

Analyse quantitative des schémas numériques pour les équations aux dérivées partielles

Daniel Bouche et William Weens



Imprimé en France

ISBN (papier) : 978-2-7598-2760-2 – ISBN (ebook) : 978-2-7598-2761-9

Tous droits de traduction, d'adaptation et de reproduction par tous procédés, réservés pour tous pays. La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (alinéa 1^{er} de l'article 40). Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles 425 et suivants du code pénal.

© EDP Sciences, 2024

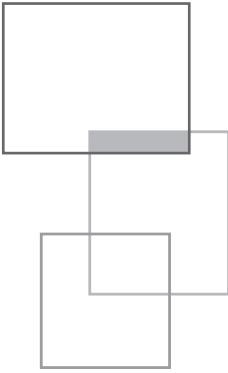


Table des matières

I	Introduction et méthodes d'analyse	1
1	Problèmes abordés	3
2	Les équations traitées	5
2.1	Équations hyperboliques	5
2.1.1	Équation d'advection	5
2.1.2	Équation des ondes	6
2.1.3	Équations non linéaires	6
2.2	Équations paraboliques	7
2.3	Fonctions de Green	7
3	Méthode des différences finies	9
3.1	Principe de la méthode	9
3.2	Ordre d'une approximation différences finies	11
3.3	Schémas pour l'équation d'advection	12
3.3.1	Ordre de consistance d'un schéma	12
3.3.2	Théorème de Lax-Richtmyer	13
3.3.3	Exemples de schémas	14

4	Méthodes d'analyse des schémas	19
4.1	Méthodes opératorielles	20
4.2	Analyse de Fourier des schémas	23
4.2.1	Les différentes formes de la transformation de Fourier	23
4.2.2	Transformation de Fourier semi-discrète	24
4.2.3	Transformation de Fourier discrète	31
4.2.4	Visualisation dans l'espace réciproque, dispersion et dissipation d'un schéma	32
4.2.5	Analyse de convergence	41
4.3	Méthode de l'équation équivalente	44
4.3.1	Première méthode : développement de Taylor et récursion	45
4.3.2	Seconde méthode : analyse de Fourier	46
4.3.3	Calcul par méthode opératorielle	48
4.3.4	Validité de l'équation équivalente à nombre fini de termes ..	49
4.3.5	Solution de l'équation équivalente à un terme pour l'équation d'advection	55
4.4	Conclusion	57
II	Applications	61
5	Schémas pour l'équation d'advection	63
5.1	Schéma upwind	63
5.1.1	Méthode opératorielle	64
5.1.2	Analyse de Fourier	67
5.1.3	Solution de l'équation équivalente	69
5.1.4	Conclusion sur les solutions exactes et approchées	75
5.2	Schéma centré semi-discrétisé	76
5.2.1	Solution exacte	76
5.2.2	Fonction de Green	76
5.3	Schéma saute-mouton	82
5.3.1	Solution par la méthode de l'équation équivalente	82
5.3.2	Solution exacte pour la fonction de Green	83
5.3.3	Approximation de la solution exacte	88

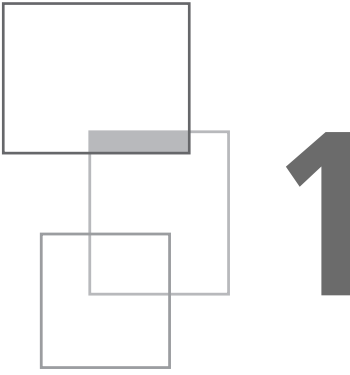
5.3.4	Solution exacte et approximation de l'équation équivalente	93
5.3.5	Conclusion sur le schéma saute-mouton	94
5.4	Conclusion sur les schémas non dissipatifs	95
5.5	Schéma de Lax-Wendroff	95
5.5.1	Solution exacte	96
5.5.2	Étude par la méthode de l'équation équivalente	98
5.6	Le schéma de Beam-Warming	105
5.6.1	Analyse du schéma	107
5.7	Propagation de fronts avec des schémas d'ordre 2 dissipatifs	109
5.7.1	Artéfacts des schémas linéaires pour une condition initiale marche	109
5.7.2	Comment obtenir des profils monotones?	111
5.8	Conclusion sur les schémas d'ordre 2	114
5.9	Schémas d'ordre 3	115
5.10	Schémas d'ordre supérieur	119
5.10.1	Schémas d'ordre élevé non dissipatifs	119
5.10.2	Schémas d'ordre élevé dissipatifs	124
5.10.3	Un exemple de schéma dissipatif d'ordre 4	125
5.10.4	Schémas de Strang d'ordre impair	128
5.10.5	Conclusion sur les schémas d'ordre élevé	128
5.11	Conclusion sur les schémas pour l'advection 1D	129
5.12	Schémas pour l'advection en 2D	130
5.12.1	Schéma upwind sur maillage carré	130
5.12.2	Schéma splitté	132
5.12.3	Fonctions de Green	133
5.12.4	Conclusion sur l'advection sur maillage carré	134
6	Schémas pour l'équation des ondes 1D	137
6.1	Équation des ondes continue	137
6.2	Schémas dérivés des schémas pour l'équation d'advection	138
6.3	Schéma de Yee en 1D	139
6.3.1	Formulation du schéma pour le système des ondes	139

6.3.2	Réinterprétation comme une discrétisation de l'équation des ondes	140
6.3.3	Conclusion	144
7	Schémas pour l'équation des ondes en 2D	145
7.1	Fonctions de Green pour l'équation des ondes continue	145
7.1.1	Source Dirac impulsive	146
7.1.2	Conditions initiales Dirac sur la dérivée	147
7.1.3	Conditions initiales Dirac sur la fonction	147
7.2	Schéma semi-discrétisé en 2D	148
7.2.1	Solution exacte en Fourier	148
7.2.2	Approximation par phase stationnaire	149
7.2.3	Solution sur l'axe des x	150
7.2.4	Solution sur la diagonale $X = Y > 0$	154
7.3	Approximation équation équivalente	159
7.4	Conclusion	162
8	Schémas pour l'équation de la chaleur	165
8.1	L'équation de la chaleur	165
8.1.1	Schéma explicite	166
8.1.2	Schéma implicite	183
8.1.3	Schéma de Crank-Nicolson	186
8.1.4	Conclusion	189
8.2	Schémas discrétisant l'équation de la chaleur en 2D	191
8.2.1	Schéma explicite sur maillage régulier	191
9	Conclusion générale	193
III	Appendice	195
10	Schémas d'ordre élevé	197
10.1	Schémas de Strang	197
10.1.1	Génération des schémas de Strang pour l'équation d'advection à vitesse a	197
10.1.2	Analyse des schémas de Strang	199

10.2	Montée en ordre en espace	203
10.2.1	Montée en ordre par corrections successives : schémas centrés	203
10.2.2	Schémas sur grilles décalées	204
10.2.3	Conclusion	205
11	Quelques fonctions spéciales	207
11.1	La fonction d'Airy et ses généralisations	207
11.1.1	La fonction d'Airy	207
11.1.2	Fonctions d'Airy généralisées	215
11.2	Fonctions de Bessel	221
12	Développements asymptotiques d'intégrales	223
12.1	Méthode de Laplace	224
12.2	Méthode de la phase stationnaire	227
12.3	Méthode du col	230
	Table des figures	235
	Index	239

Première partie

**Introduction et méthodes
d'analyse**



Problèmes abordés

Les équations aux dérivées partielles (EDP) permettent de prédire l'évolution au cours du temps t d'une quantité physique u , par exemple un champ de pression, ou de température, sur un domaine de l'espace, connaissant cette quantité à un instant initial. Pour calculer leurs solutions, il est souvent nécessaire de recourir à la simulation numérique sur ordinateur. Comme même les ordinateurs les plus puissants ne peuvent traiter qu'un ensemble certes grand, mais fini, de nombres, il faut ramener l'EDP à un problème discret : les valeurs de sa solution vont être recherchées en un ensemble discret de points du domaine, et à des valeurs discrètes du temps. On conçoit que cette opération ne puisse donner qu'une solution approximative. Le but de l'analyse numérique est de concevoir des méthodes de discrétisation, appelées schémas numériques, et de s'assurer qu'ils fournissent une solution se rapprochant de la solution de l'EDP, quand la discrétisation est assez fine en espace et temps. Par exemple, dans le cas d'un domaine monodimensionnel repéré par la coordonnée x , si l'ensemble de points a pour coordonnées les multiples jh d'un pas d'espace, et l'ensemble d'instantes les multiples $n\tau$ d'un pas de temps τ , où j et n sont des entiers, il faut s'assurer que les valeurs $u_s(jh, n\tau)$ de la solution du schéma tendent (*i.e.* convergent) en un certain sens vers les valeurs $u(x = jh, t = n\tau)$ de la solution de l'EDP, quand h et τ tendent vers zéro. Nous poursuivons l'objectif un peu différent d'analyser quantitativement les schémas, *i.e.* de prédire leurs résultats. Nous cherchons plus spécifiquement à obtenir des formules (autant que possible) explicites, exactes ou approchées, de la solution fournie par les schémas étudiés. De telles formules permettent en effet de comprendre simplement et précisément comment se comporte un schéma, en particulier comment il converge vers la solution de l'équation de départ, et quels artéfacts numériques il est susceptible d'introduire.

Nous considérons des problèmes dits de Cauchy, *i.e.* où u est calculé à partir de conditions dites initiales, pour des EDP linéaires hyperboliques (comme l'équation

des ondes : u sera un champ de pression) ou paraboliques (comme l'équation de la chaleur : u sera un champ de température). En somme, le problème continu est de calculer la solution $u(x, t)$, où x est la variable d'espace et t le temps, connaissant $u(x, 0)$. Nous analysons principalement des problèmes monodimensionnels, posés sur tout \mathbb{R} , ou sur un intervalle, mais, dans ce dernier cas, pour des conditions aux limites périodiques uniquement. Nous traitons aussi quelques problèmes bidimensionnels, posés sur \mathbb{R}^2 , de calcul de $u(x, y, t)$ à partir de $u(x, y, 0)$.

Nous prenons principalement comme condition initiale, en 1D, la distribution de Dirac $\delta(x)$, soit

$$u(x, 0) = \delta(x)$$

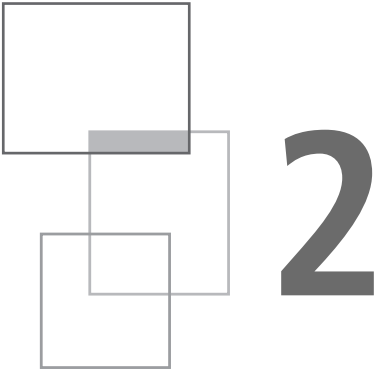
Dans ce cas, la solution exacte est la fonction de Green $G_e(x, t)$ de l'EDP. Nous définissons la fonction de Green d'un schéma, comme la solution $G_s(jh, n\tau)$ obtenue avec ce schéma du problème discret avec comme condition initiale $G_s(jh, 0) = \delta_0^j$, ce qui correspond à une condition initiale $h\delta(x)$ pour l'EDP, puisque la maille est de longueur h . Pour un schéma idéal $G_s(jh, n\tau) = hG_e(jh, n\tau)$ (nous espérons que cette différence de normalisation, qui allège les formules, ne perturbera pas le lecteur...). Mais tout schéma introduit des artéfacts numériques, non présents dans $G_e(x, t)$, comme un étalement, ou bien des oscillations. Ce problème a un intérêt intrinsèque, mais aussi plus général. En effet, pour les EDP linéaires, la solution $u_s(jh, n\tau)$ du schéma pour une condition initiale $u(jh, 0)$ s'obtient par convolution avec $G_s(jh, n\tau)$, qui caractérise en quelque sorte le schéma.

Nous prenons aussi parfois comme condition initiale la marche montante $H(x)$ ou descendante $H(-x)$, où H est la fonction de Heaviside. Dans ce cas aussi, les schémas introduisent des artéfacts. Nous analyserons aussi l'ordre de convergence (la notion d'ordre d'un schéma est définie plus loin) pour cette condition initiale. Il est le plus souvent bien inférieur à l'ordre nominal du schéma.

Ce n'est en effet que pour des conditions initiales assez régulières (*i.e.* avec un certain nombre de dérivées continues) que l'on dispose de théorèmes de convergence d'application simple, comme le théorème de Lax, pour les équations aux dérivées partielles linéaires. Ce sujet est bien documenté dans la littérature, et nous le développons moins en détail, nous bornant à rappeler les résultats essentiels.

Pour valider et illustrer les méthodes employées, les prédictions théoriques seront comparées à des simulations numériques pour un ensemble de schémas représentatifs, de différents ordres. Mais commençons par préciser quelles EDP nous allons considérer.

Nous avons regroupé dans l'Appendice, qui constitue la dernière partie de l'ouvrage, un ensemble de résultats auxquels nous nous référons dans le corps du texte.



Les équations traitées

2.1 Équations hyperboliques

Nous renvoyons à la littérature [9, 14] pour une définition précise de l'hyperbolicité. En termes simples, pour une équation hyperbolique, le problème aux conditions initiales est bien posé : il n'y a pas amplification non bornée de la condition initiale $\exp(ikx)$ en temps fini pour $k \rightarrow \infty$.

2.1.1 Équation d'advection

L'EDP hyperbolique linéaire modèle est l'équation d'advection 1D à vitesse constante $a \in \mathbb{R}^+$:

$$\partial_t u + a \partial_x u = 0, u(x, 0) = f(x) \quad (2.1)$$

Cette équation a une solution exacte explicite

$$u(x, t) = u(x - at, 0) = f(x - at) \quad (2.2)$$

La condition initiale est, selon (2.2), simplement transportée sans déformation. Cela permet des comparaisons aisées de la solution exacte de l'EDP aux résultats des schémas numériques.

Un exemple d'équation non hyperbolique est obtenu en modifiant de façon (apparemment) mineure l'équation (2.1) : $u_t + iau_x = 0$. Considérons en effet

la condition initiale $u(x, 0) = \exp(ikx)$. La solution est de la forme $u(x, t) = A(t) \exp(ikx)$ et $A(t)$ vérifie $A'(t) - kaA(t) = 0$, $A(0) = 1$ soit $A(t) = \exp(kat)$, qui n'est pas borné quand le nombre d'onde $k \rightarrow \infty$.

2.1.2 Équation des ondes

L'équation des ondes sonores est une autre équation hyperbolique. Elle s'écrit, en 1D, pour une vitesse du son égale à a

$$u_{tt} - a^2 u_{xx} = 0 \tag{2.3}$$

Comme elle est d'ordre 2, il faut préciser les conditions initiales sur la fonction $u(x, 0) = u^0(x)$ et sa dérivée $u_t(x, 0) = v^0(x)$. Comme l'équation d'advection (2.1), l'équation des ondes (2.3) a une solution explicite, qui s'écrit

$$u(x, t) = (u^0(x - at) + u^0(x + at))/2 + \int_{x-at}^{x+at} v^0(y) dy / 2c$$

En 1D, l'équation des ondes comme l'équation d'advection transporte les conditions initiales sans les déformer. L'équation des ondes génère, pour $u_t(x, 0) = 0$, $u(x, 0) = u_0(x)$ deux signaux se propageant l'un vers la droite, l'autre vers la gauche, alors que l'équation d'advection propage le signal dans une seule direction.

En dimension supérieure, la situation est plus complexe : la forme du signal évolue lors de sa propagation. Nous reviendrons sur les solutions de l'équation des ondes, en particulier sur les fonctions de Green de cette équation, dans la partie consacrée aux applications.

2.1.3 Équations non linéaires

Un exemple d'EDP hyperbolique non linéaire est l'équation de Burgers sans viscosité :

$$\partial_t u + u \partial_x u = 0$$

Elle peut générer des solutions discontinues (comme des chocs [10]), même à partir de conditions initiales régulières. En effet, la vitesse de propagation est maintenant u , si bien qu'une rampe descendante $u(x, 0) = 1, x < 0, u(x, 0) = 1 - x, 0 < x < 1, u(x, 0) = 0, x > 1$ va se transformer en marche descendante.

Nous ne traitons pas directement dans ce texte, centré sur les schémas numériques pour les équations aux dérivées partielles linéaires, de l'équation de Burgers. Toutefois, les résultats obtenus sur les conditions initiales de type marche pour l'équation d'advection illustrent assez bien le comportement des schémas au voisinage des chocs

pour cette équation non linéaire [21]. De plus, une méthode possible d'étude des équations non linéaires est de les linéariser [2].

2.2 Équations paraboliques

L'équation parabolique [9] modèle est l'équation de la chaleur (ou de la diffusion) qui, en 1D, s'écrit, avec $\alpha \in \mathbb{R}^+$ coefficient de diffusion

$$\partial_t u - \alpha \partial_{xx}^2 u = 0 \quad (2.4)$$

Contrairement aux équations précédentes, l'équation (2.4) change la forme des conditions initiales, en amortissant les composantes haute fréquence, ce qui la rend régularisante. On peut le voir en appliquant la transformation de Fourier. L'équation (2.4) s'écrit, dans l'espace réciproque

$$\widehat{u}_t + \nu k^2 \widehat{u} = 0$$

et sa solution est $\widehat{u}(k, t) = \widehat{u}(k, 0) \exp(-\alpha k^2 t)$, soit, pour la condition initiale $u(x, 0) = \exp(ikx)$, $u(x, t) = \exp(ikx) \exp(-\alpha k^2 t)$. L'amortissement (au sens de décroissance au cours du temps) est d'autant plus rapide que le nombre d'onde k est grand.

2.3 Fonctions de Green

Le terme fonction de Green est utilisé dans ce texte, par abus de langage, pour désigner la solution de l'équation sans second membre, et avec comme conditions initiales $u(x, 0) = \delta(x)$. Dans la littérature, il est plutôt défini comme la solution de l'équation avec conditions initiales nulles et second membre $\delta(x)\delta(t)$.

Par exemple, pour l'équation d'advection en 1D, la fonction de Green selon notre définition s'écrit

$$G(x, t) = \delta(x - at)$$

Pour l'équation des ondes en 1D, qui est du second ordre, il faut spécifier une condition initiale pour la fonction, une pour sa dérivée en temps. Si on choisit de définir la fonction de Green avec comme conditions initiales la distribution de Dirac sur la fonction, et 0 sur la dérivée, soit $G(x, 0) = \delta(x)$, $G_t(x, 0) = 0$,

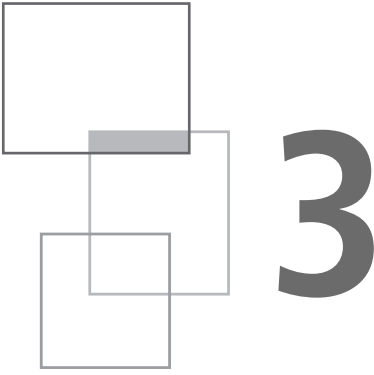
$$G(x, t) = \frac{1}{2}(\delta(x - at) + \delta(x + at))$$

Pour l'équation de la diffusion en 1D, G s'écrit

$$G(x, t) = \frac{1}{2\sqrt{\pi\alpha t}} \exp\left(-\frac{x^2}{4\alpha t}\right)$$

La fonction de Green de (2.4) est une gaussienne (donc une fonction analytique) qui s'élargit avec le temps. On voit à nouveau que pour l'équation d'advection et l'équation des ondes, la non-régularité des conditions initiales persiste dans le temps, alors qu'elle disparaît pour l'équation de la diffusion.

Les méthodes usuelles pour résoudre les EDP sont les différences, éléments, et volumes finis [26]. Nous nous focalisons sur la méthode des différences finies, très utilisée notamment pour les EDP présentées ci-dessus.



Méthode des différences finies

En 1D, un schéma aux différences finies [25] recherche une approximation, notée u_j^n , des valeurs $u(jh, n\tau)$ de la solution d'une EDP aux points d'une grille régulière en espace (pas h) et en temps (pas τ). Il se traduit par des relations entre ces quantités u_j^n caractérisées par deux indices : un indice temporel n et un indice spatial j . Ces relations sont appelées équations aux différences partielles. Si la résolution d'équations aux différences, *i.e.* à un seul indice, est le sujet de plusieurs ouvrages complets et didactiques, il n'en va pas de même des équations aux différences partielles (EDiP) : le seul ouvrage entièrement consacré à ce sujet est à notre connaissance [5]. Le chapitre sur les méthodes d'analyse des schémas présente les méthodes de résolution, exactes ou approchées, des EDiP dans un contexte différences finies. Mais, sur maillage régulier, les éléments et volumes finis se ramènent aussi à ces équations, ce qui leur donne une généralité supplémentaire. Dans le cas des volumes finis, u_j^n sera une approximation de la valeur moyenne de $u(x, t)$ sur la maille de centre $x = jh$ et le pas de temps $t = n\tau$. Les EDiP s'appliquent aussi à des problèmes multidimensionnels, il y aura autant d'indices spatiaux que de dimensions d'espace. Par exemple, en 2D, nous notons j et m ces indices et u_{jm}^n les quantités recherchées.

Avant de traiter des EDiP, nous présentons brièvement les différences finies.

3.1 Principe de la méthode

Commençons par le cas simple d'une équation différentielle ordinaire (EDO) du premier ordre, relation entre u et sa dérivée u' . La méthode des différences finies

s'applique aussi. Elle calcule une approximation u_j des valeurs $u(jh)$ de la solution u de l'EDO sur une grille régulière de pas h , en approximant u' par un quotient différentiel. Par exemple, la dérivée $u'(jh)$ peut être approchée sur une grille de pas h par

$$u'(jh) \simeq \frac{u_j - u_{j-1}}{h}$$

Pour les EDP en 1D, nous notons n l'indice du pas de temps, et, comme dit plus haut, la solution du schéma est notée u_j^n . Nous prenons toujours un pas de constant noté τ , et u_j^n est l'approximation de $u(x, t)$ en $x = jh$, $t = n\tau$. Le quotient différentiel (3.1) approche alors la dérivée partielle en x

$$u_x(jh, n\tau) \simeq \frac{u_j^n - u_{j-1}^n}{h} \quad (3.1)$$

Dans les applications, l'EDP est souvent posée sur un domaine borné, soit, en 1D, sur un intervalle, j est alors un entier naturel, compris entre 0 et $J - 1$, et une (ou des) conditions aux limites est (sont) imposée(s) en 0 (et/ou en $J - 1$). Mais, pour les études de schémas, nous considérons soit un problème posé sur tout \mathbb{R} , donc une grille (virtuelle) infinie, et un indice spatial j dans \mathbb{Z} , soit un problème posé sur un intervalle, mais avec une condition aux limites périodiques ($u(J) = u(0)$). Cela permet, comme expliqué dans le chapitre « Méthodes d'analyse des schémas », d'obtenir des solutions explicites des EDiP constitutives du schéma, par analyse de Fourier.

Dans le cas multidimensionnel, il y a plusieurs indices spatiaux. Par exemple en 2D, ces indices sont j et m , et la solution du schéma est notée u_{jm}^n , approximation de $u(x, y, t)$ en $x = jh$, $y = mh$, $t = n\tau$. Les dérivées partielles sont aussi approchées par des quotients différentiels, par exemple

$$(u_x)_{jm}^n \simeq \frac{u_{jm}^n - u_{(j-1)m}^n}{h}, \quad (u_y)_{jm}^n \simeq \frac{u_{jm}^n - u_{j(m-1)}^n}{h}$$

L'opérateur essentiel, à partir duquel sont construits les quotients différentiels, donc les schémas, est l'opérateur de décalage, ou shift en espace S . Il associe à la suite $\{u_j^n\}$ la suite $\{u_{j+1}^n\}$, et (3.1) se réécrit

$$u_x(jh, n\tau) \simeq \frac{(I - S)u_j^n}{h}$$

En multidimensionnel, il y a un opérateur de décalage spatial par dimension. De même que l'espace, le temps est discret, et nous notons T l'opérateur de décalage en temps, associant, à la suite $\{u_j^n\}$ la suite $\{u_j^{n+1}\}$.

3.2 Ordre d'une approximation différences finies

L'ordre d'une approximation différences finies mesure la précision de cette approximation : si l'erreur d'approximation est $O(h^p)$, elle est dite d'ordre p . La norme utilisée dans ce livre est essentiellement la norme du maximum : $\|u_j\| = \sup_j |u_j|$.

L'évaluation de l'ordre d'une approximation différences finies de u' dans une EDO (ou de u_x dans une EDP) repose sur des développements de Taylor. Plus précisément, la fonction u est injectée dans le quotient différentiel et on développe en supposant que cette fonction est assez régulière (*i.e.* que les dérivées à l'ordre nécessaire existent). Par exemple, si on utilise l'approximation (3.1), comme, par développement de Taylor, au point $x = (j - 1)h$

$$u((j - 1)h) = u(jh) - hu'(jh) + \frac{h^2}{2}u''(jh) + O(h^3)$$

on obtient

$$\frac{u(jh) - u((j - 1)h)}{h} = u'_j - \frac{h}{2}u''(jh) + O(h^2)$$

Le terme dominant de l'erreur d'approximation, $-\frac{h}{2}u''(jh)$, est proportionnel à h , donc d'ordre 1. On dit que l'approximation est consistante à l'ordre 1. Si on choisit maintenant comme approximation de u'_j

$$u'_j \simeq \frac{u_{j+1} - u_{j-1}}{2h} \quad (3.2)$$

comme la combinaison des développements de $u(x \pm h)$ fournit

$$\frac{u((j + 1)h) - u((j - 1)h)}{2h} = u'(jh) + O(h^2)$$

et l'approximation (3.2) est consistante à l'ordre 2. Ces écritures supposent la régularité de u , *i.e.* l'existence de dérivées bornées jusqu'à l'ordre du développement de Taylor, soit 1 pour (3.1) et 2 pour (3.2).

On note Q_p une approximation d'ordre p de l'opérateur de différentiation. Nous avons vu l'approximation (3.1) d'ordre 1, dite décentrée amont, souvent notée D_- :

$$D_-(u_j) \equiv \frac{u_j - u_{j-1}}{h}$$

et l'approximation centrée (3.2), d'ordre 2, notée D_0 :

$$D_0(u_j) \equiv \frac{u_{j+1} - u_{j-1}}{2h}$$

Il est possible d'obtenir des approximations d'ordre plus élevé. En appendice sont présentées celles utilisées dans ce livre : schémas centrés ou décalés et schémas dits de Strang.

3.3 Schémas pour l'équation d'advection

Nous allons maintenant appliquer ces approximations aux différences finies aux opérateurs de dérivation en espace et en temps pour construire des schémas pour l'équation (2.1) d'advection $1D u_t + au_x = 0$, nous en profiterons pour rappeler les notions de consistance et de stabilité d'un schéma.

3.3.1 Ordre de consistance d'un schéma

Comme vu précédemment, l'ordre p d'une discrétisation d'un opérateur différentiel est l'ordre de la différence entre l'opérateur et son approximation appliquée à une fonction régulière. Un schéma pour l'équation d'advection peut être obtenu en discrétisant séparément les dérivées partielles en espace à l'ordre p et en temps à l'ordre q . Comme précédemment, l'erreur de consistance des opérateurs de différentiation en espace et en temps est obtenue en appliquant le schéma à une fonction suffisamment régulière. L'erreur de consistance du schéma est la somme des erreurs de consistance sur chacun des opérateurs de différentiation. On obtient ainsi des schémas d'ordre p en espace et q en temps, avec une erreur $O(h^p)$ sur la discrétisation en espace et $O(\tau^q)$ sur la discrétisation en temps. Une autre manière de voir est d'appliquer le schéma à une solution exacte de l'équation d'advection (2.1) : le résultat est l'erreur de troncature. On retrouve bien la somme des erreurs de consistance sur chacun des opérateurs de différentiation, mais la méthode s'applique aussi à des schémas (par exemple, les schémas à un pas et 3 points étudiés ci-dessous), où les discrétisations en temps et en espace ne sont pas séparées.

À titre d'exemple, utilisons la discrétisation décentrée amont (3.1) pour la dérivée en espace, soit $(u_j^n - u_{j-1}^n)/h$, et la dérivée en temps, soit $(u_j^{n+1} - u_j^n)/\tau$, ce qui conduit, pour l'EDP (2.1), au schéma

$$\frac{u_j^{n+1} - u_j^n}{\tau} + a \frac{u_j^n - u_{j-1}^n}{h} = 0 \quad (3.3)$$

(3.3) s'écrit, avec le paramètre de Courant $\eta = a\tau/h$

$$u_j^{n+1} = (1 - \eta)u_j^n + \eta u_{j-1}^n \quad (3.4)$$

Le schéma (3.4), dit upwind ou décentré amont, est consistant à l'ordre 1 en espace et en temps : son erreur de troncature est $O(h) + O(\tau)$. Mais ce qui nous intéresse au final n'est pas l'erreur de troncature, mais l'erreur du schéma, définie comme la différence, dans une certaine norme, entre la solution du schéma et la solution de l'EDP. Un schéma est dit convergent si cette différence (*i.e.* l'erreur du schéma) tend vers 0 quand les pas de maillage (h en espace et τ en temps) tendent vers 0.

La relation entre erreur de troncature et erreur du schéma est exprimée par le théorème de Lax-Richtmyer.

3.3.2 Théorème de Lax-Richtmyer

Il s'énonce ainsi :

Théorème

Un schéma consistant à l'ordre p en espace et q en temps, pour une EDP linéaire, est convergent à ces mêmes ordres, pour des conditions initiales assez régulières, s'il est stable.

Les notions de consistance et de stabilité sont définies ci-après :

Consistance

Un schéma est dit consistant à l'ordre p en espace et q en temps si son erreur de troncature est en $O(h^p) + O(\tau^q)$.

Stabilité

Un schéma est dit stable s'il donne des résultats u_j^n bornés *i.e.* si, pour T donné, $\|u_j^n\| \leq D(T)$, où $\|u\|$ désigne une norme, non précisée ici, de la suite u_j^n , $j \in \{0, J\}$, $D(T)$ une fonction de T . Dans ce livre, nous utilisons essentiellement la norme du maximum, ou norme L^∞ .

Sous condition de stabilité, la différence entre la solution de l'EDP et la solution donnée par un schéma d'ordre p en espace, q en temps sera majorée par $C(h^p + \tau^q)$, si la condition initiale est assez régulière. Pour certains schémas, la stabilité implique une relation de la forme $\eta = a\tau/h < \eta_{\text{lim}}$. Il est alors cohérent de choisir $q = p$. L'erreur est alors $O(h^p)$, et le schéma dit « d'ordre p » sans plus de précision.

Le théorème de Lax-Richtmyer permet de ramener le problème de convergence de la solution du schéma vers celle de l'EDP à la vérification (relativement) facile :

- de la consistance : on injecte la solution de l'EDP dans le schéma, on fait les développements de Taylor pour calculer l'erreur de troncature et donc l'ordre du schéma ;
- de la stabilité : plusieurs méthodes sont possibles pour la vérifier, la plus simple (Von Neumann) pour les schémas linéaires consiste à s'assurer que le schéma n'amplifie pas les $\exp(ikx)$, k dans $[-\pi/h, \pi/h]$.

Les conditions de régularité seront précisées plus loin. Ces conditions de régularité ne sont pas satisfaites pour la fonction de Green G_s du schéma, définie plus haut comme la solution du schéma pour une condition initiale $u_j^0 = 1$ pour $j = 0$, $u_j^0 = 0$ pour $j \neq 0$. Rappelons qu'un des buts de ce livre est le calcul exact ou approché de cette fonction de Green. Elle permet notamment de comprendre quantitativement le comportement du schéma, pour cette condition initiale particulière, mais aussi plus généralement, car la solution pour une condition initiale générale s'obtient par convolution avec G_s .

3.3.3 Exemples de schémas

Dans cette section, pour illustrer les concepts précédents, quelques schémas types pour l'advection sont étudiés du point de vue de l'ordre et de la stabilité.

Schémas à un pas

Un schéma est dit à un pas si la solution u_j^{n+1} au temps $n+1$ s'exprime uniquement à l'aide des valeurs $u_{j'}^n$, $j' \in I$ au temps n . Le schéma est alors caractérisé par l'intervalle $I = \{j-r, j+s\}$ auquel appartient j' , et par les coefficients de pondération des $u_{j'}^n$. Le schéma général s'écrit

$$u_j^{n+1} = c_{-r} u_{j-r}^n + c_{-r+1} u_{j-r+1}^n + \dots + c_0 u_j^n + \dots + c_s u_{j+s}^n$$

Nous étudions dans cette section le cas $r = s = 1$, soit des schémas de la forme

$$u_j^{n+1} = \alpha u_{j-1}^n + \beta u_j^n + \gamma u_{j+1}^n \quad (3.5)$$

Pour vérifier la consistance et calculer l'erreur de troncature r_j^n du schéma défini par (3.5), il est appliqué à une solution u de l'équation d'advection

$$r_j^n = u(jh, (n+1)\tau) - \alpha u((j-1)h, n\tau) - \beta u(jh, n\tau) - \gamma u((j+1)h, n\tau) \quad (3.6)$$

En développant (3.6) à l'ordre 1, on obtient

$$r_j^n = (1 - \alpha - \beta - \gamma)u(jh, n\tau) + \tau u_t(jh, n\tau) + (\alpha - \gamma)h u_x(jh, n\tau) + o(h) + o(\tau)$$

puis, en utilisant l'équation d'advection vérifiée par u , l'expression (3.7) de r_j^n

$$r_j^n = (1 - \alpha - \beta - \gamma)u(jh, n\tau) + ((\alpha - \gamma)h - a\tau)u_x(jh, n\tau) + o(h) + o(\tau) \quad (3.7)$$

La consistance impose que r_j^n soit $o(h) + o(\tau)$, d'où les conditions

$$\alpha + \beta + \gamma = 1, \quad \alpha - \gamma = a\tau/h = \eta$$

Ce type de schéma dépend donc de deux paramètres, par exemple ν et le paramètre de Courant η

$$\alpha = \frac{\nu + \eta}{2}, \quad \gamma = \frac{\nu - \eta}{2}, \quad \beta = 1 - \nu$$

Le schéma peut donc se mettre sous la forme suivante, dite forme diffusive

$$u_j^{n+1} - u_j^n = \frac{\eta}{2}(u_{j-1}^n - u_{j+1}^n) + \frac{\nu}{2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n),$$

ou encore

$$\frac{u_j^{n+1} - u_j^n}{\tau} + a \frac{u_{j+1}^n - u_{j-1}^n}{2h} = \frac{v h^2}{2\tau} \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2}. \quad (3.8)$$

On reconnaît au premier membre de (3.8) les discrétisations :

- de la dérivée en temps : $(u_j^{n+1} - u_j^n) / \tau$
- de la dérivée en espace : $(u_{j+1}^n - u_{j-1}^n) / 2h$, multipliée par la vitesse a

et au second membre celle la dérivée seconde en espace : $(u_{j+1}^n - 2u_j^n + u_{j-1}^n) / h^2$, multipliée par un coefficient $d = v h^2 / (2\tau)$. Ce terme est homogène, si $v > 0$, à une diffusion. La dépendance en h^2 du coefficient d est une condition nécessaire de la consistance du schéma. Nous verrons dans la section 4.3 sur l'équation équivalente que le terme $(u_j^{n+1} - u_j^n) / \tau$ introduit *a contrario* un terme d'antidiffusion. Pour que le schéma soit stable, il faut que la diffusion globale du schéma, somme de ces deux termes, soit positive. Le schéma va introduire de la dissipation définie au sens suivant : les nombres d'onde présents dans la condition initiale sont dissipés par le schéma, *i.e.* leur amplitude décroît avec n . Une condition initiale de la forme $\exp(ikjh)$ est en effet multipliée à chaque pas de temps par un facteur (dit d'amplification) $g(k) \leq 1$ dont le module est d'autant plus petit que k est grand. Cette notion est précisée et illustrée dans la section « Visualisation dans l'espace réciproque, dispersion et dissipation » du chapitre 4. Comme illustré en particulier dans la section du chapitre 5 consacrée au schéma upwind, la diffusion génère un étalement au cours du temps de la fonction de Green du schéma, pour la condition initiale $u(x, 0) = h\delta(x)$, et du profil de la marche, pour la condition initiale $u(x, 0) = H(\pm x)$. Cet étalement est d'autant plus sensible que la diffusion du schéma, est forte, donc que v est grand.

Schéma upwind

Le schéma upwind est obtenu pour $v = \eta$. Il est stable sous condition CFL (pour Courant Friedrichs Lewy) $\eta \leq 1$. Il s'écrit

$$u_j^{n+1} - u_j^n = \frac{\eta}{2}(u_{j-1}^n - u_{j+1}^n) + \frac{\eta}{2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) = \eta(u_{j-1}^n - u_j^n)$$

Il se met, comme vu plus haut, sous la forme barycentrique (3.4)

$$u_j^{n+1} = \eta u_{j-1}^n + (1 - \eta) u_j^n$$

La valeur au pas de temps $n + 1$ est un barycentre de valeurs au pas de temps n avec poids $\eta, 1 - \eta$. Cette forme barycentrique permet, si $\eta \leq 1$, de démontrer

directement la stabilité en norme L^∞ :

$$\left| u_j^{n+1} \right| \leq \eta \left| u_{j-1}^n \right| + (1 - \eta) \left| u_j^n \right| \Rightarrow \| u^{n+1} \|_\infty \leq \| u^n \|_\infty$$

On peut aussi montrer la stabilité en norme L^p , pour $p > 1$

$$\left| u_j^{n+1} \right|^p \leq \eta \left| u_{j-1}^n \right|^p + (1 - \eta) \left| u_j^n \right|^p \quad (3.9)$$

Sommons (3.9) sur l'indice j

$$\sum_{j \in \mathbb{Z}} \left| u_j^{n+1} \right|^p \leq \eta \sum_{j \in \mathbb{Z}} \left| u_j^n \right|^p + (1 - \eta) \sum_{j \in \mathbb{Z}} \left| u_j^n \right|^p = \sum_{j \in \mathbb{Z}} \left| u_j^n \right|^p \Rightarrow \| u^{n+1} \|_p \leq \| u^n \|_p$$

La norme de la suite u_j décroît au cours du temps, le schéma est stable en norme L^p .

Schéma de Lax-Friedrichs

Ce schéma est obtenu pour $\nu = 1$, et s'écrit

$$u_j^{n+1} - u_j^n = \frac{\eta}{2}(u_{j-1}^n - u_{j+1}^n) + \frac{1}{2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n)$$

ou bien

$$u_j^{n+1} = \frac{\eta}{2}(u_{j-1}^n - u_{j+1}^n) + \frac{1}{2}(u_{j+1}^n + u_{j-1}^n)$$

Il est stable pour $\eta \leq 1$. Le coefficient ν du termes de diffusion est plus grand que celui du schéma upwind, sans gain en termes de stabilité.

Schéma de Richardson

Il est obtenu pour $\nu = 0$. Imaginé dans les années 1920 pour des simulations météorologiques, avant que les notions de stabilité n'aient été introduites par Courant, Friedrichs et Lewy dans leur article de 1928, ce schéma est instable, car son facteur d'amplification $g = 1 - i\eta \sin k$ est, en module, plus grand que 1 : $|g| = \sqrt{1 + (\eta \sin k)^2}$.

Schéma de Lax-Wendroff

Il est obtenu pour $\nu = \eta^2$ et s'écrit

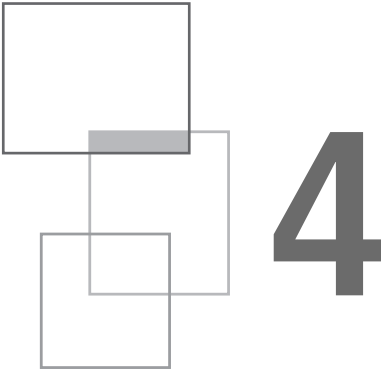
$$u_j^{n+1} - u_j^n = \frac{\eta}{2}(u_{j-1}^n - u_{j+1}^n) + \frac{\eta^2}{2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (3.10)$$

(3.10) se réécrit

$$u_j^{n+1} = \eta u_{j-1}^n + (1 - \eta) u_j^n - \frac{\eta(1 - \eta)}{2} (u_{j+1}^n - 2u_j^n + u_{j-1}^n)$$

Ce schéma peut donc être interprété de deux manières : ajout de diffusion au schéma de Richardson, ou soustraction de la diffusion du schéma upwind. Il est d'ordre 2. Son équation équivalente (notion définie dans le chapitre suivant) a un premier terme dit de dispersion : la vitesse de propagation d'une condition initiale $u_j^0 = \exp ikj$ dépend du nombre d'onde k . La notion de dispersion est précisée dans la section 4.4. Nous verrons dans la deuxième partie, consacrée aux applications, que la fonction de Green d'un schéma dispersif présente des oscillations.

Ces schémas (et d'autres) sont analysés plus en détail dans la deuxième partie. Nous présentons dans le chapitre suivant des méthodes d'analyse quantitative des schémas aux différences finies.



Méthodes d'analyse des schémas

D'un point de vue général, un schéma aux différences finies peut être défini, en 1D, comme une relation entre les opérateurs, présentés plus haut, de décalage en espace S (associant à la suite $\{u_j^n\}$ la suite $\{u_{j+1}^n\}$) et de décalage en temps T (associant à la suite $\{u_j^n\}$ la suite $\{u_j^{n+1}\}$). Ces opérateurs s'appliquent aussi en continu. Ils associent à $u(x, t)$ respectivement $Tu = u(x, t + \tau)$ et $Su = u(x + h, t)$, et s'écrivent $T = \exp \tau \partial_t$ et $S = \exp h \partial_x$ comme on peut le voir avec un développement de Taylor

$$Su(x) = u(x + h) = \sum_{p=0}^{\infty} \frac{h^p}{p!} (\partial_x)^p u = \exp h \partial_x u$$

Mais un schéma est conçu pour approcher une EDP, *i.e.* une relation entre les opérateurs différentiels ∂_t et ∂_x . On conçoit que l'introduction d'échelles de temps τ et d'espace h exogènes à l'EDP initiale ne peut pas être sans conséquences. Autrement dit, il est (en général) illusoire d'espérer représenter parfaitement l'EDP, qui relie ∂_t et ∂_x par un schéma, qui relie T et S . En particulier, les échelles spatiales de l'ordre du pas d'espace h ne seront pas correctement traitées.

Tout l'objet de l'analyse numérique est de comprendre en quel sens la solution du schéma approche celle de l'EDP. La plupart des ouvrages sont centrés sur la convergence de la solution du schéma vers celle de l'EDP pour des conditions initiales régulières. Nous cherchons plutôt à établir, comme annoncé dans l'introduction, des solutions explicites, exactes ou approchées de schémas, en nous intéressant

particulièrement aux conditions initiales δ_j^0 , *i.e.* valant 1 sur la maille $j = 0$: comme dit dans l'introduction, nous parlons, par abus de langage, de fonction de Green G_s du schéma. La solution du schéma pour des conditions initiales de hautes fréquences spatiales $\exp(ikj)$, avec k de l'ordre de l'unité diffère sensiblement de celle de l'EDP. La transformée de Fourier de δ_j^0 valant 1 sur tout le spectre, les hautes fréquences sont une composante non négligeable dans les conditions initiales pour G_s . C'est une des raisons des différences entre G_s et la fonction de Green G_e de l'équation.

Pour analyser le comportement d'un schéma, vu comme une relation entre $T = \exp \tau \partial_t$ et $S = \exp h \partial_x$, il existe au moins trois façons de procéder.

Soit on exprime directement la solution du schéma à partir de cette relation : nous appelons méthodes opératorielles ces techniques, qui, quand elles s'appliquent, donnent simplement et directement des formules totalement explicites.

Soit le schéma est analysé par transformation de Fourier dans l'espace réciproque (ou spectral), l'idée étant que, dans cet espace, T et S sont de simples multiplications. Cette méthode fournit la solution du schéma sous forme intégrale, donc moins explicite que les formules analytiques obtenues avec les méthodes opératorielles, mais est d'application plus générale. Elle peut amener aussi à des solutions explicites approchées si elles sont associées à des méthodes de développement asymptotiques d'intégrales, présentées en appendice.

Enfin, la relation entre T et S peut être approchée par une famille de relations entre les opérateurs différentiels ∂_t et ∂_x , dont la plus simple sera l'EDP de départ, *i.e.* sans second membre. Cette méthode commode et intuitive porte plusieurs noms : approximation différentielle, équation équivalente (sous-entendu au schéma), équation modifiée (sous-entendu par rapport à l'EDP de départ). Nous écrirons toujours l'équation équivalente sous la forme appelée Π dans [24], *i.e.* avec comme premier membre l'EDP de départ, et avec, au second membre, des opérateurs différentiels spatiaux. Nous verrons que les techniques opératorielles sont aussi utiles pour écrire l'équation équivalente.

Nous passons en revue ces trois méthodes, qui sont essentiellement complémentaires.

4.1 Méthodes opératorielles

Un schéma numérique s'écrit, en une dimension, à l'aide d'opérateurs agissant sur des suites u_j , si l'indice j est dans \mathbb{Z} (problème posé sur tout \mathbb{R}), dans \mathbb{N} (problème posé sur \mathbb{R}_+) ou sur des vecteurs (les opérateurs sont dans ce cas des matrices), si l'indice j est dans $I = \{0, \dots, j, \dots, J - 1\}$ (problème posé sur un intervalle I , pour des conditions aux limites périodiques).

Pour les problèmes dépendant du temps, rappelons que n est l'indice du pas de temps, et que la solution du schéma est notée u_j^n .

Dans le cas multidimensionnel, il y aura plusieurs indices spatiaux. En 2D, nous notons j et m les indices, u_{jm}^n la solution du schéma.

L'opérateur essentiel, à partir duquel sont construits tous les schémas, est l'opérateur de décalage positif d'un cran, ou shift en espace S , introduit plus haut. Il associe, à la suite $\{u_j\}$ la suite $\{u_{j+1}\}$, $j \in \mathbb{Z}$ ou au vecteur $\{u_j\}$ le vecteur $\{u_{j+1 \bmod J}\}$, $j \in I$.

Dans le cas périodique, $\{u_j\}$ est un vecteur colonne à J composantes, et S est une matrice (dite circulante) avec des 1 au-dessus de la diagonale, qui agit sur le vecteur des $\{u_j\}$. Par exemple, si $J = 4$, ce vecteur est

$$\begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{pmatrix} \text{ et la matrice } S \text{ s'écrit } S = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Nous notons S^m la puissance m de l'opérateur, qui applique à $\{u_j\}$ un décalage de m , *i.e.* lui associe $\{u_{j+m}\}$. L'entier m peut être positif ou négatif, par exemple

$$S^{-1} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Dans le cas multidimensionnel, il y a un opérateur shift par indice. Tous les opérateurs aux différences s'expriment à l'aide de l'opérateur S . Par exemple, le quotient différentiel décentré amont (3.1) $D_-(u_j) \equiv \frac{u_j - u_{j-1}}{h}$ s'écrit sous forme opérateur

$$D_- = \frac{I - S^{-1}}{h} \text{ et le quotient d'ordre 2 centré (3.2) } D_0(u_j) \equiv \frac{u_{j+1} - u_{j-1}}{2h} \text{ s'écrit}$$

$$D_0 = \frac{S - S^{-1}}{2h}$$

Donc tous les schémas, au moins à un pas, s'écrivent comme des combinaisons linéaires de puissances positives et négatives de cet opérateur, ce qui permet d'écrire leur solution sous forme totalement explicite. Nous donnons ci-dessous, à titre d'illustration, quelques exemples de solutions de schémas obtenues par méthode opératorielle.

Réécrivons par exemple le schéma upwind (3.4) en 1D $u_j^{n+1} = \eta u_{j-1}^n + (1 - \eta) u_j^n$ à l'aide de l'opérateur S . Puisque $u_{j-1}^n = S^{-1} u_j^n$,

$$u_j^{n+1} = \eta S^{-1} u_j^n + (1 - \eta) u_j^n = (\eta S^{-1} + (1 - \eta)I) u_j^n$$

Notant $(1 - \eta) = p$, $\eta = q$, en développant la puissance, on obtient

$$u_j^n = (pI + qS^{-1})^n u_j^0 = \sum_{l=0}^n p^{n-l} q^l \binom{n}{l} u_{j-l}^0 \quad (4.1)$$

La solution du schéma upwind en 1D s'exprime donc à l'aide de coefficients binomiaux. En particulier, puisque la condition initiale dans (4.1) est $u_j^0 = \delta_j^0$, le seul indice pour lequel $u_{j-l}^0 \neq 0$ est $l = j$, sa fonction de Green s'écrit

$$u_j^n = p^{n-j} q^j \binom{n}{j} \quad (4.2)$$

Elle s'étale donc sur les $n + 1$ points $\{0, \dots, j, \dots, n\}$. Tout schéma à un pas avec un « stencil » (défini comme l'ensemble des points intervenant dans le schéma) avec r points à gauche et s points à droite s'écrit sous la forme

$$\begin{aligned} u_j^n &= (a_{-r} S^{-r} + a_{-r+1} S^{-r+1} + \dots + a_0 I + \dots + a_s S^s) u_j^{n-1} \\ u_j^n &= (a_{-r} S^{-r} + a_{-r+1} S^{-r+1} + \dots + a_0 I + \dots + a_s S^s)^n u_j^0 \end{aligned} \quad (4.3)$$

La solution s'exprime maintenant, en développant dans (4.3) la puissance de la somme d'opérateurs, comme une somme pondérée de coefficients multinomiaux. Elle est donc moins simple que pour le schéma upwind.

Le cas périodique peut se traiter comme un problème d'algèbre linéaire, mais s'écrit formellement comme le cas sur \mathbb{R} . Plus précisément un schéma à un pas s'écrit, en notant u^n le vecteur des u_j^n

$$u^{n+1} = M u^n$$

M est une matrice circulante, combinaison linéaire de puissances de S . La solution du schéma upwind s'écrit par exemple

$$u^n = (pI + qS^{-1}) u^{n-1} = (pI + qS^{-1})^n u_j^0 = \sum_{l=0}^n p^{n-l} q^l \binom{n}{l} S^{-l} u_j^0 \quad (4.4)$$

(4.4) est formellement identique à (4.1), mais, dans le cas périodique, $S^J = I$. La solution est donc repliée sur son domaine de définition $\{0, \dots, j, \dots, J - 1\}$.

L'opérateur shift en temps T est parfois utilisé pour traiter des schémas à un pas, qui s'expriment comme une relation entre T et une combinaison de puissances de S . Le schéma upwind (3.4) s'écrit par exemple

$$T = pI + qS^{-1}$$

$$u_j^n = T u_j^{n-1} = T^n u_j^0 = (pI + qS^{-1})^n u_j^0$$

C'est une simple reformulation des expressions vues plus haut. Ce point de vue a plus d'intérêt quand il s'agit de calculer automatiquement l'équation équivalente d'un

schéma totalement discrétisé, comme expliqué dans la dernière section du chapitre. Le schéma est alors considéré comme une relation exacte entre $T = \exp(\tau \partial_t)$ et $S = \exp(h \partial_x)$. L'équation équivalente l'approche à l'aide d'une relation entre ∂_t et ∂_x .

Les méthodes opératorielles ont l'avantage de fournir des expressions explicites, mais celles-ci sont d'autant plus compliquées que le stencil du schéma comporte plus de points. Une autre classe de méthodes, basée sur l'analyse de Fourier, n'a pas cet inconvénient. Par contre, la solution du schéma est obtenue sous forme d'une intégrale, en domaine infini, ou d'une somme, pour des conditions périodiques, comme nous allons le voir dans la section suivante.

4.2 Analyse de Fourier des schémas

4.2.1 Les différentes formes de la transformation de Fourier

La transformée de Fourier est une application de l'espace direct, avec une variable continue (notée x en 1D) ou discrète (notée jh en 1D) de position vers un espace dit réciproque (ou spectral) avec une variable continue ou discrète notée k en 1D. Elle transforme un opérateur différentiel (cas continu) ou aux différences (cas discret) de l'espace direct, en une multiplication dans l'espace réciproque. Pour un schéma à coefficients constants, constitués d'opérateurs aux différences, la transformation de Fourier est un outil commode d'analyse. La solution du schéma est écrite dans l'espace réciproque. On revient ensuite par transformation de Fourier inverse dans l'espace réel.

Elle ne s'applique pas *stricto sensu* dans le cas où j est un entier naturel, compris entre 0 et $J - 1$, avec une (ou des) conditions aux limites imposée(s) en 0 (et/ou en $J - 1$).

Pour pouvoir utiliser la transformation de Fourier pour déterminer la solution d'un schéma, nous considérons donc

- soit une grille (virtuelle) infinie, *i.e.* j dans \mathbb{Z}
- soit une condition aux limites périodiques $u(J - 1) = u(0)$

La transformation de Fourier, pour nos applications, est donc déclinée sous deux formes :

- pour un maillage infini (j dans \mathbb{Z}), transformation de Fourier semi-discrète, associant à la suite u_j la fonction

$$\widehat{u}(k) = \sum_{j=-\infty}^{j=+\infty} u_j e^{-ikj}, \quad k \in [-\pi, \pi]$$

- pour le cas périodique, transformation de Fourier discrète, associant à la suite $u_j, j \in \{0, J - 1\}$, la suite

$$\widehat{u}_k = \sum_{j=0}^{J-1} u_j e^{-\frac{2i\pi}{J}kj}, k \in \{0, J - 1\}$$

La transformation de Fourier semi-discrète, comme sa variante discrète, permet d'obtenir, comme expliqué dans la section suivante, des représentations explicites des solutions des schémas sous forme intégrale. Une formulation un peu différente, dite transformée en z , est aussi utilisée : nous la présenterons pour être complet, et aussi car elle peut s'avérer commode.

La transformée de Fourier usuelle, associant à une fonction $u(x), x \in \mathbb{R}$, la fonction $\widehat{u}(k) = \int_{-\infty}^{+\infty} u(x) e^{-ikx} dx$ est parfois utilisée pour l'étude des schémas dans la littérature, mais il faut pour cela généraliser le schéma à une équation aux différences sur \mathbb{R} , comme nous l'avons fait plus haut pour étudier la stabilité des schémas à un pas et trois points. On introduit ainsi une imprécision. Nous utiliserons par contre cette transformation de Fourier usuelle pour calculer les solutions des équations « équivalentes » (en un sens à préciser), aux schémas, comme on le verra dans le chapitre 5. Rappelons enfin que la transformée de Fourier est une isométrie, ce que traduit la relation de Parseval.

4.2.2 Transformation de Fourier semi-discrète

TFSD en 1D

La transformation de Fourier semi-discrète est définie par

$$(u_j)_{j \in \mathbb{Z}} \in l^2(\mathbb{Z}) \rightarrow \widehat{u}(k) = \sum_{j=-\infty}^{j=+\infty} u_j e^{-ikj} \in L^2[-\pi, \pi] \quad (4.5)$$

Les espaces $l^2(\mathbb{Z})$ et $L^2[-\pi, \pi]$ sont définis par leur norme respective : $\|u_j\|^2 = \sum_{j=-\infty}^{j=+\infty} (u_j)^2, \|\widehat{u}(k)\|^2 = \int_{-\pi}^{\pi} |\widehat{u}(k)|^2 dk$. Il peut être commode dans certains cas (particulièrement quand intervient la variable $x = jb$) de faire le changement de variable $k \rightarrow kb, i.e.$ d'utiliser la forme

$$u_j \in l^2(\mathbb{Z}) \rightarrow \widehat{u}(k) = h \sum_{j=-\infty}^{j=+\infty} u_j e^{-ikjh} \in L^2 \left[-\frac{\pi}{h}, \frac{\pi}{h} \right] \quad (4.6)$$

La transformation inverse est donnée, avec la première définition, par

$$u_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{u}(k) e^{ikj} dk \quad (4.7)$$

et, avec la seconde, par

$$u_j = \frac{1}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \widehat{u}(k) e^{ikjh} dk \quad (4.8)$$

Dans la première définition (4.5), le nombre d'onde k et $\widehat{u}(k)$ sont adimensionnés. Dans la seconde (4.6), k est l'inverse d'une unité de longueur, par exemple le m^{-1} , $\widehat{u}(k)$ a la dimension d'une unité de longueur. La transformation conserve la norme L^2 : $\|\widehat{u}\|_2 = \sqrt{2\pi} \|u\|_2$ et le produit scalaire $\widehat{u} \cdot \widehat{v} = 2\pi u \cdot v$: ce sont les relations de Parseval. Notons qu'il est possible d'utiliser la transformation pour des suites qui ne sont pas dans $l^2(\mathbb{Z})$, *i.e.* pour lesquelles $\sum_j (u_j)^2$ est infini :

nous le ferons en particulier pour la condition initiale de type marche descendante. La transformation remplace l'opérateur S par une multiplication par e^{ik} . En effet

$$\widehat{Su}(k) = \sum_{j=-\infty}^{j=+\infty} u_{j+1} e^{-ikj} = \sum_{j'=-\infty}^{j'=+\infty} u_{j'} e^{-ik(j'-1)} = e^{ik} \sum_{j'=-\infty}^{j'=+\infty} u_{j'} e^{-ikj'} = e^{ik} \widehat{u}(k)$$

Donc, dans l'espace réciproque, tout schéma à un pas devient une multiplication par une somme de puissances de (*i.e.* un polynôme en) e^{ik} : cela permet d'obtenir une solution explicite de tout schéma à un pas dans le domaine spectral. L'application de la transformation inverse (4.7) fournit alors une représentation intégrale de la solution du schéma. Plus précisément, le schéma à un pas

$$u_j^n = (a_{-r} S^{-r} + a_{-r+1} S^{-r+1} + \dots + a_0 I + \dots + a_s S^s) u_j^{n-1}$$

est caractérisé par l'opérateur aux différences $Q_p = a_{-r} S^{-r} + a_{-r+1} S^{-r+1} + \dots + a_0 I + \dots + a_s S^s$, et s'écrit dans l'espace réciproque

$$\widehat{u}^n(k) = (a_{-r} e^{-ikr} + a_{-r+1} e^{-ik(r-1)} + \dots + a_0 + \dots + a_s e^{iks}) \widehat{u}^{n-1}(k)$$

soit, en notant \widehat{Q}_p le facteur d'amplification

$$\widehat{Q}_p(k) \equiv a_{-r}e^{-ikr} + a_{-r+1}e^{-ik(r-1)} + \dots + a_0 + \dots + a_s e^{iks}$$

$$\widehat{u}^n(k) = \widehat{Q}_p(k) \widehat{u}^{n-1}(k) = \widehat{Q}_p(k)^n \widehat{u}^0(k) \quad (4.9)$$

La solution dans l'espace direct s'obtient sous forme intégrale en appliquant à (4.9) la transformation de Fourier inverse

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{Q}_p(k)^n \widehat{u}^0(k) e^{ikj} dk \quad (4.10)$$

Le facteur d'amplification $\widehat{Q}_p(k)$ caractérise complètement le schéma, puisqu'il permet d'en calculer la solution selon la formule (4.10). Il est aussi possible d'obtenir par analyse de Fourier des solutions explicites pour des schémas à plusieurs pas comme on le verra dans les exemples, notamment dans la section 3 du chapitre 5 sur le schéma saute-mouton. Le principe reste le même, mais la forme de la solution est plus complexe que pour les schémas à un pas.

Transformée en z bilatérale

La transformée en z est définie, sous forme dite bilatérale (car la sommation dans (4.11) est sur \mathbb{Z}), par

$$u_j \in l^2(\mathbb{Z}) \rightarrow \widehat{u}(z) = \sum_{j=-\infty}^{j=+\infty} u_j z^{-j} \quad (4.11)$$

La série au second membre de (4.11) est une série formelle. Transformée en z bilatérale et transformation de Fourier semi-discrète sont strictement équivalentes : le changement de variable $z = e^{ik}$ dans la transformée en z (4.11) donne

$$\widehat{u}(e^{ik}) = \sum_{j=-\infty}^{j=+\infty} u_j e^{-ikj}, \text{ soit la transformation de Fourier semi-discrète (4.5). Son}$$

inverse s'écrit comme une intégrale en z

$$u_j = \frac{1}{2\pi i} \int_C \widehat{u}(z) z^{j-1} dz$$

C est le cercle de centre l'origine et de rayon 1, mais peut être remplacé par un contour, qui doit inclure l'origine. En effet, en substituant $\widehat{u}(z) = \sum_{j=-\infty}^{j=+\infty} u_j z^{-j}$ dans l'intégrale ci-dessus, on obtient par la méthode des résidus

$$\frac{1}{2\pi i} \int_C \sum_{l=-\infty}^{l=+\infty} u_l z^{j-l-1} dz = u_j$$

La contribution de résidu en $z = 0$ vient du terme $l = j$. En particulier, la transformée en z remplace l'opérateur S par une multiplication par z , et donc tout stencil spatial par une multiplication par une fraction rationnelle

$$\widehat{Su}(z) = \sum_{j=-\infty}^{j=+\infty} u_{j+1} z^{-j} = \sum_{j'=-\infty}^{j'=+\infty} u_{j'} z^{-(j'-1)} = z \widehat{u}(z)$$

On peut donc en 1D appliquer l'une ou l'autre méthode. La transformée en z aboutit à des notations plus compactes. La liberté sur le contour (qu'on peut aussi utiliser pour la transformation de Fourier) permet d'obtenir des approximations par des méthodes de développement asymptotiques d'intégrales, en particulier la méthode du col, présentées en détail en appendice. Par contre, la généralisation au 2D est possible [32,33], mais moins commode que pour la transformation de Fourier. Nous utiliserons la transformée en z seulement en 1D.

Transformée en z unilatérale

Sous forme unilatérale, la transformée en z s'écrit

$$u_j \in l^2(\mathbb{N}) \rightarrow \widehat{u}(z) = \sum_{j=0}^{j=+\infty} u_j z^{-j}$$

Elle permet, appliquée sur l'indice j , de traiter des problèmes sur des demi-droites. Elle peut aussi être utilisée sur la partie différence temporelle d'un schéma, l'indice de sommation est alors l'indice n

$$u_j^n \in l^2(\mathbb{N}) \rightarrow \widehat{u}_j(z) = \sum_{n=0}^{n=+\infty} u_j^n z^{-n}$$

Il est ainsi possible de traiter des problèmes avec maillage de pas non constant, que nous n'avons pas inclus dans cet ouvrage. Elle peut être vue comme une version discrète de la transformée de Laplace [32,33]. L'analyse de Fourier permet donc d'obtenir des représentations intégrales de la solution d'un schéma à un pas appliqué à une EDP linéaire. Elle s'applique aussi aux schémas semi-discrétisés en espace.

Représentation intégrale de la solution d'un schéma semi-discrétisé

Considérons, par exemple, la solution du schéma discrétisé seulement en espace $u_t = Q_p u$, $u(0) = u^0$, où Q_p est un opérateur aux différences d'ordre p , u est la suite des $u_j(t)$. La transformation de Fourier semi-discrète transforme l'opérateur aux différences Q_p en une multiplication par sa transformée \widehat{Q}_p . Si par exemple, Q_p est l'opérateur décentré amont $Q_1 = aD_-$, $Q_1(u_j) = a(u_j - u_{j-1})/h$, $\widehat{Q}_p = a(1 - e^{-ikh})/h$

En appliquant la TFSD au schéma et à la condition initiale, on se ramène ainsi à une équation différentielle $\widehat{u}_t = \widehat{Q}_p \widehat{u}$, $\widehat{u}(0) = \widehat{u}^0$ dont la solution est, dans l'espace réciproque, donnée par (4.12)

$$\widehat{u}(k, t) = \widehat{u}^0 \exp(\widehat{Q}_p t) \tag{4.12}$$

donc, revenant à l'espace réel en appliquant à (4.12) la transformation de Fourier inverse, on obtient la solution du schéma discrétisé en espace sous forme intégrale

$$u_j(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{u}^0(k) \exp(\widehat{Q}_p t + ijk) dk \tag{4.13}$$

$\widehat{Q}_p(k)$ est le facteur d'amplification du schéma, pour une condition initiale $\exp(ik\frac{x}{h}) = \exp(ijk)$. Il est homogène à une pulsation : posons $\widehat{Q}_p = -i\omega$, où ω est réel si le schéma est sans dissipation. Passons en variable continue $x = jh$. Ce passage n'est pas trivial, il y a plusieurs façons de procéder [25], que nous ne détaillerons pas. Faisons dans (4.13) le changement de variable $k \rightarrow kh$

$$u_j(t) = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} \widehat{u}^0(k) \exp i(kx - \omega(k)t) dk \tag{4.14}$$

si bien que la solution (4.14) du schéma s'interprète comme une superposition d'ondes planes de vitesse de phase $\frac{\omega(k)}{k}$. Cette vitesse, qui vaut a pour l'équation, dépend de k : le schéma introduit de la dispersion.

Comparaison à la solution exacte de l'équation

Comparons les solutions du schéma à la solution exacte, dans le cas de l'équation d'advection, d'abord en version semi-discrétisée. La solution exacte de l'équation d'advection (2.1) $u_t + au_x = 0$ dans l'espace de Fourier s'écrit, puisque $\widehat{u}_t + ik a \widehat{u} = 0$, $\widehat{u} = \widehat{u}^0 \exp(-ikat)$, comme une superposition d'ondes planes $\exp ik(x - at)$

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widehat{u}^0(k) \exp ik(x - at) dk$$

En appliquant la relation de Parseval pour les produits scalaires, on obtient, puisque la transformée de Fourier inverse de $\exp ik(x - at)$ est $\delta(y - (x - at))$

$$u(x, t) = \int_{-\infty}^{+\infty} u^0(y) \delta(y - (x - at)) dy = u^0(x - at)$$

Le facteur d'amplification exact est $\exp(-ikat)$, celui du schéma est $\exp(\widehat{Q}_p t)$. Il faut donc comparer \widehat{Q}_p à $-ika$. Par exemple, pour une discrétisation amont de la dérivée spatiale

$$\widehat{Q}_p = -a\widehat{D}_- = a(e^{-ikh} - 1)/h = -ika + O(h)$$

L'ordre 1 de l'approximation se traduit par une tangente à l'origine de $-a\widehat{D}_-$ à $-ika$. La différence ($\exp(-ikat) - \exp(\widehat{Q}_p t)$) au voisinage de $k = 0$ est du même ordre que la différence

$$-ikat - \widehat{Q}_p t \approx -ikat - (-ika - k^2 ah/2)t = k^2 hat/2$$

En version totalement discrétisée, le facteur d'amplification exact sur un pas de temps est $\exp(-ikat) = \exp(-ikh \frac{a\tau}{h}) = \exp(-ik\eta h)$, à comparer à \widehat{Q}_p . Par exemple, pour le schéma upwind

$$u_j^{n+1} = \eta u_{j-1}^n + (1 - \eta)u_j^n, \quad \widehat{Q}_p = \eta e^{-ik} + 1 - \eta$$

Revenons à k adimensionné, et développons le facteur d'amplification exact $\exp(-ik\eta)$

$$\exp(-ik\eta) = 1 - i\eta k - \eta^2 k^2/2 + o(k^2)$$

et celui du schéma

$$\widehat{Q}_p = \eta e^{-ik} + 1 - \eta = 1 - i\eta k - \eta k^2/2 + o(k^2)$$

d'où la différence

$$\exp(-ik\eta) - \widehat{Q}_p = \frac{1}{2}k^2\eta(\eta - 1) + o(k^2)$$

Le facteur d'amplification du schéma upwind approxime à l'ordre 1 le facteur d'amplification exact quand $k \rightarrow 0$. La différence des facteurs d'amplification au voisinage de $k = 0$ est d'ordre de $\eta k^2 = a\tau k^2 h$. Pour le schéma upwind, le terme principal de cette différence est $\eta(\eta - 1)k^2/2$. La différence des facteurs d'amplification à la puissance n est d'ordre $an\tau k^2/h = k^2 n\eta$.

La représentation intégrale de la solution d'un schéma, en version semi (4.10) ou totalement (4.13) discrétisée, est très utile pour en analyser les propriétés de convergence. En effet, la différence entre solution exacte et solution du schéma s'exprime

explicitement sous forme intégrale, et la démonstration de la convergence revient à une majoration de cette intégrale. Cette majoration s'appuie sur la convergence du facteur d'amplification du schéma vers celui de l'équation pour $k \rightarrow 0$ et sur la régularité des conditions initiales, qui se traduit par une décroissance de leur spectre pour $k \rightarrow \infty$. Ces points sont précisés plus loin (paragraphe 4.2.5), dans la démonstration du théorème de convergence de Lax-Richtmyer s'appuyant sur la représentation en Fourier.

La représentation intégrale permet également, cette fois pour les conditions initiales non régulières Dirac et Heaviside étudiées dans ce livre, de prédire le comportement du schéma, en particulier les artéfacts qu'il est susceptible de générer. Pour obtenir un résultat explicite, l'intégrale est approchée par une des méthodes de développement asymptotique d'intégrales présentées dans l'appendice.

TFSD en 2 dimensions

La transformation de Fourier semi-discrète se généralise à plusieurs dimensions. Elle est ainsi définie en 2 dimensions par

$$u_{jm} \in l^2(\mathbb{Z} \otimes \mathbb{Z}) \rightarrow \widehat{u}(k, l) = \sum_{j,m=-\infty}^{j,m=+\infty} u_{jm} e^{-ikj-il m} \in L^2([-\pi, \pi] \otimes [-\pi, \pi])$$

Sa transformation inverse est donnée par

$$u_{jm} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \widehat{u}(k, l) e^{ikj+ilm} dk dl$$

Tous les résultats présentés en 1D se généralisent. Un schéma est ainsi caractérisé par son facteur d'amplification $\widehat{Q}_p(k, l)$. Par exemple, le schéma upwind, dans la généralisation la plus immédiate du schéma 1D, s'écrit, notant a et b les composantes respectives de la vitesse suivant x et y , supposées positives

$$(u_{jm}^{n+1} - u_{jm}^n)/\tau + a(u_{jm}^n - u_{(j-1)m}^n)/h + b(u_{jlm}^n - u_{j(m-1)}^n)/h = 0$$

Sous forme barycentrique, il s'écrit, en posant $q = \frac{a\tau}{h}$, $r = \frac{b\tau}{h}$

$$u_{jm}^{n+1} = (1 - q - r)u_{jm}^n + qu_{(j-1)m}^n + ru_{j(m-1)}^n \quad (4.15)$$

Nous allons appliquer la démarche générale présentée plus haut à ce schéma. En appliquant la transformation de Fourier semi-discrète à (4.15), on obtient

$$\widehat{u}^{n+1} = (1 - q - r + qe^{-ikj} + re^{-ilm})\widehat{u}^n \quad (4.16)$$

(4.16) fournit

$$\widehat{u}^n = (1 - q - r + qe^{-ikj} + re^{-ilm})^n \widehat{u}^0 \quad (4.17)$$

Dans (4.17) \widehat{Q}_p est le facteur d'amplification

$$\widehat{Q}_p(k, l) = 1 - q - r + qe^{-ikj} + re^{-ilm}$$

En appliquant la transformation de Fourier inverse à (4.17), la solution du schéma s'écrit

$$u_{jm}^n = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \widehat{Q}_p(k, l)^n e^{ikj+ilm} \widehat{u}^0 dk dl \quad (4.18)$$

\widehat{u}^0 est comme en 1D la transformée de la condition initiale. La formule (4.18) est l'équivalent 2D de (4.10). La section suivante présente la transformation de Fourier discrète, utile pour les conditions périodiques.

4.2.3 Transformation de Fourier discrète

Elle s'applique aux conditions périodiques, et associe, à une suite finie, *i.e.* un vecteur $u \equiv \{u_0, u_1, \dots, u_j, \dots, u_{J-1}\}$ (les valeurs de u aux J points du maillage, pour nos applications), une suite finie $\widehat{u} = \{\widehat{u}_0, \widehat{u}_1, \dots, \widehat{u}_k, \dots, \widehat{u}_{J-1}\}$

$$\widehat{u}_k = \sum_{j=0}^{J-1} u_j e^{-\frac{2i\pi}{J} kj} \quad (4.19)$$

La formule (4.19) peut s'écrire sous forme matricielle $\widehat{u} = Fu$, F est la matrice de Fourier. La TFD transforme S , donc l'opérateur aux différences caractérisant un schéma, en une multiplication. Donc, comme dans la section précédente, le schéma s'écrit comme une multiplication dans l'espace réciproque.

Mais la transformation inverse, donc la solution du schéma, s'écrit maintenant non comme une intégrale, mais comme une somme

$$u_j = \frac{1}{J} \sum_{k=0}^{J-1} \widehat{u}_k e^{\frac{2i\pi}{J} kj}$$

ou, en version matricielle, comme $u = F^{-1}\widehat{u}$. Le schéma est caractérisé par une matrice circulante $u^n = Mu^{n-1}$ diagonalisée par F : $FM = DF$

La solution du schéma s'écrit donc sous la forme de la somme (4.20), équivalent de (4.10) dans le cas périodique, en notant λ_k les coefficients de D

$$u_j^n = \frac{1}{J} \sum_{k=0}^{J-1} \widehat{u}_k^0 (\lambda_k)^n e^{\frac{2i\pi}{J} kj} \quad (4.20)$$

En version matricielle (4.20) s'écrit

$$u^n = F^{-1} D^n F u^0$$

L'analyse de Fourier est un outil commode pour obtenir des solutions exactes de schémas discrétisant des équations aux dérivées partielles linéaires. Elle permet aussi, plus qualitativement, de comprendre le comportement physique, *i.e.* la dispersion et la dissipation des schémas, en s'appuyant sur une visualisation dans l'espace réciproque. La dissipation est présentée dans l'introduction : elle fait décroître dans le temps l'amplitude d'une composante de Fourier, en général d'autant plus vite que k est grand. La dispersion en physique des ondes est la dépendance en k de la vitesse de propagation d'une onde à sa fréquence.

4.2.4 Visualisation dans l'espace réciproque, dispersion et dissipation d'un schéma

Les résultats dans l'espace réciproque donnent des informations utiles sur le schéma. Une visualisation adaptée, qui peut varier suivant le type de schéma, a donc son importance. L'élément essentiel est le facteur d'amplification $\widehat{Q}_p(k)$, qui est un nombre complexe. Il peut se mettre, pour un schéma totalement discrétisé, sous la forme $\widehat{Q}_p = |\widehat{Q}_p| e^{-i\tau\omega(k)}$, $\tau\omega(k) \in \mathbb{R}$ est la phase de \widehat{Q}_p . Si $|\widehat{Q}_p(k)| < 1$, le schéma diminue l'amplitude des composantes de Fourier de la condition initiale et introduit donc de la dissipation. La dépendance en k de $\omega(k)$ est appelée relation de dispersion. En effet, si $\omega(k)$ n'est pas constant, les composantes de Fourier de la condition initiale sont déphasées (translatées) non uniformément, donc dispersées. Pour un schéma semi-discrétisé, selon (4.13), la composante de vecteur d'onde k est multipliée par $\exp(\widehat{Q}_p(k)t)$. Posons $\widehat{Q}_p(k) = -i\omega(k)$. Si $\text{Re } \omega$ n'est pas constant, le schéma disperse. Si $\text{Im } \omega \neq 0$, le schéma dissipe. Le facteur d'amplification peut être visualisé de différentes manières, que nous présentons maintenant.

Schémas non dissipatifs

Si \widehat{Q}_p est de norme 1, pour un schéma totalement discrétisé, ou si ω est réel pour un schéma semi-discrétisé, la norme des composantes de Fourier est conservée, *i.e.* il n'y a pas de dissipation. Pour l'équation d'advection, le schéma saute-mouton $u_j^{n+1} - u_j^{n-1} = -\eta(u_{j+1}^n - u_{j-1}^n)$, étudié au chapitre 5, de même que le schéma centré semi-discrétisé $\partial_t u_j = -\frac{a}{2b}(u_{j+1} - u_{j-1})$, sont des exemples de schémas non dissipatifs, mais dispersifs. En effet, appliquons ce dernier à une condition

initiale $u_j(t = 0) = \exp ikj$. Le schéma s'écrit dans ce cas, puisque $u_{j+1} - u_{j-1} = \exp ik(j+1) - \exp ik(j-1) = 2i \sin(k)u_j$,

$$\partial_t u_j = -\frac{ia \sin k}{h} u_j$$

ce qui implique

$$u_j(t) = \exp\left(-\frac{iat}{h} \sin k\right) \exp ikj$$

La condition initiale $\exp ikj$ est multipliée par un facteur de phase $\exp\left(-\frac{iat}{h} \sin k\right)$, *i.e.* traduite de $\frac{at}{h} \frac{\sin k}{k}$ en variable j , et de $at \frac{\sin k}{k}$ en variable x . Elle se déplace donc à la vitesse (de phase) $a \frac{\sin k}{k}$. La vitesse de translation, qui vaut a pour tout k dans l'EDP, dépend de k : le schéma disperse les nombres d'onde. La dispersion est caractérisée par la dépendance en k de la pulsation $\omega(k)$. Pour ce schéma, elle s'écrit $\omega(k) = \frac{a}{h} \sin k$.

Relation de dispersion en 1D

Il est commode de tracer dans ce cas la relation de dispersion $\omega(k)$ du schéma, éventuellement normalisée par celle de l'équation exacte $\omega_e(k)$. La consistance du schéma assure que $\omega(k)$ tend vers $\omega_e(k)$ quand k tend vers zéro, mais ne renseigne pas sur le comportement de $\omega(k)$ pour k d'ordre 1, en particulier près de la borne supérieure de l'intervalle spectral $k = \pi$, ce que fait la relation de dispersion. Pour le schéma centré semi-discretisé (avec k adimensionné), la relation de dispersion $\omega(k) = \frac{a}{h} \sin k$ est à comparer à celle de l'EDP $\omega_e(k) = \frac{a}{h} k$. Développons le ratio $\omega(k)$ en $k = 0$

$$\omega(k) = \frac{a}{h} \sin k = \frac{a}{h} k \left(1 - \frac{1}{6} k^2\right) + O(k^3) \quad (4.21)$$

Pour le schéma saute-mouton, la relation de dispersion du mode physique (dérivée dans la section qui lui est consacrée) est $\omega_+ = \frac{1}{\tau} \arcsin(\eta \sin k)$, en développant, on obtient

$$\omega_+(k) = \frac{a}{h} k \left(1 - \frac{1}{6} k^2 (1 - \eta^2)\right) + O(k^3)$$

Dans les deux cas, le contact à l'origine avec $\omega_e(k)$ est d'ordre 2 (celui de ces schémas). Mais, comme montré sur la figure 4.1, quand k est $O(1)$, $\omega(k)$ peut sensiblement différer de $\omega_e(k)$. En particulier, pour $k = \pi$, $\omega = 0$.

En général, la relation de dispersion d'un schéma s'écarte de celle de l'équation hors du voisinage de $k = 0$, même pour les schémas d'ordre élevé. En particulier, pour tous les schémas centrés, quel que soit leur ordre, $\omega(\pi) = 0$.

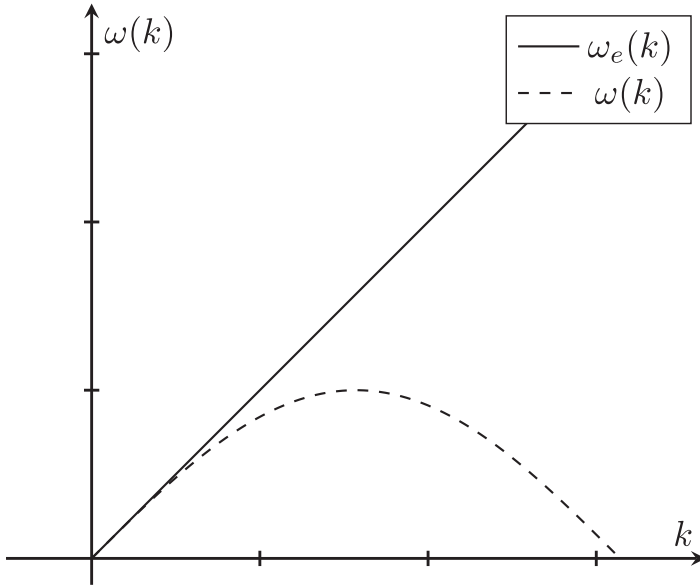


Fig. 4.1 Relation de dispersion (4.21) du schéma centré semi-discrétisé comparée à $\omega_e(k)$.

Vitesse de groupe

Introduisons enfin la notion [30] de vitesse de groupe v_g , liée à la représentation sous forme d'intégrale de Fourier de la solution d'un schéma et plus précisément à son approximation par la méthode de la phase stationnaire (présentée au chapitre 12 de l'appendice). Elle est définie par

$$v_g = \frac{d\omega}{dk}$$

Pour le schéma centré (en utilisant maintenant la variable k dimensionnée), $v_g = a \cos kh$ approche la vitesse a de l'équation pour $kh \rightarrow 0$. Une manière simple d'expliquer la notion de vitesse de groupe est la suivante. Considérons la représentation (4.14) sous forme d'intégrale de Fourier de la solution d'un schéma semi-discrétisé

$$u_j(t) = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} \widehat{u}^0(k) \exp i(kx - \omega(k)t) dk$$

Le schéma est supposé non dissipatif, donc $\omega(k)$ est réelle. Les points de phase stationnaire $k_c(x)$ annulent la dérivée en k de la phase $kx - \omega(k)t$ i.e. vérifient

$$\frac{d}{dk}(kx - \omega(k)t) = x - \frac{d\omega}{dk}t = 0$$

De par la définition de la vitesse de groupe, cette égalité implique la relation suivante entre x et k_c

$$v_g(k_c) = \frac{x}{t}$$

et la formule de la phase stationnaire (formule 12.7 de l'appendice) implique

$$u(x, t) \sim \widehat{u}^0(k_c) \exp i(k_c x - \omega(k_c)t)$$

Prenons, suivant [30], comme condition initiale un paquet d'ondes (terminologie issue de la mécanique quantique) de fréquence centrale k_0 , *i.e.* tel que $\widehat{u}^0(k)$ présente un maximum marqué en k_0 . Alors $u(x, t)$ sera petit sauf au voisinage de $k_c = k_0$, *i.e.* au voisinage de $x = tv_g(k_0)$. Le paquet d'ondes se déplace donc à la vitesse $v_g(k_0)$. En particulier, si la vitesse de groupe est nulle pour le nombre d'onde k_0 , le paquet d'ondes correspondant reste stationnaire.

Pour le schéma centré semi-discretisé étudié dans la partie Applications, par exemple, $v_g = a \cos(kh) = 0$ pour $k_0 = \pi/2h$: le paquet (ou mode) de longueur d'onde $\lambda_0 = \frac{2\pi}{k_0} = 4h$, reste stationnaire. Pour le mode de plus haute fréquence résolu sur le maillage $k_0 = \pi/h$, $v_g = a \cos(\pi) = -a$. Ce mode se propage donc à l'opposé de la vitesse d'advection. Cet exemple est souvent utilisé dans la littérature [30], car particulièrement illustratif de l'écart entre vitesse de groupe du schéma et vitesse d'advection.

Pour les conditions initiales plus générales, contenant tout le spectre de fréquences spatiales, l'approximation de la solution s'écrit sous la forme d'une onde modulée $A \exp(i k_c x - i \omega(k_c)t)$. Le nombre d'onde spatial local au point $x = tv_g(k_c)$ vaut donc k_c . Pour les schémas étudiés dans ce livre, la vitesse de groupe est une fonction décroissante de k , si bien que les oscillations des fonctions de Green des schémas ont des nombres d'onde d'autant plus grands que l'on s'éloigne du front $x = at$. Par exemple, la fonction de Green du schéma centré d'ordre 2 présente des oscillations dont le nombre d'onde croît de $k = 0$ en $x = at$ jusqu'à $k_0 = \pi/h$ en $x = -at$, comme on le verra dans la section 5.2 du chapitre 5 où il est étudié.

Relation de dispersion en 2D

En 2 dimensions, la relation de dispersion fait intervenir 2 variables. Écrivons-la, à titre d'exemple, pour le schéma centré semi-discretisé pour l'équation des ondes sur maillage carré

$$\partial_{tt} u_{jm} = a^2 \left(\frac{u_{(j+1)m} - 2u_{jm} + u_{(j-1)m}}{h^2} + \frac{u_{j(m+1)} - 2u_{jm} + u_{j(m-1)}}{h^2} \right) \quad (4.22)$$

Dans l'espace réciproque, (4.22) devient

$$\partial_{tt} \widehat{u} = 2 \frac{a^2}{h^2} ((\cos k - 1) + (\cos l - 1)) \widehat{u} \quad (4.23)$$

(4.23) conduit à la relation de dispersion

$$\omega^2(k, l) = \frac{4a^2}{b^2} \left(\sin^2 \frac{k}{2} + \sin^2 \frac{l}{2} \right)$$

à comparer à $\omega_c^2(k, l) = \frac{a^2}{b^2}(k^2 + l^2)$. Le contact à l'origine est d'ordre 2 : en effet

$$\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right) \approx \frac{1}{4}(k^2 + l^2) - \frac{1}{48}(k^4 + l^4)$$

La relation de dispersion est représentée sur les figures 4.2 et 4.3.

La vitesse de groupe du schéma s'écrit, puisque

$$\frac{d}{dk} \sqrt{\sin^2 \frac{k}{2} + \sin^2 \frac{l}{2}} = \frac{1}{4} \frac{\sin k}{\sqrt{\sin^2 \frac{k}{2} + \sin^2 \frac{l}{2}}}$$

$$v_g = \frac{a}{2\sqrt{\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right)}} (\sin k, \sin l)$$

Au voisinage de (0, 0), elle peut être approchée par

$$v_g \approx a \left(1 + \frac{1}{24} \frac{k^4 + l^4}{k^2 + l^2} \right) \left(k \left(1 - \frac{k^2}{6} \right), l \left(1 - \frac{l^2}{6} \right) \right)$$

La relation de dispersion et la vitesse de groupe ont une déviation $d = \frac{k^4 + l^4}{k^2 + l^2}$ d'ordre 2. L'isotropie de l'équation n'est respectée par le schéma qu'à son ordre. En passant en coordonnées polaires $k = r \cos \theta$, $l = r \sin \theta$, $d = r^2 (\cos^4 \theta + \sin^4 \theta) = r^2 (\frac{1}{4} \cos 4\theta + \frac{3}{4})$, soit une déviation par rapport à l'isotropie d'amplitude $\frac{r^2}{4} \cos 4\theta$. Notons aussi pour mémoire les modes stationnaires, qui sont maintenant en $k = l = \pi$.

Schémas dissipatifs

Si le facteur d'amplification est de module $\leq \exp(-c_S k^d)$, où c_S est une constante, le schéma est qualifié de dissipatif d'ordre d . Un tracé possible est celui de la courbe C parcourue dans le plan complexe par le facteur d'amplification pour $k \in [-\pi, \pi]$. Pour l'équation d'advection, il vaut, sur un pas de temps, $e^{-ika\tau}$ et C_e est donc le cercle unité centré à l'origine, d'équation $|z| = 1$. Pour un schéma dissipatif, C_s sera intérieur au cercle unité. On peut vérifier par exemple que, pour $\eta \leq 1$, pour le schéma upwind (3.4) $u_j^{n+1} = \eta u_{j-1}^n + (1 - \eta) u_j^n$, dissipatif d'ordre 2,

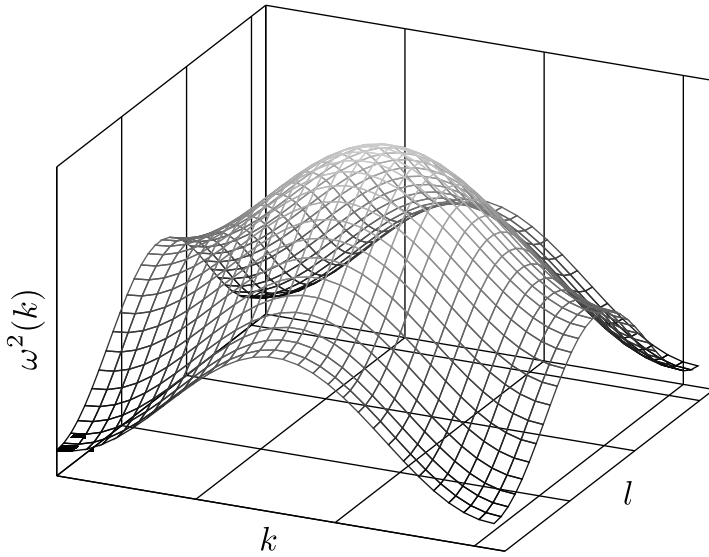


Fig. 4.2 Relation de dispersion $\omega^2(k, l) = \frac{4a^2}{h^2} (\sin^2(\frac{kh}{2}) + \sin^2(\frac{lh}{2}))$ du schéma centré semi-discrétisé.

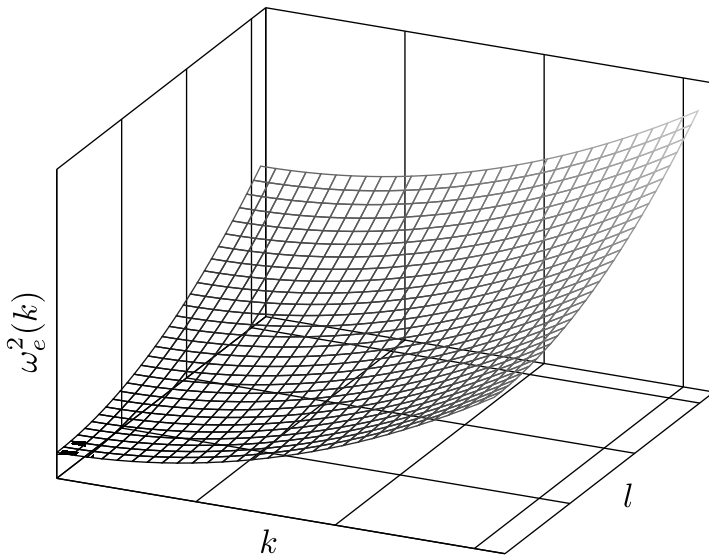


Fig. 4.3 Relation de dispersion exacte $\omega_e^2(k, l) = a^2(k^2 + l^2)$ de l'équation.

$\widehat{Q}(k) = \eta e^{-ik} + 1 - \eta$ décrit le cercle C_s de centre $1 - \eta$ de rayon η . Ce cercle est d'autant plus petit, et le schéma d'autant plus dissipatif, que η est petit, comme illustré sur la figure 4.4. Il est tangent en $k = 0$, *i.e.* en $z = \widehat{Q}(0) = 1$ au cercle unité. Pour $\eta = 1$, le schéma est exact, donc C_e et C_s sont confondus.

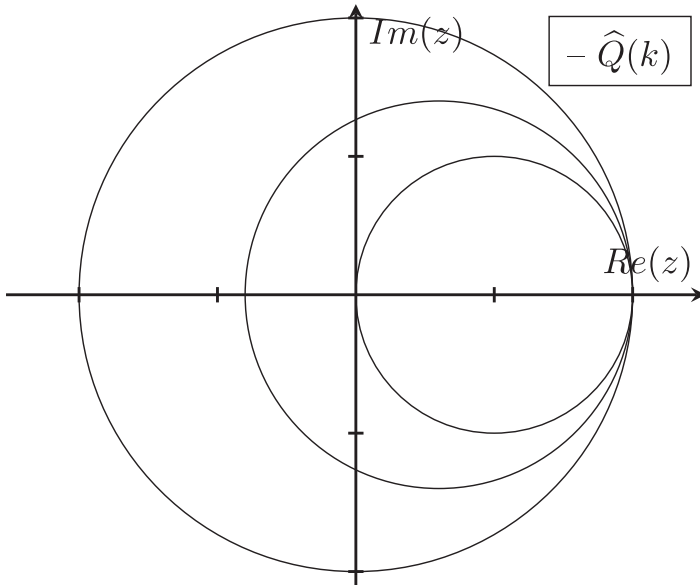


Fig. 4.4 Facteur d'amplification dans le plan complexe du schéma upwind
 $\widehat{Q}(k) = 1 - \eta + \eta e^{-ik} = 1 - \eta(1 - \cos k) - i\eta \sin k$ (pour $\eta = 0,5, 0,7, 1,0$).

Le tracé du module $g(k)$ et de l'argument du facteur d'amplification est une autre manière de visualiser respectivement dissipation et dispersion. Nous les calculons à titre d'exemple pour le schéma upwind (3.4), dont on rappelle le facteur d'amplification

$$\widehat{Q}(k) = 1 - \eta + \eta e^{-ik} = 1 - \eta(1 - \cos k) - i\eta \sin k$$

Son module vaut donc

$$g(k) = \sqrt{1 - 4\eta(1 - \eta) \sin^2(k/2)}$$

Il décroît avec k , *i.e.* les modes sont d'autant plus dissipés que leur fréquence spatiale est grande. Au voisinage de $k = 0$

$$g(k) = 1 - \frac{1}{2}k^2\eta(1 - \eta) + O(k^4) \leq \exp\left(-\frac{1}{2}k^2\eta(1 - \eta)\right)$$

Le schéma est donc dissipatif d'ordre 2. Quantifions maintenant sa dispersion en calculant la phase φ de $\widehat{Q}(k)$

$$\varphi = \arctan\left(\frac{-\eta \sin k}{1 - \eta + \eta \cos k}\right) = -k\eta + k^3\eta\left(\frac{1}{6} - \frac{1}{2}\eta + \frac{1}{3}\eta^2\right) + O(k^5)$$

Pour ce schéma d'ordre 1, la dissipation domine la dispersion, d'ordre 3. Ce serait l'inverse pour un schéma d'ordre 2.

Une approche alternative est d'écrire $\widehat{Q}(k) = \exp(\omega(k))$, où $\omega(k) = \omega'(k) + i\omega''(k)$ est un nombre complexe et de tracer la partie réelle $\omega'(k)$ (resp. imaginaire $\omega''(k)$). Ces différentes méthodes permettent de visualiser la dispersion (resp. la dissipation) du schéma, comme illustré sur les figures 4.5 et 4.6.

Choix d'une visualisation adaptée

En somme, la visualisation du facteur d'amplification ou de la relation de dispersion apporte, en une ou plusieurs dimensions, une information sur le comportement du schéma, en particulier sur sa stabilité, sa dispersion et sa dissipation. Le choix de l'une ou l'autre de ces représentations dépend du type de schéma. En outre, elle permet de comprendre comment le schéma agit sur les moyennes et hautes fréquences. Elle est donc complémentaire de l'information apportée par l'ordre, qui renseigne uniquement sur les modes de basse fréquence spatiale, au voisinage de $k = 0$.

Le facteur d'amplification d'un schéma non dissipatif parcourt le cercle unité : on voit l'intérêt particulier de tracer la relation de dispersion dans ce cas, comme nous l'avons fait pour le schéma centré sur la figure 4.1. Pour un schéma dissipatif, il est préférable de tracer le facteur d'amplification, représenté pour le schéma upwind sur les figures 4.4, 4.5 et 4.6. En deux dimensions, on peut en plus visualiser (figures 4.2 et 4.3) l'anisotropie induite par le schéma. Les exemples de schémas traités seront donc accompagnés de figures illustratives, complémentaires, pour bien comprendre leur comportement, des formules exactes ou approchées de leurs solutions. Nous renvoyons particulièrement à [32], qui présente les relations de dispersion de toute une palette de schémas.

L'analyse de Fourier permet, en plus d'obtenir des résultats explicites, d'étudier dans l'espace réciproque la convergence des schémas. Dans la section suivante, nous démontrons le théorème de Lax-Richtmyer, d'abord dans l'espace direct, puis dans l'espace réciproque.

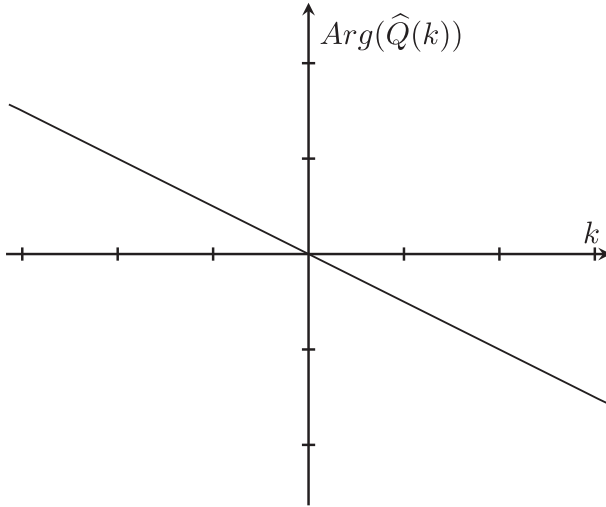


Fig. 4.5 Argument du facteur d'amplification du schéma upwind $Arg(\widehat{Q}(k)) = \arctan\left(\frac{-\eta \sin k}{1 - \eta + \eta \cos k}\right)$ (pour $\eta = 0,5$).

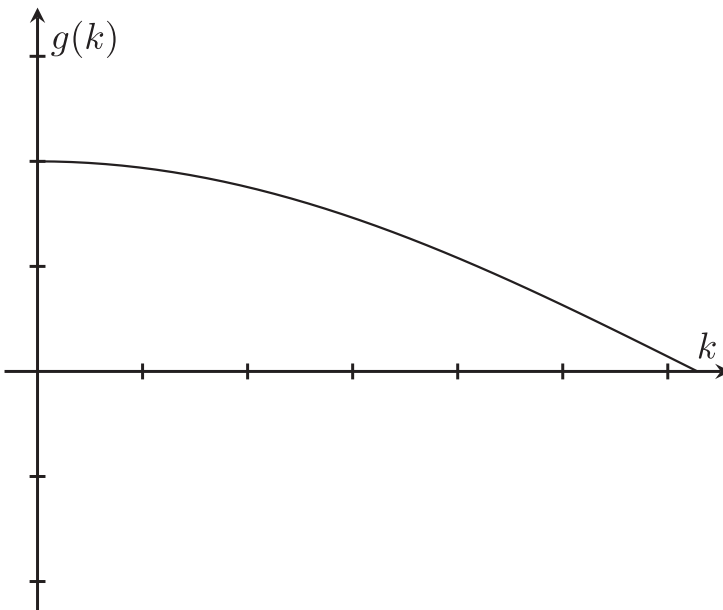


Fig. 4.6 Module du facteur d'amplification du schéma upwind $g(k) = \sqrt{(1 - 2\eta \sin^2(k/2))^2 + (\eta \sin k)^2}$ (pour $\eta = 0,5$).

4.2.5 Analyse de convergence

Analyse de convergence dans l'espace direct

Le théorème de Lax-Richtmyer a été énoncé plus haut, sans démonstration. Une démonstration fondée sur une estimation d'erreur dans l'espace réel est donnée ci-dessous. L'estimation d'erreur est donnée pour un schéma linéaire abstrait à un pas, écrit sous la forme

$$u^{n+1} = Au^n \quad (4.24)$$

u^n est la suite des u_j^n , A est un opérateur linéaire aux différences. La solution exacte, notée v^n , est injectée dans le schéma, *i.e.* dans (4.24), un développement de Taylor est réalisé, comme dans le cas des schémas à trois points (section 3.3), et on obtient

$$v^{n+1} = Av^n + r^n \tau \quad (4.25)$$

où r^n est par définition l'erreur de troncature, suite des r_j^n , $j \in \mathbb{Z}$ supposée d'ordre $h^p + \tau^q$, *i.e.* $\|r^n\| \leq C(h^p + \tau^q)$, C dépend des dérivées de la solution exacte et de l'opérateur A . Par exemple, dans le cas des schémas à trois points étudiés au chapitre 3 pour l'équation d'advection

$$r_j^n = \frac{ah}{2} \left(\eta - \frac{q}{\eta} \right) u_{xx}(jh, n\tau) + o(\tau) + o(h) \Rightarrow \|r^n\| \leq Dh \|u_{xx}\|$$

La norme $\|\cdot\|$ est une norme L^p , avec $p > 1$. L'erreur e^n est définie comme la différence $v^n - u^n$. En soustrayant (4.24) de (4.25), on obtient

$$e^{n+1} = Ae^n + r^n \tau = (A^n r_0 + \dots + A^{n-m} r^m + \dots r^n) \tau \quad (4.26)$$

Nous supposons, pour simplifier un peu la démonstration, la stabilité dite forte : $\|A\| \leq 1$, (4.26) implique

$$\|e^n\| \leq (\|r^0\| + \dots + \|r^m\| + \dots \|r^{n-1}\|) \tau \leq n\tau C(h^p + \tau^q) = CT(h^p + \tau^q)$$

ce qui démontre le théorème de Lax-Richtmyer : un schéma consistant à l'ordre p en espace et q en temps, est convergent à ces mêmes ordres s'il est stable, pour des conditions initiales assez régulières. Le niveau de régularité nécessaire apparaît via les développements de Taylor, qui doivent être menés (si par exemple $q = p$) jusqu'à l'ordre $p + 1$.

Analyse de convergence dans l'espace réciproque

Une démonstration est également possible dans l'espace réciproque. L'erreur de troncature est obtenue en insérant dans le schéma la solution exacte de l'équation, mais sous forme intégrale, par exemple en utilisant (4.10). Pour la stabilité, le facteur

d'amplification doit être ≤ 1 pour tout k : la partie réelle de $\widehat{Q}_p(\xi)$ doit être ≤ 0 . La vérification de la consistance à l'ordre p et de la stabilité en Fourier reviennent respectivement à s'assurer que $|\widehat{Q}(k) - \widehat{Q}_p(k)| = O((kh)^p)$ pour $k \rightarrow 0$ et que $\text{Re}(\widehat{Q}_p(k)) \leq 0$ pour $k \in [-\pi/h, \pi/h]$.

L'ordre de consistance est une propriété locale, pilotée par l'ordre du contact entre le facteur d'amplification de l'équation $\widehat{Q}(k)$ et celui du schéma $\widehat{Q}_p(k)$ au seul point $k = 0$, alors que la condition de stabilité est globale : elle porte sur tout l'intervalle des nombres d'onde. Bien souvent, les modes instables ont des nombres d'onde $k = O(1)$, ce qui génère l'amplification de ces nombres d'onde, *i.e.* des hautes fréquences spatiales. La démonstration est donnée ci-dessous pour un schéma à un pas, pour l'équation d'advection.

Partons de la solution explicite (4.9) du schéma discret en Fourier

$$\widehat{u}^n(k) = \widehat{Q}_p(k)^n \widehat{u}^0$$

à comparer à la solution exacte de l'équation en Fourier

$$\widehat{v}(n\tau) = \exp(\widehat{Q}n\tau) \widehat{u}^0$$

La différence entre ces deux quantités s'écrit

$$\widehat{e}(n\tau) = (\exp(\widehat{Q}n\tau) - \widehat{Q}_p(k)^n) \widehat{u}^0$$

Revenant à l'espace réel, e s'exprime donc comme la somme de deux intégrales, dont la première I_1 porte sur le spectre résolu sur le maillage

$$e = \frac{1}{\pi} \int_{-\pi/h}^{\pi/h} (\exp(\widehat{Q}t) - \widehat{Q}_p(k)^n) \widehat{u}^0(k) e^{ikx} dk + \frac{1}{\pi} \int_{\pi/h}^{+\infty} \widehat{u}^0(k) e^{ikx} dk = I_1 + I_2 \quad (4.27)$$

Majorons l'intégrande de I_1 , en utilisant l'identité

$$a^n - b^n = (a - b)(a^{n-1} + \dots + b^{n-1})$$

la stabilité du schéma implique $|\widehat{Q}_p| \leq 1$, et comme $|\exp(\widehat{Q}t)| = |e^{-ikat}| = 1$

$$|\exp(\widehat{Q}t) - \widehat{Q}_p(k)^n| \leq n |\exp(\tau \widehat{Q}(k)) - \widehat{Q}_p(k)|$$

Il est possible, comme démontré dans le chapitre 10 du livre de Strikwerda [25], de majorer sur $[-\pi/h, \pi/h]$ la différence des facteurs d'amplification en utilisant la condition d'ordre de consistance $\exp(\tau \widehat{Q}(k)) - \widehat{Q}_p(k) = O((kh)^p)$ pour $kh \rightarrow 0$. La majoration (4.28) de la différence des facteurs d'amplification est obtenue dans [25]

$$|\exp \tau \widehat{Q}(k) - \widehat{Q}_p(k)| \leq C \tau h^p (1 + |k|^\rho), \quad k \in [-\pi/h, \pi/h] \quad (4.28)$$

Le théorème 10.3.1 de [25], dont la démonstration est assez longue, montre que, pour tout schéma d'ordre p , (4.28) est vérifiée pour un certain ρ . Le nombre ρ dépend du schéma, qualifié de précis d'ordre $[p, \rho]$, mais est égal à $p + 1$ pour la majorité des schémas. Dans l'appendice consacré aux schémas de Strang pour l'équation d'advection, où nous suivons l'approche de Després [1, chapitre 10], la majoration est obtenue directement en appliquant la formule de Taylor avec reste intégral à $\exp \tau \widehat{Q}(k)$. Comme la consistance à l'ordre p implique que \widehat{Q}_p est précisément le développement de Taylor de $\exp \tau \widehat{Q}(k)$ la différence se réduit au reste intégral

$$\exp(\tau \widehat{Q}(k)) - \widehat{Q}_p(k) = \frac{1}{p!} \int_0^1 (1-x)^p \frac{d^{p+1}}{dk^{p+1}} (\exp x \tau \widehat{Q}(k)) dx$$

ce qui permet d'obtenir une majoration de la différence des facteurs d'amplification de la forme (4.28) avec $\rho = p + 1$. On obtient finalement une majoration du premier terme de l'erreur

$$\begin{aligned} I_1 &\leq \int_{-\pi/h}^{\pi/h} |\exp(\widehat{Q}_p t) - \exp(\widehat{Q} t)| |\widehat{u}_0(k)| dk \\ &\leq C n \tau h^p \int_{-\infty}^{\infty} (1 + |k|^\rho) |\widehat{u}^0(k)| dk \leq C_1 t h^p \|u(0)\|_{H^\rho} \end{aligned} \quad (4.29)$$

Considérons la seconde intégrale $I_2 = \frac{1}{\pi} \int_{\pi/h}^{+\infty} \widehat{u}_0(k) e^{ikx} dk$. Dans cet intervalle d'intégration, $1 \leq \frac{kb}{\pi}$

$$I_2 \leq \int_{\pi/h}^{\infty} |\widehat{u}_0(k)| dk \leq \left(\frac{h}{\pi}\right)^p \int_0^{\infty} k^p |\widehat{u}_0(k)| dk \leq C_2 h^p \|u\|_{H^\rho} \quad (4.30)$$

Au final, on obtient, donc pour l'erreur totale, combinant (4.29) et (4.30), la majoration

$$|e(t)| \leq C(t) h^p \|u(0)\|_{H^\rho}$$

La majoration dépend de la norme H^ρ de la condition initiale, et n'est donc valide que si u^0 est dans H^ρ , défini ici comme l'espace des fonctions telles que l'intégrale dans (4.30) converge. Pour retomber sur les espaces de Sobolev usuels, il faut utiliser, comme dans [25], la norme L^2 . Le schéma n'atteint son ordre théorique que pour des conditions initiales assez régulières (d'autant plus régulières que l'ordre du schéma est élevé).

C dépend en général de t : l'erreur croît avec le temps pour la majorité des schémas. Mais il existe aussi des schémas pour lesquels C ne dépend pas de t : il est préférable de les utiliser s'il est nécessaire de calculer avec précision aux temps longs.

La démonstration du théorème de Lax-Richtmyer dans l'espace réciproque est plus longue que dans l'espace direct, mais elle permet de bien comprendre le rôle de la régularité des conditions initiales. De plus, elle peut être adaptée aux conditions initiales non régulières considérées dans ce livre. Dans la deuxième partie, consacrée aux applications, l'ordre de convergence pour une condition initiale non régulière, de type marche, est calculé : il est, en norme L^1 , inférieur à p .

Ainsi se conclut la section consacrée à l'analyse de Fourier. Nous traitons maintenant de la méthode de l'équation équivalente, de validité moins générale, mais à même de fournir simplement des approximations explicites des solutions des schémas.

4.3 Méthode de l'équation équivalente

Nous avons introduit dans la première partie la notion d'équation équivalente à un schéma. Nous allons maintenant donner plus de précisions sur cette technique d'analyse, présentée en détail dans [24]. Une équation aux dérivées partielles est plus facile à étudier qu'une équation aux différences. Le but de la méthode est de chercher une équation aux dérivées partielles « équivalente » à l'équation aux différences (partielles) définissant le schéma. L'équation équivalente (EE), utilisée dans son domaine de validité, est une méthode puissante et commode d'analyse quantitative des schémas, mais il faut bien avoir conscience de ses limites...

Pour obtenir l'équation équivalente, l'équation aux différences définissant le schéma est d'abord étendue des $x = jh$, $t = n\tau$, discrets, aux x, t continus. Puis on « approxime » cette équation aux différences en domaine continu par une EDP. Un schéma d'ordre p est ainsi consistant à l'ordre p avec l'équation initiale $L(u) = 0$, mais à l'ordre $p + q$ avec une équation équivalente avec q termes au second membre. L'équation équivalente au schéma s'écrit ainsi

$$L(u) = h^p M_1(u) + \dots + h^{p+q-1} M_q(u)$$

Les M_1, \dots, M_q sont des opérateurs différentiels. Pour l'équation d'advection en 1D, l'équation équivalente s'écrit

$$\partial_t u + a \partial_x u = d_2 h \partial_{xx} u + d_3 h^2 \partial_{3x} u + d_4 h^3 \partial_{4x} u + \dots$$

Les termes pairs comme $d_2 h \partial_{xx} u$ et $d_4 h^3 \partial_{4x} u$ induisent de la dissipation. La diffusion $d_2 h \partial_{xx} u$ est un cas particulier de dissipation. Les termes impairs comme $d_3 h^2 \partial_{3x} u$ induisent de la dispersion. Pour un schéma d'ordre p , le premier terme non nul est $d_{p+1} h^p \partial_{px} u$.

Pour des conditions initiales assez régulières, l'application du théorème de Lax-Richtmyer montre que la différence entre la solution de l'équation équivalente et la solution du schéma est $O(h^{p+q})$. Mais l'équation équivalente est aussi utilisée pour des conditions initiales irrégulières : une marche, ou un créneau et les conditions de

régularité précédentes ne sont pas satisfaites. La méthode rend compte malgré tout assez fidèlement des artéfacts générés par certains schémas. Les artéfacts les plus courants sont l'étalement artificiel des discontinuités, induit par la dissipation, comme on l'a vu par exemple pour le schéma upwind, et les oscillations parasites au voisinage des discontinuités, générées par la dispersion. Il est souvent admis que deux termes, l'un caractérisant la dissipation, l'autre la dispersion, suffisent pour rendre compte du comportement de la plupart des schémas. Une justification de cette assertion est proposée un peu plus loin, dans la sous-section 4.3.4.

Nous présentons maintenant plusieurs méthodes pour obtenir l'équation équivalente à un schéma, sur l'exemple de l'équation d'advection. La plus simple est fondée sur un développement de Taylor.

4.3.1 Première méthode : développement de Taylor et récursion

En développant, dans l'équation aux différences définissant le schéma, les $u(j'h, n'\tau)$ au voisinage de $(jh, n\tau)$, on obtient une EE avec des dérivées en temps et en espace.

Les dérivées en temps sont ensuite exprimées récursivement en fonction des dérivées en espace, ce qui permet d'obtenir une EE (sous forme appelée Π dans [24]) avec seulement des dérivées en espace au second membre. Reprenons le schéma à trois points étudié plus haut, écrit sous forme diffusive

$$\frac{u_j^{n+1} - u_j^n}{\tau} + a \frac{u_{j+1}^n - u_{j-1}^n}{2h} = \frac{v h^2}{2\tau} \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2} \quad (4.31)$$

et calculons son EE par développement de Taylor. Le terme $\frac{u_j^{n+1} - u_j^n}{\tau}$ discrétisant la dérivée en temps dans (4.31) s'écrit, en développant en τ à l'ordre 2 $u(jh, (n+1)\tau)$

$$u(jh, (n+1)\tau) - u(jh, n\tau) = u_t(jh, n\tau)\tau + \frac{1}{2}u_{tt}(jh, n\tau)\tau^2 + o(\tau^2)$$

Exprimons la dérivée en temps u_{tt} en fonction de la dérivée en espace en dérivant en temps l'équation initiale $u_t = -au_x$

$$u_{tt}(jh, n\tau) = -au_{xt}(jh, n\tau)$$

puis, dérivant en espace $u_t = -au_x$, la dérivée croisée s'écrit $u_{xt} = -au_{xx}$ si bien que $u_{tt}(jh, n\tau) = -au_{xt}(jh, n\tau) = a^2u_{xx}(jh, n\tau)$. Finalement,

$$u(jh, (n+1)\tau) - u(jh, n\tau) = u_t(jh, n\tau)\tau + \frac{a^2}{2}u_{xx}(jh, n\tau)\tau^2 + o(\tau^2)$$

Il apparaît donc au second membre un terme d'antidiffusion $-\frac{a^2}{2}u_{xx}(jh, n\tau)\tau^2$. Le terme centré de dérivée spatiale de (4.31) s'écrit

$$u((j+1)h, n\tau) - u((j-1)h, n\tau) = 2u_x(jh, n\tau)h + o(h^2)$$

Il n'introduit pas, à cet ordre, de terme dans l'équation équivalente. Développons au second membre de (4.31) les termes $u((j+1)h, \tau)$ et $u((j-1)h, \tau)$ en h

$$u((j+1)h, \tau) - 2u(jh, n\tau) + u((j-1)h, n\tau) = u_{xx}(jh, n\tau)h^2 + o(h^3)$$

Un terme de diffusion $u_{xx}(jh, n\tau)h^2$ apparaît. En reportant ces résultats dans (4.31), cette contribution diffusive de la discrétisation spatiale se combine à celle, antidiffusive de la discrétisation temporelle et on obtient l'équation équivalente à un terme du schéma, qui s'écrit finalement

$$\partial_t u + a\partial_x u = -\frac{a^2\tau}{2}\partial_{xx}u + \frac{v h^2}{2\tau}\partial_{xx}u = \frac{ab}{2}\left(\frac{v}{\eta} - \eta\right)\partial_{xx}u \quad (4.32)$$

Pour le schéma upwind, $v = \eta$: le terme de diffusion vaut $\frac{ab}{2}(1 - \eta)u_{xx}$. Le coefficient de diffusion, donc $(1 - \eta)$, doit être positif pour que le schéma soit stable. Il introduit alors une diffusion, d'autant plus importante que $(1 - \eta)$ est grand : plus le pas de temps est petit, plus le schéma diffuse. *A contrario*, pour $\eta = 1$, il est exact et la diffusion s'annule. L'EE fournit la limite de stabilité à $\eta = 1$ du schéma.

Pour $v = \eta^2$, *i.e.* pour le schéma de Lax-Wendroff, le terme de diffusion s'annule. Des calculs analogues aux précédents (il suffit de développer à l'ordre 2) montrent que l'équation équivalente s'écrit dans ce cas

$$\partial_t u + a\partial_x u = db^2\partial_{3x}u$$

Son premier terme est maintenant dispersif. Nous calculons les équations équivalentes à plusieurs schémas d'ordre 2 dans le chapitre 5. La méthode présentée est directe, mais quelque peu laborieuse, notamment à cause de la récursion. D'autres méthodes permettent un calcul direct de l'équation équivalente.

4.3.2 **Seconde méthode : analyse de Fourier**

Partons de la solution du schéma dans l'espace réciproque, obtenue, comme expliqué plus haut par TFSD

$$\widehat{u}(k, n\tau) = \widehat{u}^0 \widehat{Q}^n = \widehat{u}^0 \exp(n \ln \widehat{Q})$$

où $\widehat{Q}(k)$ est le facteur d'amplification du schéma. Défini sur l'intervalle $[-\pi, \pi]$, il est étendu à tout \mathbb{R} . Le schéma peut être considéré comme un opérateur pseudo-différentiel. L'équation équivalente est l'approximation de cet opérateur par un opérateur différentiel. Elle est obtenue par un développement de Taylor du logarithme népérien $\ln(\widehat{Q})$ en $k = 0$.

$$\ln(\widehat{Q}) \approx (-ika + \sum_{m=2}^q r_m k^m) \tau$$

On obtient ainsi une expression polynomiale dans l'espace réciproque, correspondant à une équation aux dérivées partielles, puisqu'une multiplication par ik dans cet espace se traduit par une dérivation à l'ordre m dans l'espace direct. Si tous les termes du développement sont conservés, *i.e.* $q = \infty$, et si $\ln \widehat{Q}$ est une fonction analytique, la méthode de l'équation équivalente revient à remplacer $\ln \widehat{Q}$ par son développement en série de Taylor.

Solution de l'équation équivalente et détermination de ses coefficients

Le facteur d'amplification exact est $\exp(-ika\tau)$, celui du schéma est $\exp(\ln \widehat{Q})$, celui de l'équation équivalente (coïncidant en $k = 0$ avec celui du schéma à l'ordre q) est $\exp(-ika + \sum_{m=2}^q r_m k^m) \tau$. Autrement dit, sa solution s'écrit

$$\widehat{u}(k, t) = \widehat{u}^0 \exp(-ika + \sum_{m=2}^q r_m k^m) t$$

On reconnaît la solution de l'équation $u_t + au_x = \sum_{m=2}^q q_m u_{x^m}$ qui s'écrit dans l'espace réciproque

$$\partial_t \widehat{u} + ik a \widehat{u} = \sum_{m=2}^q q_m (ik)^m \widehat{u}$$

La solution de cette EDO est

$$\widehat{u}(k, t) = \widehat{u}^0 \exp(-ika + \sum_{m=2}^q q_m (ik)^m) t$$

Par identification des coefficients $q_m = (-i)^m r_m$. En somme, les coefficients de l'équation équivalente se calculent directement, simplement par développement de

Taylor du logarithme népérien du facteur d'amplification. La simplification par rapport à la méthode de récursion est considérable. La méthode est rapide et directe, et peut aisément s'automatiser en utilisant un logiciel de calcul formel. Une variante dans l'espace direct est présentée ci-dessous.

4.3.3 Calcul par méthode opératorielle

Un schéma numérique à un pas, comme expliqué dans la section sur les méthodes d'analyse opératorielle, peut être considéré comme une relation univoque entre l'opérateur shift en temps T et l'opérateur shift en espace S .

Pour obtenir une approximation différentielle du schéma, il faut l'exprimer comme une relation entre l'opérateur de dérivée en temps ∂_t et l'opérateur de dérivée en espace ∂_x . Pour cela, on exprime les opérateurs shift en fonction des opérateurs différentiels, soit

$$T = \exp(\tau \partial_t) \quad \text{et} \quad S = \exp(h \partial_x)$$

Un schéma à un pas avec un stencil avec r points à gauche et s points à droite s'écrit sous la forme opératorielle

$$T = a_{-r} S^{-r} + a_{-r+1} S^{-r+1} + \dots + a_0 I + \dots + a_s S^s \quad (4.33)$$

En substituant les expressions de T et de S dans l'équation (4.33), on obtient

$$\exp(\tau \partial_t) = a_{-r} \exp(-r h \partial_x) + a_{-r+1} \exp(-(r-1) h \partial_x) S^{-r+1} + \dots + a_0 I + \dots + a_s \exp(s h \partial_x)$$

En prenant le logarithme de (4.33), on obtient l'expression exacte de l'approximation de ∂_t par le schéma, soit

$$\begin{aligned} \partial_t &= \frac{1}{\tau} \ln(a_{-r} \exp(-r h \partial_x) + a_{-r+1} \exp(-(r-1) h \partial_x) S^{-r+1} \\ &\quad + \dots + a_0 I + \dots + a_s \exp(s h \partial_x)) \end{aligned} \quad (4.34)$$

Comme dans le paragraphe précédent, le développement de Taylor en $h = 0$ de l'expression (4.33) ci-dessus fournit l'approximation différentielle du schéma. Ainsi, pour le schéma upwind, qui s'écrit $T = pI + qS^{-1}$, (4.34) est simplement

$$\partial_t = \frac{1}{\tau} \ln(p + q \exp(-h \partial_x)) \quad (4.35)$$

Il suffit de développer à l'ordre souhaité (ici 2) au voisinage de 0 l'expression (4.34) formellement en $h \partial_x$, soit d'écrire

$$\ln(1 - q + q \exp(-h \partial_x)) = -h q \partial_x - \frac{1}{2} h^2 (q^2 - q) \partial_x^2 + O(h^3)$$

pour obtenir

$$\partial_t = -\frac{hq}{\tau} \partial_x + \frac{pqh^2}{2\tau} \partial_{xx} + O(h^3)$$

mais $\frac{hq}{\tau} = \frac{h}{\tau} \frac{a\tau}{h} = a$, et l'EE à 1 terme s'écrit

$$\partial_t u + a \partial_x u = \frac{pab}{2} \partial_{xx} u$$

On retrouve (4.32). Méthode opératorielle et méthode de Fourier sont équivalentes : dans la première, l'opérateur ∂_x est exprimé dans l'espace réel, dans la seconde, dans l'espace réciproque, où il devient une multiplication par ik . Ces méthodes sont plus directes et moins laborieuses que la méthode de récursion. Elles s'appliquent directement aux équations linéaires, traitées dans ce livre. Pour les équations non linéaires, l'analyse est possible [2] en linéarisant autour d'une valeur particulière de u .

Une restriction possible sur la validité de la méthode

Pour que la méthode ait un sens, il faut au moins que le développement en série de $\ln \widehat{Q}$ converge sur l'intervalle $[-\pi, \pi]$, en somme que le rayon de convergence R_C de la série soit $\geq \pi$. Si c'est le cas, l'équation équivalente (à nombre infini de termes) est strictement équivalente au schéma. Cette condition n'est pas toujours vérifiée. Calculer R_C n'est pas toujours possible, mais si $R_C < \pi$, l'EE n'est plus représentative du schéma. Par exemple, elle ne donne pas la bonne limite de stabilité et a un terme d'antidissipation alors que le schéma est stable. Cela se produit notamment pour les schémas analysés dans le chapitre 8 sur l'équation de la chaleur.

4.3.4 Validité de l'équation équivalente à nombre fini de termes

En pratique, il faut bien sûr s'arrêter à un nombre q fini de termes. La validité de la méthode se déduit de la relation entre la solution exacte du schéma et la solution de l'équation équivalente. L'équation équivalente remplace le logarithme du facteur d'amplification exact du schéma $\widehat{Q}_S(k)$ par un développement de Taylor $\widehat{Q}_{EE}(k)$, avec autant de termes qu'au second membre de l'équation, au voisinage de $k = 0$. La différence entre la solution du schéma et celle de l'équation équivalente s'écrit

$$\begin{aligned} \frac{1}{2\pi} \left(\int_{-\pi/h}^{\pi/h} (\widehat{Q}_S(k)^n - \widehat{Q}_{EE}(k)^n) \widehat{u}^0(k) e^{ikx} dk + \int_{\pi/h}^{+\infty} \widehat{Q}_{EE}(k)^n \widehat{u}^0(k) e^{ikx} dk \right. \\ \left. + \int_{-\infty}^{-\pi/h} \widehat{Q}_{EE}(k)^n \widehat{u}^0(k) e^{ikx} dk \right) \end{aligned} \quad (4.36)$$

La méthode la plus générale est d'estimer l'intégrale (4.36), mais des critères plus aisés à vérifier permettent souvent de conclure plus simplement. Nous en présentons

quelques-uns, en commençant par les plus simples. Un critère heuristique, comportant une part d'appréciation plus subjective, est que l'essentiel du spectre de la solution se trouve dans une zone où la différence entre le facteur d'amplification du schéma $\widehat{Q}_S(k)$ et celui de l'équation équivalente $\widehat{Q}_{EE}(k)$ est suffisamment petite. On peut le quantifier sous la forme

$$\int_{-\pi/h}^{\pi/h} |\widehat{Q}_S(k) - \widehat{Q}_{EE}(k)| |\widehat{u}^0(k)| dk < \text{seuil}$$

Il est aussi possible de démontrer la validité de l'équation équivalente, au moins dans un certain nombre de cas passés en revue ci-après.

Régularité des conditions initiales

Si le spectre de la condition initiale, donc de la solution, est principalement au voisinage de $k = 0$, ce qui se passe pour des conditions initiales régulières, quel que soit le schéma, pour k au voisinage de $k = 0$, $|\widehat{Q}_S(k) - \widehat{Q}_{EE}(k)|$ est petit, et ailleurs, $\widehat{u}(k)$ est petit, ce qui permet de conclure.

Par contre, pour les conditions initiales de type Dirac et marche, avec, respectivement, $\widehat{u}^0(k) = 1$ et $\widehat{u}^0(k) = 1/ik$, ce n'est plus le cas, même en prenant plusieurs termes dans l'équation équivalente.

Cas des schémas dissipatifs

Si le schéma est dissipatif, même pour des équations hyperboliques, les modes haute fréquence sont amortis aux temps longs, si bien que l'équation équivalente devient valide, même pour des conditions initiales peu régulières. Avant de le voir sur des exemples dans le chapitre suivant, consacré aux applications, nous allons estimer la différence e entre la fonction de Green du schéma G_S et celle de l'équation équivalente G_{EE} . Pour les comparer, nous étendons G_S à x réel, et définissons le facteur d'amplification Q_{EE} sur un pas de temps (d'autres choix sont possibles [25]), ce qui permet d'écrire

$$\begin{aligned} e(x, t = n\tau) = & \frac{1}{2\pi} \left(\int_{-\pi/h}^{\pi/h} (\widehat{Q}_S(k)^n - \widehat{Q}_{EE}(k)^n) e^{ikx} dk + \int_{\pi/h}^{+\infty} \widehat{Q}_{EE}(k)^n e^{ikx} dk \right. \\ & \left. + \int_{-\infty}^{-\pi/h} \widehat{Q}_{EE}(k)^n e^{ikx} dk \right) \end{aligned} \quad (4.37)$$

(4.37) est la somme d'un terme sur la bande résolue par le schéma

$$I_1 = \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} (\widehat{Q}_S(k)^n - \widehat{Q}_{EE}(k)^n) e^{ikx} dk \quad (4.38)$$

et d'un terme généré par l'extension à tout l'axe réel du spectre

$$I_2 = \frac{1}{2\pi} \int_{\pi/h}^{+\infty} \widehat{Q}_{EE}(k)^n e^{ikx} dk + \frac{1}{2\pi} \int_{-\infty}^{-\pi/h} \widehat{Q}_{EE}(k)^n e^{ikx} dk$$

Majorons l'intégrande de I_1 sur $k \in [-\pi/h, \pi/h]$, en utilisant l'identité

$$a^n - b^n = (a - b)(a^{n-1} + \dots + b^{n-1})$$

De par la dissipativité à l'ordre d du schéma $|\widehat{Q}_S| \leq \exp(-c_S |kh|^d)$, et de l'équation équivalente $|\widehat{Q}_{EE}| \leq \exp(-c_{EE} |kh|^d)$

$$|\widehat{Q}_S(k)^n - \widehat{Q}_{EE}(k)^n| \leq n |\widehat{Q}_S(k) - \widehat{Q}_{EE}(k)| \exp(-c(n-1) |kh|^d)$$

avec $c = \inf(c_S, c_{EE})$. Majorant comme dans la démonstration du théorème de Lax-Richtmyer (section 3.3) la différence des facteurs d'amplification $\widehat{Q}_S(k) - \widehat{Q}_{EE}(k)$ en utilisant la condition d'ordre de consistance pour $kh \rightarrow 0$, on obtient une majoration globale (en se référant à [25]) de la différence des facteurs d'amplification de la forme (r est l'ordre de consistance du schéma, mais par rapport à l'équation équivalente, et dépend entre autres du nombre de termes dans l'équation équivalente)

$$|\widehat{Q}_S(k) - \widehat{Q}_{EE}(k)| \leq C\tau h^r |k|^{r+1} = C\eta h^{r+1} |k|^{r+1}, \quad k \in [-\pi/h, \pi/h]$$

On obtient donc une majoration de l'intégrande par un terme de la forme $|k|^{r+1} \exp(-c(n-1) |k|^d)$ et, comme l'intégrale est convergente, une majoration de $\epsilon(t = n\tau)$.

Traitons d'abord le cas particulier du schéma upwind, et de son équation équivalente (4.32) à un terme (de diffusion), où tout s'écrit simplement. Le facteur d'amplification du schéma est

$$\begin{aligned} \widehat{Q}_S(k) &= 1 - \eta(1 - \cos kh) - i\eta \sin kh \\ kh &= 1 - ikb\eta - \frac{1}{2}k^2 b^2 \eta + \frac{1}{6}ik^3 b^3 \eta + O((kh)^4) \end{aligned}$$

Le facteur d'amplification de l'équation équivalente à un terme (4.32) s'écrit

$$\widehat{Q}_{EE}(k) = 1 - ikb\eta - \frac{1}{2}k^2 b^2 \eta + O((kh)^4)$$

Au voisinage de 0

$$\widehat{Q}_S(k) - \widehat{Q}_{EE}(k) = \frac{1}{6}i\eta k^3 b^3 + O((kh)^4)$$

La différence entre les deux facteurs d'amplification est donc (toujours en suivant la démonstration de [25], pour passer de la majoration locale à la majoration globale) majorée

$$|\widehat{Q}_S(k) - \widehat{Q}_{EE}(k)| \leq C\eta |k|^3 h^3, k \in [-\pi/h, \pi/h]$$

On notera que l'équation équivalente est consistante avec le schéma à l'ordre $r = 2$. Les facteurs d'amplification sont aussi majorés

$$|\widehat{Q}_S| \leq |\widehat{Q}_{EE}| \leq \exp\left(-\frac{1}{2}k^2 h^2 n\eta(1-\eta)\right), k \in [-\pi/h, \pi/h]$$

On peut donc majorer l'intégrale I_1

$$I_1 \leq 2Cn\eta \int_0^{\pi/h} k^3 h^3 \exp\left(-\frac{1}{2}k^2 h^2 (n-1)\eta(1-\eta)\right) dk$$

En passant dans l'intégrale à la variable sans dimension $l = k\sqrt{atb}$, on obtient

$$\begin{aligned} & n\eta \int_0^{\pi/h} k^3 h^3 \exp\left(-\frac{1}{2}k^2 h^2 (n-1)\eta(1-\eta)\right) dk \\ &= \frac{h}{(at)^2} n\eta \int_0^{\sqrt{at/b}} l^3 \exp\left(-\frac{1}{2}l^2 (1-\eta)\left(1-\frac{\tau}{t}\right)\right) dl \end{aligned}$$

$$\frac{h}{(at)^2} n\eta = \frac{h}{(at)^2} \frac{t}{\tau} \frac{a\tau}{h} = \frac{1}{at}$$

$$\int_0^{\infty} l^3 \exp\left(-\frac{1}{2}(1-\eta)\left(1-\frac{\tau}{t}\right)l^2\right) dl \leq M$$

pour $\eta < 1$ et $\frac{\tau}{t} > \frac{1}{2}$. On en déduit la majoration

$$I_1 \leq \frac{CM}{at} \tag{4.39}$$

On notera la similitude de la démonstration avec celle plus haut du théorème de Lax Richtmyer : dans les deux cas, le point essentiel est qu'une intégrale sur la variable spectrale est finie, grâce à la régularité des CI dans le premier cas, et à la dissipation dans le second. Majorons maintenant

$$I_2 = \int_{\pi/h}^{+\infty} \exp\left(-\frac{1}{2}k^2 h^2 n\eta(1-\eta)\right) e^{ikx-ikhm\eta} dk$$

$$I_2 \leq \int_{\pi/h}^{+\infty} \exp\left(-\frac{1}{2}k^2 b^2 n\eta(1-\eta)\right) dk = \frac{1}{\sqrt{ath}} \int_{\pi\sqrt{\frac{a}{b}}}^{+\infty} \exp\left(-\frac{1}{2}l^2(1-\eta)\right) dl \quad (4.40)$$

qui est donc exponentiellement petite. On obtient finalement, en combinant (4.39) et (4.40), une majoration de e de la forme

$$e(t = n\tau) \leq \frac{D}{at}$$

La fonction de Green de l'équation équivalente converge donc vers celle du schéma, mais seulement aux temps longs. Raisononnons plutôt en valeurs relatives, pour s'affranchir de la dépendance du résultat à une définition particulière de la fonction de Green. Les fonctions de Green de l'équation équivalente G_{EE} et du schéma G_S sont, avec la définition choisie dans cette section (condition initiale $\frac{1}{h}$), $O\left(\frac{1}{\sqrt{ah\tau}}\right)$. Plus précisément, G_{EE} est la fonction de Green de l'équation de la chaleur $G(x, t) = \frac{1}{2\sqrt{\pi\alpha t}} \exp\left(-\frac{x^2}{4\alpha t}\right)$ avec un α (déterminé dans la section précédente) $\frac{pab}{2}$, si bien que

$$\frac{1}{2\sqrt{\pi\alpha t}} = \frac{1}{2\sqrt{\pi\frac{pab}{2}t}} = \frac{1}{\sqrt{2\pi p}\sqrt{ah\tau}}$$

La différence relative $\frac{\epsilon}{G}$ est donc $O\left(\frac{\sqrt{ah\tau}}{at}\right) = O\left(\frac{\sqrt{h}}{\sqrt{an\tau}}\right) = O\left(\frac{1}{\sqrt{\eta n}}\right)$. En somme, la convergence est obtenue pour de grandes valeurs de n , *i.e.* pour un grand nombre de pas de temps.

Ce résultat de convergence est retrouvé par une méthode probabiliste dans la section de la partie Applications consacrée au schéma upwind. Elle se ramène dans ce cas à la convergence de la loi binomiale, qui donne la solution exacte du schéma, vers la loi normale, qui donne celle de l'équation équivalente. C'est un cas d'application du théorème central limite, d'où la convergence en $\frac{1}{\sqrt{n}}$.

Pour un schéma dont le premier terme est dissipatif d'ordre 4, le premier terme de la différence des facteurs d'amplification est en $\eta k^5 b^5$, la majoration sur les facteurs d'amplification en $\exp(-dk^4 b^4 n\eta)$. Il faut donc maintenant majorer l'intégrale

$$I = n\eta \int_0^{\pi/h} k^5 b^5 \exp(-dk^4 b^4 n\eta) dk$$

En passant à la variable $l = kb(n\eta)^{\frac{1}{4}}$

$$I = b^{-1}(n\eta)^{-\frac{1}{2}} \int_0^{\pi/h} l^5 \exp(-dl^4) dl < \frac{D}{\sqrt{ath}} \quad (4.41)$$

La différence entre les fonctions de Green du schéma et celle de l'équation équivalente est $O(t^{-\frac{1}{2}}b^{-\frac{1}{2}})$, mais, comme la fonction de Green de l'équation équivalente (nous renvoyons à la section consacrée aux schémas d'ordre 3 pour son expression explicite) est maintenant $O(t^{-\frac{1}{4}}b^{-\frac{3}{4}})$, leur différence relative est $O(t^{-\frac{1}{4}}b^{-\frac{1}{4}})$. On obtient donc aussi une convergence pour de grandes valeurs de n , mais seulement en $n^{-1/4}$.

Hedström [11] a obtenu, pour des conditions initiales marche descendante, des estimations de la différence entre la solution de l'équation équivalente et celle du schéma, en utilisant la méthode du col pour des schémas dissipatifs d'ordre supérieur. Le résultat principal est, pour un schéma dissipatif dont le premier terme de l'équation équivalente est un terme de dispersion en ik^r , une majoration de la forme

$$|u_S - u_{EE}| \leq Cb^{-\frac{2}{r}}, |x - at| \leq atn^{\frac{r-1}{r}}$$

Une application de la méthode de [11] au cas de la fonction de Green donnerait des résultats plus précis que ceux obtenus plus haut, mais au prix de démonstrations un peu plus compliquées. Formulons deux observations qualitatives : la convergence est plus précise pour $|x - at|$ petit, *i.e.* au voisinage du front, que loin du front, situé, pour l'advection, en $x = at$, et elle devient moins bonne quand l'ordre du schéma augmente. Ces observations seront confirmées dans la partie Applications, notamment dans le chapitre 5.

En somme, le point essentiel est que la fonction de Green de l'équation équivalente G_{EE} converge, en valeur relative, vers celle du schéma G_S , si le schéma *et* l'équation équivalente sont dissipatifs. La raison est que la dissipation amortit les modes haute fréquence que ce soit dans le schéma ou dans l'équation équivalente. Comme l'équation équivalente rend mal compte de l'action du schéma sur ces modes, il faut que leur contribution soit petite pour que cette méthode soit valide, ce qui se produit s'ils sont amortis par la dissipation. Les simulations numériques présentées pour les schémas dissipatifs, par exemple dans les sections 5.4 et 5.5, confirment ce résultat : la fonction de Green de l'équation équivalente, avec au moins un terme de dissipation, est une bonne approximation de celle du schéma aux n grands, avec une précision meilleure au voisinage du front.

Cas des schémas non dissipatifs

Pour ces schémas, les modes haute fréquence ne sont pas amortis, si bien que l'équation équivalente ne prédit pas correctement les résultats pour des conditions initiales non régulières. Cela se produit en particulier pour les schémas centrés d'ordre 2, traités dans les sections 5.2 et 5.3.

Nous serons donc conduits à revenir à la solution exacte par Fourier, et à approcher par phase stationnaire les intégrales, comme expliqué dans la discussion sur la vitesse de groupe de la section 4.2. Les points stationnaires respectifs vérifient $\frac{d}{dk} \widehat{Q}_S(k) = \frac{i}{n}$, $\frac{d}{dk} \widehat{Q}_{EE}(k) = \frac{i}{n}$, ce qui détermine k_S et k_{EE} . Une condition heuris-

tique est que, localement, au voisinage de ces points $\widehat{Q}_S(k) \simeq \widehat{Q}_{EE}(k)$. On retrouve le critère sur les facteurs d'amplification, mais cette fois au niveau local. Ces arguments sont développés et illustrés dans les sections 5.2 et 5.3 consacrées respectivement au schéma centré semi-discrétisé et au schéma saute-mouton. On y verra que l'équation équivalente n'est pas valide globalement, mais est correcte localement, au voisinage du front.

Quand peut-on employer l'équation équivalente ?

L'équation équivalente est un outil commode, mais il est important de bien s'assurer de sa validité. En premier lieu, comme l'équation équivalente repose sur un développement en série, un critère général de validité est que le rayon de convergence R_C de la série de Taylor du logarithme du facteur d'amplification vérifie $R_C \geq \pi$. Si ce n'est pas le cas, la méthode n'est pas valide, quel que soit le nombre de termes pris en compte. Nous verrons des exemples de cette situation dans le chapitre 8 de la partie Applications consacré à l'équation de la chaleur.

Si la condition précédente est vérifiée, l'équation équivalente à nombre fini de termes est valide pour n grand pour les schémas dissipatifs, pourvu qu'elle inclue elle-même de la dissipation. Elle n'est en revanche pas valide globalement pour des schémas non dissipatifs.

Avant de conclure cette section, nous présentons un exemple d'application de la méthode, conduisant à des résultats explicites particulièrement simples.

4.3.5 Solution de l'équation équivalente à un terme pour l'équation d'advection

Elle s'obtient, sur maillage uniforme, par transformée de Fourier. Considérons d'abord les schémas d'ordre pair $2r$. Le problème aux conditions initiales pour l'équation équivalente à un terme d'un schéma d'ordre pair $2r$ s'écrit

$$\partial_t u + a \partial_x u = \frac{1}{2r+1} C_r a b^{2r} \partial_{(2r+1)x} u, \quad u(x, 0) = u^0(x) \quad (4.42)$$

La constante sans dimension C_r dépend du schéma considéré. En appliquant à (4.42) la transformée de Fourier en x , on obtient

$$\partial_t \widehat{u} + ika \widehat{u} = \frac{1}{2r+1} C_r a b^{2r} (ik)^{2r+1} \widehat{u}, \quad \widehat{u}(k, 0) = \widehat{u}^0(k)$$

équation différentielle du premier ordre dont la solution s'écrit

$$\widehat{u}(k, t) = \exp(-ikat + \frac{1}{2r+1} C_r a b^{2r} (ik)^{2r+1} t) \widehat{u}^0(k) \quad (4.43)$$

Par Fourier inverse, la solution de notre problème est

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(ik(x - at)) + \frac{1}{2r+1} C_r ath^{2r} (ik)^{2r+1} \widehat{u}^0(k) dk \quad (4.44)$$

L'intégrale (4.44) n'est pas absolument convergente, mais bien définie au sens des intégrales oscillantes. Plus généralement, si sur le contour d'intégration, k a une petite partie imaginaire ε , l'intégrale devient alors absolument convergente et vaut I_ε . L'intégrale sur \mathbb{R} est définie comme la limite de I_ε pour $\varepsilon \rightarrow 0$.

Un des intérêts de la méthode est que toute l'information sur le schéma dans l'approximation considérée est contenue dans la constante C_r . Elle est adimensionnée, car ath^{2r} est en m^{2r+1} , et $u_{(2r+1)x}$ en m^{2r+1} . Notons $L = (C_r ath^{2r})^{1/2r+1}$. L a la dimension d'une longueur, croissante en temps. Passons dans l'intégrale à la variable adimensionnée $k' = Lk$. On obtient, pour les schémas d'ordre pair avec $C_r > 0$

$$u(x, t) = \frac{1}{2\pi L} \int_{-\infty}^{+\infty} \exp\left(ik' \frac{x - at}{L} + \frac{(ik')^{2r+1}}{2r+1}\right) \widehat{u}^0(k') dk' \quad (4.45)$$

Donc tous les schémas d'un ordre donné, dans l'approximation de l'équation équivalente à un terme, sont semblables : leur solution dépend du seul facteur d'échelle $(ath^{2r})^{1/2r+1} = L (C_r)^{-1/2r+1}$. La fonction de Green du schéma, obtenue pour $\widehat{u}^0(k') = 1$, s'écrit

$$G(x, t) = \frac{1}{2\pi L} \int_{-\infty}^{+\infty} \exp\left(ik' \frac{x - at}{L} + \frac{(ik')^{2r+1}}{2r+1}\right) dk' \quad (4.46)$$

Pour $r = 1$, *i.e.* un schéma d'ordre 2, $L = (Cath^2)^{1/3}$ et (4.46) se ramène à l'intégrale définissant la fonction d'Airy

$$G(x, t) = \frac{1}{L} \int_{-\infty}^{+\infty} \exp\left(ik' \frac{x - at}{L} + \frac{1}{3}(ik')^3\right) dk' = \frac{1}{L} Ai\left(-\frac{x - at}{L}\right) \quad (4.47)$$

Cette fonction spéciale apparaît dans plusieurs domaines de la physique et a été bien étudiée. Les fonctions de Green pour r quelconque s'expriment aussi à l'aide de fonctions d'Airy généralisées. Leurs propriétés sont présentées dans le chapitre 11 de l'Appendice.

Pour une condition initiale marche montante, où $\widehat{u}_0 = A/ik$, où A est la hauteur de la marche, on obtient, puisque $\frac{dk}{k} = \frac{dk'}{k'}$

$$u(x, t) = A \int_{-\infty}^{+\infty} \exp\left(ik' \frac{x - at}{L} + \frac{1}{2r+1}(ik')^{2r+1}\right) \frac{dk'}{ik'}$$

La solution est autosemblable : elle ne dépend que de la variable $\frac{x-at}{L}$. Des exemples d'applications à des schémas d'ordre 2 (resp. 4) sont détaillés en sections 5.4-5 (resp. 5.7). Pour les schémas d'ordre impair $2r - 1$, l'équation équivalente à un terme du schéma s'écrit

$$\partial_t u + a \partial_x u = -\frac{1}{2r} C_r a b^{2r-1} \partial_{2rx} u$$

donc la solution du problème aux conditions initiales est maintenant

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(ik(x - at) - i^{2r} C_r a b^{2r-1} \frac{k^{2r}}{2r}\right) \widehat{u}^0(k) dk$$

et la fonction de Green

$$G(x, t) = \frac{1}{\pi (C_r a b^{2r-1})^{1/2r}} \int_0^{+\infty} \cos\left(k' \frac{x - at}{(|C_r| a b^{2r-1})^{1/2r}} - \frac{k'^{2r}}{2r}\right) dk' \quad (4.48)$$

L'intégrale obtenue est convergente si $i^{2r} C_r > 0$, soit en particulier $C_1 < 0$ (terme de diffusion positif au second membre, comme vu pour le schéma upwind) et $C_2 < 0$ (terme de dissipation d'ordre 4 négatif pour un schéma d'ordre 3). On obtient, exactement comme pour les schémas d'ordre pair, des solutions dépendant du seul facteur d'échelle $(a b^{2r-1})^{1/2r}$, en particulier des solutions autosemblables pour les marches. Nous les présentons plus en détail dans la partie Applications, pour les schémas d'ordre 1 (section 5.1) et 3 (section 5.6). À partir de l'ordre 3, les fonctions de Green s'expriment à l'aide de fonctions spéciales, généralisations de la fonction d'Airy.

En conclusion, dans l'approximation de l'équation équivalente à un terme, les solutions obtenues ont un caractère universel : les résultats sont formulés non pour un schéma spécifique, mais pour une classe de schémas, en fonction d'un coefficient caractérisant le schéma. C'est un des attraits de la méthode. L'équation équivalente à un terme n'est toutefois pas toujours une bonne approximation de la solution du schéma. Comme expliqué dans le paragraphe précédent, il faut pour cela qu'elle intègre un terme de dissipation. C'est le cas pour les schémas d'ordre impair, mais pas pour les schémas d'ordre pair. Il faut pour ces derniers, pour obtenir une solution proche de celle du schéma, intégrer dans l'équation équivalente le second terme, qui est dissipatif.

4.4 Conclusion

Nous avons présenté les méthodes d'analyse des schémas les plus courantes et les plus simples (mais aussi souvent les plus efficaces) : méthodes opératorielles, analyse de Fourier, méthode de l'équation équivalente. Ces méthodes permettent de

comprendre précisément le comportement des schémas pour des équations aux dérivées partielles linéaires, sur des maillages réguliers. Il est utile pour le concepteur et l'utilisateur de schémas numériques de connaître l'ensemble de ces méthodes, leurs interrelations, et leur domaine de validité. Cette partie est essentiellement théorique, même si nous avons donné quelques exemples simples pour illustrer notre propos.

Nous passons maintenant à la partie de l'ouvrage consacrée aux applications. Son but est de présenter un ensemble représentatif d'équations et de schémas, sans aucune prétention à l'exhaustivité. Nous nous sommes limités à trois équations linéaires : l'équation d'advection, l'équation des ondes et l'équation de la chaleur. La généralisation à d'autres EDP linéaires est assez directe. Nous n'abordons pas les EDP non linéaires, nous bornant à signaler que la méthode de l'équation équivalente permet souvent d'analyser au moins qualitativement les schémas pour ces équations, comme illustré dans [2, 21]. Avant de commencer cette seconde partie, nous présentons ci-dessous les références de la première partie.

Le livre [5] est dédié aux équations aux différences partielles, particulièrement leur résolution par méthodes opératorielles. Plusieurs ouvrages présentent l'analyse de Fourier des schémas numériques : [32], dont c'est le sujet, mais aussi [29], qui y consacre un chapitre, et [25], qui l'utilise notamment pour étudier la convergence dans le cas de conditions initiales non régulières. [20] l'applique aux schémas GD d'ordre 2 pour les ondes 1D. La notion de vitesse de groupe d'un schéma est expliquée et illustrée dans [12, 30]. Le livre [24] contient une présentation détaillée et nombre d'applications, à des EDP linéaires mais aussi non linéaires de la méthode de l'équation équivalente. Elle est aussi utilisée dans [15] pour illustrer la différence de comportement entre schémas diffusifs et dispersifs, et dans [17, chapitre 5] pour comparer les prédictions de toute une palette de schémas. [33] présente la transformée en z en général, utilisée notamment par Serdyukova [22].

L'équation équivalente est aussi étudiée et appliquée dans de nombreux articles. [34] est un des articles fondateurs de la méthode de détermination par récursion de l'équation équivalente et de son application à l'étude des schémas. Nous nous sommes beaucoup appuyés dans ce livre et dans [1] sur les travaux de Chin et Hedström [7, 8, 11] où l'advection d'une marche (et notamment les oscillations parasites générées par les schémas) est étudiée avec précision. Le profil advecté est notamment caractérisé à l'aide de fonctions spéciales, primitives de celles présentées en Appendice. Les oscillations parasites ont aussi été analysées dans d'autres d'articles, dont [16, 17]. La validité de l'approche équation équivalente pour prédire la solution d'un schéma est étudiée dans [11, 27, 28, 35] et pour déterminer sa limite de stabilité dans [6, 18]. [4] montre comment obtenir l'équation équivalente par méthode opératorielle. La convergence pour des conditions initiales régulières est étudiée dans [31]. Des applications de la méthode à l'analyse des schémas sont proposées dans [19, 23]. De très nombreux autres articles utilisent l'analyse de Fourier ou l'équation équivalente pour étudier ou concevoir des schémas, dont [3, 13, 19, 23]. Parmi les nombreux ouvrages disponibles, nous renvoyons, pour les définitions des types d'EDP à [9, 14], et pour les méthodes numériques à [10, 26]. Le livre de G. Bonnaud [2] contient de nombreux résultats numériques, et aborde les équations non linéaires.

Bibliographie

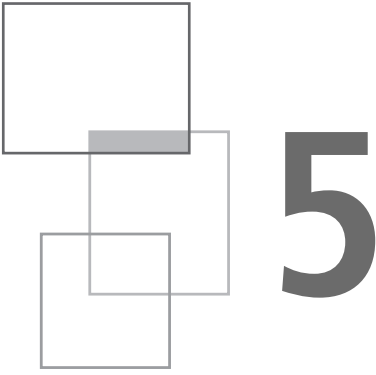
- [1] D. Bouche, G. Bonnaud, D. Ramos “Comparison of numerical schemes for solving the advection equation” *Applied Math Letters* 16, pp 147-154, 2003
- [2] G. Bonnaud « Physique de l'interaction laser plasma » *Ellipses*, 2022
- [3] E.J. Caramana “Derivation of Implicit Difference Schemes by the Method of Differential Approximation” *JCP* 96, pp 484-493, 1991
- [4] R. Carpentier, A. de la Bourdonnaye, B. Larroutourou “On the derivation of the modified equation for the analysis of linear numerical methods” *RAIRO – Modélisation mathématique et analyse numérique* 31, No. 4, pp 459-470, 1997
- [5] S.S. Cheng “Partial difference equations” Taylor and Francis, 2005
- [6] S. C. Chang “A critical analysis of the modified equation technique of Warming and Hyett” *JCP* 86, pp 117-126, 1990
- [7] R. C. Chin “Dispersion and Gibbs Phenomenon Associated with Difference Approximations to Initial Boundary-Value Problems for Hyperbolic Equations” *JCP* 18, pp 233-247, 1975
- [8] R.C. Chin, G. W. Hedström “A Dispersion Analysis for Difference Schemes : Tables of Generalized Airy Functions” *Mathematics of Computation*, Vol. 32, No. 144, pp 1163-1178, 1978
- [9] L. Evans “Partial Differential Equations” *Graduate Text in Maths* 19, AMS, 1998
- [10] E. Godlewski, P.A. Raviart “Numerical approximation of hyperbolic systems of conservation laws” troisième édition, Springer, 2021
- [11] G.W. Hedström “Models of difference schemes for $u_t + u_x = 0$ by partial differential equations” *Math. Comp.* 29, No. 132, pp 969-977, 1975
- [12] S. Karni “On the group velocity of symmetric and upwind difference schemes” *International Journal for Numerical Methods in Fluids* 18 pp 1073-1081, 1994
- [13] G. H. Klopfer “The non linear modified equation approach for analysing finite difference schemes” rapport DOE 81-1209, 1981
- [14] P. Lax “Hyperbolic Partial Differential Equations” *Courant Lecture Notes in Maths* 14, AMS, 2006
- [15] R. LeVeque “Numerical methods for conservation laws” ETH Zürich, 2012
- [16] J. Li, H. Tang “Local Oscillations in Finite difference Solutions of Hyperbolic Conservation Laws” *Math. Comp.* 78, No. 268, pp 1997-2018, 2009
- [17] J. Li, Z. Yang “Heuristic Modified Equation Analysis on Oscillations in Numerical Solutions of Conservation Laws” *SIAM J. Num. An.* 49, No. 6, pp. 2386-2406, 2011
- [18] J. Li, Z. Yang “The von Neumann analysis and modified equation approach for finite difference schemes” *Applied Mathematics and Computation* 225, pp 610-621, 2013
- [19] E. Monomiat “A Modified Equation Approach to Selecting a Nonstandard Finite Difference Scheme Applied to the Regularized Long Wave Equation” *Abstract and Applied Analysis*, 2014
- [20] A. Marica, E. Zuazua “Symmetric DG Methods for 1D Waves, Fourier Analysis, Propagation and Observability” *Springer Briefs*, 2014
- [21] A. Porubov, G. Bonnaud, D. Bouche “Description of numerical shock profiles of non-linear Burgers' equation by asymptotic solution of its differential approximations” *International Journal on Finite Volumes* 5, pp 1-16, 2008
- [22] S. I. Serdyukova “Asymptotic properties of difference schemes with maximal order of accuracy” *Mathematical Notes* 32, Issue 4, pp 739–745, 1982

- [23] R. A. Shapiro “Prediction of Dispersive Errors in Numerical Solution of the Euler Equations” *Nonlinear Hyperbolic Equations Theory, Computation Methods, and Applications*, 1989
- [24] Y. Shokin “The Method of Differential Approximation” Springer, 1983
- [25] J. Strikwerda “Finite difference schemes and partial differential equations” SIAM, 2004
- [26] E. Süli, D. Mayers “An introduction to numerical analysis” CUP, 2003
- [27] Z.H. Teng “Modified Equation for Adaptive Monotone Difference Schemes and Its Convergence Analysis” *Mathematics of Computation* 77, No. 263, pp 1453-1465, 2008
- [28] Z.H. Teng “Error bound between monotone difference schemes and their modified equation” *Math. Comp.* 79, No. 271, pp 1473–1491, 2010
- [29] L. N. Trefethen “Finite difference and spectral methods for ordinary and partial differential equations” livre en ligne sur le site de l’université d’Oxford
- [30] L. N. Trefethen “Group velocity in finite difference schemes” *SIAM Review* 24, No. 2, pp 113-136, 1982
- [31] A. I. Urusov, Yu. I. Shokin & N. N. Yanenko “Differential approximation as a qualitative test for difference schemes” *Journal of Applied Mechanics and Technical Physics* 21, pp 581–587, 1980
- [32] R. Vichnevetsky, J. B. Bowles “Fourier Analysis of Numerical Approximations of Hyperbolic Equations” SIAM, 1982
- [33] R. Vich “Z transform theory and applications” Kluwer, 1987
- [34] R. F. Warming, B. J. Hyett “The modified equation approach to the stability and accuracy analysis of finite-difference methods” *Journal of Computational Physics* 14, No. 2, pp 159–179, 1974
- [35] P. Wilders “The modified equation as a model of local errors in convective schemes” *Journal of Computational and Applied Mathematics* 16, pp 69-77 (1986)

Deuxième partie

Applications

Nous appliquons dans cette partie les méthodes présentées dans le chapitre 4 à des équations hyperboliques : advection (chapitre 5) et ondes 1D (chapitre 6) et 2D (chapitre 7) et à l'équation parabolique de la chaleur (chapitre 8).



Schémas pour l'équation d'advection

Nous présentons dans ce chapitre l'analyse de différentes catégories de schémas pour l'équation d'advection. L'essentiel du chapitre est dédié au cas monodimensionnel, soit à l'équation (2.1) $\partial_t u + a \partial_x u = 0$, à la vitesse $a > 0$. Nous commençons par le schéma upwind (3.4), déjà rencontré plusieurs fois dans la première partie comme exemple. En maillage uniforme, il est suffisamment simple pour que toutes les méthodes s'appliquent, ce qui permet de les illustrer.

5.1 Schéma upwind

Cette section présente donc plus en détail les résultats des différentes méthodes sur ce schéma, défini par (3.4), réécrite sous la forme :

$$u_j^n = p u_j^{n-1} + q u_{j-1}^{n-1} \quad (5.1)$$

La notation $p = 1 - \eta$ et $q = \eta$, où $\eta = \frac{a\tau}{h}$ est le paramètre de Courant, vient de l'interprétation probabiliste du schéma, que nous expliquons ci-dessous. Considérons une marche aléatoire à temps et états discrets sur un ensemble de cases $\{0, 1 \dots j \dots\}$: le marcheur reste en place avec la probabilité p , se déplace d'une case à droite avec la probabilité q . La solution du schéma est interprétée comme la probabilité $P(j, n) = u_j^n$ de se trouver dans la case j à l'instant n . En effet, $P(j, n)$ vérifie l'équation

$$P(j, n + 1) = P(j, n + 1 | j, n)P(j, n) + P(j, n + 1 | j - 1, n)P(j - 1, n)$$

car le marcheur se trouve en j au temps $n + 1$, soit s'il était en j au temps n (probabilité $P(j, n)$) et qu'il est resté en place (probabilité $P(j, n + 1 | j, n) = p$), soit s'il était en $j - 1$ au temps n (probabilité $P(j - 1, n)$) et qu'il s'est déplacé d'une case à droite (probabilité $P(j, n + 1 | j - 1, n) = q$). Au final

$$P(j, n + 1) = pP(j, n) + qP(j - 1, n)$$

$P(j, n)$ vérifie donc l'équation (5.1) définissant le schéma. Une manière équivalente de trouver ce résultat est de considérer la somme S_n de n variables aléatoires de Bernoulli $B(p, q)$ valant 0 avec la probabilité p et 1 avec la probabilité q

$$\begin{aligned} P(S_{n+1} = j) &= P(B_n = 0)P(S_n = j) + P(B_n = 1)P(S_n = j - 1) \\ &= pP(S_n = j) + qP(S_n = j - 1) \end{aligned}$$

Les deux interprétations sont équivalentes : les B_n sont les incréments de la marche aléatoire. Donc $P(j, n) = P(S_n = j)$ vérifie la même équation (5.1) aux différences que la solution du schéma, qui est donc simplement $u_j^n = P(S_n = j)$.

5.1.1 Méthode opératorielle

Cas monodimensionnel

La méthode opératorielle est particulièrement efficace car elle fournit une expression explicite simple, que nous rappelons ci-dessous, de la solution du schéma, notamment en 1D. En effet, selon (4.10) de la première partie

$$u_j^n = (pI + qS^{-1})u_j^{n-1} = (pI + qS^{-1})^n u_j^0 = \sum_{l=0}^n p^{n-l} q^l \binom{n}{l} u_{j-l}^0$$

La solution du schéma s'exprime donc à l'aide de coefficients binomiaux. En particulier, sa fonction de Green s'écrit, puisque la condition initiale est $u_j^0 = \delta_j^0$,

$$G_j^n = p^{n-j} q^j \binom{n}{j} \tag{5.2}$$

Dans le cas $q = 1$, le schéma est exact et (5.2) se réduit à $G_j^n = \delta_j^n$. La figure 5.1 compare la solution numérique au résultat (5.2) : la différence vient des erreurs d'arrondi. Elle est de l'ordre de la précision machine.

Comme noté plus haut, G_j^n est la probabilité $P(x = j, t = n)$ que le marcheur aléatoire sur \mathbb{N} , initialement en 0, soit en j au temps n . Il est donc resté en place

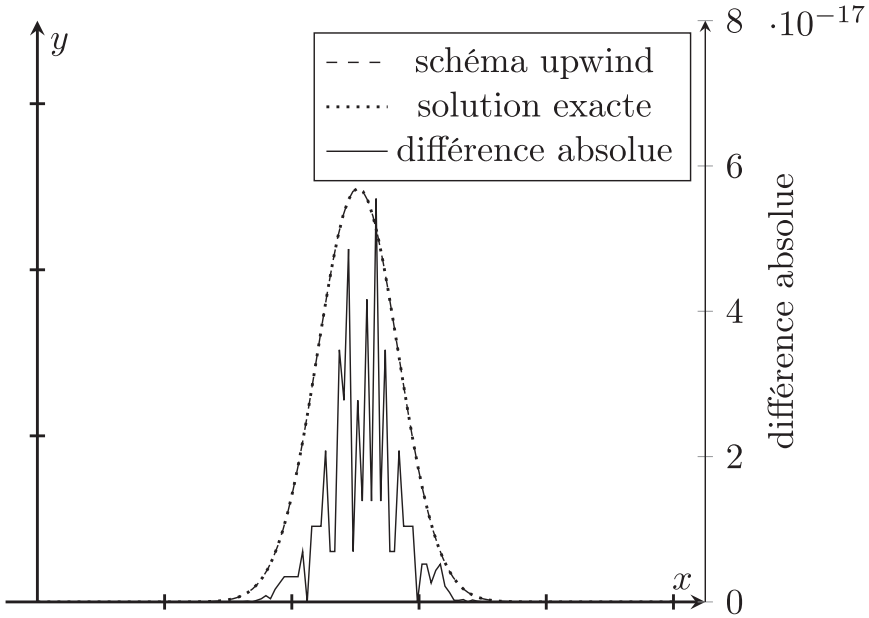


Fig. 5.1 Fonction de Green du schéma upwind, solution numérique et solution exacte.

$n - j$ fois et a avancé j fois. La probabilité d'un de ces événements est $p^{n-j} q^j$ et il y a $\binom{n}{j}$ façons de choisir les j instants parmi les n , d'où (5.2).

Revenons au résultat (5.2) : on constate que G_j^n s'étale sur $n + 1$ points : $\{0, 1, \dots, j, \dots, n\}$.

Pour la condition initiale marche descendante, $u_j^0 = 1$, pour $j \leq 0$, $u_j^0 = 0$, pour $j > 0$, donc $u_{j-l}^0 = 1$ pour $j \leq l$ et $u_{j-l}^0 = 0$ pour $j > l$, d'où

$$u_j^n = \sum_{l=j}^n p^{n-l} q^l \binom{n}{l}, \quad \text{pour } 0 < j \leq n$$

Pour $j \leq 0$, $u_{j-l}^0 = 1 \forall l$, $u_j^0 = \sum_{l=0}^n p^{n-l} q^l \binom{n}{l} = (p + q)^n = 1$

Pour $j > n$, $u_{j-l}^0 = 0$ pour $l \leq n$, donc tous les termes de la somme sont nuls et $u_j^n = 0$. Le schéma étale le front sur $n + 1$ points.

Le cas périodique sur l'intervalle $[0, J - 1]$ est formellement analogue, à ceci près que S est une matrice. Le schéma s'écrit

$$u^n = (pI + qS^{-1})u^{n-1} = (pI + qS^{-1})^n u^0 = \sum_{l=0}^n p^{n-l} q^l \binom{n}{l} S^{-l}$$

Dans ce cas, comme $S^{-J} = I$, la somme de termes peut être réécrite (ou repliée), pour $n \geq J$, sur les indices de 0 à $J - 1$. Le coefficient du terme d'indice 0 sera par exemple la somme des termes $p^n q^0 \binom{n}{0}$ venant de $l = 0$ et $p^{n-J} q^J \binom{n}{J}$ venant de $l = J$.

Cas bidimensionnel, maillage uniforme

Ce paragraphe est une courte incursion dans le cas à 2 dimensions, essentiellement traité dans le chapitre consacré à l'équation des ondes. Nous notons S_j et S_l les opérateurs S pour les indices j et l . La solution u_{jl}^n du schéma peut s'écrire à l'aide des puissances des opérateurs S_j et S_l .

Par exemple, le schéma upwind en 2d sur maillage carré s'écrit, notant a et b les composantes respectives de la vitesse suivant x et y , supposées positives

$$\frac{u_{jl}^{n+1} - u_{jl}^n}{\tau} + a \frac{u_{jl}^n - u_{(j-1)l}^n}{h} + b \frac{u_{jl}^n - u_{j(l-1)}^n}{h} = 0 \quad (5.3)$$

Sous forme barycentrique, (5.3) s'écrit, avec $q = \frac{a\tau}{h}$, $r = \frac{b\tau}{h}$, $p = 1 - q - r$

$$u_{jl}^n = (pI + qS_j^{-1} + rS_l^{-1})u_{jl}^{n-1}$$

si bien que la solution du schéma est

$$u_{jl}^n = (pI + qS_j^{-1} + rS_l^{-1})^n u_{jl}^0 = \sum_{0 \leq h_1, h_2 \leq n} p^{n-h_1-h_2} q^{h_1} r^{h_2} \frac{n!}{h_1! h_2! (n-h_1-h_2)!} u_{j-h_1, l-h_2}^0$$

En particulier, sa fonction de Green G_{jl}^n s'écrit, puisque la condition initiale est $u_{jl}^0 = \delta_j^0 \delta_l^0$ ce qui implique que $u_{j-h_1, l-h_2}^0 \neq 0$ que si $h_1 = j, h_2 = l$

$$G_{jl}^n = \frac{n!}{j! l! (n-j-l)!} p^{n-j-l} q^j r^l \quad (5.4)$$

Elle s'étale sur le triangle délimité par les axes et la droite $j + l \leq n$. Le résultat (5.4) a également une interprétation probabiliste en termes de marche aléatoire, mais dans

le plan, et peut enfin s'exprimer à l'aide de coefficients binomiaux

$$G_{jl}^n = \binom{n}{j+l} \binom{j+l}{j} p^{n-j-l} q^j r^l$$

5.1.2 Analyse de Fourier

Transformée de Fourier semi-discrète

Le facteur d'amplification du schéma a été déterminé plus haut (section 4.2), il s'écrit

$$\widehat{Q} = 1 - \eta + \eta e^{-ik} = p + qe^{-ik}$$

et sa fonction de Green s'écrit donc

$$G_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} (p + qe^{-ik})^n e^{ijk} dk \quad (5.5)$$

Le lien avec le résultat (5.2) précédent s'obtient en calculant cette intégrale. Développons dans (5.5) la puissance

$$G_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{l=j}^n \binom{n}{l} p^{n-l} q^l e^{i(j-l)k} dk = \frac{1}{2\pi} \sum_{l=j}^n p^{n-l} q^l \binom{n}{l} \int_{-\pi}^{\pi} e^{i(j-l)k} dk$$

Mais $\int_{-\pi}^{\pi} e^{i(j-l)k} dk = 2\pi \delta_j^l$, donc seul le terme $l = j$ est non nul, et on retrouve (5.2).

Transformée en z

Le résultat peut aussi être obtenu en utilisant la transformée en z . On obtient

$$\widehat{u}^n(z) = (p + q/z)^n \widehat{u}^0(z)$$

Pour la fonction de Green

$$G_j^n = \frac{1}{2\pi i} \int_C (p + q/z)^n z^{j-1} dz = \frac{1}{2\pi i} \int_C (pz + q)^n z^{j-n-1} dz \quad (5.6)$$

(5.6) se déduit de (5.5) par le changement de variable $z = e^{ik}$. Le résidu (issu du monôme en z^{-1}) vient du terme $\binom{n}{j} p^{n-j} q^j$: on retrouve à nouveau (5.2). La solution

pour la marche descendante s'exprime à l'aide de la transformée en z de la condition initiale

$$u_j^0 = 1, j \leq 0 \implies \widehat{u}^0(z) = \sum_{m=-\infty}^0 z^{-m} = \sum_{m=0}^{\infty} z^m = \frac{1}{1-z}$$

Pour assurer la convergence, l'intégration doit se faire sur un cercle de rayon < 1 . On obtient

$$u_j^n = \frac{1}{2\pi i} \int_C (pz + q)^n z^{j-n-1} \frac{dz}{1-z}$$

$$u_j^n = \frac{1}{2\pi i} \int_C \sum_{l=0}^n \sum_{m=0}^{\infty} \binom{n}{l} p^{n-l} q^l z^{j-l+m-1} dz$$

Les résidus viennent des termes $m = l - j$ positifs

$$u_j^n = \sum_{l=j}^n \binom{n}{l} p^{n-l} q^l$$

comme vu plus haut.

Cas périodique

Dans le cas périodique, en utilisant la transformée de Fourier discrète, on obtient, pour la fonction de Green du schéma, car $\widehat{u}_k^0 = 1, \forall k$

$$u_j^n = \frac{1}{J} \sum_{k=0}^{J-1} (\lambda_k)^n e^{\frac{2i\pi}{J} kj}$$

avec $\lambda_k = p + qe^{-\frac{2i\pi k}{J}}$

$$u_j^n = \frac{1}{J} \sum_{k=0}^{J-1} \sum_{l=0}^n \binom{n}{l} p^{n-l} q^l e^{-\frac{2i\pi kl}{J}} e^{\frac{2i\pi}{J} kj} = \frac{1}{J} \sum_{l=0}^n \binom{n}{l} p^{n-l} q^l \sum_{k=0}^{J-1} e^{\frac{2i\pi}{J} k(j-l)}$$

Si $n < J$, $\sum_{k=0}^{J-1} e^{\frac{2i\pi}{J} k(j-l)} = J\delta_j^l$

$$u_j^n = \frac{1}{J} \sum_{l=0}^n \binom{n}{l} p^{n-l} q^l J \delta_j^l = \binom{n}{j} p^{n-j} q^j$$

On retrouve donc bien le résultat obtenu par méthode opératoire. Si $n \geq J$, d'autres termes sont non nuls, pour $l = j + J, j + 2J$, ce qui conduit au repliement évoqué plus haut.

5.1.3 Solution de l'équation équivalente

L'équation équivalente à un terme du schéma est donnée par l'équation (4.32) de la première partie, qui s'écrit

$$\partial_t u + a \partial_x u = \frac{p a h}{2} \partial_{xx} u \quad (5.7)$$

Sa fonction de Green s'obtient à partir de celle de l'équation de la chaleur, en utilisant le changement de variable $(x, t) \rightarrow (x - at, t)$

$$G(x, t) = \frac{1}{\sqrt{2\pi p a h t}} \exp - \frac{(x - at)^2}{2 p a h t} \quad (5.8)$$

Lien avec l'interprétation probabiliste du schéma

Ce résultat coïncide avec l'approximation gaussienne de la loi binomiale. Cette approximation peut s'obtenir en approchant dans (5.2) les factorielles par la formule de Stirling, puis en effectuant un développement de Taylor. Sa précision est meilleure près du maximum $j = nq$ (voir un peu plus loin pour le détail des calculs).

$$G_j^n = p^{n-j} q^j \binom{n}{j} \simeq \frac{1}{\sqrt{2\pi n p q}} \exp - \frac{(j - nq)^2}{2 n p q}$$

en utilisant $nq = na\tau/h = at/h, j = x/h$

$$\frac{1}{\sqrt{2\pi n p q}} \exp - \frac{(j - nq)^2}{2 n p q} = \frac{\sqrt{h}}{\sqrt{2\pi p a t}} \exp - \frac{(x - at)^2}{2 p a h t} = h G(x, t)$$

Le facteur h vient de la définition de G_j^n . La condition initiale est $u_0^n = 1$, l'intégrale sur la maille vaut h . La fonction de Green $G(x, t)$ est définie avec la condition initiale $u(x, 0) = \delta(x)$, et son intégrale vaut 1.

Ce résultat peut aussi être obtenu directement par le théorème central limite en utilisant l'interprétation probabiliste du schéma comme $G_j^n = P(S_n = j)$. S_n est la

somme de variables aléatoires de Bernoulli (les incréments dans la marche aléatoire) B_n , de moyenne q et d'écart type \sqrt{pq} et converge donc, pour n grand, vers une variable aléatoire normale de moyenne nq et d'écart type \sqrt{npq} . On retrouve bien le résultat de convergence obtenu précédemment.

Lien avec la représentation intégrale de u_j^n

Le résultat peut enfin être obtenu en approximant l'intégrale donnant u_j^n en deux étapes, que nous allons détailler. La première consiste à approcher l'intégrale par la méthode du col. Elle conduira à une formule explicite faisant intervenir des puissances de n et de j .

Approximation pour n grand de l'intégrale par la méthode du col

Cette méthode permet d'approcher, pour de grandes valeurs du paramètre n , des intégrales de la forme

$$I = \int_C f(k) \exp(nS(k)) dk \quad (5.9)$$

où C est un chemin dans le plan complexe. L'idée de la méthode, détaillée dans la section 12.3 de l'Appendice, est de déformer C en un contour dit de descente rapide, où $\text{Im}(S) = \text{cte}$. La partie réelle de S varie alors rapidement sur ce contour, si bien que l'essentiel de l'intégrale vient des points dits critiques, ou points col, où S est maximale, et donc $S'' \leq 0$. On obtient, dans le cas d'un seul point col, défini par $S'(k_c) = 0$, en supposant qu'il est non dégénéré, *i.e.* que $S'' < 0$

$$I \approx \frac{\sqrt{2\pi} f}{\sqrt{-nS''}} e^{nS} \quad (5.10)$$

Les valeurs de f , S et S'' sont prises au point col k_c . La simplicité de la formule est quelque peu trompeuse. Il faut notamment s'assurer que la déformation de C vers le contour de descente rapide (CDR) est licite et définir correctement la branche de la racine de S'' . Les conditions d'application peuvent être relaxées : le contour final n'est pas nécessairement le CDR, mais doit coïncider avec le CDR aux points col. L'intégrale (5.5)

$$I = \frac{1}{2\pi} \int_{-\pi}^{\pi} (p + qe^{-ik})^n e^{ijk} dk$$

est réécrite sous la forme

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(n \ln(p + qe^{-ik}) + ijk) dk \quad (5.11)$$

soit (5.9) avec

$$\ln(p + qe^{-ik}) + i\frac{j}{n}k = S, f = 1$$

Appliquons la méthode à (5.11). Le point col k_c vérifie, par définition

$$\frac{\partial S}{\partial k} = i \left(\frac{j}{n} - \frac{qe^{-ik_c}}{p + qe^{-ik_c}} \right) = 0$$

Il est donc donné par

$$k_c = i \ln \left(1 + \frac{j - nq}{q(n-j)} \right) \quad (5.12)$$

Si $j = nq$, $k_c = 0$. C'est le cas simple où le CDR coïncide avec le contour initial (axe réel) C au point col. Si $j < nq$, k_c donné par (5.12) est dans le demi-plan inférieur, et si $j > nq$, k_c est dans le demi-plan supérieur. Calculons en ce point les quantités apparaissant dans (5.10). La dérivée seconde de S est

$$S'' = -pq \frac{e^{-ik}}{(p + qe^{-ik})^2}$$

Au point col

$$S'' = \frac{j}{n} (j - n) < 0 \quad (5.13)$$

On est donc sur un maximum réel de l'exposant nS . Le chemin de descente rapide passant par k_c a une tangente horizontale (suivant l'axe des x). En ce point $\text{Im } S = 0$, le CDR vérifie donc cette condition. Calculons maintenant en k_c

$$e^{nS} = \exp \left(n \ln \left(p + q \frac{jp}{q(n-j)} \right) + ij \ln \left(\frac{pj}{q(n-j)} \right) \right) = p^{n-j} q^j \frac{n^n}{j^j (n-j)^{n-j}} \quad (5.14)$$

Au final, en reportant (5.13) et (5.14) dans (5.10), on obtient l'approximation suivante

$$G_j^n \approx \frac{\sqrt{n}}{\sqrt{2\pi} \sqrt{j(n-j)}} p^{n-j} q^j \frac{n^n}{j^j (n-j)^{n-j}}$$

qui peut aussi être obtenue en remplaçant les factorielles par leur approximation de Stirling ($n! \rightarrow n^n e^{-n} \sqrt{2\pi n}$) dans le résultat exact (5.2), comme vu plus haut. En effet,

$$G_j^n = p^{n-j} q^j \frac{n!}{(n-j)!j!} \simeq p^{n-j} q^j \frac{n^n e^{-n} \sqrt{2\pi n}}{(n-j)^{n-j} e^{-n+j} \sqrt{2\pi(n-j)} j^j e^{-j} \sqrt{2\pi j}}$$

$$G_j^n \simeq \frac{1}{\sqrt{2\pi}} \frac{\sqrt{n}}{\sqrt{j(n-j)}} p^{n-j} q^j \frac{n^n}{(n-j)^{n-j} j^j}. \quad (5.15)$$

Cela était attendu, puisque la formule de Stirling s'obtient en approchant l'intégrale définissant la fonction Γ (qui prolonge la fonction factorielle aux réels : $\Gamma(n+1) = n!$) par la méthode du col.

Approximation gaussienne

On retrouve l'approximation gaussienne (dite normale) par développement de Taylor en $j = nq$ du logarithme de $p^{n-j} q^j \frac{n^n}{(n-j)^{n-j} j^j}$, soit de

$$E(j) = n \ln n - j \ln \left(\frac{j}{q} \right) - (n-j) \ln \left(\frac{n-j}{p} \right)$$

La dérivée, qui vaut $E' = \ln \left(-\frac{1}{p} (j-n) \right) - \ln \frac{j}{q}$, est nulle en $j = nq$. La dérivée seconde est

$$E'' = \frac{d}{dj} \left(\ln \left(-\frac{1}{p} (j-n) \right) - \ln \frac{j}{q} \right) = \frac{1}{j} \frac{n}{j-n}$$

En $j = nq$, $E'' = -\frac{1}{npq}$. Comme $E(nq) = n \ln n - j \ln(n) - (n-j) \ln(n) = 0$

$$E(j) \approx -\frac{1}{2npq} (j-nq)^2$$

En reportant dans (5.15), on obtient

$$G_j^n \approx \frac{\sqrt{n}}{\sqrt{2\pi} \sqrt{nq(n-nq)}} \exp \left(-\frac{1}{2npq} (j-nq)^2 \right) = \frac{1}{\sqrt{2\pi npq}} \exp \left(-\frac{(j-nq)^2}{2npq} \right)$$

soit l'approximation normale du coefficient binomial. En remplaçant j par $\frac{x}{h}$, n par $\frac{t}{\tau}$, on obtient

$$G_j^n \approx \frac{\sqrt{h}}{\sqrt{2\pi tpa}} \exp \left(-\frac{1}{2ahpt} (x-at)^2 \right)$$

On retrouve donc bien, avec le facteur multiplicatif h , le résultat (5.8). La figure 2 montre que l'approximation de l'équation équivalente de la fonction de Green du schéma est correcte sur l'ensemble du domaine. L'accord est optimal pour $\eta = \frac{1}{2}$, comme expliqué dans le paragraphe approche probabiliste. Dans les cas $\eta = \frac{1}{4}, \frac{3}{4}$ l'EE ne rend pas compte de l'asymétrie de G_j^n discernable sur la figure 5.2.

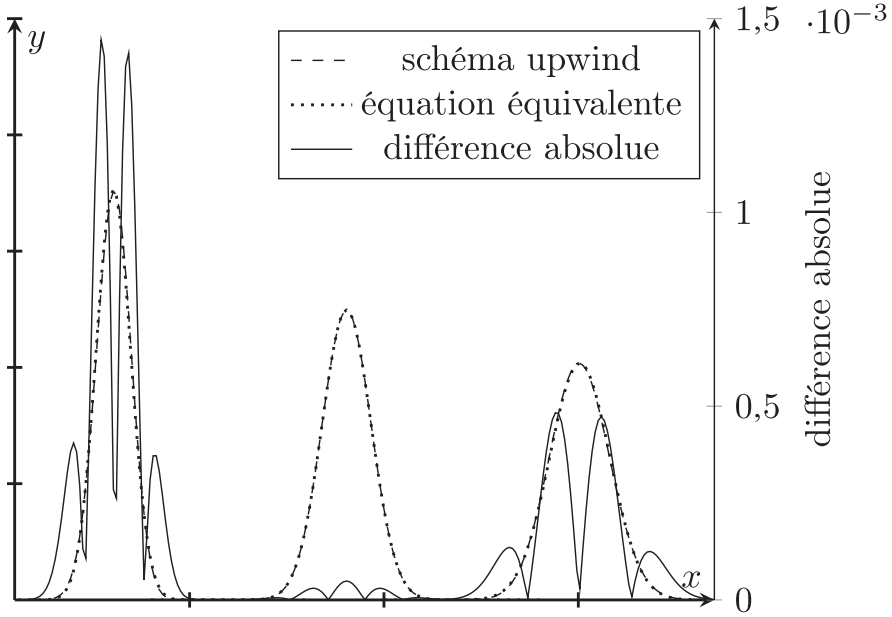


Fig. 5.2 Fonction de Green du schéma upwind et approximation de l'équation équivalente de gauche à droite, $\eta = 0,75$, $\eta = 0,50$, $\eta = 0,25$.

Considérons enfin une condition initiale marche descendante $H(-x)$ traitée par l'équation équivalente. La solution s'écrit comme une convolution de la condition initiale avec G , soit

$$u(x, t) = \int_{-\infty}^{\infty} H(-x + y) G(y, t) dy = \int_x^{\infty} G(y, t) dy = \int_x^0 G(y, t) dy + \frac{1}{2}$$

qui peut s'écrire à l'aide de la fonction d'erreur erf. En effet,

$$\int_0^x G(y, t) dy = \frac{1}{\sqrt{2\pi p a b t}} \int_0^x \exp\left(-\frac{(y - at)^2}{2 p a b t}\right) dy$$

posant $y' = \frac{y - at}{\sqrt{p a b t}}$,

$$\int_0^x G(y, t) dy = \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x - at}{\sqrt{p a b t}}} \exp\left(-\frac{y'^2}{2}\right) dy' = \text{erf}\left(\frac{x - at}{\sqrt{p a b t}}\right)$$

$$u(x, t) = \frac{1}{2} - \text{erf}\left(\frac{x - at}{\sqrt{p a b t}}\right) \quad (5.16)$$

soit, comme vu dans la section sur l'équation équivalente du chapitre sur les méthodes d'analyse, une solution autosemblable (voir (4.47) de la première partie) dépendant de la seule variable adimensionnée $\frac{x-at}{\sqrt{pakt}}$. La constante caractéristique C_1 du schéma est $C_1 = p$.

Validité de la solution équation équivalente

La solution équation équivalente est donc proche de celle du schéma, comme pour tous les schémas dissipatifs, pour n grand, tant pour la fonction de Green que pour la condition initiale marche. Les comparaisons aux résultats numériques confirment ce résultat, comme illustré sur la figure 5.2.

La raison de l'accord, comme expliqué plus haut, est la dissipation du schéma et de l'équation équivalente. Il serait possible d'améliorer encore l'accord en considérant l'équation équivalente à deux termes, qui permet de prendre en compte l'asymétrie de la distribution binomiale. Cela a été fait dans un cadre probabiliste, comme expliqué ci-dessous.

Approche probabiliste

Soit n variables aléatoires distribuées selon f , de moyenne μ et d'écart type σ . La fonction de distribution de la somme de n de ces variables est la convolution f^{*n} . Pour n grand, elle tend vers une loi normale de moyenne μ et d'écart type σ/\sqrt{n} , selon la loi des grands nombres. L'importance de ce résultat a conduit à des études détaillées de l'écart entre la fonction de distribution f^{*n} et son approximation normale. Dans le cas qui nous occupe, les variables aléatoires sont distribuées suivant la loi de Bernoulli, de moyenne p et d'écart type \sqrt{pq} , la fonction de distribution f^{*n} est la loi binomiale. Il a été démontré (principalement à l'aide des fonctions caractéristiques, donc de l'analyse de Fourier) dans [11] que, pour des variables possédant un moment d'ordre 3, l'écart relatif est en $1/\sqrt{n}$ et proportionnel au moment d'ordre 3. Celui d'une variable de Bernoulli est $pq(p^2 - q^2)$. On retrouve donc d'une part que la différence relative entre la fonction de Green du schéma et celle de l'équation équivalente à un terme, estimée en section 4.3, est en $1/\sqrt{n}$ et d'autre part l'asymétrie de la loi binomiale pour $p \neq q$, i.e. $\eta \neq \frac{1}{2}$. La dispersion du schéma, calculée plus haut, est non nulle sauf en $\eta = \frac{1}{2}$. C'est pour cette raison que, sur la figure 5.2, les deux courbes proches, mais distinctes pour $\eta = \frac{1}{4}$ et $\eta = \frac{3}{4}$, se confondent quasiment pour $\eta = \frac{1}{2}$. Plus généralement, l'écart entre la distribution binomiale, solution exacte du schéma, et ses approximations par une équation équivalente à nombre de termes déterminé peut être obtenu par méthode probabiliste. Pour plus de détails sur ces sujets, le lecteur pourra consulter le livre *An introduction to probability theory and its applications* de W. Feller [11].

5.1.4 Conclusion sur les solutions exactes et approchées

En somme, (5.2) et (5.5) sont deux représentations exactes de G_j^n : (5.2) issue de la méthode opératorielle s'exprime à l'aide de coefficients binomiaux et (5.5) issue de Fourier par une intégrale. L'intégrale redonne l'expression avec coefficients binomiaux quand on la calcule exactement. Une approximation plus maniable peut être obtenue en remplaçant dans (5.2) les factorielles par leurs approximations de Stirling, ou en approchant l'intégrale (5.5) par la méthode du col. Il y a un seul point col avec un chemin à tangente horizontale. On obtient par les 2 méthodes le même résultat (5.15), rappelé ci-dessous

$$G_j^n \approx \frac{\sqrt{n}}{\sqrt{2\pi} \sqrt{j(n-j)}} p^{n-j} q^j \frac{n^n}{j^j (n-j)^{n-j}}$$

L'approximation est valide pour n assez grand et j de l'ordre de n . En pratique, elle s'avère très précise. Une approximation supplémentaire consiste à approcher le logarithme népérien de l'exposant par un développement de Taylor à l'ordre 2 au voisinage du maximum $j = nq$, on obtient

$$G_j^n \approx \frac{1}{\sqrt{2\pi npq}} \exp\left(-\frac{(j - nq)^2}{2npq}\right)$$

Cette approximation (gaussienne) est valide pour n grand, plus précise au voisinage du maximum que sur les ailes, et coïncide avec la fonction de Green de l'équation équivalente à un terme. Elle peut aussi être obtenue grâce à l'interprétation probabiliste du schéma, à l'aide du théorème central limite. Pour une condition initiale marche $2H(x) - 1$, primitive du Dirac, on obtient comme solution de l'équation équivalente une primitive de la gaussienne, soit une fonction d'erreur.

Le schéma upwind, comme tous les schémas d'ordre 1, introduit de la diffusion. Cela se manifeste par l'étalement au cours du temps de la fonction de Green, qui est une gaussienne dans l'approximation équation équivalente, ou du profil d'une condition initiale marche, qui devient une fonction d'erreur. L'erreur en norme L^1 du schéma, calculée dans l'approximation équation équivalente est donnée dans ce dernier cas par (5.17)

$$e = 2 \int_0^\infty \left(1 - \operatorname{erf}\left(\frac{y}{\sqrt{paht}}\right)\right) dy = 2\sqrt{paht} \int_0^\infty (1 - \operatorname{erf}(x)) dx \simeq 1,128 \sqrt{paht} \quad (5.17)$$

Le schéma upwind est donc en norme L^1 d'ordre $1/2$ pour la condition initiale marche, en dessous de son ordre nominal de 1, obtenu pour des conditions initiales régulières. Il étale les discontinuités, et la largeur caractéristique de cet étalement

croît en $t^{1/2}$. Ce comportement, illustré sur l'équation d'advection, se retrouve sur les équations hyperboliques linéaires, en particulier l'équation des ondes. Le schéma n'est donc pas très adapté aux conditions initiales discontinues. Une solution possible, présentée dans la suite, est de monter en ordre. La section suivante est consacrée au schéma d'ordre 2 semi-discrétisé obtenu en centrant l'opérateur aux différences en espace. Nous étudions ensuite le schéma saute-mouton, totalement discrétisé, où l'opérateur aux différences en temps est également centré. Ces schémas, comme tous les schémas d'ordre 2, contrairement à l'upwind, n'ont pas de terme de diffusion dans leur équation équivalente. Ils n'ont pas non plus de terme de dissipation d'ordre plus élevé, *i.e.* sont non dissipatifs. Leur facteur d'amplification est de module 1.

5.2 Schéma centré semi-discrétisé

Nous analysons le schéma centré, discret en espace, mais continu en temps. En pratique, le schéma est approché en discrétisant en temps à l'aide d'une méthode de Runge-Kutta avec un petit pas de temps. Les résultats prennent une forme plus simple que pour le schéma totalement discrétisé, traité dans la section suivante.

5.2.1 Solution exacte

Le schéma s'écrit

$$\frac{\partial u_j}{\partial t} = -\frac{a}{2h}(u_{j+1} - u_{j-1}) \quad (5.18)$$

Par transformation de Fourier, (5.18) devient

$$\partial_t \widehat{u} = -\frac{ia}{h} \widehat{u} \sin k$$

et la solution du schéma s'écrit donc, dans l'espace réciproque

$$\widehat{u}(t) = \widehat{u}(0) \exp\left(-\frac{iat}{h} \sin k\right)$$

et, par transformation de Fourier inverse

$$u_j(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(-i\frac{at}{h} \sin k - ijk\right) \widehat{u}_0(k) dk \quad (5.19)$$

5.2.2 Fonction de Green

Pour $\widehat{u}_0(k) = 1$, u_j est la fonction de Green du schéma, qui s'exprime à l'aide d'une fonction de Bessel. En effet (5.19) devient

$$G_j(t) = \frac{1}{2\pi} \int_{-\pi}^0 \exp\left(-i \frac{at}{b} \sin k - ijk\right) dk + \frac{1}{2\pi} \int_0^{\pi} \exp\left(-i \frac{at}{b} \sin k - ijk\right) dk$$

comme

$$\int_{-\pi}^0 \exp\left(-i \frac{at}{b} \sin k - ijk\right) dk = \int_0^{\pi} \exp i \left(\frac{at}{b} \sin k' - ijk' \right) dk'$$

on obtient

$$G_j(t) = \frac{1}{\pi} \int_0^{\pi} \cos\left(\frac{at}{b} \sin k - jk\right) dk = J_j\left(\frac{at}{b}\right)$$

car l'intégrale est une des représentations de la fonction de Bessel J d'indice j , d'argument $\frac{at}{b}$. La transformée en z conduit à une représentation équivalente, en effet,

$$\partial_t \hat{u} = -\frac{a}{2b} \left(z - \frac{1}{z}\right) \hat{u} \implies \hat{u}(t) = \int_C \exp\left(-\frac{at}{2b} \left(z - \frac{1}{z}\right)\right) dz$$

$$G_j(t) = \frac{1}{2\pi i} \int_C \exp\left(-\frac{at}{2b} \left(z - \frac{1}{z}\right)\right) z^{j-1} dz$$

Un changement de variable $z \rightarrow 1/z$ aboutit à une autre représentation (dite d'Eckert) de J_j . La fonction de Green du schéma est donc exactement $J_j\left(\frac{at}{b}\right)$. La conservativité du schéma (conservation de la masse) se retrouve dans l'identité

$$\sum_{j=-\infty}^{+\infty} J_j = 1.$$

Évolution temporelle

Décrivons qualitativement l'évolution en temps de la solution à j fixé à partir des propriétés des fonctions $J_j(z)$. Elles sont en z^j au voisinage de $z = 0$, croissent jusqu'à z de l'ordre de j , puis décroissent en oscillant. Le développement [1, chapitre 11] de Watson $J_j(j + yj^{1/3}) \approx \left(\frac{2}{j}\right)^{1/3} Ai(-2^{1/3}y)$ rend qualitativement compte de l'ensemble de ces comportements, bien qu'il soit valide en principe seulement pour $j \rightarrow \infty$. La fonction d'Airy atteint son maximum pour $-2^{1/3}y = z_1$, premier zéro de sa dérivée. Le maximum au point j est donc atteint pour

$$j + yj^{1/3} \approx \frac{at}{b}, \quad \frac{at}{b} \approx j - 2^{-1/3}j^{1/3}z_1$$

En première approximation, $t \approx \frac{x}{a}$, ce qui correspond au déplacement de la condition initiale à la vitesse a . La figure 5.3 présente la solution exacte et le résultat numérique pour quatre valeurs de j . Elles sont comme attendu égales.

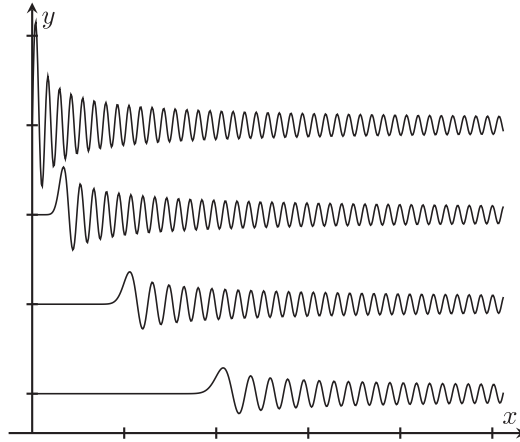


Fig. 5.3 Solution exacte à j fixé en fonction de t .

Variation spatiale de la solution

La solution exacte est, comme nous venons de le voir, $J_j(\frac{at}{h})$. La figure 5.4 présente cette solution, maintenant à temps fixé, en fonction de j , et la solution numérique. Les deux résultats sont superposés, puisqu'égaux à la précision machine.

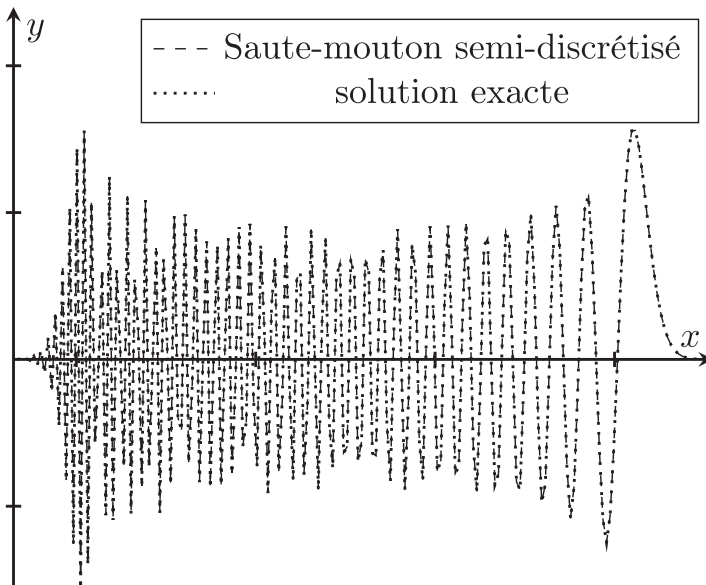


Fig. 5.4 Solution exacte et simulation numérique à t fixé.

Nous avons calculé la vitesse de groupe du schéma dans la section consacrée à la visualisation dans l'espace réciproque, et rappelons brièvement les résultats. Partons de la représentation intégrale

$$G_j(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp i \left(-\frac{at}{h} \sin k + jk \right) dk$$

et posons $x = jh$, $k = k'h$

$$G(x, t) = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} \exp i (-at \sin(k'h) + xk') dk'$$

La vitesse de phase est donc $a \sin k'h$, et la vitesse de groupe $a \cos k'h$. Comme au point stationnaire, $\cos k'h = \frac{x}{at}$, le nombre d'onde spatial local est $\frac{1}{h} \arccos \frac{x}{at}$. Il croît en s'éloignant du front de 0 pour $x = at$ à $\frac{\pi}{2h}$ pour $x = -at$. En particulier, pour les modes stationnaires, en $x = 0$, $k' = \frac{1}{h} \arccos 0 = \frac{\pi}{2h}$, soit une longueur d'onde de 4 mailles, et, pour les modes rétropropagatifs en $x = -at$, $k' = \frac{1}{h} \arccos(-1) = \frac{\pi}{h}$, soit une longueur d'onde de 2 mailles.

Nous allons maintenant approcher la fonction de Bessel par ses développements asymptotiques de Debye et de Watson, donnés notamment dans [1, chapitre 11].

Développement de Debye

Ce développement de $J_j(\frac{at}{h})$, pour l'indice j et l'argument $\frac{at}{h}$ tendant vers l'infini, est valide quand leur différence est $O(j)$. Il s'écrit, en posant

$$\cos \beta = \frac{jh}{at}, \tan \beta = \frac{at}{jh} \sqrt{1 - \left(\frac{jh}{at}\right)^2}$$

$$J_j\left(\frac{j}{\cos \beta}\right) \approx \sqrt{\frac{2}{\pi j \tan \beta}} \cos\left(j \tan \beta - j\beta - \frac{\pi}{4}\right)$$

ce qui conduit à l'approximation

$$J_j\left(\frac{at}{h}\right) \approx \sqrt{\frac{2h}{\pi at}} \left(1 - \left(\frac{x}{at}\right)^2\right)^{-1/4} \cos\left[\frac{at}{h} \sqrt{1 - \left(\frac{x}{at}\right)^2} - j \arccos\left(\frac{x}{at}\right) - \frac{\pi}{4}\right]$$

En posant $r = \frac{x}{at}$, on obtient

$$G_j \approx \sqrt{\frac{2h}{\pi at}} (1 - r^2)^{-1/4} \cos\left[\frac{at}{h} \sqrt{1 - r^2} - j \arccos r - \frac{\pi}{4}\right] \quad (5.20)$$

Le résultat (5.20) s'obtient aussi en appliquant la méthode de la phase stationnaire à la représentation intégrale de G_j . Comme vu dans le paragraphe précédent, le nombre d'onde spatial des oscillations de la solution croît quand on s'éloigne du front. Le développement de Debye est valide hors des voisinages de $x = at$ et $x = -at$, où apparaît une divergence en $(1 - (\frac{jb}{at})^2)^{\frac{1}{4}}$. La formule (5.20) garde sa validité sauf tout près de ces points. La figure 5.5 présente la solution numérique et le développement de Debye (en excluant une petite zone près du front). Elle montre que les oscillations de la solution sont bien rendues par (5.20).

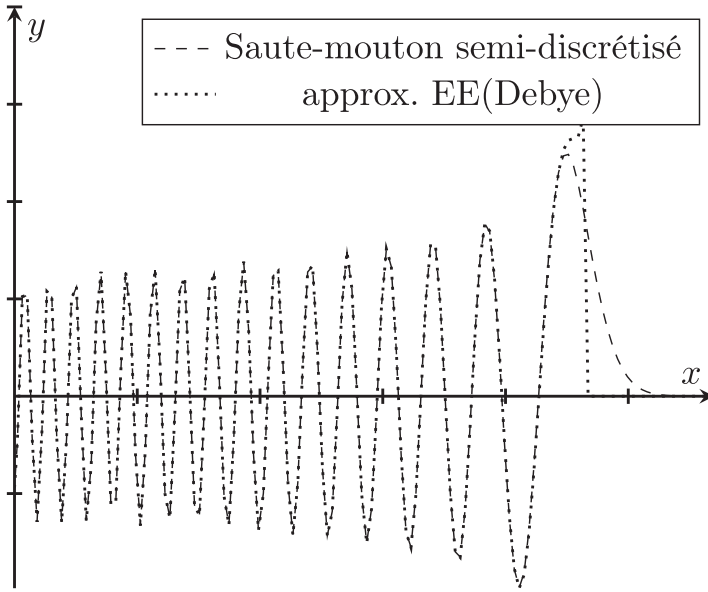


Fig. 5.5 Solution exacte et approximation par la phase stationnaire, développement de Debye (5.20).

Au voisinage du front, il faut utiliser le développement de Watson de la fonction de Bessel, valide quand la différence entre indice j et argument $\frac{at}{b}$ est $O(j^{1/3})$. Plus précisément, le profil de la solution au voisinage du front est approché par

$$u_j \approx \left(\frac{2}{j}\right)^{1/3} Ai\left(-\left(\frac{2}{j}\right)^{1/3} \left(\frac{at}{b} - j\right)\right) \quad (5.21)$$

Le résultat (5.21) peut aussi être obtenu en développant la phase dans la représentation intégrale, ce qui donne l'approximation équation équivalente, comme vu plus haut (section 4.3). On retrouve notamment la décroissance monotone pour $j > \frac{at}{b}$, et les oscillations de la solution au voisinage du front pour $j < \frac{at}{b}$. La figure 5.6 compare la solution exacte à son approximation équation équivalente. Elle est correcte

au voisinage du front, mais ne rend pas compte des oscillations de haute fréquence spatiale loin du front.

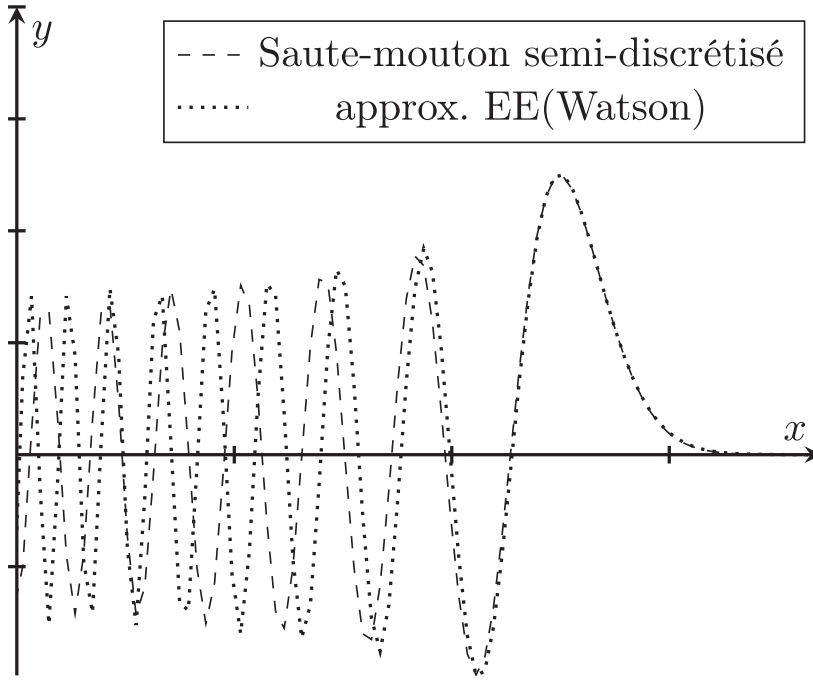


Fig. 5.6 Solution exacte et approximation équation équivalente (Watson).

Conclusion

En conclusion, la fonction de Green du schéma centré semi-discrétisé, qui s'exprime à l'aide de fonctions de Bessel J_j , permet d'illustrer le comportement des schémas d'ordre 2 non dissipatifs. Le point essentiel est que cette fonction de Green est très oscillante, avec une queue dispersive étendue. Comme on pouvait s'y attendre, puisque ce schéma est non dissipatif, l'équation équivalente, *i.e.* le développement de Watson de J_j , ne rend pas précisément compte de ces oscillations. Elles sont correctement approchées par la méthode de la phase stationnaire, *i.e.* par le développement de Debye. Une approximation uniforme de J_j , correcte partout, existe, mais conduit à des formules plus compliquées, que nous ne présentons pas. Nous étudions maintenant le schéma saute-mouton, centré en temps et en espace. C'est la version totalement discrétisée du schéma précédent.

5.3 Schéma saute-mouton

Le schéma est appliqué à l'équation d'advection $u_t + au_x = 0$ sur maillage uniforme avec un pas de temps constant, *i.e.* sur une grille régulière, $x_j = jh$, $j \in \mathbb{Z}$ en espace et $t_n = n\tau$ en temps. L'approximation numérique u_j^n de la solution satisfait l'équation aux différences partielles

$$u_j^{n+1} - u_j^{n-1} + \eta(u_{j+1}^n - u_{j-1}^n) = 0 \quad (5.22)$$

Dans (5.22), $\eta = a\tau/h$ est le paramètre de Courant. Nous prenons comme condition initiale $u_j^0 = u_j^1 = \delta_j^0$. On obtient la fonction de Green du schéma. Des résultats sur conditions initiales marche descendante unité $H(-x)$ ont été obtenus dans [1]. Dans les simulations numériques, le maillage s'étend de $-J$ à $+J$. Pour éviter des réflexions parasites, il faut que $n < J$.

5.3.1 Solution par la méthode de l'équation équivalente

C'est un cas particulier du problème traité dans la section 4.3 de la première partie sur l'équation équivalente. Nous rappelons toutefois les résultats. L'équation équivalente au schéma s'écrit

$$\partial_t u + a\partial_x u = -\frac{c}{3}\partial_{3x}^3 u$$

Le terme $\frac{c}{3}\partial_{3x}^3 u$, avec $c = h^2 a(1 - \eta^2)/2$, rend compte de la dispersion. Par transformation de Fourier de l'équation et de la condition initiale

$$\widehat{u}_t(k, t) = -ika\widehat{u} + i\frac{c}{3}k^3\widehat{u}, \quad \widehat{u}(k, 0) = \widehat{u(x, 0)}$$

$$\widehat{u}(k, t) = \widehat{u}(k, 0) \exp(-ikat + i\frac{c}{3}k^3 t)$$

et, par transformation de Fourier inverse

$$u(x, t) = u(x, 0) * G(x, t)$$

avec

$$G = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(i(x - at)k + ictk^3/3) dk = (ct)^{-1/3} Ai(X), \quad X = (x - at)(ct)^{-1/3}$$

Comme pour tous les schémas d'ordre 2 (voir (4.47) de la première partie), la fonction de Green de l'EE à un terme s'exprime à l'aide de la fonction d'Airy. La comparaison

(Figure 5.7) avec les résultats numériques montre que la fonction de Green de l'équation équivalente est proche de celle du schéma au voisinage du front de propagation, mais qu'elle ne reproduit pas correctement les oscillations de la solution numérique loin du front. C'est une illustration des limites de cette méthode en l'absence de dissipation.

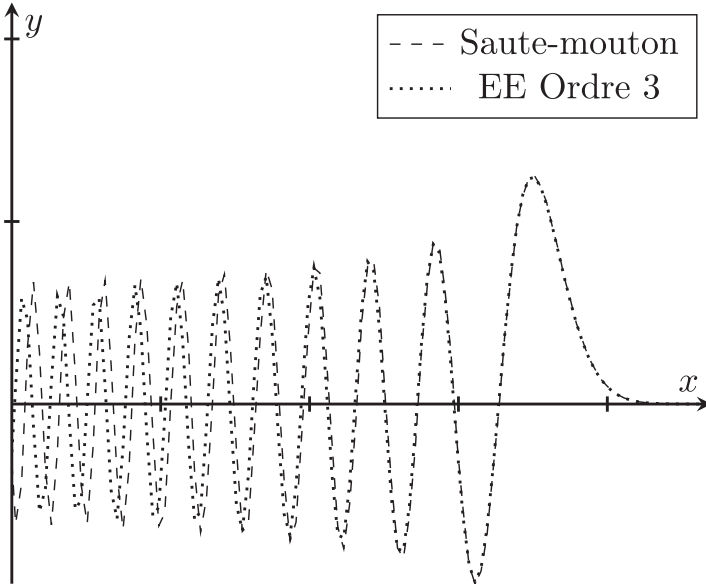


Fig. 5.7 Fonction de Green du schéma saute-mouton et de l'équation équivalente.

Dans la section suivante, nous dérivons, en suivant la même méthode que dans [1], une solution exacte pour le schéma.

5.3.2 Solution exacte pour la fonction de Green

Calcul de la solution

Nous appliquons la transformée de Fourier semi-discrète $u_j \rightarrow \widehat{u}(k)$ définie par

$$\widehat{u}^n(k) = \sum_{j=-\infty}^{j=+\infty} e^{-ijk} u_j^n$$

On obtient l'équation aux différences

$$\widehat{u}^{n+1} - \widehat{u}^{n-1} + 2i\eta\widehat{u}^n \sin k = 0$$

La solution de cette équation aux différences est une combinaison linéaire de puissances des solutions λ_{\pm} de l'équation

$$\lambda^2 + 2i\eta\lambda \sin k - 1 = 0$$

Posons $\lambda = e^{-if}$. L'équation devient $\sin(f) = \eta \sin k$, qui a pour solutions

$$f_+ = \arcsin(\eta \sin k)$$

$$f_- = \pi - \arcsin(\eta \sin k)$$

Les équations ci-dessus donnent les relations de dispersion du schéma. La relation de dispersion pour l'équation advection est $\omega = ak$. Donc $\omega_+ = \frac{f_+}{\tau} \approx \frac{1}{\tau} \arcsin\left(\frac{a\tau}{h}(k)\right) \approx ak$ pour $k \rightarrow 0$ est une approximation consistante de cette relation. La vitesse de groupe $\frac{d\omega_+}{dk}$, présentée sur la figure 5.8, vaut

$$v_+ = a \cos(k)(1 - \eta^2 \sin^2 k)^{-1/2}$$

et $v_+ \rightarrow a$ pour $kh \rightarrow 0$, comme requis. On note que, comme $\eta < 1$, $v_+ \leq a$, et que $v_+ < 0$ pour $k > \pi/2$.

Considérons maintenant la relation de dispersion de l'autre solution. Elle n'a pas le bon comportement : $\omega_- \approx \frac{\pi}{\tau} - ak$ pour $k \rightarrow 0$. Sa vitesse de groupe $v_- = -v_+$

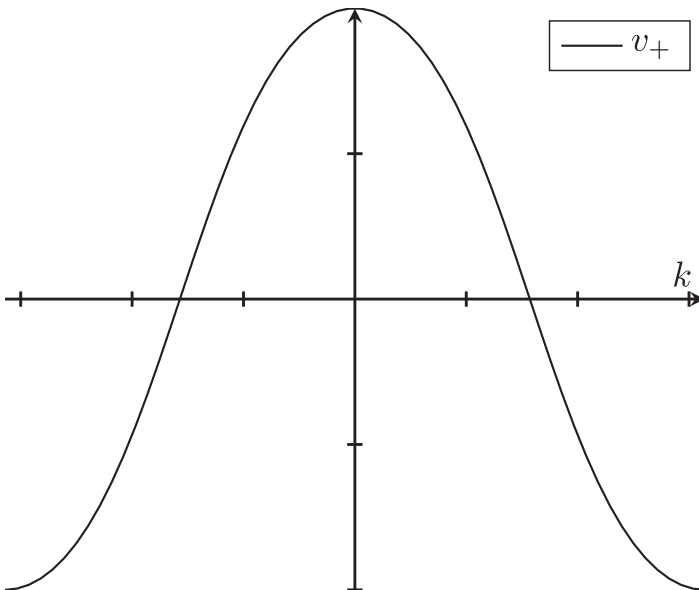


Fig. 5.8 Vitesse de groupe v_+ .

tend vers $-a$ pour $k \rightarrow 0$: le signal se propage dans la mauvaise direction. Il s'agit d'une contribution parasite introduite par le schéma. Elles sont présentes quand l'équation aux différences sur \widehat{u}^n n'est pas d'ordre 1, *i.e.* sauf pour les schémas à un pas, et ont été étudiées notamment par Strikwerda [25, première partie].

La solution dans le domaine spectral inclut les deux contributions, et s'écrit

$$\widehat{u}^n(k) = Ae^{-inf_+} + Be^{-inf_-}$$

A, B sont déterminés par les conditions initiales $\widehat{u}^0, \widehat{u}^1$: $A + B = \widehat{u}^0, Ae^{-if_+} + Be^{-if_-} = \widehat{u}^1$. On obtient donc

$$A = \frac{\widehat{u}^0 e^{-if_-} - \widehat{u}^1}{e^{-if_-} - e^{-if_+}}, B = \frac{\widehat{u}^1 - \widehat{u}^0 e^{-if_+}}{e^{-if_-} - e^{-if_+}}$$

Pour approcher un problème avec conditions initiales pour la fonction, et zéro pour sa dérivée temporelle, nous imposons, comme le schéma est à deux pas, cette condition pour les deux premiers pas de temps. Pour la fonction de Green, $u_j^0 = u_j^1 = \delta_j^0$

$$\widehat{u}^0 = \widehat{u}^1 = 1$$

La solution s'écrit, en variable k adimensionnée, comme somme d'un mode physique (5.23) et d'un mode parasite (5.28)

$$u_j^n = u_{j+}^n + u_{j-}^n$$

Mode physique

Il a pour expression

$$u_{j+}^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-if_-} - 1}{e^{-if_-} - e^{-if_+}} \exp(-inf_+ + ijk) dk \quad (5.23)$$

comme $f_+ = \arcsin(\eta \sin k), f_- = \pi - \arcsin(\eta \sin k)$, (5.23) se réécrit

$$u_{j+}^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i \arcsin(\eta \sin k)} + 1}{2 \cos \arcsin(\eta \sin k)} \exp(-in \arcsin(\eta \sin k) + ijk) dk \quad (5.24)$$

L'intégrale sur $[-\pi, 0]$ est donc la conjuguée de celle sur $[0, \pi]$, si bien que

$$u_{j+}^n = \frac{1}{2\pi} \operatorname{Re} \left(\int_0^{\pi} \frac{e^{i \arcsin(\eta \sin k)} + 1}{\cos \arcsin(\eta \sin k)} \exp(-in \arcsin(\eta \sin k) + ijk) dk \right)$$

et comme $\cos \arcsin(\eta \sin k) = \sqrt{1 - \eta^2 \sin^2 k} + 1 + i\eta \sin k$

$$u_{j+}^n = \frac{1}{2\pi} \operatorname{Re} \left(\int_0^\pi \frac{\sqrt{1 - \eta^2 \sin^2 k} + 1 + i\eta \sin k}{\sqrt{1 - \eta^2 \sin^2 k}} \exp(-in \arcsin(\eta \sin k) + ijk) dk \right) \quad (5.25)$$

Une formule alternative, en réel est $u_{j+}^n = I_+^1 + I_+^2$ avec

$$I_+^1 = \frac{1}{2\pi} \int_0^\pi \frac{\sqrt{1 - \eta^2 \sin^2 k} + 1}{\sqrt{1 - \eta^2 \sin^2 k}} \cos(-n \arcsin(\eta \sin k) + jk) dk$$

$$I_+^2 = -\frac{1}{2\pi} \int_0^\pi \frac{\eta \sin k}{\sqrt{1 - \eta^2 \sin^2 k}} \sin(-n \arcsin(\eta \sin k) + jk) dk$$

Remarque : la solution du schéma semi-discrétisé, calculée dans la section précédente, est $\frac{1}{2\pi} \int_{-\pi}^\pi \exp(-i\frac{at}{b} \sin k + ijk) dk$. Elle est plus simple, mais de même nature, et peut être obtenue comme limite $\eta \rightarrow 0$ de (5.24).

Remarque : en prenant $\sin k \sim k$, $\arcsin(\eta \sin k) \sim \eta k$, la solution devient $\frac{1}{2\pi} \int_{-\pi}^\pi \exp(-ina\tau k + ijk) dk \approx \frac{b}{2\pi} \int_{-\infty}^\infty e^{ik(x-at)} dk = b\delta(x-at)$, soit la fonction de Green de l'équation.

Le mode physique est représenté par u_{j+}^n , intégrale oscillante de la forme

$$u_{j+}^n = I_+ = \frac{1}{2\pi} \int_{-\pi}^\pi g(k) \exp(inS(k)) dk \quad (5.26)$$

avec

$$S(k) = -\arcsin(\eta \sin k) + \frac{j}{n}k, \quad g(k) = \frac{e^{-if_-} - 1}{e^{-if_-} - e^{-if_+}} \quad (5.27)$$

$$g(k) = \frac{1 + \sqrt{1 - \eta^2 \sin^2 k} + i\eta \sin k}{2\sqrt{1 - \eta^2 \sin^2 k}}$$

Mode parasite

La contribution (parasite) venant de ω_- est

$$u_{j-}^n = \frac{1}{2\pi} \int_{-\pi}^\pi \frac{1 - e^{-if_+}}{e^{-if_-} - e^{-if_+}} \exp(-in\pi + in \arcsin(\eta \sin k) + ijk) dk \quad (5.28)$$

L'intégrande peut s'écrire comme une fonction de k puisque

$$\frac{1 - e^{-if_+}}{e^{-if_-} - e^{-if_+}} = \frac{1 - \sqrt{1 - \eta^2 \sin^2 k} + i\eta \sin k}{-2\sqrt{1 - \eta^2 \sin^2 k}}$$

Une alternative à (5.28) en réel est $u_{j-}^n = I_-^1 + I_-^2$

$$I_-^1 = \frac{(-1)^n}{2\pi} \int_0^\pi \frac{1 - \sqrt{1 - \eta^2 \sin^2 k}}{-\sqrt{1 - \eta^2 \sin^2 k}} \cos(n \arcsin(\eta \sin k) + jk) dk$$

$$I_-^2 = \frac{(-1)^n}{2\pi} \int_0^\pi \frac{\eta \sin k}{\sqrt{1 - \eta^2 \sin^2 k}} \sin(n \arcsin(\eta \sin k) + jk) dk$$

Nous avons représenté sur les figures 5.9 et 5.10 mode physique et mode parasite, dont la somme est la solution du schéma. Le mode physique est dominant, mais le mode parasite n'est pas négligeable.

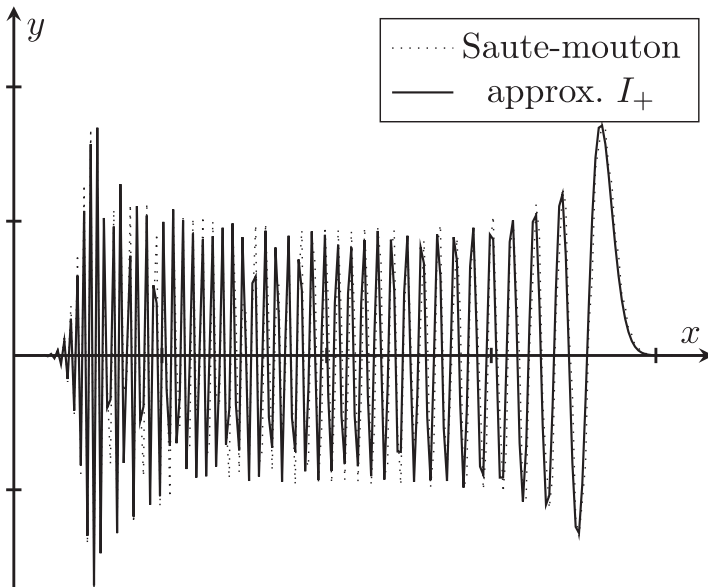


Fig. 5.9 Simulation du schéma saute-mouton et mode physique.

Dans les sections suivantes, nous dérivons des approximations de la solution exacte. Nous approcherons I_+ en utilisant la méthode de la phase stationnaire, présentée dans le chapitre 12 de l'Appendice. Cette méthode donne le développement asymptotique de cette intégrale, pour de grandes valeurs du paramètre n , comme une somme des contributions des points stationnaires de $S(k)$.

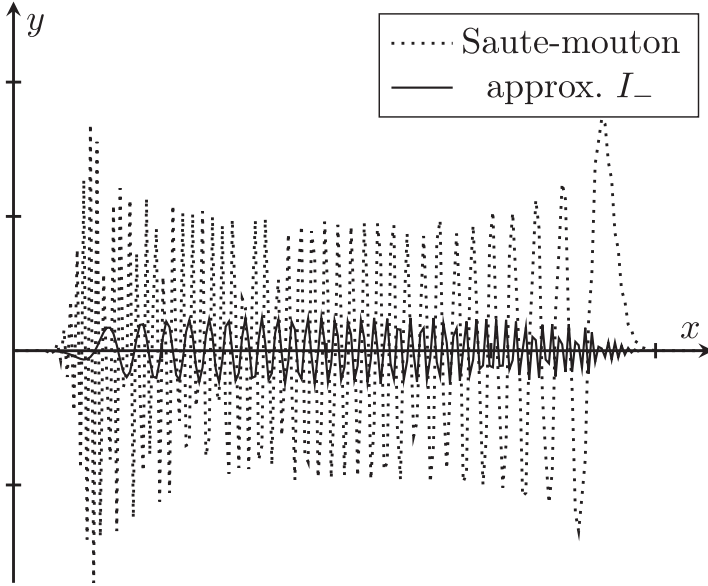


Fig. 5.10 Simulation du schéma saute-mouton et mode parasite.

5.3.3 Approximation de la solution exacte

Approximation pour I_+

Nous approximons I_+ , donnée par (5.26), par la méthode de la phase stationnaire (section 12.2 de l'Appendice) pour n grand. Pour cela, nous déterminons d'abord les points stationnaires k_c de I_+ . Ils vérifient $\frac{df_+}{dk}(k_c) = \frac{j}{n}$, soit puisque $\frac{d}{dk} \arcsin(\eta \sin k) = \eta \frac{\cos k}{\sqrt{-\eta^2 \sin^2 k + 1}}$

$$\cos(k_c)(1 - \eta^2 \sin^2(k_c))^{-1/2} = \frac{j}{n\eta} = \frac{x}{at}$$

qui implique

$$\cos^2(k_c) = \frac{r^2(1 - \eta^2)}{1 - \eta^2 r^2}$$

$r = \frac{x}{at}$ est le ratio de l'abscisse de $x = jb$ divisé par l'abscisse $at = na\tau$ de la position exacte du front. Si $\frac{r^2(1 - \eta^2)}{1 - \eta^2 r^2} \leq 1$, i.e. pour $-1 < r < 1$, i.e. $-at < x < at$, il y a deux points stationnaires opposés

$$k_c = \pm \arccos \left(r \sqrt{\frac{1 - \eta^2}{1 - \eta^2 r^2}} \right)$$

et l'approximation de la phase stationnaire de I_+ est oscillante, avec un nombre d'onde spatial local k_c , comme nous allons le voir. Pour $r > 1$, *i.e.* devant le front, et pour $r < -1$, *i.e.* $x < -at$, $r\sqrt{\frac{1-\eta^2}{1-\eta^2 r^2}} > 1$, il n'y a pas de points stationnaires réels. Notons que $k_c = 0$ sur le front ($r = 1$). Les deux points stationnaires coïncident, et la phase stationnaire n'est pas applicable.

Résultat pour I_+

La formule de la phase stationnaire (12.7) fournit une approximation explicite de I_+ comme somme des contributions des points stationnaires. Plus spécifiquement, la contribution du point $k_c = \arccos(r\sqrt{\frac{1-\eta^2}{1-\eta^2 r^2}})$ s'écrit, avec

$$S(k) = -f_+(k) + \frac{j}{n}k = -\arcsin(\eta \sin kb) + \frac{j}{n}k$$

$$I_+^1 \approx \frac{e^{+i \operatorname{sgn}(S''(k_c))\pi/4}}{\sqrt{2\pi n |S''(k_c)|}} g(k_c) \exp(inS(k_c)) \quad (5.29)$$

Calculons les différents termes apparaissant dans (5.29)

$$inS(k_c) = -in \arcsin(\eta \sin k_c) + ijk_c$$

La dérivée seconde de $S(k) = -f_+(k) + \frac{j}{n}k$ est

$$S''(k_c) = \eta(1-\eta^2)(1-\eta^2 \sin^2 k_c)^{-3/2} \sin k_c$$

$S''(k_c) > 0$, donc $\operatorname{sgn}(S''(k_c)) = 1$. Nous avons calculé plus haut

$$g(k_c) = \frac{1 + \sqrt{1 - \eta^2 \sin^2 k_c} + i\eta \sin k_c}{2\sqrt{1 - \eta^2 \sin^2 k_c}}$$

En reportant ces résultats dans (5.29), on obtient

$$I_+^1 \approx e^{+i\pi/4} \frac{(1 - \eta^2 \sin^2 k_c)^{1/4} (1 + \sqrt{1 - \eta^2 \sin^2 k_c} + i\eta \sin k_c)}{2\sqrt{2\pi n \eta (1 - \eta^2) \sin k_c}} \times \exp(-in \arcsin(\eta \sin k_c) + ijk_c) \quad (5.30)$$

La formule (5.30) n'est pas valide (et donne un résultat infini) si $k_c = 0$, ce qui se produit sur le front. C'est une expression explicite, puisque, pour $0 < r < 1$, $\sin k_c = \sqrt{\frac{1-r^2}{1-\eta^2 r^2}}$. Plus précisément, posons

$$I_+^1 \approx \frac{e^{+i\pi/4}}{2\sqrt{2\pi}} \frac{N}{D\sqrt{n}} \exp inS \quad (5.31)$$

et explicitons le résultat (5.31), en remplaçant $\sin k_c$ par sa valeur $\sqrt{\frac{1-\eta^2}{1-r^2\eta^2}}$. On obtient

$$\begin{aligned} \frac{N}{D\sqrt{n}} &= \frac{\sqrt{h}}{\sqrt{at}} (1-r^2)^{-1/4} \frac{1}{(1-\eta^2)^{1/4}} \left(\sqrt{\frac{1-\eta^2}{1-\eta^2 r^2}} + 1 + i\eta \sqrt{\frac{1-r^2}{1-\eta^2 r^2}} \right) \\ S &= - \left(\arcsin \left(\eta \sqrt{\frac{1-r^2}{1-\eta^2 r^2}} \right) - \eta r \arccos \left(r \sqrt{\frac{1-\eta^2}{1-\eta^2 r^2}} \right) \right) \end{aligned}$$

La contribution du point stationnaire $-\arccos \left(r \sqrt{\frac{1-\eta^2}{1-\eta^2 r^2}} \right)$ en est le complexe conjugué. On obtient donc au final, pour l'approximation de I_+

$$I_+ \approx 2 \operatorname{Re} I_+^1$$

Le résultat (5.31) peut s'exprimer comme le produit d'une enveloppe

$$E = \frac{\sqrt{h}}{2\sqrt{2\pi}\sqrt{at}} \frac{1}{(1-\eta^2)^{1/4}(1-r^2)^{1/4}} \sqrt{2\sqrt{\frac{1-\eta^2}{1-r^2\eta^2}} + 2}$$

modulée par une phase de la forme $\cos \left(nS + \frac{\pi}{4} + \varphi \right)$, où $\tan \varphi = \eta \frac{\sqrt{1-r^2}}{\sqrt{1-\eta^2} + \sqrt{1-\eta^2 r^2}}$

$$u \approx 2E \cos \left(nS + \frac{\pi}{4} + \varphi \right) \quad (5.32)$$

Remarque : Vérifions la cohérence avec le résultat obtenu (section précédente) pour le schéma semi-discrétisé. Pour $\eta \rightarrow 0$,

$$\begin{aligned} \frac{N}{D\sqrt{n}} &\approx 2 \frac{\sqrt{h}}{\sqrt{at}} (1-r^2)^{-1/4}, \quad nS \approx j \arccos r - \frac{at}{h} \sqrt{1-r^2} \\ I_+ &\approx \frac{2\sqrt{h}}{\sqrt{2\pi at}} (1-r^2)^{-1/4} \cos \left(j \arccos r - \frac{at}{h} \sqrt{1-r^2} + \frac{\pi}{4} \right) \end{aligned}$$

soit le résultat (5.20) pour le schéma semi-discrétisé. L'expression obtenue rend compte des oscillations de la solution numérique : le nombre d'onde spatial local est (considérant j comme une variable continue), puisque $n \frac{df_j}{dk} (k_c) = j$

$$\frac{d}{dx} nS(k_c) = -n \frac{dk_c}{dx} \frac{df_+(k_c)}{dk} + j \frac{dk_c}{dx} + \frac{dj}{dx} k_c = \frac{k_c}{h}$$

Qualitativement, les résultats sont similaires à ceux obtenus dans la section précédente. Le nombre d'onde spatial local des oscillations k_c augmente de $k_c = 0$ au front ($x = at, r = 1$) à $k_c = \pi/2$ pour $x = 0$ puis à $k_c = \pi$ en $x = -at$. Autrement dit, la longueur d'onde locale $\lambda = 2\pi h/k_c$ des oscillations décroît quand on s'éloigne du front. Par exemple, en $x = 0, k_c = \pi/2, \lambda = 4h$: la période spatiale des oscillations s'étend sur 4 mailles. En $x = -at, k_c = \pi, \lambda = 2h$: la période spatiale des oscillations s'étend sur 2 mailles. Nous comparons sur la figure 5.12 I_+ à son approximation phase stationnaire. Nous avons également tracé l'enveloppe. L'approximation est précise, sauf au voisinage immédiat du front, où elle ne peut être valide, le pôle étant alors proche du point stationnaire.

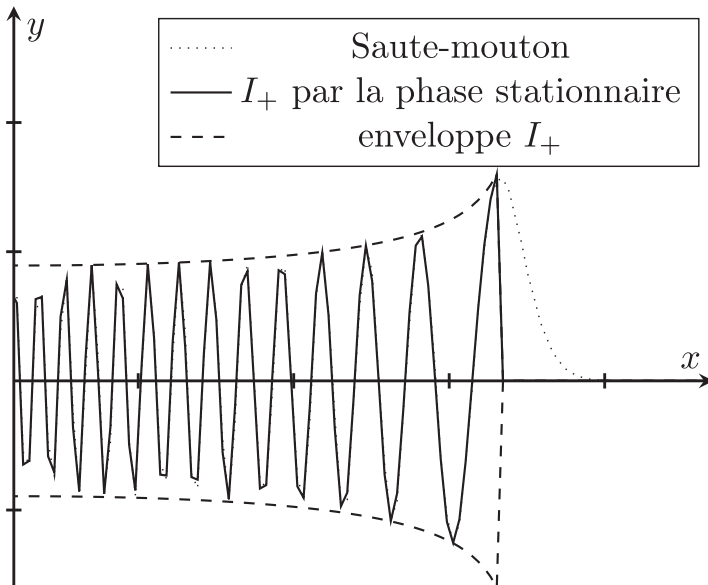


Fig. 5.11 Simulation du schéma saute-mouton et mode physique.

Résultat pour I_-

Rappelons la représentation intégrale (5.28) de I_-

$$I_- = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - e^{-if_+}}{e^{-if_-} - e^{-if_+}} \exp(-in\pi + in \arcsin(\eta \sin k) + ijk) dk$$

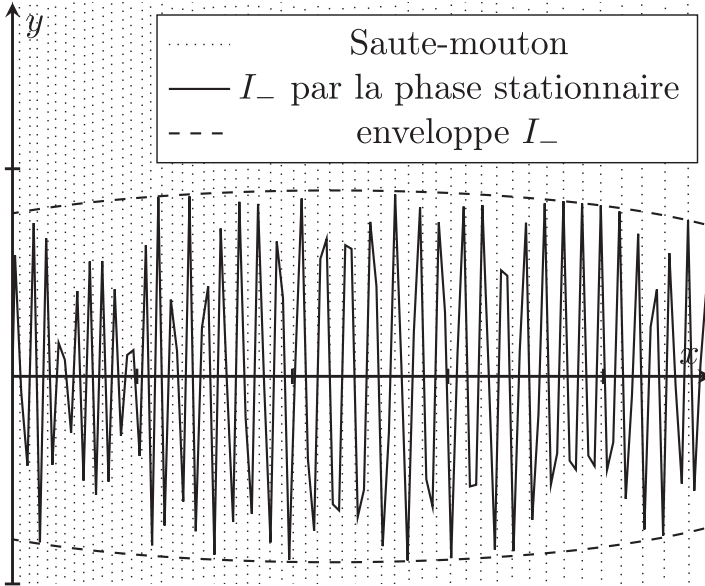


Fig. 5.12 Simulation du schéma saute-mouton et mode physique.

Les points stationnaires k'_c de I_- vérifient

$$\cos(k'_c)(1 - \eta^2 \sin^2(k'_c))^{-1/2} = -\frac{j}{n\eta} \equiv -r$$

ce qui conduit, pour $k'_c \in [0, \pi]$, à

$$k'_c = \arccos\left(-r\sqrt{\frac{1 - \eta^2}{1 - \eta^2 r^2}}\right) = \pi - k_c$$

L'autre point stationnaire est $-k'_c = k_c - \pi$. Ils sont réels pour $|x| < at$, k'_c croît de 0 à $x = -at$ à π sur le front. Calculons les fonctions de phase et d'amplitude apparaissant dans la formule (5.28). La fonction de phase en k'_c est

$$\varphi(k'_c) = -n\pi + n \arcsin(\eta \sin k_c) + j(\pi - k_c) = (j - n)\pi + n \arcsin(\eta \sin k_c) - jk_c$$

$$\frac{1 - e^{-if_+}}{e^{-if_-} - e^{-if_+}} = \frac{1 - \sqrt{1 - \eta^2 \sin^2 k} + i\eta \sin k}{-2\sqrt{1 - \eta^2 \sin^2 k}}$$

$$S''(k_c) = \eta(1 - \eta^2)(1 - \eta^2 \sin^2 k_c)^{-3/2} \sin k_c$$

$$\frac{1}{1 - e^{ik'_c}} = \frac{1}{1 + e^{ik_c}} = \frac{e^{ik_c/2}}{2 \cos(k_c/2)}$$

On obtient, pour la contribution I_-^1 du point stationnaire k'_c

$$I_-^1 \approx - \exp(i(j - n)\pi + in \arcsin(\eta \sin k_c) - ij k_c) \times \frac{e^{-i\pi/4} (1 - \eta^2 \sin^2 k_c)^{+1/4} (1 - \sqrt{(1 - \eta^2 \sin^2 k_c)} + i\eta \sin k_c)}{2\sqrt{2\pi n\eta(1 - \eta^2)} \sin k_c} \quad (5.33)$$

et, comme plus haut, $I_- \approx 2 \operatorname{Re} I_-^1$. Prenons le module de (5.33)

$$|I_-^1| \approx \frac{(1 - \eta^2 \sin^2 k_c)^{+1/4} (1 - \sqrt{(1 - \eta^2 \sin^2 k_c)} + i\eta \sin k_c)}{2\sqrt{2\pi n\eta(1 - \eta^2)} \sin k_c} \equiv E \quad (5.34)$$

Explicitons les termes apparaissant dans (5.34)

$$1 - \eta^2 \sin^2 k_c = 1 - \eta^2 \frac{1 - r^2}{1 - \eta^2 r^2} = \frac{1 - \eta^2}{1 - r^2 \eta^2} \quad (5.35)$$

$$\sqrt{\eta(1 - \eta^2) \sin k_c} = \sqrt{\eta(1 - \eta^2)} \left(\frac{1 - r^2}{1 - \eta^2 r^2} \right)^{1/4} \quad (5.36)$$

$$(1 - \sqrt{(1 - \eta^2 \sin^2 k_c)})^2 + \eta^2 \sin^2 k_c = 2 - 2\sqrt{\frac{1 - \eta^2}{1 - r^2 \eta^2}} \quad (5.37)$$

En reportant (5.35)-(5.37) dans (5.34), on obtient l'approximation d'enveloppe du mode parasite

$$E = \frac{\sqrt{h}}{2\sqrt{2\pi} \sqrt{at}} (1 - r^2)^{-1/4} \frac{1}{(1 - \eta^2)^{1/4}} \sqrt{2 - 2\sqrt{\frac{1 - \eta^2}{1 - r^2 \eta^2}}} \quad (5.38)$$

La figure 5.12 présente I_- et son approximation d'enveloppe (5.38), qui est atteinte en certains extrema, ainsi que la solution du schéma.

La contribution parasite est d'ordre $h^{1/2}$, et décroît en $t^{-1/2}$, donc assez lentement avec le temps. Elle est petite au voisinage du front.

5.3.4 Solution exacte et approximation de l'équation équivalente

Dans cette section, nous rappelons les approximations menant de la solution exacte du schéma à celle de l'équation équivalente, afin d'illustrer sur cet exemple les limites

de validité de la méthode. Nous notons comme ci-dessus.

$$f_+ = \arcsin(\eta \sin k), \quad f_- = \pi - \arcsin(\eta \sin k)$$

Partons de la formule exacte (5.23) pour u_{j+}^n ,

$$u_{j+}^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-if_-} - 1}{e^{-if_-} - e^{-if_+}} \exp(-in \arcsin(\eta \sin k) + ijk) dk$$

La solution de l'équation équivalente à un terme, comme vu dans la section 4.3 consacrée à cette méthode, s'obtient en remplaçant la phase exacte par son développement de Taylor en $k = 0$ jusqu'à l'ordre 3

$$-n \arcsin(\eta \sin k) + jk \approx \frac{(x - at)}{b} k + \frac{at}{b} (1 - \eta^2) \frac{k^3}{6}$$

et en ne gardant que le terme principal pour l'amplitude

$$\frac{e^{-if_-} - 1}{e^{-if_-} - e^{-if_+}} = \frac{\sqrt{(1 - \eta^2 \sin^2 k)} + 1 + i\eta \sin k}{2\sqrt{(1 - \eta^2 \sin^2 k)}} \approx 1$$

En reportant ces approximations dans (5.23), et en posant $c = b^2 a(1 - \eta^2)/2$, on obtient

$$u_{j+}^n \approx \frac{b}{2\pi} \int_{-\pi/b}^{\pi/b} \exp(i(x - at)l + ict \frac{l^3}{3}) dl$$

qui est la solution de l'équation équivalente présentée au début de la section, si π/b est remplacé par ∞ . Elle est donc obtenue à partir de la solution exacte en négligeant le mode parasite, en développant phase et amplitude en $k = 0$, et en prenant la limite $b \rightarrow 0$ dans les bornes de l'intégrale. Les deux premières approximations ne sont justifiées pour ce schéma qu'au voisinage du front, où le mode parasite est petit et la fréquence spatiale des oscillations de la solution faible.

5.3.5 Conclusion sur le schéma saute-mouton

Le schéma saute-mouton pour l'équation d'advection génère un certain nombre d'artefacts numériques : des oscillations, comme sa version semi-discrète, et une contribution parasite, qui lui est spécifique. Elle se propage dans la direction opposée à la vitesse d'advection. La méthode de l'équation équivalente est une approximation basse fréquence de la partie principale (non parasite) de la solution du schéma et ne l'approche correctement qu'au voisinage du front de propagation en $x = at$. Mais, loin du front, la fréquence spatiale des oscillations de la solution du schéma augmente, et l'équation équivalente ne parvient pas à les reproduire. C'est une nou-

velle illustration de la non-validité globale de cette méthode pour les schémas non dissipatifs.

5.4 Conclusion sur les schémas non dissipatifs

Nous avons étudié deux schémas de ce type : le centré semi-discrétisé et le saute-mouton. La montée à l'ordre 2 a bien permis d'éliminer les effets de diffusion, mais l'absence totale de dissipation conduit à des effets indésirables que la solution exacte de la fonction de Green du schéma centré permet de bien quantifier : oscillations faiblement amorties générées par les discontinuités, propagation sans atténuation de modes haute fréquence. De plus, comme le schéma saute-mouton fait intervenir 3 pas de temps, pour une EDP du premier ordre en temps, un mode parasite apparaît. Ces effets sont peu sensibles pour des conditions initiales régulières, mais prennent de l'importance pour des conditions initiales peu régulières, comme le Dirac, que nous venons de présenter, ou la marche, traitée dans [1]. Nous étudions dans les sections suivantes une autre classe de schémas d'ordre 2, à un pas (donc sans mode parasite), et présentant cette fois une dissipation d'ordre 4. L'ordre 2 permet toujours d'améliorer la précision par rapport au schéma upwind et notamment de ne plus être affecté par de la diffusion. Mais la dissipation d'ordre 4 réduit l'ampleur des artefacts observés sur les schémas non dissipatifs. Nous commençons par un des schémas de ce type les plus connus, celui proposé par Lax et Wendroff.

5.5 Schéma de Lax-Wendroff

Ce schéma, déjà rencontré dans la première partie (formule (3.10)) s'écrit, avec $p = \frac{a\tau}{h}$

$$u_j^{n+1} = pu_j^n + qu_{j-1}^n - \frac{1}{2}pq(u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (5.39)$$

Il s'obtient en soustrayant du schéma upwind la partie diffusion. Il est d'ordre 2, avec, comme le confirmera son équation équivalente, un terme de dissipation d'ordre 4. Comme démontré dans la section 10.1, il s'identifie au schéma de Strang S_2^1 . Pour $q = 1$, ce schéma devient $u_j^{n+1} = u_{j-1}^n$, *i.e.* il est exact. Son facteur d'amplification g se déduit de (5.39), et, comme $p = 1 - q$

$$g = \frac{q(q+1)}{2}e^{-ik} + (1-q^2) - \frac{q(1-q)}{2}e^{ik} \quad (5.40)$$

Des solutions exactes peuvent être obtenues, nous les présentons ci-dessous.

5.5.1 Solution exacte

Méthode opératorielle

L'application de la méthode opératorielle est possible, mais conduit à des sommes de coefficients multinomiaux, car le schéma est à trois points. Réécrivons-le sous forme opératorielle

$$u_j^{n+1} = (p(1+q)I + q(1 - \frac{p}{2})S^{-1} - \frac{1}{2}pqS)u_j^n$$

ce qui permet d'écrire la solution u_j^n du schéma

$$u_j^n = (p(1+q)I + q(1 - \frac{p}{2})S^{-1} - \frac{1}{2}pqS)^n u_j^0$$

$$u_j^n = \sum_{l,m} q^l (1 - \frac{p}{2})^l (-\frac{1}{2}pq)^m p^{n-l-m} (1+q)^{n-l-m} \frac{n!}{l!m!(n-l-m)!} u_{j+m-l}^0$$

Pour la fonction de Green, les termes non nuls sont pour $j + m - l = 0$, soit $j + m = l$, et on obtient

$$G_j^n = \sum_m q^{j+m} (1 - \frac{p}{2})^{j+m} (-\frac{1}{2}pq)^m p^{n-j-2m} (1+q)^{n-j-2m} \frac{n!}{(j+m)!m!(n-j-2m)!}$$

$$G_j^n = \sum_m (\frac{q}{2})^{j+2m} (-1)^m (1-q)^{n-j-m} (1+q)^{n-m} \frac{n!}{(j+m)!m!(n-j-2m)!} \quad (5.41)$$

avec $0 \leq 2m + j \leq n$, soit $0 \leq m \leq [\frac{n-j}{2}]$, la notation $[]$ désigne la partie entière. Donc, G_j^n s'exprime comme une somme de termes faisant intervenir des coefficients multinomiaux, et non avec seul coefficient binomial, comme pour le schéma upwind. Malgré cela, il est intéressant de calculer ces sommes pour comparer à la solution numérique, les différences éventuelles sont dues à des erreurs d'arrondi. Sur la figure 5.13 sont tracées la solution numérique du schéma et la solution exacte (5.41), qui se superposent.

Solution par Fourier

La formule précédente peut être aussi obtenue à partir de la solution exacte en Fourier, qui s'obtient en insérant (5.40) dans la formule (4.10) de la première partie

$$G_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{q(q+1)}{2} e^{-ik} + (1-q^2) - \frac{q(1-q)}{2} e^{ik} \right)^n e^{ijk} dk \quad (5.42)$$

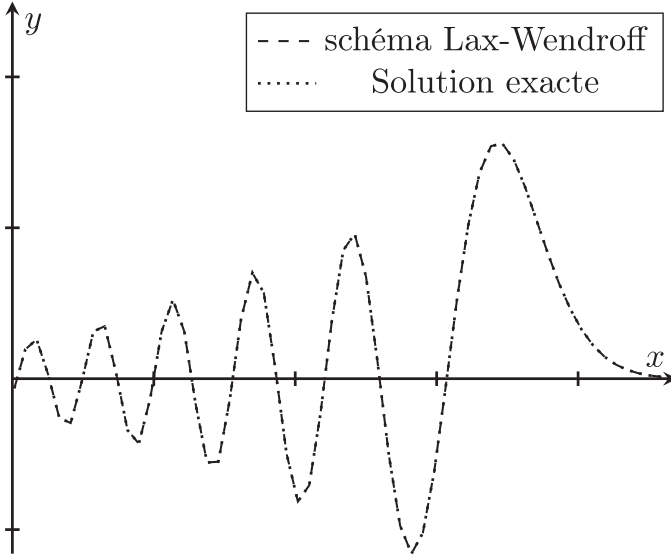


Fig. 5.13 Solution du schéma vs. solution exacte.

En développant la puissance dans (5.42), G_j^n se réécrit comme une somme de termes de la forme

$$\frac{n!}{l!m!(n-l-m)!} \left(\frac{q(q+1)}{2}\right)^l \left(\frac{q(q-1)}{2}\right)^m (1-q^2)^{n-l-m} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik(-l+m+j)} dk$$

$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik(-l+m+j)} dk = 1$, si $j+m = l$, 0 sinon, seul le terme constant $-l+m+j = 0$ contribue, si bien que

$$G_j^n = \sum_{m=0}^{\lfloor \frac{n-j}{2} \rfloor} \frac{n!}{(j+m)!m!(n-j-2m)!} (-1)^m \left(\frac{q}{2}\right)^{j+2m} (1+q)^{n-m} (1-q)^{n-j-m}$$

qui n'est autre que (5.41). En factorisant les termes indépendants de m ,

$$G_j^n = \left(\frac{q}{2(1-q)}\right)^j (1-q^2)^n \sum_{m=0}^{\lfloor \frac{n-j}{2} \rfloor} \frac{n!}{(j+m)!m!(n-j-2m)!} (-1)^m \left(\frac{q^2}{4(1-q^2)}\right)^m$$

$$G_j^n = \left(\frac{q}{2(1-q)}\right)^j (1-q^2)^n \sum_{m=0}^{\lfloor \frac{n-j}{2} \rfloor} \frac{n!}{(n-j-2m)!(2m+j)!} \frac{(2m+j)!}{(j+m)!m!} (-1)^m \left(\frac{q^2}{4(1-q^2)}\right)^m$$

ce qui permet d'écrire (5.41) avec des coefficients binomiaux

$$G_j^n = \left(\frac{q}{2(1-q)} \right)^j (1-q^2)^n \sum_{m=0}^{\lfloor \frac{n-j}{2} \rfloor} \binom{n}{2m+j} \binom{2m+j}{m} (-1)^m \left(\frac{q^2}{4(1-q^2)} \right)^m$$

En somme, les solutions exactes pour ce schéma existent et se calculent, mais il est commode de disposer de formules compactes, comme en fournit la méthode de l'équation équivalente.

5.5.2 Étude par la méthode de l'équation équivalente

La validité de la méthode peut être appréhendée en étudiant le facteur d'amplification du schéma, donné par (5.40), réécrite sous la forme

$$g = (1-q^2) + \left(\frac{q(q+1)}{2} - \frac{q(1-q)}{2} \right) \cos k - i \left(\frac{q(q+1)}{2} + \frac{q(1-q)}{2} \right) \sin k$$

mais $\frac{q(q+1)}{2} - \frac{q(1-q)}{2} = q^2$, $\frac{q(q+1)}{2} + \frac{q(1-q)}{2} = q$, et g s'écrit aussi

$$g = 1 - iq \sin k - q^2(1 - \cos k) \tag{5.43}$$

Dans le plan complexe, g parcourt l'ellipse de centre $(1 - q^2, 0)$, de demi-axe horizontal q^2 , de demi-axe vertical q . Elle est intérieure au cercle unité parcouru par le facteur d'amplification de l'équation, et lui est osculatrice (le contact est d'ordre 2) en $(0, 1)$. La figure 5.14 présente le tracé de g pour trois valeurs de η : $\eta = 0.5, 0.7, 1$. Pour $\eta = 1$, le schéma est exact et l'ellipse se confond avec le cercle unité.

Calculons le carré de son module

$$|g|^2 = gg^* = (1 - q^2 - iq \sin k + q^2 \cos k)(1 - q^2 + iq \sin k + q^2 \cos k)$$

et développons $|g|^2$

$$|g|^2 = \sqrt{1 - (q^2 - q^4)(\cos k - 1)^2} = 1 - k^4 \left(\frac{1}{8}q^2 - \frac{1}{8}q^4 \right) + O(k^5) \tag{5.44}$$

Le schéma est, selon (5.44), dissipatif d'ordre 4 si $q \neq 0, 1$. Cela implique que la méthode de l'équation équivalente est valide aux temps longs, pourvu qu'elle inclue un terme de dissipation. Sur la figure 5.15 est tracé le carré du module du facteur d'amplification : il devient notablement inférieur à 1 quand k se rapproche de π .

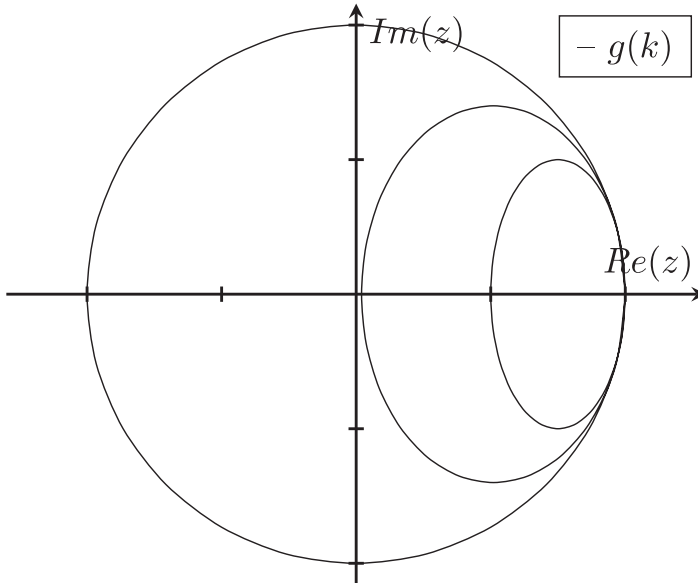


Fig. 5.14 Facteur d'amplification du schéma de LW dans le plan complexe pour $\eta = 0.5, 0.7, 1.0$.

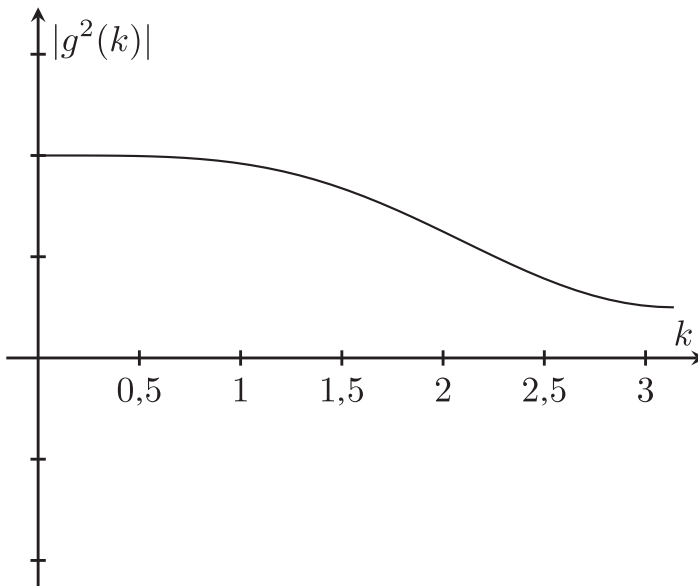


Fig. 5.15 Module au carré du facteur d'amplification du schéma de LW.

Calcul opératoire de l'équation équivalente

L'équation équivalente des schémas de Strang est donnée en appendice, mais nous la recalculons pour le schéma de Lax-Wendroff, à titre d'illustration du calcul d'une EE par méthode opératoire. Formellement, le schéma, puisque défini par (5.39), s'écrit

$$\exp(\tau \partial_t) = (1 - q) + q \exp(-h \partial_x) - pq(\cosh(h \partial_x) - 1) \quad (5.45)$$

en prenant le logarithme de (5.45), on obtient

$$\partial_t = \frac{1}{\tau} \ln(p + q \exp(-h \partial_x) - pq(\cosh(h \partial_x) - 1))$$

Calculons le développement de Taylor de cette expression

$$\begin{aligned} \ln((1 - q) + q \exp(-h \partial_x) - (1 - q)q(\cosh(h \partial_x) - 1)) \\ = -h \partial_x q + \frac{1}{6} q (q^2 - 1) h^3 \partial_{3x} + O(h^4) \end{aligned}$$

Le schéma approxime la dérivation temporelle par

$$\partial_t \approx -\frac{1}{\tau} h q \partial_x + \frac{1}{6\tau} h q (q^2 - 1) h^2 \partial_{3x}$$

ce qui conduit à l'équation équivalente à un terme

$$\partial_t u + a \partial_x u = \frac{a}{6} (q^2 - 1) h^2 \partial_{3x} u \quad (5.46)$$

Posons $c = \frac{a}{2} (1 - q^2) h^2$, l'équation équivalente (5.46) du schéma fait apparaître au second membre un terme de dispersion négatif.

$$\partial_t u + a \partial_x u = -\frac{c}{3} \partial_{3x} u$$

En développant à un ordre de plus, on obtient le deuxième terme de l'équation équivalente

$$-\frac{a}{8} q (1 - q^2) h^3 \partial_{4x}$$

C'est un terme de dissipation d'ordre 4. L'équation équivalente à 2 termes s'écrit donc

$$\partial_t u + a \partial_x u = -\frac{c}{3} \partial_x^3 u - \frac{d}{4} \partial_x^4 u, \quad c = \frac{a}{2} (1 - q^2) h^2, \quad d = \frac{a}{2} q (1 - q^2) h^3 \quad (5.47)$$

Notons que la condition de stabilité du schéma est $q < 1$. Elle est donnée par le signe du terme de dissipation, qui change pour $q = 1$. L'équation équivalente à deux termes permet donc de retrouver la limite de stabilité.

Bien que l'équation équivalente à un terme ne puisse pas être une approximation précise du schéma, puisqu'elle ne prend pas en compte la dissipation, nous rappelons la solution obtenue pour la fonction de Green, intéressante de par sa simplicité, ci-dessous.

Solution de l'équation équivalente à un terme

Le résultat est donné par la formule (4.46) de la première partie, mais nous en rappelons la dérivation. Le problème à résoudre est, comme pour le saute-mouton

$$\partial_t u + a \partial_x u = -\frac{c}{3} \partial_{3x}^3 u, \quad u(x, 0) = \delta(x)$$

En appliquant la transformation de Fourier à l'équation et aux conditions initiales, on obtient

$$\partial_t \widehat{u} = (-ika + i\frac{c}{3}k^3)\widehat{u} = 0, \quad \widehat{u}(k, 0) = 1$$

dont la solution s'écrit

$$\widehat{u} = \exp(-ikat + i\frac{c}{3}k^3 t)$$

Par transformation inverse, on obtient

$$u(x, t) = (ct)^{-1/3} Ai(y), \quad y = \frac{x - at}{(ct)^{1/3}} \quad (5.48)$$

La dimension de la solution est celle de la condition initiale Dirac : elle est en m^{-1} . L'approximation du Dirac dans le schéma consiste à prendre une condition initiale h^{-1} dans la maille 0. Le caractère conservatif du schéma est préservé par la solution de l'équation équivalente, on vérifie en effet que :

$$\int_{-\infty}^{+\infty} u(x, t) dx = \int_{-\infty}^{+\infty} Ai(y) dy = 1$$

Pour y grand, la fonction d'Airy peut être approchée par le premier terme de son développement asymptotique, comme détaillé dans le chapitre 11 de l'Appendice. En reportant ce développement dans (5.48), on obtient respectivement, pour $y > 0$ ($y < 0$), (5.49) et (5.50)

$$u(x, t) \approx \frac{1}{2\sqrt{\pi}} |y|^{-1/4} \exp\left(-\left(\frac{2}{3}|y|^{3/2} - \frac{\pi}{4}\right)\right) \quad (5.49)$$

$$u(x, t) \approx \frac{1}{\sqrt{\pi}} |y|^{-1/4} \cos\left(\frac{2}{3} |y|^{3/2} - \frac{\pi}{4}\right) \quad (5.50)$$

Les formules explicites approchées (5.49) et (5.50), valides pour $|y|$ grand, montrent respectivement une décroissance monotone de la solution pour $x > at$ et une décroissance en $|y|^{-1/4}$ des oscillations de la solution si $x < at$. Les résultats sont présentés pour la fonction de Green sur la figure 5.16. La conclusion est que l'accord est bon du côté monotone, mais n'est pas satisfaisant du côté oscillant : les oscillations sont plus amorties que ne le prévoit l'équation équivalente à 1 terme, ce qui était prévisible, puisqu'elle néglige la dissipation du schéma.

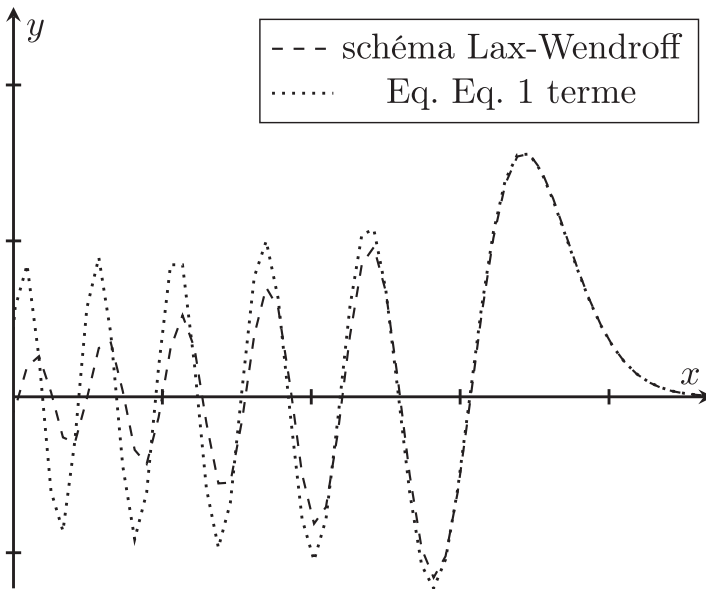


Fig. 5.16 Solution du schéma vs. équation équivalente à 1 terme.

Pour le prendre en compte, il faut intégrer cet effet dans l'équation équivalente, *i.e.* utiliser l'équation équivalente à deux termes.

Solution de l'équation équivalente à deux termes

L'équation équivalente à deux termes (5.47) s'écrit

$$\partial_t u + a \partial_x u = -\frac{c}{3} \partial_{3x}^3 u - \frac{d}{4} \partial_{4x}^4 u$$

Par transformation de Fourier sur la variable d'espace, on obtient l'équation différentielle en temps et sa condition initiale

$$\partial_t \widehat{u} = \left(-ika + i\frac{c}{3}k^3 - \frac{d}{4}k^4 \right) \widehat{u} = 0, \quad \widehat{u}(k, 0) = 1$$

dont la solution s'écrit

$$\widehat{u} = \exp \left(-ikat + i\frac{c}{3}k^3 t - \frac{d}{4}k^4 t \right)$$

Par transformation inverse, on obtient l'approximation (5.51) de la fonction de Green G_{EE2} du schéma par l'équation équivalente à deux termes

$$G_{EE2} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left(ik(x - at) + i\frac{c}{3}k^3 t - \frac{d}{4}k^4 t \right) dk \quad (5.51)$$

G_{EE2} s'exprime à l'aide d'une généralisation de la fonction d'Airy, appelée fonction d'Airy dissipative (présentée plus en détail dans le chapitre 11)

$$Ai^{3,b}(y) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left(iky + i\frac{k^3}{3} - \frac{b}{4}k^4 \right) dk$$

En posant $k' = (ct)^{1/3}k$, G_{EE2} devient

$$G_{EE2} = \frac{1}{2\pi(ct)^{1/3}} \int_{-\infty}^{+\infty} \exp \left(ik' \frac{(x - at)}{(ct)^{1/3}} + i\frac{k'^3}{3} - \frac{d}{4c(ct)^{1/3}} k'^4 \right) dk$$

$$G_{EE2} = (ct)^{-1/3} Ai^{3,b}(y), \quad y = \frac{x - at}{(ct)^{1/3}}, \quad b = \frac{d}{c(ct)^{1/3}} \quad (5.52)$$

Le coefficient b décroît en $t^{-1/3}$: le schéma reste en pratique dissipatif, même aux temps longs. Une approximation, valide sauf au voisinage du front, peut être obtenue par la méthode du col. Les points col vérifient

$$\frac{d}{dk} \left(ik(x - at) + \frac{ic}{3}k^3 t - \frac{d}{4}k^4 t \right) = 0$$

qui est une équation de degré 3

$$idk^3 t + ct k^2 + x - at = 0 \quad (5.53)$$

Ses racines peuvent être approchées par perturbation. Pour $x > at$, $k = \pm i\sqrt{\frac{(x-at)}{ct}} + \epsilon$. En reportant cette approximation dans (5.53), on obtient

$$\epsilon \simeq \frac{i}{2} q \frac{x - at}{\frac{a}{2}(1 - q^2) ht}$$

Les points col restent imaginaires purs, la solution est peu modifiée par rapport à celle obtenue dans le paragraphe précédent. Par contre, pour $x < at$, ils acquièrent une partie imaginaire, et la phase aux points col $iky + i\frac{k^3}{3} - \frac{b}{4}k^4$ a maintenant une partie réelle négative. Les oscillations de la solution de l'équation équivalente à un terme, en $|y|^{-1/4} \cos(\frac{2}{3}|y|^{3/2} - \frac{\pi}{4})$, sont donc maintenant amorties exponentiellement.

Comme on peut le voir sur la figure 5.17, l'accord aux temps longs entre la fonction de Green du schéma et celle de l'équation équivalente à 2 termes est, comme prévu, bon. L'amortissement introduit par le terme de dissipation d'ordre 4 rend bien compte de la décroissance des oscillations. L'accord est meilleur près que loin du front. L'équation équivalente, même à deux termes, perd donc en précision, comme attendu, quand on s'éloigne du front.

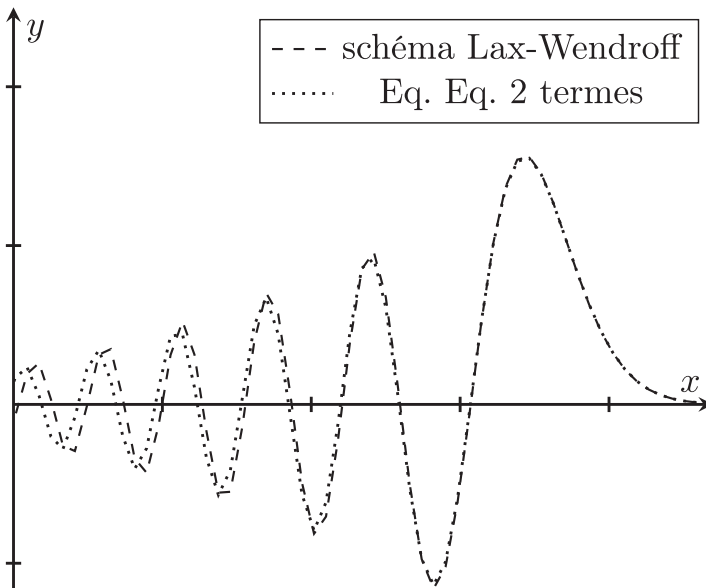


Fig. 5.17 Fonction de Green, schéma vs. EE à 2 termes.

La solution s'exprime à l'aide d'une fonction d'Airy dissipative moins connue que la fonction d'Airy usuelle. Il faut la calculer, ou l'approcher : des méthodes sont proposées dans la section 11.1 consacrée à ces fonctions.

Conclusion sur le schéma de Lax-Wendroff

Nous avons obtenu, par méthode opératoire ou par analyse de Fourier, une représentation exacte sous forme intégrale de la fonction de Green du schéma de Lax-Wendroff, et diverses approximations de cette solution.

La plus simple des approximations consiste à utiliser la méthode de l'équation équivalente à un terme. Dans la mesure où la dissipation n'est pas présente, cette méthode ne peut pas fournir une approximation précise de la fonction de Green du schéma. Elle rend toutefois compte de la décroissance monotone pour $x > at$ et de la décroissance avec oscillations pour $x < at$. La largeur caractéristique du front croît en $t^{1/3}$. On s'attend, compte tenu des approximations, à des résultats valides seulement au voisinage de $x = at$. L'accord est effectivement correct dans ce voisinage. La position et l'amplitude des premières oscillations sont bien prédites, mais l'accord se dégrade quand on s'éloigne du front.

Une approximation plus précise, convergeant vers la solution du schéma aux temps longs, consiste à prendre en compte le second terme dans l'équation équivalente. La solution est alors exprimée à l'aide d'une fonction d'Airy dissipative. Comme attendu, l'accord est amélioré, en particulier sur l'amortissement des oscillations.

D'un point de vue pratique, le schéma, grâce à sa dissipation d'ordre 4, s'affranchit d'une partie des artéfacts du schéma saute-mouton, en particulier la propagation sans atténuation de modes haute fréquence. Les oscillations générées par les discontinuités acquièrent un amortissement exponentiel. Enfin, il n'y a pas de mode parasite, puisque le schéma est à un pas. Il peut donc être utilisé même pour des solutions peu régulières, en gardant en mémoire que des artéfacts d'ampleur limitée, précisément quantifiés dans cette section pour une condition initiale $u_j^0 = \delta_j^0$, seront présents.

Analysons un autre schéma de Strang, d'ordre 2, comme le schéma de Lax-Wendroff, mais avec un stencil différent.

5.6 Le schéma de Beam-Warming

Comme expliqué dans le chapitre 10 de l'Appendice, le schéma de Strang S_n^s est construit en développant à l'ordre n l'expression

$$e^{-qb\partial_x} = (1 + (e^{-b\partial_x} - 1))^{q+s} e^{bs\partial_x}$$

Le nombre entier s permet de décaler le schéma vers la droite. Lax-Wendroff est le schéma S_2^1 . Le schéma S_2^0 est dit de Beam-Warming. Il est obtenu en développant à l'ordre 2 $(1 + (e^{-b\partial_x} - 1))^q$, soit

$$(1 + (e^{-b\partial_x} - 1))^q = \frac{1}{2}(q-1)(q-2) - q(q-2)e^{-b\partial_x} + \frac{q(q-1)}{2}e^{-2b\partial_x} + \dots$$

Ce schéma s'écrit donc, sous forme opératorielle

$$e^{\tau \partial_t} = \frac{1}{2}(q-1)(q-2) - q(q-2)e^{-b\partial_x} + \frac{q(q-1)}{2}e^{-2b\partial_x} \quad (5.54)$$

d'où son expression explicite

$$u_j^{n+1} = \frac{1}{2}(q-1)(q-2)u_j^n - q(q-2)u_{j-1}^n + \frac{1}{2}q(q-1)u_{j-2}^n$$

Son stencil est décalé d'un pas vers l'amont par rapport à celui du schéma de Lax-Wendroff. Calculons le développement du logarithme de son facteur d'amplification (5.54) écrit sous forme opératorielle

$$\begin{aligned} & \ln \left(\frac{1}{2}(q-1)(q-2) - q(q-2)e^{-b\partial_x} + \frac{q(q-1)}{2}e^{-2b\partial_x} \right) \\ &= -qb + \frac{1}{6}q(q-1)(q-2)\partial_x^3 + O(\varepsilon^4) \end{aligned} \quad (5.55)$$

(5.55) conduit à l'équation équivalente à un terme

$$\partial_t u + a\partial_x u = \frac{a}{6}(q-1)(q-2)b^2\partial_{3x}u \quad (5.56)$$

Rappelons celle (5.46) de Lax-Wendroff

$$\partial_t u + a\partial_x u = \frac{a}{6}(q^2-1)b^2\partial_{3x}u$$

Si $q < 1$, le signe du terme de dispersion, ou de la constante C_2 , est positif dans un cas, négatif dans l'autre. En effet, $C_2 = \frac{a}{6}(q^2-1) < 0$ pour LW, $C_2 = \frac{a}{6}(q-1)(q-2) > 0$ pour BW. Notons que, si $q > 1$, le terme de dispersion redevient positif. Pour $q = 1$, le schéma devient $u_j^{n+1} = u_{j-1}^n$, i.e. il est exact, comme l'upwind et le Lax-Wendroff pour cette même valeur de q . Le coefficient du terme de dissipation d'ordre 4 est obtenu en développant à un ordre de plus et vaut

$$-\frac{ab^3}{8}q(1-q)^2(2-q) \quad (5.57)$$

Il est négatif si $q < 2$, ce qui permet de retrouver la limite de stabilité du schéma. À noter que les dix premiers termes de l'équation équivalente de ce schéma ont été récemment calculés par méthode récursive [15,16].

5.6.1 Analyse du schéma

Présentons maintenant quelques résultats pour Beam-Warming. L'équation équivalente à un terme s'écrit maintenant

$$\partial_t u + a \partial_x u = \frac{c}{3} \partial_{3x} u \quad (5.58)$$

avec $c = \frac{a}{2} (q^2 - 3q + 2) h^2$. La fonction de Green se calcule exactement comme dans le cas de Lax-Wendroff. En appliquant la transformation de Fourier à (5.58) et à la condition initiale $u(x, 0) = \delta(x)$, on obtient

$$\partial_t \widehat{u} = \left(-ika - i \frac{c}{3} k^3 \right) \widehat{u} = 0, \widehat{u}(k, 0) = 1$$

dont la solution s'écrit

$$\widehat{u} = \exp \left(-ikat - i \frac{c}{3} k^3 t \right)$$

et, par transformation inverse, la fonction de Green de l'EE à un terme s'écrit

$$G_{EE1}(x, t) = (ct)^{-1/3} Ai(-y), y = \frac{x - at}{(ct)^{1/3}}$$

Considérons le cas $q < 1$. On reconnaît la fonction d'Airy, mais avec un argument de signe opposé à celui obtenu pour Lax-Wendroff. Les résultats se déduisent de ceux du schéma de Lax-Wendroff, en changeant le signe de l'argument. On obtient maintenant une croissance monotone de la solution pour $x < at$ et une décroissance avec oscillations de la solution pour $x > at$.

Le deuxième terme de l'équation équivalente est dissipatif pour $q \leq 2$. Comme pour le schéma de Lax-Wendroff, la limite de stabilité est prédite par le signe de ce terme. Le schéma a une plage de stabilité $0 \leq q \leq 2$ deux fois plus étendue que celle du Lax-Wendroff.

La figure 5.18 (resp. 5.19) compare la fonction de Green du schéma à celle de l'équation équivalente à un (resp. deux) terme(s). Comme pour le schéma de LW, il est nécessaire d'utiliser l'EE à 2 termes pour prédire la solution du schéma. La différence essentielle est que (si $q < 1$), les oscillations sont pour le schéma de BW, devant le front, alors qu'elles sont à l'arrière pour le schéma de LW. Le rapport des termes de dissipation d'ordre 4 des deux schémas vaut $\frac{2-q}{1+q}$: BW dissipe plus (resp. moins) que LW pour $q < \frac{1}{2}$ (resp. $q > \frac{1}{2}$).

Nous présentons, en vue de leur utilisation dans la section suivante, les résultats pour la marche descendante pour les 2 schémas, comparés à leur EE à un et à 2 termes. Le nombre de Courant est 1/2 dans les deux cas, si bien que les schémas ont la même dispersion et la même dissipation.

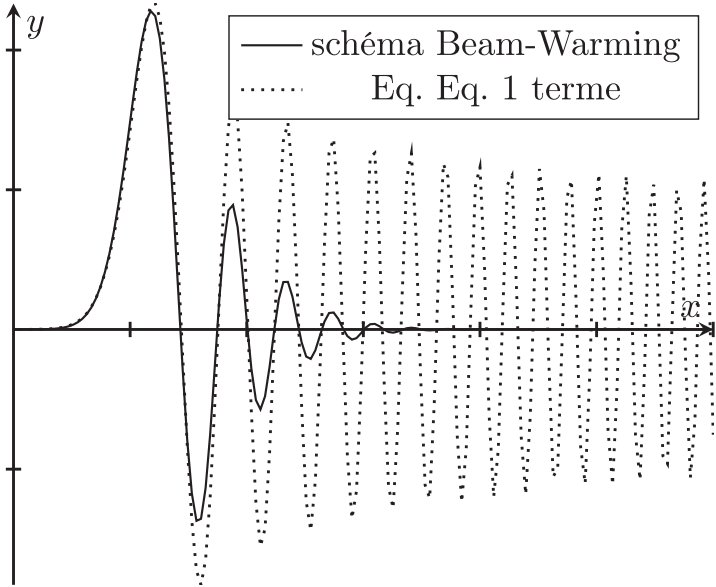


Fig. 5.18 Solution du schéma BW vs. équation équivalente à 1 terme.

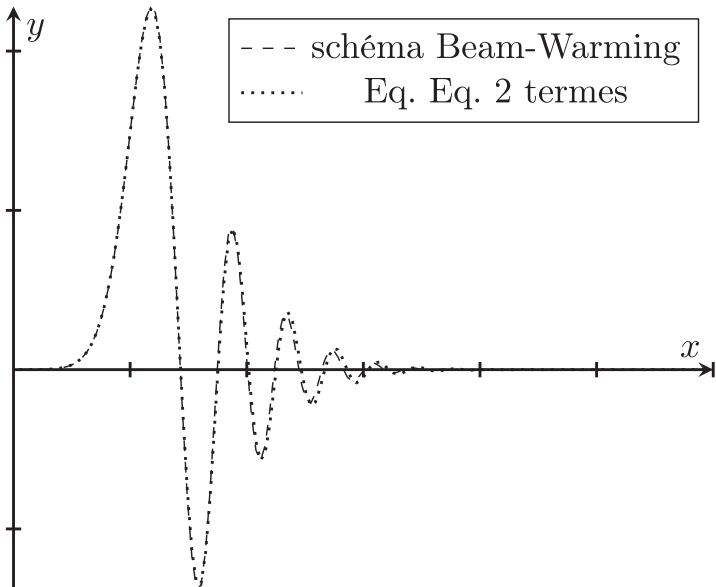


Fig. 5.19 Fonction de Green, schéma vs. EE à 2 termes.

5.7 Propagation de fronts avec des schémas d'ordre 2 dissipatifs

Nous entendons, par propagation de fronts, l'advection d'une condition initiale marche descendante, *i.e.* la CI $u(x, 0) = H(-x)$.

5.7.1 Artéfacts des schémas linéaires pour une condition initiale marche

Comme visible sur les figures 5.20 et 5.21, les schémas LW et BW génèrent des oscillations parasites. C'est une illustration et une quantification du théorème de Godunov, qui montre que tout schéma linéaire monotone est d'ordre nominal 1. La norme L^2 de l'erreur est principalement due à ces oscillations et se comporte comme leur échelle spatiale, en $(\frac{a}{2}(1-q^2)b^2t)^{1/3}$. L'ordre de ces schémas en norme L^1 sur la condition initiale marche est donc seulement de $2/3$. En norme L^∞ ils ne convergent pas.

Nous nous plaçons dans l'approximation de l'équation équivalente à un terme pour obtenir des résultats explicites simples. La solution peut être exprimée comme convolution de la condition initiale avec la fonction de Green

$$u(x, t) = \int_{-\infty}^{\infty} H(-x + y') G(y', t) dy' = \int_x^{\infty} G(y', t) dy' = 1 - \int_{-\infty}^x G(y', t) dy'$$

soit à l'aide d'une primitive PAi de $Ai(x)$ pour le schéma de Lax-Wendroff, ou de $Ai(-x)$ pour le schéma de Beam-Warming. Plus précisément, pour le schéma de Lax-Wendroff

$$u(x, t) = (ct)^{-1/3} \int_x^{\infty} Ai((x' - at)/(ct)^{1/3}) dx' = \int_{(x'-at)/(ct)^{1/3}}^{\infty} Ai(y) dy \quad (5.59)$$

où $c = \frac{a}{2}(1-q^2)b^2$. Les primitives de $Ai(x)$ ont été étudiées, on connaît notamment leurs développements asymptotiques pour

$$x \rightarrow \infty: \int_0^x Ai(x') dx' \sim \frac{1}{3} - \frac{1}{2\sqrt{\pi}x^{3/4}} \exp\left(-\frac{2}{3}x^{3/2}\right)$$

donnés dans [1, chapitre 11] (formule 10.4.82), ou retrouvés par la méthode du col (chapitre 12). La solution en pied de marche est donc monotone. Mais, en haut de la marche, des oscillations apparaissent, comme illustré sur la figure 5.19.

Comme $\int_0^x Ai(-x') dx' = -\int_0^{-x} Ai(t) dt$, on obtient le comportement de la solution pour $x < 0$ grand en valeur absolue, soit, pour la marche descendante, à

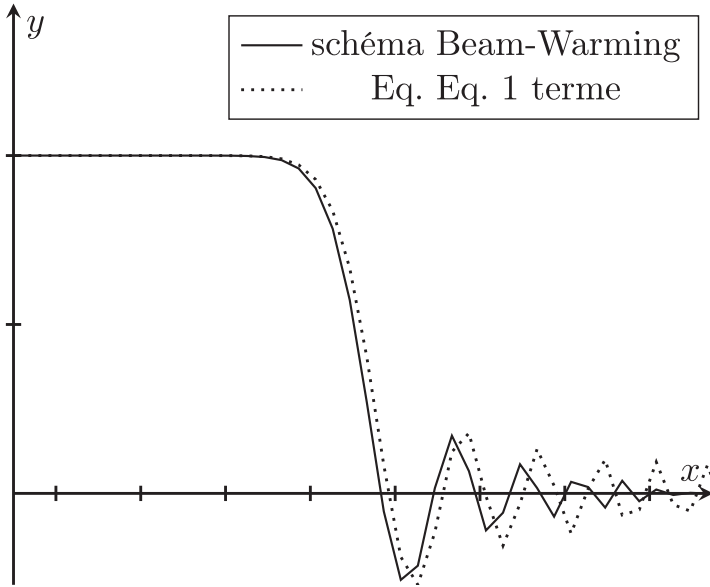


Fig. 5.20 Solution du schéma BW vs. équation équivalente à 1 terme sur une marche descendante.

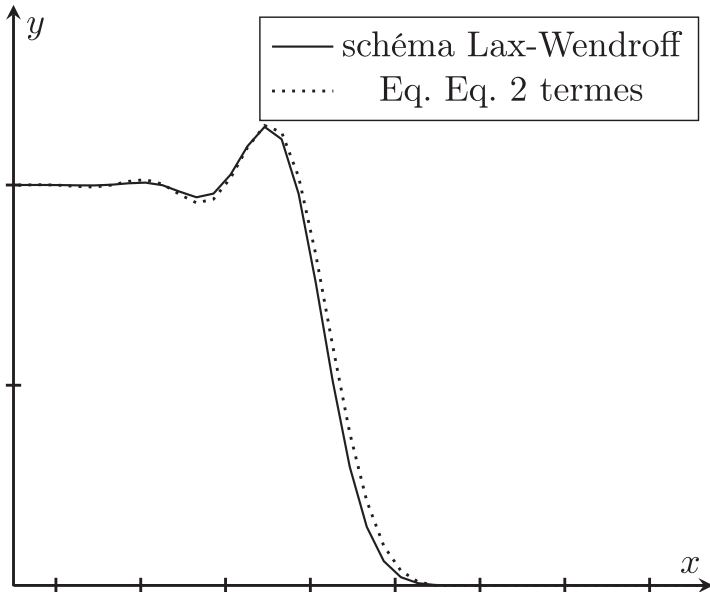


Fig. 5.21 Solution du schéma vs. équation équivalente à 2 termes sur une marche descendante.

gauche et assez loin du front. Pour le schéma de Beam-Warming, le comportement est inversé. La solution est monotone en haut de la marche, mais des oscillations apparaissent en pied de marche, comme illustré sur la figure 5.20. La solution explicite du schéma s'écrit

$$u(x, t) = \int_{(x'-at)/(ct)^{1/3}}^{\infty} Ai(-y) dy \text{ avec } c = \frac{a}{2} (q^2 - 3q + 2) b^2 \quad (5.60)$$

Caractérisons quelques points remarquables sur les profils de marche obtenus avec l'équation équivalente à un terme. Pour le schéma de Lax-Wendroff, les maxima locaux en la variable y sont, dans l'approximation de l'équation équivalente à un terme, aux points où la dérivée s'annule, soit aux zéros, tous négatifs de la fonction d'Airy. Le maximum absolu est obtenu pour le plus petit zéro z_0 en valeur absolue de la fonction d'Airy, qui vaut environ $-2,23$. Il génère un maximum au-dessus de 1 (donc pas de convergence L^∞), qui vaut approximativement 1,275. À partir de ce point, le profil est monotone décroissant, et présente un point d'inflexion pour le plus petit zéro z_1 en valeur absolue de la dérivée de la fonction d'Airy, qui vaut environ $-1,02$. Au point $y = 0$, la solution vaut $u(0, t) = \int_0^{\infty} Ai(y) dy = 1/3$.

Pour le schéma de Beam-Warming, le profil est monotone décroissant jusqu'au plus petit zéro de la fonction d'Airy de $-y$, soit environ $+2,23$, où il atteint son minimum absolu à environ $-0,275$, le point d'inflexion est à $y = -z_1 = +1,02$. Au point $y = 0$, la solution vaut, d'après (5.60)

$$u(0, t) = \int_0^{\infty} Ai(-y) dy = - \int_0^{\infty} Ai(y') dy' = \int_{-\infty}^0 Ai(y') dy' = 2/3$$

Ces résultats sont approximatifs, et peu précis du côté présentant des oscillations pour lequel il faut prendre en compte, comme illustré dans la section précédente, le terme de dissipation dans l'équation équivalente. Ils vont toutefois nous guider pour comprendre comment obtenir des profils monotones. En effet, l'équation équivalente à un terme est une approximation acceptable de la solution de ces schémas pour une condition initiale marche du côté monotone.

5.7.2 Comment obtenir des profils monotones ?

Aucun des deux schémas étudiés ne présente un profil monotone. Pour obtenir un tel profil, il faut sélectionner le schéma approprié, comme expliqué par Banks *et al.* dans [2]. Les schémas de Lax-Wendroff et de Beam-Warming peuvent être réinterprétés de la manière suivante. La solution est reconstruite à l'aide de fonctions linéaires par morceaux, caractérisées, dans chaque maille, par leur moyenne et leur pente. Le profil est ensuite advecté, ce qui conduit à un résultat linéaire par morceaux : la valeur dans la maille est calculée en prenant la moyenne.

Le schéma est déterminé par le choix de la pente p_j dans la maille j . Le choix de la pente aval $\frac{u_{j+1}-u_j}{h}$ conduit au schéma de Lax-Wendroff, celui de la pente amont $\frac{u_j-u_{j-1}}{h}$ au schéma de Beam-Warming. Un profil monotone pour la marche descendante peut s'obtenir en sélectionnant le schéma approprié. Pour Lax-Wendroff, le profil est monotone pour $y > z_1 \simeq -1,02$, pour Beam-Warming il l'est pour $y < -z_1 \simeq +1,02$.

L'idée est de choisir la pente aval si elle croît, *i.e.* si $u_{j+1} - u_j > u_j - u_{j-1}$, la pente amont si elle décroît, *i.e.* si $u_{j+1} - u_j < u_j - u_{j-1}$. Le choix est réalisé par le limiteur min mod, *i.e.* $p_j = \min \text{mod} \left(\frac{u_j - u_{j-1}}{h}, \frac{u_{j+1} - u_j}{h} \right)$, ainsi défini : si $ab > 0$, $\min \text{mod} (a, b) = a$ si $|a| < |b|$, b si $|b| < |a|$, si $ab < 0$, $\min \text{mod} (a, b) = 0$. En somme, ce limiteur choisit la pente la plus petite en module, si elles sont de même signe. Pour la marche descendante, il sélectionne le schéma de BW sur le haut de la marche, si $u_{j+1} - u_j < u_j - u_{j-1}$, *i.e.* $u_{j+1} - 2u_j + u_{j-1} < 0$, le schéma de LW sur le bas de la marche, *i.e.* si $u_{j+1} - u_j > u_j - u_{j-1}$, *i.e.* $u_{j+1} - 2u_j + u_{j-1} > 0$.

En langage équation équivalente, la dérivée de la pente est la dérivée seconde (approchée par $\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2}$) de la solution : une dérivée seconde positive (resp. négative) demande de choisir le schéma de Lax-Wendroff (resp. de Beam-Warming). L'équation équivalente vérifiée par le schéma sera donc celle de Beam-Warming pour $\partial_{xx} u < 0$, celle de Lax-Wendroff pour $\partial_{xx} u > 0$. En somme, l'équation équivalente (à un terme) au schéma composite défini par les règles ci-dessus est :

$$\partial_t u + a \partial_x u = \frac{c}{3} \partial_{3x} u, \quad c = \frac{a}{2} (q^2 - 3q + 2) h^2 \text{ si } \partial_{xx} u < 0$$

$$\partial_t u + a \partial_x u = \frac{-c}{3} \partial_{3x} u, \quad c = \frac{a}{2} (1 - q^2) h^2 \text{ si } \partial_{xx} u > 0$$

Construisons une solution de cette équation, à partir des solutions autosimilaires obtenues plus haut. Plaçons-nous dans le cas simple où les coefficients de dispersion c des deux schémas sont les mêmes en valeur absolue, soit à $q = 1/2$, $c = \frac{3a}{8} h^2$.

$$\partial_t u + a \partial_x u = \frac{c}{3} \partial_{3x} u \text{ si } \partial_{xx} u < 0 \tag{5.61}$$

$$\partial_t u + a \partial_x u = -\frac{c}{3} \partial_{3x} u \text{ si } \partial_{xx} u > 0 \tag{5.62}$$

Postulons, suivant [2], que la solution de l'équation équivalente est autosimilaire par morceaux, de la forme $u\left(\frac{x-at}{(ct)^{1/3}}\right)$. Introduisant cette expression dans l'équation (5.61), on obtient l'équation différentielle $\frac{x-at}{(ct)^{1/3}} u' = u'''$, soit

$$u'''(y) - y u'(y) = 0$$

u' vérifie l'équation d'Airy, les solutions bornées sont de la forme

$$u' = DAi(-y), u = C + DPAi(-y)$$

Pour l'équation équivalente (5.62), les solutions sont de la forme

$$u' = D' Ai(y), u = C' + D' PAi(y)$$

On obtient donc des solutions locales en termes de primitives de la fonction d'Airy. Il faut maintenant les raccorder, et pour cela, définir un point de raccordement. Dans le cas considéré dans ce paragraphe, par symétrie, il est en $y = 0$. Le cas d'un $q \neq 1/2$, pour lequel nous renvoyons à [2], demande de déterminer numériquement ce point. Choisissons comme solution u

$$\text{– pour } y < 0: u = C + D \int_y^\infty Ai(-z) dz$$

$$\text{– pour } y > 0: u = C' + D' \int_y^\infty Ai(z) dz$$

La continuité en 0 impose $C + D \int_0^\infty Ai(-z) dz = C' + D' \int_0^\infty Ai(z) dz$, soit $C + \frac{2}{3}D = \frac{1}{3}D'$. Celle de la dérivée demande $DAi(0) = D' Ai(0)$, *i.e.* $D = D'$. De plus, u doit tendre vers 1 pour $y \rightarrow -\infty$, soit $C + D \int_{-\infty}^\infty Ai(-z) dz = C + D = 1$ et doit tendre vers 0 pour $y \rightarrow +\infty$, soit $C' = 0$. Au final, C, D vérifient le système $C + \frac{1}{3}D = 0, C + D = 1$, dont la solution est $C = -\frac{1}{2}, D = \frac{3}{2}$. La solution approchée du schéma s'écrit donc

$$u = -\frac{1}{2} + \frac{3}{2} \int_y^\infty Ai(-z) dz = 1 - \frac{3}{2} \int_{-y}^\infty Ai(z) dz \text{ si } y < 0 \quad (5.63)$$

$$u = \frac{3}{2} \int_y^\infty Ai(z) dz \text{ si } y > 0 \quad (5.64)$$

u est continue en 0, $u(0) = \frac{1}{2}$, ainsi que sa dérivée $u'(0) = -\frac{3}{2}Ai(0)$, la dérivée seconde est discontinue et vaut $\pm \frac{3}{2}Ai'(0)$. Cette solution correspond à celle de [2], formulée dans cette référence à l'aide d'une fonction hypergéométrique. Elle ne se réduit pas à une juxtaposition des solutions des schémas de départ, à cause du coefficient $3/2$. Elle est comparée à la solution du schéma sur la figure 5.22. L'accord est globalement bon, avec toutefois un léger décalage entre les courbes, sans doute dû à l'utilisation de l'équation équivalente à un terme. Dans la mesure où le but de cette section est principalement d'expliquer aussi simplement que possible comment le limiteur min mod permet d'obtenir des profils monotones pour des conditions initiales marche, nous ne chercherons pas à obtenir une solution plus précise.

Il est possible de déterminer l'ordre de ce schéma composite. L'erreur en norme L^1 est approchée par la même méthode que pour le schéma upwind (voir la discussion

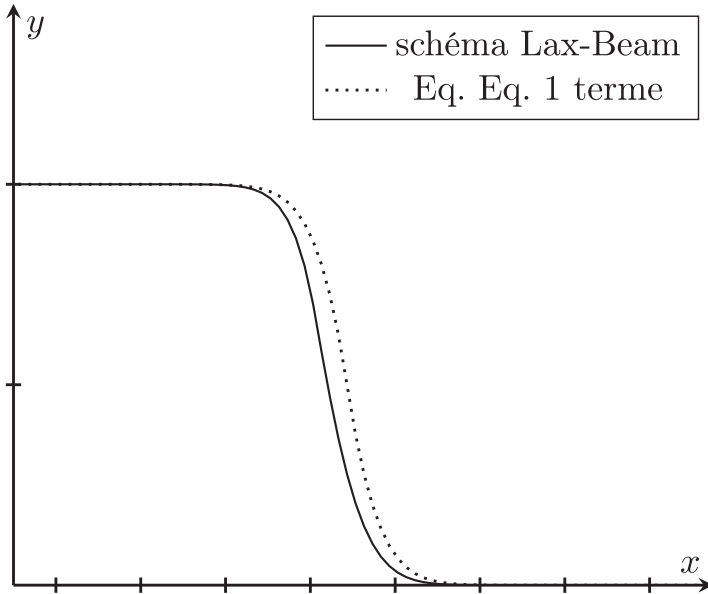


Fig. 5.22 Solution du schéma Beam-Warming avec min mod vs. équation équivalente à 1 terme sur une marche descendante.

autour de (5.17)), en soustrayant de la solution (5.63)-(5.64) du schéma la solution de l'équation $H(at - x)$. L'ordre obtenu est (comme pour les schémas de départ) égal à $\frac{2}{3}$, au lieu de $\frac{1}{2}$ pour le schéma upwind. Cet ordre un peu meilleur vient de l'étalement du profil en $(ct)^{1/3}$, avec $c = \frac{3a}{8}h^2$, donc en $h^{2/3}$, moindre que celui en $h^{1/2}$ du schéma upwind. Mais il est inférieur, comme annoncé dans l'introduction, à l'ordre nominal du schéma, qui ne s'applique que pour des CI assez régulières.

Remarque : ni les schémas de départ, ni le schéma composite ne convergent, pour la CI marche, en norme L^∞ . C'est particulièrement évident pour les schémas de départ. Par exemple, LW génère un maximum qui, pour $t \rightarrow \infty$, tend vers un nombre proche de 1,275, comme expliqué au début de ce paragraphe. Mais c'est aussi vrai pour le schéma composite, puisque sa solution prend des valeurs comprises entre 0 et 1, alors que la solution exacte vaut soit 0, soit 1.

5.8 Conclusion sur les schémas d'ordre 2

Nous avons analysé différents types de schémas d'ordre 2. Le schéma saute-mouton, non dissipatif, convient bien pour les conditions initiales régulières, mais présente des artéfacts marqués pour les conditions initiales peu régulières : oscillations de dispersion peu amorties, et s'étendant donc à assez grande distance. Ce schéma a

de plus un mode parasite en version totalement discrétisée. L'équation équivalente ne prédit correctement le comportement du schéma, pour la fonction de Green ou pour une condition initiale marche, qu'au voisinage du front. L'approximation de la solution exacte sous forme intégrale par la méthode de la phase stationnaire prédit correctement les oscillations de dispersion loin du front.

Les schémas dissipatifs ont des artéfacts moins marqués. Les oscillations de dispersion sont toujours présentes, mais sont amorties par le terme de dissipation. La méthode de l'équation équivalente à deux termes prédit correctement le comportement de ces schémas.

Une caractéristique de ces schémas d'ordre 2, pour la condition initiale marche, est leur comportement monotone en sommet ou en pied de marche, qui ne se retrouve pas aux ordres supérieurs. Le limiteur min mod en tire parti pour obtenir des profils monotones pour cette condition initiale. Cela explique en partie leur utilisation, associée à diverses techniques de limitation, de pente ou de flux, notamment en simulation des écoulements compressibles, où peuvent apparaître des chocs et des discontinuités de contact, dont une idéalisation est la marche.

Une voie alternative, moins utilisée, consiste à monter en ordre. Nous présentons dans la section suivante l'analyse d'un schéma d'ordre 3.

5.9 Schémas d'ordre 3

Par la même méthode que celle utilisée pour les schémas d'ordre 2, nous pouvons obtenir des schémas de Strang (présentés dans le chapitre 10 de l'Appendice) d'ordre 3. Nous allons étudier celui dont le stencil est aux points $\{j-2, j-1, j, j+1\}$, *i.e.* S_3^1 . Il s'obtient en développant la même expression $g = (1 + (e^{-h\partial_x} - 1))^{q+1} e^{h\partial_x}$ que pour le schéma de Lax-Wendroff mais à l'ordre 3, ce qui fournit

$$S_3^1 = LW + \frac{q(q^2 - 1)}{6} (e^{-h\partial_x} - 1)^3 e^{h\partial_x}$$

$$S_3^1 = LW + \frac{q(1 - q^2)}{6} e^{h\partial_x} + \frac{q(q^2 - 1)}{2} + \frac{q(1 - q^2)}{2} e^{-h\partial_x} + \frac{q(q^2 - 1)}{6} e^{-2h\partial_x} \quad (5.65)$$

Sous forme opératorielle, le schéma s'écrit donc, d'après (5.65) et l'expression (5.39) du schéma de LW

$$S_3^1 = -\frac{1}{6} q (q - 1)(q - 2) e^{h\partial_x} + \frac{1}{2} (q^2 - 1)(q - 2) + \frac{1}{2} q (q + 1) (q - 2) e^{-h\partial_x} + \frac{q(q^2 - 1)}{6} e^{-2h\partial_x} \quad (5.66)$$

soit sous la forme

$$u_j^{n+1} = -\frac{q(q-1)(q-2)}{6}u_{j+1}^n + \frac{1}{2}(q^2-1)(q-2)u_j^n - \frac{1}{2}q(q+1)(q-2)u_{j-1}^n + \frac{1}{6}q(q^2-1)u_{j-2}^n$$

L'analyse de Fourier permet bien sûr d'obtenir une représentation exacte de la solution du schéma, sous forme intégrale, mais elle est relativement compliquée. Nous nous limiterons donc à étudier le schéma par la méthode de l'équation équivalente.

Équation équivalente

L'équation équivalente est connue puisqu'il s'agit d'un schéma de Strang. Nous la redérivons à titre d'exercice en développant le logarithme du symbole (5.66)

$$\begin{aligned} \ln(Q) &= \ln\left(-\frac{1}{6}q(q^2-3q+2)e^{h\partial_x} + \frac{1}{2}q^3 - q^2 - \frac{1}{2}q + 1 + \frac{1}{2}q(-q^2+q+2)e^{-h\partial_x}\right. \\ &\quad \left. + \frac{q(q^2-1)}{6}e^{-2h\partial_x}\right) \\ \ln(Q) &= -hq\partial_x + \frac{h^4}{24}\partial_x^4\left(-q^4 + 2q^3 + q^2 - q\right) + O\left(h^5\right) \end{aligned} \quad (5.67)$$

Elle s'écrit donc, selon (5.67)

$$\partial_t u + a\partial_x u = \frac{a}{24}\left(-q^3 + 2q^2 + q - 2\right)h^3\partial_{4x}u \quad (5.68)$$

Le premier terme de l'équation équivalente est un terme de dissipation d'ordre 4, avec le coefficient $C_3 = \frac{1}{6}(q-1)(q-2)(q+1) > 0$ si $q < 1$. L'équation équivalente à un terme étant, ainsi que le schéma, dissipative pour $q < 1$, sa solution va être une bonne approximation de celle du schéma, sans condition sur les conditions initiales, donc y compris pour la fonction de Green. Notons au passage que la limite de stabilité $q = 1$ est bien prédite, mais que pour $q > 2$, $C_3 > 0$, alors que le schéma est instable.

Solution de l'équation équivalente à un terme

La méthode de l'équation équivalente donne une approximation explicite de la fonction de Green du schéma. Elle s'écrit, comme vu dans la section 4.3 du chapitre 4 consacrée à cette méthode

$$G_{EE}(x, t) = \frac{1}{2\pi (C_3 a t h^3)^{1/4}} \int_{-\infty}^{+\infty} \exp ik \left(\frac{x - at}{(C_3 a t h^3)^{1/4}} - \frac{k^4}{4} \right) dk$$

En notant $y = \frac{x-at}{(C_3 a t h^3)^{1/4}}$, G_{EE} s'écrit

$$G_{EE}(x, t) = \frac{1}{2\pi (C_3 a t h^3)^{1/4}} \int_{-\infty}^{+\infty} \exp \left(iky - \frac{k^4}{4} \right) dk \quad (5.69)$$

Le comportement de cette intégrale pour y assez grand peut être étudié par la méthode du col. Les points col vérifient $4k^3 = iy$. Il en existe donc 3. Pour $y > 0$, $k_1 = e^{i\pi/6} \left(\frac{y}{4}\right)^{1/3}$, $k_2 = e^{i5\pi/6} \left(\frac{y}{4}\right)^{1/3}$, $k_3 = -i \left(\frac{y}{4}\right)^{1/3}$. Le contour de descente rapide pertinent se trouve dans le demi-plan supérieur et passe donc par k_1, k_2 . La partie imaginaire non nulle de ces points conduit à une décroissance exponentielle des oscillations, qui sont donc amorties plus rapidement que pour le schéma d'ordre 2.

Dans cette approximation, la fonction de Green du schéma est paire, la solution avec condition initiale marche est impaire, donc des oscillations sont présentes aussi bien pour $x < 0$ que pour $x > 0$. Cela était bien sûr attendu d'après le théorème de Godounov.

La fonction spéciale qui permet de représenter la fonction de Green est une généralisation de la fonction d'Airy, que nous notons

$$Ai^4(y) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(iky - \frac{k^4}{4}\right) dk$$

$$G_{EE}(x, t) = \frac{1}{(C_3 at h^3)^{1/4}} Ai^4\left(\frac{x - at}{(C_3 at h^3)^{1/4}}\right), \quad C_3 = \frac{1}{6} (1 - q^2) (2 - q) \quad (5.70)$$

(5.70) est un cas particulier de la formule (4.48) de la première partie. À noter que cette fonction, ainsi que sa primitive, qui permet de décrire la solution du schéma pour des conditions initiales marche, a été tabulée ([8] de la première partie). C'est une des fonctions d'Airy généralisées présentées dans la section 11.2. La figure 5.23 compare la fonction de Green du schémas à celle de l'équation équivalente à un terme. L'accord est, comme prévu, très satisfaisant.

Il en va de même pour l'advection d'une marche, présentée sur la figure 5.24. Les oscillations parasites sont d'ampleur limitée, et conduisent respectivement, en sommet et en pied de marche, à une bosse et à un creux de faible amplitude, voisine de 0,05, donc bien moins marqués que ceux des schémas d'ordre 2.

On note toutefois sur la figure 5.23 une légère asymétrie dans la solution du schéma, alors que l'équation équivalente donne un résultat symétrique, puisque $Ai^4(-y) = Ai^4(y)$. Cette asymétrie pourrait être prise en compte en incluant le terme de dispersion quantique, mais nous ne poursuivrons pas dans cette voie, au vu de l'accord déjà satisfaisant.

Conclusion

Nous avons simulé pour le schéma de Strang d'ordre 3 S_3^1 l'advection d'une condition initiale Dirac et d'une condition initiale marche descendante. La solution du schéma est, dans les deux cas, aux temps longs, en bon accord avec la solution de l'équation équivalente à un terme. Cette dernière s'exprime, pour la condition initiale Dirac, à l'aide de la fonction Ai^4 , et, pour la condition initiale marche descendante, à l'aide d'une primitive de cette fonction. Les artéfacts sont de faible ampleur, mais

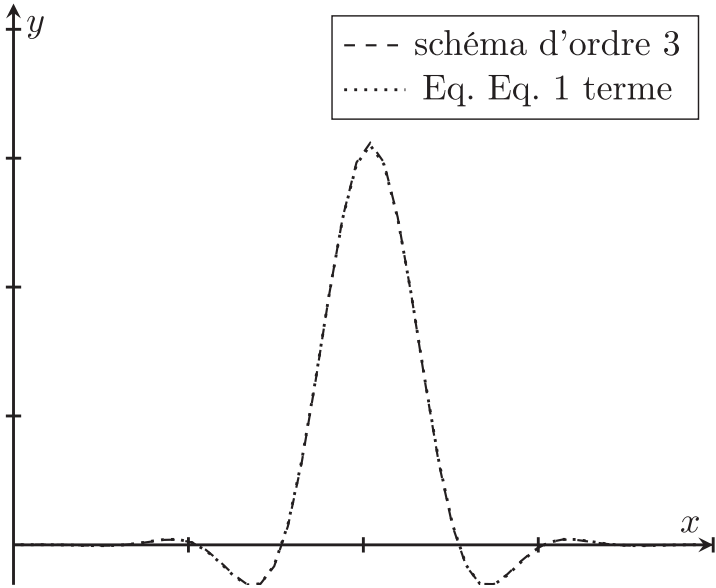


Fig. 5.23 Fonction de Green du schéma S_3^1 comparée à celle de l'équation équivalente G_{EE} .

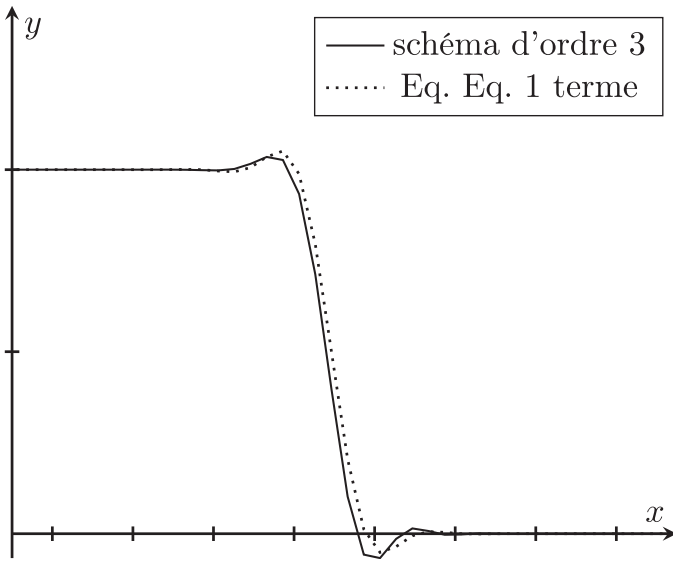


Fig. 5.24 Advection d'une marche par le schéma S_3^1 et approximation équation équivalente.

la technique pour obtenir des profils monotones utilisée pour les schémas d'ordre 2 repose sur la monotonie d'une partie du profil, qui ne se retrouve pas à l'ordre 3. Elle ne s'applique donc pas. L'accord entre la solution du schéma et son approximation par l'équation équivalente à un terme est satisfaisant. Nous poursuivons par l'analyse d'exemples de schémas d'ordre supérieur.

5.10 Schémas d'ordre supérieur

Nous renvoyons au chapitre 10 pour une présentation des types de schémas d'ordre élevé analysés dans ce livre : schémas centrés, décalés et schémas de Strang. Du point de vue de son comportement, et de la méthode d'analyse adaptée, l'essentiel est de savoir si le schéma est dissipatif ou non. Dans le premier cas, l'équation équivalente est une bonne approximation, au moins aux temps longs. Dans le second, il faut utiliser l'analyse de Fourier.

Les schémas centrés et décalés n'ont pas de dissipation : leur facteur d'amplification est de module unité. Les schémas décalés ont l'avantage d'une erreur de troncature moindre et de relations de dispersion plus favorables que celle des schémas centrés. Mais ces derniers sont plus simples à analyser, et permettent d'illustrer les artéfacts générés par ce type de schéma. Nous avons donc choisi d'étudier le schéma centré d'ordre 4, en version semi-discrétisée.

Les schémas de Strang, en revanche, ont une discrétisation couplée espace temps, ce qui les rend particulièrement simples d'utilisation. Ils sont de plus dissipatifs. Nous illustrons les effets de la montée en ordre en analysant un schéma de Strang d'ordre 4 et en présentant les fonctions de Green des schémas d'ordre impair de 3 à 11. Nous commençons pas les schémas sans dissipation.

5.10.1 Schémas d'ordre élevé non dissipatifs

Nous nous limitons aux schémas centrés en version semi-discrétisée, simples à analyser.

Schéma centré d'ordre 4

Il s'écrit

$$\partial_t u_j = -\frac{2a}{3} \frac{u_{j+1} - u_{j-1}}{h} + \frac{a}{12} \frac{u_{j+2} - u_{j-2}}{h} \quad (5.71)$$

Dans l'espace réciproque, (5.71) devient

$$\partial_t \hat{u} = \left(-\frac{4ia}{3} \frac{\sin k}{h} + \frac{ia}{6} \frac{\sin 2k}{h} \right) \hat{u}$$

dont la solution est, pour $\widehat{u}(k, 0) = 1$

$$\widehat{u} = \exp\left(\frac{4ia \sin k}{3} \frac{1}{h} + \frac{ia \sin 2k}{6} \frac{1}{h}\right)$$

et par transformation de Fourier inverse, on obtient la fonction de Green du schéma

$$G_s(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(\frac{iat}{h} \left(-\frac{4}{3} \sin k + \frac{1}{6} \sin 2k\right) + ijk\right) dk \quad (5.72)$$

Le développement du logarithme du facteur d'amplification, qui s'écrit

$$-\frac{4ia \sin k}{3} \frac{1}{h} + \frac{ia \sin 2k}{6} \frac{1}{h} = -i\frac{a}{h}k + \frac{1}{30}i\frac{a}{h}k^5 + O(k^7)$$

confirme que le schéma est bien d'ordre 4 sans dissipation. On anticipe donc que l'équation équivalente ne donnera une prédiction correcte qu'au voisinage du front en $x = jh = at$. La solution loin du front, qui fait intervenir des hautes fréquences spatiales, ne sera pas prédite avec précision. Écrivons toutefois la fonction de Green dans cette approximation

$$G_{EE} \approx \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(+\frac{1}{30}i\frac{a}{h}k^5 t + i\frac{x-at}{h}k\right) dk$$

Elle s'exprime, comme vu plus haut (partie I, (4.45)), en posant $l = \left(\frac{1}{6}\frac{a}{h}t\right)^{1/5}k$, à l'aide de la fonction généralisée Ai^5 :

$$G_{EE} \approx \frac{1}{2\pi} \left(\frac{1}{6}\frac{a}{h}t\right)^{-1/5} \int \exp\left(+\frac{1}{5}il^5 + i\frac{x-at}{h} \left(\frac{1}{6}\frac{a}{h}t\right)^{-1/5} l\right) dl$$

$$G_{EE} \approx \left(\frac{1}{6}\frac{a}{h}t\right)^{-1/5} Ai^5\left(\frac{x-at}{h} \left(\frac{1}{6}\frac{a}{h}t\right)^{-1/5}\right)$$

La fonction de Green exacte s'écrit sous la forme (5.72), que nous réécrivons

$$G_s = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\frac{iat}{h} \left(-\frac{4}{3} \sin k + \frac{1}{6} \sin 2k + \frac{jh}{at}k\right) dk$$

La figure 5.25 compare ce résultat exact à une simulation : les deux courbes se superposent, si le calcul en Fourier est assez précis. L'accord entre G_{EE} et G_s n'est bon, comme pour tous les schémas non dissipatifs, qu'au voisinage du front, comme le montre la figure 5.26. L'équation équivalente sous-estime l'amplitude des oscillations loin du front.

La vitesse de groupe du schéma est

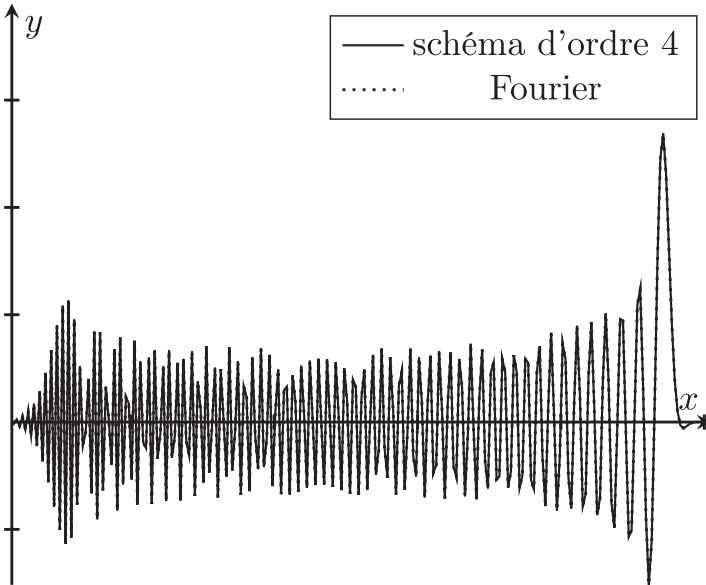


Fig. 5.25 *Fonction de Green 5.72 (calculée en Fourier) du schéma et simulation.*

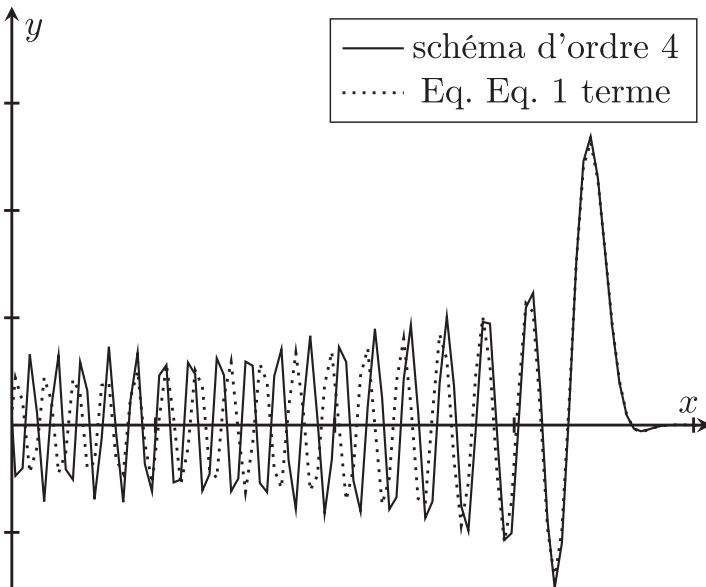


Fig. 5.26 *Fonction de Green du schéma et de l'équation équivalente à 1 terme G_{EE} .*

$$v_g = \frac{d}{dk} \left(\frac{4}{3} \sin k - \frac{1}{6} \sin 2k \right) = -\frac{1}{3} \cos 2k + \frac{4}{3} \cos k = 1 - \frac{8}{3} \sin^4 \frac{k}{2}.$$

Le contact en $k = 0$, avec la vitesse de groupe de l'équation est d'ordre 4. Elle décroît à partir de $k = 0$, où elle vaut a , passe par 0 pour $\sin^4 \frac{k}{2} = \frac{3}{8}$, générant un mode stationnaire au nombre d'onde $k = 2 \arcsin \left(\left(\frac{3}{8} \right)^{\frac{1}{4}} \right)$. Elle devient négative et vaut $-\frac{5}{3}a$ en $k = \pi$, générant un mode rétropropagatif à vitesse $-\frac{5}{3}a$. Rappelons que, pour le schéma centré d'ordre 2, la vitesse de groupe est $v_g = a \cos k$ et le mode rétropropagatif se propage à vitesse $-a$. La queue dispersive s'étend donc plus loin pour l'ordre 4 que pour l'ordre 2, comme visible sur la figure 5.25.

Schémas centrés d'ordre supérieur à 4

Les schémas centrés d'ordre élevé se comportent essentiellement comme le schéma d'ordre 4 que nous venons d'étudier. Leurs coefficients peuvent être, comme expliqué dans l'Appendice, générés à un ordre arbitraire à l'aide du développement de Taylor de la fonction $\frac{\arcsin \theta}{\theta \sqrt{1-\theta^2}}$. Leur facteur d'amplification s'écrit, pour un schéma

$$\text{d'ordre } 2p \quad \widehat{Q}_{2p} = \frac{1}{h} \sin k \sum_{v=0}^{p-1} \alpha_v 2^{2v} \left(\sin \frac{k}{2} \right)^{2v} \text{ où } \alpha_v = \frac{v}{4v+2} \alpha_{v-1}, \alpha_0 = 1$$

Leur vitesse de groupe $\frac{d}{dk} \widehat{Q}_{2p}$ s'exprime aussi explicitement. Le point essentiel est que, plus l'ordre du schéma est élevé, plus cette vitesse est proche de a dans un voisinage de $k = 0$, et même dans une bande de fréquences spatiales assez étendue autour de $k = 0$, mais plus elle s'en éloigne, et devient négative et grande en valeur absolue au voisinage de $k = \pi$, ce qui crée une pollution haute fréquence de la solution. Le schéma d'ordre 6 a par exemple le stencil à 7 points ci-dessous

$$b^{-1} \left(\frac{3}{4} (u_{j+1} - u_{j-1}) - \frac{3}{20} (u_{j+2} - u_{j-2}) + \frac{1}{60} (u_{j+3} - u_{j-3}) \right)$$

d'où la relation de dispersion et la vitesse de groupe, pour $a = 1$

$$\omega(k) = \frac{3}{2} \sin k - \frac{3}{10} \sin 2k + \frac{1}{30} \sin 3k = k - \frac{1}{140} k^7 + O(k^8)$$

$$v_g = \frac{d\omega}{dk} = \frac{3}{2} \cos k - \frac{3}{5} \cos 2k + \frac{1}{10} \cos 3k = 1 - \frac{1}{20} k^6 + O(k^8)$$

On voit sur la figure 5.27 que les vitesses de groupe des deux schémas sont, dans un voisinage assez large de $k = 0$, à quelques pourcents de celle de l'équation, et que celle du schéma d'ordre 6 en est plus proche. Mais elles s'en écartent fortement quand on se rapproche de $k = \pi$. Elles valent respectivement $-\frac{5}{3}$ et $-\frac{11}{5}$ en $k = \pi$. La queue dispersive s'étend donc, pour le schéma d'ordre 6, jusqu'à $-\frac{11}{5}at$. Le schéma

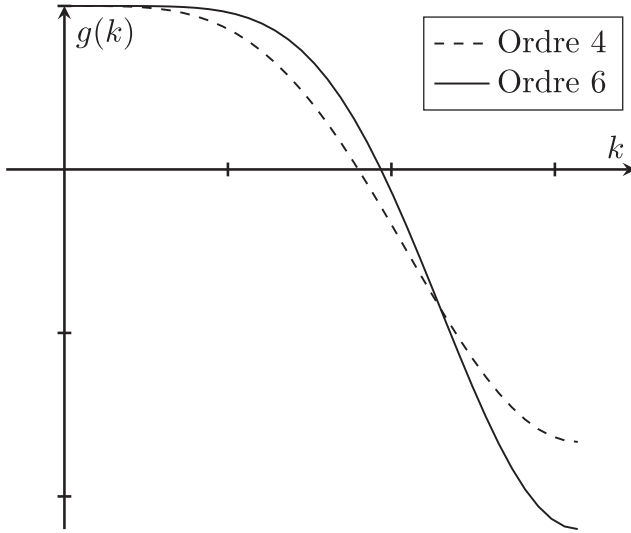


Fig. 5.27 Tracé de v_g pour les schémas centrés d'ordre 4 et 6.

d'ordre 6 présente donc les mêmes types d'artéfacts que le schéma d'ordre 4, mais ils sont encore plus marqués.

Emploi des schémas d'ordre élevé non dissipatifs

Les artéfacts de ces schémas très oscillatoires les rendent inadaptés aux conditions initiales peu régulières. Ils peuvent être utilisés pour des conditions initiales régulières, dont le spectre est rapidement décroissant. Ce point est illustré par [18], qui présente les résultats obtenus avec, comme conditions initiales, des gaussiennes de différentes largeurs $\frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{x^2}{2\sigma^2})$, dont la transformée de Fourier est $\frac{\sigma}{\sqrt{2\pi}} e^{-\frac{1}{2}k^2\sigma^2}$. Les oscillations sont amorties en $e^{-\frac{1}{2}k^2\sigma^2}$ et deviennent négligeables pour σ assez grand : c'est une illustration de la convergence de ces schémas pour ce type de conditions initiales.

Pour utiliser ces schémas pour des simulations de solutions peu régulières, il est préférable de leur ajouter de la dissipation. C'est ce qui est proposé notamment dans [18], qui étudie une dissipation de la forme $d_0 \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + d_1 \frac{u_{j+2} - 2u_j + u_{j-2}}{h^2} + d_2 \frac{u_{j+3} - 2u_j + u_{j-3}}{h^2} + \dots$ avec des coefficients optimisés. Dans l'espace réciproque, elle induit un amortissement en $d_0 \sin^2(\frac{k}{2}) + d_1 \sin^2(k) + d_2 \sin^2(\frac{3k}{2}) + \dots$. Les coefficients ont des signes alternés pour que l'amortissement soit monotone croissant, *i.e.* d'autant plus fort que la fréquence spatiale est plus grande, tout en amortissant le moins possible les modes de faible k . Une autre possibilité est d'employer des

schémas d'ordre élevé intrinsèquement dissipatifs, dont des exemples sont étudiés ci-après.

5.10.2 Schémas d'ordre élevé dissipatifs

Ce sont tous les schémas d'ordre impair, dont le terme principal de l'équation équivalente est dissipatif, et les schémas d'ordre pair non centrés.

Fonction de Green dans l'approximation équation équivalente

Il est possible, comme expliqué dans l'Appendice (chapitre 10), de concevoir des schémas d'ordre arbitraire, qui sont utiles en particulier pour réaliser des simulations précises avec un nombre raisonnable de mailles. Comme pour les schémas d'ordre 2, l'équation équivalente ne fournit une approximation correcte pour des conditions initiales peu régulières que si elle intègre de la dissipation. Les fonctions de Green de ces schémas s'expriment, dans l'approximation de l'équation équivalente, à l'aide des fonctions d'Airy généralisées.

Pour les schémas d'ordre impair $2p + 1$, un terme suffit, comme pour l'ordre 3, et la fonction à calculer est ((4.48) première partie)

$$Ai^{2p+2}(y) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(iky - \frac{k^{2p+2}}{2p+2}\right) dk$$

Pour les schémas d'ordre pair $2p$, deux termes sont nécessaires, comme pour l'ordre 2, et la fonction à calculer est

$$Ai^{2p+1,b}(y) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(iky - \frac{(ik)^{2p+1}}{2p+1} - b \frac{k^{2p+1}}{2p+2}\right) dk$$

Elle dépend du paramètre de dissipation b , qui décroît au cours du temps.

Condition initiale marche

L'approximation par méthode du col de l'intégrale représentant la solution de l'équation équivalente [3] montre que, plus le schéma est d'ordre élevé, plus les points col se rapprochent de l'axe réel, et moins les oscillations sont amorties. En particulier, l'absence d'oscillations (en sommet ou pied de marche, suivant le signe du terme de dispersion) qui permet, comme expliqué plus haut, de propager des marches avec des profils monotones se perd dès l'ordre 3. Le problème de limitation reste ouvert pour ces schémas.

Outre leur efficacité, les schémas d'ordre $2p + 1$ génèrent les fronts s'élargissant en $t^{1/2p+2}$, ce qui permet, si l'ordre est assez grand, d'obtenir des fronts sur une ou

deux mailles. Del Pino *et al.* [6] ont ainsi réalisé des simulations en acoustique avec un schéma de Strang d'ordre 17. Comme la complexité des formules augmente avec l'ordre, nous nous sommes limités à l'étude d'un schéma d'ordre 4, dû à Balakine.

5.10.3 Un exemple de schéma dissipatif d'ordre 4

Le schéma de Balakine s'écrit

$$u_j^{n+1} = -\frac{q}{24}(1-q^2)(2+q)u_{j-2}^n + \frac{q}{6}(1+q)(4-q^2)u_{j-1}^n + (1-\frac{5}{4}q^2 + \frac{1}{4}q^4)u_j^n - \frac{q}{6}(1-q)(4-q^2)u_{j+1}^n + \frac{q}{24}(1-q^2)(2-q)u_{j+2}^n \quad (5.73)$$

Sa fonction de Green ((4.10) première partie) s'écrit

$$G_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{Q}_p(k)^n e^{ikj} dk$$

Le symbole $\widehat{Q}_p(k)$ du schéma se déduit de (5.73)

$$\begin{aligned} \widehat{Q}_p(k) = & -\frac{q}{24}(1-q^2)(2+q)e^{-2ik} + \frac{q}{6}(1+q)(4-q^2)e^{-ik} + \\ & (1 - \frac{5}{4}q^2 + \frac{1}{4}q^4) - \frac{q}{6}(1-q)(4-q^2)e^{ik} + \frac{q}{24}(1-q^2)(2-q)e^{2ik} \end{aligned}$$

Le schéma peut être obtenu en développant à son ordre (ici 4) la fonction générant le schéma de Strang S_4^2 , *i.e.*

$$S_4^2 = (1 + (e^{-h\frac{\partial}{\partial x}} - 1))^{q+2} e^{2h\frac{\partial}{\partial x}} \quad (5.74)$$

Développons la puissance dans (5.74)

$$\begin{aligned} S_4^2 = & 1 + (q+2)(e^{-h\frac{\partial}{\partial x}} - 1) + \frac{(q+2)(q+1)}{2}(e^{-h\frac{\partial}{\partial x}} - 1)^2 \\ & + \frac{(q+2)(q+1)q}{6}(e^{-h\frac{\partial}{\partial x}} - 1)^3 + \frac{(q+2)(q^2-1)q}{24}(e^{-h\frac{\partial}{\partial x}} - 1)^4 \end{aligned} \quad (5.75)$$

en développant l'expression (5.75), on obtient la forme opératorielle du schéma S_4^2

$$\begin{aligned} S_4^2 = & -\frac{1}{24}q(q^2-1)(q-2)e^{2h\frac{\partial}{\partial x}} - \frac{1}{6}q(q-1)(q^2-4)e^{h\frac{\partial}{\partial x}} + \frac{1}{4}(q^2-1)(q^2-4) \\ & + \frac{1}{6}q(q+1)(q^2-4)e^{-h\frac{\partial}{\partial x}} + \frac{1}{24}q(q-1)(q-2)(q+1)e^{-2h\frac{\partial}{\partial x}} \end{aligned} \quad (5.76)$$

Elle correspond à la définition (5.73) du schéma de Balakine, qui est donc bien S_4^2 . Le symbole de l'équation équivalente à 2 termes est donc (voir Appendice, section 10.1)

$$1 - i(-1)^2 \alpha_{2,4} \left(\frac{k^5}{5} - iq \frac{k^6}{6} \right)$$

Rappelons la définition de $\alpha_{2,4}$

$$\alpha_{2,4} = \frac{1}{24} (2 + q)(1 + q)q(q - 1)(q - 2)$$

d'où les coefficients (10.12) et (10.13) des deux premiers termes de l'équation équivalente

$$C_4 = \frac{1}{24} (q - 1)(q - 2)(q + 2)(q + 1)$$

$$C_5 = \frac{1}{24} q(q - 1)(q - 2)(q + 2)(q + 1)$$

Ces résultats figurent aussi dans [BB]. La fonction de Green s'écrit, dans l'approximation équation équivalente à un terme

$$G_{EE1}(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left(ik(x - at) + \frac{1}{5} iC_4 atb^4 k^5 \right) dk$$

Posant $k' = (C_4 atb^4)^{1/5} k$

$$G_{EE1}(x, t) = \frac{1}{2\pi (C_4 atb^4)^{1/5}} \int_{-\infty}^{+\infty} \exp \left(ik' \frac{x - at}{(C_4 atb^4)^{1/5}} + \frac{1}{5} ik'^5 \right) dk'$$

Donc, dans cette approximation

$$G_{EE1} = \frac{1}{2\pi (C_4 atb^4)^{1/5}} Ai^5 \left(\frac{x - at}{(C_4 atb^4)^{1/5}} \right) \quad (5.77)$$

Mais, pour approcher correctement la fonction de Green du schéma, l'équation équivalente doit inclure le terme de dissipation, donc être à deux termes, ce qui donne

$$G_{EE2}(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp \left(ik(x - at) + \frac{1}{5} iC_4 atb^4 k^5 - \frac{1}{6} C_5 atb^5 k^6 \right) dk$$

$$G_{EE2}(x, t) = \frac{1}{2\pi (C_4 atb^4)^{1/5}} \int_{-\infty}^{+\infty} \exp\left(ik' \frac{x - at}{(C_4 atb^4)^{1/5}} + \frac{1}{5} ik'^5 - \frac{1}{6} C_5 atb^5 (C_4 atb^4)^{-6/5} k'^6\right) dk'$$

Le terme de dissipation a pour coefficient

$$d = \frac{C_5 atb^5}{(C_4 atb^4)^{6/5}} = ab^5 a^{-6/5} b^{-24/5} \frac{C_5}{(C_4)^{6/5}} t^{-1/5} = \left(\frac{b}{at}\right)^{1/5} \frac{C_5}{C_4^{6/5}}$$

Il décroît lentement, en $t^{-1/5}$: les oscillations sont amorties, même après un grand nombre de pas de temps. La fonction spéciale représentant G_{EE2} est une fonction d'Airy d'ordre 5 dissipative

$$G_{EE2}(x, t) = (C_4 atb^4)^{-1/5} Ai^{5,d}\left(\frac{x - at}{(C_4 atb^4)^{1/5}}\right) \quad (5.78)$$

Rappelons sa définition : $Ai^{5,d}(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(iky + \frac{1}{5} ik^5 - d \frac{k^6}{6}) dk'$. La figure 5.28 compare la solution du schéma à celle de l'équation équivalente à deux termes G_{EE2} : l'accord est globalement, mais se dégrade quand on s'éloigne du front, comme pour les schémas d'ordre 2.

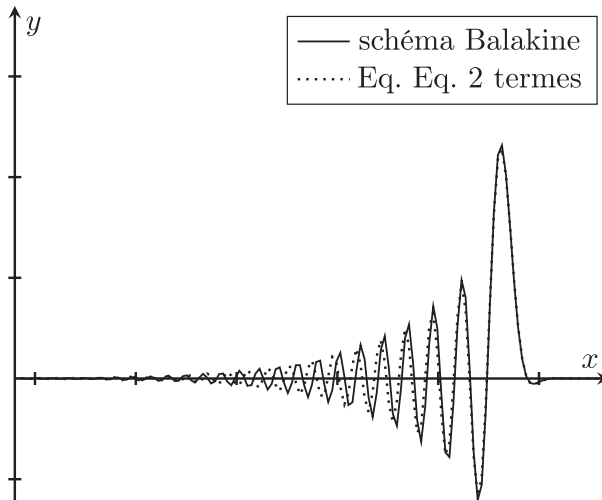


Fig. 5.28 Fonction de Green du schéma de Balakine et de l'équation équivalente à deux termes.

5.10.4 Schémas de Strang d'ordre impair

La fonction de Green d'un schéma de Strang d'ordre impair $2p - 1$ est approximable aux temps longs par la formule

$$G_{EE1}(x, t) = (C_{2p-1} at h^{2p-1})^{-1/2p} Ai^{2p} \left(\frac{x - at}{(C_{2p-1} at h^{2p-1})^{1/2p}} \right) \quad (5.79)$$

La figure 5.29 compare les fonctions d'Airy Ai^{2p} pour $p = 2, 3, 4, 5, 6$. Les oscillations deviennent plus marquées quand l'ordre du schéma croît. En effet, les points col se rapprochent de l'axe réel, ce qui conduit à diminuer l'amortissement des oscillations, comme évoqué à la fin de la sous-section précédente. Monter en ordre permet, pour des conditions initiales marche, de réduire l'étalement du front, mais au prix d'oscillations un peu plus marquées.

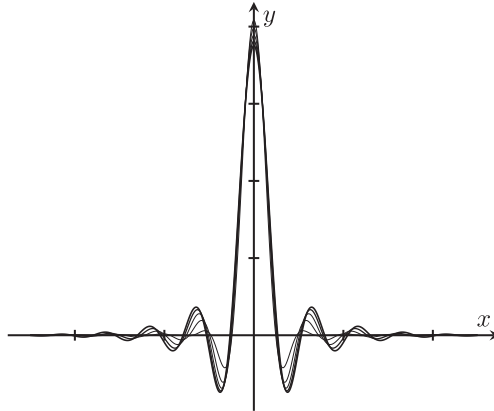


Fig. 5.29 Fonction de Green approchée des schémas de Strang d'ordre 3, 5, 7, 9 et 11 par l'équation équivalente à un terme.

5.10.5 Conclusion sur les schémas d'ordre élevé

Nous avons étudié des schémas d'ordre élevé, dissipatifs ou non. Leur intérêt n'est pas seulement théorique. Ils sont employés quand une solution précise est nécessaire : diffraction des ondes par des structures grandes en termes de longueur d'onde, aérodynamique, simulation directe d'écoulements (où la viscosité numérique doit être plus faible que la viscosité physique) par exemple, mais pour des conditions initiales régulières. Les schémas non dissipatifs sont affectés d'artéfacts très marqués pour des conditions initiales non régulières. En particulier, leur fonction de Green présente une longue queue dispersive, non reproduite par l'équation équivalente. Il est donc préférable d'utiliser des schémas dissipatifs, comme les schémas de Strang. La

fonction de Green des schémas d'ordre impair, dont le premier terme de l'équation équivalente est dissipatif, est correctement prédite, aux temps longs, par l'équation équivalente à un terme. Celles des schémas de Strang d'ordre pair doit en revanche être approchée en utilisant l'équation équivalente à deux termes, intégrant la dissipation.

5.11 Conclusion sur les schémas pour l'advection 1D

Nous avons passé en revue et analysé différents schémas discrétisant l'équation d'advection en 1D. Pour le schéma upwind, schéma d'ordre 1 très utilisé, la fonction de Green exacte a une expression simple si le maillage est uniforme. Cela permet, par comparaison avec celles des méthodes approchées, de bien en cerner les limites de validité. L'équation équivalente donne de bons résultats pour ce schéma diffusif. La diffusion, inhérente à tout schéma d'ordre 1, étale les discontinuités.

Les schémas d'ordre 2 non dissipatifs, donc sans diffusion, sont moins affectés par ce défaut, mais génèrent des oscillations pour des conditions initiales non régulières. L'analyse de Fourier fournit des solutions exactes sous forme intégrale, approximables par la méthode de la phase stationnaire. L'équation équivalente ne donne de bons résultats que dans les zones où les composantes haute fréquence spatiale sont peu importantes, soit près de $x = at$. Ces schémas, à cause de leurs artéfacts, sont surtout adaptés aux conditions initiales régulières.

L'équation équivalente à deux termes rend bien compte du comportement des schémas d'ordre 2 dissipatifs. Les oscillations qu'ils génèrent sont moins importantes que pour les schémas d'ordre non dissipatifs, et peuvent être dans le cas de la condition initiale marche, supprimées, en utilisant des limiteurs adaptés. Ils constituent donc une sorte d'optimum, ce qui explique leur utilisation intensive.

Les schémas d'ordre 3 constituent un autre optimum. Leurs artéfacts sont de faible ampleur. L'équation équivalente à un terme rend bien compte de leur comportement. Par contre, la question des limiteurs reste à notre connaissance ouverte.

Les schémas d'ordre supérieur à trois sans dissipation, par exemple les schémas centrés, génèrent des oscillations importantes, sur des zones étendues, mal rendues par l'équation équivalente.

Les schémas de Strang, qui sont dissipatifs, sont moins affectés, particulièrement s'ils sont d'ordre impair. Le terme principal de leur équation équivalente, dissipatif, amortit en effet les oscillations. Ils sont bien approchés par leur équation équivalente à un terme. Les schémas de Strang d'ordre pair ont un terme principal de l'équation équivalente dispersif, donc des artéfacts plus marqués. Une prédiction précise de leur comportement demande d'utiliser l'équation équivalente à deux termes. Ces schémas élargissent d'autant moins les discontinuités, mais génèrent par contre des oscillations au voisinage des discontinuités d'autant moins amorties que leur ordre est élevé. En particulier, les schémas d'ordre élevé impair génèrent, comme le schéma d'ordre 3, des oscillations avec un amortissement exponentiel, mais cet amortissement décroît avec

l'ordre : aller au-delà de l'ordre 3 a l'avantage de raidir de front, mais l'inconvénient d'oscillations plus marquées.

5.12 Schémas pour l'advection en 2D

Dans ce chapitre, nous nous sommes jusque-là focalisés sur les schémas 1D pour l'advection, à l'exception d'une courte incursion sur le 2D dans la section sur le schéma upwind. Nous traitons maintenant plus en détail l'équation d'advection en 2D. Nous étudierons en particulier les anisotropies induites par le maillage, en considérant une vitesse constante. Nous traiterons un cas où la solution exacte est connue, le schéma upwind sur maillage constitué de mailles carrées.

5.12.1 Schéma upwind sur maillage carré

Dans la généralisation la plus immédiate du schéma 1D, il s'écrit, notant a et b les composantes respectives de la vitesse suivant x et y , supposées positives.

$$\frac{u_{jm}^{n+1} - u_{jm}^n}{\tau} + a \frac{u_{jm}^n - u_{(j-1)m}^n}{h} + b \frac{u_{jl}^n - u_{j(m-1)}^n}{h} = 0 \quad (5.80)$$

(5.80) s'écrit sous forme barycentrique, posant $q = \frac{a\tau}{h}$, $r = \frac{b\tau}{h}$

$$u_{jm}^{n+1} = (1 - q - r)u_{jm}^n + qu_{(j-1)m}^n + ru_{j(m-1)}^n$$

Nous avons obtenu, dans la première partie, une solution exacte par méthode opératorielle pour la fonction de Green, que nous rappelons

$$G_{jm}^n = \frac{n!}{j!m!(n-j-m)!} p^{n-j-m} q^j r^m = \binom{n}{j+m} \binom{j+m}{m} p^{n-j-m} q^j r^m \quad (5.81)$$

Approximation normale

Rappelons l'approximation normale de la fonction de Green G_j^n , valide pour $j \approx nq$ dans le cas 1D :

$$G_j^n = p^{n-j} q^j \binom{n}{j} \approx \frac{1}{\sqrt{2\pi npq}} \exp\left(-\frac{(j-nq)^2}{2npq}\right) \quad (5.82)$$

Cette approximation coïncide avec celle obtenue par la méthode de l'équation équivalente. On peut obtenir le même type d'approximation pour $j \approx nq$, $m \approx nr$ dans le cas 2D. Dans le cas de l'advection suivant une diagonale, pour $j \approx nq$, $m \approx nq$,

en remplaçant les coefficients binomiaux par leurs approximations normales

$$\begin{aligned} \binom{n}{j+m} p^{n-j-m} (2q)^{j+m} &\approx \frac{1}{2\sqrt{\pi n(1-2q)q}} \exp\left(-\frac{(j+m-2nq)^2}{4n(1-2q)q}\right) \\ \binom{j+m}{m} 2^{-j-m} &\approx \frac{\sqrt{2}}{\sqrt{\pi(j+m)}} \exp\left(-\frac{2(m-(j+m)/2)^2}{(j+m)}\right) \\ &= \frac{\sqrt{2}}{\sqrt{\pi(j+m)}} \exp\left(-\frac{(m-j)^2}{2(j+m)}\right) \end{aligned}$$

et en reportant dans (5.82), on obtient

$$G_{jm}^n \approx \frac{1}{2\pi nq\sqrt{(1-2q)}} \exp\left(\frac{1}{4nq} \left((j+m-2nq)^2(1-2q)^{-1} + (m-j)^2\right)\right) \quad (5.83)$$

puis, en remplaçant dans (5.83) les indices j, m, n par leurs valeurs respectives $\frac{x}{h}, \frac{y}{h}, \frac{t}{\tau}$

$$G_{jm}^n \approx \frac{h}{2\pi at\sqrt{(1-2q)}} \exp\left(\frac{1}{4abt} \left((x+y-2at)^2(1-2q)^{-1} + (x-y)^2\right)\right)$$

Cette expression est à comparer au résultat (5.84) obtenu ci-dessous par l'approximation équation équivalente. Les deux expressions coïncident, au facteur usuel h^2 près. En somme, comme en 1D, approximation normale de la solution exacte et méthode de l'équation équivalente conduisent au même résultat. Nous allons maintenant détailler les étapes de l'approximation équation équivalente.

Résolution par équation équivalente

L'équation équivalente peut être déterminée par développement de Taylor, ou par méthode opératorielle. Le schéma s'écrit formellement.

$$\exp(\tau\partial_t) = 1 + q(\exp(-h\partial_x) - 1) + r(\exp(-h\partial_y) - 1) \quad (5.84)$$

En développant dans (5.84) les exponentielles

$$\begin{aligned} \exp(-h\partial_x) - 1 &= -h\partial_x + \frac{1}{2}h^2\partial_x^2 + O(h^3), \exp(-h\partial_y) - 1 \\ &= -h\partial_y + \frac{1}{2}h^2\partial_y^2 + O(h^3) \end{aligned}$$

et on obtient, en reportant ces expressions dans (5.84) et en prenant le logarithme

$$\tau \partial_t = \ln \left(1 - qb\partial_x - rb\partial_y + \frac{1}{2}qb^2\partial_x^2 + \frac{1}{2}rb^2\partial_y^2 \right)$$

Par développement de Taylor de cette expression, on obtient

$$\tau \partial_t = -b(q\partial_x + r\partial_y) + b^2 \left(\frac{1}{2}q\partial_x^2 - (q\partial_x + r\partial_y) \left(\frac{1}{2}q\partial_x + \frac{1}{2}r\partial_y \right) + \frac{1}{2}r\partial_y^2 \right) + O(b^3)$$

soit le terme de diffusion

$$\frac{ab}{2}(1-q)\partial_{xx}^2 + \frac{bh}{2}(1-r)\partial_{yy}^2 - ab\tau\partial_{xy}^2$$

Ce terme est anisotrope. Dans le cas d'une vitesse suivant la diagonale, la matrice de diffusion s'écrit

$$D = \frac{ab}{2} \begin{pmatrix} 1-q & -q \\ -q & 1-q \end{pmatrix}$$

avec comme valeurs propres $\frac{ab}{2}$, $\frac{bh}{2}(1-2q)$ et comme vecteurs propres respectifs $(1, -1)$, $(1, 1)$. Donc la diffusion est anisotrope, plus forte orthogonalement à la direction de propagation que selon cette direction. L'équation équivalente s'écrit

$$\partial_t u + a\partial_x u + a\partial_y u = \frac{ab}{2}((1-q)(\partial_{xx}^2 + \partial_{yy}^2) - 2q\partial_{xy}^2)u$$

Le changement de coordonnées $X = \frac{1}{\sqrt{2}}(x+y)$, $Y = \frac{1}{\sqrt{2}}(x-y)$ permet de l'écrire sous la forme

$$\partial_t u + \sqrt{2}a\partial_X u = \frac{ab}{2}((1-2q)\partial_{XX}^2 + \partial_{YY}^2)u$$

La fonction de Green est une gaussienne de centre (at, at) , avec la matrice de variance de l'équation équivalente, soit

$$G_{EE} = \frac{1}{2\pi abt\sqrt{1-2q}} \exp \left[-\frac{1}{2abt} \left((1-2q)^{-1} \left(\frac{x+y}{\sqrt{2}} - \sqrt{2}at \right)^2 + \left(\frac{x-y}{\sqrt{2}} \right)^2 \right) \right] \quad (5.85)$$

5.12.2 Schéma splitté

Une diffusion plus isotrope peut être obtenue par splitting ou par correction de coin [19]. Le schéma s'écrit dans ce cas

$$u_{jm}^{n+1} = (1-q)(1-r)u_{jm}^n + q(1-r)u_{(j-1)m}^n + r(1-q)u_{j(m-1)}^n + qru_{(j-1)(m-1)}^n$$

Il est aussi barycentrique. Une solution exacte peut aussi être obtenue par méthode opératorielle, mais elle s'exprime comme une somme de coefficients multinomiaux. L'équation équivalente peut être déterminée par développement de Taylor, ou par méthode opératorielle. Le schéma s'écrit formellement.

$$\begin{aligned} \exp(\tau \partial_t) &= (1-q)(1-r) + q(1-r) \exp(-h\partial_x) + r(1-q) \exp(-h\partial_y) \\ &\quad + qr \exp(-h\partial_x) \exp(-h\partial_y) \end{aligned} \quad (5.86)$$

Développons les exponentielles dans (5.86)

$$\begin{aligned} \exp(-h\partial_x) &= 1 - h\partial_x + \frac{1}{2}h^2\partial_x^2 + O(h^3), \exp(-h\partial_y) \\ &= 1 - h\partial_y + \frac{1}{2}h^2\partial_y^2 + O(h^3) \end{aligned}$$

$(1-q)(1-r) + q(1-r) + r(1-q) + qr = 1$, donc

$$\exp(\tau \partial_t) = 1 - qh\partial_x - rh\partial_y + \frac{1}{2}qh^2\partial_x^2 + \frac{1}{2}rh^2\partial_y^2 + qrh^2\partial_x\partial_y$$

en développant le logarithme de cette expression, on obtient

$$\begin{aligned} \tau \partial_t &= -h(q\partial_x + r\partial_y) \\ &+ h^2 \left(\frac{1}{2}q\partial_x^2 - (q\partial_x + r\partial_y) \left(\frac{1}{2}q\partial_x + \frac{1}{2}r\partial_y \right) + \frac{1}{2}r\partial_y^2 + qr\partial_x\partial_y \right) + O(h^3) \\ \tau \partial_t &= -\frac{1}{2}q^2\partial_x^2 + \frac{1}{2}q\partial_x^2 - \frac{1}{2}r^2\partial_y^2 + \frac{1}{2}r\partial_y^2 + O(h^3) \end{aligned}$$

L'équation équivalente s'écrit donc

$$\partial_t u + a\partial_x u + a\partial_y u = \frac{ah}{2} \left((1-q)\partial_{xx}^2 + (1-r)\partial_{yy}^2 \right) u \quad (5.87)$$

Sa solution présente une anisotropie moins marquée, et devient même isotrope pour une advection suivant la diagonale.

5.12.3 Fonctions de Green

Pour comparaison de la solution du schéma à la solution de l'équation équivalente, nous utilisons la fonction de Green pour la diffusion anisotrope, caractérisée par une matrice de diffusion D , donnée par exemple dans Chang, Kang, et Chen [4]

$$\frac{1}{4\pi |\det D|^{1/2} t} \exp \left(-\frac{1}{4t} (x - x(t), y - y(t)) D^{-1} \begin{pmatrix} x - x(t) \\ y - y(t) \end{pmatrix} \right)$$

Le centre $(x(t), y(t))$ est le point advecté, soit, par exemple, si on se place dans le cas d'une advection suivant la diagonale, (at, at) .

Se plaçant dans le repère tourné de $\pi/4$, l'approximation équation équivalente de la solution est donc, dans le cas non splitté, où

$$D = \frac{ah}{2} \begin{pmatrix} 1 - 2q & 0 \\ 0 & 1 \end{pmatrix}, D^{-1} = \frac{2}{ah} \begin{pmatrix} (1 - 2q)^{-1} & 0 \\ 0 & 1 \end{pmatrix}$$

$$\frac{1}{2\pi ah\sqrt{1-2q}} \exp \left[\left(-\frac{1}{2ah} \left((1-2q)^{-1}(X - \sqrt{2}at)^2 + Y^2 \right) \right) \right]$$

et, dans le cas splitté

$$\frac{1}{2\pi ah(1-q)} \exp \left(-\frac{1}{2ah(1-q)^2} ((X - \sqrt{2}at)^2 + Y^2) \right)$$

La figure 5.30 présente les fonctions de Green des deux schémas, en isocontours et niveaux de gris. Les isocontours sont elliptiques pour le schéma direct et circulaires pour le schéma splitté. En somme, le schéma splitté a, comme expliqué dans cette section, un comportement plus isotrope que le schéma direct.

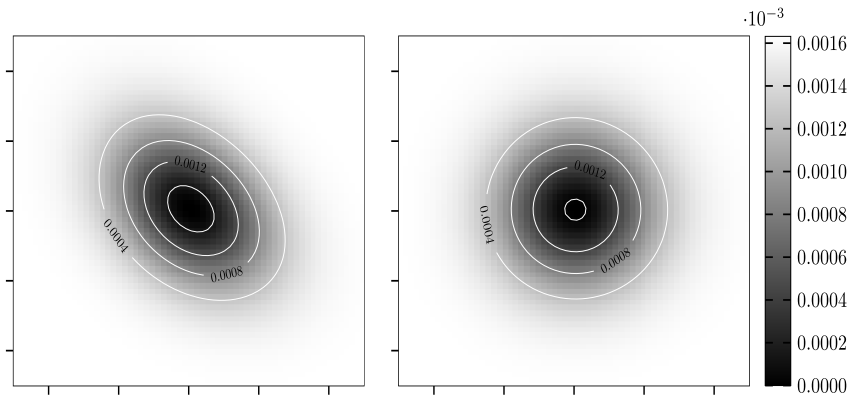


Fig. 5.30 Advection suivant la diagonale, fonction de Green du schéma direct et du schéma splitté.

5.12.4 Conclusion sur l'advection sur maillage carré

La résolution de l'équation d'advection à vitesse constante sur maillage carré avec le schéma upwind conduit, comme en 1D, à introduire de la diffusion, en général

anisotrope. Elle peut se quantifier soit par approximation de la solution exacte du schéma, soit par la méthode de l'équation équivalente. La diffusion induite par le schéma direct est plus anisotrope que celle du schéma splitté, ce qui est quelque peu contre-intuitif.

Nous poursuivrons cette partie consacrée aux applications en analysant quelques schémas discrétisant l'équation des ondes en 1D dans le chapitre 6 et en 2D dans le chapitre 7, puis celle de la chaleur dans le chapitre 8. Avant cela, nous présentons les références du chapitre 5.

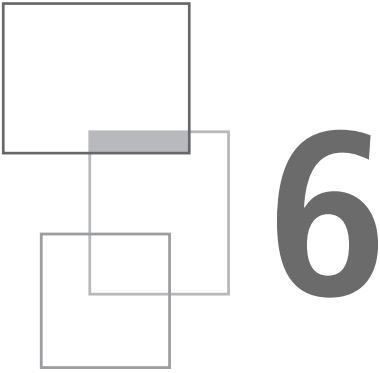
Le schéma upwind, très utilisé, est présenté dans la plupart des ouvrages consacrés aux méthodes numériques, par exemple [13], une interprétation probabiliste est détaillée dans [5]. L'advection d'une marche par un schéma non dissipatif (saute-mouton) est étudiée dans [1]. Les propriétés de ces schémas, notamment les bornes sur les solutions en norme L^p sont analysées d'un point de vue mathématique dans [10]. Les schémas de Beam–Warming et de Lax–Wendroff sont analysés à l'aide de l'équation équivalente dans [15, 16, 20] et leur combinaison pour obtenir des profils monotones pour l'advection d'une marche dans [2]. Les profils de chocs pour des schémas d'ordre élevé pour l'équation de Burgers sont prédits dans [3]. L'article [8] et les slides [7] de B. Després présentent les schémas de Strang d'un point de vue unifié. Le schéma de Balakine est un de ceux étudiés dans [24], cité en première partie.

Le livre de B. Gustafsson [12] présente les méthodes d'ordre élevé (>2) de manière systématique. Le rapport [9] analyse un ensemble de schémas d'ordre 3. Les cours en ligne [14, 19] présentent notamment l'advection sur maillage 2D.

Bibliographie

- [1] G. Arbia, D. Bouche “Closed Form Solution and Equivalent Equation Approximation of Linear Advection by a Non Dissipative Second Order Scheme for Step Initial Conditions” *Acta Applicandae Mathematicae* 130, No. 1, pp 151-162, 2014
- [2] J. Banks, T. Aslam, W. Rider “On sub-linear convergence for linearly degenerate waves in capturing schemes” *JCP* 227, pp 6985-7002, 2008
- [3] J. P. Boyd “Hyperviscous Shock Layers and Diffusion Zones Monotonicity, Spectral Viscosity, and Pseudospectral Methods for Very High Order Differential Equations” *Journal of Scientific Computing* 9, No. 1, 1994
- [4] Y. Chang, C. Kang, D. Chen “The use of fundamental Green Function for the solution of problems of heat conduction in anisotropic media” *J. Heat Mass Transfer*. 16, pp. 1905-1918, 1973
- [5] F. Delarue, F. Lagoutière “Probabilistic analysis of the upwind scheme for transport equation” *Archive for Rational Mechanics and Analysis* 199, pp. 229–268, 2011
- [6] S. Del Pino, B. Després, P. Havé, H. Jourdain, P.F. Piserchia “3D Finite Volume simulation of acoustic waves in the earth atmosphere” *Computers & Fluids* 38, pp 765-777, 2009

- [7] B. Després “Les schémas de Strang pour le transport linéaire. Histoire et Théorie” disponible sur le site du laboratoire J. L. Lions
- [8] B. Després “Finite Volume Transport Schemes” Num. Math. 108, pp 529-556, 2008
- [9] J. A. Desideri, A. Goudjo, V. Selmin “Third-order numerical schemes for hyperbolic problems” Rapport de recherche RR-0607, INRIA, 1987
- [10] D. Estep, M. Loss, J. Rauch “Boundedness of Dispersive Difference Schemes” Math. Comp. 55, No. 191 pp 55-87, 1990
- [11] W. Feller “An introduction to probability theory and its applications” Vol. 2, Wiley, 1971
- [12] B. Gustafsson “High Order Difference Methods for Time Dependent PDE” Springer, 2008
- [13] R. J. LeVeque “Finite Difference Methods for Ordinary and Partial Differential Equations Steady-State and Time-Dependent Problems” SIAM, 2007
- [14] Atmospheric and Oceanic Modeling, MIT Open Courseware, cours en ligne, 2004
- [15] Y. Shokin, I. Winnicki, J. Jasinski, S. Pietrek “High order modified differential equation of the Beam–Warming method, the dispersive features” Russ. J. Numer. Anal. Math. Modelling 35, No. 2, pp 83-94, 2020
- [16] Y. Shokin, I. Winnicki, J. Jasinski, S. Pietrek “High order modified differential equation of the Beam–Warming method, The dissipative features” Russ. J. Numer. Anal. Math. Modelling 35, No. 3, pp 1-11, 2020
- [17] R.J. Steinle “Finite difference methods for the advection equation” thèse, université d’Adélaïde, 1993
- [18] C. Tam “Computational Aeroacoustic, a wavenumber approach” Cambridge, 2011
- [19] G. Tryggvason “Introduction to computational fluid dynamics” cours en ligne, <http://www.nd.edu/~gtryggva/CFD-Course/> 2017
- [20] I. Winnicki, J. Jasinski, S. Pietrek “New approach to the Lax-Wendroff modified differential equation for linear and nonlinear advection” Numerical Methods for Partial Differential Equations 35, pp 2275–2304, 2019



Schémas pour l'équation des ondes 1D

6.1 Équation des ondes continue

Comme l'équation d'advection, l'équation des ondes est hyperbolique, mais, alors que l'équation d'advection est d'ordre 1 en temps et en espace, l'équation des ondes est d'ordre 2. L'équation des ondes en 1D peut s'écrire soit comme un système d'ordre 1

$$\begin{aligned}\partial_t u &= a \partial_x v \\ \partial_t v &= a \partial_x u\end{aligned}\tag{6.1}$$

soit comme une équation unique, obtenue en éliminant v du système (6.1) ci-dessus, en dérivant en t la première équation, en x la seconde. On obtient

$$\partial_{tt} u = a \partial_{xt} v, \partial_{tx} v = a \partial_{xx} u$$

ce qui implique $\partial_{tt} u = a^2 \partial_{xx} u$, soit l'équation des ondes (2.3). Considérons la différence $d = u - v$ et la somme $s = u + v$ des deux équations du système (6.1), on obtient le système

$$\begin{aligned}\partial_t d + a \partial_x d &= 0 \\ \partial_t s - a \partial_x s &= 0\end{aligned}$$

Donc d et s sont respectivement solutions d'une équation d'advection à vitesse a et $-a$. Cela justifie la dénomination « one way wave equation » utilisée pour l'équation d'advection. La solution du problème aux conditions initiales s'écrit

$$d(x, t) = u_0(x - at) - v_0(x - at)$$

$$s(x, t) = u_0(x + at) + v_0(x + at)$$

soit pour $u = \frac{1}{2}(s + d)$

$$u(x, t) = \frac{1}{2}(u(x + at) + u_0(x - at)) + \frac{1}{2}(v_0(x + at) - v_0(x - at))$$

On reconnaît la solution de d'Alembert de l'équation des ondes en 1D, avec conditions initiales u_0 sur la fonction et av'_0 sur sa dérivée en temps. En effet

$$u_t(x, t) = \frac{a}{2}(u'_0(x + at) - u'_0(x - at)) + \frac{a}{2}(v'_0(x + at) + v'_0(x - at)) \Rightarrow u_t(x, 0) = av'_0(x)$$

Il est possible de discrétiser soit le système (6.1), soit l'équation (2.3), ce qui conduit à différentes méthodes numériques, dont quelques-unes sont analysées ci-dessous.

6.2 Schémas dérivés des schémas pour l'équation d'advection

Tous les schémas étudiés dans le chapitre précédent pour l'équation d'advection peuvent être utilisés directement pour le système des ondes, en les appliquant avec la vitesse a sur $d = u - v$ et la vitesse $-a$ sur $s = u + v$. Les résultats obtenus pour l'ensemble des schémas étudiés dans le chapitre précédent se transposent donc immédiatement. Par exemple, le schéma upwind s'écrit sous forme système

$$\begin{aligned} d_j^{n+1} &= (1 - \eta)d_j^n + \eta d_{j-1}^n \\ s_j^{n+1} &= (1 - \eta)s_j^n - \eta s_{j+1}^n \end{aligned}$$

En appliquant les méthodes de la section 5.1 consacrée à ce schéma, on obtient aisément des solutions exactes ou approchées. Nous ne les présentons pas, par souci de concision. Nombre de schémas ont aussi été développés spécifiquement pour l'équation des ondes. Nous en étudions un des plus connus, le schéma de Yee.

6.3 Schéma de Yee en 1D

6.3.1 Formulation du schéma pour le système des ondes

Il a été proposé par S. Yee pour les équations de Maxwell. Il est décrit en détail dans le livre de Gustafson [12, chapitre 5], dont nous reprenons la présentation. Le schéma de Yee est un exemple de schéma décalé, dont la théorie est présentée dans l'Appendice sur les schémas d'ordre élevé. Le décalage en espace, qui consiste à définir certaines quantités aux centres des mailles, et d'autres aux cloisons, permet d'obtenir automatiquement l'ordre 2. Le décalage en temps suit la même idée en définissant certaines quantités aux temps entiers et d'autres aux temps demi-entiers.

Le schéma de Yee définit ainsi u aux cloisons (indice j) et aux temps entiers (indice n), v aux centres des mailles (indice $j + \frac{1}{2}$) et aux temps demi-entiers (indice $n + \frac{1}{2}$). Il est donc d'ordre 2 en espace et en temps. Il s'écrit

$$\frac{u_j^{n+1} - u_j^n}{\tau} = a \frac{v_{j+1/2}^{n+1/2} - v_{j-1/2}^{n+1/2}}{h}, \quad \frac{v_{j+1/2}^{n+1/2} - v_{j+1/2}^{n-1/2}}{\tau} = a \frac{u_{j+1}^n - u_j^n}{h} \quad (6.2)$$

Dans la première équation, les approximations de la dérivée temporelle $\frac{u_j^{n+1} - u_j^n}{\tau}$ et de la dérivée spatiale $\frac{v_{j+1/2}^{n+1/2} - v_{j-1/2}^{n+1/2}}{h}$ sont en $j, n + \frac{1}{2}$. En développant $u_j^{n+1} - u_j^n$ au point $t = (n + 1/2)\tau$, on obtient

$$\begin{aligned} u_j^{n+1} - u_j^n &= u_j^{n+1/2} + \frac{1}{2}\tau \partial_t u_j^{n+1/2} + \frac{1}{8}\tau^2 \partial_{tt} u_j^{n+1/2} + \frac{1}{48}\tau^3 \partial_{3t} u_j^{n+1/2} \\ &\quad - (u_j^{n+1/2} - \frac{1}{2}\tau \partial_t u_j^{n+1/2} + \frac{1}{8}\tau^2 \partial_{tt} u_j^{n+1/2} - \frac{1}{48}\tau^3 \partial_{3t} u_j^{n+1/2}) + o(\tau^3) \end{aligned}$$

$$\frac{u_j^{n+1} - u_j^n}{\tau} = \partial_t u_j^{n+1/2} + \frac{1}{24}\tau^2 \partial_{3t} u_j^{n+1/2} + o(\tau^2)$$

donc une approximation d'ordre 2 de $\partial_t u_j^{n+1/2}$. De même, en développant $v_{j+1/2}^{n+1/2} - v_{j-1/2}^{n+1/2}$ au point $x = jh$, on obtient

$$\frac{v_{j+1/2}^{n+1/2} - v_{j-1/2}^{n+1/2}}{h} = \partial_x v_j^{n+1/2} + \frac{1}{24}\tau^2 \partial_{3x} v_j^{n+1/2} + o(\tau^2)$$

La première équation de (6.2), que nous réécrivons

$$u_j^{n+1} - u_j^n = \eta \left(v_{j+1/2}^{n+1/2} - v_{j-1/2}^{n+1/2} \right),$$

où $\eta = \frac{a\tau}{b}$, est donc une approximation consistante à l'ordre 2 de l'équation continue $\partial_t u = a \partial_x v$. La seconde, qui s'écrit $v_{j+1/2}^{n+1/2} - v_{j-1/2}^{n+1/2} = \eta(u_{j+1}^n - u_j^n)$, est de même une approximation consistante à l'ordre 2 de l'équation continue $\partial_t v = a \partial_x u$. On notera que les premiers termes de l'erreur de troncature sont de nature dispersive, comme pour le schéma saute-mouton, dont dérive le schéma de Yee. Le schéma peut aussi s'interpréter comme une discrétisation de l'équation des ondes, comme expliqué ci-dessous.

6.3.2 Réinterprétation comme une discrétisation de l'équation des ondes

Il est possible d'éliminer v des équations (6.2), en écrivant

$$\begin{aligned} u_j^{n+1} - u_j^n - (u_j^n - u_j^{n-1}) &= \eta \left((v_{j+1/2}^{n+1/2} - v_{j-1/2}^{n+1/2}) - (v_{j+1/2}^{n-1/2} - v_{j-1/2}^{n-1/2}) \right) \\ &= \eta^2 ((u_{j+1}^n - u_j^n) - (u_j^n - u_{j-1}^n)) \end{aligned}$$

et comme

$$u_j^{n+1} - u_j^n - (u_j^n - u_j^{n-1}) = u_j^{n+1} - 2u_j^n + u_j^{n-1}$$

le schéma se réécrit sous la forme d'une seule équation aux différences partielles

$$u_j^{n+1} - 2u_j^n + u_j^{n-1} = \eta^2 (u_{j+1}^n - 2u_j^n + u_{j-1}^n) \quad (6.3)$$

On reconnaît au premier (resp. second) membre de (6.3) les approximations centrées d'ordre 2 des dérivées secondes temporelle (resp. spatiale) de u , donc de l'équation des ondes.

Solution par Fourier

En appliquant la transformée de Fourier semi-discrète sur l'indice spatial à l'équation (6.3), on obtient

$$\widehat{u}^{n+1} - 2\widehat{u}^n + \widehat{u}^{n-1} = -4\eta^2 \sin^2\left(\frac{k}{2}\right) \widehat{u}^n \quad (6.4)$$

soit une équation aux différences, d'ordre 2 comme celle obtenue plus haut pour le schéma saute-mouton. Ses solutions sont une combinaison linéaire des puissances solutions de l'équation

$$g - 2 + \frac{1}{g} = -4\eta^2 \sin^2\left(\frac{k}{2}\right)$$

soit de

$$\left(\sqrt{g} - \frac{1}{\sqrt{g}}\right)^2 - \left(2i\eta \sin\left(\frac{k}{2}\right)\right)^2 = 0$$

posant $\sqrt{g} = \lambda$, on obtient

$$\lambda^2 + 2i\eta \sin\left(\frac{k}{2}\right)\lambda - 1 = 0 \quad (6.5)$$

et sa conjuguée

$$\lambda^2 - 2i\eta \sin\left(\frac{k}{2}\right)\lambda - 1 = 0$$

On retrouve l'équation obtenue pour le schéma saute-mouton (section 5.3), *i.e.* $\lambda^2 + 2i\eta \sin(k)\lambda + 1 = 0$ à la substitution $k \rightarrow \frac{k}{2}$ près. L'équation conjuguée correspond à la vitesse $-a$. Comme déjà noté, les schémas de Yee pour les ondes et saute-mouton pour l'advection sont en étroite correspondance. L'équation des ondes, qui se factorise en $(\partial_t + a\partial_x)(\partial_t - a\partial_x)u = 0$, a les vitesses $\pm a$, d'où les deux équations conjuguées. Le remplacement $k \rightarrow \frac{k}{2}$ vient de ce que le schéma de Yee est un schéma décalé, alors que le saute-mouton est centré (voir chapitre 10). En posant comme pour le saute-mouton $\lambda = e^{-i\omega}$, l'équation (6.5) se réécrit $\sin(\omega) = \eta \sin\left(\frac{k}{2}\right)$, qui a deux solutions

$$\omega = \arcsin\left(\eta \sin\left(\frac{k}{2}\right)\right) \text{ et } \omega_p = \pi - \arcsin\left(\eta \sin\left(\frac{k}{2}\right)\right)$$

La seconde correspond au mode parasite observé sur le schéma saute-mouton. Mais ces deux solutions donnent la même solution de l'équation en $g = \lambda^2 = e^{-2i\omega}$. Le schéma de Yee ne présente donc pas de mode parasite. La solution en Fourier s'écrit

$$\widehat{u^n(k)} = A(k)e^{-2ni\omega} + B(k)e^{2ni\omega}$$

En revenant à l'espace réel, on obtient

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} A e^{-2ni\omega + ijk} dk + \frac{1}{2\pi} \int_{-\pi}^{\pi} B e^{2ni\omega + ijk} dk$$

Les coefficients sont déterminés par les conditions initiales. Comme le schéma est à 2 pas, il faut imposer ces conditions sur les pas de temps 0 et 1, ce qui s'écrit, en Fourier,

$$\widehat{u^0(k)} = A + B \text{ et } \widehat{u^1(k)} = A e^{-2i\omega} + B e^{2i\omega}$$

ce qui donne

$$A = \frac{\widehat{u}^0 e^{2i\omega} - \widehat{u}^1}{e^{2i\omega} - e^{-2i\omega}}, \quad B = \frac{\widehat{u}^1 - \widehat{u}^0 e^{-2i\omega}}{e^{2i\omega} - e^{-2i\omega}}$$

La condition initiale non nulle sur la fonction et nulle sur la dérivée en temps est simulée en imposant $\widehat{u}^0(k) = \widehat{u}^1(k)$, A, B sont déterminés par $\widehat{u}^0(k)$. Si cette quantité est réelle, $B = A^*$.

Tous les résultats de la section 5.3 sur le schéma saute-mouton se transposent directement, notamment l'approximation équation équivalente, obtenue en développant en $k = 0$ l'argument $-2ni \arcsin(\eta \sin(\frac{k}{2})) + ijk$ de l'exponentielle de l'intégrande, soit

$$-2ni \arcsin\left(\eta \sin\left(\frac{k}{2}\right)\right) + ijk = i(kj - nk\eta) + \frac{i}{24} k^3 n\eta (1 - \eta^2) + o(k^3) \quad (6.6)$$

Passons dans (6.6) aux variables continues $kj - nk\eta = k \frac{x-at}{b}$, $n\eta = \frac{at}{b}$

$$-2ni \arcsin\left(\eta \sin\left(\frac{k}{2}\right)\right) + ijk \approx k \frac{x-at}{b} + \frac{i}{24} k^3 \frac{at}{b} (1 - \eta^2)$$

Le premier terme de la solution équation équivalente s'écrit donc

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} A \exp\left(ik \frac{x-at}{b} + \frac{i(1-\eta^2)at}{24} \frac{k^3}{b}\right) dk = \frac{b}{2\pi} \int_{-\pi/b}^{\pi/b} A \exp\left(ik(x-at) + \frac{i(1-\eta^2) b^2 at}{24} k^3\right) dk$$

Le résultat est analogue à l'approximation de la solution par la méthode de l'équation équivalente à un terme pour le schéma saute-mouton appliqué à l'équation d'advection, mais le coefficient devant k^3 est $\frac{1}{24}$ au lieu de $\frac{1}{3}$. En effet, le schéma de Yee est un schéma décalé, d'où la substitution $k \rightarrow \frac{k}{2}$ par rapport au schéma saute-mouton. Les bornes de l'intégrale tendent vers l'infini quand $b \rightarrow 0$. Posons $c = \frac{1}{8} a (1 - \eta^2) b^2$

$$I \approx \frac{b}{2\pi} \int_{-\pi/b}^{\pi/b} A \exp\left(ik(x-at) + ict \frac{k^3}{3}\right) dk$$

Le second terme donne un résultat similaire, au changement de signe près de la vitesse a . Dans l'approximation équation équivalente à un terme, pour une condition initiale $u_j^0 = u_j^1 = \delta_j^0$, $\widehat{u}^0(k) = \widehat{u}^1(k) = 1$

$$A = \frac{e^{2i\omega} - 1}{e^{2i\omega} - e^{-2i\omega}} \approx \frac{e^{-i\eta k} - 1}{e^{-i\eta k} - e^{i\eta k}} \approx \frac{1}{2}, B = A^*$$

Remarque : A peut aussi s'écrire

$$A = \frac{e^{2i\omega} - 1}{e^{2i\omega} - e^{-2i\omega}} = \frac{e^{2i\omega} - 1}{(e^{i\omega} - e^{-i\omega})(e^{i\omega} + e^{-i\omega})} = \frac{e^{2i\omega}}{e^{2i\omega} + 1}$$

Dans cette approximation, la fonction de Green du schéma s'exprime donc à l'aide de deux fonctions d'Airy,

$$G_j^n \approx \frac{h}{2(ct)^{1/3}} \left(Ai \left(\frac{x - at}{(ct)^{1/3}} \right) + Ai \left(\frac{x + at}{(ct)^{1/3}} \right) \right) \quad (6.7)$$

avec $c = \frac{1}{8}a(1 - \eta^2)h^2$. Les résultats numériques obtenus pour la condition initiale $u_j^0 = u_j^1 = \delta_j^0$ sont présentés sur la figure 6.1.

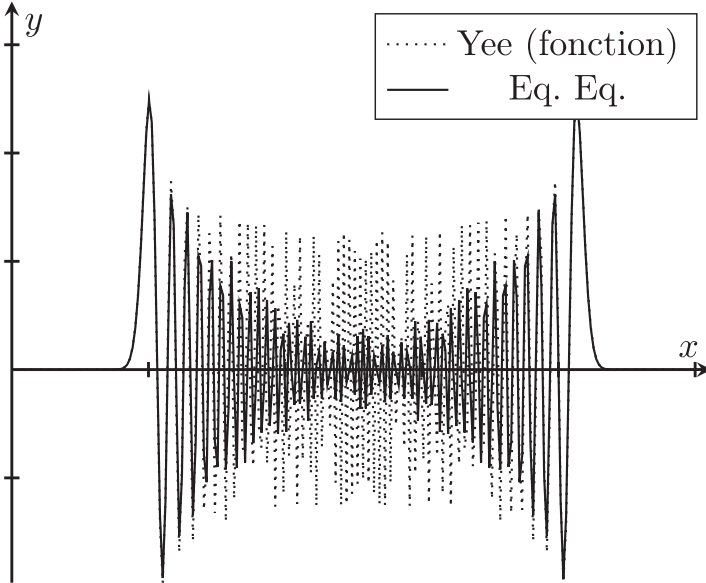
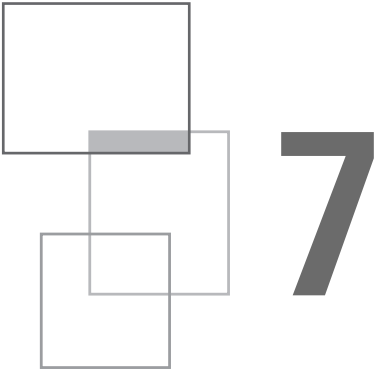


Fig. 6.1 Fonction de Green du schéma de Yee comparé à son approximation équation équivalente.

Deux signaux identiques se propagent, l'un à vitesse a , l'autre à vitesse $-a$. Au voisinage des fronts de propagation, en $\pm at$, ils sont correctement approchés par les fonctions d'Airy. Mais, cette approximation, comme dans le cas du schéma saute-mouton, sous-estime l'amplitude des oscillations de la solution du schéma quand on s'éloigne du front. Leur fréquence spatiale croît avec la distance au front.

6.3.3 *Conclusion*

En une dimension, il est possible d'obtenir des approximations explicites des solutions des schémas discrétisant l'équation des ondes, en particulier de leurs fonctions de Green. Il n'y a pas en effet de nouveauté significative par rapport à l'équation d'advection. Nous nous sommes donc limités à l'exemple du schéma de Yee. Les résultats sont proches de ceux obtenus pour les schémas d'ordre 2 non dissipatifs appliqués à l'équation d'advection, avec notamment la présence d'oscillations de l'origine aux fronts pour la fonction de Green du schéma. Nous passons maintenant au cas bidimensionnel.



Schémas pour l'équation des ondes en 2D

L'équation des ondes dans le cas multidimensionnel s'écrit, en présence d'une source $s(t)$

$$\partial_{tt}^2 u - a^2 \Delta u = s(t) \quad (7.1)$$

Toute une palette de schémas ont été proposés pour la résoudre. Nous nous focaliserons particulièrement sur l'anisotropie induite par le maillage, présente à des degrés divers dans tous les schémas. Pour la mettre en évidence, des solutions isotropes de l'équation, au sens où elles ne dépendent que de la distance r à l'origine, seront simulées. Une difficulté spécifique à l'équation des ondes bidimensionnelle apparaît : la fonction de Green, au sens de solution du problème avec conditions initiales Dirac sur la fonction, n'est pas définie.

7.1 Fonctions de Green pour l'équation des ondes continue

Pour comprendre cette difficulté, il est commode de se placer dans un cadre un peu plus général que le problème aux conditions initiales considéré jusque-là. Une méthode systématique a été proposée par Butkovitsky [2]. Elle consiste à définir une fonction w permettant de ramener le problème général avec second membre

et conditions initiales pour une EDP à un seul problème, celui de la source Dirac impulsive.

7.1.1 Source Dirac impulsive

En 2D, le problème (7.1) avec source Dirac impulsive et CI nulles s'écrit

$$\partial_{tt}^2 u - a^2(\partial_{xx}^2 u + \partial_{yy}^2 u) = \delta(x, y)\delta(t), \quad u(x, y, 0) = 0, \quad u_t(x, y, 0) = 0$$

Diverses méthodes de résolution ont été utilisées : descente à partir de la solution 3D, qui est plus simple, ou transformée de Fourier et Laplace, ce qui donne pour la double transformée :

$$s^2 \tilde{u} + a^2(k^2 + l^2) \tilde{u} = 1 \Rightarrow \tilde{u} = \frac{1}{s^2 + a^2(k^2 + l^2)}$$

En inversant la transformée de Laplace (σ est choisi pour que le contour soit à droite des singularités de l'intégrande)

$$\hat{u}(k, l, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{1}{s^2 + a^2(k^2 + l^2)} e^{st} ds$$

Il y a deux contributions de résidus, en $s = \pm ia\sqrt{k^2 + l^2}$

$$\hat{u}(k, l, t) = \frac{1}{2ia\sqrt{k^2 + l^2}} (\exp(iat\sqrt{k^2 + l^2}) - \exp(-iat\sqrt{k^2 + l^2}))$$

et, par transformation de Fourier inverse

$$u(x, y, t) = \frac{1}{4\pi^2 a} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{k^2 + l^2}} \sin(at\sqrt{k^2 + l^2}) \exp(ikx + ily) dkdl \quad (7.2)$$

En passant en coordonnées polaires dans (7.2), on obtient

$$\begin{aligned} u(r, t) &= \frac{1}{4\pi^2 a} \int_0^{\infty} \int_0^{2\pi} \sin(atk) \exp(ikr \cos \theta) dk d\theta \\ &= \frac{1}{2\pi a} \int_0^{\infty} \sin(atk) J_0(kr) dk \end{aligned} \quad (7.3)$$

L'intégrale en θ est donnée, par exemple dans les tables de Gradsteyn et Ryzhik, page 718

$$\frac{1}{2\pi a} \int_0^{\infty} \sin(atk) J_0(kr) dk = \frac{1}{2\pi a} (a^2 t^2 - r^2)^{-\frac{1}{2}} H(at - r)$$

ce qui conduit à la solution

$$u(r, t) = \frac{1}{2\pi a} (a^2 t^2 - r^2)^{-\frac{1}{2}} H(at - r) \quad (7.4)$$

7.1.2 Conditions initiales Dirac sur la dérivée

En appliquant la transformée de Fourier à l'équation des ondes $u_{tt} = a^2(u_{xx} + u_{yy})$, on obtient l'EDO $\widehat{u}_{tt} = -a^2(k^2 + l^2)\widehat{u}$, dont la solution s'écrit

$$\widehat{u} = A \cos(at\sqrt{k^2 + l^2}) + B \sin(at\sqrt{k^2 + l^2})$$

En imposant la condition initiale $u(t = 0) = 0$, $u_t(t = 0) = \delta(x, y)$, soit $\widehat{u}(0) = A = 0$, $\widehat{u}_t(0) = Ba\sqrt{k^2 + l^2} = 1$, on obtient $B = \frac{1}{a\sqrt{k^2 + l^2}}$, soit le résultat (7.2)

$$u(x, y, t) = \frac{1}{4\pi^2 a} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{k^2 + l^2}} \sin(at\sqrt{k^2 + l^2}) \exp(ikx + ily) dkdl$$

7.1.3 Conditions initiales Dirac sur la fonction

Formellement, en imposant la condition initiale $u(t = 0) = \delta(x, y)$, $u_t(t = 0) = 0$, soit $\widehat{u}(t = 0) = 1$, $\widehat{u}_t(t = 0) = 0$, on obtient $A = 1$ et $B = 0$, donc

$$u(x, y, t) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos(at\sqrt{k^2 + l^2}) e^{ikx + ily} dkdl$$

C'est la dérivée en temps de (7.2), comme on peut le voir en dérivant formellement en temps. Mais la dérivée de (7.4) n'est pas définie en $r = at$. En effet, les règles de dérivation usuelles donnent

$$\frac{d}{dt} (a^2 t^2 - r^2)^{-\frac{1}{2}} H(at - r) = a^2 t (a^2 t^2 - r^2)^{-\frac{3}{2}} H(at - r) + a (a^2 t^2 - r^2)^{-\frac{1}{2}} \delta(at - r)$$

Le second terme, produit d'un Dirac de support $r = at$ et d'une fonction singulière sur ce support, n'est pas défini.

Nous simulerons, en plus du problème avec conditions initiales sur la fonction, le problème avec conditions initiales sur la dérivée, dont la solution est donnée par (7.4). La condition initiale sur la fonction ou sa dérivée sont concentrées à l'origine, *i.e.*, en 2D, en $j = 0$, $m = 0$. Nous présentons les résultats pour la version semi-discretisée du schéma de Yee. Elle conduit à des résultats plus simples que le schéma

de Yee totalement discrétisé. Nous avons également analysé ce schéma. Les résultats sont similaires à ceux du semi-discrétisé, mais les formules deviennent nettement plus complexes. Nous avons choisi de ne pas les inclure dans cet ouvrage.

7.2 Schéma semi-discrétisé en 2D

Nous présentons dans ce paragraphe quelques résultats pour le schéma semi-discrétisé, qui s'écrit

$$\partial_{tt}^2 u_{jm} = \frac{a^2}{b^2} \left\{ (u_{(j+1)m} - 2u_{jm} + u_{(j-1)m}) + (u_{j(m+1)} - 2u_{jm} + u_{j(m-1)}) \right\} \quad (7.5)$$

En pratique, les simulations sont menées avec le schéma de Yee (dont la limite pour $\eta \rightarrow 0$ est le schéma semi-discrétisé) avec un η petit.

7.2.1 Solution exacte en Fourier

La solution de (7.5) s'obtient, comme usuellement, avec une TFSD, ce qui donne l'équation différentielle sur \hat{u}

$$\partial_{tt}^2 \hat{u} + 4 \frac{a^2}{b^2} \left(\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right) \right) \hat{u} = 0 \quad (7.6)$$

Pour une CI $\delta_j^0 \delta_m^0$ sur la fonction, nulle sur la dérivée, la solution de (7.6) est

$$\hat{u} = \cos \left(\frac{2at}{b} \sqrt{\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right)} \right) \quad (7.7)$$

et pour une CI $\delta_j^0 \delta_m^0$ sur la dérivée, nulle sur la fonction

$$\hat{u} = \frac{b}{2a \sqrt{\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right)}} \sin \left(\frac{2at}{b} \sqrt{\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right)} \right) \quad (7.8)$$

La CI $\delta_j^0 \delta_m^0$ sur la fonction est simulée numériquement en prenant comme condition, sur le pas de temps $n = 0$, $u_{00} = 1$ et, sur le pas de temps $n = 1$, $u_{00} = 1$. La CI $\delta_j^0 \delta_m^0$ sur la dérivée est simulée en prenant sur le pas de temps $n = 0$, $u_{00} = -\frac{\tau}{2}$ et, sur le pas de temps $n = 1$, $u_{00} = +\frac{\tau}{2}$. En posant $\omega = \sqrt{\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right)}$, et en appliquant la transformée de Fourier inverse à (7.7)

et (7.8), les solutions avec CI sur la fonction et sur la dérivée s'écrivent respectivement

$$u_{jm} = \frac{h}{8\pi^2 a} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} (e^{-i\frac{2at}{h}\omega + ijk + iml} - e^{+i\frac{2at}{h}\omega + ijk + iml}) dkdl \quad (7.9)$$

$$u'_{jm} = \frac{h}{8\pi^2 a} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{i}{\omega} (e^{-i\frac{2at}{h}\omega + ijk + iml} - e^{+i\frac{2at}{h}\omega + ijk + iml}) dkdl \quad (7.10)$$

7.2.2 Approximation par phase stationnaire

Comme le schéma est non dissipatif, nous appliquons la méthode de la phase stationnaire, comme nous l'avons fait pour le schéma saute-mouton, mais maintenant en 2 dimensions. Plus précisément, nous l'appliquons au premier terme des expressions (7.9) et (7.10), qui est de la forme

$$u_{jm}^+ = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} A e^{i\lambda S} dkdl$$

Le grand paramètre est $\lambda = \frac{at}{h}$. S s'écrit

$$S = -2\sqrt{\sin^2\left(\frac{k}{2}\right) + \sin^2\left(\frac{l}{2}\right)} + \frac{jh}{at}k + \frac{mh}{at}l \quad (7.11)$$

$A = \frac{i}{8\pi^2}$ pour (7.9) avec CI sur la fonction, $A = \frac{i}{8\pi^2 a \sqrt{\sin^2(\frac{k}{2}) + \sin^2(\frac{l}{2})}}$ pour (7.10)

avec CI sur la dérivée. Écrivons la formule (12.8), reprise du chapitre 12, donnant la contribution, pour $d = 2$, d'un point stationnaire k_s de composantes (k, l)

$$I \approx \frac{2\pi}{\lambda |\det(H(k_s))|^{1/2}} A(k_s) \exp(ikS(k_s) + i\frac{\pi}{4} \text{sgn}(H(k_s))) \quad (7.12)$$

La première étape est de trouver les points stationnaires, la seconde de calculer en ces points A et la hessienne H , matrice des dérivées secondes de S .

Détermination des points stationnaires

Les points stationnaires (k, l) vérifient

$$\frac{\partial S}{\partial k} = - \left(\frac{\sqrt{2}at \sin k}{2\sqrt{2 - \cos l - \cos k}} - \frac{hj}{at} \right) = 0$$

$$\frac{\partial S}{\partial l} = - \left(\frac{\sqrt{2}at \sin l}{2\sqrt{2 - \cos l - \cos k}} - \frac{hm}{at} \right) = 0$$

Posant $X = \frac{jb}{at}$, $Y = \frac{mb}{at}$, (k, l) sont donc solution du système

$$\frac{\sin k}{\sqrt{2 - \cos l - \cos k}} = \sqrt{2}X, \frac{\sin l}{\sqrt{2 - \cos l - \cos k}} = \sqrt{2}Y \quad (7.13)$$

Il peut se ramener, en écrivant la première équation de (7.13)

$$2 \sin \frac{k}{2} \cos \frac{k}{2} = \frac{x}{at} 2 \sqrt{\sin^2 \frac{k}{2} + \sin^2 \frac{l}{2}}$$

puis en l'élevant au carré, et en procédant de même pour la seconde, au système sur les inconnues $\kappa = \sin^2 \frac{k}{2}$ et $\lambda = \sin^2 \frac{l}{2}$

$$\frac{\kappa(1 - \kappa)}{\kappa + \lambda} = X^2, \frac{\lambda(1 - \lambda)}{\kappa + \lambda} = Y^2$$

reportant $\lambda = -\kappa - \frac{\kappa}{\xi} (\kappa - 1)$ dans la seconde équation, on obtient $\frac{\lambda(1-\lambda)}{\kappa+\lambda} = \chi$, soit l'équation de degré 4

$$\frac{(\kappa\xi + \kappa(\kappa - 1))((1 + \kappa)\xi + \kappa(\kappa - 1))}{\kappa(\kappa - 1)} = \chi\xi$$

qui n'a pas de solution explicite simple en général, mais se simplifie sur la diagonale $\lambda = \kappa$, $\lambda = 1 - \kappa$ et les axes $\lambda = 0, 1, \frac{l}{2} = 0, \frac{\pi}{2}$. Nous allons calculer la solution du schéma dans ces directions particulières.

7.2.3 Solution sur l'axe des x

Détermination des points stationnaires

Pour $Y = 0$, la seconde équation de (7.13) implique $\sin l = 0$, donc $l = 0$ ou $l = \pi$. Traitons d'abord le point stationnaire $(k, 0)$, pour lequel, puisque $\cos l = 0$, elle s'écrit

$$\frac{\sin k}{\sqrt{1 - \cos k}} = \sqrt{2}X \quad (7.14)$$

Point stationnaire $(k, 0)$

Pour $X > 0$, (7.14) implique que $\sin k > 0$ donc $k \in]0, \pi[$, $\cos \frac{k}{2} > 0$, donc k vérifie simplement $\cos \frac{k}{2} = X$ et

$$k = 2 \arccos X$$

Il est réel pour $X < 1$, i.e. $x < at$, pour calculer sa contribution, calculons le hessien H et la phase S .

Calcul de H

Calculons les dérivées secondes de S (notation $S_{kk} = \frac{\partial^2 S}{\partial k^2}$)

$$S_{kk} = \frac{\partial}{\partial k} \left(\frac{1}{2} \frac{\sqrt{2} \sin k}{\sqrt{2 - \cos l - \cos k}} \right) = \frac{1}{4} \frac{\sqrt{2}}{(2 - \cos l - \cos k)^{\frac{3}{2}}} \\ \times (2 \cos k \cos l - 4 \cos k + \cos^2 k + 1)$$

Comme $\cos l = 1$, $S_{kk} = \frac{1}{4} \frac{\sqrt{2}}{(1 - \cos k)^{\frac{3}{2}}} (-2 \cos k + \cos^2 k + 1)$

$$S_{kk} = \frac{1}{4} \sqrt{2} \sqrt{1 - \cos k} = \frac{1}{2} \sqrt{1 - X^2}$$

de même,

$$S_{kl} = \frac{1}{4} \sqrt{2} (\sin k) \frac{\sin l}{(2 - \cos l - \cos k)^{\frac{3}{2}}} = 0$$

$$S_{ll} = \frac{1}{4} \frac{\sqrt{2}}{(2 - \cos l - \cos k)^{\frac{3}{2}}} (2 \cos k \cos l - 4 \cos l + \cos^2 l + 1)$$

$$S_{ll} = -\frac{\sqrt{2}}{2} \frac{1}{(1 - \cos k)^{\frac{1}{2}}} = -\frac{1}{2\sqrt{1 - X^2}}$$

La matrice H a des valeurs propres S_{kk}, S_{ll} de signes opposés, donc $\text{sgn}(H) = 0$. Son déterminant vaut

$$\det H = S_{kk} S_{ll} = -\frac{1}{4}$$

Calcul de S

En prenant $l = 0$ dans (7.11), on obtient $S = -2\sqrt{\sin^2(\frac{k}{2})} + Xk$, et, comme $\cos \frac{k}{2} = X$

$$S = -2\sqrt{1 - X^2} + 2X \arccos X \quad (7.15)$$

Points stationnaires (k, π)

La seconde équation de (7.13), puisque $\cos l = \pi$, s'écrit $\frac{\sin k}{\sqrt{3-\cos^2 k}} = \sqrt{2}X$. Pour $X > 0$, $\sin k > 0$ donc $k \in]0, \pi[$. L'équation s'écrit aussi $1 - \cos^2 k = 2X^2(3 - \cos k)$. Si son discriminant est positif, soit pour $X < \sqrt{3 - 2\sqrt{2}} \approx 0,41421$, elle a 2 solutions $\cos k = \left(\pm \sqrt{-6X^2 + X^4 + 1} + X^2 \right)$. Il y a donc 2 valeurs de k : $\arccos \left(\sqrt{-6X^2 + X^4 + 1} + X^2 \right)$ et $\arccos \left(-\sqrt{-6X^2 + X^4 + 1} + X^2 \right)$. Mais pour $X > \sqrt{3 - 2\sqrt{2}}$, elle n'a pas de solution réelle : seule la contribution précédente existe. Comme notre objectif est de caractériser l'anisotropie, pour éviter de complexifier les formules, nous ne calculerons pas cette contribution.

CI $\delta_j^0 \delta_m^0$ pour la fonction

Nous nous limitons au cas $X > \sqrt{3 - 2\sqrt{2}}$, où seule la contribution du point $k = 2 \arccos X$ intervient. En reportant dans (7.12) les valeurs de $\det H = \frac{1}{4}$, de S donnée par (7.15), et de $A = \frac{1}{8\pi^2}$, on obtient l'approximation phase stationnaire

$$\begin{aligned} u_{jm}^+ &\approx \frac{h}{at} \frac{1}{2\pi} \exp \left(i \frac{at}{h} (-2\sqrt{1-X^2} + 2X \arccos X) \right) \\ u_{jm} &\approx \frac{h}{at} \frac{1}{\pi} \cos \left(\frac{at}{h} (-2\sqrt{1-X^2} + 2X \arccos X) \right) \end{aligned} \quad (7.16)$$

Comme on peut le voir sur la figure 7.1, l'accord est correct dans la zone $X > \sqrt{3 - 2\sqrt{2}}$, sauf au voisinage du front $X = 1$, où il faudrait une approximation uniforme. Pour $X < \sqrt{3 - 2\sqrt{2}}$, (7.16) ne peut pas être en accord avec la solution du schéma, puisque les points stationnaires (k, π) ne sont pas pris en compte.

CI $\delta_j^0 \delta_m^0$ pour la dérivée

Seul change, par rapport au paragraphe précédent

$$A = \frac{ih}{16\pi^2 a \sqrt{\sin^2 \left(\frac{k}{2} \right)}} = \frac{ih}{16\pi^2 a \sqrt{1-X^2}}$$

et on a maintenant

$$C^+ \approx \frac{h^2}{4\pi a^2 t} \frac{i}{\sqrt{1-X^2}} \exp \left(i \frac{at}{h} (-2\sqrt{1-X^2} + 2X \arccos X) \right)$$

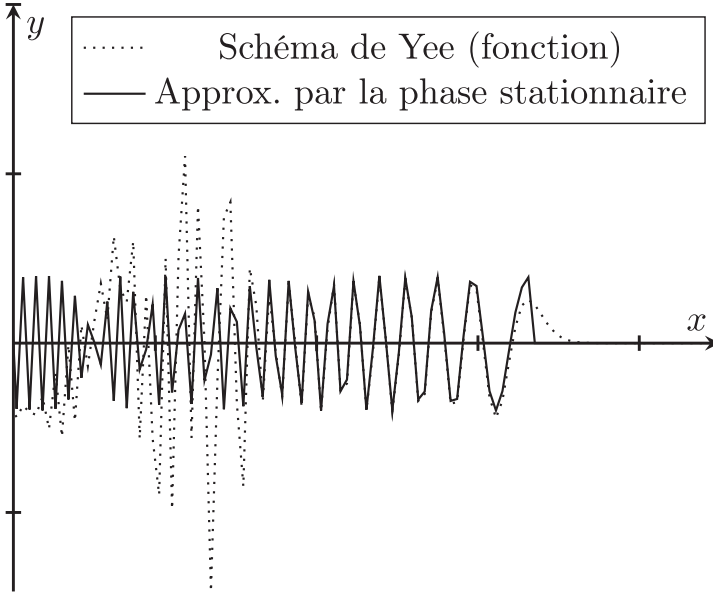


Fig. 7.1 Solution du schéma et approximation phase stationnaire (7.16).

Le calcul de l'approximation phase stationnaire de

$$u_{jm}^- = \frac{h}{8\pi^2 a} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{-i}{\omega} e^{i \frac{2at}{b} \omega + ijk + iml} dk dl$$

est analogue, ce qui donne

$$C^- \approx \frac{h^2}{4\pi a^2 t} \frac{-i}{\sqrt{1-X^2}} \exp\left(-i \frac{at}{b} (-2\sqrt{1-X^2} + 2X \arccos X)\right)$$

Ces contributions s'ajoutent, d'où la contribution phase stationnaire à u_{jm}

$$u_{jm}^{ps} \approx \frac{-h^2}{2\pi a^2 t} \frac{1}{\sqrt{1-X^2}} \sin\left(\frac{2at}{b} (-\sqrt{1-X^2} + X \arccos X)\right)$$

à comparer à la fonction de Green de l'équation pour $r < at$

$$G = \frac{1}{2\pi a} (a^2 t^2 - r^2)^{-\frac{1}{2}} = \frac{1}{2\pi a^2 t} \frac{1}{\sqrt{1-X^2}}$$

La contribution phase stationnaire a comme enveloppe la fonction de Green exacte, dans cette approximation. Mais il faut lui ajouter la contribution de la singularité en

$\frac{1}{\omega}$. Nous n'avons pas obtenu de formule exacte, et l'avons approchée par la fonction de Green de l'équation (multipliée par h^2), où seule cette contribution existe, ce qui donne, pour l'approximation de la solution du schéma

$$u_{jm} \approx \frac{h^2}{2\pi a^2 t} \frac{1}{\sqrt{1-X^2}} \left(\sin \left(\frac{2at}{b} (\sqrt{1-X^2} - X \arccos X) \right) + 1 \right) \quad (7.17)$$

La figure 7.2 compare la solution du schéma à cette approximation pour $X > 0$. L'accord dans la zone $X > \sqrt{3 - 2\sqrt{2}}$ est correct, en dépit d'une approximation *a priori* imprécise de la contribution de la singularité.

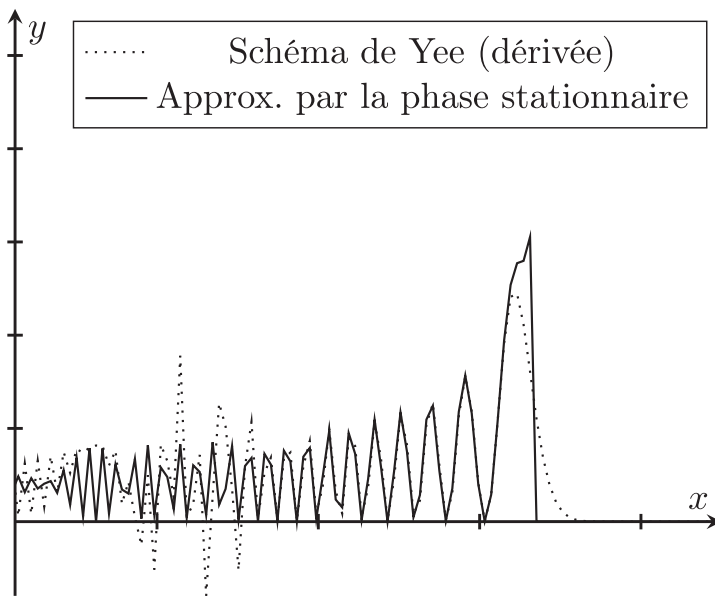


Fig. 7.2 Solution du schéma (obtenu avec un petit η) et approximation (7.17) de la fonction de Green de l'équation avec conditions initiales sur la dérivée.

7.2.4 Solution sur la diagonale $X = Y > 0$

7.2.4.1 Détermination des points stationnaires

Les points stationnaires vérifient (7.13), qui s'écrit

$$\frac{\sin k}{\sqrt{2 - \cos l - \cos k}} = \frac{\sin l}{\sqrt{2 - \cos l - \cos k}} = \sqrt{2}X$$

Pour $X > 0$, $\sin k > 0$, ces équations impliquent que $\sin l = \sin k$ donc $l = k$ ou $l = \pi - k$

Cas $l = \pi - k$

Comme $\cos l + \cos k = 0$, les équations se simplifient en

$$\frac{\sin k}{\sqrt{2}} = \sqrt{2}X, \frac{\sin l}{\sqrt{2}} = \sqrt{2}Y \rightarrow \sin k = \sin l = \sqrt{2}\rho$$

Ces points stationnaires $(\arcsin(\sqrt{2}\rho), \pi - \arcsin(\sqrt{2}\rho))$ $(\pi - \arcsin(\sqrt{2}\rho), \arcsin(\sqrt{2}\rho))$ n'existent que pour $\rho \leq \sqrt{2}/2$. Ils coalescent pour $\rho = \sqrt{2}/2$ en $(\frac{\pi}{2}, \frac{\pi}{2})$.

Cas $l = k$

On a maintenant $\cos l + \cos k = 2 \cos k$, donc

$$\frac{\sin k}{\sqrt{2} - 2 \cos k} = \sqrt{2}X$$

$$\frac{\sin k}{\sqrt{2} - 2 \cos k} = \frac{\sin k/2 \cos k/2}{\sqrt{\sin^2 k/2}} = \cos \frac{k}{2} \text{ donc}$$

$$\cos \frac{k}{2} = \sqrt{2}X = \frac{r}{at} = \rho$$

et le point stationnaire a pour coordonnée k

$$k = 2 \arccos \rho, \rho = \frac{r}{at} \quad (7.18)$$

Nous nous limiterons au cas $\rho > \sqrt{2}/2$, où n'intervient que ce point, pour lequel $l = k$.

Calcul du Hessien

Calculons S_{kk}

$$S_{kk} = \frac{1}{4} \frac{\sqrt{2}}{(2 - \cos l - \cos k)^{\frac{3}{2}}} (2 \cos k \cos l - 4 \cos k + \cos^2 k + 1)$$

mais $k = l$

$$S_{kk} = \frac{1}{4} \frac{\sqrt{2}}{(2 - 2 \cos k)^{\frac{3}{2}}} (3 \cos k - 1) (\cos k - 1) = -\frac{1}{8} \frac{3 \cos k - 1}{\sqrt{1 - \cos k}}$$

Le terme de dérivée croisée S_{kl} vaut

$$S_{kl} = \frac{1}{4} \sqrt{2} \frac{(1 - \cos k)(1 + \cos k)}{(2 - 2 \cos k)^{\frac{3}{2}}} = \frac{1}{8\sqrt{1 - \cos k}} (\cos k + 1)$$

et, comme $S_{ll} = S_{kk}$, la matrice hessienne s'écrit

$$H = \frac{1}{8\sqrt{1 - \cos k}} \begin{pmatrix} 1 - 3 \cos k & \cos k + 1 \\ \cos k + 1 & 1 - 3 \cos k \end{pmatrix}$$

Ses valeurs propres sont $2 - 2 \cos k > 0$, $-4 \cos k = -4(2\rho^2 - 1) < 0$ et, comme $\rho > \frac{1}{\sqrt{2}}$, $\text{sgn}H = 0$.

Calcul de det H et de S

Calculons les quantités apparaissant dans (7.12)

$$\det H = \frac{1}{64 (1 - \cos k)} ((3 \cos k - 1)^2 - (1 + \cos k)^2) = -\frac{1}{8} \cos k$$

$$\cos k = \cos(2 \arccos \rho) = 2\rho^2 - 1$$

$$S = -\frac{2at}{h} \omega + jk + ml, \quad jk = ml = \frac{\sqrt{2}}{2} \frac{r}{h} 2 \arccos \rho, \quad \omega = \sqrt{2} \sin \frac{k}{2} = \sqrt{2} \sqrt{1 - \rho^2}$$

$$S = -\frac{2\sqrt{2}at}{h} \sqrt{1 - \rho^2} + 2\sqrt{2} \frac{r}{h} \arccos \rho = \frac{2\sqrt{2}at}{h} (-\sqrt{1 - \rho^2} + \rho \arccos \rho)$$

Solution avec conditions initiales $\delta_j^0 \delta_m^0$ sur la fonction

Reportons dans (7.12)

$$|\det H(k)|^{-1/2} A(k) = \frac{\sqrt{2}}{4\pi^2 a \sqrt{|(2\rho^2 - 1)|}}$$

on obtient

$$u_{jm}^+ \approx \frac{h\sqrt{2}}{2\pi at} \frac{1}{\sqrt{2\rho^2 - 1}} \exp \left(i \frac{2\sqrt{2}at}{h} (-\sqrt{1 - \rho^2} + \rho \arccos \rho) \right)$$

$$u_{jm} \approx \frac{h\sqrt{2}}{\pi at} \frac{1}{\sqrt{2\rho^2 - 1}} \cos\left(\frac{2\sqrt{2}at}{h} \left(-\sqrt{1 - \rho^2} + \rho \arccos \rho\right)\right) \quad (7.19)$$

La figure 7.3 compare la solution du schéma à son approximation phase stationnaire (7.19) pour $\rho \in [0, 1]$. L'accord est bon dans la zone $\rho > \frac{1}{\sqrt{2}}$. Au voisinage de $\rho = \frac{1}{\sqrt{2}}$ le résultat (7.19), non uniforme, a une divergence qui se manifeste par l'augmentation de l'amplitude des oscillations quand on se rapproche de ce point. Ce phénomène ne se produit pas sur les axes (figure 7.1). Cette différence entre les axes et la diagonale est une illustration locale de l'anisotropie du schéma.

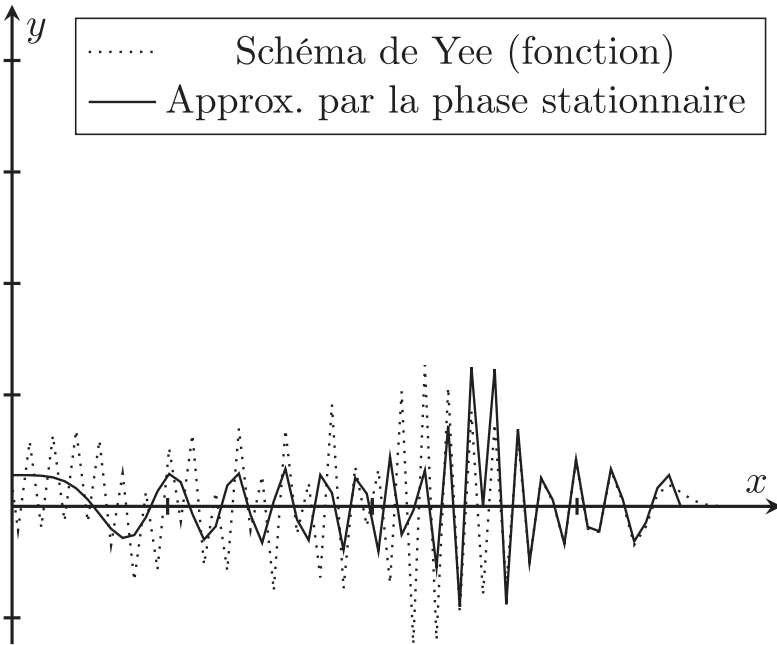


Fig. 7.3 Solution du schéma et approximation phase stationnaire (7.19).

CI $\delta_j^0 \delta_m^0$ pour la dérivée

On a dans ce cas

$$8\pi^2 aA = \frac{h}{\sqrt{\sin^2\left(\frac{k}{2}\right) + \sin^2\left(\frac{l}{2}\right)}} = \frac{h}{\sqrt{2 - 2\cos^2\left(\frac{k}{2}\right)}} = \frac{h}{\sqrt{2 - 2\rho^2}}$$

$$\begin{aligned} |\det H(k)|^{-1/2} A(k) &= \frac{2\sqrt{2}}{8\pi^2 a \sqrt{|1-2\rho^2|}} \frac{h}{\sqrt{2-2\rho^2}} \\ &= \frac{1}{4\pi^2 a} \frac{h}{\sqrt{|2\rho^2-1|}\sqrt{1-\rho^2}} \end{aligned}$$

ce qui conduit, pour la contribution phase stationnaire C^+ à

$$\begin{aligned} C^+ &\approx \frac{2\pi h}{at} |\det H(k)|^{-1/2} A(k) \exp\left(i \frac{at}{h} S(k)\right) \\ C^+ &\approx \frac{ib^2}{2\pi a^2 t} \frac{1}{\sqrt{1-\rho^2}\sqrt{2\rho^2-1}} \exp\left(i \frac{2\sqrt{2}at}{h} \left(-\sqrt{1-\rho^2} + \rho \arccos \rho\right)\right) \end{aligned}$$

La contribution phase stationnaire à la solution s'écrit, puisque $\text{sgn}(H) = 0$ si $\rho > \frac{1}{\sqrt{2}}$

$$u_{jm}^{ps} \approx \frac{-b^2}{\pi a^2 t} \frac{1}{\sqrt{(2\rho^2-1)\sqrt{1-\rho^2}}} \sin\left(\frac{2\sqrt{2}at}{h} \left(-\sqrt{1-\rho^2} + \rho \arccos \rho\right)\right) \quad (7.20)$$

L'enveloppe $\frac{1}{2\pi a^2 t} \frac{1}{\sqrt{1-\rho^2}}$ de la fonction de Green de l'équation n'est plus respectée.

Le résultat (7.20), non uniforme, comporte un facteur $\frac{1}{\sqrt{(2\rho^2-1)}}$. Comme dans la sous-section précédente, il faut ajouter à (7.20) la contribution de la singularité, approchée de la même manière. L'augmentation de l'amplitude des oscillations quand on se rapproche de ce point est visible sur la figure 7.4. Comme précédemment, l'accord est satisfaisant dans la zone $\left[\frac{1}{\sqrt{2}}, 1\right]$, sauf au voisinage des extrémités, car (7.20) n'est pas uniforme.

Conclusion

Notre objectif relativement modeste de montrer l'anisotropie du schéma semi-discrétisé localement, en analysant les différences entre les résultats de ce schéma sur les axes et sur la diagonale, est atteint : les oscillations générées par le schéma sont de plus grande amplitude sur la diagonale quand on se rapproche du point $\rho = \frac{1}{\sqrt{2}}$. Ce résultat est particulièrement visible par comparaison des figures 7.1 et 7.3, pour les conditions initiales sur la fonction. Nous nous sommes contentés de proposer des formules approchant la solution du schéma dans les seules zones où la contribution de phase stationnaire prend une forme simple. L'autre option, qui consiste à prendre en compte l'ensemble des contributions de phase stationnaire, et à utiliser des formules uniformes, serait préférable sur le plan théorique, mais conduit

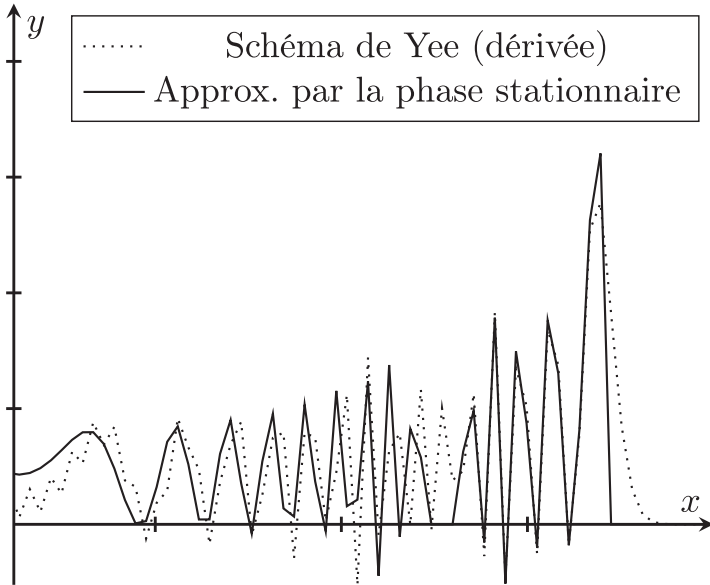


Fig. 7.4 Solution du schéma et approximation (phase stationnaire + singularité), pour $\rho > 0$.

à des résultats qui nous ont paru trop complexes pour figurer dans cet ouvrage. Nous allons maintenant caractériser plus globalement l'anisotropie du schéma totalement discrétisé.

7.3 Approximation équation équivalente

Dans ce paragraphe, nous utilisons l'approximation équation équivalente, valide au voisinage du front d'onde, pour étudier le schéma d'un point de vue global. Rappelons la forme de la solution pour la partie indicée + du schéma totalement discrétisé

$$u_{jm}^{n+} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} A \exp \left(-2ni \arcsin \left(\eta \sqrt{\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right)} \right) + ijk + iml \right) dkdl \quad (7.21)$$

Développons, pour obtenir l'approximation équation équivalente, le terme

$$-2n \arcsin \left(\eta \sqrt{\sin^2 \left(\frac{k}{2} \right) + \sin^2 \left(\frac{l}{2} \right)} \right)$$

apparaissant dans la phase de l'intégrande de (7.21) en $k, l = 0$. En développant les sinus

$$\sin^2\left(\frac{k}{2}\right) = \frac{1}{4}k^2 - \frac{1}{48}k^4 + O(k^5), \sin^2\left(\frac{l}{2}\right) = \frac{1}{4}l^2 - \frac{1}{48}l^4 + O(l^5)$$

on obtient

$$\sqrt{\sin^2\left(\frac{k}{2}\right) + \sin^2\left(\frac{l}{2}\right)} = \sqrt{k^2 + l^2 - \frac{1}{12}(k^4 + l^4) + O(l^5)}$$

Posons $k = s \cos \theta$, $l = s \sin \theta$

$$\begin{aligned} \sqrt{k^2 + l^2 - \frac{1}{12}(k^4 + l^4)} &= \sqrt{s^2 - \frac{1}{12}s^4(\cos^4 \theta + \sin^4 \theta)} \\ \sqrt{s^2 - \frac{1}{12}s^4(\cos^4 \theta + \sin^4 \theta)} &= s - \frac{s^3}{24}(\cos^4 \theta + \sin^4 \theta) + O(s^4) \end{aligned}$$

On obtient

$$\sqrt{\sin^2\left(\frac{k}{2}\right) + \sin^2\left(\frac{l}{2}\right)} \simeq \frac{1}{2} \left(s - \frac{s^3}{24}(\cos^4 \theta + \sin^4 \theta) \right)$$

puis

$$\arcsin\left(\frac{\eta}{2} \left(s - \frac{s^3}{24}(\cos^4 \theta + \sin^4 \theta) \right)\right) = \frac{1}{2}s\eta + \frac{1}{48}\eta s^3(\eta - (\cos^4 \theta + \sin^4 \theta)) + O(s^4)$$

Passons à l'approximation continue $n\eta = n\frac{a\tau}{b} = \frac{at}{b}$

$$-2n\eta \left(\frac{1}{2}s + \frac{1}{48}s^3(\eta^2 - (\cos^4 \theta + \sin^4 \theta)) \right) = -s\frac{at}{b} - \frac{1}{24}s^3\frac{at}{b}(\eta^2 - (\cos^4 \theta + \sin^4 \theta))$$

Considérons maintenant le terme $jk + ml$, passons à l'approximation continue :

$$j \simeq \frac{r \cos \varphi}{b}, m \simeq \frac{r \sin \varphi}{b}, n\eta \simeq \frac{t}{\tau} \frac{a\tau}{b} = \frac{at}{b}$$

$$jk + ml \simeq \frac{rs}{b} \cos(\theta - \varphi)$$

On obtient donc l'approximation de la phase de l'intégrande de (7.21)

$$\frac{s}{b}(r \cos(\theta - \varphi) - at) - \frac{1}{24}s^3\frac{at}{b}(\eta^2 - (\cos^4 \theta + \sin^4 \theta))$$

Dans l'approximation équation équivalente, la borne sur s est étendue à l'infini, si bien que, dans l'approximation équation équivalente

$$u_{jm}^{n+} \simeq \frac{1}{4\pi^2} \int_0^\infty \int_0^{2\pi} A \exp i \left(\frac{s}{b} (r \cos(\theta - \varphi) - at) - \frac{1}{24} s^3 \frac{at}{b} \right) \times \left(\eta^2 - \left(\cos^4 \theta + \sin^4 \theta \right) \right) s ds d\theta$$

L'imposition de la condition initiale, ou la comparaison à la fonction de Green de l'équation des ondes, permet de montrer que $A = \frac{1}{2}$. Au final, nous obtenons l'approximation de u_{jm}^{n+}

$$u_{jm}^{n+} \simeq \frac{1}{8\pi^2} \int_0^\infty \int_0^{2\pi} \exp i \left(\frac{s}{b} (r \cos(\theta - \varphi) - at) - \frac{1}{24} s^3 \frac{at}{b} \left(\eta^2 - \frac{3}{4} - \frac{1}{4} \cos 4\theta \right) \right) s ds d\theta \quad (7.22)$$

Nous n'avons pas trouvé d'approximation simple de cette intégrale, mais nous pouvons déduire de (7.22) des résultats qualitatifs. En particulier, le coefficient du terme dispersif induit, à cause du facteur $(\eta^2 - \frac{3}{4} - \frac{1}{4} \cos 4\theta)$, une anisotropie de période $\pi/2$. À noter que, dans le premier quadrant, respectivement sur l'axe des x et sur la diagonale,

$$u_{jm}^{n+} \simeq \frac{1}{8\pi^2} \int_0^\infty \int_0^{2\pi} \exp i \frac{at}{b} \left[s(\rho \cos(\theta) - 1) - \frac{1}{24} s^3 \left(\eta^2 - \frac{3}{4} - \frac{1}{4} \cos 4\theta \right) \right] s ds d\theta \quad (7.23)$$

$$u_{jm}^{n+} \simeq \frac{1}{8\pi^2} \int_0^\infty \int_0^{2\pi} \exp i \frac{at}{b} \left[s(\rho \cos(\theta) - 1) - \frac{1}{24} s^3 \left(\eta^2 - \frac{3}{4} + \frac{1}{4} \cos 4\theta \right) \right] s ds d\theta \quad (7.24)$$

Les coefficients du terme $\cos 4\theta$ dans (7.23) et (7.24) sont donc opposés : cela explique qualitativement les différences entre les axes et la diagonale analysées plus haut. Nous allons voir que les simulations numériques globales confirment aussi ces résultats. Nous avons calculé la solution du schéma approchant la fonction de Green précédente par la même technique qu'en 1D, *i.e.* en imposant $u_{jm}^0 = u_{jm}^1 = \delta_j^0 \delta_m^0$. L'anisotropie des résultats numériques, avec une période angulaire $\pi/2$, en lien avec le facteur $\frac{3}{4} + \frac{1}{4} \cos 4(\theta + \varphi)$ est visible sur la figure 7.5. Cette figure est divisée en deux parties : à gauche les amplitudes, à droite les lignes de niveau de la solution. Ces deux graphiques complémentaires permettent d'appréhender le comportement global du schéma, plus spécifiquement son anisotropie. Celui de gauche montre aussi que les oscillations de plus forte amplitude mise en évidence sur la diagonale sont limitées à un voisinage de cette dernière.

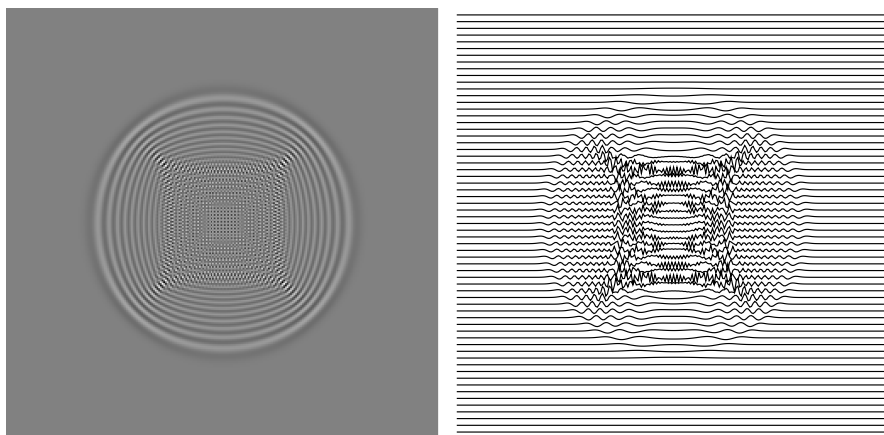


Fig. 7.5 *Solution numérique globale, amplitude et lignes de niveau.*

7.4 Conclusion

En plusieurs dimensions, un des intérêts de l'équation des ondes est d'avoir des solutions analytiques relativement simples à symétrie cylindrique en 2D, sphérique en 3D, utiles pour étudier l'anisotropie générée par le maillage. Nous nous sommes limités au schéma de Yee en version semi-discrétisée, pour obtenir des résultats explicites relativement simples. Nous avons mis en évidence l'anisotropie globalement : un facteur dépendant de l'angle apparaît dans la solution de l'équation équivalente, et localement : le schéma donne des résultats différents sur les axes et la diagonale, où il génère des oscillations de grande amplitude.

Nous nous sommes appuyés pour les schémas discrétisant l'équation des ondes sur les références listées ci-après. [2, 6] donnent des informations sur les fonctions de Green des EDP, dont l'équation des ondes. [1] présente des résultats pour des schémas décalés de différents ordres. Le livre [3] est consacré aux méthodes de Galerkin discontinu, associées à des discrétisations temporelles d'ordre élevé en temps.

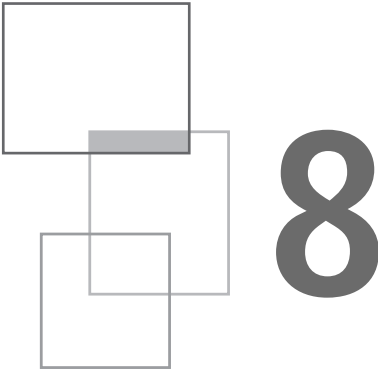
L'article [5] utilise des solutions exactes de l'équation des ondes, notamment les moyennes sphériques en 3D, et les compare à des solutions numériques pour illustrer les effets d'anisotropie, pour des schémas d'ordre 2 et 3.

Enfin, la formulation en système d'ordre 1, avec splitting par directions, permet d'utiliser les schémas de Strang d'ordre élevé pour obtenir des fronts très peu étalés, comme illustré dans l'article [4].

Dans le chapitre suivant, nous quittons le domaine des équations hyperboliques, pour nous intéresser à une équation parabolique simple, l'équation de la chaleur.

Bibliographie

- [1] J. W. Banks, W. D. Henshaw “Upwind schemes for the wave equation in second-order form” *Journal of Computational Physics* 231, pp 5854–5889, 2012
- [2] A. G. Butkovskiy “Green’s functions and transfer functions handbook” Halsted Press-Wiley, 1982
- [3] G. Cohen, S. Pernet “Finite Element and Discontinuous Galerkin Methods for Transient Wave Equations” Springer, 2016
- [4] S. Del Pino, B. Després, P. Havé, H. Jourdain, P.F. Piserchia “3D Finite Volume simulation of acoustic waves in the earth atmosphere” *Computers & Fluids* Vol. 38, pp 765–777, 2009
- [5] T. Eymann, P. Roe “Multidimensional Active Flux Schemes” 21th CFD Conference AIAA San Diego, 2013
- [6] S. Salsa “Partial Differential Equations in Action, from Modelling to Theory” Springer, 2008



Schémas pour l'équation de la chaleur

8.1 L'équation de la chaleur

L'équation de la chaleur, ou de diffusion, du premier ordre en temps, donc associée à une condition initiale u^0 , et du second ordre en espace est une équation parabolique. Rappelons son expression en 1D

$$\partial_t u = \alpha \partial_{xx} u$$

α est le coefficient de diffusion, il est en ms^{-1} . Par transformation de Fourier en x de l'équation et de la CI on obtient

$$\partial_t \widehat{u}(k, t) = -\alpha k^2 \widehat{u}(k, t), \quad \widehat{u}(k, 0) = \widehat{u}^0(k)$$

qui implique $\widehat{u} = \exp(-\alpha k^2 t) \widehat{u}^0$, d'où la solution du problème de Cauchy

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-\alpha k^2 t + ikx) \widehat{u}^0(k) dk$$

La fonction de Green est obtenue en prenant $\widehat{u}^0(k) = 1$

$$G_e = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-\alpha k^2 t + ikx) dk$$

C'est une intégrale gaussienne. La fonction de Green est donc

$$G_e = \frac{1}{2\sqrt{\pi\alpha t}} \exp\left(-\frac{x^2}{4\alpha t}\right)$$

Elle est indéfiniment dérivable. L'équation de la chaleur est régularisante, c'est-à-dire que la solution du problème aux conditions initiales, même non régulières, est régulière. C'est une différence essentielle par rapport aux équations hyperboliques, équation d'advection et équation des ondes, étudiées dans les chapitres précédents.

Du point de vue de l'analyse des schémas, les méthodes sont toutefois les mêmes que pour les équations hyperboliques. Le caractère régularisant pourrait faire penser que l'équation équivalente est suffisante pour une approximation précise de la solution du schéma. Mais ce n'est pas le cas : nous verrons que l'équation équivalente peut conduire à un problème mal posé. Nous commençons par le plus simple : le schéma explicite. Outre son intérêt théorique, la nature et la puissance des machines actuelles lui ont redonné vie dans les codes de calcul.

8.1.1 Schéma explicite

Ce schéma calcule u_j^{n+1} par la formule

$$\frac{u_j^{n+1} - u_j^n}{\tau} = \alpha \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2} \quad (8.1)$$

notant $\eta = \frac{\alpha\tau}{h^2}$, (8.1) s'écrit

$$u_j^{n+1} = (1 - 2\eta)u_j^n + \eta u_{j-1}^n + \eta u_{j+1}^n \quad (8.2)$$

Pour $\eta < 1/2$, le schéma est barycentrique, comme le schéma upwind pour l'équation d'advection, donc stable. Il est d'ordre 1 en temps et 2 en espace [4]. Toutes les méthodes d'analyse s'appliquent. Nous présentons les résultats, en commençant par l'analyse de Fourier.

Analyse de Fourier

Solution du schéma par transformation de Fourier

En Fourier, (8.2) devient

$$\widehat{u^{n+1}} = ((1 - 2\eta) + 2\eta \cos k) \widehat{u^n} = \left(1 - 4\eta \sin^2 \frac{k}{2}\right) \widehat{u^n}$$

Pour $\eta > 1/2$, le schéma est instable, le mode le plus instable est $k = \pi$, avec un facteur d'amplification $(1 - 4\eta)$. Pour $\eta < 1/2$, il est dissipatif d'ordre 2. Revenant dans l'espace réel par transformation de Fourier inverse, la solution du schéma s'écrit

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(1 - 4\eta \sin^2 \frac{k}{2}\right)^n e^{ijk} \widehat{u_0} dk$$

soit, pour une condition initiale $u_j^0 = \delta_j^0$

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(1 - 4\eta \sin^2 \frac{k}{2}\right)^n e^{ijk} dk \quad (8.3)$$

La fonction de Green du schéma G_s s'écrit donc (en faisant le changement de variable usuel $k' = \frac{k}{h}$ pour se comparer à fonction de Green G_e de l'équation)

$$G_s = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} \left(1 - 4\eta \sin^2 \frac{hk'}{2}\right)^n e^{ijk'} dk'$$

Convergence vers la fonction de Green de l'équation

La différence entre cette fonction de Green (divisée par h) et celle de l'équation est donc $h^{-1}G_s - G_e(jh, n\tau) = I_1 + I_2$, avec

$$I_1 = \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} \left(\left(1 - 4\eta \sin^2 \frac{kh}{2}\right)^n - \exp(-\alpha k^2 n\tau) \right) \exp(+ikjh) dk$$

$$I_2 = \frac{1}{\pi} \int_{-\pi/h}^{+\infty} \exp(-\alpha k^2 n\tau) \cos kx dk$$

Majorons le second terme I_2

$$I_2 = \frac{\sqrt{\alpha t}}{\pi} \int_{\pi\sqrt{\alpha n\tau}/h}^{+\infty} \exp(-k'^2) dk' < D \exp\left(-\frac{\pi^2 \alpha n\tau}{h^2}\right)$$

Le premier terme I_1 peut être majoré comme dans la démonstration du théorème de Lax-Richtmyer dans l'espace de Fourier donnée dans la première partie (section 3.3) de l'ouvrage. Les 2 points essentiels pour ce théorème sont que la différence entre le facteur d'amplification du schéma et celui de l'équation est $O(k^2)$ au voisinage de $k = 0$ (consistance) et que pour k grand, le spectre décroît si la condition initiale est régulière.

Le premier point est inchangé : la consistance du schéma à l'ordre 2 implique l'identité des développements de Taylor des deux facteurs d'amplification jusqu'à cet ordre

$$\exp[\ln(1 - 4\eta \sin^2 \frac{kh}{2})] = 1 - k^2 \alpha \tau + \frac{1}{12} \alpha \tau h^2 k^4 + O(h^4)$$

$$\exp(-\alpha k^2 \tau) = 1 - k^2 \alpha \tau + \frac{1}{2} \eta \alpha \tau h^2 k^4 + O(h^4)$$

Les premiers termes qui diffèrent dans les développements de Taylor sont en $\alpha \tau k^4 h^2$. Cela permet de majorer leur différence sur l'intervalle d'intégration, comme démontré dans Strikwerda [3].

$$\left| \left(1 - 4\eta \sin^2 \frac{kh}{2} \right) - \exp(-\alpha k^2 \tau) \right| \leq C \alpha \tau k^4 h^2$$

On en déduit

$$\begin{aligned} & \left| \left(1 - 4\eta \sin^2 \frac{kh}{2} \right)^n - \exp(-n\alpha k^2 \tau) \right| \\ & \leq C \alpha n \tau k^4 h^2 \max \left(\left(1 - 4\eta \sin^2 \frac{kh}{2} \right), \exp(-\alpha k^2 \tau) \right)^{n-1} \\ |I_1| & < \left| \frac{1}{2\pi} C \alpha t h^2 \int_{-\pi/h}^{\pi/h} k^4 \max \left(\left(1 - 4\eta \sin^2 \frac{kh}{2} \right), \exp(-\alpha k^2 \tau) \right)^{n-1} dk \right| \end{aligned}$$

Soit β tel que $(1 - 4\eta \sin^2 \frac{kh}{2}) \leq \exp(-\alpha \beta k^2 \tau)$. L'intégrale reste donc bornée pour $\frac{\pi}{h} \rightarrow \infty$, puisque l'intégrande est majorée par le produit de k^4 par un facteur $\exp(-\alpha \gamma k^2 \tau)$, $\gamma = \min(1, \beta)$. Posant $l = \sqrt{\alpha \tau'} k$, $\tau' = (n-1)\tau$

$$\int_{-\pi/h}^{\pi/h} k^4 [\exp(-\alpha \gamma k^2 \tau)]^{n-1} dk = \frac{1}{(\alpha \tau')^{\frac{5}{2}}} \int_{-\pi \sqrt{\alpha \tau'} / h}^{\pi \sqrt{\alpha \tau'} / h} l^4 \exp(-\gamma l^2) dl$$

On en déduit, puisque $\int_{-\infty}^{+\infty} l^4 \exp(-\gamma l^2) dl$ est fini

$$\left| h^{-1} G_s - G_e(jh, n\tau) \right| \leq C \frac{h^2}{(\alpha \tau)^{\frac{3}{2}}}$$

Autrement dit, la fonction de Green du schéma (divisée par h) converge, à t fixé, en h^2 vers celle de l'équation. Comme cette dernière est $O\left(\frac{1}{\sqrt{\alpha \tau}}\right)$, la différence relative

est en $\frac{b^2}{\alpha t} = \frac{b^2}{\alpha \tau n} = \frac{1}{n\eta}$. De ce point de vue, la convergence se produit pour n grand. Nous renvoyons au chapitre 10 de [3] et à [4] pour une analyse plus détaillée.

Pour l'équation d'advection, l'intégrale ne reste bornée que pour des conditions initiales dont le spectre en Fourier décroît suffisamment vite, donc assez régulières, ce qui exclut la fonction de Green. Les chapitres précédents ont d'ailleurs donné nombre d'exemples montrant sa non-convergence en norme L^∞ , et les artéfacts associés. Mais, dans le cas présent, la décroissance est assurée indépendamment de la régularité des conditions initiales, ce qui permet d'obtenir la convergence pour la fonction de Green. Il n'y a donc plus à proprement parler d'artéfacts numériques. Nous poursuivons toutefois, en restant focalisés sur l'établissement de résultats analytiques exacts ou approchés pour les fonctions de Green des schémas étudiés.

Cas particulier de la limite de stabilité

À la limite de stabilité (8.3) devient

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos k)^n e^{ijk} dk$$

Le facteur d'amplification ne peut plus être majoré comme plus haut. Le résultat précédent n'est plus valable dans ce cas. Il n'y a plus convergence, du moins au sens usuel.

Expression alternative de la solution

La variante transformée en z de la solution s'écrit pour mémoire :

$$u_j^n = \frac{1}{2i\pi} \int_C ((1 - 2\eta)z + \eta(z^2 + 1))^n z^{j-1-n} dz$$

Nous allons maintenant obtenir une approximation explicite de la représentation intégrale en Fourier pour n grand.

Approximation par la méthode du col

Réécrivons (8.3) sous une forme commode pour appliquer la méthode du col (section 12.3) pour n grand à l'intégrale (de la forme (12.9)) donnant u_j^n

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left(n \left(\ln(1 - 2\eta + 2\eta \cos k) + \frac{ijk}{n} \right) \right) dk \quad (8.4)$$

Nous traitons le cas $\eta < \frac{1}{2}$, où $1 - 2\eta + 2\eta \cos k = 1 - 4\eta \sin^2 \frac{k}{2} \geq 0$. Le grand paramètre est n . Les points cols de l'exposant $S = \ln(1 - 2\eta + 2\eta \cos k) + \frac{ijk}{n}$ vérifient

$$\frac{d}{dk} \ln((1 - 2\eta) + 2\eta \cos k) = -2\eta \frac{\sin k}{2\eta \cos k - 2\eta + 1} = -\frac{ij}{n}$$

Posons $z = e^{ik}$, $\gamma = \frac{1-2\eta}{2\eta} \geq 1$, $\frac{i}{n} = y$, l'équation précédente devient

$$\frac{z - \frac{1}{z}}{z + \frac{1}{z} + 2\gamma} = -y$$

z est solution de l'équation du second degré $(1 + y)z^2 + 2\gamma yz - (1 - y) = 0$. Son discriminant réduit est $D = (\gamma^2 - 1)y^2 + 1$. Comme $y^2 \leq 1$ et $\gamma > 0$, $D > 0$. Donc, pour $\eta < \frac{1}{2}$, il existe deux solutions réelles :

$$z_- = \frac{-\gamma y - \sqrt{(\gamma^2 - 1)y^2 + 1}}{1 + y}, \quad z_+ = \frac{-\gamma y + \sqrt{(\gamma^2 - 1)y^2 + 1}}{1 + y}$$

$-1 - 2\gamma \frac{i}{n} < z_- < 0$, $0 < z_+ \leq 1$. Il faut savoir laquelle doit être prise en compte. Calculons la dérivée seconde. On obtient

$$S'' = -\frac{2\eta((1 - 2\eta) \cos k + 2\eta)}{(2\eta \cos k - 2\eta + 1)^2} = -\frac{\gamma \cos k + 1}{(\cos k + \gamma)^2} \quad (8.5)$$

Au point z_+ , $\cos k > 0$, $S'' < 0$. Ce point col correspond donc à un maximum de l'intégrande et le développement asymptotique de l'intégrale doit être calculé à partir de sa contribution. L'autre point col correspond à un minimum de l'intégrande.

Dans l'espace des z , le point $z_+ \leq 1$ est à l'intérieur du cercle unité. Dans l'espace des k comme $e^{ik} = z_+ \leq 1$, $k = ik''$ est imaginaire positif $z_+ = e^{i(ik'')} = e^{-k''} < 1$. Le chemin de descente rapide passe par ce point ik'' , avec une tangente horizontale, puisque S'' est réel. Ce chemin n'est pas simple à déterminer exactement, mais, comme, si $|k| \rightarrow \infty$, $|\cos k|^2 = (\cos k' \cosh k'')^2 + (\sin k' \sinh k'')^2 \rightarrow \infty$, sauf si $k'' = 0$, ses seules asymptotes possibles sont les demi-droites de l'axe réel. La seule contribution est bien celle associée à z_+ , *i.e.* $k'' = -\ln z_+$.

Pour estimer u_j^n , il suffit de substituer dans la formule de l'appendice sur les développements asymptotiques d'intégrales, qui se réduit dans notre cas à

$$I \approx \frac{1}{2\pi} \frac{\sqrt{2\pi}}{\sqrt{-nS''}} e^{nS} \quad (8.6)$$

Dans (8.6), il faut substituer les valeurs obtenues en $k = -i \ln z_+$ pour $e^{nS} = (1 - 2\eta + 2\eta \cos k)^n z_+^j$ et pour S'' donné par (8.5), avec $\cos k = \frac{1}{2}(z_+ + \frac{1}{z_+})$.

La formule obtenue est peu parlante. Mais elle se simplifie au voisinage de l'origine. En effet, si $\frac{j}{n} = y$ est petit, on obtient, en développant en y , les approximations suivantes

$$j \ln z_+ \approx j \ln \left(1 - \frac{y}{2\eta} + \frac{y^2}{8\eta^2} \right) \approx -\frac{1}{2} \frac{y}{\eta} j = -\frac{ny^2}{2\eta}$$

$$1 - 2\eta + 2\eta \cos k = 1 + \frac{y^2}{4\eta} + O(y^3), \ln(1 - 2\eta + 2\eta \cos k) = \frac{y^2}{4\eta} + O(y^3)$$

$$\begin{aligned} nS &= n \ln(1 - 2\eta + 2\eta \cos k) + iny k \approx -\frac{ny^2}{4\eta} - \frac{nj^2 b^2}{4n^2 \alpha \tau} = -\frac{x^2}{4\alpha t} \\ -nS'' &\approx 2n\eta = 2n\alpha \tau h^{-2} = 2\alpha t h^{-2} \end{aligned}$$

En reportant ces approximations dans (8.6), on obtient donc, au voisinage de l'origine, $I \approx \frac{b}{2\sqrt{\pi\alpha t}} \exp(-\frac{x^2}{4\alpha t})$. C'est la fonction de Green G_e de l'équation, ce qui permet de vérifier que le résultat de la méthode du col est correct dans cette limite. En poussant le développement à un ordre de plus, on obtient

$$S \approx -\frac{1}{4} \frac{y^2}{\eta} - \frac{1}{32\eta^2} y^4 \left(\eta - \frac{1}{6} \right)$$

et donc le terme correctif sur S

$$-\frac{1}{32\eta^2} ny^4 \frac{6\eta - 1}{6\eta} = -\frac{n}{32\eta^3} \left(\eta - \frac{1}{6} \right) \left(\frac{j}{n} \right)^4$$

En revenant aux variables continues x et t , ce terme correctif s'écrit $-\frac{1}{32} \left(\eta - \frac{1}{6} \right) \frac{x^4 b^2}{\alpha^3 t^3}$. La fonction de Green du schéma, dans cette approximation, intègre donc un facteur multiplicatif de la forme

$$\exp \left(-\frac{1}{32} \left(\eta - \frac{1}{6} \right) \frac{x^4 b^2}{\alpha^3 t^3} \right) \quad (8.7)$$

Il est < 1 si $\eta > \frac{1}{6}$. Développons l'expression de S'' . On obtient $S'' = -2\eta + y^2 \frac{1}{4\eta} (6\eta - 1) + O(y^3)$, donc

$$\frac{1}{\sqrt{-S''}} = \frac{1}{\sqrt{2\eta}} + \frac{3}{8} \frac{y^2}{\eta^2 \sqrt{2\eta}} \left(\eta - \frac{1}{6} \right) + O(y^3)$$

ce qui introduit un terme multiplicatif

$$1 + \frac{3}{8} \frac{y^2}{\eta^2} \left(\eta - \frac{1}{6} \right) = 1 + \frac{3}{8} \left(\eta - \frac{1}{6} \right) \left(\frac{xb}{\alpha t} \right)^2 \quad (8.8)$$

Il est > 1 si $\eta < \frac{1}{6}$. Au final, en reportant (8.7) et (8.8) dans (8.6), la méthode du col fournit l'approximation suivante, valide au voisinage de l'origine

$$u_j^n \simeq bG_e \left(1 + \frac{3}{8} \frac{b^2}{t^2} \frac{x^2}{\alpha^2} \left(\eta - \frac{1}{6} \right) \right) \exp \left(-\frac{1}{32} \left(\eta - \frac{1}{6} \right) \frac{b^2 x^4}{(\alpha t)^3} \right) \quad (8.9)$$

Nous allons améliorer cette approximation en utilisant la méthode de Laplace. Avant de le faire, nous présentons le cas $\eta = \frac{1}{4}$.

Cas $\eta = \frac{1}{4}$

Les résultats de la méthode du col prennent une forme simple. En effet,

$$S = \ln \frac{1}{2} + \ln(1 + \cos k) + iyk = \ln \frac{1}{2} + 2 \ln \left(\cos \frac{k}{2} \right) + iyk$$

L'équation du CDR s'écrit, puisque $\text{Im} \cos(k + il) = -\sin k \text{Im}(\sin il)$
 $= -\sin k \text{Im} \left(\frac{\exp -l - \exp l}{2i} \right) = \sin k \sinh l$

$$\sin k \sinh l - 2l = 0$$

On obtient alors, puisque $\gamma = \frac{1-2\eta}{2\eta} = 1$, $z_+ = \frac{1-\gamma}{1+\gamma}$, $\cos k = \frac{1}{2} \left(z_+ + \frac{1}{z_+} \right) = \frac{1+\gamma^2}{1-\gamma^2}$, ce qui donne, $1 + \cos k = -\frac{2}{y^2-1}$. Les quantités S et S'' s'écrivent

$$S'' = -\frac{1}{1 + \cos k} = -\frac{1}{2}(1 - y^2)$$

$$\frac{1 + \cos k}{2} = \frac{1}{1 - y^2}$$

$$e^{nS} = \left(\frac{1 + \cos k}{2} \right)^n z_+^j = \frac{1}{(1 - y^2)^n} \frac{(1 - y)^j}{(1 + y)^j} = \frac{1}{(1 - y)^{n-j} (1 + y)^{n+j}}$$

ce qui permet d'écrire l'approximation méthode du col

$$G_s \approx \frac{1}{2\pi} \frac{\sqrt{2\pi}}{\sqrt{-nS''}} e^{nS} = \frac{1}{\sqrt{\pi n}} \frac{1}{\sqrt{1 - y^2}} \frac{1}{(1 - y)^{n-j} (1 + y)^{n+j}}$$

qui contrairement à la précédente, est valide sans la limite de proximité de l'origine. En $y = 0$, puisque $\eta = \frac{\alpha\tau}{b^2} = \frac{1}{4}$

$$G_s(0, n) \approx \frac{1}{\sqrt{\pi n}} = \frac{1}{2\sqrt{\pi \frac{t}{\tau} \frac{\alpha\tau}{b^2}}} = \frac{b}{2\sqrt{\pi\alpha t}} = G_e$$

Mais on a aussi

$$u_0^n = \frac{1}{\pi} \int_0^\pi \left(\cos \frac{k}{2}\right)^{2n} dk = \frac{2}{\pi} \int_0^{\pi/2} (\cos k)^{2n} dk = \frac{2}{\pi} W_{2n}$$

W_{2n} est une intégrale de Wallis, qui se calcule exactement.

$$W_{2n} = \frac{\pi}{2} \frac{(2n)!}{2^n (n!)^2}$$

Calcul par la méthode de Laplace

L'intégrale peut aussi s'écrire : $u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(n \ln(1 - 2\eta + 2\eta \cos k)) \cos(jk) dk$.

Elle est de la forme

$$I = \int_a^b \exp(-np(k)) q(k) dk \quad (8.10)$$

où $p = \ln(1 - 2\eta + 2\eta \cos k)$ et $q = \cos(jk)$ sont des fonctions réelles et peut être calculée par la méthode de Laplace. Comme $\cos(jk)$ est maintenant dans la fonction, et non dans l'exponentielle, cette méthode ne permet que de traiter le cas $j = O(1)$. Le point d'annulation de la dérivée de $p(k)$ est $k = 0$. L'intérêt de cette méthode est qu'il est plus facile d'écrire explicitement le développement asymptotique en puissances inverses du grand paramètre que dans la méthode du col. En se limitant aux deux premiers termes, (8.10) est approximée par

$$I \approx \sqrt{\frac{\pi}{np_2}} \exp(-np_0) \left(a_0 + \frac{1}{n} a_2 \left(\frac{1}{2}\right)_1 \right) \quad (8.11)$$

avec $\left(\frac{1}{2}\right)_1 = \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{1}{2})} = \frac{1}{2}$. Les quantités a_0 et a_2 s'écrivent à partir des coefficients des développements de Taylor de p et q : les formules, données dans [6, chapitre 12] s'écrivent

$$a_0 = q_0, \quad a_2 = \frac{1}{p_2} \left(q_2 - \frac{3}{2} \frac{p_4}{p_2} q_0 \right)$$

Le développement de p fournit $p_0 = 0$, $p_2 = \eta$, $p_4 = \frac{1}{2}\eta(\eta - \frac{1}{6})$, celui de q donne $q_0 = 1$, $q_2 = -\frac{1}{2}j^2$, $q_4 = +\frac{1}{24}j^4$. On obtient $a_0 = q_0 = 1$, $a_2 = -\frac{1}{8\eta}(4j^2 + 6\eta - 1)$. Reportant ces valeurs dans l'approximation (8.11) de I , on

obtient

$$\begin{aligned} I &\approx \sqrt{\frac{\pi}{n\eta}} \left(1 - \frac{1}{n} \frac{1}{16\eta} (4j^2 + 6\eta - 1) \right) \\ I &\approx \sqrt{\frac{\pi}{n\eta}} \left(1 - \frac{x^2}{4\alpha t} - \frac{3b^2}{8\alpha t} \left(\eta - \frac{1}{6} \right) \right) \end{aligned} \quad (8.12)$$

Dans la formule (8.12), le terme $\exp\left(-\frac{x^2}{4\alpha t}\right)$ de G_e apparaît sous la forme $1 - \frac{x^2}{4\alpha t}$. Au final, la méthode donne le résultat ci-dessous, valide dans un voisinage $O(b)$ de l'origine

$$u_j^n \approx bG_e \left(1 - \frac{3b^2}{8\alpha t} \left(\eta - \frac{1}{6} \right) \right) \quad (8.13)$$

Les deux méthodes sont complémentaires : la méthode de Laplace permet d'estimer la différence près de l'origine entre les deux fonctions de Green. Elle est, comme prévue par l'estimation ci-dessus, d'ordre $\frac{b^2}{\alpha t}$. Combinant (heuristiquement) les résultats des deux méthodes, *i.e.* les équations (8.9) et (8.13), on obtient l'approximation de la fonction de Green du schéma

$$G_s \simeq bG_e \left(1 - \frac{3b^2}{8\alpha t} \left(\eta - \frac{1}{6} \right) \right) \left(1 + \frac{3b^2}{8t^2} \frac{x^2}{\alpha^2} \left(\eta - \frac{1}{6} \right) \right) \exp \left(-\frac{1}{32} \left(\eta - \frac{1}{6} \right) \frac{b^2 x^4}{(\alpha t)^3} \right) \quad (8.14)$$

Cette équation s'écrit aussi, notant $b_e = \frac{1}{6} - \eta$

$$G_s \simeq bG_e \left(1 + \frac{3b^2}{8\alpha t} b_e \right) \left(1 - \frac{3b^2}{8t^2} \frac{x^2}{\alpha^2} b_e \right) \exp \left(+\frac{1}{32} b_e \frac{b^2 x^4}{(\alpha t)^3} \right)$$

Nous allons maintenant utiliser la méthode de l'équation équivalente, et obtenir à nouveau les résultats précédents.

Équation équivalente

Développons, pour obtenir l'approximation de l'équation équivalente, l'exposant dans (8.3) en $k = 0$

$$\ln \left(1 - 4\eta \sin^2 \frac{k}{2} \right) = -k^2 \eta + \frac{1}{2} k^4 \eta \left(\frac{1}{6} - \eta \right) + O(k^5)$$

En substituant ce développement dans la représentation intégrale de u_j^n , et en passant aux variables continues, on obtient

$$u(x, t) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left(-\frac{\alpha t}{b^2} k^2 - \frac{1}{2} \frac{\alpha t}{b^2} \left(\eta - \frac{1}{6} \right) k^4 + i \frac{x}{b} k \right) dk \quad (8.15)$$

avec la variable d'intégration $k' = k/h$,

$$u(x, t) \approx \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} \exp\left(-\alpha tk'^2 - \frac{1}{2}\alpha tb^2\left(\eta - \frac{1}{6}\right)k'^4 + ixk'\right) dk' \quad (8.16)$$

Il est aussi possible de déterminer l'équation équivalente par calcul formel. Le schéma s'écrit en effet, sous forme opératorielle

$$e^{\tau \frac{\partial}{\partial t}} = (1 - 2\eta) + 2\eta \cosh\left(h \frac{\partial}{\partial x}\right) \Rightarrow \tau \frac{\partial}{\partial t} = \ln\left((1 - 2\eta) + 2\eta \cosh\left(h \frac{\partial}{\partial x}\right)\right)$$

Par développement de Taylor de cette expression, on obtient

$$\frac{\partial}{\partial t} \approx \alpha \partial_{xx} + \frac{1}{2}\alpha h^2 \left(\frac{1}{6} - \eta\right) \partial_{4x}$$

L'équation équivalente à un terme s'écrit donc

$$\partial_t u - \alpha \partial_{xx} u = \frac{1}{2}\alpha h^2 \left(\frac{1}{6} - \eta\right) \partial_{4x} u \quad (8.17)$$

Sa solution est donnée par l'intégrale (8.16), avec les bornes à l'infini. Le schéma est d'ordre 2 en espace. Le premier terme de l'équation équivalente est un terme de dissipation d'ordre 4, si $\eta > \frac{1}{6}$, d'antidissipation d'ordre 4, si $\eta < \frac{1}{6}$. Son coefficient fait apparaître $b_e = \frac{1}{6} - \eta$, paramètre caractérisant le schéma. Dans le cas $\eta = 1/6$, ce paramètre, et donc le second membre de 8.17 est nul. Le schéma est en effet d'ordre 4. Le premier terme de son équation équivalente est alors en h^4 .

Cas $\eta < \frac{1}{6}$

Un problème apparaît dans le cas $\eta < \frac{1}{6}$, l'antidissipation rend le problème de l'équation équivalente mal posé. L'intégrale n'est pas définie.

Cas $\eta > \frac{1}{6}$

Le terme de dissipation d'ordre 4 est négatif pour $\eta > 1/6$, donc le problème redevient bien posé, et l'intégrale a un sens. Mais une nouvelle difficulté apparaît. Pour les schémas traitant l'équation d'advection analysés dans le chapitre 5, le signe de ce terme permet de prévoir le seuil de stabilité du schéma. Si nous appliquons ce critère de stabilité au cas présent, nous déduisons que le schéma explicite est stable $\forall \eta > 1/6$. Mais ce n'est bien sûr pas le cas : le schéma devient instable pour $\eta > 1/2$. Le mode le plus instable est $k = \pi$, donc n'est pas à basse fréquence spatiale (*i.e.* $k \ll 1$) dans la zone de validité de la méthode.

Cette question a été étudiée avec précision dans l'article [2], auquel nous renvoyons le lecteur intéressé. Le point essentiel, évoqué aussi dans la section 4.3 du chapitre 4 traitant de la méthode de l'équation équivalente, est que, pour que cette méthode soit valide, et prédise notamment la stabilité d'un schéma, il faut que le rayon de convergence R_c de la série donnant le facteur d'amplification soit $> \pi$, ce qui n'est pas le cas pour $\eta = 1/2$. Il vaut $\frac{\pi}{2}$, comme expliqué dans [2].

Le cas $\eta = 1/4$ est étudié en détail dans cette référence. La série s'exprime dans ce cas explicitement à l'aide des nombres de Bernoulli $\cos^2 \frac{k}{2} = \sum_{p=1}^{\infty} (-1)^{p+1} b_p k^p$, avec

$b_p = \frac{2^{2p}-1}{p(2p)!} B_{2p}$. La limite $\lim_{p \rightarrow \infty} (b_p)^{\frac{1}{p}} = \frac{1}{\pi^2}$ permet de conclure que $R_c = \pi$. La série converge sur tout le spectre.

L'équation équivalente doit donc prédire correctement la fonction de Green du schéma dans ce cas. Une adaptation de la démonstration de la convergence de la fonction de Green du schéma vers celle de l'équation explicitée en début de chapitre permettrait de prouver que la différence entre ces deux fonctions est $O(h^4)$. Plutôt que de donner une démonstration de ce résultat, nous allons estimer la fonction de Green G_{EE} de l'équation équivalente. Le but est de retrouver plus simplement les résultats de la section précédente. Partons de l'expression (8.16) de G_{EE}

$$G_{EE} = \frac{h}{2\pi} \int_{-\pi/h}^{\pi/h} \exp\left(-\alpha tk^2 - \frac{1}{2}\alpha th^2\left(\eta - \frac{1}{6}\right)k^4 + ixk\right) dk$$

Cette intégrale n'a pas d'expression explicite, mais on peut lui appliquer la méthode du col. Le grand paramètre est αt . Les points col vérifient

$$k + h^2\left(\eta - \frac{1}{6}\right)k^3 = i\frac{x}{2\alpha t}$$

Sans le terme cubique,

$$k = \frac{ix}{2\alpha t}, -\alpha tk^2 + ixk = -\frac{x^2}{4\alpha t}, e^{nS} = \exp\left(-\frac{x^2}{4\alpha t}\right)$$

Comme $S'' = -2\alpha t < 0$, $u_j^n \approx \frac{h}{2\sqrt{\pi\alpha t}} \exp\left(-\frac{x^2}{4\alpha t}\right)$. On retrouve la fonction de Green de l'équation de la chaleur, avec le facteur h usuel. Si on traite le terme cubique en perturbation, ce qui est licite au voisinage de l'origine,

$$k \approx \frac{ix}{2\alpha t} \left(1 + \left(\eta - \frac{1}{6}\right)\left(\frac{xh}{2\alpha t}\right)^2\right)$$

Il suffit de remplacer $S = -(k^2 + \frac{1}{2}h^2(\eta - \frac{1}{6})k^4)$ et S'' par leurs développements de Taylor dans la formule $I \approx \frac{\sqrt{2\pi f}}{\sqrt{-nS''}} e^{nS}$ pour obtenir l'approximation méthode du col de G_{EE}

$$G_{EE} \simeq hG_e \left(1 + \frac{3}{8} \frac{h^2}{t^2} \frac{x^2}{\alpha^2} \left(\eta - \frac{1}{6} \right) \right) \exp \left(-\frac{1}{32} \left(\eta - \frac{1}{6} \right) \frac{b^2 x^4}{(\alpha t)^3} \right)$$

On retrouve, comme attendu, l'équation (8.9) obtenue précédemment, avec des calculs plus simples.

Mais une approximation plus précise peut être obtenue. Nous nous sommes limités à calculer le premier terme du développement asymptotique pour αt grand par la méthode du col, alors que le second terme de ce développement est du même ordre. Plutôt que de calculer explicitement ce terme dans le cas général, nous allons le calculer en $x = 0$, ce qui permettra de déterminer l'écart entre les deux fonctions de Green à leur maximum. Les termes multiplicatifs précédents valent en effet 1 en $x = 0$. Calculons donc, dans le cas $\frac{1}{6} < \eta \leq \frac{1}{4}$, où l'équation équivalente est valide et le schéma stable

$$u_0^n \simeq \frac{b}{2\pi} \int_{-\infty}^{\infty} \exp \left(-\alpha t k^2 - \frac{1}{2} \alpha t h^2 \left(\eta - \frac{1}{6} \right) k^4 \right) dk \quad (8.18)$$

La méthode du col se réduit dans ce cas à la méthode de Laplace. Le point critique est en $k = 0$. On obtient, si on ne prend que le premier terme, $S(0) = 0$, $S''(0) = -2$, ce qui entraîne $u_0^n \approx \frac{b}{2\sqrt{\pi\alpha t}} \exp \left(-\frac{x^2}{4\alpha t} \right)$. Le résultat est indépendant du coefficient du terme en k^4 , ce qui confirme son imprécision. Comme vu plus haut, il faut prendre en compte le terme suivant. Posons $c = \frac{1}{2} \frac{h^2}{\alpha t} \left(\eta - \frac{1}{6} \right)$ et passons à la variable $l = \sqrt{\alpha t} k$ dans l'intégrale (8.18) approchant u_0^n

$$u_0^n \simeq \frac{b}{\pi \sqrt{\alpha t}} \int_0^{\infty} \exp(-l^2 - cl^4) dl$$

Cette intégrale s'exprime à l'aide de fonctions de Bessel K , mais nous allons nous contenter d'une approximation linéaire. Calculons la dérivée de $I(bc) = \int_0^{\infty} \exp(-k^2 - ck^4) dk$ en $b = 0$

$$\frac{d}{db} \int_0^{\infty} \exp(-k^2 - ck^4) dk \Big|_{b=0} = - \int_0^{\infty} k^4 \exp(-k^2) dk = -\frac{3}{8} \sqrt{\pi}$$

et approximations linéairement : $I(b) \approx I(0) - \frac{3}{8} \sqrt{\pi} b$. On obtient finalement le résultat (8.13) de la section précédente

$$u_0^n \simeq \frac{b}{2\sqrt{\pi\alpha t}} \left(1 - \frac{3}{8} \frac{b^2}{\alpha t} \left(\eta - \frac{1}{6} \right) \right)$$

Regroupant les résultats précédents, on obtient à nouveau l'équation (8.14), certes, par une méthode non rigoureuse, puisque le terme correctif, qui correspond au second terme du développement asymptotique, a été calculé seulement en $x = 0$.

$$G_{EE} \simeq bG_e \left(1 - \frac{3}{8} \frac{b^2}{\alpha t} \left(\eta - \frac{1}{6} \right) \right) \left(1 + \frac{3}{8} \frac{b^2}{t^2} \frac{x^2}{\alpha^2} \left(\eta - \frac{1}{6} \right) \right) \exp \left(-\frac{1}{32} \left(\eta - \frac{1}{6} \right) \frac{b^2 x^4}{(\alpha t)^3} \right) \quad (8.19)$$

Nous obtenons donc le même résultat, pour n grand, pour la fonction de Green G_{EE} de l'équation équivalente et celle G_s du schéma. Qualitativement, cette fonction de Green est, au moins dans un voisinage de l'origine, au-dessus de celle de l'équation de la chaleur si $\eta < \frac{1}{6}$, au-dessous si $\eta > \frac{1}{6}$. On notera que tous les termes correctifs s'expriment à l'aide de $b_e = \frac{1}{6} - \eta$, paramètre caractéristique du schéma, qui est déterminé par le premier terme de l'équation équivalente.

La figure 8.1 présente les résultats numériques obtenus avec $\eta = \frac{1}{4}$. La figure compare la fonction de Green G_s du schéma à l'approximation (8.19). Elle est quasiment superposée à G_s . Nous avons donc aussi tracé leur différence, de l'ordre de quelques 10^{-8} .

Nous avons obtenu dans ce paragraphe, par une méthode heuristique, associant la méthode du col et la méthode de Laplace, une approximation de la fonction de Green G_s du schéma valide pour n grand. Nous avons ensuite obtenu la même approximation pour la fonction de Green G_{EE} . La comparaison aux simulations numériques montre que cette approximation est précise. Nous allons maintenant utiliser la méthode opératorielle pour obtenir un résultat exact.

Méthode opératorielle

Le schéma s'écrit, sous forme opératorielle

$$u_j^{n+1} = (1 - 2\eta)u_j^n + \eta u_{j-1}^n + \eta u_{j+1}^n = ((1 - 2\eta)I + \eta(S + S^{-1}))u_j^n$$

On obtient donc

$$u_j^n = ((1 - 2\eta)I + \eta(S + S^{-1}))^n u_j^0 = \sum_{m,p} (1 - 2\eta)^{n-m-p} \eta^{m+p} \frac{n!}{(n-m-p)!m!p!} u_{j+m-p}^0$$

Pour la fonction de Green, les termes non nuls vérifient $p = j + m$,

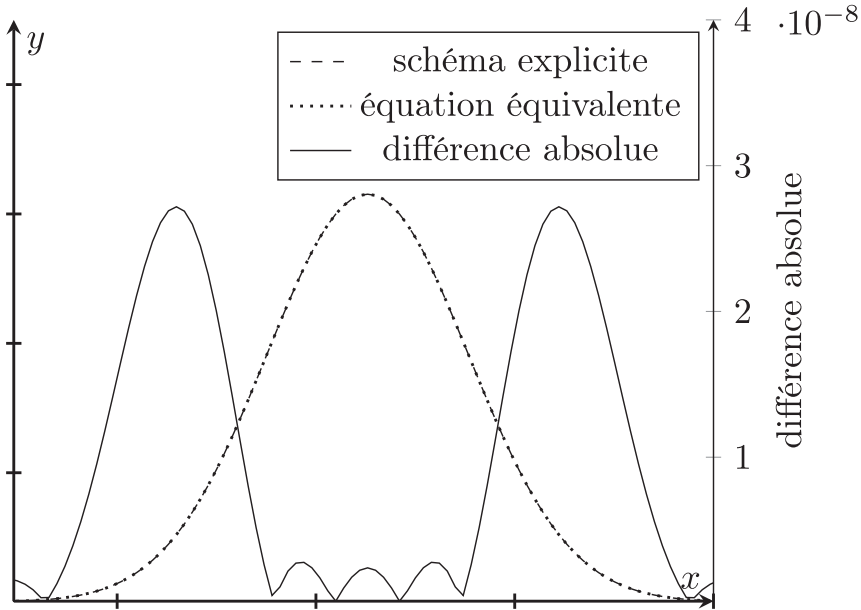


Fig. 8.1 Fonctions de Green G_s du schéma explicite, son approximation par (8.19), et leur différence.

$$G_s = \sum_{m, 2m+j \leq n} (1-2\eta)^{n-2m-j} \eta^{2m+j} \frac{n!}{(n-2m-j)! m! (m+j)!}$$

$$G_s = \sum_{m, 2m+j \leq n} (1-2\eta)^{n-2m-j} (2\eta)^{2m+j} \binom{n}{2m+j} 2^{-(2m+j)} \binom{2m+j}{m}$$

L'intérêt de la seconde écriture fait apparaître les lois binomiales de probabilité 2η et $1/2$. On s'assure que la transformée en z

$$G_s = \frac{1}{2\pi i} \int_C \left((1-2\eta) + \eta \left(z + \frac{1}{z} \right) \right)^n z^{j-1} dz$$

$$= \frac{1}{2\pi i} \int_C \sum_{m,p} (1-2\eta)^{n-m-p} \eta^{m+p} \frac{n!}{(n-m-p)! m! p!} z^{m-p+j-1} dz$$

donne bien le même résultat : en effet, les seuls termes non nuls, *i.e.* les contributions de résidu, viennent des couples m, p vérifiant $m-p+j-1 = -1 \rightarrow p = j+m$.

Le schéma a une interprétation probabiliste en termes de marche aléatoire : si le marcheur reste en place avec la probabilité $(1-2\eta)$, avance ou recule d'une case avec la probabilité η , la probabilité $P(j, n+1)$ qu'il se trouve sur la case j au temps

n vérifie

$$P(j, n+1) = (1 - 2\eta)P(j, n) + \eta P(j-1, n) + \eta P(j+1, n)$$

$P(j, n+1)$ et u_j^{n+1} satisfont la même équation, ce qui permet d'identifier la solution du schéma u_j^n à $P(j, n)$.

L'événement « être en j au temps n » est réalisé si le marcheur stationne $(n - 2m - j)$ fois, recule m fois, avance $(m + j)$ fois. Chacun de ces événements a une probabilité $(1 - 2\eta)^{n-2m-j} \eta^{2m+j}$. Ils sont au nombre de $\frac{n!}{(n-2m-j)!m!(m+j)!}$. Donc $P(j, n) = u_j^n$ est la somme de ces probabilités, d'où la formule pour G_5 . On vérifie bien la conservation de $\sum u_j^n = 1$

$$((1 - 2\eta) + \eta + \eta)^n = \sum_{m,p} (1 - 2\eta)^{n-m-p} \eta^{m+p} \frac{n!}{(n-m-p)!m!p!} = 1$$

Enfin, selon le théorème central limite, $P(j, n)$ converge, comme loi de probabilité d'une somme de n variables aléatoires indépendantes de moyenne nulle et de variance $\frac{1}{3}(2\eta) = \frac{2}{3}\eta = \sigma^2$ vers une gaussienne de variance $n\sigma^2$. Cela implique la convergence de la solution du schéma vers celle de l'équation, comme pour le schéma upwind pour l'équation d'advection. En somme, la méthode opératoire donne un résultat exact explicite, relativement facile à calculer et interprétable en termes de probabilité. Ce résultat se simplifie encore dans le cas particulier $\eta = \frac{1}{2}$.

Cas particulier $\eta = \frac{1}{2}$

Dans ce cas, le schéma s'écrit

$$u_j^{n+1} = \frac{1}{2}(u_{j-1}^n + u_{j+1}^n)$$

La formule explicite donnant u_j^n se simplifie : le marcheur est en j au temps n s'il a avancé $\frac{n+j}{2}$ fois et reculé $\frac{n-j}{2}$ fois

$$u_j^n = \frac{1}{2^n} \frac{n!}{\left(\frac{n+j}{2}\right)! \left(\frac{n-j}{2}\right)!}, \quad -n \leq j \leq n, \quad u_j^n = 0, \quad |j| > n \quad (8.20)$$

On notera que, pour n pair et j impair et, pour n impair et j pair, *i.e.* pour n, j de parités opposées, $u_j^n = 0$. En effet, $\frac{n+j}{2}$ et $\frac{n-j}{2}$ ne sont pas des entiers. Donc, u_j^n a des valeurs non nulles seulement pour n, j de même parité, $\frac{n+j}{2}$ et $\frac{n-j}{2}$ sont alors des entiers. La cohérence des représentations de la solution du schéma est présentée ci-dessous. La formule donnant u_j^n se réécrit

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos k)^n e^{ijk} dk = \frac{1}{2i\pi} \int_C \frac{1}{2^n} \left(z + \frac{1}{z}\right)^n z^{j-1} dz$$

En développant la puissance, on obtient

$$u_j^n = \frac{1}{2i\pi} \int_C \frac{1}{2^n} \sum \binom{n}{p} z^{2p+j-n-1} dz$$

Le seul terme non nul est celui devant z^{-1} , soit $2p + j - n = 0$, $p = \frac{n-j}{2}$. Il n'existe que pour n, j de même parité, et on retrouve (19), i.e $u_j^n = \frac{1}{2^n} \binom{n}{\frac{n-j}{2}}$ pour n, j de même parité, $u_j^n = 0$ pour n, j de parités opposées. En particulier, pour $j = 0$, $u_0^n = \frac{1}{2^n} \binom{n}{\frac{n}{2}}$, pour n pair. En approchant les factorielles par la formule de Stirling,

$$u_0^n \approx \frac{1}{2^n} \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{\pi n \left(\frac{n}{2e}\right)^n} = \frac{\sqrt{2}}{\sqrt{\pi n}}$$

La fonction de Green exacte pour $\tau = 1, h = 1, \eta = \frac{1}{2}$ donc $\alpha = \frac{1}{2}$ vaut en $j = 0$,

$$G(0) = \frac{1}{2\sqrt{\pi\alpha t}} = \frac{1}{\sqrt{2\pi n}} = \frac{1}{2} u_0^n$$

Pour $j \neq 0$, l'utilisation de la formule de Stirling pour les factorielles donne

$$u_j^n \approx \frac{2}{\sqrt{2\pi n}} \sqrt{\frac{n}{n^2 - j^2}} \left(1 + \frac{j}{n}\right)^{-(n+j)/2} \left(1 - \frac{j}{n}\right)^{-(n-j)/2}$$

En développant au voisinage de $\frac{j}{n} = 0$ la quantité

$$\ln \left(\left(1 + \frac{j}{n}\right)^{-(n+j)/2} \left(1 - \frac{j}{n}\right)^{-(n-j)/2} \right)$$

on obtient l'approximation de u_j^n pour n, j de même parité

$$u_j^n \approx \frac{2}{\sqrt{2\pi n}} \exp\left(-\frac{j^2}{2n}\right)$$

mais $\frac{j^2}{2n} = \frac{x^2\tau}{2h^2t}$ et $\eta = \frac{\alpha\tau}{h^2} = \frac{1}{2}$, donc $\frac{j^2}{2n} = \frac{x^2}{4\alpha t}$, et $\sqrt{\frac{2}{\pi n}} = \sqrt{\frac{h^2}{\pi\alpha n\tau}} = \frac{h}{\sqrt{\pi\alpha t}}$. Au final

$$u_j^n \approx \frac{h}{\sqrt{\pi\alpha t}} \exp\left(-\frac{x^2}{4\alpha t}\right) = 2G_e \quad (8.21)$$

soit le double de la fonction de Green G_e de l'équation. En complément, $u_0^n = \frac{1}{\pi} \int_0^\pi (\cos k)^n dk = \frac{1}{\pi} \int_0^{\pi/2} (\cos k)^n dk + \frac{(-1)^n}{\pi} \int_0^{\pi/2} (\cos k)^n dk$. Donc, si $n = 2p$,

$$u_0^{2p} = \frac{2}{\pi} W_n = \frac{(2p)!}{2^{2p}(p!)^2} \approx \frac{2}{\sqrt{2\pi n}}$$

La fonction de Green du schéma est représentée sur la figure 8.2. L'annulation à une valeur sur 2 de j lui donne un aspect très oscillant. De plus, comme démontré seulement au voisinage de l'origine, elle vaut partout approximativement 2 fois celle de l'équation aux points où elle est non nulle.

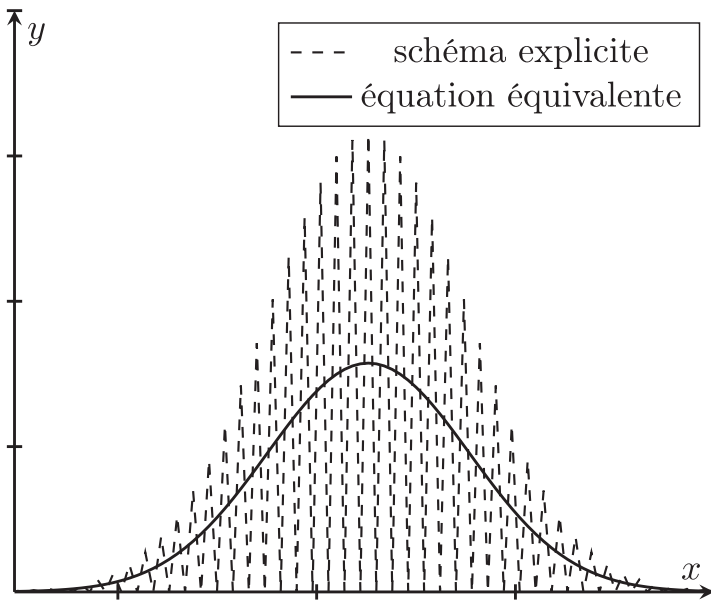


Fig. 8.2 Fonction de Green du schéma pour $\eta = \frac{1}{2}$ comparée à (8.21).

Conclusion

Nous avons étudié dans cette section le schéma explicite pour l'équation de la chaleur 1D, et obtenu des formules exactes par méthode opératorielle et par analyse de Fourier pour la fonction de Green du schéma. Nous avons ensuite obtenu une approximation de cette fonction pour n grand, d'abord directement, puis en passant par l'équation équivalente, qui a l'avantage (en dépit de sa non-validité pour une plage de valeurs de η ...) de conduire à des calculs plus simples, et de permettre de résumer le

comportement du schéma à l'aide du seul paramètre b_e . L'accord entre les simulations numériques et les formules (quelque peu heuristiques) proposées est bon.

Pour la valeur particulière $\eta = \frac{1}{2}$, la fonction de Green du schéma s'exprime sous forme simple. La convergence vers la fonction de Green de l'équation n'a pas lieu en un sens classique, mais en moyenne sur deux valeurs consécutives de j . Pour la valeur particulière $\eta = \frac{1}{4}$, la valeur à l'origine est donnée par une intégrale de Wallis.

Le schéma explicite a toutefois une condition de stabilité restrictive, et il peut être préférable d'impliciter, pour ne plus avoir à la satisfaire. La section suivante est consacrée à l'analyse du schéma implicite.

8.1.2 Schéma implicite

Considérons le schéma implicite, qui calcule u_j^{n+1} par la formule

$$\frac{u_j^{n+1} - u_j^n}{\tau} = \alpha \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{h^2} \quad (8.22)$$

Il est d'ordre 1 en temps et 2 en espace [4]. (8.22) s'écrit aussi

$$u_j^{n+1} = u_j^n - 2\eta u_j^{n+1} + \eta u_{j-1}^{n+1} + \eta u_{j+1}^{n+1} \quad (8.23)$$

où, comme plus haut, $\eta = \frac{\alpha\tau}{h^2}$. Il demande la résolution d'un système tridiagonal à chaque pas de temps, donc le pas de temps est plus coûteux que pour le schéma explicite. Par contre, il est inconditionnellement stable, ce qui permet d'utiliser de grands pas de temps. En effet, en Fourier, (8.23) devient

$$\widehat{u}^{n+1} = \widehat{u}^n - 2\eta \widehat{u}^{n+1} + \eta e^{-ik} \widehat{u}^{n+1} + \eta e^{ik} \widehat{u}^{n+1}$$

ce qui permet d'écrire

$$\widehat{u}^{n+1} = \frac{1}{1 + 4\eta \sin^2(k/2)} \widehat{u}^n \quad (8.24)$$

et le facteur d'amplification est toujours plus petit que 1, d'où la stabilité inconditionnelle. La fonction de Green s'écrit, d'après (8.24)

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 + 4\eta \sin^2(k/2))^{-n} e^{ijk} dk \quad (8.25)$$

ou bien, en utilisant la transformée en z

$$u_j^n = \frac{1}{2i\pi} \int_C (z - \eta(z-1)^2)^{-n} z^{n+j-1} dz$$

Équation équivalente

L'équation équivalente à un terme peut être calculée en développant le logarithme du facteur d'amplification au premier ordre

$$\ln\left(\frac{1}{(1 + 4\eta \sin^2(k/2))}\right) = -k^2\eta + k^4\left(\frac{1}{12}\eta + \frac{1}{2}\eta^2\right) + O(k^5)$$

L'équation équivalente à un terme s'écrit donc

$$\partial_t u - \alpha \partial_{xx} u = \frac{1}{2} \alpha h \left(\eta + \frac{1}{6}\right) \partial_{4x} u \tag{8.26}$$

Par comparaison à (8.17), on déduit le paramètre caractéristique du schéma implicite $b_i = \eta + \frac{1}{6}$ (rappelons celui du schéma explicite : $b_e = -\eta + \frac{1}{6}$). Le terme de dérivée d'ordre 4 est positif, ce qui rend le problème de l'équation équivalente à un terme mal posé, alors que le schéma est stable. Il apparaît donc la même difficulté que dans le cas du schéma explicite. La fonction de Green dans l'approximation équation équivalente à un terme s'écrit, en gardant les bornes finies $u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-n\eta k^2 + \frac{1}{2}n\eta(\eta + \frac{1}{6})k^4 + ijk) dk$, mais il s'agit d'une écriture formelle. Une estimation de u_0^n peut toutefois être obtenue comme précédemment, par la méthode de Laplace. Le maximum de l'intégrande est en $k = 0$, et, comme le résultat ne dépend que des coefficients du développement de Taylor de S au voisinage de $k = 0$, on obtient le résultat en substituant b_i à b_e

$$u_0^n \simeq \frac{h}{2\sqrt{\pi\alpha t}} \left(1 + \frac{3}{8} \frac{h^2}{\alpha t} \left(\eta + \frac{1}{6}\right)\right)$$

Vérifions ce résultat en appliquant la méthode de Laplace à l'intégrale (8.25) exprimant la solution du schéma, réécrite sous la forme

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(-n \ln(1 + 4\eta \sin^2(k/2))) \cos jk dk$$

qui s'écrit sous la forme (8.10), que nous rappelons ci-dessous

$$I = \int_a^b \exp(-np(k)) q(k) dk$$

En suivant la même démarche que pour le schéma explicite, on obtient le développement asymptotique pour n grand

$$I \approx \sqrt{\frac{\pi}{np_2}} \exp(-np_0) \left(a_0 + \frac{1}{n} a_2 \left(\frac{1}{2}\right)_1\right) \tag{8.27}$$

a_0 et a_2 s'écrivent à partir des coefficients du développement de Taylor de $p = \ln(1 + 4\eta \sin^2(k/2))$ et $q = \cos jk$. Calculons ces coefficients. En développant

$$\ln\left(1 + 4\eta \sin^2\left(\frac{k}{2}\right)\right) = k^2\eta - k^4\left(\frac{1}{2}\eta^2 + \frac{1}{12}\eta\right) + O(k^6)$$

on obtient $p_0 = 0$, $p_2 = \eta$, $p_4 = -\frac{1}{2}\eta(\eta + \frac{1}{6})$. De même,

$$\cos(jk) = 1 - \frac{1}{2}j^2k^2 + \frac{1}{24}j^4k^4 + O(k^6)$$

donc

$$q_0 = 1, \quad q_2 = -\frac{1}{2}j^2, \quad q_4 = +\frac{1}{24}j^4$$

Comme $(\frac{1}{2})_n = \frac{\Gamma(\frac{1}{2}+n)}{\Gamma(\frac{1}{2})}$, $(\frac{1}{2})_1 = \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{1}{2})} = \frac{1}{2}$, on obtient, pour a_0 et a_2

$$a_0 = q_0 = 1, \quad a_2 = \frac{1}{p_2}\left(q_2 - \frac{3}{2}\frac{p_4}{p_2}q_0\right) = -\frac{1}{2\eta}j^2 + \frac{3}{4\eta}\left(\eta + \frac{1}{6}\right)$$

En reportant ces résultats dans (8.27)

$$I \approx \sqrt{\frac{\pi}{n\eta}}\left(1 + \frac{1}{2n}\left(-\frac{1}{2\eta}j^2 + \frac{3}{4\eta}\left(\eta + \frac{1}{6}\right)\right)\right)$$

et, en revenant aux variables continues, $-\frac{1}{4\eta n}j^2 + \frac{3}{8\eta n}\left(\eta + \frac{1}{6}\right) = -\frac{1}{4\alpha t}x^2 + \frac{3b^2}{8\alpha t}\left(\eta + \frac{1}{6}\right)$. Dans cette formulation, le terme $\exp(-\frac{x^2}{4\alpha t})$ apparaît sous la forme $1 - \frac{x^2}{4\alpha t}$. Au final

$$u_j^n \approx \frac{1}{2\sqrt{\pi n\eta}}\left(1 + \frac{3b^2}{8\alpha t}\left(\eta + \frac{1}{6}\right)\right)\exp\left(-\frac{x^2}{4\alpha t}\right) \quad (8.28)$$

Ce résultat s'obtient en substituant dans (8.13) $b_i = \eta + \frac{1}{6}$ à $b_e = -\eta + \frac{1}{6}$. Nous pouvons donc inférer de l'approximation (8.14) pour le schéma explicite celle pour le schéma implicite

$$G_s/G_e \approx \left(1 + \frac{3}{8}\frac{b^2}{\alpha t}\left(\eta + \frac{1}{6}\right)\right)\left(1 - \frac{3}{8}\frac{b^2}{t^2}\frac{x^2}{\alpha^2}\left(\eta + \frac{1}{6}\right)\right)\exp\left(\frac{1}{32}\left(\eta + \frac{1}{6}\right)\frac{b^2x^4}{(\alpha t)^3}\right) \quad (8.29)$$

Nous présentons des résultats de simulations numériques, et les comparons à cette formule. Pour des valeurs de η de l'ordre de l'unité ou plus petites, les courbes sont quasiment superposées, comme on peut le voir sur la figure 8.3 qui compare, à titre d'exemple, G_s obtenue numériquement pour $\eta = 1$ à (8.29). La différence est toutefois plus grande (mais η est aussi plus grand) que dans le cas du schéma

explicite. Elle est plus petite au centre que sur les bords, ce qui est attendu pour une approximation équation équivalente.

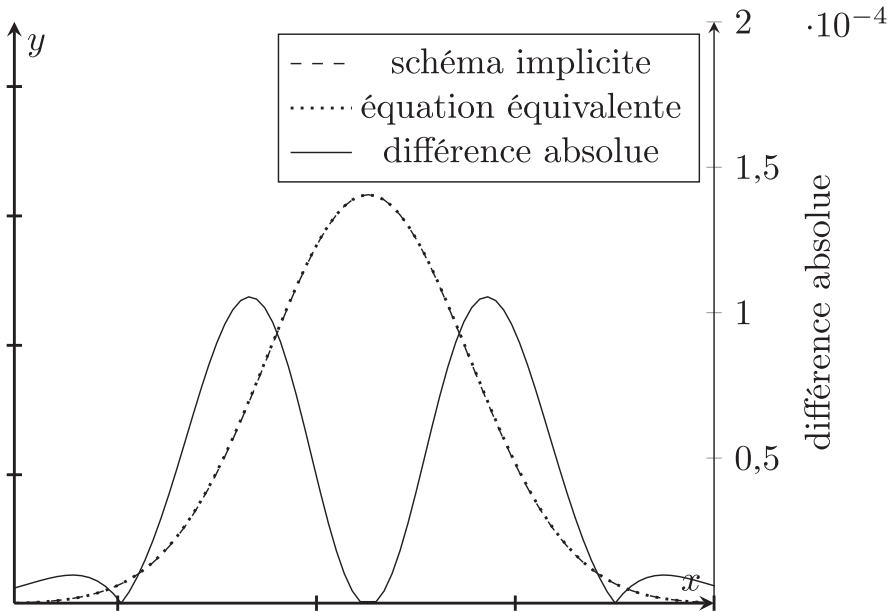


Fig. 8.3 G_S , son approximation (8.29) et leur différence.

Le schéma implicite étant stable, il est possible d'utiliser des valeurs de η arbitraires. La figure 8.4 présente les fonctions de Green pour de grandes valeurs de η obtenues en allant au temps final en un petit nombre de pas de temps. L'écart relatif par rapport à la fonction de Green de l'équation devient significatif pour η grand, mais le schéma reste monotone, au sens où il ne génère pas d'oscillations.

Nous étudions maintenant le schéma de Crank-Nicolson, inconditionnellement stable et plus précis en principe que le schéma implicite.

8.1.3 Schéma de Crank-Nicolson

Ce schéma, d'ordre 2 en temps et en espace [4], est la demi-somme des schémas explicite et implicite, soit

$$\frac{u_j^{n+1} - u_j^n}{\tau} = \alpha \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{2h^2} + \alpha \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{2h^2} \quad (8.30)$$

En Fourier, il s'écrit

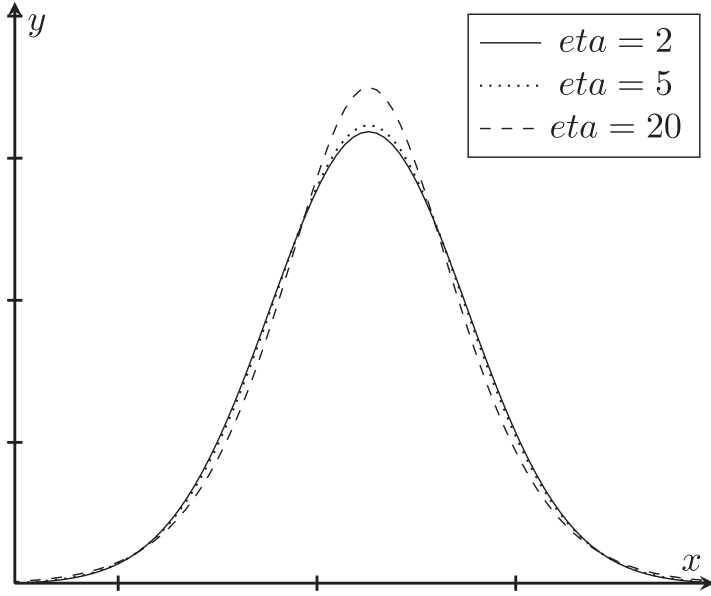


Fig. 8.4 Fonction de Green du schéma implicite pour de grandes valeurs de η .

$$\widehat{u^{n+1}}(1 + \eta(1 - \cos k)) = (1 - \eta(1 - \cos k))\widehat{u^n}$$

et sa solution est donc, avec le facteur d'amplification

$$g(k) = \frac{1 - \eta(1 - \cos k)}{1 + \eta(1 - \cos k)}$$

$$u_j^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(k)^n e^{ijk} dk \quad (8.31)$$

Comme $g(k) \leq 1$, le schéma est stable pour tout η . Le développement de son logarithme

$$\ln g(k) = -k^2\eta + \frac{1}{12}k^4\eta + O(k^6)$$

donne l'équation équivalente, qui détermine le paramètre du schéma, $b_c = \frac{1}{6}$, ce qui est suffisant pour écrire l'approximation de sa fonction de Green. Appliquons toutefois, à titre de vérification, la méthode de Laplace à (8.31), réécrite sous la forme

$$u_j^n = \frac{1}{\pi} \int_{-\pi}^{\pi} \exp\{n[\ln(1 - \eta(1 - \cos k)) - \ln(1 + \eta(1 - \cos k))]\} \cos(jk) dk \quad (8.32)$$

Les extrema de $g(k)$ vérifient

$$\frac{d}{dk} \frac{1 - \eta(1 - \cos k)}{1 + \eta(1 - \cos k)} = -2\eta \frac{\sin k}{(\eta - \eta \cos k + 1)^2} \sin k = 0$$

soit $k = 0$ ou π . Pour n grand, la contribution dominante est celle en $k = 0$. Les coefficients p_i intervenant dans l'approximation par la méthode de Laplace de (8.32) sont maintenant

$$p_0 = 0, \quad p_2 = \eta, \quad p_4 = -\frac{1}{12}\eta; \quad q_0 = 1, \quad q_2 = -\frac{1}{2}j^2, \quad q_4 = +\frac{1}{24}j^4$$

Ils permettent de calculer

$$a_0 = q_0 = 1, \quad a_2 = \frac{1}{p_2} \left(q_2 - \frac{3}{2} \frac{p_4}{p_2} q_0 \right) = \frac{1}{\eta} \left(-\frac{1}{2}j^2 + \frac{1}{8} \right)$$

On obtient

$$u_j^n \approx \sqrt{\frac{\pi}{n\eta}} \exp(-np_0) \left(1 + \frac{1}{2n} \left(\frac{1}{\eta} \left(-\frac{1}{2}j^2 + \frac{1}{8} \right) \right) \right) = \sqrt{\frac{\pi}{n\eta}} \left(1 - \frac{j^2}{4n\eta} + \frac{1}{16n\eta} \right)$$

et, revenant aux variables continues

$$u_j^n \approx \frac{1}{2\sqrt{\pi n\eta}} \left(1 + \frac{1}{16} \frac{b^2}{\alpha t} \right) \exp\left(-\frac{x^2}{4\alpha t}\right)$$

Cette formule confirme la valeur $b_c = \frac{1}{6}$. On notera que $b_c = \frac{1}{2}(b_e + b_i)$: ce résultat était prévisible, au vu de la définition (8.30) du schéma.

Une approximation de G_s valide au voisinage de l'origine peut être inférée comme pour les schémas précédents. Le paramètre du schéma est $b_c = \frac{1}{6}$, d'où l'approximation

$$G_s \approx G_e \left(1 + \frac{1}{16} \frac{b^2}{\alpha t} \right) \left(1 - \frac{1}{16} \frac{b^2}{t^2} \frac{x^2}{\alpha^2} \right) \exp\left(-\frac{1}{192} \frac{b^2 x^4}{(\alpha t)^3}\right) \quad (8.33)$$

Nous présentons des résultats de simulations numériques, et les comparons à cette formule. Pour des valeurs de η de l'ordre de l'unité ou plus petites, les courbes sont quasiment superposées, comme on peut le voir sur la figure 8.5, qui compare, à titre d'exemple, G_s obtenue numériquement pour $\eta = 1$ à (8.33). Nous avons tracé leur différence, qui est du même ordre que pour le schéma implicite. Elle est plus petite au centre que sur les bords, comme pour l'implicite.

Le schéma diffère des précédents au sens où l'approximation (8.33) de G_s ne dépend pas de η . L'écart relatif au voisinage de l'origine entre la fonction de Green du schéma et celle de l'équation doit donc être plus faible que pour le schéma implicite. La figure 8.6 présente la fonction de Green pour différentes valeurs de η . La prédiction est validée, avec un écart quasi indépendant de η , jusqu'à une valeur seuil, où le schéma

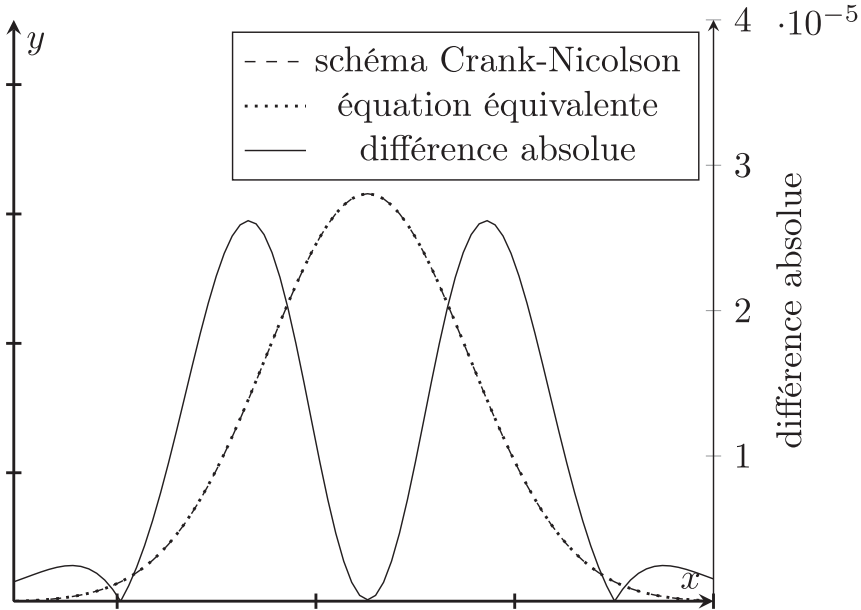


Fig. 8.5 G_s , son approximation (8.33) et leur différence.

génère des oscillations. Ces oscillations sont évoquées par Strikwerda, qui les attribue à la faible dissipativité du schéma pour de grands pas de temps [3]. Elles ont aussi été étudiées, ainsi que des techniques pour les amortir, dans [1]. Qualitativement, elles sont générées par le facteur d'amplification $g(k) = \frac{1-\eta(1-\cos k)}{1+\eta(1-\cos k)}$ qui, pour $k = \pi$ devient $g(k) = \frac{1-2\eta}{1+2\eta} \approx -1$ pour de grandes valeurs de η .

8.1.4 Conclusion

Il existe un grand nombre de schémas discrétisant l'équation de la chaleur en 1D. Nous en avons analysé trois parmi les plus classiques : schéma explicite, souvent utilisé, en dépit de sa limite de stabilité restrictive, car le coût de calcul unitaire du pas de temps est très faible, schéma implicite et schéma de Crank-Nicolson, inconditionnellement stables. Nous nous sommes attachés à établir des résultats approximatifs, mais explicites, en particulier l'écart relatif à l'origine entre les fonctions de Green des schémas G_s et celle de l'équation G_e . Il s'écrit

$$G_s/G_e \approx 1 + \frac{3}{8} b \frac{b^2}{\tau \alpha}$$

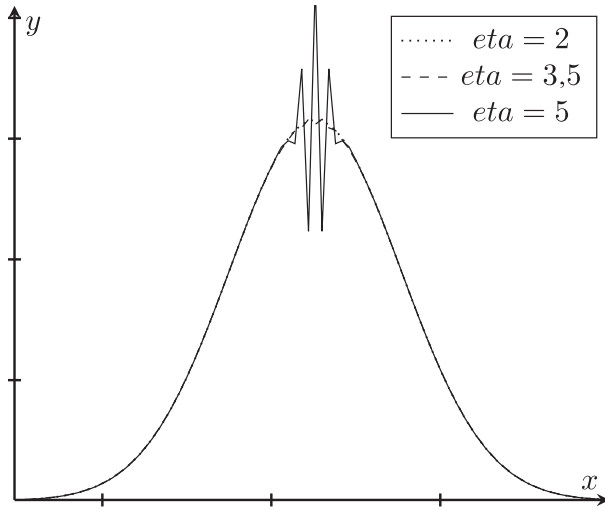


Fig. 8.6 Fonction de Green G_s du schéma de Crank-Nicolson pour de grandes valeurs de η .

Le paramètre b dépend du schéma et vaut respectivement $b_e = -\eta + \frac{1}{6}$, $b_i = \eta + \frac{1}{6}$, pour les schémas explicite, implicite et $b_c = \frac{1}{6}$ pour Crank-Nicolson. Il est indépendant de η pour ce dernier schéma. Une approximation valide en principe seulement au voisinage de l'origine a été proposée.

$$G_s/G_e \approx \left(1 + \frac{3}{8}b \frac{h^2}{\alpha t}\right) \left(1 - \frac{3}{8}b \frac{h^2 x^2}{(\alpha t)^2}\right) \exp\left(\frac{1}{32}b \frac{h^2 x^4}{(\alpha t)^3}\right) \quad (8.34)$$

Elle s'avère précise pour les trois schémas étudiés, comme illustré sur les figures.

L'équation équivalente à un terme ne permet pas de prédire la stabilité des schémas, et conduit à des problèmes mal posés, alors que les schémas correspondants sont stables. Elle peut toutefois être utilisée pour le schéma explicite, dans le cas $\frac{1}{6} < \eta \leq \frac{1}{4}$. Nous l'avons aussi utilisée, associée à la méthode du col, comme moyen heuristique d'obtention de la formule (8.34) ci-dessus.

Nous terminons ce chapitre par une courte section sur l'équation de la chaleur en 2D. Le but est de montrer que des effets d'anisotropie peuvent apparaître, mais qu'ils sont peu sensibles sur maillage régulier, composé de carrés.

8.2 Schémas discrétisant l'équation de la chaleur en 2D

8.2.1 Schéma explicite sur maillage régulier

En 2D, le schéma explicite s'écrit, avec $\eta = \frac{\alpha \tau}{h^2}$

$$u_{jm}^{n+1} = (1 - 4\eta)u_{jm}^n + \eta u_{(j-1)m}^n + \eta u_{(j+1)m}^n + \eta u_{j(m-1)}^n + \eta u_{j(m+1)}^n \quad (8.35)$$

Il est barycentrique, donc stable, pour $\eta < \frac{1}{4}$. Comme il fait intervenir le point (j, m) et ses 4 voisins, il est appelé schéma à 5 points.

Dans ce cas, u_{jm}^n peut être interprété comme la probabilité qu'un marcheur aléatoire dans le plan se trouve dans la case (j, m) au temps n . Il reste en place avec la probabilité $(1 - 4\eta)$, avance ou recule d'une case suivant x ou y avec la probabilité η . Pour la fonction de Green, le marcheur part de la case $(0, 0)$. C'est donc la somme de n variables aléatoires vectorielles. Le théorème central limite permet comme dans le cas 1D de conclure à la convergence vers une loi normale vectorielle (vecteur gaussien), *i.e.* vers la fonction Green de l'équation en 2D.

Les résultats du 1D se généralisent, toutes les méthodes s'appliquent. La méthode opératorielle conduit à une expression de la solution comme une somme à deux indices pondérée de coefficients multinomiaux, qui a peu d'utilité pratique. Nous nous limiterons à présenter les résultats de l'analyse de Fourier, qui donne des résultats explicites simples.

Analyse de Fourier

Dans l'espace réciproque, (8.35) s'écrit

$$\widehat{u}^{n+1} = ((1 - 4\eta) + 2\eta \cos k + 2\eta \cos l) \widehat{u}^n$$

soit le facteur d'amplification

$$g(k, l) = (1 - 4\eta) + 2\eta \cos k + 2\eta \cos l = 1 - 4\eta \sin^2\left(\frac{k}{2}\right) - 4\eta \sin^2\left(\frac{l}{2}\right)$$

Le développement du logarithme du facteur d'amplification au voisinage de $(k, l) = 0$ s'écrit, en notant $(k, l) = \kappa(\cos \theta, \sin \theta)$, $\kappa = \sqrt{k^2 + l^2}$

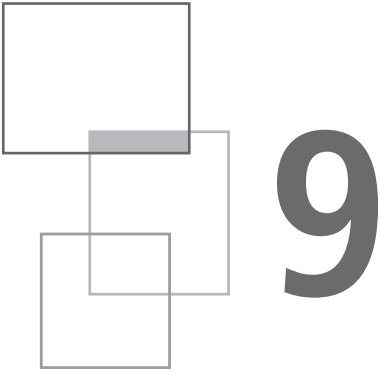
$$\ln g(k, l) = -\kappa^2 \eta - \kappa^4 \frac{\eta}{2} \left(\eta + \frac{1}{6} (\cos^4 \theta + \sin^4 \theta) \right) + O(r^6)$$

Une anisotropie de mode 2 apparaît, comme pour l'équation des ondes traitée avec le schéma de Yee. En pratique, elle est peu sensible.

Listons les références du chapitre 8. [2] étudie la validité de l'équation équivalente pour l'équation de la chaleur. Le chapitre [3] du livre de Strikwerda et le livre [4] de Thomée, très complet sur les équations paraboliques, analysent un ensemble de schémas, dont ceux présentés dans ce chapitre, mais aussi des schémas d'ordre élevé, et présentent des résultats de convergence en norme L^2 . [1] analyse les oscillations du schéma de Crank-Nicolson.

Bibliographie

- [1] D. Britz, O. Østerby, J. Strutwolf "Damping of Crank-Nicolson error oscillations" Computational Biology and Chemistry 27, pp. 253-263, 2003
- [2] F. Dhaouadi, E. Duval, S. Tkachenko, J.P. Vila, "Stability theory for some scalar finite difference schemes : validity of the modified equations approach" ESAIM 70, pp. 124-136, 2021
- [3] J. Strikwerda "Finite difference schemes and partial differential equations" chapitre 10, SIAM, 2004
- [4] V. Thomée "Finite difference methods for linear parabolic equations" Elsevier, 1990



Conclusion générale

Nous avons présenté dans le chapitre 4 de cet ouvrage un ensemble de méthodes permettant de comprendre et d'analyser quantitativement le comportement de schémas numériques pour des EDP linéaires. La description théorique des méthodes est complétée par un ensemble que nous espérons représentatif, d'applications à des schémas. Le chapitre 5, consacré à l'équation d'advection, présente une palette assez large de schémas, et les analyses les plus détaillées. Le chapitre 6 sur l'équation des ondes 1D se limite à un seul schéma, car les analyses du chapitre 5 se transposent sans difficulté particulière à cette équation. Le chapitre 7 sur l'équation des ondes 2D est focalisé sur un schéma d'ordre 2 suffisamment simple pour mettre en évidence avec des formules explicites les effets d'anisotropie. Enfin, le chapitre 8 présente trois schémas classiques pour l'équation de la chaleur. Nous espérons que ce livre aidera le lecteur non seulement à comprendre comment se comportent les schémas présentés, mais aussi à analyser d'autres schémas, appliqués à d'autres EDP.

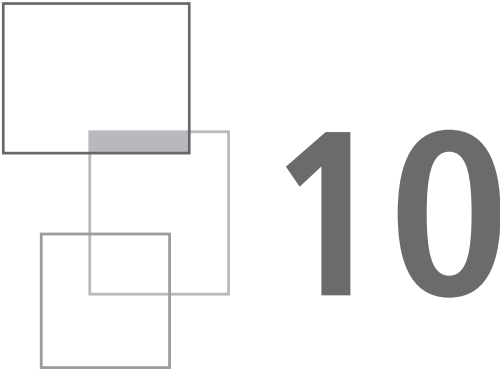
Remerciements

Nous remercions G. Bonnaud, G. Colin de Verdière, S. Del Pino, E. Labourasse, G. Naulin et B. Rebourcet, du CEA, et B. Després, de Sorbonne université, pour leur relecture attentive de la version préliminaire du manuscrit. Leurs remarques nous ont permis de rectifier un certain nombre d'erreurs et d'imprécisions.

Troisième partie

Appendice

L'appendice comprend trois chapitres techniques. Le chapitre 10 présente les trois types de schémas d'ordre élevé analysés dans le chapitre 5. Le chapitre 11 présente les fonctions spéciales intervenant dans la solution des schémas, et plus particulièrement la fonction d'Airy et ses généralisations. Le dernier chapitre est consacré aux méthodes de développement asymptotiques d'intégrales, fréquemment utilisées dans les chapitres 5 à 8 pour établir des formules explicites approchant les fonctions de Green des schémas étudiés. Ces résultats, le plus souvent disponibles dans la littérature, sont inclus pour que le lecteur puisse y accéder facilement. Ils sont présentés assez brièvement, en se référant à des ouvrages dédiés que les explorent plus en détail.



Schémas d'ordre élevé

Pour construire un schéma d'ordre élevé pour une EDP dépendant du temps, plusieurs méthodes peuvent être employées. La discrétisation simultanée en temps et en espace conduit aux schémas de Strang, présentés dans la première partie du chapitre. La seconde partie traite de l'opérateur de dérivée spatiale Q_p , *i.e.* s'intéresse aux schémas semi-discrétisés. Nous nous limitons aux schémas utilisés dans le livre : schémas centrés et schémas décalés.

10.1 Schémas de Strang

Les schémas de Strang sont des schémas à un pas, donc faciles à mettre en œuvre, et d'ordre arbitraire. Ils ont été proposés par G. Strang dans les années 1960. Une méthode, simplifiant considérablement la présentation et l'analyse de ces schémas, a été proposée par B. Després [1]. Nous suivons son approche dans cet appendice.

10.1.1 *Génération des schémas de Strang pour l'équation d'advection à vitesse a*

Le symbole exact en Fourier de l'opérateur de transport sur un pas de temps τ est, avec $\nu = \frac{a\tau}{h}$, puisque $\widehat{u}_t = -ika\widehat{u}$

$$\exp(-ika\tau) = \exp(-ik\nu h) = \exp(-ikh)^\nu$$

Un schéma à un pas approxime cet opérateur $\exp(-ikh)^\nu = z^\nu$ par un polynôme $P(z)$ en $z = \exp(-ikh)$, par exemple, pour le schéma upwind

$$\begin{aligned} u_j^{n+1} &= (1-\nu)u_j^n + \nu u_{j-1}^n \\ \widehat{u}^{n+1} &= (1-\nu)\widehat{u}^n + \nu \exp(-ikh)\widehat{u}^n \\ P(z) &= (1-\nu) + \nu z \end{aligned}$$

Plus généralement, $P(z)$ et z^ν doivent avoir le même développement de Taylor en $z = 1$ jusqu'à l'ordre p du schéma. Un schéma de Strang S_p^s est ainsi caractérisé par son ordre p et son décalage s . En partant des égalités (10.1), paramétrées par le décalage s ,

$$z^\nu = (1 + (z - 1))^{s+\nu} z^{-s} \quad (10.1)$$

les schémas sont obtenus par développement de Taylor autour de $z = 1$ à l'ordre p du second membre de (1) : on obtient ainsi le polynôme P_p^s associé au schéma. Les contraintes de stabilité limitent les valeurs possibles des couples (p, s) à $(s, 2s)$, $(s, 2s + 1)$, $(s, 2s + 2)$. Donnons quelques exemples : S_1^0 est obtenu en développant à l'ordre 1 $(1 + (z - 1))^\nu$, soit

$$(1 + (z - 1))^\nu \approx 1 + \nu(z - 1) = (1 - \nu) + \nu z \quad (10.2)$$

(10.2) implique que S_1^0 s'écrit, puisque $z = \exp(-ikh)$

$$u_j^{n+1} = (1-\nu)u_j^n + \nu u_{j-1}^n$$

S_1^0 est donc le schéma upwind. Pour S_2^0 , on développe aussi (10.2), mais à l'ordre 2

$$(1 + (z - 1))^\nu \approx 1 + \nu(z - 1) - \frac{1}{2}\nu(1 - \nu)(z - 1)^2 = P_2^0$$

Le schéma S_2^0 s'écrit donc

$$u_j^{n+1} = (1-\nu)u_j^n + \nu u_{j-1}^n - \frac{1}{2}\nu(1-\nu)(u_{j-2}^n - 2u_{j-1}^n + u_j^n)$$

soit le schéma de Beam-Warming (étudié dans la section 5.5). De même, S_2^1 est obtenu en développant à l'ordre 2 $(1 + (z - 1))^{\nu+1} z^{-1}$

$$(1 + (z - 1))^{\nu+1} z^{-1} \approx z^{-1} + (\nu + 1)(1 - z^{-1}) + \frac{1}{2}\nu(\nu + 1)(z - 2 + z^{-1}) \quad (10.3)$$

Le schéma S_2^1 s'écrit donc

$$u_j^{n+1} = (1 - \nu)u_j^n + \nu u_{j-1}^n - \frac{1}{2}\nu(1 - \nu)(u_{j-1}^n - 2u_j^n + u_{j-1}^n)$$

soit le schéma de Lax-Wendroff (étudié dans la section 5.4). D'autres exemples de ces schémas sont étudiés dans le chapitre 5 sur l'équation d'advection. Le point essentiel est qu'ils sont générés simplement par une procédure facile à automatiser. Un autre intérêt de l'approche est de permettre leur analyse quantitative, comme expliqué dans le paragraphe suivant.

10.1.2 Analyse des schémas de Strang

Expression du facteur d'amplification

Reprenons les notations de Després $z = e^{i\theta}$. Le facteur d'amplification g du schéma de Strang S_p^s est, comme expliqué ci-dessus, le développement $P_p^s(e^{i\theta} - 1)$ de Taylor à l'ordre p du symbole exact. Donc, en utilisant la formule de Taylor avec reste intégral pour ce symbole exact, ou plus précisément pour la fonction

$$f_s(e^{i\theta} - 1) = (1 + (e^{i\theta} - 1))^{s+\nu} \quad (10.4)$$

On obtient, en développant (10.4)

$$f_s(e^{i\theta} - 1) = P_p^s(e^{i\theta} - 1) + \frac{1}{p!} \int_0^1 (1-t)^p f_k^{(p+1)}(t(e^{i\theta} - 1)) dt$$

si bien que le facteur d'amplification g est la somme de celui de l'équation et d'un terme intégral

$$g = \left[f_s(e^{i\theta} - 1) - \frac{1}{p!} \int_0^1 (1-t)^p f_k^{(p+1)}(t(e^{i\theta} - 1)) dt \right] e^{-is\theta} \quad (10.5)$$

L'expression explicite (10.5) de g permet en particulier d'estimer l'écart entre la solution de l'équation de transport et celle donnée par le schéma. Le premier travail est de réécrire sous une forme plus maniable le reste intégral

$$I = \int_0^1 (1-t)^p f_k^{(p+1)}(t(e^{i\theta} - 1)) dt \quad (10.6)$$

Nous allons le faire en détaillant les étapes, qui sont loin d'être évidentes. Notons $e^{i\theta} - 1 = y$, l'intégrale (10.6) se réécrit

$$I = \frac{1}{p!} \int_0^1 (1-t)^p f_k^{(p+1)}(ty) dt$$

En dérivant la fonction f_s , i.e. (10.4), à l'ordre $p + 1$, on obtient

$$f_s^{(p+1)}(ty) = (s+v)\dots(s+v-p)(1+ty)^{s+v-p-1}z^{p+1} = p!\alpha_{s,p}(1+ty)^{s+v-p-1}y^{p+1}$$

avec $\alpha_{s,p} = \frac{1}{p!} \prod_{r=0}^p (s+v-r)$, donc (10.6) devient

$$I = \alpha_{s,p} \int_0^1 (1-t)^p (1+ty)^{s+v-p-1} y^{p+1} dt = \alpha_{s,p} A$$

$$A = \int_0^1 (1-t)^p (1+ty)^{s+v-p-1} y^{p+1} dt \quad (10.7)$$

Mais $y = e^{i\theta} - 1$, donc (10.7) devient

$$A = \int_0^1 \frac{(1-t)^p (e^{i\theta} - 1)^{p+1}}{(1+t(e^{i\theta} - 1))^{\rho+1-s-v}} dt \quad (10.8)$$

Mais l'intégrale a une singularité en $t = \frac{1}{2}$ pour $\theta = \pi$, $1+t(e^{i\theta} - 1) = 1 + \frac{1}{2}(-1 - 1) = 0$. Effectuons, pour éviter cette singularité, le changement de variable ci-dessous dans l'intégrale (10.8)

$$z = 1 + t(e^{i\theta} - 1) \Rightarrow 1-t = \frac{e^{i\theta} - z}{e^{i\theta} - 1}, \quad dt = \frac{dz}{e^{i\theta} - 1}$$

Les bornes sont : $t = 0 \rightarrow z = 1$, $t = 1 \rightarrow z = e^{i\theta}$, si bien que (10.8) devient

$$A = \int_1^{e^{i\theta}} \frac{(e^{i\theta} - z)^p (e^{i\theta} - 1)^{p+1}}{(e^{i\theta} - 1)^p z^{\rho+1-s-v}} \frac{dz}{e^{i\theta} - 1} = \int_1^{e^{i\theta}} \frac{(e^{i\theta} - z)^p}{z^{\rho+1-s-v}} dz$$

Il n'y a plus de singularité, mais l'intégrale est dans le plan complexe. Pour se ramener à une intégrale sur un intervalle réel, effectuons un nouveau changement de variable $\varphi \in [0, \theta]$ tel que $z = e^{i\varphi}$, soit $dz = ie^{i\varphi} d\varphi$, les bornes deviennent $z = 1 \rightarrow \varphi = 0$, $z = 0 \rightarrow \varphi = \theta$. On obtient donc

$$A = \int_0^\theta \frac{(e^{i\theta} - e^{i\varphi})^p}{e^{i(p+1-s-v)\varphi}} ie^{i\varphi} d\varphi = i \int_0^\theta \frac{(e^{i\theta} - e^{i\varphi})^p}{e^{i(p-s-v)\varphi}} d\varphi \quad (10.9)$$

Mais $e^{i\theta} - e^{i\varphi} = 2i \sin\left(\frac{\theta-\varphi}{2}\right) e^{i\frac{\theta+\varphi}{2}}$, donc (10.9) se réécrit

$$A = 2^p i^{p+1} e^{ip\frac{\theta}{2}} \int_0^\theta \frac{\sin\left(\frac{\theta-\varphi}{2}\right)^p e^{i\frac{\theta+\varphi}{2}}}{e^{i(p-s-v)\varphi}} d\varphi = 2^p i^{p+1} e^{ip\frac{\theta}{2}} \int_0^\theta \sin\left(\frac{\theta-\varphi}{2}\right)^p e^{i(s+v-\frac{p}{2})\varphi} d\varphi \quad (10.10)$$

En reportant (10.10) dans l'expression (10.5) du facteur d'amplification g , on obtient

$$g(\theta) = e^{i\nu\theta} - 2^p i^{p+1} \alpha_{s,p} \int_0^\theta \sin\left(\frac{\theta-\varphi}{2}\right)^p e^{i(s+\nu-\frac{p}{2})\varphi} e^{-i(s-\frac{p}{2})\theta} d\varphi$$

$$g(\theta) = e^{i\nu\theta} (1 - 2^p i^{p+1} \alpha_{s,p} \int_0^\theta \sin\left(\frac{\theta-\varphi}{2}\right)^p e^{i(s+\nu-\frac{p}{2})(\varphi-\theta)} d\varphi) \quad (10.11)$$

Le terme intégral au second membre de (10.11), différence entre le facteur d'amplification $g(\theta)$ du schéma et celui de l'équation $e^{i\nu\theta}$, a maintenant une forme explicite relativement simple. Elle permet notamment de montrer que les schémas d'ordre impair sont dissipatifs d'ordre $p+1$, *i.e.* $|g(\theta)| \leq 1 - c(s)\nu(1-\nu)\theta^{p+1}$.

Després en déduit la stabilité en norme L^2 de ces schémas. Comme leur équation équivalente à un terme est dissipative au même ordre, ce résultat permet aussi de prouver (chapitre 4) la convergence aux temps longs de la solution de l'équation équivalente, à un ou deux termes, vers celle du schéma. Des exemples d'analyse de schémas de Strang de différents ordres, avec des comparaisons aux résultats numériques, sont donnés dans le chapitre 5.

Coefficients de l'équation équivalente

Le résultat (10.11) permet également de déterminer directement les coefficients de l'équation équivalente. Il suffit de remplacer, dans l'intégrale (exprimant la différence entre le facteur d'amplification du schéma et celui de l'équation) au second membre de (10.11), l'intégrande $\sin\left(\frac{\theta-\varphi}{2}\right)^p e^{i(s+\nu-\frac{p}{2})(\varphi-\theta)}$ par son développement de Taylor. Comme

$$\sin\left(\frac{\theta-\varphi}{2}\right) = \frac{\theta-\varphi}{2} + o(\theta^2)$$

$$e^{i(s+\nu-\frac{p}{2})(\varphi-\theta)} = 1 + i\left(s+\nu-\frac{p}{2}\right)(\varphi-\theta) + o(\theta^2)$$

on obtient

$$\int_0^\theta \sin\left(\frac{\theta-\varphi}{2}\right)^p e^{i(s+\nu-\frac{p}{2})(\varphi-\theta)} d\varphi \approx \int_0^\theta \left(\frac{\theta-\varphi}{2}\right)^p (1 + i\left(s+\nu-\frac{p}{2}\right)(\varphi-\theta)) d\varphi$$

Les intégrales se calculent explicitement

$$\int_0^\theta \left(\frac{\theta-\varphi}{2}\right)^p d\varphi = -\int_{-\theta}^0 \left(\frac{y}{2}\right)^p dy = 2^{-p} \int_0^{-\theta} y^p dy = \frac{1}{p+1} (-1)^p 2^{-p} \theta^{p+1}$$

donc le coefficient du premier terme de l'équation équivalente s'écrit

$$\frac{1}{p+1} (-i)^{p+1} \alpha_{s,p} \quad (10.12)$$

et le second

$$\frac{1}{p+2}(-i)^{p+2}\left(s+v-\frac{p}{2}\right)\alpha_{s,p} \quad (10.13)$$

Les coefficients des termes de dissipation et de dispersion de l'équation équivalente sont ainsi calculés sans passer par la procédure, quelque peu laborieuse pour les schémas d'ordre élevé, de leur calcul direct à partir de l'expression du schéma et du développement du logarithme de son facteur d'amplification. Par exemple, pour le schéma de Beam-Warming, on a $\alpha_{0,2} = \frac{1}{2}v(v-1)(v-2)$, le coefficient du terme de dispersion vaut

$$(-i)^3\frac{1}{6}v(v-1)(v-2) = \frac{i}{6}v(v-1)(v-2)$$

et celui du terme de dissipation

$$\frac{1}{8}v(v-1)^2(v-2)$$

Pour le schéma LW, $\alpha_{1,2} = \frac{1}{2}v(1-v)^2$ le coefficient du terme de dispersion vaut

$$\frac{i}{6}v(1-v)^2$$

et celui du terme de dissipation

$$\frac{1}{8}v^2(1-v)^2$$

Pour les schémas d'ordre impair $(s, 2s+1)$, le terme de dissipation domine, il vaut $\frac{1}{p+1}(-1)^s\alpha_{s,2s+1}\theta^{p+1}$. Le terme de dispersion $\frac{i}{p+2}(-1)^{k+1}(s+v-\frac{1}{2})\alpha_{k,2k+1}\theta^{p+2}$ est secondaire, dans la mesure où l'équation équivalente à un terme rend déjà bien compte de la solution du schéma.

Pour les schémas d'ordre pair $(s, 2s)$, le terme de dispersion domine, il vaut $\frac{i}{p+1}(-1)^{s+1}\alpha_{s,2s}\theta^{p+1}$. Le terme de dissipation $\frac{1}{p+2}(-1)^{k+1}\alpha_{k,2k+1}\theta^{p+2}$ doit être pris en compte pour bien prédire la solution du schéma. Il en va de même pour les schémas d'ordre pair $(s, 2s+2)$.

La plupart des schémas utilisés dans le corps du texte sont des schémas de Strang. Mais, comme des schémas centrés et décalés apparaissent aussi, il nous a paru utile de présenter brièvement aussi ces deux manières de monter en ordre.

10.2 Montée en ordre en espace

10.2.1 Montée en ordre par corrections successives : schémas centrés

L'idée la plus simple pour monter en ordre, partant de l'opérateur D_0 , d'ordre 2, défini dans le chapitre 3

$$D_0 u = u_x + \frac{h^2}{6} u_{xxx} + O(h^4)$$

est de soustraire de D_0 le terme dominant $\frac{h^2}{6} u_{xxx}$ de l'erreur, u_{xxx} est approché par $D_0 D_+ D_-$, ce qui donne l'opérateur « corrigé » d'ordre 4 : $Q_4 = D_0 - \frac{h^2}{6} D_0 D_+ D_-$. Dans l'espace réciproque, $\widehat{D}_0 = \frac{i}{h} \sin \xi$, $\widehat{Q}_4 = \frac{i}{h} \sin \xi (1 + \frac{2}{3} \sin^2(\frac{\xi}{2}))$: on a corrigé le second terme du développement de Taylor de $\sin \xi \simeq \xi - \frac{\xi^3}{6}$ en lui ajoutant $\frac{2}{3} \sin^2(\frac{\xi}{2})$ qui compense ce second terme, et fournit un opérateur d'ordre 4. Par corrections successives, on obtient ainsi des schémas d'ordre 6, 8, etc. Systématisons cette idée en écrivant l'opérateur d'ordre p (pair) sous la forme

$$Q_p = D_0 \sum_{m=0}^{p/2-1} a_m (-1)^m (h D_+ D_-)^m$$

et déterminons les a_m pour que Q_p approxime l'opérateur dérivation en x à l'ordre p , ce qui s'écrit, en Fourier (\widehat{Q}_p est un simple opérateur multiplicatif)

$$\widehat{Q}_p = ik + O(h^p k^{p+1})$$

Avec $\xi = kh$

$$\sin \xi \sum_{m=0}^{p/2-1} a_m (-1)^m \left(2 \sin \left(\frac{\xi}{2} \right) \right)^{2m} = \xi + O(\xi^{p+1})$$

soit, en posant $\theta = \sin(\frac{\xi}{2})$,

$$\theta \sqrt{1 - \theta^2} \sum_{m=0}^{p/2-1} a_m (-1)^m (2\theta)^{2m} = \arcsin \theta + O(\theta^{p+1})$$

ou encore

$$\sum_{m=0}^{p/2-1} a_m (-1)^m (2\theta)^{2m} = \frac{\arcsin \theta}{\theta \sqrt{1 - \theta^2}} + O(\theta^p)$$

En somme, les a_m sont les coefficients du développement de Taylor en 0 de $\frac{\arcsin \theta}{\theta \sqrt{1-\theta^2}}$. Ensuite, les coefficients du schéma se calculent à partir des a_m et de l'expression de Q_p . Par exemple, pour le schéma d'ordre 6, on obtient un schéma à 7 points, avec les coefficients $(-\frac{1}{60}, \frac{3}{20}, -\frac{3}{4}, 0, \frac{3}{4}, -\frac{3}{20}, \frac{1}{60})$. \widehat{Q}_p est imaginaire, donc le facteur d'amplification du schéma semi-discret est de la forme $\exp(iR(\xi))$, $R \in \mathbb{R}$, donc de norme 1. Les schémas obtenus sont stables et non dissipatifs (en version semi-discrète) : ils conservent la norme L^2 . Le premier terme de l'équation équivalente pour ce type de discrétisation spatiale est un terme d'ordre impair, en $\frac{\partial^{p+1}}{\partial x^{p+1}}$. L'erreur porte sur la phase : ces schémas sont dispersifs.

10.2.2 Schémas sur grilles décalées

L'idée est maintenant de stocker certaines quantités aux cloisons (indices j) et d'autres aux points milieu (indices $j + \frac{1}{2}$) pour obtenir le bon centrage. Ces schémas décalés (ou *staggered*) sont surtout utilisés pour des systèmes. Prenons l'exemple de l'équation des ondes $\partial_{tt} u - a^2 \partial_{xx} u = 0$, réécrite comme un système $\partial_t p = a^2 \partial_x u$, $\partial_t u = \partial_x p$. Centrons les p aux points $j + 1/2$ et les u aux points j , soit

$$p_t = a^2 u_x \Rightarrow \frac{d}{dt} p_{j+1/2} = a^2 \frac{u_{j+1} - u_j}{h}, \quad u_t = p_x \Rightarrow \frac{d}{dt} u_j = a^2 \frac{p_{j+1/2} - p_{j-1/2}}{h}$$

On obtient ainsi automatiquement l'ordre 2 avec une erreur divisée par 4 par rapport au schéma centré. En effet, $D_0 u(x) = \frac{u(x+h) - u(x-h)}{2h} \approx u_x + \frac{1}{6} h^2 u_{3x}$, alors que $D_s u(x) = \frac{u(x+h/2) - u(x-h/2)}{2h} \approx u_x + \frac{1}{24} h^2 u_{3x}$. Un exemple de schéma décalé est le schéma de Yee, très utilisé pour les équations de Maxwell, et étudié dans le chapitre 6. Dans ce schéma, la grille est décalée en espace et en temps. Il s'écrit donc (ordre 2 en espace et en temps)

$$\frac{p_{j+1/2}^{n+1/2} - p_{j+1/2}^{n-1/2}}{\tau} = a^2 \frac{u_{j+1}^n - u_j^n}{h}, \quad \frac{u_j^{n+1} - u_j^n}{\tau} = \frac{p_{j+1/2}^{n+1/2} - p_{j-1/2}^{n+1/2}}{h}$$

Il est possible de monter en ordre en espace comme pour les schémas classiques. Par exemple, le schéma d'ordre 4 est obtenu par correction du schéma d'ordre 2, *i.e.* en soustrayant le terme d'erreur principal $\frac{1}{24} h^2 u_{3x}$. L'opérateur de triple différentiation est approché par une composition d'opérateurs différentiels, comme précédemment.

$$D_s^2 u(x) = \frac{u(x+h/2) - u(x-h/2)}{2h} \approx u_x + \frac{1}{24} h^2 u_{3x}$$

$$\Rightarrow D_s^4 u(x) = D_s^2 u(x) - \frac{1}{24} h^2 D_+ D_-^2 u(x+h/2)$$

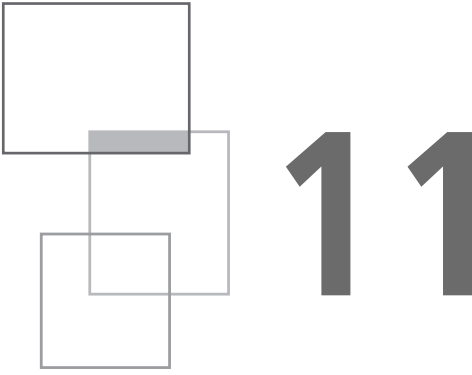
On obtient ainsi les coefficients pour des schémas décalés d'ordre arbitraire que nous donnons pour l'ordre 4 : $-\frac{1}{24}, -\frac{9}{8}, \frac{9}{8}, \frac{1}{24}$.

10.2.3 Conclusion

Nous avons présenté, dans la première partie de ce chapitre les schémas de Strang, discrétisation simultanée en temps et en espace, en suivant [1], et, dans la seconde, deux manières de monter en ordre en espace : schémas centrés et schémas sur grilles décalées. Il y en a bien sûr beaucoup d'autres : schémas d'interpolation et schémas spectraux, schémas compacts notamment, pour lesquelles nous renvoyons, ainsi que pour les techniques de montée en ordre en temps à [2].

Bibliographie

- [1] B. Després “Uniform Asymptotic Stability of Strang’s Explicit Compact Schemes for Linear Advection”, *SIAM J. Num. An.*, Vol. 47, issue 5, 2009
- [2] B. Gustafsson “High Order Difference Methods for Time Dependent PDE” Springer, 2008



Quelques fonctions spéciales

Nous avons rencontré dans le cours du texte un certain nombre de fonctions spéciales apparaissant dans les solutions exactes ou approchées des schémas. Nous présentons en particulier la fonction d'Airy, qui est une approximation de la fonction de Green des schémas d'ordre 2 pour les équations hyperboliques, et ses généralisations, utiles pour les schémas d'ordre plus élevé.

11.1 La fonction d'Airy et ses généralisations

11.1.1 *La fonction d'Airy*

Cette fonction, introduite par l'astronome George Biddell Airy, apparaît dans plusieurs domaines de la physique, dont celui de la diffraction des ondes, où elle décrit le champ au voisinage des enveloppes de rayons, appelées caustiques. Au moins un ouvrage complet [6], et des chapitres dans la plupart des livres sur les fonctions spéciales lui ont été consacrés. Nous nous bornerons à en rappeler les principales propriétés, et à décrire les méthodes que nous avons utilisées pour la calculer.

Définition de la fonction

Elle est définie pour $t \in \mathbb{R}$ par l'intégrale (11.1), impropre, car le module de l'intégrande est constant

$$Ai(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(izt + \frac{i}{3}z^3\right) dz \quad (11.1)$$

Une autre représentation, montrant que Ai est réelle, est parfois utilisée

$$Ai(t) = \frac{1}{\pi} \int_0^{+\infty} \cos\left(zt + \frac{1}{3}z^3\right) dz$$

La fonction d'Airy est bornée sur \mathbb{R} . Elle est oscillante pour $t < 0$, exponentiellement décroissante pour $t > 0$. Elle apparaît (sur \mathbb{R}) comme fonction de Green des schémas d'ordre 2 dans l'approximation équation équivalente. C'est une fonction analytique sur \mathbb{C} . Pour $t \in \mathbb{C}$, elle est définie comme

$$Ai(t) = \frac{1}{2\pi i} \int_C \exp\left(-wt + \frac{1}{3}w^3\right) dw$$

Le contour C part de $\infty e^{-i\pi/3}$ et finit à $\infty e^{i\pi/3}$. Avec le changement de variable $w = -iz$

$$Ai(t) = \frac{1}{2\pi} \int_C \exp\left(izt + \frac{i}{3}z^3\right) dz$$

Le contour part de $\infty e^{i5\pi/6}$ et finit à $\infty e^{i\pi/6}$, il est équivalent à l'axe réel, et on retrouve la définition précédente de la fonction.

Calcul de la fonction

Des algorithmes de calcul sont disponibles pour cette fonction [4]. Ils utilisent, pour de petites (resp. grandes) valeurs de l'argument t , la série de Taylor (resp. les développements asymptotiques), et ont recours à l'intégration numérique pour $t = O(1)$. Mais la fonction à intégrer oscille et ne décroît pas, ce qui rend délicat le calcul de l'intégrale. Même pour des valeurs modestes de t , l'intégration numérique est imprécise. Le chemin d'intégration initial suivant l'axe réel sera donc déformé dans le plan complexe, pour se ramener à une intégrale moins oscillante, et plus rapidement décroissante.

Le contour d'intégration peut être déformé pour $|z| \rightarrow \infty$ dans les secteurs où $|\exp(\frac{i}{3}z^3)| = \exp(-\frac{1}{3}|z|^3 \sin 3\theta) \rightarrow 0$, soit $\sin 3\theta > 0$. Deux options sont possibles. La première est d'utiliser le contour de descente rapide (CDR). L'avantage est que l'intégrande n'oscille pas et décroît rapidement, mais il faut le déterminer puis calculer le dz suivant ce contour. La seconde est de choisir un contour plus simple, moins optimal pour l'intégration, au sens où l'intégrande décroît moins vite et oscille plus que sur le CDR. L'avantage est que le contour est donné, ainsi que le dz . Nous présentons les deux options, en commençant par la seconde.

Contour composé de demi-droites

L'intérêt de se limiter à des demi-droites est que l'intégrale s'écrit plus simplement que pour un contour général, car le dz dans l'intégrale a un argument fixe.

Le nouveau contour d'intégration est composé des demi-droites $D(\theta)$ et $D(\pi - \theta)$ d'angle θ et $\pi - \theta$ passant par l'origine, avec $\theta \in [0, \pi/3]$. L'intégrale sur cette demi-droite $D(\theta)$ est notée

$$I_+ = \int_{D(\theta)} \exp\left(izt + \frac{i}{3}z^3\right) dz \quad (11.2)$$

celle sur $D(\pi - \theta)$

$$I_- = \int_{D(\pi-\theta)} \exp\left(izt + \frac{i}{3}z^3\right) dz \quad (11.3)$$

Sur $D(\theta)$, $z = |z|(\cos \theta + i \sin \theta)$, $\exp(izx + \frac{i}{3}z^3) = \exp(i|z|x(\cos \theta + i \sin \theta) + \frac{i}{3}|z|^3(\cos 3\theta + i \sin 3\theta))$. La phase de l'intégrande a donc

- une partie imaginaire : $|z|x \cos \theta + \frac{1}{3}|z|^3 \cos 3\theta$. C'est elle qui génère les oscillations de l'intégrande ;
- une partie réelle : $-|z|x \sin \theta - \frac{1}{3}|z|^3 \sin 3\theta$. La décroissance de la partie réelle facilite l'intégration numérique.

Le choix $\theta = \pi/6$ permet, d'une part de minimiser les oscillations en annulant $|z|^3 \cos 3\theta$ et, d'autre part de maximiser la décroissance en maximisant $|z|^3 \sin 3\theta$. Sur la demi-droite $D(\pi/6)$,

$$iz = itl \left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6} \right), iz^3 = -\frac{1}{3}l^3, dz = \left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6} \right) dl$$

L'intégrale (11.2) s'écrit donc

$$\begin{aligned} I_+ &= \frac{1}{2\pi} \int_0^{+i\frac{\pi}{6}\infty} \exp\left(izt + \frac{i}{3}z^3\right) dz \\ &= \frac{1}{2\pi} \int_0^{+\infty} \exp\left(itl \left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6} \right) - \frac{1}{3}l^3\right) \left(\cos \frac{\pi}{6} + i \sin \frac{\pi}{6} \right) dl \end{aligned}$$

La partie réelle de l'intégrande de I_+ vaut $\frac{1}{2}e^{-\frac{1}{2}tl}(-\sin \frac{1}{2}\sqrt{3}lt + \sqrt{3} \cos \frac{1}{2}\sqrt{3}lt)$. On obtient le même résultat pour I_-

$$\begin{aligned} I_- &= \frac{1}{2\pi} \int_{+i\frac{5\pi}{6}\infty}^0 \exp\left(izt + \frac{i}{3}z^3\right) dz \\ &= \frac{1}{2\pi} \int_0^{+\infty} \exp\left(itl \left(-\cos \frac{\pi}{6} + i \sin \frac{\pi}{6} \right) - \frac{1}{3}l^3\right) \left(\cos \frac{\pi}{6} - i \sin \frac{\pi}{6} \right) dl \end{aligned}$$

Au final, il suffit donc de calculer, pour obtenir $Ai(t)$

$$2 \operatorname{Re}(I_+) = \frac{1}{2\pi} \int_0^{+\infty} \left(-\sin \frac{\sqrt{3}}{2} lt + \sqrt{3} \cos \frac{\sqrt{3}}{2} lt \right) \exp\left(-\frac{1}{2} tl\right) \exp\left(-\frac{1}{3} l^3\right) dl \quad (11.4)$$

Nous nous sommes assurés que le calcul numérique de l'intégrale permet de retrouver avec au moins 4 chiffres significatifs la valeur tabulée de la fonction pour quelques valeurs de t : 1, $\frac{1}{2}$, 0, -1, -2 de l'ordre de l'unité. Nous avons comparé, pour des valeurs négatives de t , le résultat de l'intégration numérique de 11.2 à celui du développement asymptotique (11.12) à un $\frac{1}{\sqrt{\pi}} \cos\left(\frac{2}{3}|t|^{\frac{3}{2}} - \frac{\pi}{4}\right)$ ou deux termes, qui intègre une correction multiplicative :

$$\left(1 + \frac{5}{48} |t|^{-\frac{3}{2}} \sin\left(\frac{2}{3}|t|^{\frac{3}{2}} - \frac{\pi}{4}\right) \right).$$

t	Intégration	Développement à 1 terme	Développement à 2 termes
-5	0,350 76	0,349 68	0,350 9
-7	0,184 28	0,185 97	0,185 09
-10	$4,024 1 \times 10^{-2}$	$3,920 9 \times 10^{-2}$	$3,933 8 \times 10^{-2}$
-20	-0,176 41	-0,176 64	-0,176 80
60	$5,621 5 \times 10^{-2}$	$5,610 6 \times 10^{-2}$	$5,610 6 \times 10^{-2}$
100	$4,646 9 \times 10^{-2}$	$4,577 8 \times 10^{-2}$	$4,577 8 \times 10^{-2}$

En pratique, l'intégration est faite jusqu'à une borne B . La borne de l'intégrale est choisie pour que la partie \int_B^{∞} soit petite. Grâce au facteur $\exp\left(-\frac{1}{3}l^3\right)$, $B = 5$ est largement suffisant.

Intégration le long du CDR

Le contour de descente rapide (CDR) est celui où la partie imaginaire de la phase est constante (on peut donc l'extraire) et donc, puisque la phase de l'intégrande est une fonction analytique, où sa partie réelle varie rapidement, *i.e.* décroît de part et d'autre d'un maximum. Il est utilisé dans la méthode du col pour calculer une approximation explicite de l'intégrale pour de grandes valeurs de l'argument. Les résultats de cette section sont présentés d'une part pour donner une méthode de grande précision de calcul de la fonction d'Airy, et d'autre part comme exemple de mise en œuvre de la méthode du col.

Cas $t > 0$

Traisons d'abord le cas $t > 0$. Le changement de variable $z' = t^{1/2}z$ permet de réécrire l'intégrale sous la forme la plus classique pour appliquer la méthode du col.

$$Ai(t) = \frac{t^{1/2}}{2\pi} \int_{-\infty}^{+\infty} \exp it^{3/2} \left(z + \frac{1}{3}z^3 \right) dz \quad (11.5)$$

Les chemins de descente rapide passent par les points col vérifiant $z_c^2 + 1 = 0$: $z_c = \pm i$. Le CDR passant par $z_c = +i$ a pour asymptotes les demi-droites de décroissance rapide du terme dominant de $-\frac{1}{3}|z|^3 \sin 3\theta$ soit $\theta = \pi/6, \theta = 5\pi/6$. La déformation de l'axe réel sur ce CDR est licite, *i.e.* il est équivalent au contour initial. En effet, les sections de cercle $e^{i\theta}R, \theta \in [0, \pi/6[, [5\pi/6, \pi[$ permettent de fermer le contour, ce n'est pas le cas pour le CDR passant par $z_c = -i$.

Au point col i , $iz_c + \frac{i}{3}z_c^3 = -\frac{2}{3}$, la partie imaginaire $\text{Im}(iz + \frac{i}{3}z^3)$ est nulle. Le CDR a donc pour équation $\text{Im}(iz + \frac{i}{3}z^3) = 0$, soit

$$\text{Im}(i(x + iy) + \frac{i}{3}(x + iy)^3) = \frac{1}{3}x^3 - xy^2 + x = 0$$

Cette équation a 2 racines $y_{\pm} = \pm\sqrt{1 + \frac{1}{3}x^2}$. Le CDR est décrit par y_+ : pour $x \rightarrow \infty, y_+ \sim \frac{\sqrt{3}}{3}x$. Par symétrie, nous nous ramenons à l'intégration sur le demi-CDR C_+ situé dans le demi-plan $x \geq 0$, et (11.5) devient

$$Ai(t) = \frac{t^{1/2}}{\pi} \text{Re} \int_{-\infty}^{+\infty} \exp it^{3/2} \left(z + \frac{1}{3}z^3 \right) dz$$

Le long de $C_+, \text{Re}(i(z + \frac{1}{3}z^3)) = -\frac{2}{3}\sqrt{1 + \frac{1}{3}x^2} \left(\frac{4}{3}x^2 + 1 \right)$, d'où l'expression (11.6) de la fonction, pour $t > 0$

$$Ai(t) = \frac{t^{1/2}}{\pi} \int_0^{+\infty} \exp \left(-\frac{2}{3}t^{3/2} \sqrt{1 + \frac{1}{3}x^2} \left(\frac{4}{3}x^2 + 1 \right) \right) dx \quad (11.6)$$

Par exemple, $Ai(1) = \frac{1}{\pi} \int_0^{+\infty} \exp \left(-\frac{2}{3}\sqrt{1 + \frac{1}{3}x^2} \left(\frac{4}{3}x^2 + 1 \right) \right) dx = 0,135\,29$. On retrouve le résultat des tables [1].

Cas $t < 0$

Pour $t < 0$,

$$Ai(t) = \frac{|t|^{1/2}}{2\pi} \int_{-\infty}^{+\infty} \exp i|t|^{3/2} \left(-z + \frac{1}{3}z^3 \right) dz$$

Les points col sont donc $z_c = \pm 1$. Nous nous ramenons, par symétrie, au demi-CDR noté C_d , passant par $z_c = +1$, situé dans le demi-plan de droite $x \geq 0$. Le

demi-CDR passant par $z_c = -1$ est son symétrique miroir, si bien que la fonction est deux fois la partie réelle de l'intégrale le long de la courbe C_d

$$Ai(t) = \frac{|t|^{1/2}}{\pi} \operatorname{Re} \int_{C_d} \exp i |t|^{3/2} \left(-z + \frac{1}{3} z^3 \right) dz \quad (11.7)$$

La courbe C_d est composée de deux branches ayant pour équation

– pour $x \in [0, 1]$

$$y = -\sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2} \quad (11.8)$$

– pour $x \in [1, \infty[$

$$y = \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2} \quad (11.9)$$

Les asymptotes respectives sont les demi-droites $\theta = -\pi/2, \theta = \pi/6$. L'exposant dans (11.7) s'écrit

$$i \left(-z + \frac{1}{3} z^3 \right) = y - x^2 y + \frac{1}{3} y^3 - \frac{2}{3} i = -\frac{2}{9x} (2x + 1)^2 (x - 1) y - \frac{2}{3} i$$

Sa partie imaginaire est constante, et vaut $-\frac{2}{3}$, d'où un facteur $\cos \frac{2}{3} |t|^{3/2} - i \sin \frac{2}{3} |t|^{3/2}$. La partie réelle de cette quantité introduit une décroissance exponentielle $\exp(-|t|^{3/2} \frac{2}{9x} (2x + 1)^2 (x - 1) y)$ de l'intégrande de (11.7).

Considérons la partie de C_d allant de 1 à $\infty e^{i\pi/6}$, sa contribution à (11.7) s'écrit donc

$$C = \frac{1}{\pi} \operatorname{Re} \int_1^{+\infty} \exp(-|t|^{3/2} \frac{2}{9x} (2x + 1)^2 (x - 1) y) \times \left(\cos \frac{2}{3} |t|^{3/2} - i \sin \frac{2}{3} |t|^{3/2} \right) (dx + idy)$$

Mais, puisque $x \in [1, \infty[$, y est donné par (11.9), donc

$$dy = \frac{d}{dx} \left(\sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2} \right) = \frac{1}{3} \frac{\sqrt{3}}{x^2} \frac{x^3 - 1}{\sqrt{\frac{1}{x} (x - 1)^2 (x + 2)}}$$

C est donc la somme des deux intégrales

$$I_1 = \frac{1}{\pi} \int_1^{+\infty} \exp \left(-|t|^{3/2} \frac{2}{9x} (2x + 1)^2 (x - 1) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2} \right) \left(\cos \frac{2}{3} |t|^{3/2} \right) dx$$

$$I_2 = \frac{1}{\pi} \int_1^{+\infty} \exp\left(-|t|^{3/2} \frac{2}{9x} (2x+1)^2 (x-1) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) \\ \times \left(\sin \frac{2}{3} |t|^{3/2}\right) \frac{\sqrt{3}}{3} \frac{x^3 - 1}{\sqrt{x^3 (x-1)^2 (x+2)}} dx$$

Considérons maintenant la partie de C_d allant de $\infty e^{-i\pi/2}$ à 1, où $y = -\sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}$, y est donné par (11.9), donc

$$\frac{d}{dx} \left(-\sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) = -\frac{1}{3} \frac{\sqrt{3}}{x^2} \frac{x^3 - 1}{\sqrt{\frac{1}{x} (x-1)^2 (x+2)}} \\ i \left(-z + \frac{1}{3}z^3\right) = -\frac{2}{9x} (2x+1)^2 (x-1) y - \frac{2}{3}i$$

Sa contribution à (11.7) s'écrit comme la somme des deux intégrales

$$I_3 = \frac{1}{\pi} \int_0^1 \exp\left(-|t|^{3/2} \frac{2}{9x} (2x+1)^2 (1-x) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) \left(\cos \frac{2}{3} |t|^{3/2}\right) dx$$

$$I_4 = \frac{1}{\pi} \int_0^1 \exp\left(-|t|^{3/2} \frac{2}{9x} (2x+1)^2 (1-x) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) \\ \times \left(\sin \frac{2}{3} |t|^{3/2}\right) \frac{\sqrt{3}}{3} \frac{1 - x^3}{\sqrt{x^3 (1-x)^2 (x+2)}} dx$$

En somme, la fonction est la somme des 4 intégrales ci-dessous, multipliée par $|t|^{1/2}$

$$Ai(t) = \frac{|t|^{1/2}}{\pi} (I_1 + I_2 + I_3 + I_4) \tag{11.10}$$

Calculons à titre d'exemple $Ai(-1)$. Intégrant numériquement, on obtient

$$\int_1^{+\infty} \exp\left(-\frac{2}{9x} (2x+1)^2 (x-1) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) \left(\cos \frac{2}{3}\right) dx = 0,471\ 68$$

$$\int_1^{+\infty} \exp\left(-\frac{2}{9x} (2x+1)^2 (x-1) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) \left(\sin \frac{2}{3}\right) \frac{\sqrt{3}}{3} \frac{x^3-1}{\sqrt{x^3(x-1)^2(x+2)}} dx = 0,315\ 84$$

$$\int_0^1 \exp\left(-\frac{2}{9x} (2x+1)^2 (1-x) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) \left(\cos \frac{2}{3}\right) dx = 0,408\ 70$$

$$\int_0^1 \exp\left(-\frac{2}{9x} (2x+1)^2 (1-x) \sqrt{\frac{2}{3x} - 1 + \frac{1}{3}x^2}\right) \left(\sin \frac{2}{3}\right) \frac{\sqrt{3}}{3} \frac{1-x^3}{\sqrt{x^3(1-x)^2(x+2)}} dx = 0,486\ 29$$

$Ai(-1) \approx \frac{1}{\pi}(0,471\ 68 + 0,315\ 84 + 0,408\ 70 + 0,486\ 29) = 0,535\ 56$, ce qui est bien la valeur des tables.

La méthode est adaptée pour la fonction d'Airy classique, car l'équation du CDR est explicite. Elle permet, associée à une méthode d'intégration numérique adaptée, de calculer précisément la fonction. Nous n'avons toutefois pas trouvé, pour $t < 0$, de différence significative avec la méthode de la section précédente, *i.e.* entre (11.4) et (11.10). Pour les fonctions d'Airy généralisées, il faut d'abord déterminer le CDR, ensuite intégrer dans le plan complexe avec un dz suivant la tangente au contour, donc calculer une intégrale avec un $dz = dx + idy$ général. C'est ce qui motive les contours simplifiés proposés dans la section précédente, et l'utilisation de développements asymptotiques, qui fournissent des formules explicites.

Développements asymptotiques

Le calcul d'intégrale est nécessaire pour $t = O(1)$. Pour $|t|$ assez grand, la fonction peut être approchée par son développement asymptotique, obtenu par la méthode du col. Le développement valable pour t complexe s'écrit

$$Ai(t) \approx \frac{1}{2\sqrt{\pi}} t^{-1/4} e^{-\frac{2}{3}t^{3/2}}, \text{ pour } -\pi + \epsilon < \arg t < \pi - \epsilon$$

La branche de coupure pour $t^{1/2}$ et $t^{-1/4}$ est le demi-axe réel < 0 , et nous choisissons la détermination principale : si $t = |t| e^{i\theta}$, $t^{1/2} = |t|^{1/2} e^{i\theta/2}$. Nous utilisons principalement le développement pour t réel donné ci-dessous.

Pour t réel positif, le résultat précédent s'applique avec $\arg t = 0$, $Ai(t) \approx \frac{1}{2\sqrt{\pi}} t^{-1/4} e^{-\frac{2}{3}t^{3/2}}$. C'est la contribution du point col $z = +i\sqrt{t}$.

Pour t réel négatif, le résultat précédent ne s'applique pas car $\arg t = \pi$. Le développement est la somme des contributions des points col $z = \pm\sqrt{-t}$. Il s'obtient aussi par la méthode de la phase stationnaire. Son premier terme fournit l'approximation

$$Ai(t) \approx \frac{1}{\sqrt{\pi}} |t|^{-1/4} \sin\left(\frac{2}{3}|t|^{3/2} + \frac{\pi}{4}\right) = \frac{1}{\sqrt{\pi}} |t|^{-1/4} \cos\left(\frac{2}{3}|t|^{3/2} - \frac{\pi}{4}\right) \tag{11.11}$$

Posant $\frac{2}{3}|t|^{3/2} = \zeta$, le développement complet s'écrit

$$Ai(t) \approx \frac{1}{\sqrt{\pi}} |t|^{-1/4} \left(\cos\left(\zeta - \frac{\pi}{4}\right) \sum_0^\infty (-1)^k c_{2k} \zeta^{-2k} + \sin\left(\zeta - \frac{\pi}{4}\right) \sum_0^\infty (-1)^k c_{2k+1} \zeta^{-2k-1} \right)$$

Les coefficients valent $c_k = \frac{\Gamma(3k + \frac{1}{2})}{54^k \Gamma(k + \frac{1}{2}) k!}$ donc : $c_1 = \frac{\Gamma(\frac{7}{2})}{54 \Gamma(\frac{3}{2})} = \frac{5}{72}$. Le second terme du développement introduit donc un facteur correctif multiplicatif $\left(1 + \frac{5}{72} \frac{3}{2} |t|^{-\frac{3}{2}} \sin\left(\frac{2}{3}|t|^{3/2} - \frac{\pi}{4}\right)\right)$.

L'erreur est majorée par le terme suivant du développement, dans ce cas une erreur relative $c_2 \zeta^{-2k} = \frac{\Gamma(\frac{13}{2})}{54^2 \Gamma(\frac{5}{2})^2} \zeta^{-4}$. Elle est faible même pour des valeurs de l'argument de quelques unités (pour $t = -5$, $\frac{\Gamma(\frac{13}{2})}{54^2 \Gamma(\frac{5}{2})^2} \left(\frac{2}{3} |5|^{\frac{3}{2}}\right)^{-4} \approx 1,2031 \times 10^{-5}$).

11.1.2 Fonctions d'Airy généralisées

Elles sont définies spécifiquement pour décrire la solution des schémas dans l'approximation de l'équation équivalente, et n'ont pas la portée générale de la fonction d'Airy.

Fonction d'Airy dissipative

Elle est définie par

$$Ai^{3,b}(t) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(izt + i\frac{z^3}{3} - \frac{b}{4}z^4\right) dz \quad (11.12)$$

et apparaît pour calculer la fonction de Green de schémas d'ordre 2 dissipatifs (sections 5.4 et 5.5). Calculons quelques valeurs dans un cas tabulé dans [3] $b = 0,5$:

$$Ai^{3,0,5}(1) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(iz + i\frac{z^3}{3} - \frac{0,5}{4}z^4\right) dz = 0,1732$$

$$Ai^{3,0,5}(-1) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(-iz + i\frac{z^3}{3} - \frac{0,5}{4}z^4\right) dz = 0,42603$$

Elle peut être aussi ramenée à une intégrale réelle. Le chemin de descente rapide a maintenant pour asymptote l'axe réel positif. Mais, comme b est petit, le comportement de l'intégrale est piloté par le terme cubique, même pour k assez grand. Une solution possible est d'intégrer sur une demi-droite d'angle intermédiaire. Chin et Hedström [3] prennent par exemple $\theta = \frac{1}{2\pi(3+4)} = \frac{\pi}{14}$.

Fonction d'Airy Ai^4

Elle est définie par

$$Ai^4(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(izt - \frac{1}{4}z^4\right) dz = \frac{1}{\pi} \int_0^{+\infty} \cos(zt) \exp\left(-\frac{1}{4}z^4\right) dz \quad (11.13)$$

C'est une fonction paire, pour $t > 0$, $Ai^4(t) = \frac{t^{1/3}}{2\pi} \int_{-\infty}^{+\infty} \exp t^{4/3} \left(iz - \frac{1}{4}z^4 \right) dz$.

Elle est plus simple à calculer numériquement que la fonction d'Airy, puisque, sur le contour initial, l'intégrande est décroissant.

Les points col vérifient $z^3 = i$, soit $z = e^{i\pi/6}$, $z = e^{i5\pi/6}$, $z = -i$ pour $t > 0$. Les points à retenir pour le calcul du développement asymptotique sont les 2 premiers. En effet, le CDR équivalent au contour initial passe par ces points.

$\text{Im} \left(i(x + iy) - \frac{1}{4}(x + iy)^4 \right) = x - x^3y + xy^3$. L'équation du CDR est donc $x - x^3y + xy^3 = \frac{3}{8}\sqrt{3}$. Il a pour asymptotes les demi-axes. Appliquant la méthode du col, la contribution du point $e^{i\pi/6}$ vaut

$$C = \frac{1}{2\pi} \sqrt{2\pi} t^{-1/3} (-S'')^{-1/2} \exp t^{4/3} S$$

Reportons dans cette expression les valeurs de S et de S'' , soit

$$S = iz - \frac{1}{4}z^4 = ie^{i\pi/6} - \frac{1}{4}(e^{i2\pi/3}) = \frac{3}{8}i\sqrt{3} - \frac{3}{8}$$

$$-S'' = 3z^2, (-S'')^{-1/2} = (\sqrt{3}e^{i\pi/6})^{-1}$$

On obtient

$$C = \frac{1}{\sqrt{2\pi} t^{1/3}} (\sqrt{3}e^{i\pi/6})^{-1} \exp \left(i\frac{3}{8}\sqrt{3}t^{4/3} \right) \exp \left(-\frac{3}{8}t^{4/3} \right)$$

La contribution du point $e^{-i\pi/6}$ en est le conjugué. Le premier terme du développement asymptotique s'écrit donc

$$\frac{2}{\sqrt{6\pi} t^{1/3}} \cos \left(\frac{3}{8}\sqrt{3}t^{4/3} - \frac{\pi}{6} \right) \exp \left(-\frac{3}{8}t^{4/3} \right) \quad (11.14)$$

Nous comparons dans le tableau ci-dessous le développement asymptotique à l'intégration numérique pour quelques valeurs de t : la précision relative est de l'ordre du % à partir de $t = 10$

t	$Ai^4(t)$, d.a.(14)	$Ai^4(t)$ intégration (13)
2	$6,280\ 2 \times 10^{-2}$	$6,632\ 3 \times 10^{-2}$
5	$3,405\ 2 \times 10^{-3}$	$3,177\ 6 \times 10^{-3}$
10	$4,101\ 4 \times 10^{-5}$	$4,121\ 5 \times 10^{-5}$
20	$-2,405\ 4 \times 10^{-10}$	$-2,402\ 4 \times 10^{-10}$

Fonctions d'Airy d'ordre plus élevé

Elles servent à décrire les fonctions de Green des schémas d'ordre élevé. On distinguera le cas pair

$$Ai^{2p}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(izt - \frac{1}{2p}z^{2p}\right) dz$$

et le cas impair, où deux termes sont nécessaires dans l'exposant

$$Ai^{2p+1,b}(t) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(izt + i\frac{z^{2p+1}}{2p+1} - \frac{b}{2p+2}z^{2p+2}\right) dz$$

Dans le cas impair avec $b = 0$, l'intégrale définissant la fonction est une intégrale oscillante : son module est constant. Son calcul effectif passe donc par une déformation du contour, comme expliqué plus haut. Une solution simple est d'intégrer suivant la demi-droite d'angle θ où $iz^{2p+1} = ie^{(2p+1)\theta}$ réel négatif, soit $\theta = \frac{1}{2(2p+1)}\pi$.

Fonction Ai^5

Cette fonction est donnée par

$$Ai^5(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(izt + i\frac{z^5}{5}\right) dz \tag{11.15}$$

Intégration numérique

L'intégration est suivant $\theta = \frac{1}{10}\pi, i.e.$

$$Ai^5(t) = \frac{1}{\pi} \int_0^{+\infty} \exp\left(itl \left(\cos \frac{\pi}{10} + i \sin \frac{\pi}{10}\right) - \frac{1}{5}l^5\right) \left(\cos \frac{\pi}{10} + i \sin \frac{\pi}{10}\right) dl \tag{11.16}$$

Comme

$$\exp itl \left(\cos \frac{\pi}{10} + i \sin \frac{\pi}{10}\right) = \left(\cos\left(tl \cos \frac{\pi}{10}\right) + i \sin\left(tl \cos \frac{\pi}{10}\right)\right) \exp\left(-tl \sin \frac{\pi}{10}\right),$$

(11.16) devient

$$Ai^5(t) = \frac{1}{\pi} \int_0^{+\infty} \left(\cos \frac{\pi}{10} \cos\left(tl \cos \frac{\pi}{10}\right) - \sin \frac{\pi}{10} \sin\left(tl \cos \frac{\pi}{10}\right)\right) \exp\left(-tl \sin \frac{\pi}{10} - \frac{1}{5}l^5\right) dl \tag{11.17}$$

Quelques valeurs ont été calculées numériquement, et reportées dans le tableau plus bas. Elles sont proches de celles calculées par Chin et Hedström [3].

Développement asymptotique

Pour $t < 0$, il y a 2 points de phase stationnaire sur l'axe réel en ± 1 . Le changement de variable $u = |t|^{1/4} z$ permet d'écrire 11.15

$$Ai(t) = \frac{1}{2\pi} |t|^{1/4} \int_{-\infty}^{\infty} \exp\left(\frac{1}{5} i |t|^{5/4} (z^5 - 5z)\right) dz$$

Le développement asymptotique à un terme s'écrit donc

$$Ai(t) \approx \frac{1}{\sqrt{2\pi} |t|^{3/8}} \cos\left(\frac{4}{5} |t|^{5/4} - \frac{\pi}{4}\right)$$

$-t$	Développement	Intégration
6	0,184 01	0,186 21
10	0,107 91	0,109 87
20	$-8,265 6 \times 10^{-3}$	$-7,403 9 \times 10^{-3}$
40	$-4,081 4 \times 10^{-2}$	$-4,106 8 \times 10^{-2}$
60	$5,610 6 \times 10^{-2}$	$5,621 5 \times 10^{-2}$
100	$4,577 8 \times 10^{-2}$	$4,646 9 \times 10^{-2}$

Contour de descente rapide

Le calcul du CDR peut être fait explicitement, mais les formules font plusieurs lignes. Nous n'avons donc pas poursuivi dans cette voie.

Primitive de la fonction d'Airy

Elle est définie à partir de la représentation intégrale

$$PAi(t) = \frac{1}{\pi} \int_0^{+\infty} \frac{1}{z} \sin\left(zt + \frac{1}{3}z^3\right) dz$$

Avec cette définition, la primitive calculée varie entre $-\frac{1}{2}$ et $\frac{1}{2}$. En effet

$$\frac{1}{\pi} \int_0^{+\infty} \frac{1}{z} \sin(zt) dz = \frac{1}{2} \operatorname{sgn}(t)$$

Après rotation du contour de $\pi/6$, l'intégrale à calculer est

$$\int_0^{+\infty} \frac{1}{iz} \exp\left(izt + \frac{i}{3}z^3\right) dz = \int_0^{+\infty} \frac{1}{iz} \exp\left(e^{i2\pi/3}zt - \frac{1}{3}z^3\right) dz,$$

sa partie réelle $\text{Re} = \frac{1}{\pi} \int_0^{+\infty} \frac{1}{z} \sin\left(\frac{1}{2}\sqrt{3}zt\right) \exp\left(-\frac{1}{2}zt - \frac{1}{3}z^3\right) dz$

$$PAi(t) = \frac{1}{\pi} \int_0^{+\infty} \frac{1}{z} \sin\left(\frac{1}{2}\sqrt{3}zt\right) \exp\left(-\frac{1}{2}zt - \frac{1}{3}z^3\right) dz + \frac{1}{6}$$

La primitive de la fonction d'Airy dissipative s'écrit

$$PAi^{3,b}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(izt + i\frac{z^3}{3} - \frac{b}{4}z^4\right) dz = \frac{1}{\pi} \int_0^{+\infty} \frac{1}{z} \sin\left(zt + \frac{1}{3}z^3\right) \exp\left(-\frac{b}{4}z^4\right) dz$$

L'intégration numérique selon le demi-axe réel fonctionne, grâce à la décroissance exponentielle.

Conditions initiales gaussiennes et fonction d'Airy

Considérons l'équation d'advection discrétisée avec un schéma d'ordre 2. L'équation équivalente à un terme du schéma s'écrit

$$u_t + au_x = -\frac{1}{3}cu_{xxx}$$

Considérons une condition initiale gaussienne,

$$u(x, 0) = \frac{1}{2\sqrt{b\pi}} \exp\left(-\frac{x^2}{4b}\right)$$

de transformée de Fourier

$$\widehat{u}(k) = e^{-k^2 b}$$

Suivant la méthode usuelle, la solution est obtenue par transformation de Fourier

$$\widehat{u}_t + ik a \widehat{u} = ik^3 c \widehat{u}, \widehat{u} = \exp\left(i\frac{1}{3}ck^3 t - bk^2 - ikat\right)$$

Par transformation de Fourier inverse

$$u = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\left(i\frac{1}{3}ctk^3 - bk^2 + ik(x - at)\right) dx$$

L'exposant se réécrit

$$e = i\frac{1}{3}ctk^3 - bk^2 + ik(x - at) = i\frac{1}{3}ct\left(k^3 + i\frac{3b}{ct}k^2\right) + ik(x - at)$$

Complétons le cube

$$e = i \frac{1}{3} ct \left(k + i \frac{b}{ct} \right)^3 + i \frac{b^2}{c} \frac{k}{t} + ik(x - at) - \frac{1}{3} \frac{b^3}{c^2 t^2}$$

$$u = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left(i \frac{1}{3} ct \left(k + i \frac{b}{ct} \right)^3 + i \frac{b^2}{c} \frac{k}{t} + ik(x - at) - \frac{1}{3} \frac{b^3}{c^2 t^2} \right) dk$$

Prenons $k + i \frac{b}{ct} = l$

$$u = \frac{1}{2\pi} \exp \left(-\frac{1}{3} \frac{b^3}{c^2 t^2} \right) \int_{-\infty + i \frac{b}{ct}}^{\infty + i \frac{b}{ct}} \exp \left(i \frac{1}{3} c t l^3 + i \frac{b^2}{c} \frac{(l - i \frac{b}{ct})}{t} \right. \\ \left. + i \left(l - i \frac{b}{ct} \right) (x - at) \right) dl$$

$$u = \frac{1}{2\pi} \exp \left(\frac{2}{3} \frac{b^3}{c^2 t^2} + \frac{b}{ct} (x - at) \right) \int_{-\infty + i \frac{b}{ct}}^{\infty + i \frac{b}{ct}} \exp \left(i \frac{1}{3} c t l^3 + i \frac{b^2}{c} \frac{l}{t} + il(x - at) \right) dl$$

On peut ramener le contour sur l'axe réel, si, quand $R \rightarrow \infty$, $I \rightarrow 0$

$$I = \int_R^{R + i \frac{b}{ct}} \exp \left(i \frac{1}{3} c t l^3 + i \frac{b^2}{c} \frac{l}{t} + il(x - at) \right)$$

$$dl = \int_0^{\frac{b}{ct}} \exp \left(i \frac{1}{3} c t (R + iy)^3 + i \frac{b^2}{c} (R + iy) + i(R + iy)(x - at) \right) dy$$

Le terme dominant de la partie réelle de l'exposant est $-ctR^2y \rightarrow -\infty$ pour $R \rightarrow \infty$, donc $I \rightarrow 0$ et

$$u = \frac{1}{2\pi} \exp \left(\frac{2}{3} \frac{b^3}{c^2 t^2} + \frac{b}{ct} (x - at) \right) \int_{-\infty}^{\infty} \exp \left(i \frac{1}{3} c t l^3 + il \left((x - at) + \frac{b^2}{ct} \right) \right) dl$$

La solution s'exprime à l'aide d'une fonction d'Airy d'argument $\frac{1}{(ct)^{1/3}} \left((x - at) + \frac{b^2}{ct} \right)$

$$u = \frac{1}{(ct)^{1/3}} \exp \left[\frac{b}{ct} \left((x - at) + \frac{2}{3} \frac{b^2}{ct} \right) \right] Ai \left(\frac{1}{(ct)^{1/3}} \left((x - at) + \frac{b^2}{ct} \right) \right)$$

Le cas particulier $b = 1/4$, $u(x, 0) = \frac{1}{\sqrt{\pi}} \exp(-x^2)$, $(ct)^{1/3} = \alpha$ est présenté dans le livre de Vallée et Soarès [6], avec plusieurs figures. On obtient dans ce cas

$$u = \frac{1}{\alpha} \exp \left[\frac{1}{4\alpha^3} \left((x - at) + \frac{1}{24\alpha^3} \right) \right] Ai \left(\frac{1}{\alpha} \left((x - at) + \frac{1}{16\alpha^3} \right) \right)$$

Comme la condition initiale est régulière, la solution du schéma converge vers la solution de l'équation, *i.e.* la condition initiale translatée de at . Vérifions ce résultat au voisinage de $x = at$.

En $x = at$, le développement asymptotique de la fonction d'Airy $Ai\left(\frac{b^2}{(ct)^{4/3}}\right) \approx \frac{(ct)^{1/3}}{2\sqrt{\pi}b} \exp\left(-\frac{2}{3}\frac{b^3}{(ct)^2}\right)$ donne $u(x = at) \approx \frac{1}{2\sqrt{\pi}b} = \frac{1}{\sqrt{2\sigma\pi}}$. On retrouve donc $u(at, t) \approx \frac{1}{\sqrt{2\sigma\pi}} = u(0, 0)$.

Pour $(x - at) + \frac{b^2}{ct} > 0$, si $(x - at) \ll \frac{b^2}{ct}$, utilisant le développement asymptotique de la fonction d'Airy $Ai\left(\frac{1}{(ct)^{1/3}}\left((x - at) + \frac{b^2}{ct}\right)\right)$

$$\frac{1}{(ct)^{1/2}} \frac{2}{3} \left((x - at) + \frac{b^2}{ct} \right)^{\frac{3}{2}} = \frac{2}{3} \frac{b^3}{c^2 t^2} + \frac{b}{ct} (x - at) + \frac{1}{4b} (x - at)^2 + O\left((x - at)^3\right)$$

$$Ai\left(\frac{1}{(ct)^{1/3}}\left((x - at) + \frac{b^2}{ct}\right)\right) \approx \frac{(ct)^{1/3}}{2\sqrt{\pi}b} \exp\left(\frac{2}{3} \frac{b^3}{c^2 t^2} + \frac{b}{ct} (x - at) + \frac{1}{4b} (x - at)^2\right)$$

On obtient au final

$$u \approx \frac{1}{2\sqrt{\pi}b} \exp\left(-\frac{1}{4b} (x - at)^2\right)$$

soit la condition initiale translatée de at .

En conclusion, la solution de l'équation équivalente à un terme pour un schéma d'ordre 2 avec conditions initiales gaussiennes a une expression explicite, produit d'une exponentielle et d'une fonction d'Airy. Bien que nous n'ayons pas considéré ce cas dans la partie Applications, il nous a semblé utile de le présenter.

11.2 Fonctions de Bessel

Elles ont été intensivement étudiées, notamment pour résoudre des problèmes aux limites à symétrie cylindrique. Elles apparaissent dans la solution (section 5.2) du schéma centré semi-discrétisé. La fonction d'indice entier j et d'argument z est définie par

$$J_j(z) = \frac{1}{\pi} \int_0^\pi \cos(z \sin \theta - j\theta) d\theta = \frac{1}{\pi} \operatorname{Re} \int_0^\pi \exp i(z \sin \theta - j\theta) d\theta$$

Pour de grandes valeurs de leur argument ou de leur indice, elles peuvent être approchées par leur développement asymptotique, obtenu par la méthode de la phase stationnaire. Par exemple pour z grand, le point stationnaire vérifie $\cos \theta = \frac{j}{z}$. Il existe pour $j < z$, et la phase stationnaire donne le développement de Debye ([1])

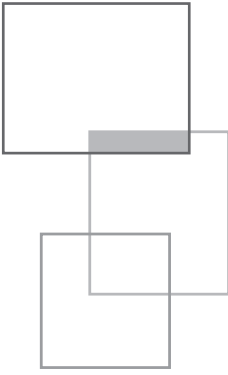
formule 9.3.3). Ce développement est valide pour des valeurs de j et z pas trop proches. Sinon, il faut utiliser le développement de Watson ([1] formule 9.3.4).

De nombreuses représentations intégrales existent, dont celle d'Eckert [1].

La référence classique sur les fonction spéciales est [1]. L'ouvrage [6] est dédié à la fonction d'Airy et à ses nombreuses applications, les articles [3] (contenant des tables numériques et graphiques), [2, 5] à ses généralisations. Enfin, le livre [4] présente un large choix de méthodes de calcul numérique pour les fonctions spéciales.

Bibliographie

- [1] M. Abramowitz, I. Stegun "Handbook of Mathematical Functions", Dover Publications, 1964
- [2] G. Biondini, T. Trogdon, "Gibbs Phenomenon for Dispersive PDEs on the Line" SIAM J. App. Mat. 77, 2007
- [3] R.C. Chin, G.W. Hedström, "A Dispersion Analysis for Difference Schemes: Tables of Generalized Airy Functions", Mathematics of Computation 32, No. 144, pp 1163-1163, 1978
- [4] A. Gil, G. Segura, N. Temme "Numerical methods for special functions" SIAM, 2007
- [5] G.W. Hedström, "Generalized Airy functions and artificial viscosities" UCRL Preprint, 1978
- [6] O. Vallée, M. Soarès "Airy functions and application to physics" Imperial College Press, 2004



12

Développements asymptotiques d'intégrales

Trois méthodes de développements asymptotiques d'intégrales, ayant des points communs, mais aussi des spécificités, sont présentées dans cet appendice : méthode de Laplace, méthode de la phase stationnaire et méthode du col. Toutes ont pour but d'estimer des intégrales, dépendant d'un grand paramètre λ , de la forme

$$I = \int_C f(x) \exp(\lambda S(x)) dx \quad (12.1)$$

Plus précisément, elles permettent d'obtenir le développement asymptotique de I pour $\lambda \rightarrow \infty$. Les hypothèses sur les fonctions S et f et sur le contour C seront définies pour chaque méthode. L'idée générale est que les contributions essentielles à l'intégrale viennent de voisinages de points particuliers (dits « critiques »), où S a un extremum. Il est donc possible d'approcher l'intégrale en se limitant à ces voisinages. L'intégrande sera approximée en remplaçant f et S par leurs développements de Taylor, et en étendant les bornes à l'infini, pour se ramener à une intégrale « étalon » explicitement calculable. Au final, des formules d'approximation explicites, ne dépendant que des valeurs de f et S et de leurs dérivées en ces points critiques seront ainsi obtenues. Ces méthodes permettent en particulier d'obtenir des approximations sous forme fermée des intégrales donnant les solutions des schémas discrétisant l'équation d'advection, obtenues par analyse de Fourier, ou par la méthode de l'équation équivalente, sous les formes respectives

$$u_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{Q}_p(k)^n \widehat{u}^0(k) e^{ikj} dk$$

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp[ik(x - at) + \sum_{m=2}^q q_m(ik)^m t] \widehat{u}^0(k) dk$$

Le grand paramètre est n ou t . Le contour est tout ou une partie de l'axe réel. Les trois méthodes sont complémentaires. La méthode de Laplace est la plus ancienne et la plus simple conceptuellement, le contour est dans \mathbb{R} et la phase S est réelle. La méthode de la phase stationnaire traite aussi de contours dans \mathbb{R} , mais avec une phase iS imaginaire. Ces deux méthodes sont en général d'application simple, et se généralisent au cas multidimensionnel. La méthode du col traite du cas où C est un contour dans le plan complexe, et où S et f sont des fonctions analytiques. Sa mise en œuvre peut s'avérer délicate, particulièrement la phase de déformation du contour initial vers le contour dit de descente rapide, passant par un ou plusieurs cols. Mais elle permet de traiter des cas hors de portée des autres méthodes. Le but du chapitre est de présenter les grandes lignes de ces trois méthodes, de rassembler les formules les plus utiles, en renvoyant aux références pour des démonstrations plus détaillées et des exemples. Nous commençons par la méthode de Laplace.

12.1 Méthode de Laplace

Elle permet d'estimer, pour des valeurs grandes du paramètre λ , des intégrales, sur un intervalle fini ou non de l'axe réel, de la forme

$$I(\lambda) = \int_a^b f(x) \exp(\lambda S(x)) dx \tag{12.2}$$

$S(x), f(x)$ sont des fonctions réelles régulières. Supposons dans un premier temps, pour présenter simplement l'idée de la méthode, que S atteint son maximum non dégénéré en un seul point, noté x_0 , *i.e.* $S'(x_0) = 0, S''(x_0) < 0$. La fonction $\exp \lambda S(x)$ a un maximum d'autant plus piqué en ce point que λ est grand. L'intégrale peut alors être approchée en ne considérant qu'un petit voisinage du point x_0 . En effet, les contributions des intervalles $]x_0 + \epsilon, \infty[$, $] -\infty, x_0 - \epsilon[$ sont d'autant plus petites, et l'approximation sera d'autant plus précise que λ est grand.

$$I(\lambda) \approx \int_{x_0 - \epsilon}^{x_0 + \epsilon} f(x) \exp(\lambda S(x)) dx$$

Remplaçons f, S par leur développement de Taylor,

$$f(x) \approx f(x_0), S(x) \approx S(x_0) + \frac{1}{2} S''(x_0)(x - x_0)^2$$

$$I(\lambda) \approx f(x_0) \exp(\lambda S(x_0)) I_g$$

où I_g est l'intégrale

$$I_g = \int_{x_0-\varepsilon}^{x_0+\varepsilon} \exp \frac{1}{2} \lambda S''(x_0) (x - x_0)^2 dx$$

Avec le changement de variable $t = \sqrt{-\lambda S''(x_0)}(x - x_0)$

$$I_g = \frac{1}{\sqrt{-\lambda S''(x_0)}} \int_{-\varepsilon \sqrt{-\lambda S''(x_0)}}^{\varepsilon \sqrt{-\lambda S''(x_0)}} \exp \left(-\frac{1}{2} t^2 \right) dt \approx \frac{\sqrt{2\pi}}{\sqrt{-\lambda S''(x_0)}}$$

Les bornes de l'intégrale tendent vers l'infini, ce qui permet de se ramener à une intégrale (gaussienne) $\int_{-\infty}^{+\infty} \exp \left(-\frac{1}{2} t^2 \right) dt = \sqrt{2\pi}$, et au final, pour $\lambda \rightarrow \infty$

$$I(\lambda) \approx \sqrt{\frac{2\pi}{-\lambda S''(x_0)}} f(x_0) \exp \lambda S(x_0) \tag{12.3}$$

Le résultat (12.3) peut être démontré plus rigoureusement comme suit. Effectuons, ce qui est possible au moins localement, au voisinage du maximum, le changement de variable $S(x) - S(x_0) = -y^2$, $x = \varphi(y)$. Restreignons l'intervalle d'intégration à $[-\delta, \delta]$. C'est licite puisque les contributions des intervalles négligés sont exponentiellement petites.

$$I(\lambda) = \exp(\lambda S(x_0)) \int_{-\delta}^{\delta} \exp(-\lambda y^2) f(\varphi(y)) \varphi'(y) dy$$

$$I(\lambda) \approx \exp(\lambda S(x_0)) \int_0^{\delta} \exp(-\lambda y^2) (f(\varphi(y)) \varphi'(y) + f(\varphi(-y)) \varphi'(-y)) dy$$

Appliquons à cette intégrale le lemme de Watson. Dans sa forme la plus générale, ce lemme fournit le développement asymptotique d'intégrales de la forme

$$\Phi(\lambda) = \int_0^a x^{\beta-1} f(x) \exp(-\lambda x^\alpha) dx$$

Il s'énonce ainsi

$$\Phi(\lambda) \approx \frac{1}{\alpha} \sum_{k=0}^{\infty} \lambda^{-(k+\beta)/\alpha} \Gamma \left(\frac{k+\beta}{\alpha} \right) \frac{f^{(k)}(0)}{k!}$$

Il est fondé sur l'identité $\int_0^\infty x^{\beta-1} \exp(-\lambda x^\alpha) dx = \lambda^{-\beta/\alpha} \Gamma(\frac{\beta}{\alpha})$, $\lambda > 0$, et le développement de Taylor de f . Nous renvoyons aux références pour la démonstration de ce lemme.

Le cas à traiter correspond à $\alpha = 2, \beta = 1$. L'intérêt de la méthode rigoureuse est qu'elle permet d'obtenir les termes successifs du développement asymptotique. On obtient ainsi :

$$I(\lambda) \approx \sqrt{\frac{2\pi}{-\lambda S''(x_0)}} \exp \lambda S(x_0) \left(f(x_0) + \sum_{k=1}^\infty a_k \frac{\Gamma(\frac{1}{2} + k)}{\Gamma(\frac{1}{2})} \lambda^{-k} \right) \quad (12.4)$$

Les coefficients s'expriment récursivement :

$$a_k = \frac{\Gamma(k + \frac{1}{2})}{2k!} \left(\frac{d}{dx} \right)^k \left[f(x) \left(\frac{2(S(x_0) - S(x))}{(x - x_0)^2} \right)^{-k-1/2} \right]_{x=x_0}$$

Par exemple, notant f_j, S_j les coefficients de Taylor respectifs de f, S au point col

$$a_1 = \frac{1}{S_2} \left(f_2 - \frac{3 S_3 f_1}{2 S_2} - \frac{3 S_4 f_0}{2 S_2} + \frac{15 S_3^2 f_0}{8 S_2^2} \right)$$

Donnons un exemple simple d'application. Considérons l'intégrale définissant la fonction Γ

$$\Gamma(n + 1) = \int_0^\infty t^n e^{-t} dt = n!$$

Calculons le premier terme de son développement asymptotique pour n grand. Le changement de variable $t = nx$ permet d'écrire l'intégrale sous la forme désirée

$$\Gamma(n + 1) = n^{n+1} \int_0^\infty \exp(n(\ln x - x)) dx$$

Le maximum de $\ln x - x = S(x)$ est atteint si $S'(x) = \frac{1}{x} - 1 = 0$, donc en $x = 1$. En ce point, $S''(1) = -1, f = 1, S(1) = -1$. On obtient la formule de Stirling

$$n! \approx n^{n+1} \sqrt{\frac{2\pi}{-nS''(1)}} f(1) \exp(nS(1)) = \sqrt{2\pi n} n^n e^{-n}$$

La méthode se généralise au cas où la première dérivée non nulle de S est $S^{[2n]}$, on obtient

$$I(\lambda) \approx \frac{1}{n} \Gamma\left(\frac{1}{2n}\right) \left(\frac{(2n)!}{-\lambda S^{[2n]}(x_0)} \right)^{1/2n} f(x_0) \exp \lambda S(x_0)$$

La méthode heuristique pour obtenir ce résultat est de remplacer les fonctions par leur développement de Taylor. Le changement par rapport au cas précédent est que $S(x) \approx S(x_0) + \frac{1}{(2n)!} S^{[2n]}(x_0)(x - x_0)^{2n}$. L'intégrale étalon est maintenant

$$\int_{-\infty}^{+\infty} \exp(-t^{2n}) dt = \frac{1}{2n} \Gamma\left(\frac{1}{2n}\right).$$

S'il y a plusieurs points col, leurs contributions s'ajoutent. Cela se produit par exemple pour les fonctions d'Airy généralisées étudiées dans le chapitre 11. Mais, en général, sauf s'il existe des symétries, une des contributions est dominante. La méthode se généralise aussi au cas multidimensionnel. En notant H le Hessien de S

$$I \approx \frac{(2\pi)^{d/2}}{\lambda^{d/2} |\det(H(x_s))|^{1/2}} f(x_s) \exp(\lambda S(x_s)) \quad (12.5)$$

12.2 Méthode de la phase stationnaire

Elle est utilisée pour calculer le développement asymptotique pour $\lambda \rightarrow \infty$ d'intégrales de la forme

$$I = \int_{-\infty}^{+\infty} f(x) \exp(i\lambda S(x)) dx \quad (12.6)$$

Le contour est l'axe réel, $f(x)$ et $S(x)$ sont des fonctions régulières, S est réelle. Pour simplifier, f est supposée à support compact, mais les résultats restent applicables même si ce n'est pas le cas. L'intégrale I doit alors être considérée comme une intégrale oscillante. Un facteur de convergence est introduit, et I est définie comme $\lim_{\epsilon \rightarrow 0} I_\epsilon$ avec $I_\epsilon = \int_{-\infty}^{+\infty} f(x) \exp(i\lambda S(x) - \epsilon x^2) dx$. Cette définition permet notamment de donner un sens à des intégrales non absolument convergentes comme $\int_{-\infty}^{+\infty} x^j \exp(ix^2) dx$. La dérivation des formules est très semblable à celle de la section précédente.

Par définition, x_s est un point de phase stationnaire si $S'(x_s) = 0$. Hors des voisinages de ces points, la phase $\lambda S(x)$ varie rapidement, car sa dérivée est $\lambda S'(x)$. Il se produit une compensation par oscillations, et la contribution à l'intégrale des intervalles sans point de phase stationnaire est petite. Plus précisément, on peut démontrer que I définie par (12.6) décroît plus vite que toute puissance de λ . C'est le lemme de la phase non stationnaire. En effet, si $S'(x) \neq 0$ sur tout le domaine d'intégration, en intégrant par parties

$$I = \int_{-\infty}^{+\infty} \frac{f(x)}{i\lambda S'(x)} i\lambda S'(x) \exp(i\lambda S(x)) dx = \frac{i}{\lambda} \int_{-\infty}^{+\infty} \left(\frac{f(x)}{S'(x)} \right)' \exp(i\lambda S(x)) dx$$

Après n intégrations par parties, licites puisque f et S sont indéfiniment dérivables,

$$I = \left(\frac{i}{\lambda}\right)^n \int_{-\infty}^{+\infty} B_n(x) \exp(i\lambda S(x)) dx, B_n = \left(\frac{B_{n-1}}{S'(x)}\right)'$$

B_n est, comme f , à support compact, $\int_{-\infty}^{+\infty} B_n(x) \exp(i\lambda S(x)) dx < C_n$, donc $|I| < C_n \lambda^{-n}, \forall n : I$ décroît donc plus vite que toute puissance négative de λ .

Ce résultat n'est pas valide si les fonctions (resp. de leur dérivée nième) ont une discontinuité. On peut montrer dans ce cas que I décroît seulement en λ^{-1} (resp. λ^{-n-1}). Cela se produit par exemple (discontinuité de la fonction f) quand on intègre sur un intervalle $[a, b]$. L'intégration par parties génère alors un terme supplémentaire :

$$\begin{aligned} I &= \int_a^b \frac{f(x)}{i\lambda S'(x)} (\exp(ikS(x)))' dx \\ &= \frac{i}{\lambda} \int_a^b \left(\frac{f(x)}{S'(x)}\right)' \exp(i\lambda S(x)) dx - \frac{i}{\lambda} \left[\frac{f(x)}{S'(x)} \exp(i\lambda S(x))\right]_a^b \end{aligned}$$

S'il existe des points de phase stationnaires, le d.a. de I est déterminé par les valeurs prises par les fonctions et leurs dérivées en ces points. Commençons par le cas où il y a un seul de ces points x_s , et où $S''(x_s) \neq 0$. L'idée est toujours, pour obtenir un résultat explicite, d'approximer f et S au voisinage de x_s par leurs développements de Taylor :

$$S(x) \approx S(x_s) + S''(x_s)(x - x_s)^2/2, f(x) \approx f(x_s)$$

Reportons ces approximations dans l'intégrale

$$I \approx f(x_s) \exp(i\lambda S(x_s)) \int_{-\infty}^{+\infty} \exp\left(i\lambda S''(x_s) \frac{(x - x_s)^2}{2}\right) dx$$

On reconnaît une intégrale gaussienne

$$\begin{aligned} \int_{-\infty}^{+\infty} \exp\left(i\lambda S''(x_s) \frac{x^2}{2}\right) dx &= \sqrt{\frac{2\pi}{\lambda |S''(x_s)|}} \exp i \frac{\pi}{4} \text{sgn}(S''(x_s)) \\ I &\approx \sqrt{\frac{2\pi}{\lambda |S''(x_s)|}} f(x_s) \exp(i\lambda S(x_s) + i \frac{\pi}{4} \text{sgn}(S''(x_s))) \end{aligned} \quad (12.7)$$

Le résultat (12.7) peut être démontré plus rigoureusement comme dans la partie consacrée à la méthode de Laplace, en effectuant le changement de variable $S(x) - S(x_0) = y^2$, puis en appliquant une variante du lemme de Watson due à Erdelyi, donnant le développement asymptotique de $\Phi(\lambda) = \int_0^a x^{\beta-1} f(x) \exp(-i\lambda x^\alpha) dx$

Il s'énonce ainsi (noter le facteur $\exp(i \frac{k+\beta}{2\alpha})$)

$$\Phi(\lambda) \approx \frac{1}{\alpha} \sum_{k=0}^{\infty} \lambda^{-(k+\beta)/\alpha} \Gamma\left(\frac{k+\beta}{\alpha}\right) \exp\left(i\frac{k+\beta}{2\alpha}\right) \frac{f^{(k)}(0)}{k!}$$

Cette méthode permet d'obtenir les coefficients successifs du développement ([3], p. 164).

S'il existe plusieurs points de phase stationnaire : $S'(x_s^i) = 0$ en x_s^i , le d.a. est la somme des « contributions de phase stationnaire ». Comme ces contributions sont toutes d'ordre $\lambda^{-\frac{1}{2}}$, elles doivent être toutes prises en compte.

Ce résultat se démontre en utilisant une partition de l'unité couvrant le support F de f : $1_F = \sum_i \varphi_i(x) + \sum_r \psi_r(x)$. Les φ_i valent 1 au voisinage de x_s^i , les ψ_r sont nulles au voisinage de tous les points stationnaires. Par le lemme de la phase non stationnaire, $\int_{-\infty}^{+\infty} \left(\sum_r \psi_r(x)\right) f(x) \exp(i\lambda S(x)) dx$ est négligeable et

$$\begin{aligned} I_i &= \int_{-\infty}^{+\infty} \varphi_i(x) f(x) \exp(i\lambda S(x)) dx \\ &\approx \sqrt{\frac{2\pi}{\lambda |S''(x_s^i)|}} f(x_s^i) \exp\left(i\lambda S(x_s^i) + i\frac{\pi}{4} \operatorname{sgn}(S''(x_s^i))\right) \\ I &\approx \sum_i I_i \approx \sum_i \sqrt{\frac{2\pi}{\lambda |S''(x_s^i)|}} f(x_s^i) \exp\left(i\lambda S(x_s^i) + i\frac{\pi}{4} \operatorname{sgn}(S''(x_s^i))\right) \end{aligned}$$

Appliquons ce résultat à la représentation intégrale de la fonction d'Airy (voir chapitre précédent) pour un argument négatif $Ai(-t) = \frac{t^{1/2}}{2\pi} \int_{-\infty}^{+\infty} \exp it^{3/2}(-x + \frac{1}{3}x^3) dx$ qui se met sous la forme précédente, avec $f = 1$ et $S(x) = -x + x^3/3$. Les points stationnaires vérifient par définition $S'(x) = x^2 - 1 = 0$. Donc il y a deux contributions de phase stationnaire en $x = \pm 1$. Comme $S(\pm 1) = \mp \frac{2}{3}$, $S''(\pm 1) = \pm 2$, on obtient

$$\sqrt{\frac{2\pi}{t^{3/2} |S''(\pm 1)|}} \exp\left(iS(\pm 1) + i\frac{\pi}{4} \operatorname{sgn}(S''(\pm 1))\right) = t^{-3/4} \sqrt{\pi} \exp\left(\mp \frac{2}{3} it^{3/2} \pm i\frac{\pi}{4}\right)$$

Donc (résultat obtenu par descente rapide dans le chapitre précédent, formule (11.11)), pour $t \rightarrow \infty$

$$Ai(-t) \approx \frac{1}{\sqrt{\pi}} t^{-1/4} \cos\left(\frac{2}{3} t^{3/2} - \frac{\pi}{4}\right)$$

La méthode est notamment utilisée dans le chapitre 5, et plus spécifiquement dans les sections 5.2 et 5.3 sur les schémas d'ordre 2 non dissipatifs. Une version de la méthode est parfois utilisée quand le grand paramètre n'apparaît pas sous forme multiplicative :

$$I \approx \sqrt{\frac{2\pi}{|S''(x_s)|}} f(x_s) \exp(iS(x_s) + i\frac{\pi}{4} \text{sgn}(S''(x_s^i)))$$

C'est alors directement $|S''(x_s)|$ qui sert de grand paramètre. Par exemple, partant de la représentation la plus usuelle de la fonction d'Airy, le résultat plus haut est obtenu directement. La méthode se généralise au cas où la première dérivée non nulle de S au point stationnaire est d'ordre m . Les intégrales étalons auxquelles on se ramène sont maintenant $\int_0^\infty \exp(ix^m) dx$. Par exemple, pour m impair

$$I(\lambda) \approx \frac{2}{m} \Gamma\left(\frac{1}{m}\right) \left(\frac{m!}{\lambda |S^{[m]}(x_s)|}\right)^{1/m} f(x_0) \cos(\lambda S(x_s) + \frac{\pi}{2m} \text{sgn}(S^{[m]}(x_0)))$$

Exercice : montrer que $Ai(0) = \frac{1}{3\pi} \Gamma(\frac{1}{3}) \frac{\sqrt[3]{6}}{\sqrt[3]{2}} \cos \frac{\pi}{6} \approx 0,35503$

La méthode se généralise au cas multidimensionnel, *i.e.* pour calculer $I = \int_{R^d} f(x) \exp(i\lambda S(x)) d^d x$. En utilisant l'approximation $S(x) \approx S(x_s) + (x - x_s)H(x_s)(x - x_s)/2$, où H est la matrice hessienne, supposée de rang d , de S , on obtient

$$I \approx \frac{(2\pi)^{d/2}}{\lambda^{d/2} |\det(H(x_s))|^{1/2}} f(x_s) \exp(ikS(x_s) + i\frac{\pi}{4} \text{sgn}(H(x_s))) \quad (12.8)$$

Où $\text{sgn}(H(x_s)) = n_+ - n_-$, n_+ (*resp* n_-) : nombre de valeurs propres > 0 (*resp.* < 0) de $H(x_s)$. Le cas où une ou plusieurs valeurs propres sont nulles est assez peu documenté.

Nous nous sommes limités au premier terme du d.a., ce qui est suffisant dans la majorité des applications. Nous utilisons (12.8) dans le chapitre 7. Les formules générales sont présentées dans les références.

12.3 Méthode du col

Rappelons la forme de l'intégrale à approximer :

$$I = \int_C f(z) \exp(\lambda S(z)) dz \quad (12.9)$$

Le contour C est maintenant dans le plan complexe, f et S sont des fonctions analytiques. La méthode de Laplace est un cas particulier de méthode du col, mais aussi un ingrédient essentiel de cette méthode. En effet, le principe de la méthode est de déformer C (au sens de le remplacer par) sur le contour dit de descente rapide (CDR), où la partie imaginaire v de $S = u + iv$ est constante, si bien que $I = \exp(i\lambda v) \int_C f(z) \exp(\lambda u(z)) dz$ et que l'intégrale peut alors être approximée par la méthode de Laplace.

Une interprétation géométrique de la méthode est utile à sa compréhension. Par les relations de Cauchy, les lignes de niveau de v sont les courbes intégrales du gradient de u . Nous suivons dans les grandes lignes la présentation de [7]. Posant $z = x + iy$, et traitant u comme le troisième axe orthogonal à x et y , l'équation $u = u(x, y)$ définit une surface U dans l'espace (x, y, u) . Si $z_c = x_c + iy_c$ est un point critique, *i.e.* $S'(z_c) = 0$, $u_x(z_c) = u_y(z_c) = 0$. Mais u est, en tant que partie réelle d'une fonction analytique, harmonique, donc n'a pas de minimum ni de maximum absolu. Le point (x_c, y_c) est donc un point col de u , au sens où la surface U a la forme d'un col (ou d'une selle, ou d'un parabolôïde hyperbolique) au voisinage de ce point. Dans le cas $S''(z_c) \neq 0$, $S(z) \approx S(z_c) + \frac{1}{2}S''(z_c)(z - z_c)^2$,

$$S''(z_c)(z - z_c)^2 = |S''(z_c)| |z - z_c|^2 \exp(2i \arg(z - z_c) + i \arg(S''(z_c)))$$

$$u(z) - u(z_c) \approx \frac{1}{2} |S''(z_c)| |z - z_c|^2 \cos(2 \arg(z - z_c) + \arg(S''(z_c)))$$

En somme, la courbe section de U par le plan passant par z_c et contenant l'axe vertical u et la droite horizontale $\arg(z - z_c) = \theta$ est approximativement une parabole

– montante si $\cos(2\theta + \arg(S''(z_c))) > 0$

– descendante si $\cos(2\theta + \arg(S''(z_c))) < 0$

En particulier, si $2 \arg(z - z_c) + \arg(S''(z_c)) = \pi$, notant s l'abscisse curviligne sur cette courbe D , en prenant l'origine en z_c , $u(s) - u(0) \approx -\frac{1}{2} |S''(z_c)| s^2 u(s)$. La courbe D est celle où u décroît le plus rapidement de part et d'autre de z_c , ce qui définit le chemin de descente rapide D sur U . Le point col est à l'altitude u maximum sur D .

Si $2\theta + \arg(S''(z_c)) = 0$, $u(z) - u(z_c) \approx \frac{1}{2} |S''(z_c)| |z - z_c|^2$. Comme $u(s) - u(0) \approx \frac{1}{2} |S''(z_c)| s^2$, $u(s)$ croît donc le plus rapidement possible de part et d'autre de z_c . M est donc le chemin de montée rapide, ou la ligne de faite sur U . Le point col est à l'altitude u minimum sur M .

Globalement, le plan complexe se divise en deux secteurs de collines, $-\frac{\pi}{4} - \arg(S''(z_c)) < \arg(z - z_c) < \frac{\pi}{4} - \arg(S''(z_c))$ et son symétrique par rapport à z_c , où $u(z) > u(z_c)$ et deux secteurs de vallées $\frac{\pi}{4} - \frac{1}{2} \arg(S''(z_c)) < \arg(z - z_c) < \frac{3\pi}{4} - \frac{1}{2} \arg(S''(z_c))$ et son symétrique par rapport à z_c . Le chemin de descente rapide est tangent à la bissectrice des secteurs de vallées $\arg(z - z_c) = \pm \frac{\pi}{2} - \frac{1}{2} \arg(S''(z_c))$.

La partie imaginaire de S vaut approximativement

$$v(z) - v(z_c) \approx \frac{1}{2} |S''(z_c)| |z - z_c|^2 \sin(2 \arg(z - z_c) + \arg(S''(z_c)))$$

Sur le chemin de descente rapide $\sin(2 \arg(z - z_c) + \arg(S''(z_c))) = \sin(\pm\pi) = 0$ ce qui confirme sa définition comme courbe, où la partie imaginaire de S est constante.

Prenons l'exemple où S est réelle et a un seul maximum, $\arg(S''(z_c)) = \pi$. La surface U est approximativement au voisinage de $z_c = x_c$, le paraboloidé hyperbolique

$$u(x, y) = u(x_c) - \frac{1}{2} |S''(x_c)| ((x - x_c)^2 - (y - y_c)^2)$$

Le CDR a une tangente d'angle défini par $2 \arg(z - z_c) + \pi = \pi$, $\arg(z - z_c) = 0 \pmod{\pi}$, donc c'est l'axe réel. Le contour de montée rapide, projection de M , est l'axe imaginaire.

La méthode du col consiste, comme dit plus haut, à intégrer le long du CDR. Comme $v = cte$ et $u \approx -\frac{1}{2} |S''(z_c)| s^2$, il est possible d'appliquer la méthode de Laplace. Le seul changement est une rotation globale d'angle $\frac{\pi}{2} - \frac{1}{2} \arg(S''(z_c))$. Formellement, la formule est identique à (12.3) obtenue précédemment avec la méthode de Laplace

$$I(\lambda) \approx \sqrt{\frac{2\pi}{-\lambda S''(x_c)}} f(x_0) \exp \lambda S(x_c) \quad (12.10)$$

Mais, maintenant, $-S''(x_c)$ est un nombre complexe. Appliquons à titre d'illustration la méthode dans le cas où $S = iT$, qui peut être traité par la phase stationnaire. Si $T'' > 0$,

$$\sqrt{\frac{1}{-S''(x_c)}} = \sqrt{\frac{1}{-iT''(x_c)}} = e^{i\frac{\pi}{4}} \frac{1}{\sqrt{T''(x_c)}}$$

et on retrouve bien sûr le résultat de la phase stationnaire, avec un CDR d'angle $\frac{\pi}{4}$ avec l'axe des x . Il est plus direct dans ce cas d'utiliser la phase stationnaire. Mais, certaines intégrales ne peuvent être traitées que par la méthode du col. Un exemple simple est la

fonction d'Airy pour un argument positif $Ai(t) = \frac{t^{1/2}}{2\pi} \int_{-\infty}^{+\infty} \exp it^{3/2}(x + \frac{1}{3}x^3) dx$

qui se met sous la forme (2), avec $f = 1$ et $S(x) = i(x + x^3/3)$. Il est traité en détail dans la sous-section 11.1.1 sur les fonctions d'Airy. Les points cols vérifient par définition $S'(x) = x^2 + 1 = 0$. Donc il y a deux cols en $x = \pm i$, mais le contour initial ne peut être déformé que sur le CDR passant par i , où $S''(i) = 2ix = -2$, $S(i) = -1 + \frac{1}{3} = -\frac{2}{3}$. La formule (12.10) donne le premier terme du développement asymptotique de la fonction (voir chapitre précédent)

$$Ai(t) \approx \frac{t^{1/2}}{2\pi} \sqrt{\frac{2\pi}{2t^{3/2}}} \exp\left(-\frac{2}{3}t^{3/2}\right) = \frac{1}{2\sqrt{\pi}} t^{-1/4} e^{-\frac{2}{3}t^{3/2}}$$

Cette intégrale n'est traitable que par la méthode du col, car il n'y a pas de point critique sur l'axe réel. Le lemme de la phase non stationnaire ne donne qu'une information qualitative de décroissance plus rapide que toute puissance de t , ce que confirme la formule ci-dessus. La méthode, et notamment (12.10), est utilisée pour approximer la solution du schéma upwind (section 5.1) et celle des schémas discrétisant l'équation de la chaleur (chapitre 8).

Comme pour la méthode de Laplace, la méthode du col peut fournir les termes suivants du développement asymptotique, et être généralisée au cas où la première dérivée non nulle de S est $S^{[2n]}$. Nous renvoyons aux références pour les formules explicites. Le point délicat de la méthode est la détermination du CDR approprié : comme le souligne M. Fedoriouk, il s'agit d'un problème global, et il n'existe pas de méthode générale pour le résoudre.

Une autre difficulté apparaît dans le cas multidimensionnel : la généralisation de la méthode reste un sujet ouvert. Malgré tout, la qualifier de méthode obsolète [5] est quelque peu excessif...

Nous espérons que ce bref chapitre aidera le lecteur à se familiariser avec les techniques de développements asymptotiques d'intégrales. Nous n'avons pas abordé le sujet important, mais difficile, des développements asymptotiques uniformes : les formules présentées ne sont valides que sous condition : par exemple (12.7) diverge si f a un pôle au point stationnaire. Les références contiennent des présentations détaillées et rigoureuses, et de nombreux exemples. Les points de vue adoptés sont différents, si bien qu'il est difficile d'en recommander une seule. L'article [5] est particulièrement clair. Les livres [3] de M. Fedoriouk, disponibles qu'en russe, contiennent énormément de résultats, repris en partie dans le livre de R. Wong, complet et pédagogique, tout comme celui de N. Temme. L'ouvrage de Paris [4] est plus spécialisé sur l'hyperasymptotique, mais son premier chapitre présente clairement la méthode du col, celui de Butler [1] contient un grand nombre d'applications en statistique. Enfin, le rapport [2] est très pédagogique.

Bibliographie

- [1] R.W. Butler "Saddlepoint Approximations with Applications" Cambridge, 2007
- [2] E. Falck "Asymptotic Expansions of Integrals and the Method of Steepest Descent" rapport, université d'Uppsala, 2016
- [3] M. Fedoriouk « Méthode du point selle » et « Asymptotique : intégrales et séries », Nauka 1986 et 1992
- [4] R.B. Paris "Hadamard Expansions and Hyperasymptotic Evaluation, an extension of the method of steepest descent", Cambridge, 2016

- [5] T. Tao “Oscillatory integrals, single-variable phases” Almost Originality, a mathematical journal, 2010
- [6] N. Temme “Asymptotic Methods for integrals”, World scientific, 2014
- [7] R. Wong “Asymptotic approximations of integrals”, Academic Press, 2001

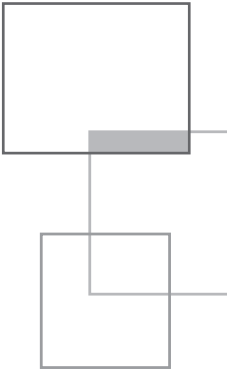
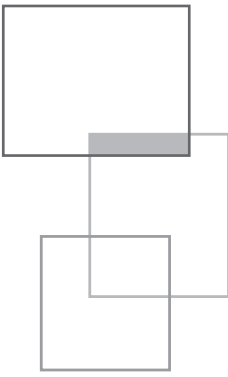


Table des figures

4.1	Relation de dispersion (4.21) du schéma centré semi-discrétisé comparée à $\omega_e(k)$	34
4.2	Relation de dispersion $\omega^2(k, l) = \frac{4a^2}{b^2} (\sin^2(\frac{kb}{2}) + \sin^2(\frac{lb}{2}))$ du schéma centré semi-discrétisé.	36
4.3	Relation de dispersion exacte $\omega_e^2(k, l) = a^2(k^2 + l^2)$ de l'équation.	37
4.4	Facteur d'amplification dans le plan complexe du schéma upwind $\widehat{Q}(k) = 1 - \eta + \eta e^{-ik} = 1 - \eta(1 - \cos k) - i\eta \sin k$ (pour $\eta = 0, 0, 5, 0, 7, 1, 0$).	38
4.5	Argument du facteur d'amplification du schéma upwind $Arg(\widehat{Q}(k)) = \arctan(\frac{-\eta \sin k}{1 - \eta + \eta \cos k})$ (pour $\eta = 0, 5$).	39
4.6	Module du facteur d'amplification du schéma upwind $g(k) = \sqrt{(1 - 2\eta \sin^2(k/2))^2 + (\eta \sin k)^2}$ (pour $\eta = 0, 5$).	40
5.1	Fonction de Green du schéma upwind, solution numérique et solution exacte.	65
5.2	Fonction de Green du schéma upwind et approximation de l'équation équivalente de gauche à droite, $\eta = 0, 75, \eta = 0, 50, \eta = 0, 25$	73
5.3	Solution exacte à j fixé en fonction de t	78
5.4	Solution exacte et simulation numérique à t fixé.	78
5.5	Solution exacte et approximation par la phase stationnaire, développement de Debye (5.20).	80
5.6	Solution exacte et approximation équation équivalente (Watson).	81

5.7	Fonction de Green du schéma saute-mouton et de l'équation équivalente.	83
5.8	Vitesse de groupe v_+	84
5.9	Simulation du schéma saute-mouton et mode physique.	87
5.10	Simulation du schéma saute-mouton et mode parasite.	88
5.11	Simulation du schéma saute-mouton et mode physique.	91
5.12	Simulation du schéma saute-mouton et mode physique.	92
5.13	Solution du schéma <i>vs.</i> solution exacte.	97
5.14	Facteur d'amplification du schéma de LW dans le plan complexe pour $\eta = 0.5, 0.7, 1.0$	99
5.15	Module au carré du facteur déamplification du schéma de LW.	99
5.16	Solution du schéma <i>vs.</i> équation équivalente à 1 terme.	102
5.17	Fonction de Green, schéma <i>vs.</i> EE à 2 termes.	104
5.18	Solution du schéma BW <i>vs.</i> équation équivalente à 1 terme.	108
5.19	Fonction de Green, schéma <i>vs.</i> EE à 2 termes.	108
5.20	Solution du schéma BW <i>vs.</i> équation équivalente à 1 terme sur une marche descendante.	110
5.21	Solution du schéma <i>vs.</i> équation équivalente à 2 termes sur une marche descendante.	110
5.22	Solution du schéma Beam-Warming avec min mod <i>vs.</i> équation équivalente à 1 terme sur une marche descendante.	114
5.23	Fonction de Green du schéma S_3^1 comparée à celle de l'équation équivalente G_{EE}	118
5.24	Advection d'une marche par le schéma S_3^1 et approximation équation équivalente.	118
5.25	Fonction de Green 5.72 (calculée en Fourier) du schéma et simulation.	121
5.26	Fonction de Green du schéma et de l'équation équivalente à 1 terme G_{EE}	121
5.27	Tracé de v_g pour les schémas centrés d'ordre 4 et 6.	123
5.28	Fonction de Green du schéma de Balakine et de l'équation équivalente à deux termes.	127
5.29	Fonction de Green approchée des schémas de Strang d'ordre 3, 5, 7, 9 et 11 par l'équation équivalente à un terme.	128
5.30	Advection suivant la diagonale, fonction de Green du schéma direct et du schéma splitté.	134

6.1	Fonction de Green du schéma de Yee comparé à son approximation équivalente.	143
7.1	Solution du schéma et approximation phase stationnaire (7.16).	153
7.2	Solution du schéma (obtenu avec un petit η) et approximation (7.17) de la fonction de Green de l'équation avec conditions initiales sur la dérivée.	154
7.3	Solution du schéma et approximation phase stationnaire (7.19).	157
7.4	Solution du schéma et approximation (phase stationnaire + singularité), pour $\rho > 0$	159
7.5	Solution numérique globale, amplitude et lignes de niveau.	162
8.1	Fonctions de Green G_s du schéma explicite, son approximation par (8.19), et leur différence.	179
8.2	Fonction de Green du schéma pour $\eta = \frac{1}{2}$ comparée à (8.21).	182
8.3	G_s , son approximation (8.29) et leur différence.	186
8.4	Fonction de Green du schéma implicite pour de grandes valeurs de η	187
8.5	G_s , son approximation (8.33) et leur différence.	189
8.6	Fonction de Green G_s du schéma de Crank-Nicolson pour de grandes valeurs de η	190



Index

A

Airy (fonction de), 56, 57, 77, 82, 101, 103-105, 107, 111-113, 117, 124, 127, 128, 143, 207, 208, 210, 214-222, 227, 229, 230, 232
 anisotropie, 39, 130, 132, 133, 135, 145, 152, 157-159, 161, 162, 190, 192
 artéfact, 3-5, 9, 30, 45, 94, 105, 109, 114, 115, 117, 119, 123, 128, 129, 169

B

Bessel (fonction de), 76, 77, 79-81, 177, 221

C

consistance, 11-15, 33, 41-44, 51, 52, 84, 140, 167
 convergence, 4, 13, 19, 29, 30, 39, 41, 44, 49, 53-55, 58, 68, 70, 111, 123, 135, 167, 169, 176, 180, 183, 191, 192, 201, 227

D

descente rapide, 70, 71, 117, 170, 208, 210, 211, 215, 218, 224, 229, 231, 232
 diffusif, 14, 45, 46, 58, 129
 diffusion, 7, 8, 15-17, 44, 46, 51, 57, 75, 76, 95, 129, 132-135, 165
 Dirac, 4, 7, 30, 50, 75, 95, 101, 117, 145-147
 discrétisation, 3, 12, 15, 29, 46, 140, 162, 197, 204, 205
 dispersif, 17, 32, 46, 58, 129, 161, 204
 dispersion, 15, 17, 28, 32-36, 38, 39, 44, 45, 54, 74, 82, 84, 100, 106, 107, 112, 114, 115, 117, 119, 122, 124, 202
 dissipatif, 32, 34, 36, 38, 39, 50, 53-55, 57, 74, 76, 81, 95, 98, 103-105, 107, 109, 114-116, 119, 120, 123-125, 127-129, 135, 144, 149, 167, 201, 204, 215, 219, 230
 dissipation, 15, 28, 32, 38, 39, 44, 45, 49, 52, 54, 55, 57, 74, 76, 83, 95, 98, 100-102, 104-107, 111, 115, 116, 119, 120, 123, 124, 126, 127, 129, 175, 202

G

gaussienne, 8, 69, 72, 75, 123, 132, 166, 180, 191, 219, 221, 225, 228

Green (fonction de), 4, 6-8, 13, 15, 17, 20, 22, 35, 50, 53, 54, 56, 57, 64, 66-69, 72, 74-77, 81-83, 85, 86, 95, 96, 101-105, 107, 109, 115-117, 119-121, 124-126, 128-130, 132-134, 143-145, 153, 154, 158, 161, 162, 165-169, 171, 174, 176-179, 181-184, 186-189, 191, 207, 208, 215, 217

H

Heaviside (fonction de), 4, 30
hyperbolique, 3, 5, 6, 50, 76, 137, 162, 166, 207, 231, 232

L

Laplace, 27, 146, 172-174, 177, 178, 184, 187, 188, 223, 224, 228, 231-233

M

méthode du col, 27, 54, 72, 75, 103, 117, 124, 169, 171-173, 176-178, 190, 210, 214, 224, 230-233

P

phase stationnaire, 34, 35, 54, 80, 81, 87-89, 91, 115, 129, 149, 153, 157, 158, 218, 221, 223, 224, 227-229, 232

S

shift, 10, 21, 22, 48
stabilité, 12-14, 16, 24, 39, 41, 42, 46, 49, 58, 101, 106, 107, 116, 169, 175, 176, 183, 189, 190, 198, 201

T

troncature, 12-14, 41, 119, 140

U

upwind, 12, 15-17, 21, 22, 29, 30, 36, 38, 39, 45, 46, 48, 51, 53, 57, 63, 66, 75, 76, 95, 96, 106, 113, 114, 129, 130, 134, 135, 138, 166, 180, 198, 233