
PROJET DE FIN D'ETUDES

Méthodes d'analyse asymptotique et
d'approximation numérique de modèles dissipatifs
multi-échelles pour EDO à variété centrale

MENTION DE CONFIDENTIALITÉ

Rapport non confidentiel et publiable sur Internet

AUTEUR : LÉOPOLD TRÉMANT
FILIÈRE : MODÉLISATION ET SIMULATION
PROMOTION : 2018

ENSEIGNANT RÉFÉRENT :
Sonia FLISS

TUTEURS INRIA :
Philippe CHARTIER
Mohammed LEMOU

Stage effectué du 12 mars 2018 au 10 août 2018

INRIA RENNES - BRETAGNE ATLANTIQUE
CAMPUS UNIVERSITAIRE DE BEAULIEU – 35042 RENNES CEDEX

RÉSUMÉ

On commence par présenter une classe de modèle à variété centrale à détente rapide et les problèmes qu'elle engendre d'un point de vue numérique avec les approches usuelles. En s'inspirant de travaux récents sur les systèmes hautement oscillants, on effectue un développement double-échelle pour séparer la dynamique rapide de la lente. Après cette séparation, on étudie le nouveau problème engendré et on exhibe des conditions pour que le problème soit bien posé, régulier avec des dérivées non raides. À partir de ces conditions, on construit un schéma numérique uniformément convergent qu'on implémente et dont on évalue les performances et limites en le comparant à d'autres approches.

Mots-clés : *méthodes numériques, simulation, systèmes raides, développement double-échelle, uniformément précis, problème de transport, variété centrale*

ABSTRACT

We first present a class of equations with a center manifold which relax very quickly, and the numerical shortcomings of usual approaches when dealing with this system. Inspired by recent advances in numerical schemes for highly oscillating problems, a new problem is derived from a two-scale decomposition separating the quick and slow dynamics. Non trivial conditions are then found to assure that this new problem is well-posed, with regular and non stiff solutions. We then devise a uniformly accurate numerical scheme, analyse and implement it. Finally the implementation is studied in terms of convergence and performance and is compared to other schemes and approaches.

Keywords: *numerical methods, simulation, stiff systems, two-scale expansion, uniformly accurate, transport equation, centre manifold*

REMERCIEMENTS

Apparemment, beaucoup de gens n'écrivent pas de remerciements pour leur projet de fin d'études. Peut-être voient-ils ce rapport comme une simple épreuve, une obligation parmi tant d'autres à passer avant l'entrée dans la vie active. Peut-être que d'autres le voient comme une étape presque négligeable devant un projet plus ambitieux en thèse. Pour ma part, j'ai eu envie de prendre le temps de réfléchir aux trois années qui ont précédé ce stage et ce rapport, non seulement pour exprimer ma gratitude envers tous ceux qui m'ont accompagné, mais aussi pour me forcer à considérer la fin de ces études comme une vraie étape.

Ainsi j'aimerais en premier lieu remercier mes tuteurs Philippe Chartier et Mohamed Lemou qui m'ont donné la chance d'effectuer ce stage, et qui ont su faire preuve de patience lorsque j'ai bloqué et douté. J'avais choisi ce sujet parce qu'il me sortait de ma zone de confort, donc je n'aurais pas dû être surpris lorsque j'ai commencé à tourner en rond sur les mêmes idées, mais face à la difficulté, j'ai eu au début beaucoup de mal à me structurer. Merci donc à eux pour leurs lumières, et leur fermeté parfois, qui m'ont été cruciales pour remettre mon travail et ma pensée à l'endroit. Je suis honoré qu'ils m'aient accepté comme doctorant l'année prochaine.

Je souhaite aussi remercier mes professeurs Patrick Ciarlet et Sonia Fliss, qui m'ont initié aux méthodes numériques en m'appâtant avec de jolies images. En me faisant découvrir ce domaine, ils m'ont rendu ma motivation pour les mathématiques, et m'ont ouvert plus de portes de je ne pourrai jamais en ouvrir. Je remercie les professeurs qui m'ont montré les différentes facettes qui peuvent et doivent être abordées pour faire du calcul numérique, qu'il s'agisse d'approches ingénieuses avec Bertrand Maury, de preuves mathématiques plus complexes avec Patrick Joly ou encore de recherche de performance avec Axel Modave et Marc Massot. Parmi toutes ces personnes formidables, et bien qu'elle est déjà mentionnée, j'aimerais insister sur l'importance de l'apport et du soutien que m'a fourni Sonia Fliss. Au-delà de ses capacités pédagogiques et de sa gentillesse évidente, elle s'est montrée attentive et de très bon conseil lorsque je lui ai fait part de mes doutes quant à mes choix d'avenir.

Merci aussi à ma famille, qui a depuis quelques temps arrêté d'essayer de comprendre les détails de mes études, ce qui m'épargne bien des explications laborieuses. Surtout, merci à eux de m'aimer malgré le peu de temps que j'ai su leur accorder ces dernières années. J'espère pouvoir leur consacrer plus de temps dans les années à venir. Plus spécifiquement, je remercie mes parents qui ont toujours soutenu mes décisions, en me faisant souvent part de leur avis éclairé. Des partages artistiques aux discussions politiques, notre relation m'aide à me construire encore aujourd'hui. Un gros merci à

mon frère Jules de me faire assez confiance pour me confier ses doutes, de partager sa culture et sa passion vidéoludique, et aussi, lorsque les astres s'alignent, de discuter avec moi de ses réflexions humaines.

Je remercie Maël et Ece pour m'avoir écouté et accueilli lors de moments difficiles. Il y a peu de gens avec qui j'arrive à garder un lien aussi proche que celui-ci, et je compte maintenir cela aussi longtemps que possible. Enfin, je remercie tous mes amis de l'ENSTAlibur. Julien et Benoît toujours prêts à m'offrir un pied-à-terre en région parisienne, Élise et Benoît au bonheur contagieux, Théo plein de surprises, Léa qui m'a supporté tout le M2 durant et qui me donne une excuse pour retourner en Angleterre, et tous les autres : Florian (les deux), Diane, Emilien (un seul des deux), Lucas, Pierre, Rémi, Quentin, Aurore, Youssef, Adrien, Manu, BPJ, ainsi que tous les autres petits jeunes dont je n'ai pas encore fait la connaissance. Cette grande famille me rend toujours très heureux et j'espère qu'on restera soudés pour les nombreuses années à venir.

TABLE DES MATIÈRES

Introduction	9
1 Présentation du problème et motivation	11
1.1 Modèle, propriétés et soucis numériques	11
1.1.1 Problème à variété centrale	11
1.1.2 Méthodes numériques usuelles	12
1.1.3 Méthodes asymptotiques	15
1.2 Introduction au développement double-échelle	16
1.2.1 Nouvelles problématiques	16
1.2.2 Caractère bien posé dans C^2	17
1.2.3 Condition initiale ou condition au bord?	19
2 Mise en place d'un schéma numérique uniformément convergent	21
2.1 Choix de la condition initiale	21
2.1.1 Approche initiale	21
2.1.2 Développement de Chapman-Enskog	22
2.2 Bornes uniformes de la solution	24
2.2.1 Théorème sur $U = (X, Z)$	24
2.2.2 Forme explicite des conditions initiales et au bord sur X, Z . . .	26
2.3 Présentation du schéma d'ordre 1	28
2.3.1 Propriétés du schéma	28
2.3.2 Erreur locale du schéma	29
2.3.3 Erreur globale	30
3 Implémentations et résultats numériques	33
3.1 Implémentations « simples » et convergences	34
3.1.1 Schéma classiques sur des séries	34
3.1.2 Implémentation par différences finies	35
3.2 Implémentation avec la librairie chebfun	36
3.2.1 Erreur globale sur l'équation à variété centrale	37
3.2.2 Erreur locale sur l'équation de transport	39
3.2.3 Coût de l'implémentation	40
3.3 Pistes de recherche	40
3.3.1 Discrétisation du domaine de Laplace	41
3.3.2 Schéma local à coût plus faible	42
Conclusion	43

A	Présentation des cas tests	45
A.1	Équations scalaires couplées	45
A.2	Oscillations lentes de x	45
B	Preuve du lemme 2 (p. 17)	47
C	Erreurs des schémas usuels	49
C.1	Schémas à séparation des vitesses	49
C.2	Schéma à pas de temps adaptatif	50
C.3	Développement asymptotique	52
	Bibliographie	55
	Table des figures	57

INTRODUCTION

De nombreuses études ont été effectuées sur des équations différentielles ordinaires (EDO) impliquant deux dynamiques : une lente et une périodique très rapide. Ce type d'équation peut être dérivé de nombreux problèmes physiques hautement oscillants, comme l'équation de Klein-Gordon ou l'équation de Schrödinger [5] et donc la résolution numérique de ce type d'équations est un enjeu clé en physique.

Avec des méthodes de résolution classiques, la dynamique rapide impose d'utiliser un pas de temps Δt petit devant la période ε , ce qui est généralement impossible techniquement (temps de calcul trop long, plus d'erreurs d'arrondis, etc). Depuis le début des années 80, des méthodes d'homogénéisation [1, sec. I.2] sont développées. Elles consistent à trouver un nouveau système qui approche le problème d'origine dans le cas où ε est très petit (devant une grandeur caractéristique qui dépend du problème). L'homogénéisation est encore utilisée et développée aujourd'hui, et elle est très utile autant au monde académique qu'à l'industrie, mais elle n'est pas très pratique ni efficace si ε est trop grand.

Plus récemment (2015), des méthodes *uniformément précises* (UP) [5,6] ont été développées pour certains de ces problèmes à deux dynamiques. Celles-ci permettent d'obtenir une erreur indépendante de ε , c'est-à-dire qu'à pas de temps fixé, faire varier ε ne fait pas varier l'erreur. Elles sont aujourd'hui moins développées et répandues que les méthodes classiques ou asymptotiques, mais leur étude et leur développement sont particulièrement intéressants, puisqu'elles ne dépendent pas de la taille de ε qui peut varier au sein d'un même problème.

L'objectif de ce stage (et de la thèse par la suite) était de développer une méthode uniformément précise pour une classe de modèles à *variété centrale*. Il s'agit d'EDO qui présentent une détente rapide (de durée d'ordre ε) vers une variété à dynamique lente.

De manière similaire au cas périodique, les méthodes usuelles nécessitent d'utiliser un pas de temps court devant ε pour ne pas perdre d'information pendant cette détente. Récemment, des méthodes asymptotiques ont été développées pour cette classe de modèles [4] et permettent d'obtenir un modèle non-raide qui approche l'état du système après détente.

Notre classe de problèmes trouve certaines applications en écologie et en chimie, mais le but est surtout de pouvoir ensuite considérer des équations aux dérivées partielles du type équation de Vlasov avec collisions raides. Même si le lien n'est pas direct, on se familiarise ainsi avec les phénomènes à détente rapide, et on pourra sûrement adapter certaines méthodes à ces nouvelles équations.

Pour trouver un schéma uniformément précis, nous nous sommes basé sur ce qui avait été fait dans le cas périodique dans [5] avec un développement double-échelle.

La difficulté est alors de s'approprier la méthode et de trouver comment résoudre le problème obtenu après développement.

On commence par présenter notre problème et son comportement, avec en plus une étude préliminaire du nouveau problème suite au développement double-échelle. Ensuite, on démontre que ce nouveau problème peut être bien posé avec certaines propriétés qui nous permettent de construire un schéma numérique uniformément convergent. Enfin, on se penche sur l'implémentation de ce schéma —son comportement, les difficultés qui y sont liées, les résultats de convergence, le coût...

PARTIE 1

PRÉSENTATION DU PROBLÈME ET MOTIVATION

On présente d'abord le modèle, son comportement et les problématiques numériques associées, en détaillant l'ordre et le coût de certains schémas. On introduit ensuite le développement double-échelle qui va être la base de la prochaine partie, en précisant certains des problèmes qu'il pose et en évoquant certaines de ses propriétés.

1.1 Modèle, propriétés et soucis numériques

On considère le modèle suivant

$$\begin{cases} \dot{x}^\varepsilon = f(x^\varepsilon, z^\varepsilon), & x^\varepsilon(0) = x_0 \in \mathbb{R}^n \\ \dot{z}^\varepsilon = -\frac{1}{\varepsilon}z^\varepsilon + g(x^\varepsilon, z^\varepsilon), & z^\varepsilon(0) = z_0 \in \mathbb{R}^m \end{cases} \quad (1.1)$$

avec $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n, g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ analytiques et ε un petit paramètre.

Ce système est une simplification d'un autre, équivalent, qui fait apparaître $\frac{1}{\varepsilon}\Lambda z^\varepsilon$ (Λ diagonale positive) au lieu de $\frac{1}{\varepsilon}z^\varepsilon$. Le système (1.1) peut ainsi être dérivé d'un laplacien, e.g. d'un problème de réaction-diffusion.

On présente d'abord certaines propriétés de ce système, puis quelques méthodes numériques usuelles, en précisant leurs avantages et leurs défauts, et enfin on décrit le principe d'une méthode asymptotique moderne de [4] pour saisir encore un peu mieux le comportement du modèle.

1.1.1 Problème à variété centrale

Le comportement de la solution est résumé dans le *théorème de variété centrale* et le *principe de réduction* [3] suivant

Théorème 1. Soit B_R la boule de rayon R et de centre 0 dans $\mathbb{R}^n \times \mathbb{R}^m$. On suppose $f(0, 0) = 0$ et $g(0, 0) = 0$. Alors pour tout $R > 0$, il existe $\varepsilon^* > 0$ et $T > 0$ tels que la solution $(x^\varepsilon(t), z^\varepsilon(t))$ de (1.1) existe pour tout $\varepsilon \in]0, \varepsilon^*]$, tout $t \in [0, T]$ et toute condition initiale $(x_0, z_0) \in B_R$.

En outre, pour tout $\varepsilon \in]0, \varepsilon^*]$, il existe une fonction $h^\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}^m$ analytique de sorte que la variété

$$\mathcal{M} = \{(x, \varepsilon h^\varepsilon(x)), x \in \mathbb{R}^n\}$$

est invariante pour (1.1) dans le sens $(x_0, z_0) \in \mathcal{M} \cap B_R \Rightarrow \forall t \in [0, T], (x(t), z(t)) \in \mathcal{M}$. En particulier, $0_{\mathbb{R}^n \times \mathbb{R}^m} \in \mathcal{M}$ donc on l'appelle « variété centrale ». Par souci de simplicité, on confond l'ensemble \mathcal{M} et la fonction $\varepsilon h^\varepsilon$ par cette appellation.

Si on note φ_t le flot en t de l'EDO réduite

$$\dot{v} = f(v, \varepsilon h^\varepsilon(v)) \quad (1.2)$$

alors il existe $\mu > 0$ tel que pour tous $(x_0, z_0) \in B_R, \varepsilon \in]0, \varepsilon^*]$, il existe une condition initiale dite réduite $x_0^\varepsilon \in \mathbb{R}^n$ telle que

$$\forall t \in [0, T], \quad x(t) = \varphi_t(x_0^\varepsilon) + \mathcal{O}(e^{-\mu t/\varepsilon}) \quad \text{et} \quad z(t) = \varepsilon h(\varphi_t(x_0^\varepsilon)) + \mathcal{O}(e^{-\mu t/\varepsilon}).$$

En d'autres termes, le système converge très rapidement (à vitesse $e^{-t/\varepsilon}$) vers un état « réduit » où z^ε est entièrement déterminé par x^ε via la relation $z^\varepsilon = \varepsilon h^\varepsilon \circ x^\varepsilon$. On ne démontrera pas ce théorème bien connu mais on l'illustre par des exemples en annexe A.

Présentons un exemple précis et vérifions la convergence vers la variété. On choisit $n, m = 1$,

$$\begin{cases} f(x, z) = -x^3(z - z^3/3), \\ g(x, z) = x(1 - z^2/2) \end{cases} \quad \text{et} \quad \begin{cases} x_0 = 0,8 \\ z_0 = 0,05 \end{cases} \quad (1.3)$$

cet exemple étant traité chez [4]. On trace la solution sur $[0, 1]$ pour quelques valeurs de ε ainsi que la variété associée afin de vérifier ce comportement. Comme on ne connaît pas de solution analytique, on utilise une implémentation MATLAB du schéma Radau IIa [2] avec une tolérance absolue de 10^{-16} et une tolérance relative de 10^{-13} pour calculer une solution quasi-exacte.

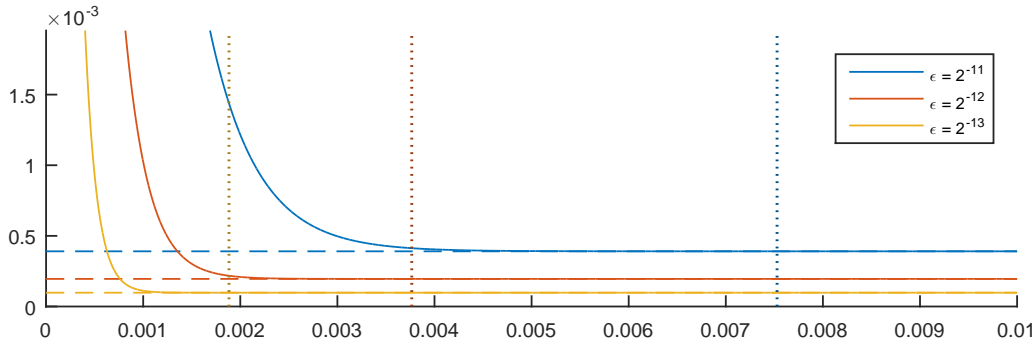


FIGURE 1.1 – Solution z en fonction du temps et la solution sur la variété associée $z_\infty : t \mapsto \varepsilon h^\varepsilon \circ \varphi_t(x_0^\varepsilon)$, pour $\varepsilon = 2^{-11}, 2^{-12}, 2^{-13}$. La fin du régime transitoire est indiquée en pointillés comme l'instant t_0 à partir duquel $\forall t \geq t_0, |z(t) - z_\infty(t)| \leq 10^{-8}$.

On observe bien dans cet exemple une convergence exponentielle de z vers le régime permanent qu'est la variété centrale $z_\infty = \varepsilon h^\varepsilon(x)$, avec un régime transitoire de durée d'ordre ε . En effet on voit que diviser ε par 2 divise également la durée du régime transitoire par 2. Il y a bien deux dynamiques en jeu : celle du régime transitoire, en $e^{-t/\varepsilon}$ et celle du régime permanent, qui dépend de ε de façon *non raide*.

1.1.2 Méthodes numériques usuelles

On considère ici principalement les méthodes numériques à pas de temps uniforme. On suppose dans nos exemples que le système (1.1) admet une solution bornée

sur $[0, T]$, et on définit le pas de temps $\Delta t = T/N$ avec N le nombre de sous-divisions de l'intervalle. On note $t_n = n\Delta t$, $n = 0, \dots, N$. Pour un schéma donné, on s'intéressera aux erreurs locales $\ell e_{n+1}^{(x)}, \ell e_{n+1}^{(z)}$ définies par

$$\ell e^{(x)} = |x_1 - x(\Delta t)|, \quad \ell e^{(z)} = |z_1 - z(\Delta t)|$$

où $|\cdot|$ est la norme euclidienne de \mathbb{R}^n ou de \mathbb{R}^m et x_1, z_1 sont obtenus par le schéma à partir de x_0, z_0 .

Les méthodes numériques usuelles à pas fixe permettent d'obtenir des erreurs en $\Delta t^p/\varepsilon^q$, $p, q > 0$. Lors de la première partie de mon stage, j'ai implémenté et expérimenté avec quelques uns de ces schémas. J'ai aussi évalué les performances d'un schéma implicite-explicite *ad hoc*, d'un splitting de Strang et d'un schéma à pas de temps adaptatif.

On présente certains de ces schémas de façon plus poussée en annexe C, avec des études théoriques et numériques d'erreurs. En particulier, des graphes d'erreurs sont disponibles.

Schémas explicites à pas fixe

L'erreur locale $\ell e^{(x)}, \ell e^{(z)}$ d'un schéma d'Euler explicite avec $\varepsilon < 1$ fixé est

$$\ell e^{(x)} = \mathcal{O}(\Delta t^2/\varepsilon), \quad \ell e^{(z)} = \mathcal{O}(\Delta t^2/\varepsilon^2).$$

Il est intéressant de noter que cela ne génère pas nécessairement une erreur finale en ε^{-2} sur z , puisque z cesse rapidement d'être raide.

Implicite-explicite *ad hoc*

Une idée qui peut sembler naturelle est d'impliciter la partie raide pour garder explicite la partie non-linéaire (dans f, g). Nos calculs et simulations (en annexe C.1) donnent les erreurs locales

$$\ell e^{(x)} = \begin{cases} \mathcal{O}(\Delta t) & \text{si } \varepsilon \ll \Delta t \\ \mathcal{O}(\Delta t^2/\varepsilon) & \text{si } \Delta t \ll \varepsilon \end{cases}, \quad \ell e^{(z)} = \begin{cases} \mathcal{O}(\varepsilon/\Delta t) & \text{si } \varepsilon \ll \Delta t \\ \mathcal{O}(\Delta t^2/\varepsilon^2) & \text{si } \Delta t \ll \varepsilon \end{cases}$$

On voit qu'il faut toujours imposer $\Delta t \ll \varepsilon$ pour avoir une erreur raisonnable. L'avantage est que le schéma est bien plus stable que le schéma complètement explicite, et il est aussi plus rapide qu'un schéma d'Euler implicite usuel.

Splitting de Strang

La partie raide du système est essentiellement contenue dans le terme $-\frac{1}{\varepsilon}z$, et donc on propose de séparer le système en deux parties pour le splitting

$$\begin{cases} \dot{x} = 0 \\ \dot{z} = -\frac{1}{\varepsilon}z \end{cases} \quad \begin{cases} \dot{x} = f(x, z) \\ \dot{z} = g(x, z) \end{cases}$$

On voit qu'on peut résoudre la première partie exactement à moindre coût, et la deuxième avec un schéma Radau IIA avec une tolérance faible. On peut alors supposer que l'erreur provient uniquement du splitting, et donc on calcule une erreur locale théorique

d'ordre $\mathcal{O}(\Delta t^3/\varepsilon^2)$ sur (x, z) dans le cas f, g linéaires. Cette vitesse de convergence est bien vérifiée pour x sur nos exemples, mais l'ordre est dégénéré pour z (sûrement à cause des non-linéarités). Ainsi on obtient

$$\ell e^{(x)} = \mathcal{O}(\Delta t^3/\varepsilon^2) \quad \text{et} \quad \ell e^{(z)} = \mathcal{O}(\Delta t^2/\varepsilon)$$

Les erreurs globales $e^{(x)}, e^{(z)}$ respectent

$$e^{(x)} = \mathcal{O}(\Delta t^2/\varepsilon) \quad \text{et} \quad e^{(z)} = \mathcal{O}(\Delta t^2/\varepsilon^2)$$

ce qui est bien une convergence d'ordre 2.

Pas de temps adaptatif

Une dernière méthode usuelle que nous avons considérée est une méthode de Runge-Kutta imbriqués d'ordre 3, à partir de RK4 avec la règle des 3/8. Son tableau de Butcher est donné par

0	0	0	0	0	0
1/3	1/3	0	0	0	0
2/3	-1/3	1	0	0	0
1	1	-1	1	0	0
1	1/8	3/8	3/8	1/8	0
	1/8	3/8	3/8	1/8	0
	1/12	1/2	1/4	0	1/6

qui nous donne deux estimations $\hat{x}_{n+1}, \hat{z}_{n+1}$ et $\tilde{x}_{n+1}, \tilde{z}_{n+1}$ (notons que la colonne des coefficients en temps à gauche n'est pas utile dans notre cas). On calcule la différence entre ces deux estimations pour obtenir un pas de temps adapté

$$\Delta t_{\text{opt}} = 0.9 \Delta t \left(\frac{\text{Tol}}{|\hat{u}_{n+1} - \tilde{u}_{n+1}|} \right)^{1/4}$$

où $u = \begin{pmatrix} x \\ z \end{pmatrix}$.

Ce choix de pas de temps est classique et est présenté à maintes reprises dans [8]. Le cas échéant, il est expliqué en annexe C.2. Après implémentation, on observe un coût en $\mathcal{O}(1/\varepsilon)$ indépendant du facteur de tolérance pour ε petit (fig. C.3 en sec. C.2), ce qui rend finalement ce schéma aussi (mal) adapté que les autres.

Schéma Radau IIA

On utilise ce schéma implicite à pas de temps adaptatif pour calculer nos solutions de référence. Pour l'implémentation, on utilise un code MATLAB [2] qui est une traduction du code FORTRAN de E. Hairer et G. Wanner appliquant la méthode Radau IIA [8, sec. IV.5].

On se sert de ce code pour calculer nos solutions de références et on n'évalue pas sa convergence, mais on peut mesurer son coût en fonction de ε . Il est évident que la valeur de ε a une influence sur le temps de calcul, mais il est impossible de déceler une relation nette.

On pourrait se dire que ce schéma est suffisant, mais nous allons poursuivre notre étude dans une autre direction pour

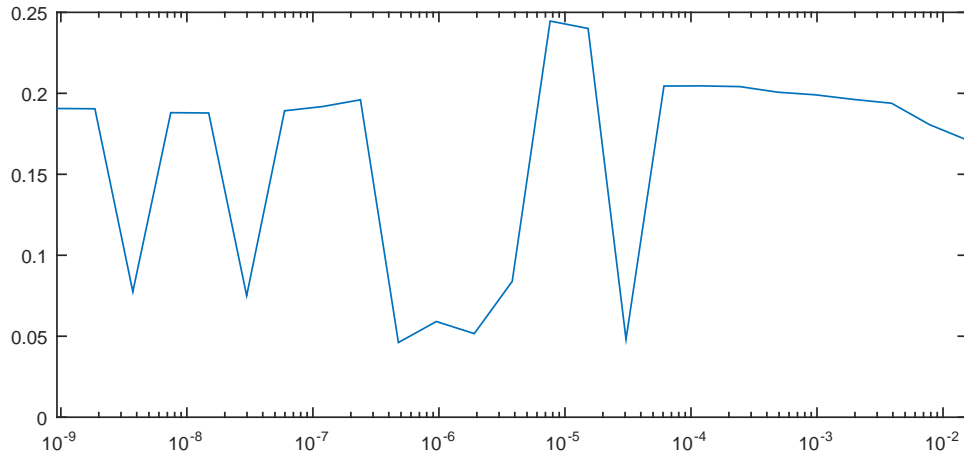


FIGURE 1.2 – Temps du calcul d’une solution u^ε de référence avec le schéma Radau IIA pour différentes valeurs de ε .

1. essayer de trouver des méthodes au coût indépendant de ε et
2. trouver des méthodes *uniformément précises* qui pourront nous aider à attaquer d’autres problèmes raides plus complexes par la suite.

1.1.3 Méthodes asymptotiques

Des méthodes plus modernes et plus formelles cherchent à déterminer l’expression quasi-exacte de la variété $\varepsilon h^\varepsilon$ ainsi que la condition initiale modifiée x_0^ε .

Il est intéressant de noter que h^ε respecte la relation

$$\forall x \in \mathbb{R}^n, \quad \varepsilon(h^\varepsilon)'(x) \cdot f(x, \varepsilon h^\varepsilon(x)) = -h^\varepsilon(x) + g(x, \varepsilon h^\varepsilon(x)) \quad (1.4)$$

et ainsi, à l’aide de développements limités successifs sur la seconde variable de f et de g , on obtient une formulation en puissances de ε pour h^ε . En particulier, on a

$$h^\varepsilon(x) = g(x, 0) + \varepsilon(\partial_z g(x, 0) \cdot g(x, 0) - \partial_x g(x, 0) \cdot f(x, 0)) + \mathcal{O}(\varepsilon^2).$$

Il est bien plus difficile de déterminer x_0^ε , puisque cela demande de connaître l’état du système après la période transitoire pour ensuite « remonter » le flot de l’équation différentielle réduite (1.2). Une méthode a été trouvée récemment dans [4] et permet d’obtenir un développement de cette condition modifiée jusqu’à un certain ordre ε^k . Cette approche est très utile lorsqu’on ne cherche pas à capturer la période transitoire du système et que ε est petit. Elle permet de réduire la complexité du problème *a priori* par principe de réduction (i.e. en se limitant à la variété, où z est déterminé par x), et fournit un système dont la seule raideur provient de f . Cependant, elle est difficile à implémenter du fait de sa nature algébrique complexe, et n’est pas adaptée si l’on veut capturer la phase transitoire dans une simulation.

On obtient néanmoins une erreur en $\mathcal{O}(\Delta t^p + \varepsilon^q)$ sur (x, z) après la phase transitoire.

Dans [4], la méthode permet également de trouver un nouveau modèle asymptotiquement stable où x est indépendante, peu raide et où on peut calculer z avec le nouveau modèle. L’équation sur z peut être raide, en revanche, et les calculs restent assez formels et complexes. On utilise les résultats de cet article en annexe C.3 pour vérifier que l’erreur est la bonne et pour illustrer cette approche.

1.2 Introduction au développement double-échelle

De la même manière que pour le cas périodique dans [5], on considère $u = (x, z)$ comme évaluation particulière d'une fonction $U = (X, Z)$ à deux variables t, τ ,

$$u(t) = U(t, t/\varepsilon)$$

pour laquelle on doit trouver un bon ensemble de définition. Cette décomposition est aussi à la base de la méthode d'homogénéisation. Dans le cas périodique de période P , il était naturel de faire évoluer la seconde variable τ dans $\mathbb{R}/P\mathbb{Z}$, puisqu'elle représentait la partie périodique du problème. On verra quel ensemble choisir ici.

Quel que soit cet ensemble, on trouve que X, Z sont solutions d'un problème de transport

$$\begin{aligned} \partial_t X + \frac{1}{\varepsilon} \partial_\tau X &= f(X, Z) \\ \partial_t Z + \frac{1}{\varepsilon} \partial_\tau Z &= -\frac{1}{\varepsilon} Z + g(X, Z) \end{aligned} \tag{1.5}$$

où la seule donnée qu'on connaît est $X(0, 0) = x_0$ et $Z(0, 0) = z_0$.

On va d'abord discuter des conditions supplémentaires dont on a besoin pour que le problème soit bien posé avec τ dans un intervalle ou ensemble qu'on choisira, puis on exhibera des conditions pour que le problème soit bien posé et que sa solution soit d'une certaine régularité qu'on précisera.

1.2.1 Nouvelles problématiques

Comme nous l'avons déjà remarqué, la première difficulté qui se pose est de choisir l'ensemble de définition pour τ . Ce choix est en réalité assez simple. Dans le cas périodique, l'ensemble de définition de τ était indépendant de ε . On souhaite garder cette propriété dans le cas non-périodique, et on observe que si x, z sont définies sur $[0, T]$, τ parcourt $[0, T/\varepsilon]$. Ainsi on choisit \mathbb{R}_+ comme ensemble pour τ , i.e. on aimerait que U soit définie sur $[0, T] \times \mathbb{R}_+$, i.e.

$$U : [0, T] \times \mathbb{R}_+ \rightarrow E := \mathbb{R}^n \times \mathbb{R}^m.$$

On peut maintenant analyser le problème (1.5) dans $[0, T] \times \mathbb{R}_+$. Il s'agit d'un problème de transport à vitesse constante $1/\varepsilon$, donc l'analyse se fera surtout par méthode des caractéristiques. Considérons dans un premier temps un problème indépendant de ε ,

$$\partial_t u + \partial_\tau u = F(t, \tau, u)$$

avec $F : [0, T] \times \mathbb{R}_+ \times E \rightarrow E$. On voit qu'on peut obtenir un problème similaire à partir de (1.5) avec le changement de variable $\tau \leftarrow \varepsilon \tau$, et en fixant ε pour obtenir F indépendante de ε . Ce problème est donc, à ε fixé, plus général que celui sur X, Z .

L'aspect de $[0, T] \times \mathbb{R}_+$ fait qu'on doit s'intéresser tout particulièrement à la droite $(t, t), t \in [0, T]$. Les droites caractéristiques sont de la forme $(t, \tau + t), \tau \in \mathbb{R}_+, t \in [0, T]$ sous cette diagonale et $(t + t_0, t), t_0 \in [0, T], t \in [0, T - t_0]$ au dessus. On voit ainsi que certaines caractéristiques ont leur origine sur la demi-droite $(0, \tau), \tau \geq 0$ et d'autre sur le segment

$(t_0, 0), t_0 \in [0, T]$, donc pour pouvoir résoudre le problème il faut nécessairement une condition initiale *et* une condition au bord.

C'est une différence majeure avec le problème périodique, qui ne nécessitait pas de condition au bord. Sans condition au bord, la régularité de la condition initiale est transmise naturellement à la solution. Nous allons voir très rapidement que ce n'est pas le cas ici. Considérons le problème avec une régularité presque minimale
Trouver $(\bar{T}, u) \in]0, T] \times C^0([0, \bar{T}] \times \mathbb{R}_+)$ tel que pour tout $(t, \tau) \in [0, \bar{T}] \times \mathbb{R}_+$ on ait

$$\partial_t u + \partial_\tau u = F(t, \tau, u) \quad (1.6)$$

avec les conditions initiales $u(0, \tau) = u_I(\tau), \tau \in \mathbb{R}_+$ et les conditions au bord $u(t, 0) = u_B(t), t \in [0, \bar{T}]$. La restriction à $\bar{T} \leq T$ permet d'assurer qu'on ne cherche pas nécessairement une solution sur $[0, T] \times \mathbb{R}_+$ entier, où l'existence ne serait pas toujours garantie à cause des non-linéarités de F .

Admettons qu'on choisisse u_I et u_B quelconques dans $C^0(\mathbb{R}_+)$ et $C^0([0, T])$ respectivement. Il est évident que si $u_I(\tau = 0) \neq u_B(t = 0)$, la solution (si elle existe) ne peut pas être continue, puisqu'alors $\lim_{\delta \rightarrow 0^+} u(\delta, 0) \neq u(0, 0)$. Réciproquement,

Lemme 2. Si $u_B \in C_b^0([0, T])$, $u_I \in C_b^0(\mathbb{R}_+)$ avec $u_B(t = 0) = u_I(\tau = 0) =: u_0$, si F est localement lipschitzienne par rapport à u , continue de $[0, T] \times \mathbb{R}_+ \times E$ dans E , bornée par rapport à τ , alors le problème (1.6) est bien posé au sens suivant :

Pour tout facteur $\kappa > 1$, il existe un temps $T_\kappa \in]0, T]$ tel que le problème (1.6) a une unique solution u dans $C^0([0, T_\kappa] \times \mathbb{R}^+)$, qui respecte l'inégalité

$$\forall t \in [0, T_\kappa], \quad \|u(t, \cdot)\|_{L_t^\infty} \leq \kappa \max\{\|u_I\|_{L_t^\infty}, \|u_B\|_{L_t^\infty}\}. \quad (1.7)$$

Si en outre F respecte $\forall (t, \tau, u) \in [0, T] \times \mathbb{R}_+ \times E, |F(t, \tau, u)| \leq C_F|u| + D_F$ pour des constantes $C_F, D_F \geq 0$, alors il existe une unique solution u dans $C^0([0, T] \times \mathbb{R}_+)$ qui vérifie

$$\forall t \in [0, T], \|u(t, \cdot)\|_{L_t^\infty} \leq \left(\max\{\|u_I\|_{L_t^\infty}, \|u_B\|_{L_t^\infty}\} + tD_F\right)e^{tC_F}. \quad (1.8)$$

Démonstration. La preuve utilise les droites caractéristiques du problème et ne pose pas de difficultés particulières. Néanmoins, sa bonne compréhension nécessite de séparer les fonctions caractéristiques qui ont comme origine une condition initiale de celle qui prennent une condition au bord. Cette distinction rend la preuve un peu longue et elle se trouve donc en annexe B, p. 47. \square

Exhibons maintenant des conditions supplémentaires sur u_B et u_I pour que la solution u du problème soit de classe au moins C^2 .

1.2.2 Caractère bien posé dans C^2

Pour simplifier l'étude, on suppose maintenant F indépendante de t, τ , i.e. $F : E \rightarrow E$. Avec κ fixé, quitte à ce que ce soit au sens des distributions dans $L^\infty([0, T_\kappa] \times \mathbb{R}_+)$, on peut dériver le problème de transport (1.6) selon t pour obtenir que $\partial_t u^\varepsilon =: v^\varepsilon$ vérifie

$$\begin{aligned} \partial_t v + \partial_\tau v &= \partial_u F(u) \cdot v \\ v(0, \tau) &= F(u_I(\tau)) - u'_I(\tau) \quad \text{et} \quad v(t, 0) = u'_B(t) \end{aligned} \quad (1.9)$$

D'après l'étude précédente, il apparaît immédiatement que pour que v soit continue, il faut non seulement avoir $u_I \in C^1(\mathbb{R}_+)$, $u_B \in C^1([0, T_\kappa])$ et $\partial_u F$ continue de E dans $L(E, E)$, mais aussi

$$u'_B(t=0) + u'_I(\tau=0) = F(u_0) \quad (1.10)$$

où $u_0 := u_I(\tau=0) = u_B(t=0)$. Admettons que nos conditions initiale et au bord sont bien choisies de sorte que v ainsi définie soit continue (i.e. de sorte que la relation (1.10) soit vérifiée). Alors on peut dériver à nouveau pour obtenir une équation sur $\partial_t v =: w$

$$\partial_t w + \partial_\tau w = \partial_u^2 F(u) \cdot (v, v) + \partial_u F(u) \cdot w \quad (1.11)$$

$$w(0, \tau) = \partial_u F(u_I) \cdot (v(0, \tau) - u'_I) + u''_I \quad \text{et} \quad w(t, 0) = u''_B(t)$$

où u_I, u'_I et u''_I sont évaluées en τ dans la condition initiale. On voit que la condition de continuité peut alors s'écrire

$$\partial_u F(u_0) \cdot u'_B(0) - u''_B(0) = \partial_u F(u_0) \cdot u'_I(0) - u''_I(0). \quad (1.12)$$

On peut alors exhiber des conditions suffisantes pour que la solution soit de classe C^2 dans le problème (1.6), avec F indépendante de t, τ .

Théorème 3. Si $u_B \in C_b^2([0, T])$, $u_I \in C_b^2(\mathbb{R}_+)$, si F (resp. $\partial_u F$ et $\partial_u^2 F$) est continue de E dans E (resp. $L(E, E)$ et $L(E^2, E)$), et enfin si

$$u_B(0) = u_I(0) =: u_0, \quad u'_B(0) + u'_I(0) = F(u_0),$$

$$\text{et} \quad \partial_u F(u_0) \cdot u'_B(0) - u''_B(0) = \partial_u F(u_0) \cdot u'_I(0) - u''_I(0)$$

alors le problème (1.6) est bien posé et régulier au sens suivant :

Pour tout facteur $\kappa > 1$, il existe un temps $T_\kappa \in]0, T]$ tel que le problème (1.6) a une unique solution u dans $C^2([0, T_\kappa] \times \mathbb{R}^+)$, qui respecte l'inégalité

$$\forall t \in [0, T_\kappa], \quad \|u(t, \cdot)\|_{L_t^\infty} \leq \kappa \max\{\|u_B\|_{L_t^\infty}, \|u_I\|_{L_\tau^\infty}\}. \quad (1.13)$$

et avec $C'_0 = \sup\{\|\partial_u F(u(t, \tau))\|, (t, \tau) \in [0, T_\kappa] \times \mathbb{R}_+\}$ on a

$$\forall t \in [0, T_\kappa], \quad \|\partial_t u(t, \cdot)\|_{L_t^\infty} \leq \max\{\|u'_B\|_{L_t^\infty}, \|v_I\|_{L_\tau^\infty}\} e^{tC'_0}, \quad (1.14)$$

où $v_I = F(u_I) - u'_I$. De la même manière, avec $C''_1 = \|\partial_u^2 F(u) \cdot (v, v)\|_{L_{t,\tau}^\infty}$, on a

$$\forall t \in [0, T_\kappa], \quad \|\partial_t^2 u(t, \cdot)\|_{L_t^\infty} \leq \left(\max\{\|u''_B\|_{L_t^\infty}, \|w_I\|_{L_\tau^\infty}\} + tC''_1\right) e^{tC'_0} \quad (1.15)$$

où $w_I = \partial_u F(u_I) \cdot (v_I - u'_I) + u''_I$.

Démonstration. Soit $\kappa > 1$. Le lemme 2 assure l'existence de $(T_\kappa, u) \in]0, T] \times C_b^0([0, T_\kappa] \times \mathbb{R}_+)$ vérifiant l'inégalité souhaitée. On peut ensuite appliquer le cas affine du lemme 2 successivement aux problèmes (1.9) puis (1.11) en considérant u puis v comme des paramètres, et on obtient $\partial_t u \in C_b^0([0, T_\kappa] \times \mathbb{R}_+)$ puis $\partial_t^2 u \in C_b^0([0, T_\kappa] \times \mathbb{R}_+)$, avec les bornes souhaitées.

On remarque enfin

$$\partial_\tau u = F(u) - \partial_t u, \quad \partial_\tau \partial_t u = \partial_u F(u) \cdot \partial_t u - \partial_t^2 u \quad \text{et} \quad \partial_\tau^2 = \partial_u F(u) \cdot \partial_\tau u - \partial_t \partial_\tau u$$

donc la régularité de $u, \partial_t u$ et $\partial_t^2 u$ se propage aux dérivées première et seconde de u pour assurer $u \in C^2([0, T_\kappa] \times \mathbb{R}_+)$.

□

On voit ainsi que les conditions initiale et au bord sont intimement liées. Puisqu'on est libre de choisir les deux, on se dit que le plus simple est de fixer l'une des deux pour ensuite choisir l'autre de manière adaptée. On choisit de se concentrer d'abord sur la condition initiale, pour des raisons qui vont apparaître rapidement.

Remarque 1. Pour avoir une solution d'ordre C^3 , on doit respecter

$$u_B''' + u_I''' = \partial_u^2 F(u_0) \cdot (u_B', u_B') + \partial_u^2 F(u_0) \cdot (u_I', u_I') - \partial_u^2 F(u_0) \cdot (u_B', u_I') + \partial_u F(u_0) \cdot (u_B'' + u_I'').$$

1.2.3 Condition initiale ou condition au bord ?

Le but est de trouver de bonnes données (conditions initiale et au bord) de sorte à avoir $X, \partial_t X, \partial_t^2 X$ et $Z, \partial_t Z, \partial_t^2 Z$ uniformément bornés sur $[0, T_\kappa] \times \mathbb{R}_+$ afin de pouvoir trouver un schéma numérique uniformément précis (UP) d'ordre 1, i.e. d'erreur globale $\mathcal{O}(\Delta t)$ à constante indépendante de ε . L'inspiration provient naturellement des méthodes d'Euler pour lesquelles l'erreur locale est d'ordre $\Delta t^2(\|\partial_t^2 X\| + \|\partial_t^2 Z\|)$.

Mettons d'abord le problème (1.5) sous la forme

$$\partial_t U + \frac{1}{\varepsilon} L U = F(U) \tag{1.16}$$

avec $U = \begin{pmatrix} X \\ Z \end{pmatrix}$ et $F(U) = \begin{pmatrix} f(X, Z) \\ g(X, Z) \end{pmatrix}$, des conditions initiale U_I et au bord U_B , et L un opérateur bien choisi. En l'occurrence,

$$L U := \begin{pmatrix} L_x X \\ L_z Z \end{pmatrix} := \begin{pmatrix} \partial_\tau X \\ e^{-\tau} \partial_\tau [e^\tau Z] \end{pmatrix}.$$

On admet pour le moment qu'on peut borner $\partial_t U$ en fonction de $\frac{1}{\varepsilon} L U_I$ puisque $\partial_t U(0, \tau) = F(U_I) - \frac{1}{\varepsilon} L U_I$ et que le lemme 3 va dans ce sens. On voit alors qu'un mauvais choix de la condition initiale peut faire apparaître un facteur $1/\varepsilon$ dans $\partial_t U$.

On voit en revanche que pour la valeur au bord de $\partial_t U$, seule la dérivée U_B' compte : il n'y a pas de facteur $1/\varepsilon$ qui apparaît. Même pour les dérivées en τ , on a $\partial_\tau U(t, 0) = -U_B + \varepsilon(F(U_B) - U_B')$, donc U_B semble ne pouvoir faire apparaître un facteur $1/\varepsilon$ que si sa dérivée en fait apparaître un.

En outre, si U_I est telle que $\partial_t U(0, \cdot) =: V_I(\cdot)$ et $\partial_t^2 U(0, \cdot) =: W_I(\cdot)$ sont définies (il suffit d'avoir $U_I \in C^2$) et uniformément bornées, i.e.

$$\exists C > 0, \forall \varepsilon \in]0, \varepsilon^*], \|V_I\|_{L_t^\infty}, \|W_I\|_{L_t^\infty} \leq C$$

alors on peut définir

$$U_B(t) = U_I(0) + t V_I(0) + \frac{1}{2} t^2 W_I(0) \tag{1.17}$$

pour obtenir une solution de classe C^2 qui, on l'espère (et on le montrera ensuite), est uniformément bornée sur $[0, T] \times \mathbb{R}_+$ avec ses dérivées $\partial_t U, \partial_t^2 U$ qui le sont aussi.

PARTIE 2

MISE EN PLACE D'UN SCHÉMA NUMÉRIQUE UNIFORMÉMENT CONVERGENT

Maintenant qu'on sait comment obtenir un problème régulier, on cherche des conditions initiales qui permettent d'obtenir une condition au bord uniformément bornée avec la formule (1.17). On montre ensuite que le problème est alors bien posé et donne $U, \partial_t U$ et $\partial_t^2 U$ uniformément bornées, et on traduit les conditions initiales et au bord explicitement pour X, Z . Enfin, on met en place un schéma numérique uniformément convergent dont on prouve les convergences locales et globales.

Par la suite on ne considère plus la variété centrale $\varepsilon h^\varepsilon$, si bien que les notations h, h^ε pourront représenter n'importe quel type de fonctions.

2.1 Choix de la condition initiale

Cette section décrit une analyse du problème à l'aide d'outils utiles au cas périodique. À cet effet, les méthodes présentées partent d'hypothèses de régularité qui ne sont pas explicitées.

D'abord on présente une approche peu formelle par développement en puissances de ε qui a fait le sujet d'une bonne partie du stage. Face aux limites de cette approche, on exhibe une meilleure méthode pour trouver les bonnes conditions initiales pour avoir des dérivées ∂_t^α uniformément bornée jusqu'à un ordre $\alpha \in \mathbb{N}$ arbitraire.

2.1.1 Approche initiale

Dans un premier temps, on a essayé de faire le lien entre un développement multi-échelles en puissances de ε et la régularité sous la diagonale. On écrit ce développement

$$X^\varepsilon(t, \tau) = \sum_{n \in \mathbb{N}} \varepsilon^n X_n(t, \tau) \quad \text{et} \quad Z^\varepsilon(t, \tau) = \sum_{n \in \mathbb{N}} \varepsilon^n Z_n(t, \tau)$$

en supposant les X_n, Z_n et leurs dérivées uniformément bornées en ε .

Analysons le comportement de la solution approchée $X = X_0, Z = Z_0$. On a par développement selon les puissance de ε

$$\begin{aligned} \partial_\tau X_0 &= 0, & \partial_\tau Z_0 + Z_0 &= 0 & (\varepsilon^{-1}) \\ \partial_t X_0 &= f(X_0, Z_0), & \partial_t Z_0 &= g(X_0, Z_0) & (\varepsilon^0) \end{aligned}$$

La condition initiale associée à ce problème « tronqué » est alors évidente : on a

$$X(0, \tau) = x_0, \quad Z(0, \tau) = e^{-\tau} z_0. \quad (2.1)$$

On a alors effectivement $\partial_t X(0, \tau) = f(x_0, e^{-\tau} z_0)$ et $\partial_t Z(0, \tau) = g(x_0, e^{-\tau} z_0)$, ce qui est uniformément borné sur \mathbb{R}_+ , mais en dérivant à nouveau on obtient $\partial_t^2 X(0, \tau) = \mathcal{O}(1/\varepsilon)$ et $\partial_t^2 Z(0, \tau) = \mathcal{O}(1/\varepsilon)$.

On tronque maintenant à l'ordre 1 : $X = X_0 + \varepsilon X_1, Z = Z_0 + \varepsilon Z_1$. Comme précédemment, on a

$$\begin{aligned} \partial_\tau X_0 &= 0, & \partial_\tau Z_0 + Z_0 &= 0 & (\varepsilon^{-1}) \\ \partial_\tau X_1 &= f(X_0, Z_0) - \partial_t X_0, & \partial_\tau Z_1 + Z_1 &= g(X_0, Z_0) - \partial_t Z_0 & (\varepsilon^0) \\ \partial_t X_1 &= \Delta_\varepsilon f(X_0, Z_0) \cdot (X_1, Z_1), & \partial_t Z_1 &= \Delta_\varepsilon g(X_0, Z_0) \cdot (X_1, Z_1) & (\varepsilon^1) \end{aligned}$$

où $\Delta_\varepsilon f(X_0, Z_0) \cdot (X_1, Z_1) = \frac{1}{\varepsilon} [f(X_0 + \varepsilon X_1, Z_0 + \varepsilon Z_1) - f(X_0, Z_0)] = f_x(X_0, Z_0) \cdot X_1 + f_z(X_0, Z_0) \cdot Z_1 + \mathcal{O}(\varepsilon)$, une sorte de dérivée discrète.

Puisque $X(0, 0)$ et $Z(0, 0)$ sont fixés comme indépendant de ε , on pose $X_1(0, 0) = 0$ et $Z_1(0, 0) = 0$. Alors on a toujours $X_0(t, \tau) = x_0(t), Z_0(t, \tau) = e^{-\tau} z_0(t)$ (avec l'abus de notation où x_0, z_0 sont des fonctions du temps qui coïncident avec leur définition précédente en 0), mais aussi

$$\begin{aligned} X_1(t, \tau) &= \int_0^\tau \left[f(x_0(t), e^{-\sigma} z_0(t)) - \dot{x}_0(t) \right] d\sigma + x_1(t) \\ Z_1(t, \tau) &= \int_0^\tau e^{-(\tau-\sigma)} \left[g(x_0(t), e^{-\sigma} z_0(t)) - e^{-\sigma} \dot{z}_0(t) \right] d\sigma + e^{-\tau} z_1(t) \end{aligned}$$

avec $x_1(0) = 0, z_1(0) = 0$ et \dot{x}_1, \dot{z}_1 donnés par le système d'ordre ε^1 (ces fonctions n'influencent de toute manière pas les conditions initiales).

On voit qu'on n'a pas encore de relation pour \dot{x}_0 ni \dot{z}_0 . On peut cependant en extraire une pour \dot{x}_0 . En effet, puisqu'on veut avoir X_1 bornée, il faut nécessairement $\dot{x}_0 = f(x_0, 0)$. Le choix de \dot{z}_0 , est arbitraire, et on voit par le calcul que quel que soit notre choix, on obtient une condition initiale et des dérivées $X, \partial_t X, \partial_t^2 X$ et $Z, \partial_t Z, \partial_t^2 Z$ uniformément bornées sur $(0, \tau), \tau > 0$.

Lorsqu'on veut étendre le développement à l'ordre 2 en ε , i.e. $X = X_0 + \varepsilon X_1 + \varepsilon^2 X_2$ et $Z = Z_0 + \varepsilon Z_1 + \varepsilon^2 Z_2$, on se rend compte que les calculs deviennent plus compliqués et qu'on a encore beaucoup de choix à faire sur $\ddot{x}_0, \ddot{z}_0, \dot{x}_1, \dot{z}_1$. On va donc chercher un cadre formel qu'on pourra utiliser pour trouver des conditions initiales permettant d'avoir des $\partial_t^\alpha X(0, \tau), \partial_t^\alpha Z(0, \tau)$ uniformément bornées jusqu'à α arbitraire.

2.1.2 Développement de Chapman-Enskog

On reprend la formulation (1.16) :

$$\partial_t U + \frac{1}{\varepsilon} L U = F(U).$$

Étudions l'opérateur L dans l'ensemble $\mathcal{E} = \{h : \mathbb{R}^+ \rightarrow E \mid \exists (h_n)_{n \in \mathbb{N}}, \forall \tau \in \mathbb{R}^+, h(\tau) = \sum_n h_n e^{-n\tau}\}$, l'ensemble des séries exponentielles. Cet ensemble est plus contraignant que les preuves suivantes ne le nécessitent, mais cela rendra les raisonnements plus

clairs, et est en accord avec l'analogie des fonctions périodiques qui sont aussi des séries, de la forme $\sum_n h_n e^{in\tau}$. Pour assurer que \mathcal{E} est stable par F , on suppose F analytique.

Les noyaux de L_x et L_z sont respectivement $Vect(\tau \mapsto 1_{\mathbb{R}^n})$ et $Vect(\tau \mapsto e^{-\tau} 1_{\mathbb{R}^m})$. On note Π_x et Π_z les projections sur ces noyaux. Sur le complémentaire des noyaux respectifs de L_x et L_z (i.e. sur les noyaux respectifs de Π_x et Π_z), on a

$$\begin{aligned}(L_x^{-1}a)(\tau) &= -\int_{\tau}^{\infty} a(\sigma) d\sigma = (I - \Pi_x) \int_0^{\tau} a(\sigma) d\sigma, \\ (L_z^{-1}b)(\tau) &= b_0 - \int_{\tau}^{\infty} e^{-\tau+\sigma} (b(\sigma) - b_0) d\sigma = (I - \Pi_z) \int_0^{\tau} e^{-\tau+\sigma} b(\sigma) d\sigma.\end{aligned}$$

Il est intéressant de remarquer qu'on doit modifier un peu l'intégrande pour pouvoir définir l'inverse de L_z . En définissant $o_x : \tau \mapsto 1_{\mathbb{R}^n}$ et $o_z : \tau \mapsto e^{-\tau} 1_{\mathbb{R}^m}$ les fonctions de base de $\text{Ker } L_x$ et $\text{Ker } L_z$ respectivement, on remarque

$$(L_x^{-1}a)(\tau) = (I - \Pi_x)(o_x * a),$$

$$(L_z^{-1}b)(\tau) = (I - \Pi_z)(o_z * b).$$

où $*$ représente le produit de convolution. De manière générale, si $Th = e^{-p\tau} \partial_{\tau} [e^{p\tau} h]$ avec $h = \sum_{k \neq p} h_k e^{-k\tau} \in \mathcal{E} / \text{Ker } T$ on inverse $(T^{-1}h)(\tau) = \sum_{k \neq p} \frac{1}{p-k} h_k e^{-k\tau} = \sum_{k < p} \frac{1}{p-k} h_k e^{-k\tau} - \int_{\tau}^{\infty} e^{-k(\tau-\sigma)} (h(\sigma) - \sum_{k < p} h_k e^{-k\sigma}) d\sigma$.

Supposons connaître la solution U^{ε} . Le développement de Chapman-Enskog consiste à la décomposer en deux parties

$$U^{\varepsilon}(t, \tau) = \underline{U}^{\varepsilon}(t, \tau) + h^{\varepsilon}(t, \tau)$$

où $\underline{U}^{\varepsilon} := \Pi U^{\varepsilon}$ et $\Pi h^{\varepsilon} = 0$ (on rappelle que h^{ε} n'est pas la variété du théorème 1). On peut injecter cette décomposition de U^{ε} dans le système de départ, puis projeter sur le noyau de L pour obtenir les équations

$$\partial_t \underline{U}^{\varepsilon} = \Pi F(\underline{U}^{\varepsilon} + h^{\varepsilon}),$$

$$\partial_t h^{\varepsilon} + \frac{1}{\varepsilon} L h^{\varepsilon} = (I - \Pi) F(\underline{U}^{\varepsilon} + h^{\varepsilon}).$$

Comme $\Pi h^{\varepsilon} = 0$ on déduit de la dernière relation

$$h^{\varepsilon} = \varepsilon A F(\underline{U}^{\varepsilon} + h^{\varepsilon}) - \varepsilon L^{-1}(\partial_t h^{\varepsilon}).$$

avec $A := L^{-1}(I - \Pi)$ (défini sur \mathcal{E} tout entier). En supposant U^{ε} suffisamment régulière, on peut décomposer h en puissances de ε :

$$h^{\varepsilon}(t, \tau) = \varepsilon h_1(t, \tau, \underline{U}^{\varepsilon}) + \varepsilon^2 h_2(t, \tau, \underline{U}^{\varepsilon}) + \dots$$

Et alors en supposant $h^{\varepsilon}, \partial_t h^{\varepsilon} = \mathcal{O}_{\varepsilon \rightarrow 0}(1)$ on obtient $h^{\varepsilon} = \mathcal{O}(\varepsilon)$. Si on suppose en outre $\partial_t^2 h^{\varepsilon} = \mathcal{O}(1)$ on obtient

$$h^{\varepsilon} = \varepsilon h_1 + \mathcal{O}(\varepsilon^2) := \varepsilon A F(\underline{U}^{\varepsilon}) + \mathcal{O}(\varepsilon^2).$$

On s'attend à avoir une solution régulière à l'ordre 2 si les conditions initiales respectent ce développement $U_0 = \underline{U}_0 + \varepsilon h_1$, c'est-à-dire avec

$$U_0^\varepsilon(\tau) = \underline{U}_0^\varepsilon(\tau) + \varepsilon AF(\underline{U}_0^\varepsilon(\tau)). \quad (2.2)$$

où $\underline{U}_0^\varepsilon = \Pi \underline{U}_0^\varepsilon = \Pi U_0^\varepsilon$ pourra être déterminé à partir de $U_0^\varepsilon(0,0) = u_0$ (seul degré de liberté qui correspond aux conditions initiales du problème (1.1)).

Remarque 2. On pourrait aller un ordre au dessus en supposant $\partial_t^3 h^\varepsilon = \mathcal{O}(1)$ et alors on obtiendrait

$$\begin{aligned} U_0^\varepsilon(\tau) = & \underline{U}_0^\varepsilon(\tau) + \varepsilon AF(\underline{U}_0^\varepsilon) \\ & + \varepsilon^2 \left(A \partial_u F(\underline{U}_0^\varepsilon) AF(\underline{U}_0^\varepsilon) - A^2 \partial_u F(\underline{U}_0^\varepsilon) \Pi F(\underline{U}_0^\varepsilon) \right). \end{aligned}$$

2.2 Bornes uniformes de la solution

L'étude de la régularité en section 1.2.2 nous indique qu'on peut définir $U_\alpha^\varepsilon(\tau) = (\partial_t^\alpha U^\varepsilon)(0, \tau)$, $\alpha = 0, 1, 2$ à partir de F et des dérivées de U_0^ε . On choisit alors la condition au bord

$$U_B^\varepsilon(t) = U_0^\varepsilon(0) + t U_1^\varepsilon(0) + \frac{t^2}{2} U_2^\varepsilon(0). \quad (2.3)$$

Si $U_0^\varepsilon(0), U_1^\varepsilon(0), U_2^\varepsilon(0) = \mathcal{O}_{\varepsilon \rightarrow 0}(1)$, alors $U_B^\varepsilon, \partial_t U_B^\varepsilon$ et $\partial_t^2 U_B^\varepsilon$ sont uniformément bornées sur $[0, T]$. Montrons qu'avec ce choix de conditions initiale et au bord, on obtient une solution U^ε avec les propriétés souhaitées.

2.2.1 Théorème sur $U = (X, Z)$

Théorème 4. Considérons le problème de transport (1.16) avec les conditions initiales (2.2)

$$U_0^\varepsilon(\tau) = \underline{U}_0^\varepsilon(\tau) + \varepsilon L^{-1}(I - \Pi)F(\underline{U}_0^\varepsilon(\tau))$$

et les conditions au bord (2.3). On suppose F analytique. Alors

- (i) Les conditions initiales et au bord sont uniformément bornées en norme infinie, et il existe pour tout $\kappa > 1$ un temps T_κ tel que pour tout $\varepsilon \in]0, \varepsilon^*]$ une unique solution U^ε dans $C^2([0, T_\kappa] \times \mathbb{R}_+)$ qui vérifie

$$\forall t \in [0, T_\kappa], \quad \sup_{\varepsilon \in]0, \varepsilon^*]} \|U^\varepsilon(t, \cdot)\|_{L_t^\infty} \leq \kappa \sup_{\varepsilon \in]0, \varepsilon^*]} \left(\max \left\{ \|U_0^\varepsilon\|_{L_t^\infty}, \|U_B^\varepsilon\|_{L_t^\infty} \right\} \right) < +\infty. \quad (2.4)$$

- (ii) Pour tout T_κ assurant l'inégalité (2.4), il existe une constante C_κ telle que la solution U^ε vérifie

$$\forall t \in [0, T_\kappa], \quad \sup_{\varepsilon \in]0, \varepsilon^*]} \|\partial_t^\alpha U^\varepsilon(t, \cdot)\|_{L_t^\infty} \leq C_\kappa, \quad \alpha = 0, 1, 2. \quad (2.5)$$

Démonstration. Le théorème 3 nous permet d'affirmer que le problème est bien posé dans $C^2([0, T_\kappa] \times \mathbb{R}_+)$ à ε et κ fixé. La difficulté va être de montrer que T_κ et les bornes peuvent être indépendantes de ε .

Étude de la condition au bord. Montrons d'abord que la condition au bord est uniformément bornée (i.e. $\sup_\varepsilon \|U_B\|_{L_t^\infty} < \infty$). On remarque d'abord $U_0^\varepsilon(0) = \mathcal{O}_{\varepsilon \rightarrow 0}(1)$ et on calcule

$$U_1^\varepsilon = F(U_0^\varepsilon) - \frac{1}{\varepsilon} L U_0^\varepsilon = F_0^\varepsilon - (I - \Pi) \underline{F}_0^\varepsilon$$

avec $F_0^\varepsilon = F(U_0^\varepsilon)$ et $\underline{F}_0^\varepsilon = F(\underline{U}_0^\varepsilon)$. Cette condition initiale est uniformément bornée, donc on a $U_1^\varepsilon(0) = \mathcal{O}_{\varepsilon \rightarrow 0}(1)$. Remanions un peu les termes pour trouver une expression plus simple à dériver. Par développement limité autour de $\underline{U}_0^\varepsilon$ dans $F(U_0^\varepsilon)$ on trouve

$$\begin{aligned} F_0^\varepsilon - \underline{F}_0^\varepsilon &= \partial_u F(\underline{U}_0^\varepsilon)(U_0^\varepsilon - \underline{U}_0^\varepsilon) + \mathcal{O}(\|U_0^\varepsilon - \underline{U}_0^\varepsilon\|^2) \\ &= \varepsilon \partial_u F(\underline{U}_0^\varepsilon) A \underline{F}_0^\varepsilon + \mathcal{O}(\varepsilon^2) \end{aligned}$$

$$\text{d'où} \quad U_1^\varepsilon = \Pi \underline{F}_0^\varepsilon + \varepsilon \partial_u F(\underline{U}_0^\varepsilon) A \underline{F}_0^\varepsilon + \varepsilon^2 r_1^\varepsilon$$

avec le terme correctif $r_1^\varepsilon \in \mathcal{E}$ qui vient prendre en compte le reste dans le développement autour de $\underline{U}_0^\varepsilon$. On a ensuite

$$U_2^\varepsilon = \partial_u F(U_0^\varepsilon) \cdot U_1^\varepsilon - \frac{1}{\varepsilon} L U_1^\varepsilon = \partial_u F(U_0^\varepsilon) - L \partial_u F(\underline{U}_0^\varepsilon) A \underline{F}_0^\varepsilon + \mathcal{O}(\varepsilon)$$

et donc $U_2^\varepsilon(0) = \mathcal{O}_{\varepsilon \rightarrow 0}(1)$. On a ainsi montré que $U_0^\varepsilon, U_1^\varepsilon, U_2^\varepsilon$ et $U_B^\varepsilon, (U_B^\varepsilon)', (U_B^\varepsilon)''$ sont uniformément bornées.

Existence de T_κ et de U , et borne uniforme. On écrit le problème sous la forme

$$\partial_t U^\varepsilon + \partial_\theta U^\varepsilon = -\frac{1}{\varepsilon} J U^\varepsilon + F(U^\varepsilon),$$

avec J diagonale, de coefficients 0 ou 1 selon que le coefficient agisse sur X ou Z . Ce problème est presque celui de la partie 1 à l'exception du terme $-(1/\varepsilon)J U^\varepsilon$ et de l'indépendance de F par rapport à t, τ (et d'un changement de variable $\tau \leftarrow \varepsilon \tau$).

Soit $\kappa > 1$. Montrons que les résultats du lemme 2 sont maintenus, avec un temps final T_κ et une borne à la solution tous les deux indépendants de ε .

On s'intéresse d'abord aux caractéristiques sous la diagonale pour nous ramener au cas de la partie précédente. Soit $\tau \geq 0$, on pose $\varphi^\varepsilon(t, \tau) = U^\varepsilon(t, \tau + t/\varepsilon)$, alors $\varphi^\varepsilon(t, \tau)$ vérifie l'EDO

$$\partial_t \varphi^\varepsilon(t, \tau) = -\frac{1}{\varepsilon} J \varphi^\varepsilon(t, \tau) + F(\varphi^\varepsilon(t, \tau)) \quad \text{et} \quad \varphi^\varepsilon(0, \tau) = U_0^\varepsilon(\tau)$$

Sous forme intégrale, le problème s'écrit

$$\varphi^\varepsilon(t, \tau) = e^{-tJ} \varphi^\varepsilon(0, \tau) + \int_0^\tau e^{(s-t)J} F(\varphi^\varepsilon(s, \tau)) ds$$

soit, par inégalité triangulaire et croissance de l'intégrale, en remarquant $\|e^{-tJ}\| \leq 1$,

$$|\varphi^\varepsilon(t, \tau)| \leq |\varphi^\varepsilon(0, \tau)| + \int_0^\tau |F(\varphi^\varepsilon(s, \tau))| ds. \quad (2.6)$$

On est alors exactement dans le même cas que dans la démonstration du lemme 2 (voir annexe B). Il en est de même *au dessus* de la diagonale $(t, t/\varepsilon)$. Notons $R := \sup_\varepsilon \max\{\|U_0^\varepsilon\|_{L_t^\infty}, \|\underline{U}_0^\varepsilon\|_{L_t^\infty}\}$. On peut alors définir un temps final T_κ indépendant de ε tel que U^ε est bien définie et est bornée par κR sur $[0, T_\kappa] \times \mathbb{R}_+$ pour tout ε . En particulier, on passe au supremum pour obtenir l'inégalité (2.4) du théorème.

Continuité et caractère borné des dérivées. On rappelle que $\partial_t U^\varepsilon$ (resp. $\partial_t^2 U^\varepsilon$) vérifie un problème de transport linéaire avec pour conditions initiale U_1 (resp. U_2) et au

bord U_B' (resp. U_B''). De même que précédemment, on peut montrer que le cas affine du lemme 2 reste valide, avec une borne indépendante de ε . Détaillons le raisonnement pour $\partial_t U^\varepsilon =: V^\varepsilon$ afin de s'en convaincre. On a

$$\begin{aligned}\partial_t V^\varepsilon + \frac{1}{\varepsilon} \partial_t V^\varepsilon &= -\frac{1}{\varepsilon} J V^\varepsilon + \partial_u F(U^\varepsilon) \cdot V^\varepsilon, \\ V^\varepsilon(0, \tau) &= U_1(\tau) \quad \text{et} \quad V(t, 0) = (U_B^\varepsilon)'(t).\end{aligned}$$

En notant $G^\varepsilon : t, \tau, v \mapsto \partial_u F(U^\varepsilon(t, \tau)) \cdot v$, cela s'écrit sous forme intégrale pour une caractéristique $\psi^\varepsilon(t, \tau) = V^\varepsilon(t, \tau + t/\varepsilon)$ sous la diagonale, $\psi^\varepsilon(t, \tau) = e^{-\frac{t}{\varepsilon} J} \psi^\varepsilon(0, \tau) + \int_0^t e^{\frac{(s-t)}{\varepsilon} J} G^\varepsilon(s, \tau + s/\varepsilon, \psi^\varepsilon(s, \tau)) ds$. Comme précédemment, on peut majorer $|\psi^\varepsilon(t, \tau)| \leq |\psi^\varepsilon(0, \tau)| + \int_0^t |G^\varepsilon(s, \tau + s/\varepsilon, \psi^\varepsilon(s, \tau))| ds$.

Comme U^ε est bornée sur $[0, T_\kappa] \times \mathbb{R}_+$, $\partial_u F(U^\varepsilon)$ l'est aussi en norme d'opérateur. On pose $A_\kappa, B_\kappa > 0$ tels que $\forall(\varepsilon, t, \tau, v) \in]0, \varepsilon^*] \times [0, T_\kappa] \times \mathbb{R}_+ \times E, |G^\varepsilon(t, \tau, v)| \leq A_\kappa |v| + B_\kappa$ (en l'occurrence on pourrait choisir $B_\kappa = 0$ mais ce n'est pas le cas pour $\partial_t^2 U^\varepsilon$ donc conservons le). L'inégalité de Grönwall assure alors

$$\forall t \in [0, T_\kappa], \quad \|V^\varepsilon(t, \cdot)\|_{L_t^\infty} \leq \left(\max \left\{ \|(U_B^\varepsilon)'\|_{L_t^\infty}, \|U_1^\varepsilon\|_{L_\tau^\infty} \right\} + t B_\kappa \right) e^{t A_\kappa}.$$

On peut effectuer le même raisonnement sur $\partial_t^2 U^\varepsilon$ qui donne

$$\forall t \in [0, T_\kappa], \quad \|\partial_t^2 U^\varepsilon(t, \cdot)\|_{L_t^\infty} \leq \left(\max \left\{ \|(U_B^\varepsilon)''\|_{L_t^\infty}, \|U_2^\varepsilon\|_{L_\tau^\infty} \right\} + t B'_\kappa \right) e^{t A_\kappa}$$

pour une certaine constante B'_κ . Comme $U_1^\varepsilon, U_2^\varepsilon, (U_B^\varepsilon)'$ et $(U_B^\varepsilon)''$ sont uniformément bornées, on peut passer au supremum dans l'inégalité pour obtenir la deuxième partie du théorème. □

Remarque 3. Le résultat reste vrai si on prend des conditions initiales $\tilde{U}_0 = U_0 + \varepsilon^2 r_0^\varepsilon$ avec $r_0^\varepsilon \in \mathcal{E}$ uniformément bornée. Cependant, la condition au bord doit changer en conséquence.

2.2.2 Forme explicite des conditions initiales et au bord sur X, Z

Traduisons l'expression des conditions initiales (2.2) dans le système en X, Z . On a par simple traduction

$$\begin{aligned}X^\varepsilon(0, \tau) &= X_0^\varepsilon(\tau) = \underline{x}_0 - \varepsilon \int_\tau^\infty [f(\underline{x}_0, e^{-\sigma} \underline{z}_0) - f(\underline{x}_0, 0)] d\sigma, \\ Z^\varepsilon(0, \tau) &= Z_0^\varepsilon(\tau) = e^{-\tau} \underline{z}_0 + \varepsilon g(\underline{x}_0, 0) - \varepsilon \int_\tau^\infty e^{-(\tau-\sigma)} [g(\underline{x}_0, e^{-\sigma} \underline{z}_0) - g(\underline{x}_0, 0) \\ &\quad - \partial_z g(\underline{x}_0, 0) \cdot e^{-\sigma} \underline{z}_0] d\sigma.\end{aligned} \tag{2.7}$$

On note

$$f_k = \partial_z^k f(x_0, 0) \cdot \underbrace{(z_0, \dots, z_0)}_{k \text{ fois}} \tag{2.8}$$

ou \underline{f}_k si les arguments sont $\underline{x}_0, \underline{z}_0$. On définit g_k, \underline{g}_k de la même manière. Dans les applications on fixe $X(0, 0) = x_0$ et $Z(0, 0) = z_0$, donc on doit inverser un système en $\underline{x}_0, \underline{z}_0$ qui s'écrit

$$\begin{aligned}\underline{x}_0 &= x_0 + \varepsilon \int_0^\infty [f(\underline{x}_0, e^{-\sigma} \underline{z}_0) - \underline{f}_0] d\sigma \\ \underline{z}_0 &= z_0 - \varepsilon \underline{g}_0 + \varepsilon \int_0^\infty e^\sigma [g(\underline{x}_0, e^{-\sigma} \underline{z}_0) - \underline{g}_0 - e^{-\tau} \underline{g}_1] d\sigma\end{aligned}$$

$$\begin{aligned} \text{d'où } \underline{x}_0 &= x_0 + \varepsilon \int_0^\infty [f(x_0, e^{-\sigma} z_0) - f_0] d\sigma + \mathcal{O}(\varepsilon^2), \\ \underline{z}_0 &= z_0 - \varepsilon g_0 + \varepsilon \int_0^\infty e^\sigma [g(x_0, e^{-\sigma} z_0) - g_0 - e^{-\tau} g_1] d\sigma + \mathcal{O}(\varepsilon^2) \end{aligned}$$

grâce au caractère analytique de f et g . En injectant cela dans les définitions de X_0, Z_0 on obtient après calcul

$$\begin{aligned} X_0^\varepsilon(\tau) &= x_0 + \varepsilon \int_0^\tau [f(x_0, e^{-\sigma} z_0) - f_0] d\sigma + \mathcal{O}(\varepsilon^2), \\ Z_0^\varepsilon(\tau) &= e^{-\tau} z_0 + \varepsilon e^{-\tau} \int_0^\tau e^\sigma [g(x_0, e^{-\sigma} z_0) - \partial_z g(x_0, 0) \cdot e^{-\sigma} z_0] d\sigma + \mathcal{O}(\varepsilon^2). \end{aligned}$$

La remarque 3 nous indique que ce niveau de précision (à ε) est suffisant et donc on tronque les termes d'ordre ε^2 . Ainsi on modifie les définitions de $X_0^\varepsilon, Z_0^\varepsilon$ en

$$\begin{aligned} X_0^\varepsilon(\tau) &= x_0 + \varepsilon \int_0^\tau [f(x_0, e^{-\sigma} z_0) - f(x_0, 0)] d\sigma, \\ Z_0^\varepsilon(\tau) &= e^{-\tau} z_0 + \varepsilon \int_0^\tau e^{-\tau+\sigma} [g(x_0, e^{-\sigma} z_0) - \partial_z g(x_0, 0) \cdot e^{-\sigma} z_0] d\sigma. \end{aligned} \tag{2.9}$$

Déterminons maintenant les grandeurs nécessaires à la condition au bord telle que donnée par la relation (2.3). On a

$$\begin{aligned} X_1^\varepsilon(\tau) &= f(X_0^\varepsilon(\tau), Z_0^\varepsilon(\tau)) - \frac{1}{\varepsilon} (X_0^\varepsilon)'(\tau) = f(X_0^\varepsilon(\tau), Z_0^\varepsilon(\tau)) - f(x_0, e^{-\tau} z_0) + f_0 \\ Z_1^\varepsilon(\tau) &= g(X_0^\varepsilon(\tau), Z_0^\varepsilon(\tau)) - \frac{1}{\varepsilon} ((Z_0^\varepsilon)'(\tau) + Z_0^\varepsilon(\tau)) = g(X_0^\varepsilon(\tau), Z_0^\varepsilon(\tau)) - g(x_0, e^{-\tau} z_0) + e^{-\tau} g_1 \end{aligned}$$

d'où

$$X_1^\varepsilon(0) = f(x_0, 0), \quad Z_1^\varepsilon(0) = \partial_z g(x_0, 0) \cdot z_0.$$

Par définition on a ensuite

$$\begin{aligned} X_2 &= \partial_x f(X_0, Z_0) \cdot X_1 + \partial_z f(X_0, Z_0) \cdot Z_1 - \frac{1}{\varepsilon} X_1', \\ Z_2 &= \partial_x g(X_0, Z_0) \cdot Z_1 + \partial_z g(X_0, Z_0) \cdot Z_1 - \frac{1}{\varepsilon} (Z_1' + Z_1). \end{aligned}$$

Après quelques calculs sans difficulté on peut finalement évaluer $X_2^\varepsilon(0), Z_2^\varepsilon(0)$ de façon exacte

$$\begin{aligned} X_2(0) &= \partial_x f(x_0, z_0) \cdot (2f(x_0, 0) - f(x_0, z_0)) + \partial_z f(x_0, z_0) \cdot (2\partial_z g(x_0, 0) \cdot z_0 - g(x_0, z_0)) \\ Z_2(0) &= \partial_x g(x_0, z_0) \cdot (2f(x_0, 0) - f(x_0, z_0)) + \partial_z g(x_0, z_0) \cdot (2\partial_z g(x_0, 0) \cdot z_0 - g(x_0, z_0)) \end{aligned}$$

On définit enfin les conditions au bord

$$\begin{aligned} X_B(t) &= x_0 + t f_0 + \frac{t^2}{2} [\partial_x f(x_0, z_0) \cdot (2f_0 - f(x_0, z_0)) + \partial_z f(x_0, z_0) \cdot (2g_1 - g(x_0, z_0))] \\ Z_B(t) &= z_0 + t g_1 + \frac{t^2}{2} [\partial_x g(x_0, z_0) \cdot (2f_0 - f(x_0, z_0)) + \partial_z g(x_0, z_0) \cdot (2g_1 - g(x_0, z_0))] \end{aligned} \tag{2.10}$$

où $f_0 = f(x_0, 0)$ et $g_1 = \partial_z g(x_0, 0) \cdot z_0$ comme définis par (2.8). On remarque que ces conditions au bord sont indépendantes de ε (bien que ce n'ait pas d'influence sur la suite). On annonce alors le théorème

Théorème 5. Avec les conditions initiales (2.9) et au bord (2.10), et si f et g sont analytiques, alors le problème (1.5) est bien posé au sens

- (i) Les conditions initiales et au bord sont uniformément bornées en norme infinie, et il existe pour tout $\kappa > 1$ un temps T_κ tel que pour tout $\varepsilon \in]0, \varepsilon^*]$ une unique solution $X^\varepsilon, Z^\varepsilon$ dans $C^2([0, T_\kappa] \times \mathbb{R}_+)$ qui vérifie

$$\forall t \in [0, T_\kappa], \quad \sup_{\varepsilon \in]0, \varepsilon^*]} \|X^\varepsilon(t)\|_{L_t^\infty} \leq \kappa R_X \quad \text{et} \quad \sup_{\varepsilon} \|Z^\varepsilon(t)\|_{L_t^\infty} \leq \kappa R_Z$$

où R_X (resp. R_Z) est un majorant de $\sup_{\varepsilon} \|X_0^\varepsilon\|_{L_t^\infty}$ et de $\|X_B\|_{L_t^\infty}$ (resp. de $\sup_{\varepsilon \in]0, \varepsilon^*]} \|Z_0^\varepsilon\|_{L_t^\infty}$ et de $\|Z_B\|_{L_t^\infty}$).

- (ii) Pour tout T_κ de la relation précédente, on a

$$\forall t \in [0, T_\kappa], \quad \sup_{\varepsilon \in]0, \varepsilon^*]} \|\partial_t^\alpha X^\varepsilon(t, \tau)\|_{L_t^\infty} \leq C_\kappa \quad \text{et} \quad \sup_{\varepsilon \in]0, \varepsilon^*]} \|\partial_t^\alpha Z^\varepsilon(t, \cdot)\|_{L_t^\infty} \leq C_\kappa$$

pour $\alpha = 0, 1, 2$, avec $C_\kappa > 0$ une constante.

Démonstration. C'est une application directe du théorème 4 avec la remarque 3. \square

2.3 Présentation du schéma d'ordre 1

On cherche maintenant à calculer numériquement la solution du problème (1.5). On ne discrétise que la variable t et on garde la variable τ continue. On définit une grille uniforme $t_n = n\Delta t$ sur $[0, T_\kappa]$, $n = 0, \dots, N$ avec $N\Delta t = T_\kappa$. On note $X_n^\varepsilon \simeq X^\varepsilon(t_n, \cdot)$ et $Z_n^\varepsilon \simeq Z^\varepsilon(t_n, \cdot)$, et le schéma qu'on utilise est

$$\begin{aligned} \frac{X_{n+1}^\varepsilon - X_n^\varepsilon}{\Delta t} + \frac{1}{\varepsilon} \partial_\tau X_{n+1}^\varepsilon &= f(X_n, Z_n), \\ \frac{Z_{n+1}^\varepsilon - Z_n^\varepsilon}{\Delta t} + \frac{1}{\varepsilon} (\partial_\tau Z_{n+1}^\varepsilon + Z_{n+1}^\varepsilon) &= g(X_n, Z_n), \end{aligned}$$

2.3.1 Propriétés du schéma

Pour X on voit qu'on peut écrire, avec $\mu = \varepsilon/\Delta t$ et $f_n^\varepsilon = f(X_n^\varepsilon, Z_n^\varepsilon)$,

$$\mu X_{n+1}^\varepsilon + \partial_\tau X_{n+1}^\varepsilon = \mu(\Delta t f_n^\varepsilon + X_n^\varepsilon)$$

ou encore, en remarquant $\mu X(\tau) + \partial_\tau X(\tau) = e^{-\mu\tau} \partial_\tau [e^{\mu\tau} X(\tau)]$ et en intégrant entre 0 et τ ,

$$X_{n+1}^\varepsilon(\tau) = e^{-\mu\tau} X_B^\varepsilon(t_{n+1}) + \mu \int_0^\tau e^{-\mu(\tau-\sigma)} (\Delta t f(X_n^\varepsilon(\sigma), Z_n^\varepsilon(\sigma)) + X_n^\varepsilon(\sigma)) d\sigma. \quad (2.11)$$

Ceci définit précisément le schéma, qu'on cherchera à implémenter une fois sa convergence prouvée. De la même manière on a

$$Z_{n+1}^\varepsilon(\tau) = e^{-(\mu+1)\tau} Z_B^\varepsilon(t_{n+1}) + \mu \int_0^\tau e^{-(\mu+1)(\tau-\sigma)} (\Delta t g(X_n^\varepsilon(\sigma), Z_n^\varepsilon(\sigma)) + Z_n^\varepsilon(\sigma)) d\sigma. \quad (2.12)$$

On remarque que les opérateurs $Q_\mu = Id + \frac{1}{\mu} \partial_\tau$ et $S_\mu = Id + \frac{1}{\mu} (Id + \partial_\tau)$ sont centraux à notre étude, et donc on propose d'étudier quelques unes de leurs propriétés tout de suite. Pour simplifier cette étude, on pose $T_\mu = \begin{pmatrix} Q_\mu \\ S_\mu \end{pmatrix}$.

Lemme 6. L'opérateur $T_\mu[h_0]$ défini de $C^1(\mathbb{R}_+ | h_0) := \{h \in C^1(\mathbb{R}_+) \mid h(0) = h_0\}$ dans $C^0(\mathbb{R}_+)$ par

$$T_\mu[h_0]h = h + \frac{1}{\mu}(\partial_\tau h + Jh)$$

est inversible, et l'inverse de $C^0(\mathbb{R}_+)$ dans $C^1(\mathbb{R}_+ | h_0)$ s'écrit explicitement

$$\forall \tau \in \mathbb{R}_+, \quad (T_\mu[h_0]^{-1}h)(\tau) = e^{-\tau(\mu I + J)}h_0 + \mu \int_0^\tau e^{-(\tau-\sigma)(\mu I + J)}h(\sigma)d\sigma.$$

En outre, cet inverse vérifie l'inégalité

$$\forall h \in C^0(\mathbb{R}_+), \quad \|T_\mu[h_0]^{-1}h\|_{L^\infty_\tau} \leq \|h\|_{L^\infty_\tau}.$$

Démonstration. On a $T_\mu[h_0]h = \begin{pmatrix} Q_\mu[\pi_x h_0]h_x \\ S_\mu[\pi_z h_0]h_z \end{pmatrix}$, avec π_x, π_z les projections de E sur $\mathbb{R}^n, \mathbb{R}^m$ respectivement, et $h_x = \pi_x h$, $h_z = \pi_z h$. La formule d'inversion a été montrée pour $Q_\mu[\pi_x h_0]$, et celle pour $S_\mu[\pi_z h_0]$ en (2.12), se trouve exactement de la même manière. L'énoncé provient simplement de la mise en commun de ces deux relations en une seule. L'inégalité apparaît par inégalité triangulaire sur les intégrales dans (2.11) et (2.12) en remarquant

$$\mu \int_0^\tau e^{-\mu(\tau-\sigma)}d\sigma = 1 - e^{-\mu\tau} \leq 1 \quad \text{et} \quad \mu \int_0^\tau e^{-(\mu+1)(\tau-\sigma)}d\sigma \leq \frac{\mu}{\mu+1} \leq 1$$

□

On remarque par ailleurs que l'ensemble de départ de $T_\mu[h_0]$ n'est pas un espace vectoriel. On a $T_\mu[\alpha]h_1 + T_\mu[\beta]h_2 = T_\mu[\alpha + \beta](h_1 + h_2)$.

2.3.2 Erreur locale du schéma

Théorème 7. Soit $U^\varepsilon = (X^\varepsilon, Z^\varepsilon)$ la solution du problème (1.16) avec les conditions initiales $X_0^\varepsilon, Z_0^\varepsilon$ de (2.9) et les conditions au bord X_B, Z_B de (2.10). Alors l'erreur locale ℓe_{n+1}^ε du schéma donné par (2.11) et (2.12), définie pour $n = 0, \dots, N-1$ par

$$\ell e_{n+1}^\varepsilon := U^\varepsilon(t_{n+1}) - \hat{U}_{n+1}^\varepsilon$$

avec $(T_\mu[U_B(t_{n+1})]\hat{U}_{n+1}^\varepsilon)(\cdot) = U^\varepsilon(t_n, \cdot) + \Delta t F(U^\varepsilon(t_n, \cdot))$. Alors

$$\sup_{\varepsilon \in]0, \varepsilon^*]} \|\ell e_{n+1}^\varepsilon\|_{L^\infty_\tau} \leq C_\kappa \Delta t^2$$

pour une certaine constante C_κ indépendante de n .

Démonstration. Par la suite, on omet la variable τ sauf en cas de nécessité. Par définition de T_μ on a

$$T_\mu[U_B(t_{n+1})]U^\varepsilon(t_{n+1}) = U^\varepsilon(t_{n+1}) + \Delta t (F(U^\varepsilon(t_{n+1})) - \partial_t U^\varepsilon(t_{n+1}))$$

et alors

$$T_\mu[0](\ell e_{n+1}^\varepsilon) = T_\mu[U_B^\varepsilon(t_{n+1})]U^\varepsilon(t_{n+1}) - T_\mu[U_B(t_{n+1})]\hat{U}_{n+1}^\varepsilon = \alpha_{n+1}^\varepsilon$$

avec

$$\alpha_{n+1}^\varepsilon := U^\varepsilon(t_{n+1}) - U^\varepsilon(t_n) + \Delta t (F(U^\varepsilon(t_{n+1})) - F(U^\varepsilon(t_n)) - \partial_t U^\varepsilon(t_{n+1})).$$

Par développement de Taylor à reste intégral de $U^\varepsilon(t_n)$ autour de t_{n+1} , on a

$$U^\varepsilon(t_{n+1}) - U^\varepsilon(t_n) = \Delta t \partial_t U^\varepsilon(t_{n+1}) + \int_{t_n}^{t_{n+1}} (t_n - t) \partial_t^2 U^\varepsilon(t) dt$$

$$\text{et } F(U^\varepsilon(t_{n+1})) - F(U^\varepsilon(t_n)) = \int_{t_n}^{t_{n+1}} \partial_u F(U^\varepsilon(t)) \cdot \partial_t U^\varepsilon(t) dt$$

d'où la relation

$$\alpha_{n+1}^\varepsilon = \Delta t \int_{t_n}^{t_{n+1}} \left(\partial_u F(U^\varepsilon(t)) \cdot \partial_t U^\varepsilon(t) + \frac{t_n - t}{\Delta t} \partial_t^2 U^\varepsilon(t) \right) dt.$$

D'après le théorème 5 et par continuité de $\partial_u F$, les quantités dans l'intégrale sont uniformément bornées, et donc

$$\sup_{\varepsilon \in]0, \varepsilon^*]} \|\alpha_{n+1}^\varepsilon\|_{L^\infty_\tau} \leq C_\kappa \Delta t^2.$$

D'après le lemme 6 et comme $\ell e_{n+1}^\varepsilon = T_\mu[0]^{-1} \alpha_{n+1}$, on obtient le théorème souhaité. \square

2.3.3 Erreur globale

Théorème 8. Soit $U^\varepsilon = (X^\varepsilon, Z^\varepsilon)$ la solution du problème (1.16) avec les conditions initiales X_0, Z_0 de (2.9) et les conditions au bord X_B, Z_B de (2.10). On pose $(U_n^\varepsilon)_{n \in [0, N]}$ la solution numérique calculée grâce au schéma donné par (2.11) et (2.12) sur \mathbb{R}_+ avec U_0^ε donnée par les conditions initiales sur $X^\varepsilon, Z^\varepsilon$ (2.9). Alors il existe $\Delta t_0 > 0$ et $C > 0$ telles que

$$\forall \Delta t < \Delta t_0, \quad \sup_{\varepsilon \in]0, \varepsilon^*]} \|U^\varepsilon(t_n) - U_n^\varepsilon\|_{L^\infty_\tau} \leq C \Delta t$$

pour tout $n = 0, \dots, N$, où $T_\kappa = N \Delta t$ est le temps final.

Démonstration. On peut décomposer l'erreur $e_{n+1}^\varepsilon := U^\varepsilon(t_{n+1}) - U_{n+1}^\varepsilon$ en deux parties

$$e_{n+1}^\varepsilon = \underbrace{U^\varepsilon(t_{n+1}) - \hat{U}_{n+1}^\varepsilon}_{\text{erreur locale } \ell e_{n+1}^\varepsilon} + \underbrace{\hat{U}_{n+1}^\varepsilon - U_{n+1}^\varepsilon}_{\text{erreur transportée } te_{n+1}^\varepsilon}$$

avec comme précédemment $(T_\mu[U_B(t_{n+1})] \hat{U}_{n+1}^\varepsilon)(\cdot) = U^\varepsilon(t_n, \cdot) + \Delta t F(U^\varepsilon(t_n, \cdot))$. On a

$$T_\mu[0] e_{n+1}^\varepsilon = T_\mu[0] \ell e_{n+1}^\varepsilon + T_\mu[0] te_{n+1}^\varepsilon$$

(on remarque que e_{n+1}^ε est nulle en 0 par définition du schéma, exact au bord). On sait $T_\mu[0] \ell e_{n+1}^\varepsilon \leq C_0 \Delta t^2$, mais on ne sait pas encore quantifier $T_\mu[0] te_{n+1}^\varepsilon$. On remarque

$$T_\mu[0] te_{n+1}^\varepsilon = e_n^\varepsilon + \Delta t (F(U^\varepsilon(t_n)) - F(U_n^\varepsilon)).$$

Par continuité de $\partial_u F$, formule d'intégration et inégalité triangulaire, on a

$$\|F(U^\varepsilon(t_n)) - F(U_n^\varepsilon)\|_{L_\tau^\infty} \leq \int_0^1 \left\| \partial_u F(\theta U^\varepsilon(t_n) + (1-\theta)U_n^\varepsilon) \cdot e_n^\varepsilon \right\|_{L_\tau^\infty} d\theta$$

Fixons $R > 0$. On suppose $\sup_\varepsilon \|e_n^\varepsilon\|_{L_\tau^\infty} < R$, alors pour tout $\theta \in [0, 1]$,

$$\|\theta U^\varepsilon(t_n) + (1-\theta)U_n^\varepsilon\|_{L_\tau^\infty} \leq \|U^\varepsilon(t_n)\| + \theta\|e_n^\varepsilon\| \leq C_1 + R$$

d'où par continuité de $\partial_u F$ de E dans $L(E, E)$,

$$\|F(U^\varepsilon(t_n)) - F(U_n^\varepsilon)\|_{L_\tau^\infty} \leq C_2 \|e_n^\varepsilon\|_{L_\tau^\infty}$$

pour une certaine constante $C_2 > 0$ qui dépend de R (et de C_1). On obtient enfin, avec le lemme 6,

$$\|e_{n+1}^\varepsilon\|_{L_\tau^\infty} \leq (1 + C_2 \Delta t) \|e_n^\varepsilon\|_{L_\tau^\infty} + C_0 \Delta t^2$$

ce qui fournit, avec l'inégalité de Grönwall discrète,

$$\|e_{n+1}^\varepsilon\|_{L_\tau^\infty} \leq \frac{C_0 \Delta t}{C_2} (\exp(C_2 T_\kappa) - 1).$$

Par récurrence on assure $\|e_n^\varepsilon\| < R$ pour tout $n \in \llbracket 0, N \rrbracket$ pour tout $\Delta t < \Delta t_0$ en posant

$$\Delta t_0 := \frac{RC_2}{C_0} e^{-C_2 T_\kappa}$$

et la démonstration est terminée. □

On a ainsi montré que ce schéma similaire au schéma implicite-explicite 1.1.2, *a priori* très simple, est uniformément précis. On peut maintenant l'implémenter et vérifier numériquement cette convergence.

PARTIE 3

IMPLÉMENTATIONS ET RÉSULTATS NUMÉRIQUES

Dans ce schéma numérique, on suppose pouvoir représenter parfaitement les fonctions de τ . Cela paraît impossible avec une implémentation usuelle. On a alors le choix entre plusieurs possibilités :

1. Discrétiser \mathbb{R}_+ et faire des intégrations numériques
2. Identifier des fonctions de base pour avoir une approche algébrique
3. Trouver un moyen de représenter très précisément les fonctions de τ , sans exhiber une base particulière

Quelle que soit l'approche finale, il peut être intéressant d'identifier les fonctions de base qui forment $(U_n^\varepsilon)_n$. Il est facile de voir que si on a quelques fonctions de bases, tous leurs multiples sont aussi des fonctions de base, puisque f et g forment des multiples de ces fonctions. Ainsi, d'après la condition initiale, les fonctions $(\tau \mapsto e^{-k\tau})_{k \in \mathbb{N}}$ sont des fonctions de base.

Mais ce n'est pas tout. En effet, la première itération fait apparaître une composante $e_\mu : \tau \mapsto e^{-\mu\tau}$ dans $X_1^\varepsilon(\tau)$, or $Q_\mu[0]^{-1}e_\mu(\tau) = \tau e^{-\mu\tau}$, donc les fonctions de base sont les

$$\tau \mapsto \tau^i e^{-(j\mu+k)\tau} \quad \text{pour } (i, j, k) \in \mathbb{N}^3 \text{ avec } j = 0 \Rightarrow i = 0$$

Notons que si $\mu \in \mathbb{Q}$, la famille décrite ci-dessus n'est pas libre, ce qui peut poser problème algébriquement.

En réalité, ce qui pose le plus problème pour une implémentation algébrique, c'est qu'on peut avoir i grand pour j, k faibles, e.g. on peut rapidement avoir une composante $\tau^{10^3} e^{-\mu\tau}$ qui prend des valeurs très élevées sur \mathbb{R}_+ si μ est faible, ce qui la rend difficile à évaluer. Le coefficient de cette composante serait suffisamment faible pour compenser les différents ordres de grandeur, mais on devrait quand même stocker beaucoup de valeurs et traiter avec des ordres très variés. Pour ces raisons, on ne va pas s'intéresser à une implémentation algébrique.

On présente maintenant quelques unes des implémentations qu'on a mises en place. Hors mention explicite d'un cas particulier, tous les tests de convergence sont effectués sur le cas particulier (A.1). Pour rappel, ce cas correspond à $n, m = 1$, $f(x, z) = -x^3(z - z^3/3)$ et $g(x, z) = x(1 - z^2/2)$, avec $x_0 = 0,8$ et $z_0 = 0,05$.

Sur les figures, on prend pour référence des droites de « pente s » où s correspondrait à l'ordre de la méthode.

3.1 Implémentations « simples » et convergences

On étudie d'abord une approche différente sur des séries exponentielles. Celle-ci n'est pas fructueuse, mais on présente ainsi la base des raisonnements spectraux qui pourront être intéressants dans de futures recherches. Ensuite, on étudie une implémentation par différences finies du schéma.

3.1.1 Schéma classiques sur des séries

On peut facilement montrer qu'il existe une suite de fonctions $(u_k^\varepsilon : [0, T_\kappa] \rightarrow E)_{k \in \mathbb{N}}$ telle que

$$\forall (t, \tau) \in [0, T_\kappa] \times \mathbb{R}_+, \tau \geq t/\varepsilon \Rightarrow u^\varepsilon(t, \tau) = \sum_{k \in \mathbb{N}} u_k^\varepsilon(t) e^{-k\tau}.$$

Cette propriété découle de la stabilité de \mathcal{E} , l'ensemble des séries exponentielles, par f, g et du fait que $X_0^\varepsilon, Z_0^\varepsilon \in \mathcal{E}$. À travers des simulations, on se rend compte que les coefficients u_k^ε ne sont pas uniformément bornés. On les « normalise » alors par $e^{kt/\varepsilon}$: on effectue le changement de variable $u_k^\varepsilon(t) \leftarrow u_k^\varepsilon e^{-kt/\varepsilon}$ ce qui transforme la relation précédente en

$$\forall (t, \tau) \in [0, T_\kappa] \times \mathbb{R}_+, \tau \geq t/\varepsilon \Rightarrow u^\varepsilon(t, \tau) = \sum_{k \in \mathbb{N}} u_k^\varepsilon(t) e^{-k(\tau - t/\varepsilon)}. \quad (3.1)$$

En particulier $u^\varepsilon(t, t/\varepsilon) = \sum_k u_k^\varepsilon(t)$. Par développement de Taylor sur F autour de u_0^ε , on trouve des EDO sur les u_k^ε .

$$\partial_t u_0^\varepsilon + \frac{1}{\varepsilon} J u_0^\varepsilon = F(u_0^\varepsilon) \quad (3.2)$$

$$\partial_t u_1^\varepsilon + \frac{1}{\varepsilon} J u_1^\varepsilon = F'(u_0^\varepsilon) \cdot u_1^\varepsilon \quad (3.3)$$

$$\partial_t u_k^\varepsilon + \frac{1}{\varepsilon} J u_k^\varepsilon = F'(u_0^\varepsilon) \cdot u_k^\varepsilon + R_k(u_1^\varepsilon, \dots, u_{k-1}^\varepsilon), \quad k \geq 2 \quad (3.4)$$

où les fonction R_k sont à déterminer à partir du développement de Taylor (il s'agit de produits de Cauchy imbriqués qu'on ne détaillera pas ici).

Avec le schéma Radau IIa on calcule l'amplitude des coefficients en fonction de k pour décider d'un rang à partir duquel on ne calcule plus les u_k^ε . On voit alors qu'on peut choisir un rang maximal $k_{\max} = 20$ dans notre cas. Appliquons maintenant un schéma RK4 au polynôme $(u_k^\varepsilon)_{0 \leq k \leq 20}$ à Δt fixé et pour diverses valeurs de ε pour étudier si le bon choix de coefficients suffit à diminuer la raideur du problème pour obtenir des schémas uniformément précis par défaut. On voit sur la figure 3.2 que ce n'est absolument pas le cas, et que la convergence est la même qu'en considérant le système d'origine, sans séries. La stabilité est directement liée à $\Delta t/\varepsilon$. On observe cela sur l'erreur de z en fonction de ε , où tracer une droite $\sim 1/\varepsilon$ passe par le dernier point avant que la solution ne diverge pour différentes valeurs de Δt .

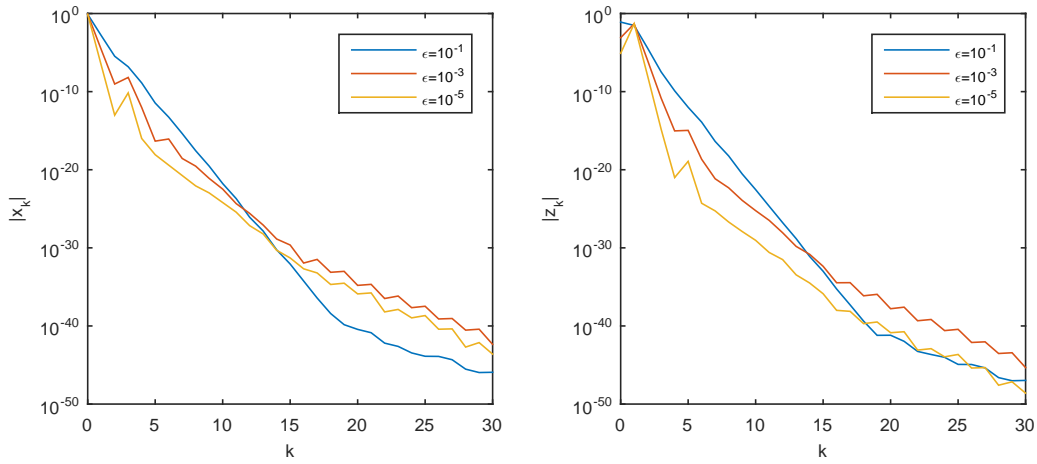


FIGURE 3.1 – Norme infinie des coefficients (x_k^ϵ) et (z_k^ϵ) en fonction de k dans le cas (1.3).

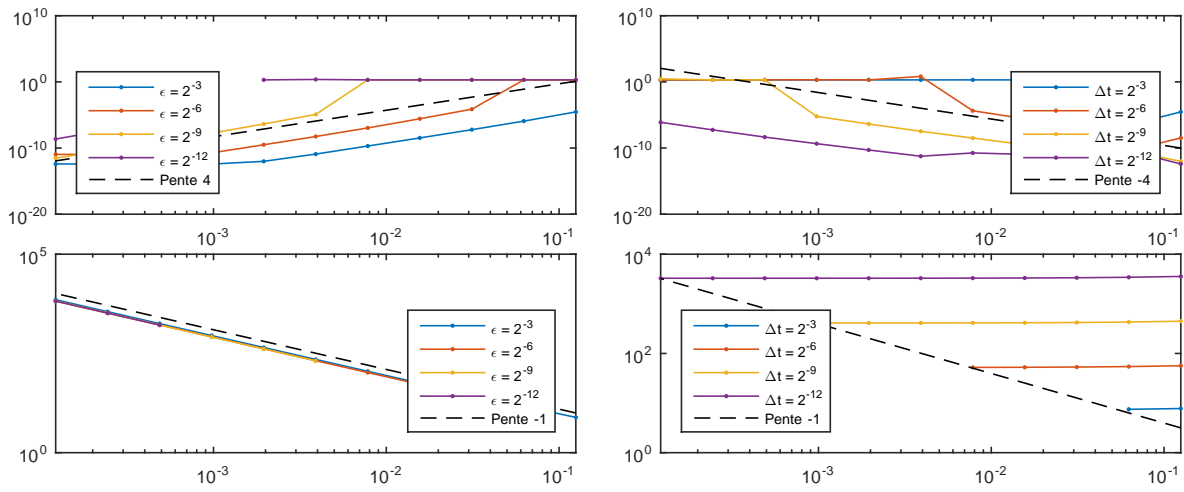


FIGURE 3.2 – Erreur relative en norme infinie sur x (en haut) et erreur absolue en norme infinie sur z (en bas) en fonction de Δt (à gauche) et de ϵ (à droite) pour le cas (1.3) avec un schéma RK4 sur les (u_k^ϵ) .

3.1.2 Implémentation par différences finies

On voit qu'on peut restreindre l'ensemble de définition de τ à $[0, T_\kappa/\epsilon]$. Il faut donc discrétiser ce segment. Afin de profiter de la forme exponentielle des fonctions, on choisit une discrétisation de la forme

$$\tau_n = -C \ln(\xi_n)$$

avec $(\xi_n)_{0 \leq n \leq N_\tau}$ une discrétisation uniforme de $[e^{-T_\kappa/\epsilon}, 1]$ (décroissante pour avoir (τ_n) croissante), et C une constante positive qui permet d'avoir une meilleure résolution sur les fonctions de base $(\tau \mapsto e^{-k\tau})$, $k \in \mathbb{N}$ et $(\tau \mapsto e^{-j\mu\tau})$, $j \in \mathbb{N}^*$. On choisit $C = \max(1/\mu, 1)$ pour un bon compromis après une rapide étude de l'erreur d'intégration (avec la formule du trapèze) sur ces fonctions.

Intégration du schéma

On voit que la seule partie un peu difficile à implémenter pour le schéma sont les produit de convolution par $\tau \mapsto e^{-\mu\tau}$ et $\tau \mapsto e^{-(\mu+1)\tau}$. On approxime

$$\begin{aligned} \int_0^{\tau_n} e^{-\mu(\tau_n-\sigma)} h(\sigma) d\sigma &= \sum_{k=0}^{n-1} \int_{\tau_k}^{\tau_{k+1}} e^{-\mu(\tau_n-\sigma)} h(\sigma) d\sigma \\ &\simeq \sum_{k=0}^{n-1} \left(\int_{\tau_k}^{\tau_{k+1}} e^{\mu(\sigma-\tau_n)} d\sigma \right) \frac{h(\tau_k) + h(\tau_{k+1})}{2} \end{aligned}$$

soit pour le schéma,

$$\begin{aligned} Q_\mu[0]^{-1} h(\tau) &\simeq \sum_{k=0}^{n-1} e^{-\mu\tau_n} (e^{\mu\tau_{k+1}} - e^{\mu\tau_k}) \frac{h(\tau_k) + h(\tau_{k+1})}{2}, \\ S_\mu[0]^{-1} h(\tau) &\simeq \frac{\mu}{\mu+1} \sum_{k=0}^{n-1} e^{-(\mu+1)\tau_n} (e^{(\mu+1)\tau_{k+1}} - e^{(\mu+1)\tau_k}) \frac{h(\tau_k) + h(\tau_{k+1})}{2}. \end{aligned}$$

Remarque 4. Si on voulait effectuer ce produit de convolution entre deux fonctions discrétisées, l'implémentation serait compliquée puisque la discrétisation n'est pas uniforme. En effet, dans le cas d'une discrétisation régulière on aurait $\int_0^{\tau_n} a(\tau-\sigma)b(\sigma)d\sigma = \sum_{k=0}^{n-1} \frac{\Delta\tau}{2} (a(\tau_n-\tau_k)b(\tau_k) + a(\tau_n-\tau_{k+1})b(\tau_{k+1})) + \mathcal{O}(\tau_n\Delta\tau^2)$ par règle du trapèze. Dans le cas d'une discrétisation non uniforme, on n'a pas forcément accès aux valeurs en $\tau_n - \tau_k$, ce qui rend cela impossible.

Une fois ces calculs faits, on implémente le schéma et on peut évaluer l'erreur.

Étude de l'erreur

On veut s'assurer que le schéma permet effectivement de résoudre le système (1.1) avec une précision uniforme. Pour cela, on calcule l'erreur pour diverses valeurs de Δt et de ε sur le cas particulier (1.3), avec $N_\tau = 2^{10}$ (sachant que dans [5], la discrétisation est à $32 = 2^5$ points et les résultats sont bons).

On voit sur la figure 3.3 que le nombre de points N_τ est insuffisant pour espérer avoir une quelconque convergence, malgré le fait qu'il soit élevé. On se rend compte que diminuer Δt et ε nécessite d'augmenter N_τ , ce qui rend l'utilisation de cette implémentation presque impossible : on peut avoir Δt et ε arbitrairement faibles donc il n'est pas possible techniquement d'augmenter N_τ en accord —et c'est absurde étant donné qu'on souhaite avoir une complexité indépendante de ε .

3.2 Implémentation avec la librairie chebfun

Puisque l'approche par discrétisation n'a pas fonctionné, on décide d'utiliser la bibliothèque chebfun [7] sur MATLAB. Cette bibliothèque utilise des interpolations par morceaux avec des polynômes de Chebychev pour effectuer des opérations quasi-exactes sur les fonctions. Son utilisation est simple et elle est très utile pour ce qu'on

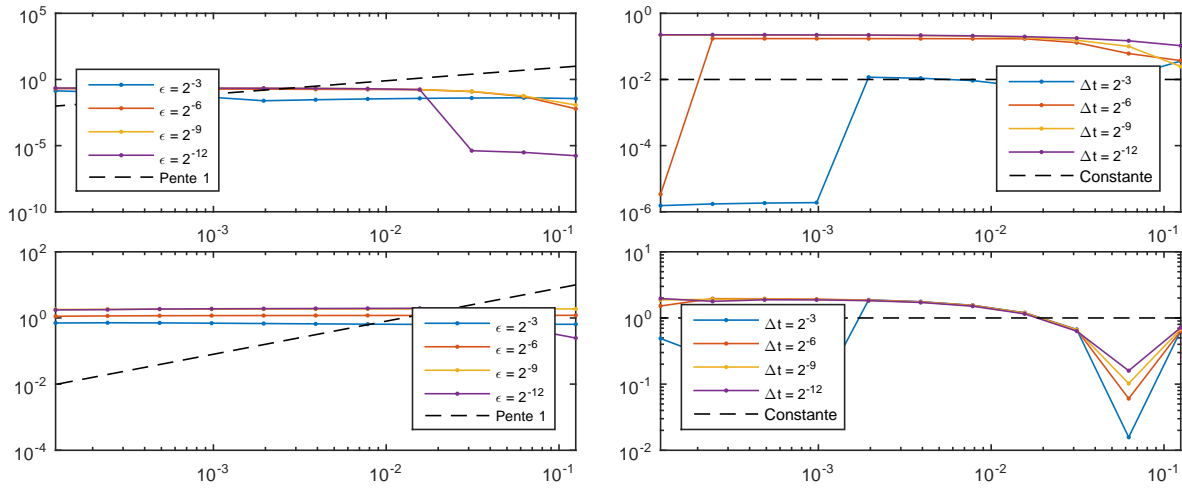


FIGURE 3.3 – Erreur relative en norme absolue sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ϵ (à droite) pour le cas (1.3) avec un schéma par différences finies.

veut faire, mais on ne contrôle pas la complexité des opérations qui peut devenir très élevée dans le cas de fonctions raides.

Cette fois-ci le calcul du produit de convolution se fait simplement avec la commande `conv`, ce qui rend l'implémentation aisée.

3.2.1 Erreur globale sur l'équation à variété centrale

De même qu'avec les différences finies, on calcule l'erreur pour diverses valeurs de Δt et de ϵ sur le cas particulier (1.3).

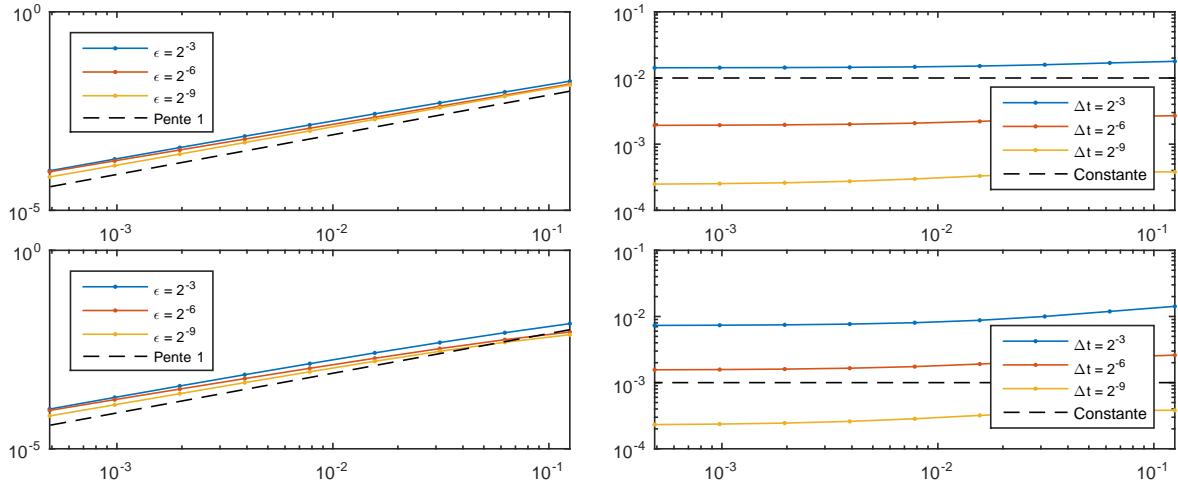


FIGURE 3.4 – Erreur relative en norme infinie sur x^ϵ (en haut) et sur z^ϵ (en bas) en fonction de Δt (à gauche) et de ϵ (à droite) pour le cas (1.3) avec le schéma utilisant `chebfun`.

On voit que les résultats obtenus (en figure 3.4) sont ceux escomptés, à l'exception d'une légère perte d'ordre. On observe bien une convergence d'ordre 1 avec une erreur

indépendante de ε comme assuré par le théorème 8. Observons le comportement sur un autre exemple, en choisissant $n = 2, m = 1$,

$$\begin{cases} f(x, z) = (1 - z) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x, \\ g(x, z) = x_1^2 x_2^2 \end{cases} \quad \text{et} \quad \begin{cases} x_0 = \begin{pmatrix} 0, 1 \\ 0, 7 \end{pmatrix}, \\ z_0 = 0,05. \end{cases} \quad (3.5)$$

On trace les mêmes courbes d'erreur en figure 3.4. Le schéma est uniformément pré-

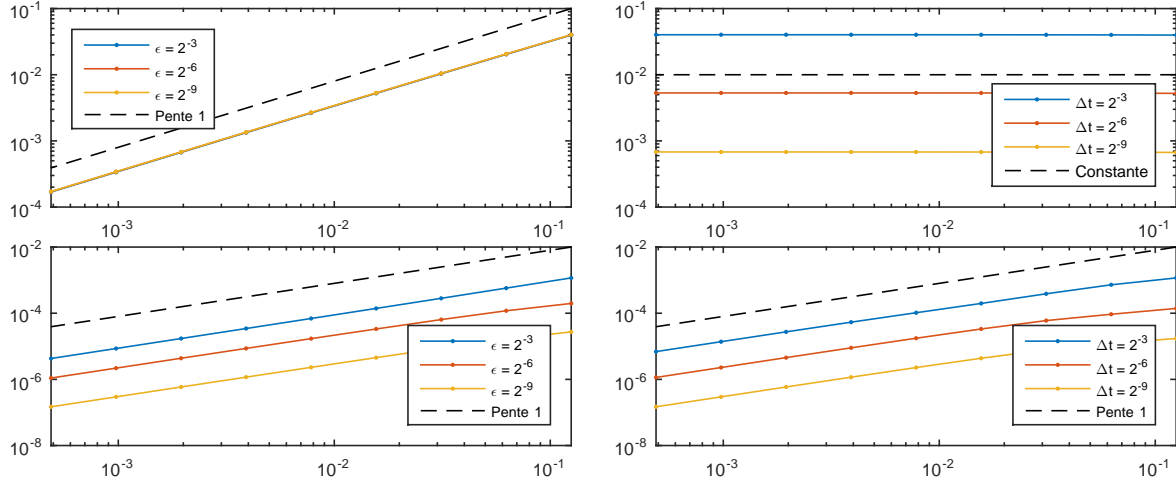


FIGURE 3.5 – Erreur absolue en norme infinie sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (3.5) avec le schéma utilisant chebfun.

cis d'ordre 1 pour le problème à variété centrale avec cette implémentation. On trace l'erreur absolue car la fonction z^ε s'annule en certains points et donc l'erreur relative est mal calculée, comme on le voit sur la figure 3.6. Dans cet exemple, l'erreur relative,

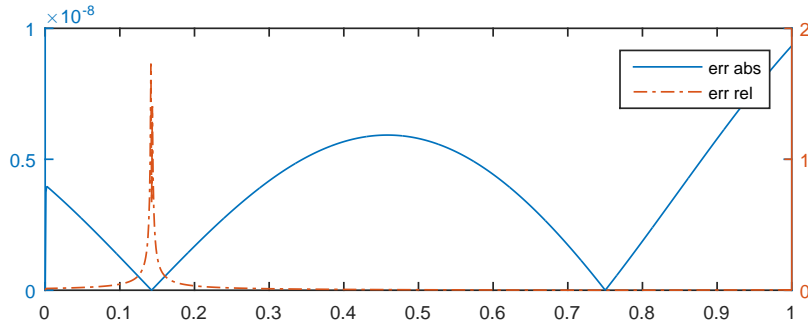


FIGURE 3.6 – Erreurs absolue (trait plein) et relative (pointillés) et leur échelle respectives à gauche et à droite en fonction du temps après calcul avec l'implémentation chebfun avec $\Delta t = 2^{-10}$ et $\varepsilon = 2^{-14}$.

localisée sur un zéro, est d'ordre 1 alors que l'erreur absolue est d'ordre 10^{-8} . On voit néanmoins sur la figure 3.5 que l'erreur absolue sur z^ε évolue en $\mathcal{O}(\varepsilon \Delta t)$, ce qui est cohérent avec le fait que z^ε converge vers une fonction d'ordre ε en temps long.

Observons l'efficacité de cette approche utilisant chebfun sur les fonctions de τ directement, pour vérifier la robustesse de notre implémentation.

3.2.2 Erreur locale sur l'équation de transport

On suppose que dans le premier cas test, les solutions obtenues avec chebfun sont exactes et on cherche à vérifier l'erreur locale en $\mathcal{O}(\Delta t^2)$.

On calcule de façon symbolique les fonctions $\tilde{X}^\varepsilon(\Delta t, \cdot), \tilde{Z}^\varepsilon(\Delta t, \cdot)$ obtenues après un seul pas de temps et on les compare à $X^\varepsilon(\Delta t, \cdot), Z^\varepsilon(\Delta t, \cdot)$ calculées avec notre schéma en utilisant un pas de temps $\Delta t_{qe} \ll \Delta t$. Si $X^\varepsilon(\Delta t, \cdot), Z^\varepsilon(\Delta t, \cdot)$ sont effectivement quasi-exactes, alors l'erreur sera d'ordre $\mathcal{O}(\Delta t^2)$.

Pour le calcul symbolique, on passe par la transformée de Laplace puis la transformée inverse. On pourrait effectuer une intégration directe, mais cela permet d'introduire rapidement une nouvelle approche, qu'on pourra développer pour une autre implémentation après ce stage. En effet, puisque $\mathcal{L}\{\partial_\tau h\}(\xi) = \xi \mathcal{L}\{h\}(\xi) - h(0)$, on a

$$\mathcal{L}\{X_1^\varepsilon\}(\xi) + \frac{1}{\mu} (\xi \mathcal{L}\{X_1^\varepsilon\}(\xi) - X_B(\Delta t)) = \mathcal{L}\{\Delta t f(X_0^\varepsilon, Z_0^\varepsilon) + X_0^\varepsilon\}(\xi)$$

d'où

$$\mathcal{L}\{X_1^\varepsilon\}(\xi) = \frac{1}{\xi + \mu} (X_B(\Delta t) + \mu \mathcal{L}\{\Delta t f(X_0^\varepsilon, Z_0^\varepsilon) + X_0^\varepsilon\}(\xi)) \quad (3.6)$$

et on peut appliquer une transformée de Laplace inverse pour obtenir X_1^ε . De la même manière on a

$$\mathcal{L}\{Z_1^\varepsilon\}(\xi) = \frac{1}{\xi + \mu + 1} (Z_B(\Delta t) + \mu \mathcal{L}\{\Delta t g(X_0^\varepsilon, Z_0^\varepsilon) + Z_0^\varepsilon\}(\xi)). \quad (3.7)$$

Avec cette approche, on est absolument certain que la solution numérique après un pas de temps est exacte. On effectue l'étude pour quelques valeurs (Δt_i) , la solution quasi-exacte étant calculé avec un pas de temps $\Delta t_{qe} = 2^{-10} \min(\Delta t_i)$.

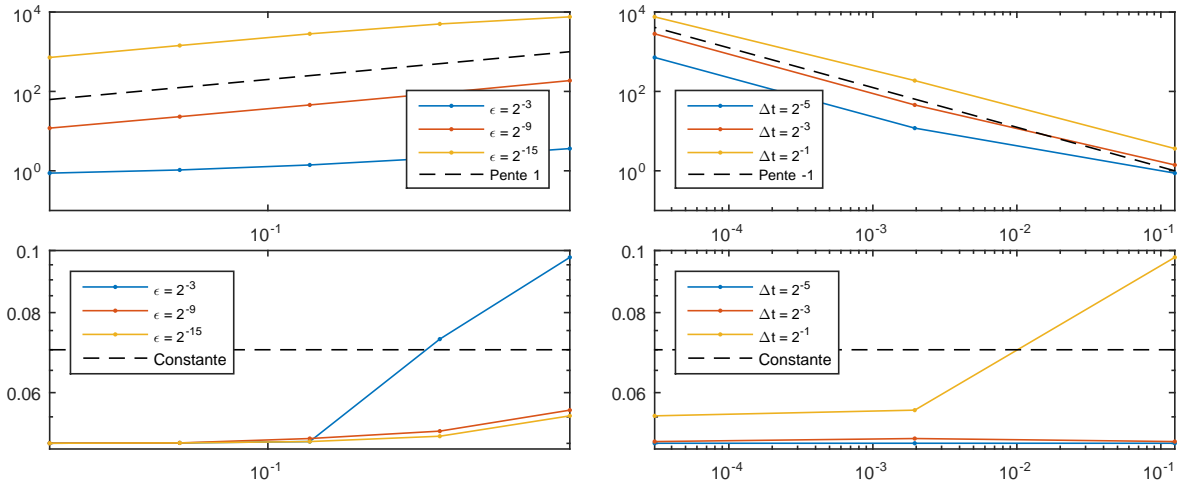


FIGURE 3.7 – Différence en norme L_τ^∞ entre la solution « quasi-exacte » calculée avec chebfun et la solution calculée symboliquement avec un seul pas de temps pour $X^\varepsilon(\Delta t, \cdot)$ (en haut) et $Z^\varepsilon(\Delta t, \cdot)$ (en bas) en fonction de Δt (à gauche) et de ε (à droite).

On voit que le résultat n'est pas du tout celui auquel on s'attendait : on n'observe pas vraiment de convergence, l'erreur est toujours très grande, avec un plateau pour ε

grand et Δt petit. En outre, même pour $\Delta t, \varepsilon$ grands, au lieu d'avoir une convergence en $\mathcal{O}(\Delta t^2)$ on observe une dépendance en $\mathcal{O}(\Delta t/\varepsilon)$.

Malheureusement, les temps de calcul sont très long et donc on ne peut pas effectuer les calculs avec Δt_{qe} plus faible. Une observation plus précise des fonctions de τ calculées nous montre qu'elles deviennent de plus en plus raides en $\tau = 0$ au fil des itérations. On ne sait pas si cela est dû à l'implémentation ou au schéma directement. En effet, on n'a pas de relation pour évaluer $\|\partial_\tau U_n^\varepsilon\|$ en fonction de n , donc le schéma pourrait causer des fonctions difficiles à représenter numériquement, ce qui pourrait accentuer ce type d'erreurs.

Cette interprétation n'explique cependant pas pourquoi l'erreur dépend de ε . Probablement que le problème est accentué lorsque μ est petit.

Étant donné que le schéma converge quand même pour la solution du système d'origine (1.1), on peut se pencher sur les coûts et les conséquences liées à son implémentation dans un cadre « réel ».

3.2.3 Coût de l'implémentation

Notre implémentation n'est pas forcément très pratique, puisque la librairie `chebfun` est puissante mais coûteuse (en termes d'opérations), et elle n'est disponible que pour `MATLAB`. On souhaite néanmoins en évaluer le coût opérationnel en fonction de $\Delta t, \varepsilon$. Notre méthode est la suivante :

- On applique le schéma avec Δt et ε fixés et on mesure le temps pris par les calculs.
- On mesure les erreurs absolue et relative sur $u = (x, z)$ grâce à une solution quasi-exacte calculée avec le schéma Radau IIA.
- On calcule une solution avec le schéma Radau en utilisant des tolérances absolue et relative de l'ordre des erreurs obtenues par le schéma et on mesure le temps de calcul.

Les calculs sont effectués sur un processeur Intel Core i5-6500U @ 2.30 GHz, à 2 cœurs. À notre connaissance, un seul processus était utilisé, même pour les opérations de `chebfun`. Cette procédure est répétée pour différentes valeurs de $\Delta t, \varepsilon$ pour obtenir des courbes similaires à celles d'erreur en figure 3.8.

La première remarque qu'on peut faire est que le temps de calcul de la solution avec Radau IIA est toujours au moins 2 ordres de grandeur en dessous de celui pris pour notre schéma. On voit aussi que le temps de calcul avec Radau est presque indépendant de la tolérance (proportionnellement liée à Δt) et de ε . On voit que le temps de calcul augmente un peu à mesure que ε diminue pour le schéma Radau et pour l'implémentation avec `chebfun`, mais de façon très faible.

3.3 Pistes de recherche

Ce stage se poursuivant en thèse, il nous reste quelques idées à étudier plus profondément qu'on n'a pas encore pu développer.

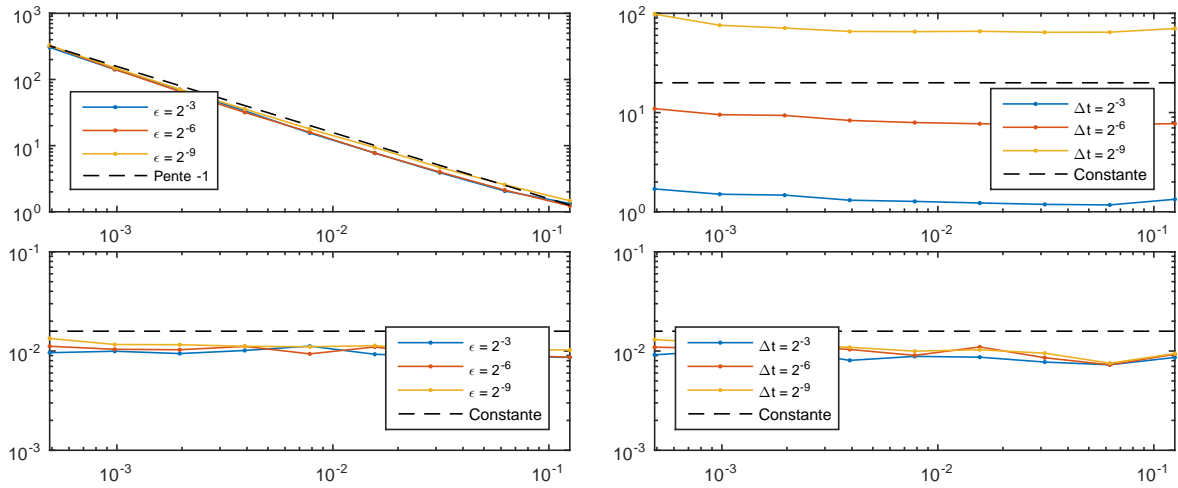


FIGURE 3.8 – Temps de calcul en secondes de la solution avec chebfun (en haut) et avec Radau IIA (en bas) en fonction de Δt (à gauche) et de ε (à droite).

3.3.1 Discrétisation du domaine de Laplace

On a vu rapidement qu'on pouvait utiliser la transformée de Laplace pour résoudre le schéma numérique. Il y a quelques soucis liés à cette approche :

1. Le domaine de Laplace est de dimension 2, ce qui rend sa discrétisation plus coûteuse qu'en 1D,
2. La bibliographie autour de la transformée de Laplace numérique et son inverse est assez peu référencée,
3. Les schémas impliquant cette transformée sont très rares, ce qui peut rendre l'approche difficile à accepter par la communauté,
4. On a besoin d'un calcul direct et inverse rapide, avec une discrétisation de $[0, T_k/\varepsilon]$ adaptée pour pouvoir calculer $f(X_n, Z_n)$ et $g(X_n, Z_n)$ à chaque itération. Cela amène des questions :
 - Comment définir « adaptée » ?
 - Comment trouver une telle discrétisation de façon dynamique ?

On a vu qu'une discrétisation fixe n'était pas adaptée au problème avec l'approche par différences finies.

On voit ainsi que cette direction apporte de nombreuses problématiques à creuser, ce qui la rend à la fois attractive et difficile à aborder.

L'argument principal en faveur de cette approche est que la transformée de Laplace peut être vue comme l'équivalent de la transformée de Fourier du cas périodique dont on s'inspire. En effet, la transformée de Laplace permettrait d'extraire facilement les coefficients d'une série $\sum_{k \geq 0} u_k e^{-k\tau}$ pour $\tau \in \mathbb{R}_+$ de la même manière que la transformée de Fourier avec les séries $\sum_{k \geq 0} u_k e^{-ik\theta}$ pour $\theta \in \mathbb{R}/P\mathbb{Z}$. Or on a vu qu'on pouvait essentiellement se ramener à des séries exponentielles pour résoudre le système (1.1), ce qui permettrait peut-être de fabriquer des méthodes « spectrales » (il en existe pour le cas hautement oscillant).

3.3.2 Schéma local à coût plus faible

Face à ce problème, on essaie d'appliquer un autre schéma, qui ne calcule pas $X^\varepsilon, Z^\varepsilon$ directement, mais qui peut peut-être servir pour $x^\varepsilon, z^\varepsilon$. L'idée est simple, même si l'étude théorique n'a pas été effectuée : il s'agit d'appliquer seulement la première étape du schéma, en boucle. Ainsi on pose $X_n^\varepsilon(\tau) = x_n^\varepsilon + \varepsilon \int_0^\tau [f(x_n^\varepsilon, e^{-\sigma} z_n^\varepsilon) - f(x_n^\varepsilon, 0)] d\sigma$ et $Z_n^\varepsilon(\tau) = e^{-\tau} z_n^\varepsilon + \varepsilon \int_0^\tau e^{\sigma-\tau} [g(x_n^\varepsilon, e^{-\sigma} z_n^\varepsilon) - \partial_z g(x_n^\varepsilon, 0) \cdot e^{-\sigma} z_n^\varepsilon] d\sigma$ et on peut calculer

$$x_{n+1}^\varepsilon = e^{-\mu\Delta t/\varepsilon} b_{\Delta t}^{(x)}(x_n^\varepsilon, z_n^\varepsilon) + \mu \int_0^{\Delta t/\varepsilon} e^{-\mu(\Delta t/\varepsilon - \sigma)} (\Delta t f(X_n^\varepsilon(\sigma), Z_n^\varepsilon(\sigma)) + X_n^\varepsilon(\sigma)) d\sigma$$

$$z_{n+1}^\varepsilon = e^{-(\mu+1)\Delta t/\varepsilon} b_{\Delta t}^{(z)}(x_n^\varepsilon, z_n^\varepsilon) + \mu \int_0^{\Delta t/\varepsilon} e^{-(\mu+1)(\Delta t/\varepsilon - \sigma)} (\Delta t g(X_n^\varepsilon(\sigma), Z_n^\varepsilon(\sigma)) + Z_n^\varepsilon(\sigma)) d\sigma$$

où $b_{\Delta t}^{(x)}(x_n^\varepsilon, z_n^\varepsilon)$ et $b_{\Delta t}^{(z)}(x_n^\varepsilon, z_n^\varepsilon)$ donnent la valeur des conditions au bord calculées comme dans (2.10) en remplaçant x_0, z_0 par $x_n^\varepsilon, z_n^\varepsilon$, ces conditions au bord étant évaluées en Δt .

En cas de convergence, cette approche permet une implémentation algébrique simple, au coût indépendant de ε . En effet, ces nouvelles fonctions $X_n^\varepsilon, Z_n^\varepsilon, f(X_n^\varepsilon, Z_n^\varepsilon), g(X_n^\varepsilon, Z_n^\varepsilon)$ sont naturellement des séries exponentielles d'après leur définition, donc le seul cas particulier à considérer pour l'approche algébrique est $\mu \in \mathbb{N}$, ce qui n'est pas un frein.

Le problème opérationnel de ce schéma est qu'il implique deux intégrations par pas de temps : une pour calculer $X_n^\varepsilon, Z_n^\varepsilon$ et une pour $x_{n+1}^\varepsilon, z_{n+1}^\varepsilon$. Ces intégrations étant simples, il semble plutôt efficace quand même. Cependant on aurait souhaité limiter le nombre d'intégrations étant donné que les schémas dans le cas hautement oscillant n'en effectuant qu'une par pas de temps.

Néanmoins, ce schéma se présente comme assez simple à étudier théoriquement et à implémenter tout en pouvant fournir des résultats convaincants. On a en effet pu l'implémenter très facilement en utilisant chebfun et on a observé une précision uniforme sur nos deux cas particuliers. Cela en fait une piste de recherche prioritaire.

CONCLUSION

On a présenté le modèle et les problèmes qu'il engendrait d'un point de vue numérique. On a ensuite effectué un développement double-échelle pour séparer la dynamique rapide $e^{-t/\varepsilon}$ de la dynamique lente. Après cette séparation, on a exhibé des conditions nécessaires à la régularité du nouveau problème, et on a montré que sous ces conditions, le problème était bien posé. Une fois cela fait, on a trouvé des conditions initiales et au bord qui permettent d'assurer que X, Z et leurs dérivées première et seconde en t sont uniformément bornées. À partir de ces données, on a construit un schéma numérique uniformément précis qu'on a implémenté et dont on a évalué les performances et les limites en le comparant à d'autres approches moins coûteuses mais moins (voire pas) robustes. Les résultats étant insatisfaisant en termes de performance et d'exportabilité, on a rapidement présenté deux de nos pistes de recherche qui seront poursuivies en thèse.

Il est très facile de se perdre dans les différentes grandeurs en jeu dans ce problème : on a affaire à des séries en puissance de ε dont les coefficients sont des séries exponentielles en τ , et on doit faire varier tout cela en fonction de t . On peut aussi facilement confondre régularité et caractère uniformément borné des dérivées, ce qui pose problème étant donné l'importance des conditions de compatibilité.

Comme on l'a évoqué rapidement, il peut être intéressant de travailler avec des séries exponentielles, mais il n'est pas évident de construire une implémentation se basant uniquement là-dessus, et il faut se rendre compte que les coefficients en jeu sont plutôt instables —d'où notre normalisation en section 3.1.1.

Tous ces « pièges » sont autant de raisons qui ont rendu le début de ce stage un peu laborieux. Au bout d'un moment, la formalisation s'est mise en place et donc on a pu progresser relativement rapidement vers la fin. Les notions restent cependant complexes et nécessitent de prendre encore un peu de recul.

Le mélange d'algèbre, analyse, calcul formel et implémentations numériques de ce sujet est ce qui le rend tout à fait attirant pour une poursuite en thèse. Il est en outre apparent qu'il y a encore beaucoup de pistes à creuser en matière d'approche et d'implémentation et que donc ce sujet « préliminaire » à l'étude des EDP est déjà très intéressant en soi et motive à l'approfondir.

ANNEXE A

PRÉSENTATION DES CAS TESTS

A.1 Équations scalaires couplées

On rappelle le problème (1.3) :

Pour ε donné, trouver $x \in C^\infty([0, T]; \mathbb{R})$ et $z \in C^\infty([0, T])$ telles que

$$\begin{cases} \dot{x} = -x^3(z - z^3/3), \\ \dot{z} = -\frac{1}{\varepsilon}z + x(1 - z^2/2) \end{cases} \quad \text{et} \quad \begin{cases} x(0) = x_0 = 0,8 \\ z(0) = z_0 = 0,05. \end{cases} \quad (\text{A.1})$$

Traçons la solution de ce système sur $[0; 0,7]$ pour diverses valeurs de ε .

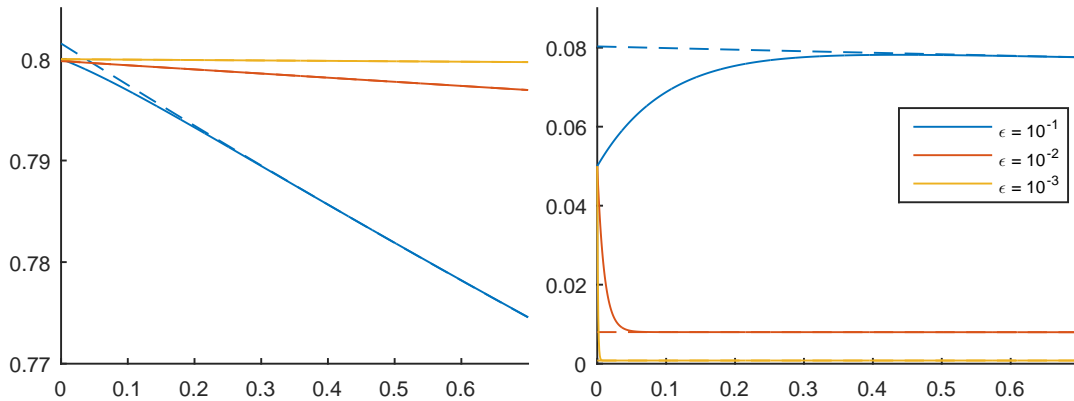


FIGURE A.1 – Solution x (à gauche) et z (à droite) du problème (A.1) sur $[0; 0,7]$, avec en pointillé l'approximation de la variété associée.

Il est apparent sur la figure qu'on a en temps long $\dot{x} = \mathcal{O}(\varepsilon)$. On voit aussi très clairement $z_\infty = \mathcal{O}(\varepsilon)$.

A.2 Oscillations lentes de x

On rappelle le problème (3.5) :

Pour ε donné, trouver $x \in C^\infty([0, T]; \mathbb{R}^2)$ et $z \in C^\infty([0, T]; \mathbb{R})$ telles que

$$\begin{cases} \dot{x} = (1 - z) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x, \\ \dot{z} = \frac{1}{\varepsilon}z + x_1^2 x_2^2 \end{cases} \quad \text{et} \quad \begin{cases} x(0) = x_0 = \begin{pmatrix} 0,1 \\ 0,7 \end{pmatrix}, \\ z(0) = z_0 = 0,05. \end{cases} \quad (\text{A.2})$$

On ne trace pas la solution en fonction du temps, mais plutôt son parcours dans \mathbb{R}^3 .

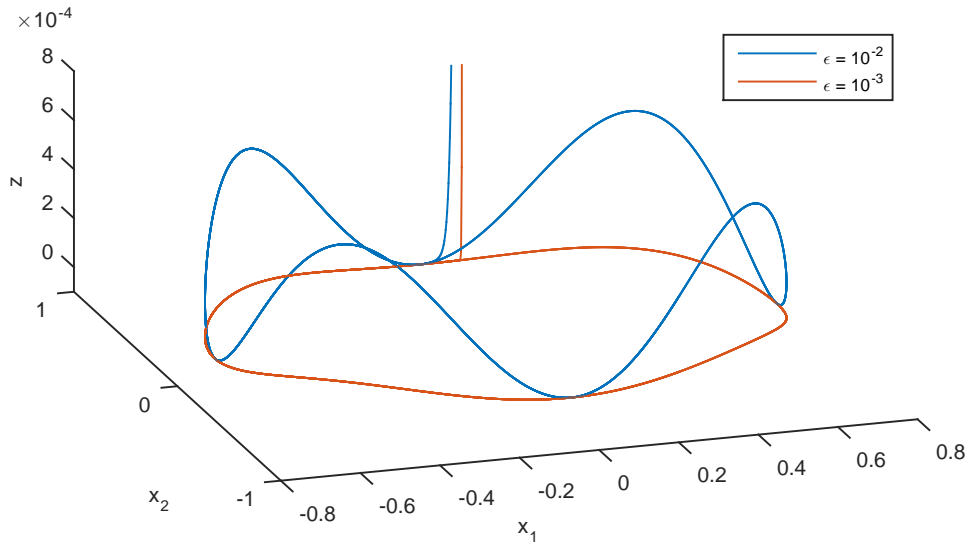


FIGURE A.2 – Parcours de (x_1, x_2, z) dans \mathbb{R}^3 où x, z sont solutions du problème (A.2) sur $[0; 10]$.

Le système semble converger rapidement vers un système périodique où x_1, x_2 et z oscillent. On trace z pour vérifier que la période ne dépend qualitativement pas de ε .

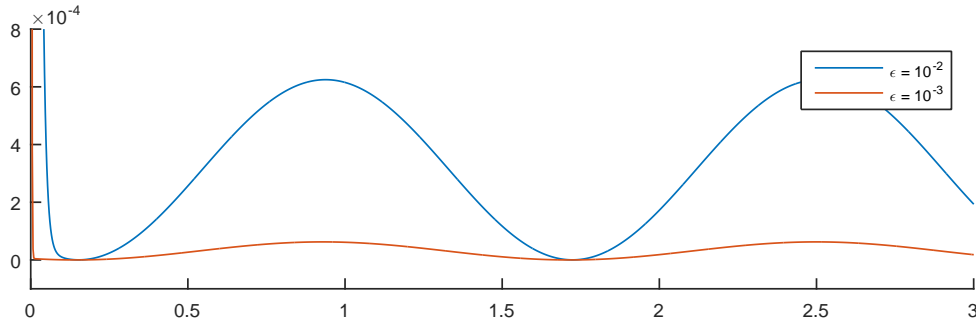


FIGURE A.3 – Solution z du problème (A.2) sur $[0; 3]$ pour $\varepsilon = 10^{-2}, 10^{-3}$.

En effet sur la variété on a $\dot{x} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x + \mathcal{O}(\varepsilon)$, ce qui cause l'oscillation de x_1, x_2 . Le terme dominant pour déterminer la période est donc forcément indépendant de ε .

ANNEXE B

PREUVE DU LEMME 2 (P. 17)

On rappelle le problème

Trouver $(\bar{T}, u) \in]0, T] \times C^0([0, \bar{T}] \times \mathbb{R}_+)$ tel que pour tout $(t, \tau) \in [0, \bar{T}] \times \mathbb{R}_+$ on ait

$$\partial_t u + \partial_\tau u = F(t, \tau, u) \quad (\text{B.1})$$

et le lemme qui assure qu'il est bien posé

Lemme. Si $u_B \in C_b^0([0, T])$, $u_I \in C_b^0(\mathbb{R}_+)$ avec $u_B(t=0) = u_I(\tau=0) =: u_0$, si F est localement lipschitzienne par rapport à u , continue de $[0, T] \times \mathbb{R}_+ \times E$ dans E , bornée par rapport à τ , alors le problème (B.1) est bien posé au sens suivant :

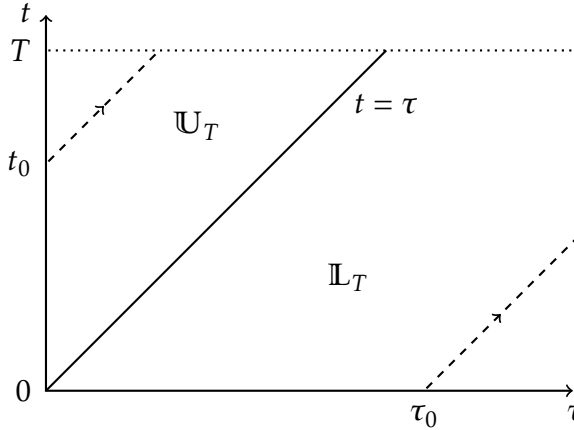
Pour tout facteur $\kappa > 1$, il existe un temps $T_\kappa \in]0, T]$ tel que le problème (B.1) a une unique solution u dans $C^0([0, T_\kappa] \times \mathbb{R}^+)$, qui respecte l'inégalité

$$\forall t \in [0, T_\kappa], \quad \|u(t, \cdot)\|_{L_t^\infty} \leq \kappa \max\{\|u_I\|_{L_\tau^\infty}, \|u_B\|_{L_t^\infty}\}. \quad (\text{B.2})$$

Si en outre F respecte $\forall (t, \tau, u) \in [0, T] \times \mathbb{R}_+ \times E, |F(t, \tau, u)| \leq C_F|u| + D_F$ pour des constantes $C_F, D_F \geq 0$, alors il existe une unique solution u dans $C^0([0, T] \times \mathbb{R}_+)$ qui vérifie

$$\forall t \in [0, T], \|u(t, \cdot)\|_{L_t^\infty} \leq \left(\max\{\|u_I\|_{L_\tau^\infty}, \|u_B\|_{L_t^\infty}\} + tD_F\right)e^{tC_F}. \quad (\text{B.3})$$

Démonstration. Intéressons nous aux droites caractéristiques du problème.



Sous la diagonale (t, t) , $t \in [0, T]$, la donnée initiale est en $(0, \tau_0)$, $\tau_0 \in \mathbb{R}_+$ et la droite est décrite par $(t, t + \tau_0)$. Au-dessus de la diagonale, elle commence en $(t_0, 0)$ pour $t_0 \in [0, T]$ et est décrite par $(t + t_0, t)$, $t \in [0, T - t_0]$.

Regardons d'abord le comportement de la solution sous la diagonale (t, t) . À partir d'une solution u , on pose $\phi(t, \tau) = u(t, t + \tau)$ la solution sur la caractéristique démarrant en $(0, \tau)$ qui respecte l'EDO

$$\partial_t \phi(t, \tau) = F(t, t + \tau, \phi) \quad \text{et} \quad \phi(0, \tau) = u_I(\tau) \quad (\text{B.4})$$

On sait que ce problème de Cauchy admet une solution unique locale sur $[0, t_m]$ (où $t_m > 0$ peut dépendre de τ) par caractère localement lipschitzien de F .

On peut mettre le problème (B.4) sous forme intégrale

$$\phi(t, \tau) = \phi(0, \tau) + \int_0^t F(s, \tau + s, \phi(s, \tau)) ds$$

pour $t \in [0, t_m]$. On cherche à déterminer un minorant de t_m qui soit indépendant de τ . Par inégalité triangulaire on obtient

$$|\phi(t, \tau)| \leq |\phi(0, \tau)| + \int_0^t |F(s, \tau + s, \phi(s, \tau))| ds. \quad (\text{B.5})$$

Soit $\kappa > 1$, on pose $R = \max\{\|u_I\|_{L_t^\infty}, \|u_B\|_{L_t^\infty}\}$ et

$$M_\kappa = \sup\{|F(t, \tau, u)| \mid (t, \tau, u) \in [0, T] \times \mathbb{R}^+ \times E, |u| \leq \kappa R\}.$$

Par continuité et caractère borné de F , $M_\kappa < \infty$. Par définition, $|\phi(0, \tau)| \leq R$ et donc tant que $|\phi(t, \tau)| \leq \kappa R$ on a $|\phi(t, \tau)| \leq R + tM_\kappa$ ce qui est assuré tant que $R + tM_\kappa \leq \kappa R$. Ainsi $\phi(t, \tau)$ est bien définie pour $t \leq T_\kappa$ avec

$$T_\kappa = \min\left\{T, \frac{(\kappa - 1)R}{M_\kappa}\right\} \quad (\text{B.6})$$

qui vérifie toujours $T_\kappa > 0$. Sur $\mathbb{L}_{T_\kappa} := \{(t, \tau) \in [0, T_\kappa] \times \mathbb{R}_+, t \leq \tau\}$, on a $u(t, \tau) = \phi(t, \tau - t)$ et donc on peut définir une solution $u_{\mathbb{L}}$ sur le domaine réduit \mathbb{L}_{T_κ} .

On remarque que $\phi(0, \tau)$ dépend continument de τ par continuité de u_I , et F est continue par rapport à t dans l'EDO sur $\phi(\cdot, \tau)$ donc la continuité se propage et $u_{\mathbb{L}}$ est continue sur \mathbb{L}_{T_κ} .

Au dessus de la diagonale, le problème est presque le même, on a $\partial_t \varphi(t, t_0) = F(t + t_0, t, \varphi)$, $\varphi(0, t_0) = u_B(t_0)$ pour $t_0 \in [0, T_\kappa]$, d'où $|\varphi(t, t_0)| \leq |\varphi(0, t_0)| + \int_0^t |F(s + t_0, s, \varphi(s, t_0))| ds$. La définition de T_κ nous permet d'assurer de façon naturelle que $\varphi(\cdot, t_0)$ est bien définie sur $[0, T_\kappa - t_0]$. On peut alors définir une solution $u_{\mathbb{U}}$ sur le domaine réduit $\mathbb{U}_{T_\kappa} := \{(t, \tau) \in [0, T_\kappa] \times \mathbb{R}_+, t \geq \tau\}$ en posant $u_{\mathbb{U}}(t, \tau) = \varphi(\tau, t - \tau)$. Comme précédemment, la continuité de u_I et de F se propage et donc $u_{\mathbb{U}} \in C^0(\mathbb{U}_{T_\kappa})$.

On peut maintenant procéder au « recollement » des solutions pour obtenir u sur $[0, T_\kappa] \times \mathbb{R}_+$. On pose

$$u(t, \tau) = \begin{cases} u_{\mathbb{L}}(t, \tau) & \text{si } t \leq \tau \\ u_{\mathbb{U}}(t, \tau) & \text{si } t > \tau \end{cases}$$

pour avoir u solution du problème dans $L^\infty([0, T_\kappa] \times \mathbb{R}_+)$ et qui vérifie nécessairement l'inégalité (B.2). La continuité est assurée sur \mathbb{L}_{T_κ} et sur $\mathbb{U}_{T_\kappa} \setminus \{(t, t), t \in [0, T_\kappa]\}$, la seule difficulté est d'assurer la continuité de part et d'autre de la diagonale. Or on remarque que pour tout t dans $[0, T_\kappa]$, on a l'égalité $\phi(t, 0) = \varphi(t, 0)$ puisque $\phi(\cdot, 0)$ et $\varphi(\cdot, 0)$ vérifient la même EDO avec la même condition initiale car $\phi(0, 0) = u_I(0) = u_B(0) = \varphi(0, 0)$, et la solution est unique sur $[0, T_\kappa]$. Ainsi $u_{\mathbb{L}}(t, t) = u_{\mathbb{U}}(t, t)$ et donc u est continue de part et d'autre de la diagonale. Finalement, $u \in C^0([0, T_\kappa] \times \mathbb{R}_+)$.

Si on est dans le cas de F bornée par une fonction affine, on applique le théorème de Grönwall à l'inégalité (B.5) pour obtenir la relation (B.3), et le résultat de régularité est maintenu. □

ANNEXE C

ERREURS DES SCHÉMAS USUELS

C.1 Schémas à séparation des vitesses

Schéma implicite-explicite

On rappelle le schéma implicite-explicite

$$\frac{1}{\Delta t}(x_{n+1} - x_n) = f(x_n, z_n),$$

$$\frac{1}{\Delta t}(z_{n+1} - z_n) = -\frac{1}{\varepsilon}z_{n+1} + g(x_n, z_n).$$

Pour étudier l'erreur locale on utilise la formule de Taylor à reste intégral

$$\begin{aligned} \ell e^{(x)} &= x_1 - x(\Delta t) = x_1 - x_0 - \Delta t f(x_0, z_0) - R_2^{(x)}(\Delta t) \\ &= \int_0^{\Delta t} (t - \Delta t) \left[\partial_x f(x, z) \cdot f(x, z) + \partial_z f(x, z) \cdot \left(-\frac{1}{\varepsilon}z + g(x, z) \right) \right] dt \end{aligned}$$

où x, z sont évaluées en t dans l'intégrande. Étant donné le comportement de la solution grâce au théorème de variété centrale (thm. 1), on peut supposer approximer $x(t) = x_0 + \mathcal{O}(\varepsilon)$ et $z(t) = z_0 e^{-\alpha t/\varepsilon} + \mathcal{O}(\varepsilon)$. En se concentrant uniquement sur le terme en $e^{-\alpha t/\varepsilon}/\varepsilon$ dans l'intégrale et en majorant $\|\partial_z f(x, z)\|$ par une constante (indépendante de ε), on obtient finalement

$$\ell e^{(x)} = \mathcal{O}\left(\frac{\varepsilon}{\alpha}(e^{-\alpha \Delta t/\varepsilon} - 1 + \alpha \Delta t/\varepsilon)\right), \quad \varepsilon, \Delta t \rightarrow 0^+.$$

Ainsi avec $\Delta t \ll \varepsilon$, on a $\ell e^{(x)} = \mathcal{O}(\Delta t^2/\varepsilon)$ et avec $\varepsilon \ll \Delta t$ on a $\ell e^{(x)} = \mathcal{O}(\Delta t)$.

Sur z , on utilise directement la formule $z(\Delta t) = z_0 e^{-\alpha \Delta t/\varepsilon}$ puisque le développement de Taylor ne traduit pas le comportement de la solution. On obtient alors

$$\ell e^{(z)} = z_0 - z(\Delta t) = \left(\frac{1}{1 + \Delta t/\varepsilon} - e^{-\alpha \Delta t/\varepsilon} \right) z_0 + \mathcal{O}(\varepsilon).$$

Dans les faits on observe une convergence $\mathcal{O}(\Delta t^2/\varepsilon^2)$ lorsque $\Delta t \ll \varepsilon$, et $\mathcal{O}(\varepsilon/\Delta t)$ quand $\varepsilon \ll \Delta t$, ce qui suggère que la variété est bien approchée par le schéma (si bien que le terme $\mathcal{O}(\varepsilon)$ n'a pas d'influence) et qu'on a $\alpha = 1$ (sinon la convergence $\Delta t \ll \varepsilon$ serait $\mathcal{O}(\Delta t/\varepsilon)$).

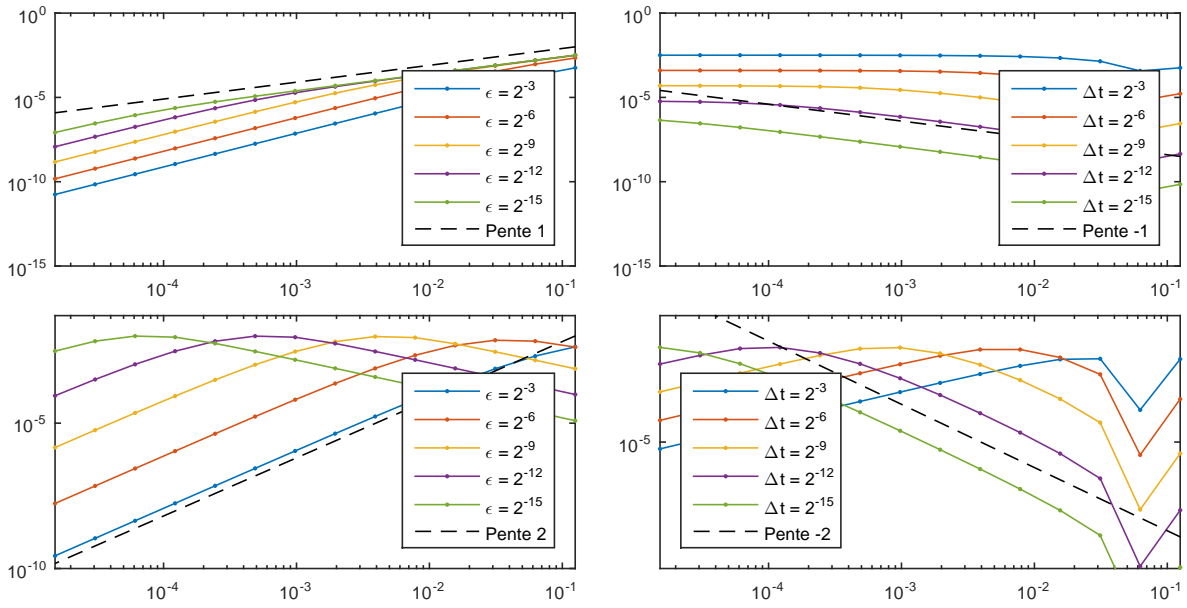


FIGURE C.1 – Erreur relative en norme absolue sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (1.3) avec un schéma implicite-explicite.

Splitting de Strang

On peut calculer l'erreur avec le splitting de Strang dans le cas f, g linéaires (pour pouvoir faire de l'algèbre d'opérateurs de façon simple). On sépare alors l'EDO $\dot{u} = -\frac{1}{\varepsilon}Ju + Fu$ en deux systèmes $\dot{u} = -\frac{1}{\varepsilon}Ju$ et $\dot{u} = Fu$. L'erreur d'opérateur due au splitting est alors

$$e^{-\frac{\Delta t}{2\varepsilon}J}e^{\Delta tF}e^{-\frac{\Delta t}{2\varepsilon}J} - e^{\Delta t(-\frac{1}{\varepsilon}J+F)} = \frac{\Delta t^3}{4!} \left(\frac{1}{\varepsilon^2} [J, F], J \right) - \frac{2}{\varepsilon} [J, F], F \Big) = \mathcal{O}\left(\frac{\Delta t^3}{\varepsilon^2}\right).$$

Dans les faits on n'observe cette erreur que pour la variable x et pas pour la variable z . En effet, sur nos exemples on a

$$\ell e^{(x)} = \mathcal{O}\left(\Delta t^3/\varepsilon^2\right), \quad \ell e^{(z)} = \mathcal{O}\left(\Delta t^2/\varepsilon\right).$$

C.2 Schéma à pas de temps adaptatif

Expliquons le tableau de Butcher du schéma à pas de temps adaptatif. De manière générale, le tableau pour une méthode explicite imbriquée est

0	0				
c_2	a_{21}	0			
c_3	a_{31}	a_{32}	0		
\vdots	\vdots		\ddots	0	
c_s	a_{s1}	a_{s2}	\cdots	$a_{s,s-1}$	0
	b_1	b_2	\cdots	b_{s-1}	b_s
	\tilde{b}_1	\tilde{b}_2	\cdots	\tilde{b}_{s-1}	\tilde{b}_s

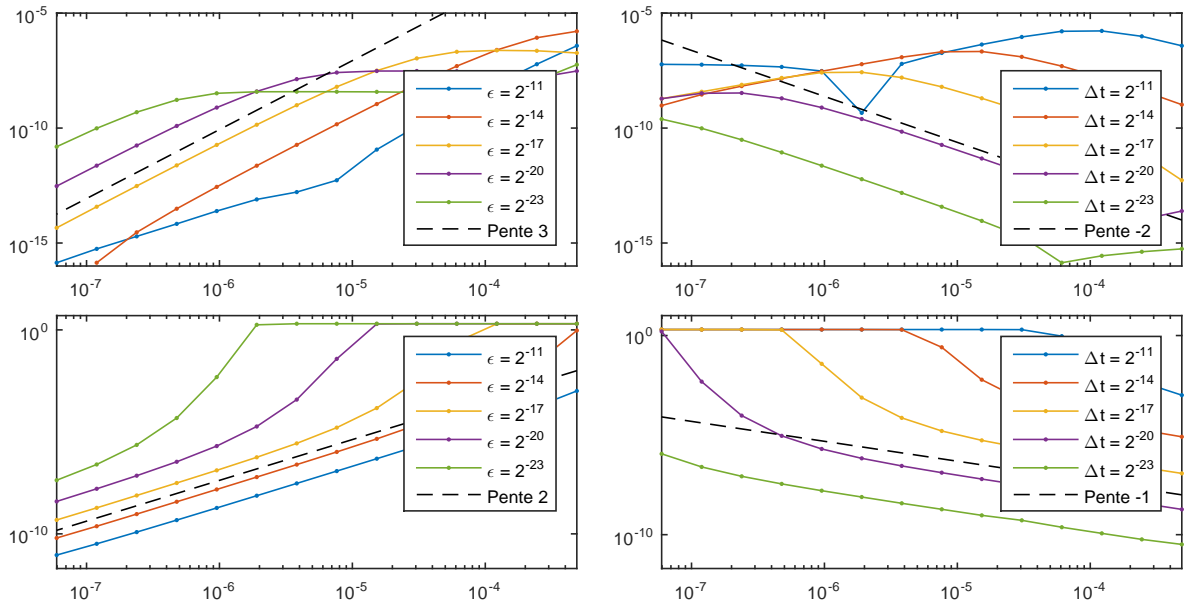


FIGURE C.2 – Erreur relative en norme absolue sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ϵ (à droite) pour le cas (1.3) avec un splitting de Strang.

On calcule alors les coefficients k_s pour une EDO $\dot{u} = f(t, u)$ avec

$$\begin{aligned} k_1 &= f(t_n, u_n), \\ k_2 &= f(t_n + c_2 \Delta t, u_n + (a_{21} \Delta t) k_1), \\ &\vdots \\ k_s &= f(t_n + c_s \Delta t, u_n + \Delta t (a_{s1} k_1 + a_{s2} k_2 + \dots + a_{s,s-1} k_{s-1})), \end{aligned}$$

puis les deux estimations de la solution $\hat{u}_{n+1}, \tilde{u}_{n+1}$ avec

$$\hat{u}_{n+1} = b_1 k_1 + b_2 k_2 + \dots + b_s k_s,$$

$$\tilde{u}_{n+1} = \tilde{b}_1 k_1 + \tilde{b}_2 k_2 + \dots + \tilde{b}_s k_s.$$

Par convention, si le schéma avec $(b_i)_{1 \leq i \leq s}$ est d'ordre p alors celui avec $(\tilde{b}_i)_{1 \leq i \leq s}$ est d'ordre $p - 1$. Si $|\hat{u}_{n+1} - \tilde{u}_{n+1}| \leq Tol$ alors on admet le pas de temps Δt et on pose $u_{n+1} = \hat{u}_{n+1}$. On estime $|\hat{u}_{n+1} - \tilde{u}_{n+1}| \leq |\hat{u}_{n+1} - u(t_{n+1})| + |\tilde{u}_{n+1} - u(t_{n+1})| = \mathcal{O}(\Delta t^{p+1}) + \mathcal{O}(\Delta t^p) = \mathcal{O}(\Delta t^p)$. On suppose alors $|\hat{u}_{n+1} - \tilde{u}_{n+1}| \simeq C \Delta t^p$ et $Tol \simeq C (\Delta t_{opt})^p$ (avec la même constante C dans les deux relations), d'où l'adaptation de pas de temps

$$\Delta t \leftarrow 0,9 \Delta t \left(\frac{Tol}{|\hat{u}_{n+1} - \tilde{u}_{n+1}|} \right)^{1/p}$$

le facteur 0,9 étant appelé *facteur de sécurité*.

Dans notre cas particulier, on a

0	0	0	0	0	0
1/3	1/3	0	0	0	0
2/3	-1/3	1	0	0	0
1	1	-1	1	0	0
1	1/8	3/8	3/8	1/8	0
	1/8	3/8	3/8	1/8	0
	1/12	1/2	1/4	0	1/6

On implémente ce schéma et on mesure le temps de calcul et l'erreur en fonction de ε pour quelques valeurs de Tol .

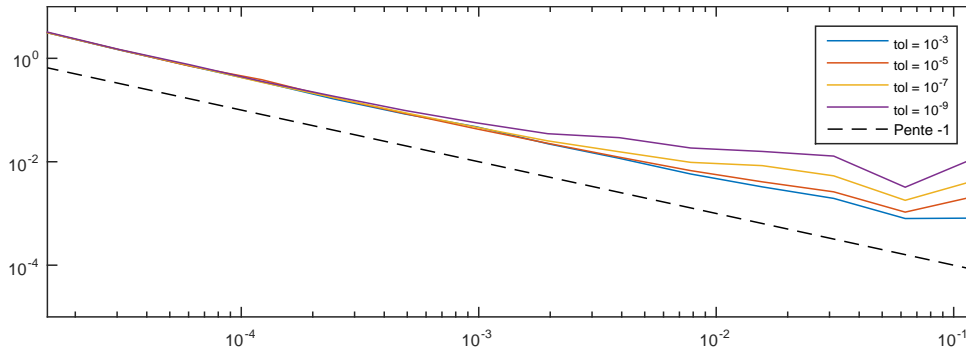


FIGURE C.3 – Temps de calcul (en s) avec notre méthode de Runge-Kutta imbriqués en fonction de ε .

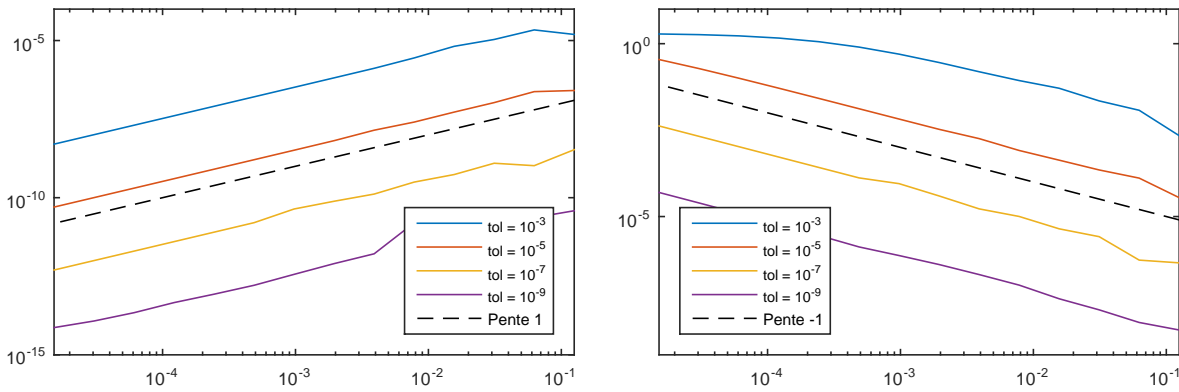


FIGURE C.4 – Erreur relative en norme absolue sur x (à gauche) et sur z (à droite) en fonction de ε pour le cas (1.3) avec une méthode de Runge-Kutta imbriqués.

On voit alors que l'erreur sur z augmente avec ε lorsque ε diminue, ce qui va à l'encontre de notre objectif. On rappelle en outre que le

C.3 Développement asymptotique

Dans [4], on fournit les développements à l'ordre 3 en ε de la variété centrale $\varepsilon h^\varepsilon(\cdot)$ et de la condition initiale réduite x_0^ε dans les cas particuliers (1.3) et (3.5). En particulier,

dans le cas (3.5), on définit la variété tronquée et la condition initiale réduite tronquée

$$\begin{aligned}\varepsilon h^{[3]}(x) &= \varepsilon x_1^2 x_2^2 - 2\varepsilon^2 x_1 x_2 (x_1^2 - x_2^2) + 2\varepsilon^3 (x_1 x_2 (x_1^2 - x_2^2 - 2) + (x_1^2 - x_2^2)^2) \\ &= \varepsilon h^\varepsilon(x) + \mathcal{O}(\varepsilon^4),\end{aligned}$$

$$\begin{aligned}x_0^{[3]} &= \left(1 - \frac{\varepsilon^2}{2} z_0^2 + \frac{\varepsilon^3}{2} x_{0,1}^2 x_{0,2}^2 z_0\right) x_0 + \left[-\varepsilon z_0 + \varepsilon^2 x_{0,1}^2 x_{0,2}^2\right. \\ &\quad \left.+ \varepsilon^3 \left(\frac{1}{6} z_0^3 - 2(1 + 2z_0) x_{0,1} x_{0,2} (x_{0,1}^2 - x_{0,2}^2)\right)\right] \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x_0 + x_0^\varepsilon + \mathcal{O}(\varepsilon^4).\end{aligned}$$

On définit la solution sur la variété approchée

$$\dot{x}_\infty^{[3]} = f(x_\infty^{[3]}, \varepsilon h^{[3]} \circ x_\infty^{[3]}), \quad x_\infty^{[3]}(0) = x_0^{[3]}.$$

On peut alors la calculer et la comparer à la solution réelle en fonction du temps et on voit que l'écart diminue avec ε . Vérifions que lors d'une simulation numérique, on obtient une erreur d'ordre $\mathcal{O}(\varepsilon^4 + \Delta t^p)$ où p est l'ordre de la méthode. L'erreur $\mathcal{O}(\varepsilon^4)$ provient du modèle, de l'approximation de x_0^ε et h^ε , tandis que l'erreur $\mathcal{O}(\Delta t^p)$ provient du calcul numérique de $x_\infty^{[3]}$.

Pour la vérification, on commence par faire une résolution quasi-exacte sur la variété et sur le système initial pour vérifier la partie $\mathcal{O}(\varepsilon^4)$, puis on fixe ε très faible et on résout le problème sur la variété approximée avec une méthode d'Euler explicite pour vérifier la partie $\mathcal{O}(\Delta t)$. Dans le cas (1.3), on a $\dot{x} = \mathcal{O}(\varepsilon)$ sur la variété, donc cette approche est impossible (c'est pourquoi on utilise le cas (3.5) pour ces tests).

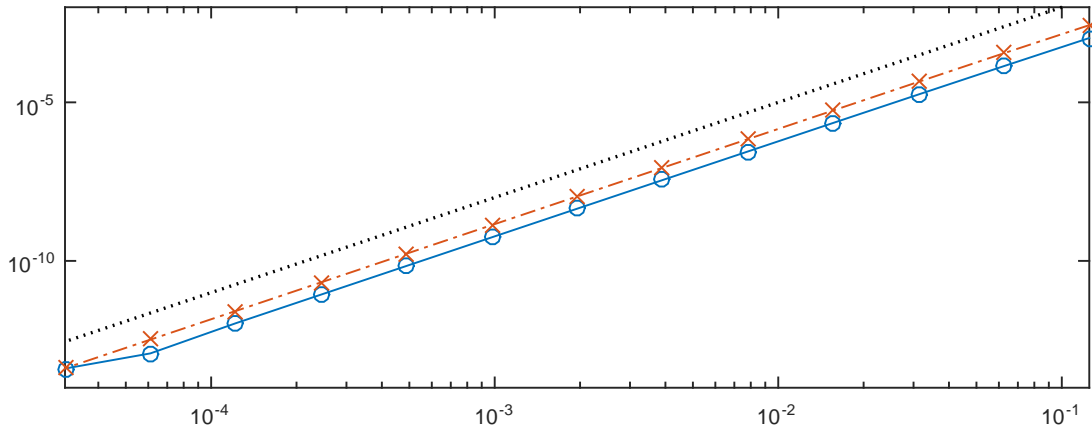


FIGURE C.5 – Différences entre x^ε et $x_\infty^{[3]}$ (ronds, trait plein bleu) et entre z^ε et $\varepsilon h^{[3]} \circ x_\infty^{[3]}$ (croix, trait pointillé rouge) en fonction de ε en $t = 1$, comparées à une pente ε^3 (pointillés noirs) dans le cas test (3.5).

On voit alors que l'erreur est en $\mathcal{O}(\varepsilon^3)$, ce qui n'est pas le résultat auquel on s'attendait. Il se peut que les formules de l'article soient erronées. En effet, en effectuant notre propre développement sur h^ε en remarquant $\partial_z g = 0$, on obtient

$$\begin{aligned}h^\varepsilon(x) &= g(x, 0) - \varepsilon \partial_x [g(x, 0) - \varepsilon \partial_x g(x, 0) \cdot f(x, 0)] \cdot f(x, \varepsilon g(x, 0)) + \mathcal{O}(\varepsilon^3) \\ &= g(x, 0) - \varepsilon \partial_x g(x, 0) \cdot f(x, 0) + \varepsilon^2 \partial_x^2 g(x, 0) \cdot (f(x, 0), f(x, 0)) \\ &\quad + \varepsilon^2 \partial_x g(x, 0) \cdot [\partial_x f(x, 0) \cdot f(x, 0) - \partial_z f(x, 0) \cdot g(x, 0)] + \mathcal{O}(\varepsilon^3)\end{aligned}$$

soit d'après nos calculs

$$\varepsilon h^{[3]}(x) = \varepsilon x_1^2 x_2^2 - 2\varepsilon^2 x_1 x_2 (x_1^2 - x_2^2) + 2\varepsilon^3 (x_1^2 x_2^2 (x_1^3 x_2 - x_1 x_2^3 - 2) + (x_2^2 - x_1^2)^2).$$

Effectuer cette correction ne change pas les résultats, donc il se peut que la formule pour $x_0^{[3]}$ soit mal reportée dans l'article. On peut néanmoins vérifier la convergence en $\mathcal{O}(\Delta t)$ avec un schéma d'Euler explicite pour calculer la solution sur la variété approchée, dans le cas ε très petit (on choisit $\varepsilon = 2^{-12}$).

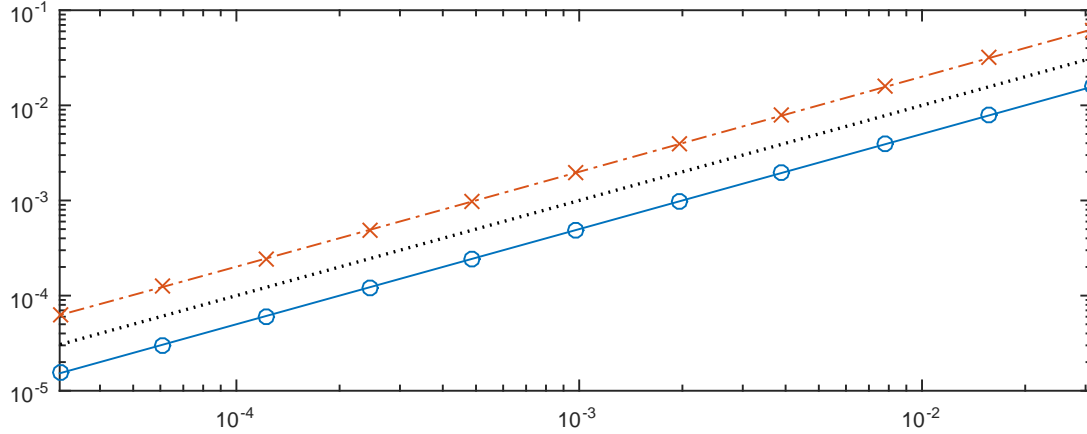


FIGURE C.6 – Différences entre x^ε et $x_\infty^{[3]}$ (ronds, trait plein bleu) et entre z^ε et $\varepsilon h^{[3]} \circ x_\infty^{[3]}$ (croix, trait pointillé rouge) en fonction de Δt en $t = 1$ lorsque $x_\infty^{[3]}(1)$ est calculée par méthode d'Euler explicite de pas de temps Δt . La droite pointillée noire est de pente 1.

On voit alors que la convergence $\mathcal{O}(\Delta t)$ est bien vérifiée.

BIBLIOGRAPHIE

- [1] Alain Bensoussan, Jacques-Louis Lions, and George Papanicolaou. *Asymptotic analysis for periodic structures*, volume 374 of *Studies in Mathematics and Its Applications*. American Mathematical Soc., 1978.
- [2] Denis Bischel. Implémentation MATLAB du schéma radau iia. <https://www.mathworks.com/matlabcentral/fileexchange/56162>. Code mis en ligne par Meysam Mahooti, téléchargé en mars 2018.
- [3] Jack Carr. *Applications of centre manifold theory*, volume 35 of *Applied Mathematical Sciences*. Springer-Verlag New York, 1982.
- [4] François Castella, Philippe Chartier, and Julie Sauzeau. A formal series approach to the center manifold theorem. *Foundations of Computational Mathematics*, pages 1–38, 2016.
- [5] Philippe Chartier, Nicolas Crouseilles, Mohammed Lemou, and Florian Méhats. Uniformly accurate numerical schemes for highly oscillatory klein–gordon and nonlinear schrödinger equations. *Numerische Mathematik*, 129(2) :211–250, 2015.
- [6] Philippe Chartier, Mohammed Lemou, Florian Méhats, and Gilles Vilmart. A new class of uniformly accurate numerical schemes for highly oscillatory evolution equations. *Foundations of Computational Mathematics*, 2018.
- [7] T. A Driscoll, N. Hale, and L. N. Trefethen. Chebfun guide. <http://www.chebfun.org/docs/guide/>, 2014. Version utilisée : 5.6.0.
- [8] Ernst Hairer, Syvert P. Norsett, and Gerhard Wanner. *Solving Ordinary Differential Equations II. Stiff and differential-algebraic problems*, volume 8 of *Springer Series in Computational Mathematics*. Springer-Verlag Berlin Heidelberg, 1982.

TABLE DES FIGURES

1.1	Solution z en fonction du temps et la solution sur la variété associée $z_\infty : t \mapsto \varepsilon h^\varepsilon \circ \varphi_t(x_0^\varepsilon)$, pour $\varepsilon = 2^{-11}, 2^{-12}, 2^{-13}$. La fin du régime transitoire est indiquée en pointillés comme l'instant t_0 à partir duquel $\forall t \geq t_0, z(t) - z_\infty(t) \leq 10^{-8}$	12
1.2	Temps du calcul d'une solution u^ε de référence avec le schéma Radau IIA pour différentes valeurs de ε	15
3.1	Norme infinie des coefficients (x_k^ε) et (z_k^ε) en fonction de k dans le cas (1.3).	35
3.2	Erreur relative en norme infinie sur x (en haut) et erreur absolue en norme infinie sur z (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (1.3) avec un schéma RK4 sur les (u_k^ε)	35
3.3	Erreur relative en norme absolue sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (1.3) avec un schéma par différences finies.	37
3.4	Erreur relative en norme infinie sur x^ε (en haut) et sur z^ε (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (1.3) avec le schéma utilisant chebfun.	37
3.5	Erreur absolue en norme infinie sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (3.5) avec le schéma utilisant chebfun.	38
3.6	Erreurs absolue (trait plein) et relative (pointillés) et leur échelle respectives à gauche et à droite en fonction du temps après calcul avec l'implémentation chebfun avec $\Delta t = 2^{-10}$ et $\varepsilon = 2^{-14}$	38
3.7	Différence en norme L_t^∞ entre la solution « quasi-exacte » calculée avec chebfun et la solution calculée symboliquement avec un seul pas de temps pour $X^\varepsilon(\Delta t, \cdot)$ (en haut) et $Z^\varepsilon(\Delta t, \cdot)$ (en bas) en fonction de Δt (à gauche) et de ε (à droite).	39
3.8	Temps de calcul en secondes de la solution avec chebfun (en haut) et avec Radau IIA (en bas) en fonction de Δt (à gauche) et de ε (à droite).	41
A.1	Solution x (à gauche) et z (à droite) du problème (A.1) sur $[0; 0,7]$, avec en pointillé l'approximation de la variété associée.	45
A.2	Parcours de (x_1, x_2, z) dans \mathbb{R}^3 où x, z sont solutions du problème (A.2) sur $[0; 10]$	46
A.3	Solution z du problème (A.2) sur $[0; 3]$ pour $\varepsilon = 10^{-2}, 10^{-3}$	46
		57

C.1	Erreur relative en norme absolue sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (1.3) avec un schéma implicite-explicite.	50
C.2	Erreur relative en norme absolue sur x (en haut) et sur z (en bas) en fonction de Δt (à gauche) et de ε (à droite) pour le cas (1.3) avec un splitting de Strang.	51
C.3	Temps de calcul (en s) avec notre méthode de Runge-Kutta imbriqués en fonction de ε	52
C.4	Erreur relative en norme absolue sur x (à gauche) et sur z (à droite) en fonction de ε pour le cas (1.3) avec une méthode de Runge-Kutta imbriqués.	52
C.5	Différences entre x^ε et $x_\infty^{[3]}$ (ronds, trait plein bleu) et entre z^ε et $\varepsilon h^{[3]} \circ x_\infty^{[3]}$ (croix, trait pointillé rouge) en fonction de ε en $t = 1$, comparées à une pente ε^3 (pointillés noirs) dans le cas test (3.5).	53
C.6	Différences entre x^ε et $x_\infty^{[3]}$ (ronds, trait plein bleu) et entre z^ε et $\varepsilon h^{[3]} \circ x_\infty^{[3]}$ (croix, trait pointillé rouge) en fonction de Δt en $t = 1$ lorsque $x_\infty^{[3]}(1)$ est calculée par méthode d'Euler explicite de pas de temps Δt . La droite pointillée noire est de pente 1.	54