soit la méthode temporelle employée), et la forte instabilité des méthodes centrées combinées au schéma d'Euler. On observe aussi l'instabilité nettement plus faible du schéma de Runge-Kutta d'ordre 2 et la stabilité du schéma d'ordre 4. Remarquons que ce dernier montre une dissipation extrêmement faible (une analyse attentive montre que $\log(\psi_{max}^f / \psi_{max}^0) \simeq -6.10^{-3}$). La conclusion pourrait être qu'il faut utiliser le schéma centré avec la méthode Runge-Kutta d'ordre 4. Cependant, nous allons voir que cette combinaison présente certains désavantages, plus discrets que ceux étudiés jusqu'à présent, qui vont nous pousser vers l'utilisation de méthodes d'ordres plus élevés.

§ 20. Intérêt des formules d'ordre élevés

On pourra remarquer que les arguments pour l'utilisation d'une méthode plutôt qu'une autre étaient fondés sur l'amplitude de la solution (18.77). De cette amplitude découle la stabilité du

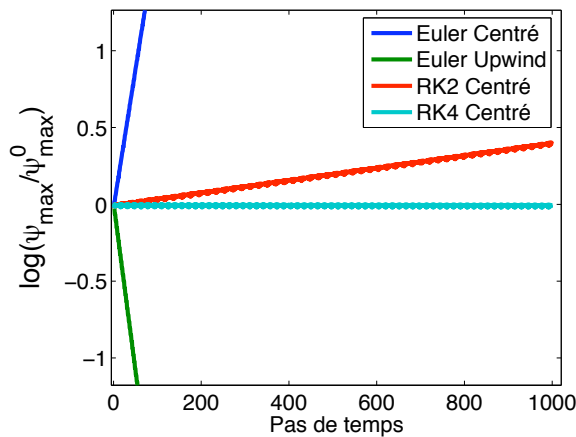


FIG. 17. Tracé du maximum de ψ en fonction du temps pour différents schémas d'advection. On voit clairement que les schémas centrés et upwind sont inexploitable avec la méthode d'Euler. L'instabilité du schéma Runge-Kutta d'ordre 2 reste très contenue et le schéma Runge-Kutta d'ordre 4 est stable.

schéma ainsi que sa dissipation intrinsèque. Cependant, l'advection de la solution est contrôlée par la *phase* de cette solution. On comprend alors que même si une méthode donne une solution d'amplitude *a priori* idéale, la physique de l'advection peut être très approximative, en raison de fortes erreurs de phase.

Cette remarque nous pousse à analyser plus en détail le problème de la phase pour les schémas utilisés jusqu'à présent. Pour simplifier, considérons que nous disposons désormais d'un schéma numérique temporel idéal (limite où $\delta t \rightarrow 0$). Dans ce cas, en utilisant une solution du type (18.77) et une formulation différences finies de la forme (18.84), on peut écrire l'équation d'advection sous la forme :

$$\partial_t \psi_j = -c \psi_j \frac{f(k\delta x)}{\delta x} \quad (20.89)$$

et la solution sera donnée de manière évidente par :

$$\psi_j(t) = \Gamma^0 \exp \left(-ct \frac{f(k\delta x)}{\delta x} + ikj\delta x \right) \quad (20.90)$$

On pourra alors noter formellement $ik_{\text{eff}} = f(k\delta x)/\delta x$ ce qui permet de comparer cette solution à la solution formelle du problème (17.73), ce qui revient finalement à comparer les vecteurs d'onde des 2 solutions k_{eff} et k . La comparaison des parties réelles donne la vitesse d'advection effective alors que la comparaison des parties imaginaires donne le taux de dissipation numérique de la formule utilisée. Une telle analyse est faite des formules différences finies centrées et upwind à des ordres différents sur la figure (18).

Remarquons tout d'abord que les vitesses d'advection des méthodes centrées d'ordre n et des méthodes upwind d'ordre $n - 1$ sont identiques. Ceci vient du fait que les méthodes upwind peuvent être vues comme des méthodes centrées auxquelles on rajoute un terme purement dissipatif, sous la forme d'un terme de différences finies représentant une puissance de ∂_x^2 . Ainsi, toutes les méthodes upwind font apparaître une certaine dissipation ce qui explique leur stabilité comparativement aux schémas centrés. De plus, on remarque que l'on obtient des vitesses d'advection plus exactes avec des formules d'ordre élevé. En particulier, la formule d'ordre 4 upwind présente un très bon comportement sur ce point, malgré une diffusion importante aux grands nombres d'onde. Ceci montre que l'on aura intérêt, pour tenir compte des petites échelles, à prendre l'ordre de différentiation le plus élevé possible compatible avec les contraintes techniques. En effet, l'utilisation de ces méthode est coûteuse en temps de calcul, et diminue l'efficacité de la parallélisation, comme nous le verrons par la suite.

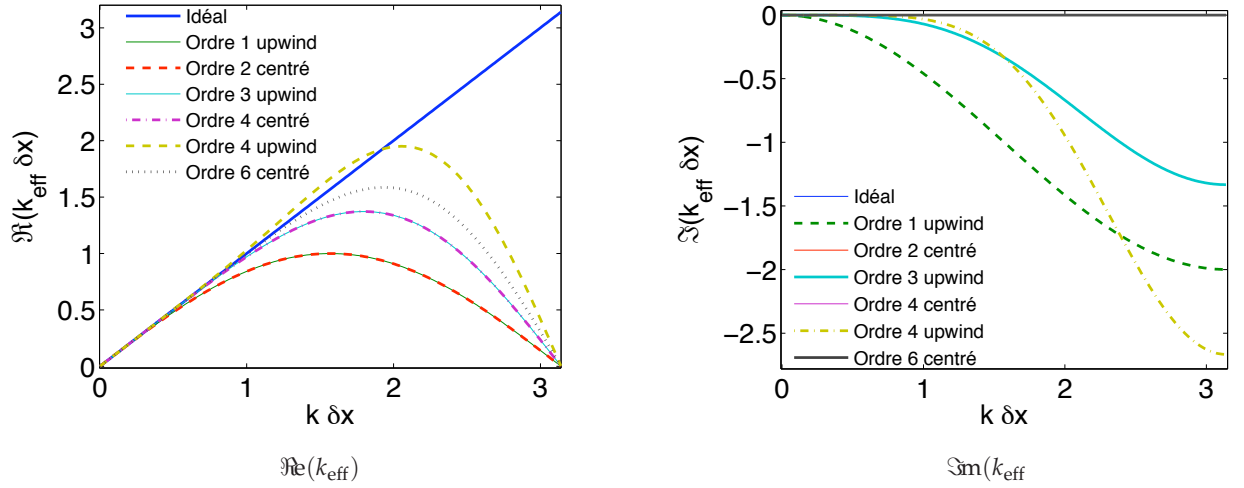


FIG. 18. Tracés de k_{eff} pour différentes formules différences finies. Notons que les schémas centrés ont systématiquement une partie imaginaire nulle. De même les schémas centrés d'ordre n et upwind d'ordre $n - 1$ sont superposés sur le schéma $\Re(k_{\text{eff}})$.

§ 21. Advection d'une discontinuité

Jusqu'à présent, nous nous sommes contentés d'utiliser des fonctions tests simples et sans discontinuités telles que les fonctions trigonométriques. Cependant, un champ de vitesse réel peut présenter des discontinuités, même pour un écoulement largement subsonique. On pourra par exemple citer les discontinuité de cisaillement qui peuvent se former dans des écoulements turbulents. Ainsi, il est important de connaître et de tester le comportement d'un schéma face à l'advection d'une discontinuité.

Une des difficultés dans l'advection d'un signal discontinu vient du fait que les formules aux différences finies advectent naturellement des polynômes. Ainsi, une formule aux différences finies d'ordre 4 pourra advecter, en théorie sans erreur, un polynôme du même ordre. De manière générale, le calcul d'une dérivée en différences finies à l'ordre $n - 1$ correspond au calcul de la dérivée du polynôme interpolateur d'ordre n passant par les points utilisés dans le schéma aux différences finies. On pourra ainsi écrire le polynôme interpolateur P_i sous la forme d'un polynôme de Lagrange :

$$P_i(x) = \sum_{j=1}^n \psi_j \prod_{\substack{k=1 \\ k \neq j}}^n \frac{x - x_j}{x_k - x_j}$$

Au voisinage d'une discontinuité, les polynômes interpolateurs ne convergent pas vers la fonction discontinue : c'est le phénomène de Gibbs (voir Fig. 19). Cette oscillation des polynômes interpolateurs se retrouve donc naturellement dans le calcul des dérivées en différences finies, ce qui engendre des oscillations parasites dans l'advection d'un choc.

Ainsi, on peut trouver sur la figure (20) le résultat de l'advection d'un choc par différentes formulations en différences finies. On remarque que toutes les méthodes, exceptée celle d'ordre 1, donnent lieu à la formation d'oscillations parasites. On peut aussi noter l'avantage des méthodes upwind qui dissipent naturellement ces oscillations au cours de l'advection,

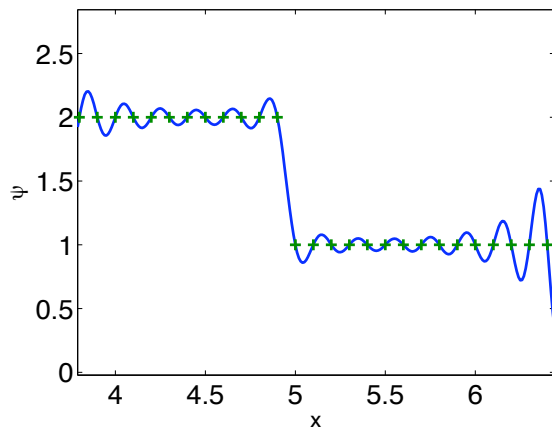


FIG. 19. Interpolation d'une fonction constante par morceaux par des polynômes de Lagrange. Le phénomène d'oscillation apparaissant est nommé phénomène de Gibbs.

contrairement aux méthodes centrées. Enfin, il apparaît que les méthodes d'ordres élevés sont plus sensibles à ce phénomène que les autres.

On notera que l'existence d'oscillation dans le résultat est relié à la monotonie d'un schéma. En effet, considérons un schéma linéaire que l'on peut écrire sous la forme :

$$\psi_i^{n+1} = \sum_{j=i-k}^{i+m} \beta_j \psi_j^n. \quad (21.91)$$

Ce schéma sera dit monotone si tous les coefficients β_j sont positifs. Un tel schéma présente l'avantage de ne pas créer de nouveaux extrema à chaque pas de temps (on montre en effet que si ψ^n est monotone, alors ψ^{n+1} l'est aussi). Ainsi, un schéma monotone permet de garantir qu'il ne créera pas d'oscillation parasite. On pourra ainsi remarquer que le schéma upwind avec une intégration d'Euler (18.79) est un schéma monotone. Malheureusement, d'après le théorème de Godunov, la formule de différences finies upwind d'ordre 1 est la seule méthode linéaire [c'est-à-dire de la forme (21.91)] qui donne un schéma monotone (voir par exemple Toro 1999 p. 444). Ceci montre que nous ne pourrons jamais éliminer complètement les oscillations dans les schémas d'ordre élevé.

Ainsi, une possibilité pour atténuer ces oscillations est l'utilisation de méthodes upwind d'ordre élevé comme on le voit sur la figure (20). Ces méthodes éliminent les oscillations formées par l'advection au voisinage du choc de manière assez efficace grâce à leur composante dissipative, et c'est cet argument qui nous poussera à choisir une méthode upwind plutôt qu'une méthode centrée dans la suite.

§ 22. Conclusions

Nous avons vu dans cette section quelques principes fondamentaux sur le problème de l'advection dans les codes numériques. En particulier, on a montré par des considérations de stabilité qu'il était préférable d'utiliser une méthode d'intégration temporelle d'ordre élevé (Runge Kutta d'ordre 4). De plus, les formules spatiales d'ordre élevé nous permettent d'obtenir des vitesses d'advection plus précises que les formules classiques. Enfin, on éliminera les oscillations parasites apparaissant au voisinage des zones de discontinuité en utilisant des formules spatiales upwind, qui intègrent naturellement une forme de dissipation numérique.

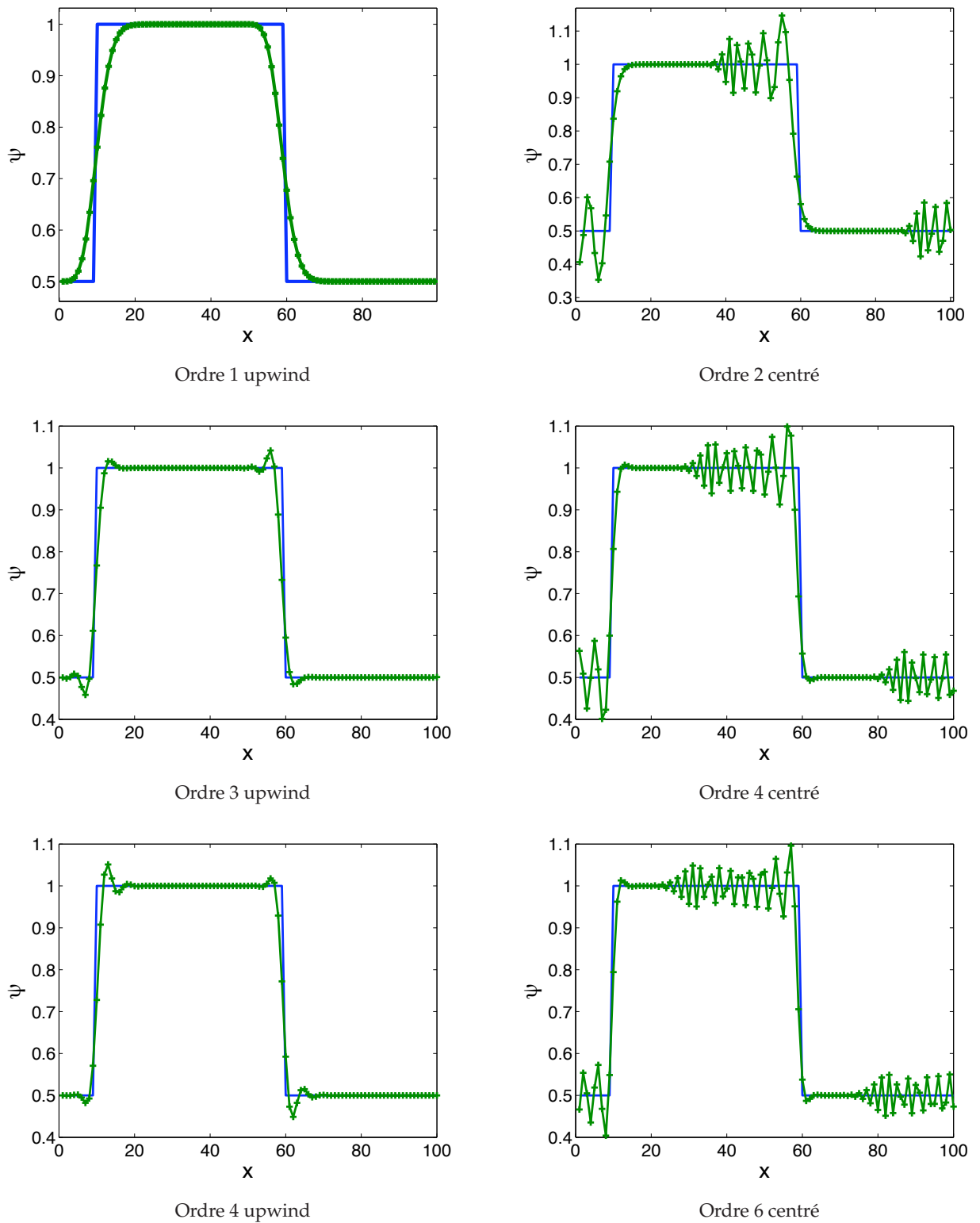


FIG. 20. Tests d'advection d'un créneau avec différentes formulations différences finies et une intégration temporelle de Runge-Kutta à l'ordre 4 (20 pas de temps avec $\delta t = 0.5$ et $c = 1$). On remarque que les méthodes upwind permettent de réduire de manière significative les oscillations, même à des ordres élevés.

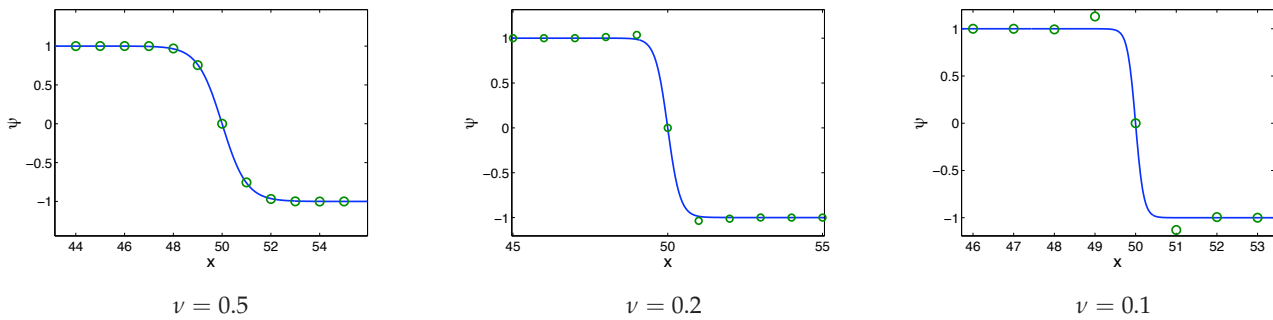


FIG. 21. Test d'un algorithme utilisant une intégration temporelle de Runge Kutta d'ordre 4 et des dérivées spatiales aux différences finies upwind d'ordre 4 sur l'équation de Burgers. L'intégration a été effectuée sur 50 pas de temps pour $U_0 \Delta t / \Delta x = 0,5$.

Compte tenu de ces résultats, nous utiliserons dans la suite une méthode de Runge-Kutta d'ordre 4 combinée avec la formule différence finie upwind d'ordre 4.

3. Transport non linéaire

§ 23. L'équation de Burgers

Jusqu'à présent, l'équation étudiée était linéaire et les solutions étaient trouvées facilement de manière analytique. Malheureusement, les équations de la physique font couramment intervenir des termes non linéaires, sources de nombreux problèmes. Ainsi, toutes les équations fluides font intervenir un terme de transport non linéaire. C'est le cas par exemple du champ de vitesse qui est transporté par lui même. Ce type de terme peut engendrer des phénomènes très violents comme des chocs et excite naturellement toute la gamme des fréquences spatiales disponibles : une analyse en harmoniques telle que celle utilisée jusqu'à présent ne nous sera donc d'aucun secours.

Pour étudier les interactions non linéaires d'un point de vue numérique, on peut utiliser une équation modèle mettant en jeu les principales caractéristiques de la physique que l'on veut étudier : c'est l'équation de Burgers, que l'on écrira sous la forme :

$$\partial_t u = -u \partial_x u + \nu \partial_x^2 u \quad (23.92)$$

On peut y reconnaître un terme de transport et un terme dissipation, comparables à une équation de Navier-Stokes à une dimension. On peut vérifier que :

$$\psi(x) = U_0 \tanh\left(\frac{U_0 x}{2\nu}\right) \quad (23.93)$$

est une solution stationnaire de cette équation. De plus, elle peut être vue comme un choc d'épaisseur $2\nu/U_0$, ce qui permet de tester le comportement non linéaire d'un code face à un gradient très fort.

Ainsi, nous avons testé le comportement de notre algorithme d'intégration sur l'équation de Burgers pour différentes valeurs de ν . Les résultats sont regroupés sur la figure (21). On voit que le comportement du code face à de forts gradients reste correct, même pour des viscosités faibles ($\nu = 0,1$). Notons cependant l'apparition d'une oscillation parasite dans les cas où la

taille typique du gradient est en dessous de la taille de grille : c'est la signature de l'utilisation d'une méthode non monotone.

4. Implémentation d'un code hydrodynamique aux différences finies

§ 24. Équations

Je vais développer succinctement dans cette section l'implémentation en différences finies de l'un des codes que j'ai développé et utilisé dans ce travail de thèse. Pour se faire, commençons par développer les équations de la mécanique des fluides telles qu'elles sont résolues par le code.

$$\begin{aligned}\frac{D \ln(\rho)}{Dt} &= -\nabla \cdot \mathbf{v} \\ \frac{D \mathbf{v}}{Dt} &= -\frac{\nabla P}{\rho} - 2\boldsymbol{\Omega} \times \mathbf{v} + \frac{1}{\rho} \nabla \cdot \mathcal{T} \\ \frac{D \ln(P)}{Dt} &= -\gamma \nabla \cdot \mathbf{v} + \left[\frac{(\gamma - 1)}{P} \mathcal{T} \cdot \nabla \mathbf{v} \right]\end{aligned}$$

Ce système d'équations appelle à plusieurs remarques. Tout d'abord, notons que l'on préfère faire évoluer les logarithmes des quantités scalaires plutôt que leurs valeurs réelles. Cela présente plusieurs avantages. En effet, une telle méthode garantit que ρ et P sont positifs. De plus, des simulations faisant intervenir une large amplitude en densité et en pression (cas d'un disque et sa couronne par exemple) seront traitées avec plus de précision.

D'autre part, je fais appel dans ce code à une dissipation physique sous la forme du tenseur des contraintes \mathcal{T} , décrit en introduction (Eq. 6.24, p. 17). Dans l'équation d'énergie, on introduit donc en plus de l'évolution isentropique classique, un terme de chauffage (terme entre crochets). Cependant, pour les raisons explicités en introduction, on négligera systématiquement ce terme de chauffage, excepté dans le cas du test de choc, où l'évolution entropique a une importance particulière (Fig. 25).

J'ai de plus introduit la force de Coriolis dans son expression générale. Dans les faits, l'axe de rotation sera systématiquement sur l'axe z , ce qui simplifiera l'expression de cette dernière. Enfin, j'utilise un coefficient adiabatique $\gamma = c_p/c_v$ dans l'équation d'énergie, ce qui suppose une équation d'état de type gaz parfait. On utilisera systématiquement $\gamma = 5/3$, sauf précision contraire. On pourra dans certains cas simplifier le système en supposant que le fluide est isotherme, on posera alors $P \propto \rho$ et on ne résoudra que l'équation de conservation de la masse.

§ 25. Conditions aux limites

Jusqu'à présent, je n'ai pas discuter des conditions aux limites. Elles font cependant partie intégrante du problème différentiel et ne doivent pas être négligées. Je vais donc présenter ici un aperçu des conditions aux limites utilisées dans mes simulations.

Tout d'abord, rappelons que les formules différences finies ne sont pas locales au sens où on l'entend en physique. En effet, un calcul de dérivée nécessite de connaître la valeur de la fonction aux points voisins du point de calcul. En pratique, sur une grille donnée, le calcul

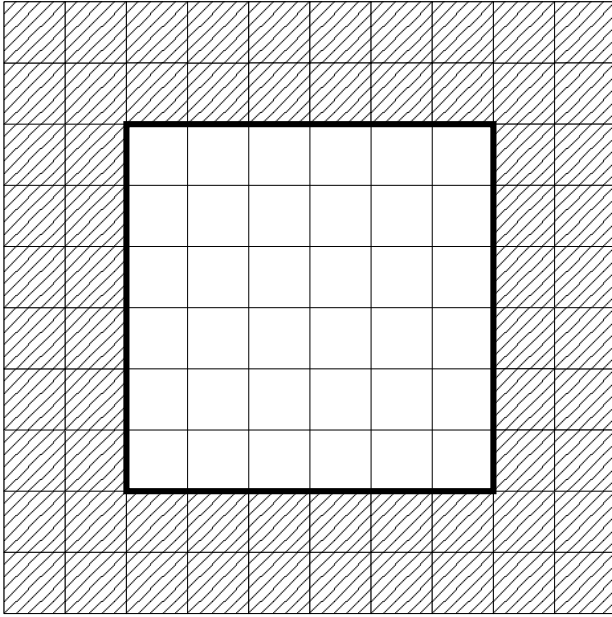


FIG. 22. Principe des zones fantômes pour traiter les conditions aux limites. L'espace physique sur lequel les équations sont effectivement résolues est délimité en gras, et les zones fantômes sont hachurées. On a représenté ici une grille avec 2 zones fantômes, bien que ce nombre puisse varier suivant l'ordre des formules différences finies utilisées.

des dérivées en bordure de grille fait appel à des points qui sont *a priori* en dehors de la grille. Pour résoudre ce problème, on fait donc appel à une série de zones fantômes ou *ghost zones*, qui sont réparties autour de la grille de calcul (voir Fig. 22). Ces zones fantômes sont initialisées manuellement à chaque pas de temps, et jouent donc le rôle des conditions aux limites que l'on rencontre en physique. Le nombre de zones à avoir au voisinage d'un point dépend directement de la formule aux différences finies utilisée, et sera à traiter au cas par cas. En pratique, plus l'ordre de la formule est élevé, plus le nombre de zones fantômes devra être important.

Je n'utiliserai que deux types de conditions aux limites, les conditions aux limites rigides, ainsi que les conditions aux limites shearing sheet, qui sont un type particulier de conditions aux limites périodiques. Je vais donc décrire ici les méthodes d'initialisation numériques pour ces deux types de conditions aux limites.

§ 25.1. Conditions aux limites rigides

Dans le cas des conditions aux limites rigides, on considérera que le fluide est « collé » à la paroi, c'est-à-dire que la vitesse parallèle au mur est égale à celle du mur. Ainsi, dans les zones fantômes, on imposera :

$$v^{\parallel} = v_{\text{mur}} \quad (25.94)$$

Cependant, lorsque les murs imposent un cisaillement linéaire au fluide, on ne pourra pas utiliser cette relation. En effet, en théorie, un cisaillement linéaire annule le terme de viscosité de l'équation de Navier-Stokes. Cependant, si on utilise la prescription proposée ci-dessus, on induit une cassure du profil linéaire dans les zones fantômes, ce qui induit une légère dissipation au voisinage des murs. Dans ce cas, on choisira donc d'imposer le profil linéaire attendu dans les zones fantômes. On écrira donc dans ces dernières :

$$v_i^{\parallel} = S y_i \quad (25.95)$$

où S est le cisaillement de l'écoulement laminaire, et y_i est la position spatiale de la zone fantôme dans la direction du cisaillement.

Pour la vitesse perpendiculaire à l'écoulement, on utilisera des conditions de réflexion sur le mur, de la même manière que Stone & Norman (1992). En supposant que $i \geq 0$ délimite les zones actives de la grille et $i < 0$ les zones fantômes, on écrira pour $i > 0$:

$$v_{-i}^{\perp} = -v_i^{\perp} \quad (25.96)$$

De même, la condition de réflexion sur les murs se traduit pour la pression et la densité par :

$$\rho_{-i} = \rho_i \quad (25.97)$$

$$P_{-i} = P_i \quad (25.98)$$

§ 25.2. Conditions aux limites shearing sheet

Les conditions aux limites shearing sheet sont une classe particulière de conditions aux limites périodiques, où l'on tient compte du cisaillement moyen du fluide. Rappelons que les conditions aux limites périodiques s'écrivent très simplement sur une grille numérique. Dans le principe, on recopie un bord de la zone active dans les zones fantômes du bord opposé. Ainsi, si on suppose que la grille s'étend des indices 0 à $n - 1$, on écrira pour les zones fantômes à gauche de la zone active :

$$\psi_{-i} = \psi_{n-i} \quad (25.99)$$

avec $i > 0$. On aura de la même manière à droite de la zone active :

$$\psi_{i+n-1} = \psi_{i-1} \quad (25.100)$$

avec là aussi $i > 0$.

Lorsque l'écoulement est cisailé, on ne peut pas appliquer directement les conditions aux limites périodiques dans la direction du cisaillement. En effet, ces dernières ne tiennent pas compte de l'advection par le cisaillement moyen. On peut alors utiliser des murs dans cette direction : c'est l'écoulement de Couette plan. Cependant, dans un vrai disque d'accrétion, les particules fluides sont libres de se déplacer dans la direction radiale. Il faut donc trouver des conditions aux limites qui autorisent à l'écoulement de sortir ou d'entrer radialement : ce sont les conditions *shearing sheet*. Ces conditions sont très similaires à des conditions aux limites périodiques, mais tiennent compte de l'advection due au cisaillement. Elles ont été développées initialement par Goldreich & Lynden-Bell (1965) pour l'étude locale des ondes spirales galactiques. Elles ont ensuite été utilisées numériquement par Hawley *et al.* (1995) qui les ont appliquées aux disques.

Le principe de ces conditions aux limites est assez simple : de la même manière que pour des conditions aux limites périodiques, on recopie la boîte de simulation de chaque côté de celle ci, dans la direction du cisaillement. Cependant, on tient compte du cisaillement moyen en décalant les boîtes recopiées (Fig. 23). Il convient alors de remarquer que ce décalage induit une nouvelle erreur numérique. En effet, les valeurs des zones fantômes dépendent de deux points de la zone active (voir Fig. 24). Il faut donc effectuer une interpolation des points de la zone active pour pouvoir initialiser correctement les zones fantômes. En pratique, il conviendra d'utiliser une méthode d'interpolation d'ordre au moins égale à l'ordre spatial du code utilisé, sous peine d'introduire de nouvelles erreurs d'arrondis. De plus, pour les mêmes raisons que l'équation linéaire d'advection, il conviendra d'utiliser une méthode d'interpolation donnant naissance au minimum d'oscillations possible (l'idéal étant une interpolation monotone).

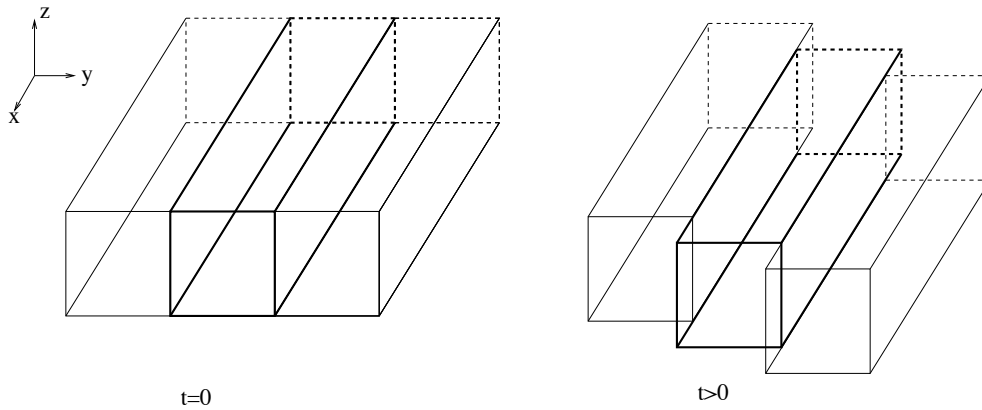


FIG. 23. Principe des conditions aux limites shearing sheet : la boîte de simulation est recopiée de part et d'autre de la boîte calculée (en gras), en prenant en compte un décalage dû au cisaillement moyen.

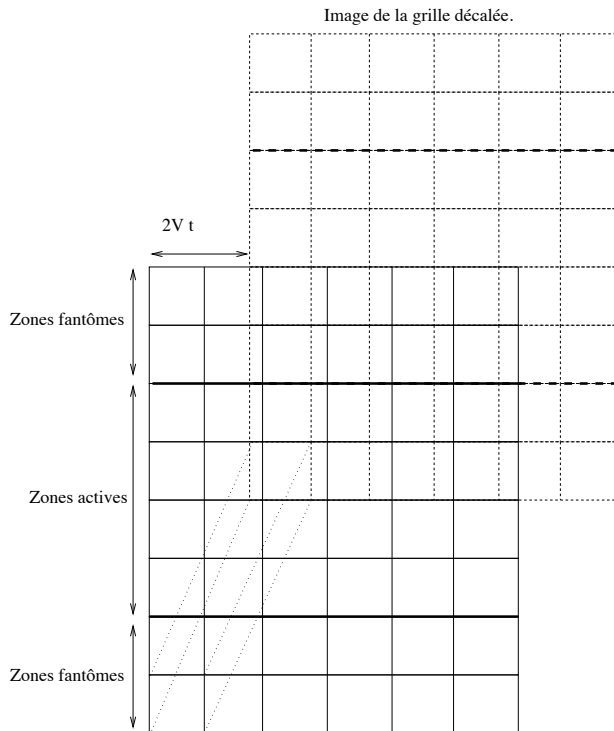


FIG. 24. Exemple de mise en œuvre des conditions aux limites shearing sheet avec un code utilisant 2 zones fantômes. On voit que la zone fantôme en bas à gauche doit être mise à jour à l'aide d'une interpolation entre 2 zones actives (en pointillé).

§ 26. Tests

Les tests pouvant être effectués sur ce type de code sont nombreux (voir par exemple Stone & Norman 1992). A titre d'exemple, je montrerai ici le résultat du test du tube de choc unidimensionnel (Sod 1978), couramment utilisé pour discriminer les méthodes aux volumes finis. Ce test permet d'obtenir le comportement d'un code vis-à-vis de différentes ondes de choc rencontrées dans le problème de Riemann. Il est certain que mon objectif n'étant pas de traiter des chocs, mes résultats ne pourront pas concurrencer les codes Godunov adaptés à ce type de physique. Les résultats de ces tests sont donc fournis à titre purement indicatif sur la figure (25).

On remarquera une fois de plus l'apparition d'oscillations dues à la méthode spatiale d'ordre élevé. Cependant, l'essentiel de la physique contenue dans ces chocs reste correct. En particulier l'onde de choc (première discontinuité sur la droite) avance à la vitesse prévue et le saut

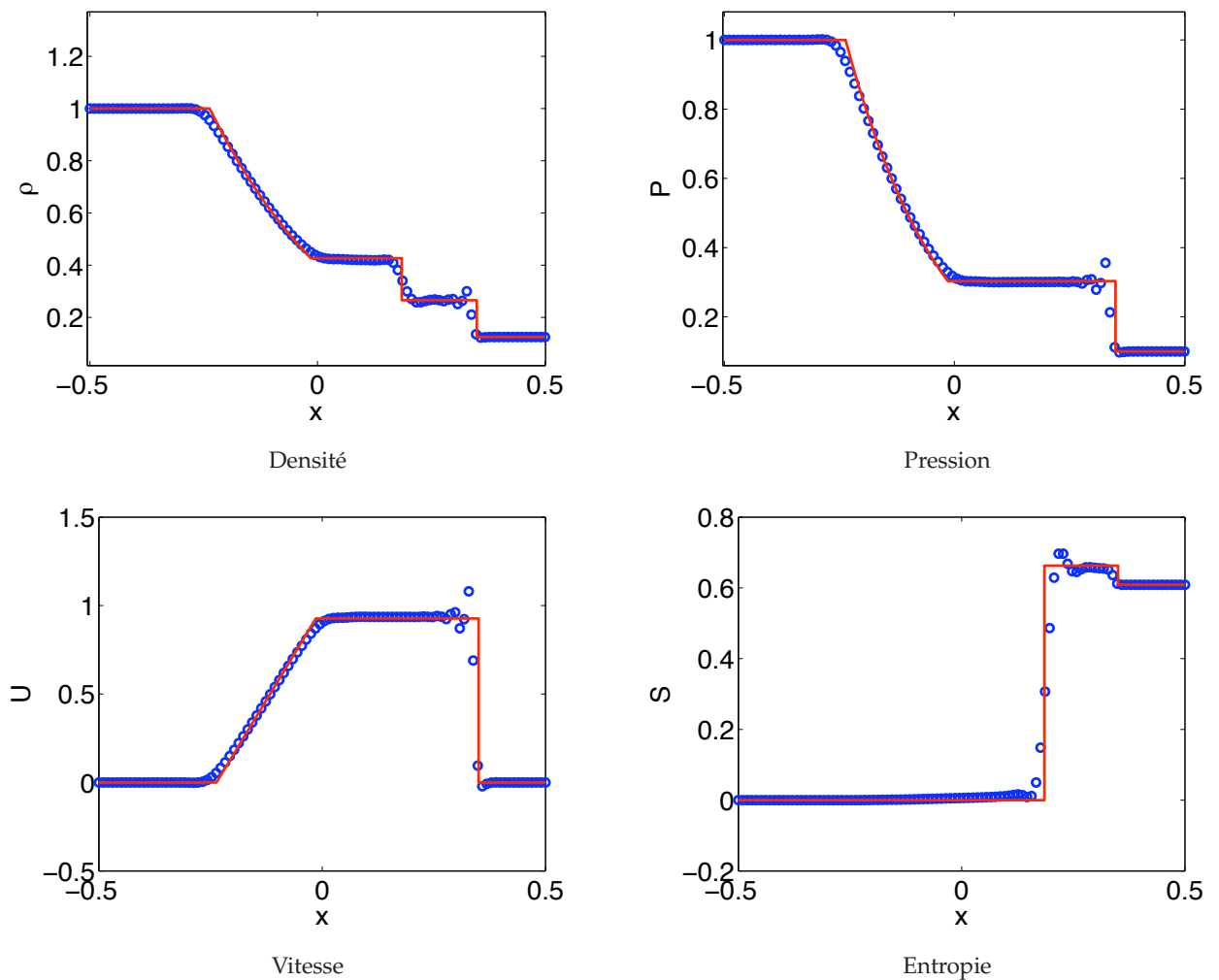


FIG. 25. Test du tube de choc pour $\gamma = 1.4$ et 100 points de grille à $t = 0.2$. La limite entre les deux milieux est fixée initialement en $x = 0$. On a choisi comme conditions initiales $\rho = 1$, $P = 1$ à gauche et $\rho = 0.125$, $P = 0.1$ à droite.

d'entropie en aval du choc correspond à la théorie. Notons que cet effet est obtenu grâce à l'utilisation d'une viscosité physique faible mais non nulle ($\nu = 10^{-3}$) dans la simulation, qui engendre naturellement un chauffage.

5. Magnétohydrodynamique

§ 27. Le problème de la divergence de B

La plupart des gaz astrophysiques sont ionisés, comme on l'a vu dans la première partie. Ainsi, la plupart des simulations effectuées en astrophysique font intervenir un couplage entre la matière et le champ magnétique sous la forme des équations de la MHD. Une des plus grandes difficultés dans ces simulations vient de la condition sur le champ magnétique :

$$\nabla \cdot B = 0 \quad (27.101)$$

Cette condition peut être satisfaite *a priori* sur les conditions initiales. De plus, l'équation de Faraday pour l'évolution du champ magnétique devrait conserver la divergence de B . Ceci n'est

pas vrai numériquement et les erreurs de troncature engendrent une croissance progressive de la divergence, ce qui forme des « monopoles » magnétiques, sources de nombreux problèmes numériques. De nombreuses méthodes ont été développées depuis les années 1980 pour contrôler ce type de phénomène. Citons par exemple le transport contraint (Evans & Hawley 1988) qui a rencontré un grand succès avec le code Zeus. Cependant, cette méthode est peu adaptée aux schémas d'ordre élevé, et j'ai préféré utiliser une méthode basée sur le potentiel vecteur.

En effet, on peut remarquer que l'on peut écrire l'équation de Faraday en fonction d'un potentiel vecteur \mathbf{A} sous la forme :

$$\frac{\partial \mathbf{A}}{\partial t} = \mathbf{v} \times \mathbf{B} + v_m \Delta \mathbf{A} \quad (27.102)$$

en imposant la condition $\mathbf{B} = \nabla \times \mathbf{A}$. Ainsi, en faisant évoluer \mathbf{A} plutôt que \mathbf{B} , on évite le problème de la divergence de \mathbf{B} , qui est maintenue naturellement au niveau de précision de la machine. Cependant, on remarquera que \mathbf{A} n'est pas unique : il dépend d'un choix de jauge tel que :

$$\mathbf{A}' = \mathbf{A} + \nabla f \quad (27.103)$$

où f est une fonction quelconque. Ce choix de jauge doit être cohérent avec la physique du problème, sans quoi des instabilités numériques peuvent apparaître, comme je vais le montrer par la suite.

§ 28. Choix de Jauge

En réécrivant l'équation (27.102) en composantes avec $v_m = 0$, il vient naturellement :

$$\partial_t A_i + v_j \partial_j A_i = v_j \partial_i A_j \quad (28.104)$$

On y retrouve un terme de transport et un terme mixte qui correspond à l'élongation des tubes de champ. Cependant, cette équation montre clairement que chaque composante est transportée *uniquement* dans les directions qui lui sont transverses. Ceci pose problème dans un disque d'accrétion. En effet, un disque est un fluide en rotation différentielle : quel que soit le référentiel que l'on choisira, il y aura toujours du fluide en mouvement dans la direction azimutale. En utilisant l'équation (28.104), la composante azimutale du potentiel vecteur ne sera pas transportée avec le fluide ! Ceci posera de graves problèmes si l'on utilise des conditions aux limites locales du type shearing sheet, qui supposent que toutes les quantités sont transportées par l'écoulement cisailé. Il convient donc de trouver un choix de jauge plus adéquat à ce problème. Pour se faire, on peut remarquer que :

$$v_j \partial_i A_j = \partial_i v_j A_j - A_j \partial_i v_j \quad (28.105)$$

Ce qui permet de réécrire (28.104) en suivant Brandenburg *et al.* (1995) avec la nouvelle jauge $f = \mathbf{v} \cdot \mathbf{A}$:

$$\partial_t A'_i + v_j \partial_j A'_i = -A'_j \partial_i v_j \quad (28.106)$$

qui inclut cette fois ci un terme de transport dans toutes les directions pour chaque composante de \mathbf{A} .

Il se trouve qu'une telle formulation implémentée en différences finies ne marche pas. En effet, l'équation (28.105), bien que valide analytiquement, n'est pas vérifiée numériquement (on

pourra facilement s'en convaincre en utilisant les formules centrées d'ordre 2). Numériquement, un changement de jauge introduit donc des termes non physiques dont le contrôle est très hasardeux. Une simple simulation d'une onde d'Alfvén avec la jauge (28.106) conduit à une explosion du code extrêmement rapidement. Il faut donc garder au maximum la physique de la MHD dans la jauge que l'on utilise. Pour se faire, on peut décomposer l'écoulement en une composante laminaire et une composante perturbative (on ne suppose pas la perturbation petite). On écrira donc pour le champ de vitesse :

$$\mathbf{v} = \mathbf{V} + \mathbf{u} \quad (28.107)$$

où \mathbf{V} est l'écoulement laminaire. Je propose alors d'utiliser la jauge (28.106) pour l'écoulement laminaire et la jauge (28.104) pour les perturbations, ce que l'on écrira

$$\partial_t A_i + V_j \partial_j A_i = -A_j \partial_i V_j + (\mathbf{u} \times \mathbf{B})_i \quad (28.108)$$

Ce type de formulation permet d'introduire un transport du potentiel vecteur par l'écoulement laminaire cohérent avec les autres champs, tout en calculant de manière aussi précise que possible la physique de l'interaction entre les perturbations et le champ magnétique. De plus les tests que j'ai effectué sur cette formulation ont montré une stabilité bien meilleure que la simple formulation (28.106).

6. Parallélisation

§ 29. Choix d'une méthode de parallélisation

L'emploi de plusieurs processeurs en parallèle est aujourd'hui nécessaire si l'on souhaite obtenir les résultats de simulations en des temps raisonnables. On fait alors appel à des méthodes de parallélisation qui dépendent de la nature du code, mais aussi de l'architecture des machines à notre disposition. Globalement, il existe 2 principes de parallélisation :

Méthodes à mémoire partagée: Dans ces méthodes, chaque processeur a accès à la totalité de la mémoire. On peut alors faire exécuter chaque boucle séparément par chaque processeur, lesquels stockeront leur résultat dans la mémoire centrale. C'est la méthode de parallélisation la plus simple mais aussi la moins efficace : elle devient inintéressante au delà de 8 processeurs. De plus, elle nécessite une architecture à mémoire partagée assez coûteuse. Le protocole *OpenMP* est l'implémentation la plus connue de méthode de parallélisation à mémoire partagée.

Méthodes à mémoire distribuée: Ces méthodes supposent que chaque processeur constitue une machine indépendante avec sa mémoire propre. On doit alors séparer le code en différents morceaux indépendants qui peuvent être résolus séparément. Chacun des processeurs doit évidemment communiquer une partie de ses résultats aux autres : c'est la méthode de parallélisation par messages. Cette méthode, bien que nécessitant des modifications plus profondes dans la construction d'un code, est aussi la plus efficace (des parallélisations sur plusieurs milliers de processeurs ont été effectués avec succès). Remarquons de plus que ce type de méthode peut être utilisé aussi avec des machines à mémoire partagée. L'implémentation la plus connue de ce type de méthode est le protocole *MPI* (Message Passing Interface).

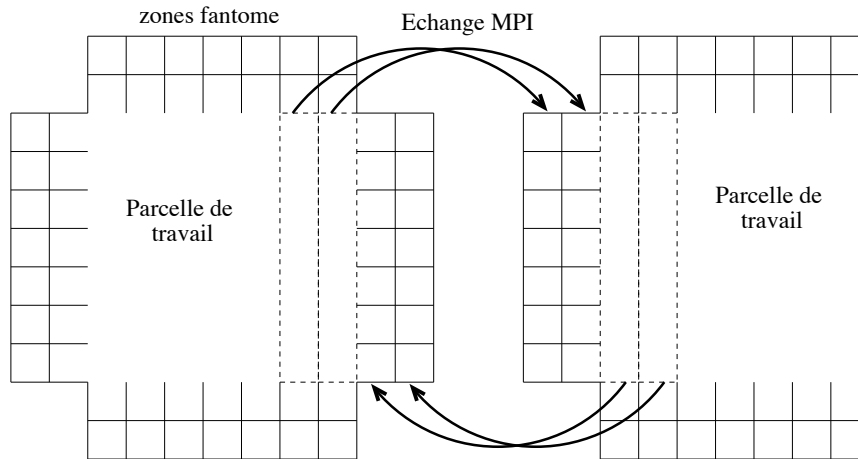


FIG. 26. Principe de la décomposition de domaine : à chaque pas de temps, les zones fantômes sont mises à jour à partir des informations aux frontières des parcelles voisines.

Des premiers tests ont été effectués en parallélisant notre code avec une méthode *OpenMP*. Il est alors apparu rapidement que les gains en vitesse d'exécution étaient trop faibles pour pouvoir répondre aux questions posées dans le temps imparti pour ce travail de thèse. J'ai donc décidé d'utiliser une parallélisation MPI basée sur la décomposition de domaine.

§ 30. Décomposition de domaine

La méthode de décomposition de domaine consiste à décomposer l'espace physique discrétisé en parcelles identiques. La résolution des équations sur chacune des parcelles est alors dévolue à un processeur. De la même manière que pour les conditions aux limites, les versions numériques des équations de la dynamique ne sont pas rigoureusement locales, et la résolution des équations dans une parcelle fait *nécessairement* intervenir des points situés dans d'autres parcelles : c'est là qu'intervient le protocole d'échange de messages.

Ainsi, on définit une série de zones fantômes entourant chaque parcelle. Ces zones fantômes sont l'image des frontières des parcelles voisines¹² et ne sont qu'un support permettant la résolution de l'équation différentielle à l'intérieur d'une parcelle. A chaque pas de temps, les données des zones fantômes sont mises à jour en récupérant le résultat de l'évolution de chaque parcelle voisine (voir Fig. 26).

On comprend alors facilement que le nombre de zones fantômes à échanger est proportionnel à l'ordre de la méthode différences finies utilisée dans la résolution de l'équation d'évolution. Ainsi, pour la méthode upwind d'ordre 4 que nous utilisons, on a besoin de connaître jusqu'à trois points en amont du point de calcul, on utilisera donc 3 zones fantômes dans chaque direction.

Ce point montre que l'on ne peut pas utiliser des méthodes aux différences finies d'ordre très élevé tout en utilisant une parallélisation de type *MPI*. En effet, ces méthodes requièrent

¹²Sauf pour les parcelles situées sur les limites du domaine de résolution, où les zones fantômes sont initialisées à partir des conditions aux limites.

l'échange d'un grand volume de données, ce qui est très pénalisant pour la vitesse finale d'exécution d'un code.

5

Méthodes spectrales

Plan du chapitre

1. Fondements.....	59
§ 31. Présentation générale des méthodes spectrales.....	60
§ 32. Base de Fourier.....	60
§ 32.1. Définition.....	60
§ 32.2. Comparaison entre méthode de Galerkin et méthode de collocation.....	61
2. L'équation d'advection.....	62
§ 33. Stabilité et condition CFL.....	62
§ 34. Test d'advection.....	62
3. Non linéarités et méthodes spectrales.....	63
§ 35. Méthodes pseudo-spectrales.....	63
§ 36. Repliement spectral et dealiasing.....	64
4. Traitement du cisaillement.....	66
§ 37. Système de coordonnées cisillées.....	66
§ 38. Procédure de remappage.....	67
§ 38.3. Pertinence physique des ondes cisillées.....	67
§ 38.4. Description du remappage.....	69

1. Fondements

LES MÉTHODES SPECTRALES ont commencé à être utilisées lors de l'apparition des premiers algorithmes de transformée de Fourier rapide (Cooley & Turkey 1965). On notera cependant que l'algorithme avait déjà été décrit au XIX^e siècle (Gauss 1866). Depuis, elles ont souvent été utilisées dans les problèmes de turbulence homogène isotrope (Orszag & Patterson 1972) car elles permettent d'obtenir une vitesse de convergence très élevée et donc une excellente résolution des petites échelles. Je ne présenterai ici que les caractéristiques générales des méthodes spectrales, ainsi que les algorithmes que j'ai spécifiquement développés pour la problématique des disques. On pourra trouver une documentation complète sur ces méthodes appliquées à la mécanique des fluides dans Canuto *et al.* (1988) et Peyret (2002) dans le cas incompressible.

§ 31. Présentation générale des méthodes spectrales

Le principe de base des méthodes spectrales est de décomposer la solution u d'une équation différentielle sur une base (infinie) de fonctions orthogonales :

$$u(x) = \sum_{k=0}^{\infty} \tilde{u}_k \phi_k(x) \quad (31.109)$$

où les fonctions ϕ_k sont une base orthogonale de l'espace des fonctions sur (x_0, x_1) munies d'un produit scalaire, c'est-à-dire :

$$(\phi_k, \phi_l) = \int_{x_0}^{x_1} \phi_k(x) \phi_l^*(x) w(x) dx = \delta_{kl} \quad (31.110)$$

où w est la fonction de poids associée au produit scalaire. On peut alors vérifier aisément que les \tilde{u}_k s'obtiennent par projection de u sur la base :

$$\tilde{u}_k = \int_{x_0}^{x_1} u(x) \phi_k^*(x) w(x) dx \quad (31.111)$$

La discrétisation numérique consiste alors à tronquer la base des ϕ_k de sorte que l'on écrive la solution sous la forme :

$$u_d(x) = \sum_{k=0}^N \hat{u}_k \phi_k(x) \quad (31.112)$$

où on notera que les coefficients \hat{u}_k sont *a priori* différents de ceux de la décomposition exacte (31.109). On peut alors obtenir 2 types d'approximation u_d de la solution exacte u .

Méthode de Galerkin: Dans cette méthode, on cherche à obtenir la projection exacte de u sur la base tronquée, c'est à dire $\hat{u}_k = \tilde{u}_k$ pour $k \in [0, N]$. Cette méthode présente l'avantage de fournir comme solution une fonction de l'espace réel, qui peut donc être calculée sur une grille spatiale aussi fine que voulue.

Méthode de collocation: Cette méthode se base sur une approche similaire aux méthodes de différences finies : on se fixe une famille de points dans l'espace x_k , $k = 0 \dots N$ et on impose $u_d(x_k) = u(x_k)$. Ces points sont alors appelés points de collocation. Cette approche permet une utilisation intuitive et assez simple des bases spectrales comme nous le verrons dans la suite.

Ces deux méthodes ne sont pas rigoureusement identiques et ne donneront donc pas les mêmes résultats. Le choix de l'une ou l'autre dépendra de l'équation à traiter mais aussi de la nature de la base utilisée. Dans la suite, nous utiliserons les bases de Fourier pour illustrer les différences et expliquer les choix qui ont été fait dans ce travail de thèse.

§ 32. Base de Fourier

§ 32.1. Définition

La base de Fourier est un ensemble de fonctions trigonométriques, périodiques sur $[0, 1]$, que l'on peut l'écrire sous la forme :

$$\phi_k(x) = \frac{1}{\sqrt{2\pi}} \exp(i2\pi kx) \quad (32.113)$$

Cette base est très utile lorsque les solutions d'une équation sont recherchées avec des conditions aux limites périodiques. Étant donné la décomposition (31.112), cette condition de périodicité

est automatiquement satisfaite *sans contrainte supplémentaire*. On pourra vérifier aisément que la relation d'orthogonalité (31.110) est respectée par la base de Fourier avec une fonction de poids $w = 1$.

§ 32.2. Comparaison entre méthode de Galerkin et méthode de collocation

On l'a vu, les coefficients \hat{u}_k^G de la méthode de Galerkin sont directement donnés par :

$$\hat{u}_k^G = \tilde{u}_k = \int_0^1 u(x) \exp(-i2\pi kx) dx \quad (32.114)$$

Par ailleurs, définissons une série de points de collocation x_j tels que :

$$x_j = \frac{j}{N} \quad (32.115)$$

On peut alors décrire en chacun de ces points, en utilisant la méthode de collocation :

$$u_d(x_j) = u(x_j) \quad (32.116)$$

$$\sum_{k=-N/2}^{N/2} \hat{u}_k^C \exp\left(i2\pi \frac{kj}{N}\right) = \sum_{k=-\infty}^{\infty} \tilde{u}_k \exp\left(i2\pi \frac{kj}{N}\right) \quad (32.117)$$

où les \hat{u}_k^C sont les coefficients de la série de Fourier issus de la décomposition par collocation. On peut alors utiliser une intégration discrète pour projeter (32.117) sur ϕ_m :

$$\frac{1}{N} \sum_{j=1}^N \sum_{k=-N/2}^{N/2} \hat{u}_k^C \exp\left(i2\pi \left(\frac{kj - jm}{N}\right)\right) = \frac{1}{N} \sum_{j=1}^N \sum_{k=-\infty}^{\infty} \tilde{u}_k \exp\left(i2\pi \frac{kj - jm}{N}\right) \quad (32.118)$$

Or

$$\frac{1}{N} \sum_{j=1}^N \exp\left(i2\pi j \frac{k - m}{N}\right) = \begin{cases} 1 & \text{si } k - m = pN, p \in \mathbb{Z} \\ 0 & \text{sinon} \end{cases} \quad (32.119)$$

d'où l'on déduit la relation entre les coefficients de la décomposition exacte et ceux de la collocation :

$$\hat{u}_m^C = \tilde{u}_m + \sum_{p=1}^{\infty} \tilde{u}_{pN+m} \quad (32.120)$$

Ce résultat se comprend assez facilement en remarquant qu'un signal de fréquence $pN + m$ effectuera p oscillations de plus qu'un signal de fréquence m entre les points j et $j + 1$. Ainsi, ces deux signaux, bien que de fréquences différentes pourront avoir les mêmes valeurs aux points de collocation. Ce phénomène est à l'origine du problème d'aliasing dans les méthodes spectrales, dont je reparlerai par la suite.

On voit donc que les méthodes de collocation, bien que pratiques car utilisant l'espace réel des solutions, font apparaître des termes spectraux de haute fréquence qui ne sont pas forcément pertinents physiquement. On préférera donc, lorsque c'est possible, utiliser une méthode de Galerkin, au contenu fréquentiel beaucoup mieux contrôlé. C'est le choix qui a été fait pour ce travail.

Dans la suite du texte, on considérera par simplicité une base sous la forme $\phi_k(x) = \exp(ikx)$, sans que ceci ne modifie les résultats.

2. L'équation d'advection

§ 33. Stabilité et condition CFL

Comme dans le cas différences finies, nous allons nous intéresser dans un premier temps à l'équation linéaire d'advection, en utilisant la méthode de Galerkin. Cette méthode présente l'avantage de transformer tous les opérateurs de dérivation spatiale en produit algébrique. Ainsi, on écrira l'équation d'advection (17.73) pour $u(x)$ avec la méthode de Galerkin sous la forme :

$$\partial_t \sum_{k=0}^N \hat{u}_k \phi_k(x) = -c \partial_x \sum_{k=0}^N \hat{u}_k \phi_k(x) \quad (33.121)$$

En remarquant que $\partial_x \phi_k(x) = ik \phi_k(x)$, une projection sur chacune des fonctions ϕ_k nous permet d'écrire pour tout k :

$$\partial_t \hat{u}_k = -ikc \hat{u}_k \quad (33.122)$$

L'équation ainsi obtenue est très similaire à celle dérivée avec les solutions tests de Von Neumann (18.77). Cependant, nous n'avons pas considéré ici de solution particulière comme dans le cas d'une analyse de Von Neumann mais bien une solution générale u , ce qui accroît notablement la portée de l'analyse proposée ici.

La méthode de Galerkin permet donc de se ramener à une équation différentielle ordinaire sur chacun des \hat{u}_k , d'une forme similaire à (18.85). L'étude de la stabilité de cette équation pour différents schémas temporels ayant déjà été faite pour le cas des schémas différences finies (voir section § 18.2), nous ne rappellerons que les résultats principaux applicables aux schémas spectraux.

En premier lieu, notons qu'un schéma spectral suppose $f = ik\delta x$ dans (18.84) de sorte que l'argument de \mathcal{T} dans (18.86) sera toujours imaginaire pur. Ainsi, les figures (15) et (16) montrent clairement qu'une méthode spectrale ne sera stable que pour un schéma de Runge Kutta d'ordre 4 (on notera qu'un schéma Runge Kutta d'ordre 3 peut être stable pour les mêmes raisons). Cette stabilité, comme tous les schémas explicites en temps, est une stabilité conditionnelle. En effet, on peut vérifier que pour le schéma temporel de Runge Kutta d'ordre 4, le taux de croissance d'un pas de temps au suivant est donné par :

$$\tau^2 = 1 + \frac{(ck\delta t)^6}{72} \left(\frac{(ck\delta t)^2}{8} - 1 \right) \quad (33.123)$$

En définissant le plus grand vecteur d'onde accessible sous la forme $k_{\max} = 2\pi/\delta x$, ce résultat montre que la condition CFL pour un code spectral utilisant un algorithme de Runge-Kutta d'ordre 4 peut s'écrire :

$$\delta t < \frac{\sqrt{8}}{2\pi} \frac{\delta x}{c} \quad (33.124)$$

On voit donc que cette nouvelle condition CFL est 2 fois plus restrictive que la condition classique $\delta t < \delta x/c$. Ceci augmente donc naturellement le temps de calcul comparativement à une méthode de différentiation spatiale type différences finies.

§ 34. Test d'advection

Afin de comparer la capacité d'un code spectral à répondre aux différentes discontinuités pouvant être présentes dans une simulation, nous avons testé l'advection d'un créneau, dans

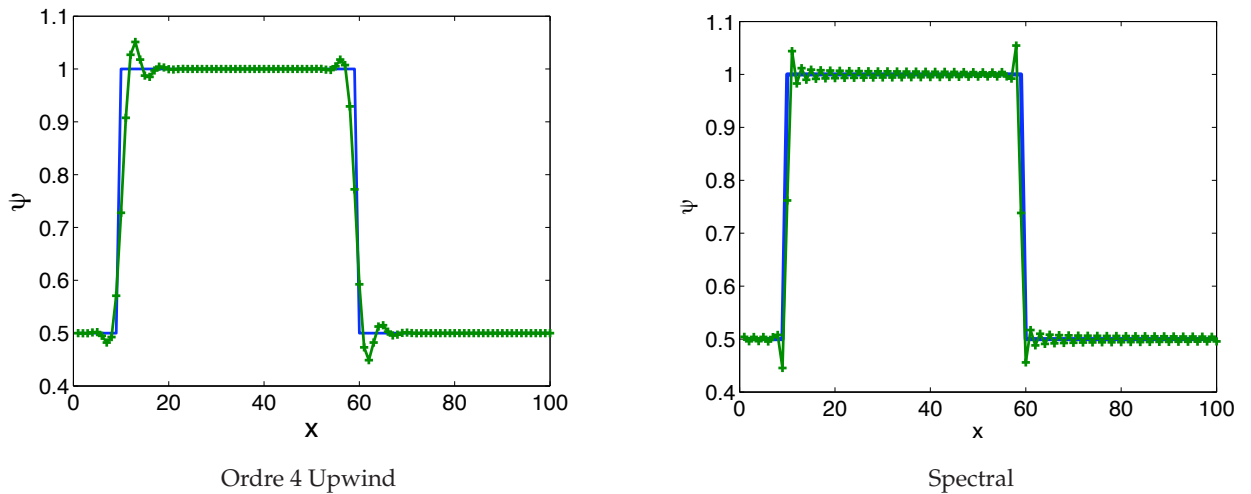


FIG. 27. Tests d'advection d'un créneau avec une méthode spectrale et un schéma de Runge-Kutta à l'ordre 4 (20 pas de temps avec $\Delta t = 0.5$ et $c = 1$). La même advection effectuée à l'aide d'un schéma différences finies est donnée à titre de comparaison.

des conditions similaires aux tests du § 21. Ces tests, présentés sur la figure (27) montrent clairement l'apparition d'oscillations dues à la présence de discontinuités. Comparativement aux schémas aux différences finies centrés (Fig. 20), le schéma arrive à maintenir un niveau d'oscillations très faible. Cependant, il semble que qualitativement, les schémas upwind d'ordre élevé arrivent à obtenir des résultats moins pollués par des problèmes numériques. Notons cependant que l'avantage du schéma spectral est son absence presque totale de diffusion (une légère diffusion venant de l'algorithme d'intégration temporelle de Runge Kutta reste néanmoins présente comme le montre l'équation 33.123). Cet avantage décisif des méthodes spectrales aura une grande importance dans l'étude de la turbulence.

3. Non linéarités et méthodes spectrales

§ 35. Méthodes pseudo-spectrales

On l'a vu dans la section précédente, les méthodes spectrales permettent de transformer très simplement une équation linéaire aux dérivées partielles en une équation ordinaire, en introduisant très peu de diffusion dans le schéma numérique. Cependant, les équations de la mécanique des fluides font appel très souvent à des termes non linéaires, plus difficiles à modéliser spectralement. Ainsi, considérons une équation modèle non linéaire du même type que l'équation de Burgers (23.92) avec $\nu = 0$:

$$\partial_t u = -u \partial_x u \quad (35.125)$$

On peut alors écrire en utilisant la représentation de Galerkin :

$$\sum_{k=0}^N (\partial_t \hat{u}_k) \phi_k = -i \sum_{l=0}^N \sum_{m=0}^N m \hat{u}_l \hat{u}_m \phi_l \phi_m \quad (35.126)$$

En remarquant que pour une base Fourier, $\phi_l \phi_m = \phi_{l+m}$, nous pouvons déduire de (35.126) l'expression de l'évolution temporelle d'un mode sous la forme :

$$\partial_t \hat{u}_k = -i \sum_{m=-N/2}^{N/2} m \hat{u}_m \hat{u}_{k-m} \quad (35.127)$$

Cette équation est donc la retranscription de l'équation de Burgers sans dissipation dans l'espace Fourier en représentation de Galerkin. On remarque par ailleurs que cette transcription fait appel à un produit de convolution : le calcul de l'évolution de chaque onde fait appel au produit de toutes les autres ondes. Le temps de calcul d'une telle équation d'évolution est donc proportionnel à N^2 , ce qui est extrêmement désavantageux en comparaison des méthodes à grille spatiale (c'est d'ailleurs la raison principale de l'absence des méthodes spectrales en mécanique des fluides dans les années 50).

Pour contourner cette difficulté, on peut faire appel aux méthodes pseudo-spectrales : plutôt que de calculer directement le produit de convolution (35.127) dans l'espace spectral, on choisit de calculer les termes non linéaires dans l'espace réel. L'équation (35.127) se transforme alors en un simple produit algébrique. Ce type de méthode est avantageuse dans la mesure où la transformée permettant le passage de l'espace réel à l'espace spectral est plus rapide que le produit de convolution. C'est le cas pour les transformées de Fourier rapides dont le temps de calcul évolue en $N \log N$.

A titre d'exemple, supposons que l'on souhaite calculer dans l'espace spectral le produit de deux fonctions f et g , dont on ne connaît que les coefficients de Galerkin \hat{f}_k et \hat{g}_k . Plutôt que de calculer la convolution de f et g , longue en terme de temps de calcul, on applique la procédure pseudo-spectrale suivante :

- (1) Passage de \hat{f} et \hat{g} dans l'espace réel à l'aide d'une transformée inverse. On obtient ainsi f et g .
- (2) Calcul du produit dans l'espace réel fg .
- (3) Passage du produit fg dans l'espace spectral à l'aide d'une transformée directe. On obtient donc finalement les coefficients $(\widehat{fg})_k$.

Cette méthode semble à priori efficace. Cependant, l'« aller-retour » dans l'espace réel engendre des effets de bord qu'il est important de maîtriser.

§ 36. Repliement spectral et dealiasing

Pour comprendre les problèmes posés par les méthodes pseudo-spectrales, considérons deux fonctions u et v dont nous connaissons la représentation de Galerkin \hat{u}_k et \hat{v}_k , et calculons leur produit en utilisant la méthode pseudo-spectrale. Tout d'abord, la transformée discrète inverse nous permet d'écrire pour u et v :

$$u(x) = \sum_{k=-N/2}^{N/2} \hat{u}_k \exp(ikx) \quad (36.128)$$

$$v(x) = \sum_{k=-N/2}^{N/2} \hat{v}_k \exp(ikx) \quad (36.129)$$

le produit de u et v dans l'espace réel s'écrit alors :

$$uv(x) = \sum_{j=-N/2}^{N/2} \sum_{m=-N/2}^{N/2} \hat{u}_j \hat{v}_m \exp[i(j+m)x] \quad (36.130)$$

Finalement, on effectue une transformée de Fourier discrète sur ce résultat en utilisant les points $x_k = k/N$. Le résultat obtenu par la méthode pseudo-spectrale s'écrit ainsi :

$$(\widehat{uv})_n = \sum_{j=-N/2}^{N/2} \sum_{m=-N/2}^{N/2} \sum_{k=1}^N \hat{u}_j \hat{v}_m \exp\left(2i\pi(j+m-n)\frac{k}{N}\right) \quad (36.131)$$

Cette somme sera non nulle uniquement pour les termes tels que :

$$j + m - n = pN \quad (36.132)$$

où $p \in N$. Remarquons alors que l'on ne retrouvera l'expression de la convolution (35.127) que si la seule solution à (36.132) est $p = 0$. Or j , m et n varient dans l'intervalle $[-N/2, N/2]$, on peut donc obtenir des combinaisons telles que p soit un entier non nul : c'est le problème du repliement spectral.

Ce phénomène peut être vu sous un angle physique assez simple : les termes non linéaires, tels que ceux présents dans les équations de transport, combinent deux à deux les ondes pour en produire de nouvelles. Les vecteurs d'onde ainsi obtenus sont soit la somme, soit la différence des vecteurs des ondes initiales. Les ondes créées qui ont une fréquence spatiale supérieure à la fréquence de Nyquist sortent alors du domaine spectral $[-N/2, N/2]$ et devraient en toute rigueur disparaître. En fait, la méthode pseudo-spectrale garde une trace de ces ondes en les faisant réapparaître dans le domaine spectral à un vecteur d'onde inférieur $k' = k - pN$ (voir Fig. 28).

Ce phénomène donne alors lieu à l'apparition d'ondes parasites de fréquences élevées dans la simulation, pouvant mener jusqu'au *crash* du code utilisé. Il est donc important de contrôler, voir d'éliminer ces ondes hautes fréquences : c'est le dealiasing.

La méthode de dealiasing que j'ai utilisée est la « règle des 3/2 ». Cette méthode présente l'avantage d'être très efficace et relativement peu coûteuse en temps de calcul. L'idée est d'étendre le domaine spectral jusqu'à une taille $M > N$. Si M est suffisamment grand, les ondes créées par le terme non linéaire dans l'espace de taille N peuvent toutes être représentées dans l'espace de taille M . En pratique, on se définit donc un espace spectral de vecteurs d'onde $[-M/2, M/2]$. Les ondes appartenant aux domaines $[-M/2, -N/2 - 1]$ et $[N/2 + 1, M/2]$ voient leur amplitude fixée à 0. On obtient ainsi la même représentation que dans l'espace N . Un calcul similaire à (36.131) montre alors que la méthode pseudo-spectrale donnera des ondes d'amplitudes non nulles si :

$$j + m - n = pM \quad (36.133)$$

où $j, m, n \in [-N/2, N/2]$, cette équation étant l'extension de (36.132). On comprend alors que si on fixe M tel que $M > 3N/2$, le repliement spectral sera effectivement inexistant car la seule solution à (36.133) sera $p = 0$: c'est la règle des « 3/2 » (voir la deuxième partie de la figure 28).

En pratique, pour une résolution effective N , on fera tous les calculs en dimension $3N/2$, en ayant rempli de 0 le tableau de travail dans les intervalles $[-3N/4, -N/2 - 1]$ et $[N/2 + 1, 3N/4]$. A l'issue du calcul d'un terme non linéaire par une méthode pseudo-spectrale, on prendra soin de réinitialiser ces intervalles à 0. On pourrait en théorie utiliser des tableaux de

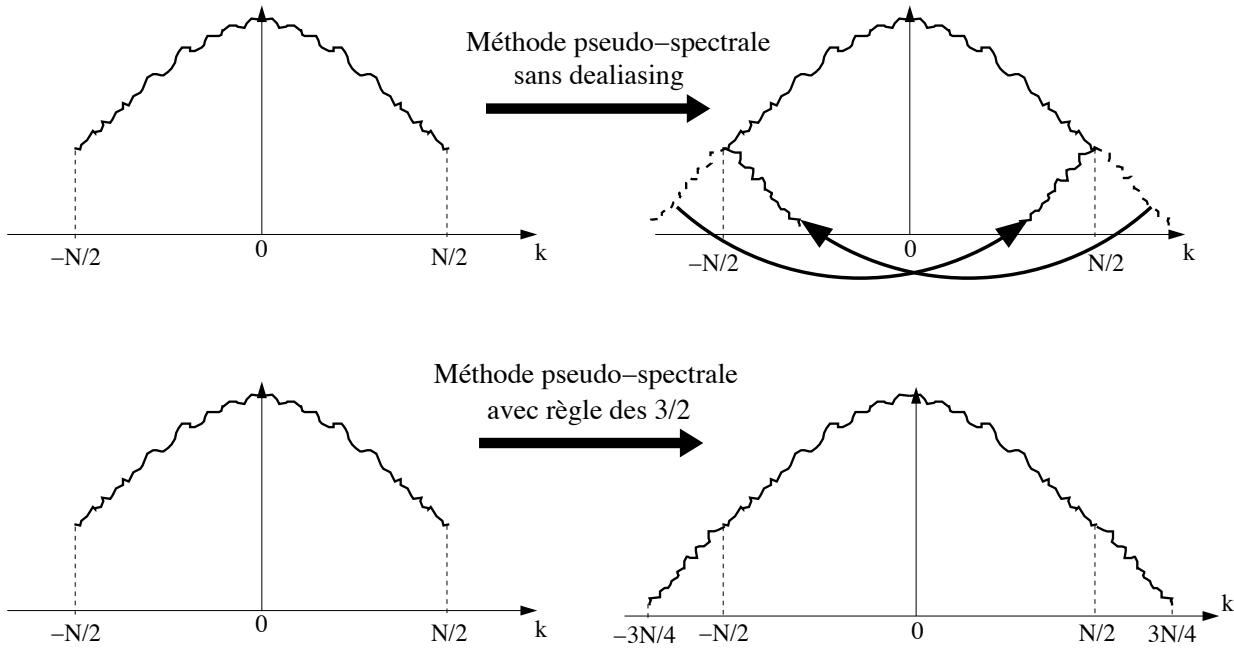


FIG. 28. Comparaison des méthodes pseudo-spectrales sans dealiasing et avec la règle des «3/2». A gauche, on donne le spectre d'une fonction dont on veut calculer le carré. A droite, dans le calcul sans dealiasing, l'énergie spectrale devant apparaître à haute fréquence se voit «repliée» vers les fréquences plus basses (flèches) : c'est l'aliasing. La règle des 3/2 contourne ce problème en allouant un espace supplémentaire pour ces ondes hautes fréquences.

taille N , et utiliser transitoirement des tableaux de dimension $3N/2$ pour calculer les termes non linéaires. Dans les faits, le passage d'une représentation de taille N à $3N/2$ est extrêmement coûteux en terme d'échange de données dans un code parallélisé, et n'est pas utilisable en pratique.

Remarquons enfin que cette méthode introduit une nouvelle dissipation numérique : les ondes hautes fréquences qui apparaissent dans le domaine spectral réinitialisé sont perdues. Cependant, le contrôle du bilan énergétique montrera que cette dissipation est en principe largement inférieure à la diffusion induite par les termes de viscosité et résistivité, et donc tout à fait négligeable.

4. Traitement du cisaillement

§ 37. Système de coordonnées cisillées

Comme nous l'avons vu, le code développé pour ce travail de thèse utilise les bases de Fourier, qui sont de nature périodique. On suppose donc *nécessairement* des conditions aux limites périodiques dans toutes les simulations effectuées avec un tel code. Cependant, la physique des disques d'accrétion s'accommode mal de ces conditions aux limites, le problème principal venant du cisaillement. Comme nous l'avons vu dans la première partie, les équations fluides dans un disque font systématiquement intervenir un terme de cisaillement de la forme :

$$\partial_t \psi + S y \partial_x \psi = \dots \quad (37.134)$$

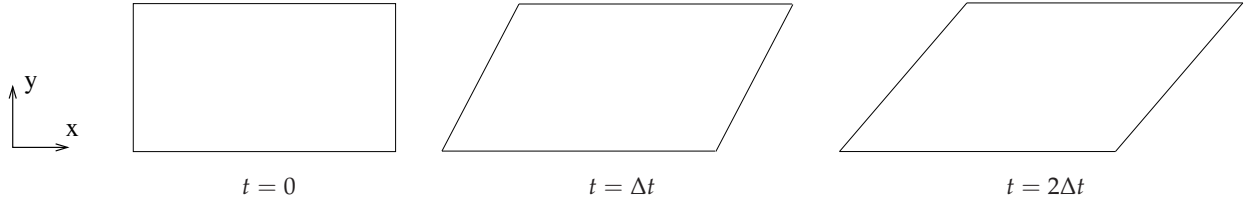


FIG. 29. Évolution en fonction du temps d'une boîte définie à partir des coordonnées cisillées (37.135).

Où $S = -rd\Omega/dr$ est le cisaillement local. Si l'on considère à présent une structure périodique dans la direction radiale à $t = 0$, le cisaillement du disque d'accrétion va naturellement déformer cette structure qui va alors immédiatement perdre sa périodicité radiale. On ne pourra donc pas approcher localement la structure d'un disque d'accrétion par une simple boîte périodique. D'un point de vue analytique, la non compatibilité des équations cisillées telles que (37.134) avec des conditions aux limites périodiques vient de la présence explicite de la coordonnée y . On remarquera en effet qu'un tel terme ne peut être représenté de manière correcte dans l'espace de Fourier.

Une méthode permettant de contourner cette difficulté est d'utiliser un référentiel cisillé défini par :

$$x' = x - Syt \quad y' = y \quad z' = z \quad (37.135)$$

Dans ce nouveau référentiel, les équations fluides peuvent être écrites sous la forme :

$$\partial_{t'}\psi = \dots \quad (37.136)$$

dans laquelle on aura pris soin de remplacer les dérivées radiales ∂_y par $\partial_{y'} - St\partial_{x'}$. La dépendance explicite en espace a donc disparue au profit d'une dépendance explicite en temps. Notons cependant que cette dernière ne pose pas de problème technique particulier pour une méthode spectrale. Physiquement, ce changement de référentiel revient à utiliser une boîte cisillée comme représentée sur la figure (29). On suppose alors que les conditions aux limites dans cette boîte sont périodiques, ce qui nous permet d'obtenir des simulations locales de disques d'accrétion (ce type d'approche est en fait en tout point identique à l'utilisation des conditions aux limites shearing sheet que l'on a discuté au chapitre précédent).

§ 38. Procédure de remappage

On pourrait *a priori* se contenter du changement de variable évoqué précédemment pour effectuer une simulation. Cependant, comme je vais le montrer, les ondes simulées par le code deviennent non pertinentes pour la physique qui nous intéresse. Il convient alors d'effectuer une procédure de rééchantillonnage que je vais décrire ici. Commençons donc par identifier le problème physique.

§ 38.1. Pertinence physique des ondes cisillées

Dans la section précédente, j'ai montré qu'il était nécessaire de résoudre les équations de la mécanique des fluides dans un système de coordonnées cisillées que l'on notera en 2D (x', y') .

Les ondes de la base de Fourier peuvent ainsi être définies avec les vecteurs d'ondes discrets $k'_x = 2\pi p/L_x$ et $k'_y = 2\pi q/L_y$ tels que :

$$\phi_{pq}(x', y') = \exp [i(k'_x x' + k'_y y')] \quad (38.137)$$

où p et q sont des entiers relatifs. En pratique, on choisira une résolution $(N_x \times N_y)$ telle que $-N_x/2 < p < N_x/2$ et $-N_y/2 < q < N_y/2$, ce qui définira un ensemble discret d'ondes (k'_x, k'_y) qui seront résolues dans une simulation. Malheureusement, comme je vais le montrer, ce schéma simple n'est pas suffisant pour obtenir une simulation physiquement pertinente.

Pour commencer, replaçons nous dans le système de coordonnées initial non cisailé (x, y) . Les ondes de la base Fourier ϕ prennent alors la forme :

$$\phi_{pq}(x, y) = \exp \left(i[k'_x x + (k'_y - Stk'_x) y] \right) \quad (38.138)$$

Les ondes cisillées (ou *shwaves*, pour « shearing waves ») ainsi obtenues peuvent être définies à partir d'un vecteur d'onde \mathbf{k} du référentiel (x, y) tel que :

$$\phi_{pq}(x) = \exp(i2\pi \mathbf{k} \cdot \mathbf{x}) \quad (38.139)$$

On a alors :

$$k_x = k'_x \quad (38.140)$$

$$k_y = k'_y - k'_x St$$

Ce résultat est l'équivalent spectral de (37.135). Il montre en effet que l'espace spectral est cisillé dans la direction k_x . En pratique, les ondes de la base de Fourier utilisées dans la simulation précédente se déplacent dans l'espace spectral défini par (k_x, k_y) selon la figure (30).

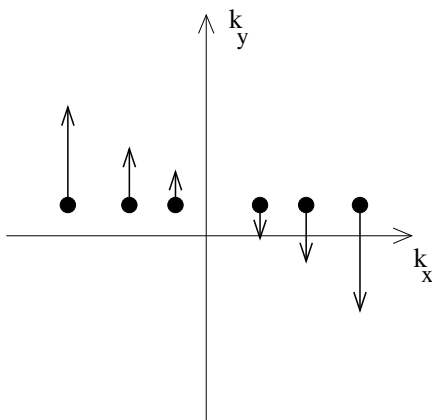


FIG. 30. Cisaillement de l'espace spectral (k_x, k_y) . Les points représentent les ondes de la base Fourier (38.137) utilisées dans le code. Ces ondes se déplacent dans l'espace spectral (k_x, k_y) (flèches) et montrent que cet espace spectral est cisillé.

De manière générale, les ondes que l'on simule peuvent être vues comme un ensemble de points de grille dans l'espace spectral. Cette grille de simulation est alors cisillée de la même manière que précédemment, et l'évolution temporelle montre que l'on aboutit à une représentation incomplète de l'espace physique (Fig. 31).

Tout d'abord, remarquons que l'on peut séparer sur cette figure deux types d'ondes : les ondes de tête, dont la fréquence spatiale $|k_y|$ diminue au cours du temps, et les ondes de traîne pour lesquels $|k_y|$ augmente¹³. On peut aussi définir une fréquence spatiale de dissipation, à la manière de ce que l'on rencontre en turbulence homogène (voir § 40 p. 79). Pour des fréquences

¹³En fait, les ondes de tête finissent forcément par devenir des ondes de traîne, mais ceci ne change rien au raisonnement.

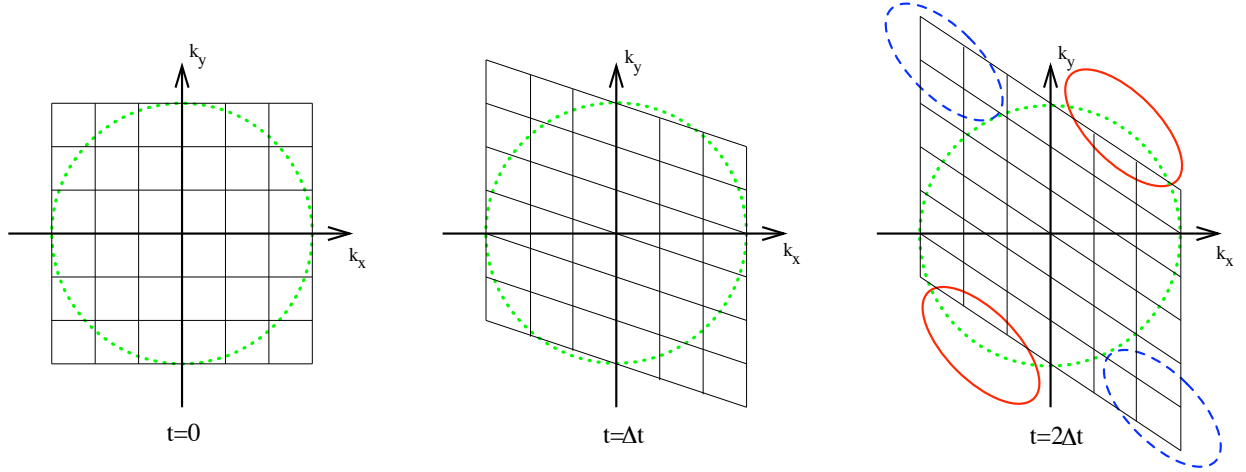


FIG. 31. Le cisaillement de la grille spectrale induit 2 effets de bords au cours du temps. D'une part, la fréquence spatiale des ondes de traîne augmente, et certaines d'entre elles passent dans le domaine dissipatif de l'écoulement (cercles tiretés). D'autre part, des ondes de tête sortent du domaine dissipatif et deviennent *a priori* pertinentes pour la physique de l'écoulement (cercles pleins). Elles ne sont cependant pas traitées par la simulation. NB : On a représenté en pointillés les limites du domaine dissipatif de l'écoulement.

spatiales plus élevées que cette fréquence de dissipation, les ondes sont essentiellement diffusées. On parlera alors de domaine dissipatif de l'écoulement.

En reprenant la figure (31), on voit que les ondes de traîne qui étaient à des fréquences élevées à $t = 0$ se retrouvent à très haute fréquence, et terminent donc dans le domaine dissipatif de l'écoulement : elles deviennent négligeables. Dans le même temps, des ondes de tête, qui étaient négligeable à $t = 0$ pour les mêmes raisons, voient leur fréquence spatiale diminuer. Elles sortent donc du domaine dissipatif et deviennent pertinentes pour la physique de l'écoulement. Cependant, n'étant pas incluses dans la simulation à $t = 0$, ces ondes ne sont pas traitées numériquement. Au bout d'un certain temps, on comprend donc que la grille de simulation définie plus haut ne représente plus la physique de l'écoulement. Il faut alors faire appel à une procédure, dite de remappage (ou *remap*), qui permet de maintenir la grille de simulation au voisinage du domaine spectral qui est physiquement pertinent.

§ 38.2. Description du remappage

Pour commencer, on définit une grille fixe dans l'espace (k_x, k_y) , qui est superposée à la grille de simulation à $t = 0$. En laissant évoluer cette dernière, les points de grille viendront se superposer de nouveau sur la grille fixe tous les $t_{\text{MAP}} = L_x / SL_y$ (Fig. 32).

La procédure de remappage consiste alors à changer les ondes représentées par la grille de simulation à $t = t_{\text{MAP}}$. On enlève donc les ondes de traîne qui sont dissipées par l'écoulement, et on les remplace par des ondes de tête. Ainsi, en utilisant la définition (38.137), on dira qu'à l'issue du remappage, la simulation représente les ondes de vecteur \mathbf{K}' , défini à partir de (38.140) par :

$$K'_x = k'_x \quad (38.141)$$

$$K'_y = k'_y - k'_x S t_{\text{MAP}} \quad (38.142)$$

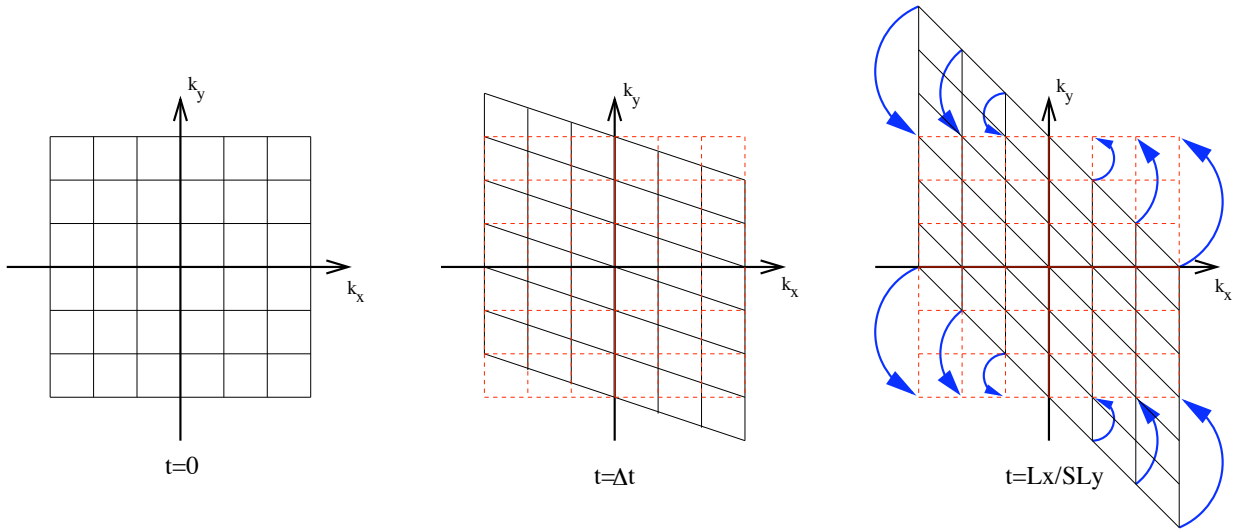


FIG. 32. Principe de la procédure de remappage : on se fixe une grille fixe superposée à la grille de simulation à $t = 0$ (en pointillés) . La grille de simulation, cisailée, vient ensuite se superposer tous les $t_{\text{MAP}} = L_x / S L_y$ sur les points de la grille fixe. On peut alors effectuer un remappage en changeant les ondes représentées par la grille de simulation (flèches).

En notant $K'_x = 2\pi p' / L_x$ et $K'_y = 2\pi q' / L_y$, la relation précédente revient à poser :

$$p' = p \quad (38.143)$$

$$q' = q - p \quad (38.144)$$

Le remappage revient donc finalement à effectuer une réallocation des ondes sur la grille, telle qu'elle est décrite sur la figure (33). On remarque alors que certaines ondes sortent de la grille pendant la procédure. Ces ondes correspondent en pratique aux ondes de traîne qui passent dans le domaine dissipatif à $t \sim t_{\text{MAP}}$ et sont donc effectivement négligeables. D'autre part, l'ajout des zéros est dû à l'inclusion des ondes de tête décrites précédemment. Si la résolution de la simulation est suffisante, ces ondes seront encore dans le domaine dissipatif à $t \sim t_{\text{MAP}}$. Ainsi, approximer leur amplitude à 0 ne sera pas pénalisant pour la suite de la simulation.

L'un des problèmes de ces simulations réside naturellement dans la dissipation induite par cette procédure. En particulier, les ondes de traîne qui sont éliminées de la simulation lors du remappage constituent une forme de dissipation numérique. Néanmoins, si la résolution est suffisante, cette perte d'énergie devrait être faible comparativement à la dissipation physique, car ces ondes seront déjà largement diffusées avant d'être éliminées de la simulation.

Un autre problème est la pertinence physique des ondes de tête qui sont incluses au fur et à mesure. En particulier, on peut s'interroger sur le comportement de ces ondes, si on les avait suivies depuis leur origine à très haute fréquence spatiale. Malheureusement, il est difficile de contrôler les artefacts introduits sur ces ondes, sinon en faisant un test de résolution. On peut néanmoins supposer que dans le domaine dissipatif, à fréquences spatiales égales, ondes de traîne et ondes de tête devraient avoir une amplitude du même ordre. Ainsi, si le remappage introduit peu de dissipation numérique, on peut supposer que l'amplitude des ondes de tête est suffisamment faible pour pouvoir être négligée. Le contrôle de la dissipation numérique introduit ainsi un contrôle partiel sur l'approximation des ondes de tête.

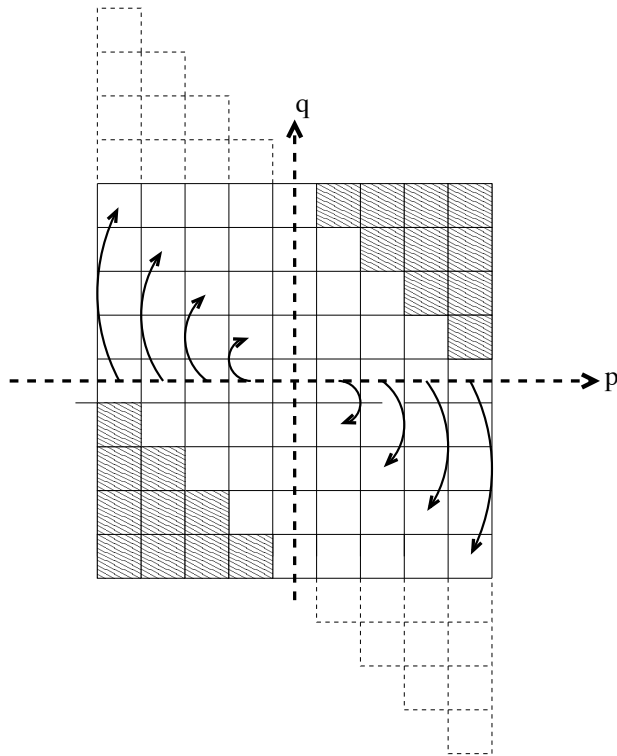


FIG. 33. Exemple de remappage d'une grille de simulation. Les éléments sont déplacés sur toutes les colonnes $p = cte$ en suivant les flèches. Les éléments qui sortent de la grille (pointillés) sont les ondes de traîne à haute fréquence, situées dans le domaine dissipatif. Les zones hachurées sont *a priori* inconnues : ce sont les ondes de tête discutées précédemment. Il faut alors que la résolution soit suffisante pour qu'elles apparaissent lorsqu'elles sont encore dans le domaine dissipatif. On peut ainsi initialiser leur amplitude à 0.

En pratique, dans les simulations que je proposerai ici, le taux de dissipation numérique totale (incluant les effets de dealiasing, *etc.*) sera systématiquement inférieur à 3% du taux de dissipation physique. On pourra donc conclure que l'influence de cette procédure est négligeable sur la dynamique générale de l'écoulement.

