



# THÈSE

En vue de l'obtention du

**DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

---

---

Présentée et soutenue le 2016 par :

**ASMA TOUMI**

**Méthode numérique asynchrone pour la modélisation de phénomènes multi-échelles**

---

---

## JURY

EDWIGE GODLEWSKI  
JULIEN DIAZ  
PHILIPPE PONCET  
GUILLAUME DUFOUR  
HÉLÈNE BARUCQ  
RONAN PERRUSSEL

Professeur d'Université  
Chargé de Recherche  
Professeur d'Université  
Ingénieur de Recherche  
Directeur de Recherche  
Chargé de Recherche

Rapporteur  
Rapporteur  
Directeur de thèse  
Directeur de thèse  
Examineur  
Examineur

---

**École doctorale et spécialité :**

*MITT : Domaine Mathématiques : Mathématiques appliquées*

**Unité de Recherche :**

*Equipe d'accueil ONERA-ISAE, MOIS : MODélisation et Ingénierie des Systèmes*

**Directeur(s) de Thèse :**

*Philippe PONCET et Guillaume DUFOUR*

**Rapporteurs :**

*Edwige GODLEWSKI et Julien DIAZ*



*A ma chère Maman Salha  
A la mémoire de mon père Khalifa*



# Remerciements

Je veux ici exprimer toute ma reconnaissance à celles et ceux qui grâce auxquels j'ai pu mener à bien cette thèse.

Je tiens tout d'abord à remercier les membres du jury pour le temps qu'ils ont consacré à mon travail, leurs questions et leurs remarques. Je remercie Hélène BARUCQ d'avoir présidé mon jury de thèse. Je remercie vivement les rapporteurs Edwige GODLEWSKI et Julien DIAZ pour avoir accordé du temps à une lecture attentive et détaillée de mon manuscrit ainsi que pour leurs remarques encourageantes et constructives. Un grand merci pour Philippe PONCET mon directeur de thèse. Je remercie chaleureusement Ronan PERRUSSEL.

J'exprime toute ma gratitude à mon encadrant Guillaume DUFOUR. Merci pour ta disponibilité, ton soutien, ta patience et la confiance que tu m'as accordée.

Je remercie sincèrement François Rogier pour son aide très précieuse, sa gaieté naturelle et ses encouragements qui m'ont beaucoup aidé à surmonter les difficultés. Je me dois aussi de remercier Thomas Unfer.

Je remercie l'équipe DITIM de m'avoir accueilli chaleureusement durant cette thèse. Des remerciements tout particuliers à mes camarades de thèse et stagiaires qui ont rendu ces trois ans plus doux. Merci pour les nombreuses pauses cafés, les repas tous les midis au RU et les jeux de cartes. Je remercie ainsi : Jeanne, Alexandra, Tomasz, Marc, Jonathan, Kevin, Emanuele, Matthieu, Hugo, Fabrice, Quentin, David, Guillaume, Simon...

Durant ma thèse j'ai eu la chance de rencontrer des vrais amis d'autres départements, avec qui j'ai partagé des moments agréables : Ahmad, Mohamed et Ramzi. Nous avons réussi à créer notre petit groupe au sein de l'ONERA. J'ai bien apprécié pouvoir discuter de différents sujets avec vous même si je n'étais pas d'accord avec vos idées. D'ailleurs vous ne m'avez jamais convaincu ! Il n'empêche que je me suis bien marrée avec vous. C'était un grand plaisir de vous rencontrer.

Je me tourne maintenant vers ma famille que la distance géographique n'a pas découragé pour suivre l'évolution de ma thèse. Rien de tout cela aurait été possible sans votre amour, votre soutien et votre patience. Mes remerciements les plus profonds vont à ma mère Salha qui m'a soutenu durant mes nombreuses années d'études et qui a toujours été là pour moi. Je profite de ce paragraphe pour te dire à quel point je t'aime. Je te dédie ce travail en témoignage de mon profond amour et de mon immense reconnaissance. J'espère que tu es contente et fière de moi. Je dédie ce travail aussi à la mémoire de mon père qui attendait avec impatience ce jour où sa "petite" fille devient docteur mais dont la maladie en a décidé autrement. Un grand merci à mes frères : Mohamed, Sami et Hatem et à mes soeurs : Raja, Aicha et Houda qui m'ont soutenu moralement durant la réalisation de ce mémoire. Ils n'ont eu de cesse de m'encourager et de m'aider durant cette période. En particulier Houda qui a suivi l'évolution de ma thèse, qui

m'a aidé à surmonter les moments les plus difficiles et qui m'a remonté le moral dans les moments de doute. Je te remercie beaucoup pour ta qualité d'écoute et ta grande patience.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1.	Contexte : le projet MACOPA . . . . .	1
1.2.	La modélisation par équations aux dérivées partielles . . . . .	2
1.3.	Discrétisation spatiale des EDPs . . . . .	5
1.3.1.	Les différences finies . . . . .	5
1.3.2.	Les volumes finis . . . . .	5
1.3.3.	Les éléments finis . . . . .	5
1.3.4.	Galerkin discontinu . . . . .	6
1.4.	Discrétisation temporelle des EDPs . . . . .	7
1.4.1.	Les schémas implicites . . . . .	7
1.4.2.	Les schémas implicites-explicites . . . . .	8
1.4.3.	Les schémas explicites . . . . .	9
1.4.3.1.	Les schémas à pas de temps locaux : Les schémas LTS . . . . .	10
1.4.3.2.	Le schéma LTS-RK développé dans [35] . . . . .	11
1.4.3.3.	Les méthodes asynchrones . . . . .	12
1.5.	Objectifs de la thèse . . . . .	13
1.6.	Structure du manuscrit . . . . .	14
<b>I</b>	<b>Schéma numérique asynchrone d'ordre un pour la discrétisation des équations de transport</b>	<b>17</b>
<b>2</b>	<b>Étude du schéma numérique asynchrone pour les équations de transport</b>	<b>19</b>
2.1.	Définition du schéma numérique asynchrone . . . . .	19
2.1.1.	La méthodologie asynchrone . . . . .	19
2.1.2.	L'algorithme asynchrone . . . . .	20
2.2.	Application de la méthode asynchrone pour la discrétisation de l'équation de transport en dimension un d'espace . . . . .	21
2.2.1.	Discrétisation en espace de l'équation d'advection . . . . .	22
2.2.2.	Discrétisation en temps : le schéma asynchrone . . . . .	22
2.2.2.1.	L'algorithme asynchrone . . . . .	23
2.3.	Propriétés du schéma asynchrone . . . . .	25
2.3.1.	Stabilité du schéma asynchrone . . . . .	26

2.3.2.	Consistance du schéma asynchrone . . . . .	30
2.3.3.	Convergence du schéma asynchrone . . . . .	33
2.4.	Validation numérique . . . . .	35
2.5.	Conclusion . . . . .	38
<b>3</b>	<b>Application de la méthode asynchrone pour la discrétisation de l'équation de transport en multi-dimensions</b>	<b>41</b>
3.1.	Discrétisation de l'équation de transport en multi-dimensions . . . . .	41
3.2.	Stabilité du schéma asynchrone . . . . .	42
3.3.	Consistance du schéma asynchrone . . . . .	44
3.3.1.	Consistance : maillage constant . . . . .	45
3.3.2.	Consistance : maillage régulier . . . . .	49
3.4.	Convergence : méthode de l'erreur corrigée . . . . .	49
3.5.	Conclusion . . . . .	56
<b>4</b>	<b>Propriétés additionnelles du schéma asynchrone</b>	<b>57</b>
4.1.	Schéma Runge Kutta 2 asynchrone (A1RK2) . . . . .	57
4.1.1.	L'algorithme asynchrone . . . . .	58
4.1.2.	Propriétés du schéma RK2 asynchrone (A1RK2) . . . . .	59
4.1.2.1.	Stabilité du schéma A1RK2 . . . . .	59
4.1.2.2.	Consistance du schéma A1RK2 . . . . .	62
4.2.	Schéma MUSCL asynchrone . . . . .	63
4.2.1.	Limiteurs de pente . . . . .	63
4.2.1.1.	Exemples de limiteurs de pente . . . . .	63
4.2.1.2.	Caractéristiques de limiteur $\varphi$ . . . . .	64
4.2.2.	Algorithme asynchrone . . . . .	65
4.2.3.	Stabilité du schéma MUSCL asynchrone . . . . .	66
4.2.4.	Consistance du schéma MUSCL asynchrone . . . . .	68
4.3.	Schéma Lax Wendroff asynchrone . . . . .	68
4.3.1.	Consistance du schéma Lax Wendroff asynchrone . . . . .	69
4.4.	Tests numériques . . . . .	70
4.4.1.	Convergence numérique : perte d'ordre . . . . .	70
4.4.2.	Gain en temps de calcul pour une erreur fixée . . . . .	71
4.4.3.	L'asynchrone et les solutions irrégulières . . . . .	72
4.5.	Conclusion . . . . .	77
<b>II</b>	<b>Schéma Runge-Kutta asynchrone d'ordre deux pour les équations différentielles linéaires</b>	<b>79</b>
<b>5</b>	<b>Schéma asynchrone du second ordre</b>	<b>81</b>
5.1.	L'algorithme asynchrone . . . . .	81
5.1.1.	L'algorithme ARK2 . . . . .	82

5.2.	Étude de la convergence du schéma ARK2 . . . . .	83
5.2.1.	Notations . . . . .	83
5.2.2.	Comparaison du schéma ARK2 avec le schéma LTS-RK2 défini dans [35] . . . . .	96
5.3.	Validation numérique . . . . .	98
5.3.1.	Tests numériques en dimension un d'espace . . . . .	98
5.3.2.	Tests numériques en deux dimensions . . . . .	99
5.4.	Conclusion . . . . .	102
<b>6</b>	<b>Mise en oeuvre de la méthode asynchrone d'ordre deux : application pour les équations de Maxwell en dimensions deux et trois</b>	<b>105</b>
6.1.	L'expérience numérique . . . . .	105
6.2.	Exemple du fil en deux dimensions . . . . .	107
6.2.1.	Premiers tests : maillages multi-échelles . . . . .	107
6.2.1.1.	Le premier test pour le domaine $\Omega_1$ . . . . .	107
6.2.1.2.	Le premier test pour le domaine $\Omega_2$ . . . . .	111
6.2.2.	Tests avec des maillages fortement multi-échelles . . . . .	113
6.2.2.1.	Le deuxième test pour le domaine $\Omega_1$ . . . . .	113
6.2.2.2.	Le deuxième test pour le domaine $\Omega_2$ . . . . .	117
6.3.	Exemple du fil en trois dimensions . . . . .	119
6.3.1.	Le fil cylindrique . . . . .	121
6.3.2.	Le fil parallélépipédique . . . . .	122
6.4.	Performance du schéma asynchrone . . . . .	126
6.4.1.	La pièce . . . . .	128
6.4.2.	La sphère . . . . .	132
6.4.3.	Le fil avec défaut . . . . .	135
6.5.	Conclusion . . . . .	137
<b>7</b>	<b>Conclusion et perspectives</b>	<b>139</b>
7.1.	Conclusion . . . . .	139
7.2.	Perspectives . . . . .	143
<b>A</b>	<b>Preuves du Chapitre 4</b>	<b>145</b>
<b>B</b>	<b>Méthode de Galerkin discontinue pour la discrétisation des équations de Maxwell 3D</b>	<b>151</b>
	<b>Bibliographie</b>	<b>157</b>



# Résumé

La simulation numérique est devenue un outil central dans la modélisation de nombreux systèmes physiques tels que la dynamique des fluides, les plasmas, l'électromagnétisme, etc. L'existence de phénomènes multi-échelles rend l'intégration numérique de ces modèles très difficile du point de vue de la précision et du temps de calcul. En effet, dans les méthodes classiques d'intégration temporelle, le pas de temps est limité par la taille des plus petites mailles au travers d'une relation de type CFL. De plus, la forte disparité entre le pas de temps effectif et la condition CFL favorise les phénomènes de diffusion numérique. Dans la littérature, des nombreux algorithmes à pas de temps locaux (LTS) ont été développés. Pour la plupart des algorithmes LTS, les pas de temps locaux doivent être choisis parmi les fractions du pas de temps global.

Nous présentons dans cette thèse une méthode asynchrone pour l'intégration explicite des équations différentielles multi-échelles. Cette méthode repose sur l'utilisation de critères de stabilité locaux, critères déterminés non pas globalement mais à partir de conditions CFL locales. De plus, contrairement aux schémas LTS, l'algorithme asynchrone permet la sélection de pas de temps indépendants pour chaque cellule de maillage.

Cette thèse comporte plusieurs volets. Le premier concerne l'étude mathématique des propriétés du schéma asynchrone. Le deuxième a pour objectif d'étudier la montée en ordre, à la fois temporelle et spatiale, des méthodes asynchrones. De nombreux développements dans le cadre des méthodes de haute précision en temps ou en espace, telles que les méthodes de type Galerkin Discontinu, peuvent offrir un cadre naturel pour l'amélioration de la précision des méthodes asynchrones. Toutefois, les estimations garantissant l'ordre de précision de ces méthodes peuvent ne pas être directement compatibles avec l'aspect asynchrone. L'objectif de cette thèse est donc de développer un schéma numérique asynchrone d'ordre élevé mais qui permet également de limiter la quantité de calculs à effectuer. Le troisième volet de cette thèse se focalise sur l'application numérique puisqu'il concerne la mise en oeuvre de la méthode asynchrone dans la simulation des cas-tests représentatifs de problèmes multi-échelles.

**Mots clefs** : méthode numérique asynchrone, équations différentielles linéaires, ordre élevé, phénomènes multi-échelles.



# Abstract

Numerical simulation has become a central tool for the modeling of many physical systems (Fluid dynamics, plasmas, electromagnetism, etc). Multi-scale phenomena make the integration of these physical systems difficult in terms of accuracy and computational time. Numerical time-stepping integration techniques used for modeling such problems generally fall into two categories : explicit and implicit schemes. In the explicit schemes, all unknown variables are computed at the current time level from quantities already available. The time step used is then limited by the most restrictive CFL condition over the whole computation domain. In the implicit method the time step is no longer limited by the CFL conditions. However the scheme is generally not suitable for strongly coupled problems. To solve such problems, a number of local time-stepping (LTS) approaches have been developed. These methods are restricted by a local CFL condition rather than the traditional global CFL condition. For most of these LTS algorithms, local time steps are usually selected to be fractions of the global time step so that regular meeting points in time exist, and only little work is available on LTS methods with independent time steps.

We present in this thesis an asynchronous method for the explicit integration of multi-scale partial differential equations. This method is restricted by a local CFL condition rather than the traditional global CFL condition. Moreover, contrary to other LTS methods, the asynchronous algorithm permits the selection of independent time steps in each mesh element.

Our work consists of several components. The first one concerns the mathematical study of the properties of the asynchronous method. the objective of the second part is to study the improvement of the convergence rate for asynchronous methods. Many approaches in the context of high precision methods in time or in space, such as the Discontinuous Galerkin methods, may offer a natural setting to improve the precision of the asynchronous methods. However, the estimates ensuring the order of the accuracy of the method may not be directly compatible with the asynchronous aspect. Then, the objective is to develop a high order asynchronous numerical scheme which also preserves the computational time reduction. Finally, the third part is focused on the implementation of the asynchronous method and illustrate the advantages of the method on test-cases representative of multiscale problems.

**Key words** : Numerical asynchronous method, linear differential equations, high order accuracy, multiscale phenomena.



# Chapitre 1

## Introduction

### 1.1. Contexte : le projet MACOPA

Les travaux présentés dans ce manuscrit ont été menés à l'Office National d'Etudes et de Recherches Aérospatiales (ONERA) dans le contexte du projet MACOPA (Méthodes Asynchrones pour la Combustion et les Plasmas Atmosphériques), lequel est partiellement financé par l'Agence Nationale de la Recherche et mené en partenariat avec l'Institut de Mécanique des Fluides de Toulouse (IMFT), le Laboratoire PLAsma et Conversion d'Energie (LAPLACE) et l'Institut de Recherche en informatique de Toulouse (IRIT). L'objectif du projet MACOPA est le développement d'une plate-forme logicielle, basée sur le formalisme asynchrone, pour la résolution numérique de problèmes de transport multi-échelles et multi-physiques, l'existence de petites échelles provenant de termes sources très raides et très localisés. Cette description couvre un large domaine d'applications (simulation de processus naturels ou de cas-tests industriels) pour lesquelles il est nécessaire de pouvoir décrire des effets à grande échelle (d'espace ou de temps) résultant de processus caractérisés par des échelles beaucoup plus petites. Le développement de la plate-forme logicielle MACOPA s'appuie sur un noyau numérique existant, fruit d'une collaborations précédente entre l'ONERA, le laboratoire LAPLACE et l'IMFT. L'adaptation de ce noyau numérique dans le cadre du projet doit lui permettre de prendre en charge l'intégration temporelle d'un système générique d'équations aux dérivées partielles.

Le programme scientifique du projet MACOPA s'articule autour de deux types de tâches. Des tâches de type amont qui consistent à alimenter la plateforme avec des outils génériques. C'est le cas de la tâche 1, dont est responsable l'ONERA, qui consiste à développer et à étudier formellement la possibilité de développer des schémas asynchrones d'ordre élevé. Les travaux de cette thèse s'inscrivent dans cette optique. C'est également le cas de la tâche 2 (responsable IRIT) qui est concentrée sur les options de parallélisation pour les algorithmes asynchrones. Un second groupe de tâches rassemble les tâches dites applicatives. En particulier, la tâche 3 (gérée par le LAPLACE) consiste à créer les modules applicatifs de la physique des plasmas de décharge et de la résolution des équations de Maxwell. L'étude du couplage Plasma/Micro-ondes entre également dans les objectifs des travaux réalisés dans cette tâche. La tâche 4 (IMFT) est centrée sur l'intégration dans la plate-forme des modèles physiques liées à la combustion. Enfin la tâche 5 rassemble les tâches de coordination du programme, ainsi que les tâches d'intégration des différentes bibliothèques et leur validation au niveau de la plateforme ; voir Fig. 1.1.

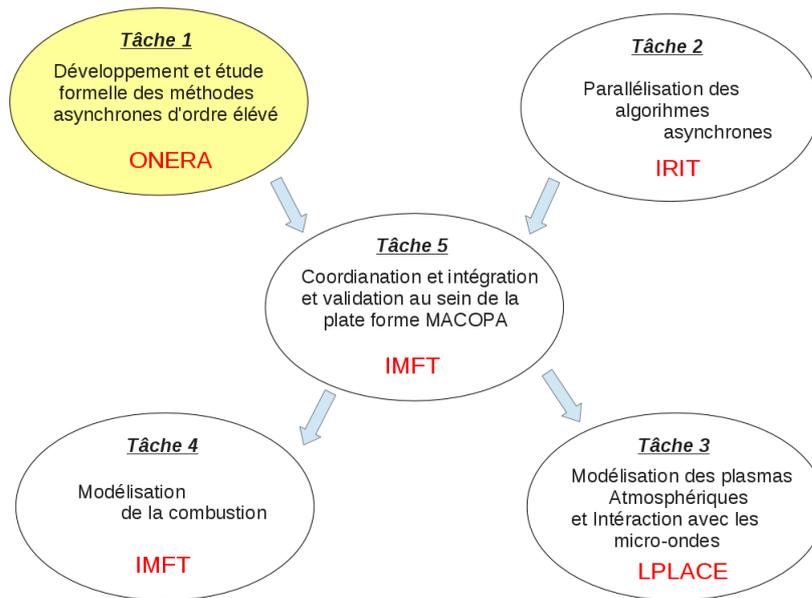


FIGURE 1.1 – Organisation du projet MACOPA.

## 1.2. La modélisation par équations aux dérivées partielles

La modélisation de très nombreux problèmes issus de la physique, de la mécanique et des sciences de l'ingénieur, font qu'ils se présentent sous forme d'équations aux dérivées partielles (EDP) : *"All in nature reduces to differential equations"* (Newton, Isaac : 1642-1727), *"...Present day physics, as far as it is theoretically organized, is completely governed by a system of space-time differential equations"* (Planck Max : 1858-1947) ; voir aussi Fig. 1.2 ( Les exemples présentés dans Fig. 1.2 ainsi que d'autres exemples peuvent être trouvés dans [74]).

On retrouve l'utilisation de modèles mathématiques à base d'équations aux dérivées partielles dans des domaines d'application divers et variés. Nous pouvons par exemple citer les domaines de :

1. l'ingénierie : optimiser la forme d'un véhicule pour réduire sa traînée.
2. l'environnement : modélisation thermique des bâtiments, météo, océan, séisme.
3. la communication : ondes sonores, électromagnétiques.
4. la défense : furtivité (comment construire un avion ne renvoyant aucune onde)
5. la finance : équation de Black-Scholes.

Rappelons que les équations aux dérivées partielles sont classées en trois grandes familles fondamentales d'équations : les équations elliptiques qui servent typiquement à décrire des phénomènes d'équilibre en physique, les équations paraboliques qui permettent en particulier de décrire des phénomènes de diffusion et les équations hyperboliques qui permettent de modéliser les phénomènes de propagation. La théorie des équations aux dérivées partielles fait l'objet d'une très vaste littérature, mais rares sont les situations réalistes pour lesquelles on sait calculer une solution analytique. Généralement, l'objectif de l'analyse mathématique d'une EDP est de garantir l'existence d'une solution, voire l'unicité de celle-ci, d'identifier les espaces fonctionnels adaptés et également de d'identifier quelques propriétés qualitatives

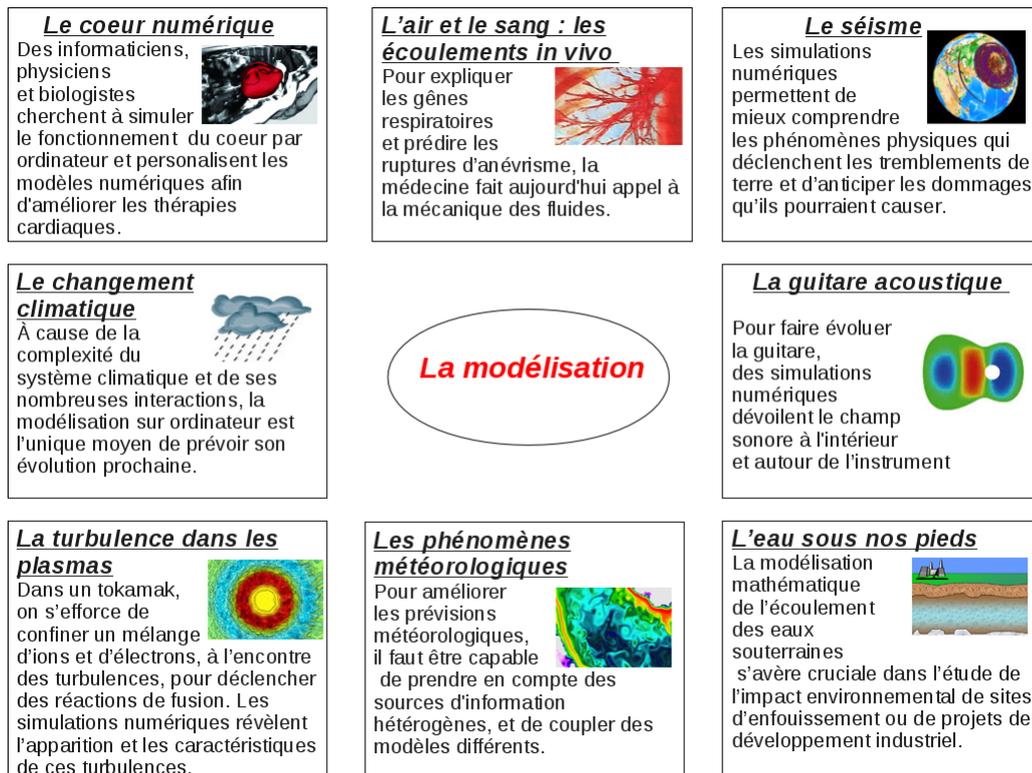


FIGURE 1.2 – Quelques exemples de la modélisation dans différents domaines, présentés dans [74].

des solutions de cette équation. Pour ce qui est de l'aspect quantitatif (voire prédictif) de la solution d'une EDP, il devient nécessaire de se tourner vers une discrétisation du problème pour avoir une solution approchée et donc de faire appel à une méthode numérique appropriée Fig. 1.3.



FIGURE 1.3 – Le passage d'un modèle réel vers un modèle discrétisé.

Une méthode numérique nécessite de transformer le problème continu en un problème discret en utilisant des méthodes de discrétisation spatiales et temporelles. Nous nous intéresserons en particulier aux méthodes qui s'appuient sur un maillage de l'équation physique. Un maillage est vu comme une division du domaine de résolution en petits volumes de contrôle. La discrétisation de l'EDP doit alors être adaptée au maillage. De notre point de vue dans ce manuscrit, une méthode numérique comporte donc les éléments suivants :

1. Le maillage : le maillage permet de décomposer un élément géométrique complexe (une surface ou un volume) en un assemblage de formes élémentaires (triangles, quadrangles, tétraèdres, etc.). La figure 1.4 donne un exemple de maillage permettant de décrire la forme d'un avion.
2. La discrétisation : c'est le procédé qui consiste à remplacer le problème initial continu (décrit par une EDP) par un problème approché discret, posé dans un espace de dimension finie. Nous

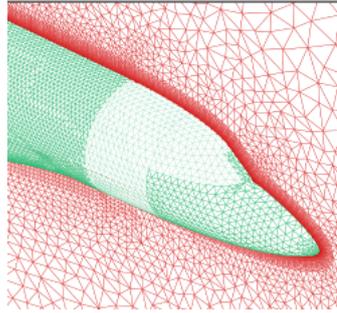


FIGURE 1.4 – Maillage d’un avion.

distinguerons deux types de discrétisation.

- (a) La discrétisation spatiale : Elle sera considérée comme donnée, et pourra être indifféremment issue des méthodes de différences finies, de volumes finis ou d’éléments finis. Ces trois formalismes seront explicités rapidement dans la section suivante.
- (b) **La discrétisation temporelle** : les méthodes d’intégration temporelle utilisées pour la modélisation des EDPs sont réparties en deux catégories : les schémas explicites et implicites. Alternativement, des méthodes implicite–explicit ont été développées. Il s’agit de la partie à laquelle nous nous intéressons spécifiquement dans cette thèse en développant des approches asynchrones de type explicite. Cet aspect sera détaillé plus avant dans la section 1.4..

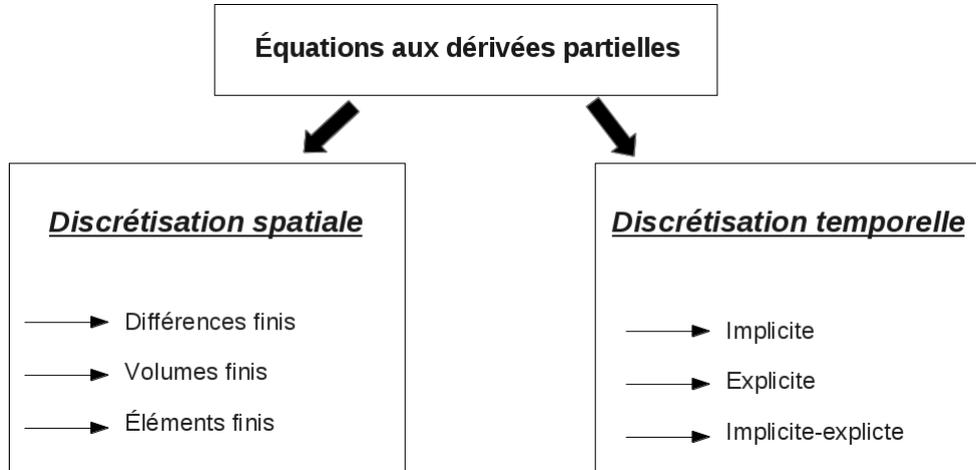


FIGURE 1.5 – La discrétisation des EDPs.

La modélisation des phénomènes physiques est confrontée à des difficultés liées principalement à la complexité des phénomènes traités (phénomènes couplés, multi-échelles, de très grande taille, etc). Parmi les points-clés de la simulation numérique, le calcul parallèle ainsi que la question du rendu graphique des calculs sont des éléments nécessitant une étude dédiée qui ne sera pas menée ici. Nous nous focaliserons sur le choix des modèles physiques, leur discrétisation (conception de méthodes numériques adaptées) ainsi que sur la possibilité d’en extraire une implémentation efficace.

## 1.3. Discrétisation spatiale des EDPs

Nous rappelons brièvement dans cette section les méthodes numériques les plus utilisées pour la résolution des équations aux dérivées partielles. Il existe trois grandes familles de méthodes numériques : les méthodes de différences finies, de volumes finis et enfin les méthodes d'éléments finis qui contiennent en particulier les approches de type Galerkin Discontinues qui nous intéresseront plus spécifiquement dans ce manuscrit.

### 1.3.1. Les différences finies

La méthode des différences finies est de loin la plus intuitive parmi les méthodes d'approximation des EDPs. Concrètement, l'idée de base consiste à approximer les dérivées (spatiales ou temporelles) des équations aux dérivées partielles en utilisant des développements de Taylor. Un des avantages majeurs de cette méthode est sa simplicité de mise en oeuvre. De plus, le passage à des approximations d'ordre élevé peut se faire à un coût de calcul faible. L'inconvénient de la méthode des différences finies est la précision et la manque de flexibilité. Cela provient du fait que l'application de cette méthode nécessite plus de régularité de la solution et du maillage [63, 64].

### 1.3.2. Les volumes finis

La méthode des volumes finis consiste à effectuer un bilan local sur des volumes élémentaires (les éléments du maillage), en intégrant sur ces éléments les équations écrites sous forme de loi de conservation. Cette approche exprime les échanges qui s'opèrent entre les différents éléments du maillage *via* la définition de flux numériques aux interfaces. Ce formalisme se révèle alors particulièrement adapté à la résolution des équations de conservation (et en particulier celles issues de la mécanique des fluides) car les formulations discrètes qui en résultent font intrinsèquement apparaître la conservation des quantités physiques telles que la masse ou l'énergie.

Avec un choix de formulation des flux adaptés, la méthode des volumes finis peut également facilement traiter les chocs dans le système physique sans engendrer d'oscillations parasites. De plus, étant facile à mettre en oeuvre sur des éléments quelconques (en particulier les simplexes), elle permet de traiter des géométries complexes avec des volumes de forme quelconque. Néanmoins cette méthode est difficilement étendue à des ordres élevés car la préservation de la stabilité et de principes du maximum tels que la positivité d'une densité imposent d'utiliser des limitations dans les reconstructions des formes des solutions au sein d'une cellule ce qui implique une perte d'ordre locale [21, 22]. Des méthodes s'affranchissant des contraintes de principe du maximum et autorisant des oscillations ont alors été développées, l'idée étant de maîtriser l'amplitude des oscillations [23, 24]. Toutefois, peu de résultats théoriques de convergence pour les méthodes générales de type volumes finis sont disponibles, la plupart d'entre eux limitant l'ordre maximal accessible des méthodes utilisées lorsque l'on veut garder de bonnes propriétés de stabilité [30, 31].

### 1.3.3. Les éléments finis

L'idée de base de la méthode des éléments finis repose sur une discrétisation de l'espace fonctionnel dans lequel le problème continu est posé. A partir du problème initial écrit sous forme variationnelle (ou

faible) dans un espace fonctionnel de dimension infinie, il s'agit de chercher une caractérisation d'une solution approchée, choisie dans un sous-espace de dimension finie [61]. Cette forme variationnelle peut en général être interprétée comme la minimisation d'une énergie, ce qui rend ce type d'approche particulièrement bien adaptée aux problèmes d'équilibre ainsi qu'à la résolution des équations elliptiques pour lesquelles la dérivation d'une formulation variationnelle est possible. Cette approche peut se généraliser sans obstacle théorique à des maillages avec des éléments quelconques, la difficulté technique étant reportée sur le choix des fonctions de base qui constituent l'espace de dimension finie retenu. Le choix de ces fonctions de base est délicat et permet de définir plusieurs classes de méthode de type éléments finis. Les fonctions à support local simplifient grandement leur construction [60]. Contrairement aux différences finies, la méthode des éléments finis permet de traiter assez simplement des géométries complexes. Pour augmenter l'ordre d'approximation de la méthode, de nombreuses méthodes ont été développées en utilisant des fonctions de base particulières d'ordre plus élevé. Cependant, plus l'ordre d'approximation est grand, plus cette méthode d'éléments finis demande la résolution de systèmes linéaires de grande dimension, ce qui implique alors une forte demande en mémoire et en temps de calcul.

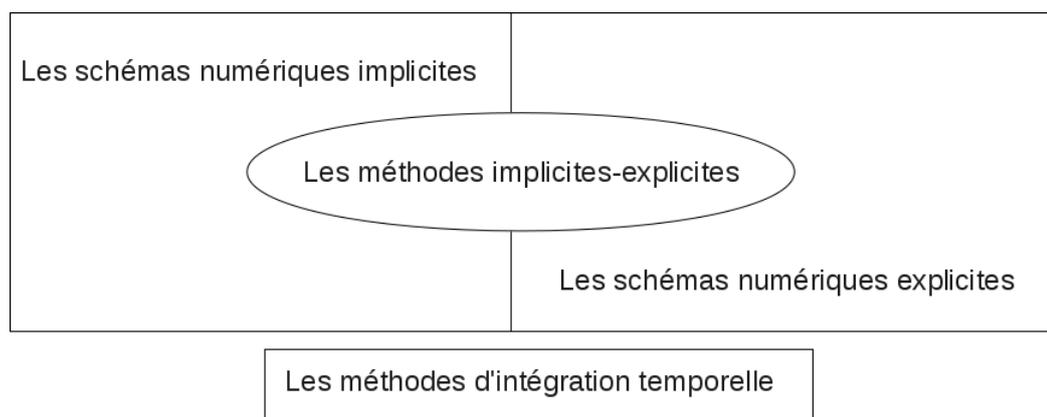
#### **1.3.4. Galerkin discontinu**

Les méthodes de Galerkin Discontinues (GD) peuvent être considérées comme des méthodes de type éléments finis particulières, permettant à la solution approchée d'être discontinue. Elles peuvent également être considérées comme des méthodes de volumes finis dans le sens où l'approximation de la solution sur chaque élément par une fonction polynomiale permet de définir un flux aux interfaces du maillage. Il est alors nécessaire de caractériser les discontinuités de l'approximation (les sauts aux interfaces) pour pouvoir garantir la convergence de la méthode, ce qui fait l'objet de nombreuses formulations [71, 72, 73]. D'un point de vue pratique, le travail avec des espaces discrets discontinus offre une flexibilité considérable ce qui rend cette approche attirante pour les simulations multi-physiques et multi-domaines. L'application des méthodes Galerkin Discontinues couvre un vaste domaine des sciences de l'ingénieur [47, 48]. Les méthodes GD ont existé sous diverses formes depuis plus de 30 ans et ont connu un développement significatif durant cette dernière décennie.

La première méthode GD pour l'approximation des équations aux dérivées partielles du premier ordre a été introduite par Reed et Hill en 1973 [55]. L'estimation d'erreur correspondante a été effectuée en 1986 par Johnson et Pitkäranta [54]. Quelques années plus tard, la méthode a été appliquée à la simulation des équations de la couche limite en dimensions trois pour l'écoulements de fluides compressibles par Caussignac et Touzani [49, 50]. Dans [51], Chavent et Cockburn ont présenté une extension de la méthode GD pour les équations aux dérivées partielles hyperboliques en utilisant la méthode d'Euler explicite pour la discrétisation en temps. Pour la montée en ordre de convergence, Cockburn et Shu ont utilisé dans [52, 53] la méthode de Runge-Kutta pour la discrétisation en temps. La convergence a été étudiée par Jaffré, Johnson et Szepessy dans [56]. Les extensions ont été discutées par Cockburn, Shu and al. dans une série des publications [57, 58, 59].

## 1.4. Discrétisation temporelle des EDPs

Les méthodes précédentes sont considérées comme des approximations en espace, approximations qui doivent - dans le cas d'équations d'évolution - être couplées avec un algorithme d'intégration temporelle. Les méthodes classiques d'intégration temporelle utilisées pour la modélisation des EDPs sont réparties en deux catégories : les schémas explicites et implicites. Alternativement, des méthodes implicit-explicit essayant de combiner le meilleur des deux mondes ont également été développées [36]. Pour situer les méthodes asynchrones, qui nous intéressent dans ces travaux de thèse, nous revenons brièvement sur ces trois approches.



### 1.4.1. Les schémas implicites

Une méthode implicite consiste à calculer la solution à l'instant courant en utilisant des valeurs calculées à l'instant précédent - connues explicitement - mais également des valeurs à l'instant courant. La valeur recherchée est donc définie en tant que solution d'une équation qu'il est nécessaire de résoudre. L'avantage majeur des schémas implicites réside dans la stabilité accrue du schéma numérique, nombre d'entre eux étant inconditionnellement stables, ce qui permet d'utiliser des pas de temps relativement grands en comparaison avec les schémas explicites (abordés plus loin). Cependant, une méthode implicite conduit à la résolution d'une équation et donc généralement à l'inversion d'une matrice, ce qui est très coûteux en termes d'espace mémoire et de temps de calcul. De plus, l'utilisation de plus grands pas de temps s'accompagne généralement d'une augmentation de la diffusion numérique. Pour pallier ce dernier problème, des stratégies de schémas implicites permettant la montée en ordre, tels que les schémas de type Runge-Kutta Implicites ont été développés. Parallèlement, avec l'amélioration des performances des méthodes itérative pour résoudre des systèmes linéaires - citons par exemple les algorithmes de Krylov, Newton et Schwartz - les schémas implicites sont devenus des outils importants pour l'étude de phénomènes multi-échelles décrits par les EDPs non linéaires. Parmi ces phénomènes, citons les systèmes de réaction-diffusion [41, 42], de radiation-diffusion [43], magnétohydrodynamique [44] et hydrodynamiques de radiation [45].

Afin de décrire correctement l'évolution du système, les schémas implicites utilisent des stratégies adéquates, basées sur la physique, pour la sélection des pas de temps qui sont en rapport avec la taille du pas de temps dominant du problème. Le choix de la taille des pas de temps qui permet l'efficacité en terme

du temps de calcul peut souvent être obtenu par l'application de la condition de stabilité CFL [46]. Malgré les progrès récents dans ce domaine, les schémas implicites avec des pas de temps relativement grand montrent un coût de calcul très grand. En effet, atteindre la convergence pour des problèmes avec des non-linéarités nécessite d'utiliser un grand nombre d'itérations. De plus, dans certains cas, les schémas implicites sont malgré tout encore incapables de bien représenter les phénomènes physiques locaux. Ces problèmes rendent ces méthodes inefficaces pour la modélisation du comportement de ces systèmes en temps long, par exemple, rentrent dans ce cas la magnétosphère de la Terre, les modèles climatiques, etc. Pour ce genre de problèmes on a généralement besoin de traiter des différentes échelles de temps avec la même précision sur de longues périodes de temps, ce qui est le cas de problèmes fortement couplés mais avec des échelles de temps très différentes, phénomène souvent illustré en physique des plasma [15].

### 1.4.2. Les schémas implicites-explicites

Le raffinement local du maillage consiste à prendre des petits éléments précisément là où se situent les longueurs caractéristiques les plus petites (et généralement de façon corollaire les pas de temps les plus petits également), et des plus grand éléments ailleurs ; voir Fig. 1.6 . Dans ce contexte, les éléments du maillage sont répartis sur deux régions distinctes : une région qualifiée de grossière et une région dite raffinée. L'utilisation d'un schéma numérique implicite ou explicite dans tout le domaine de calcul conduit alors dans les deux cas à un coût très élevé.

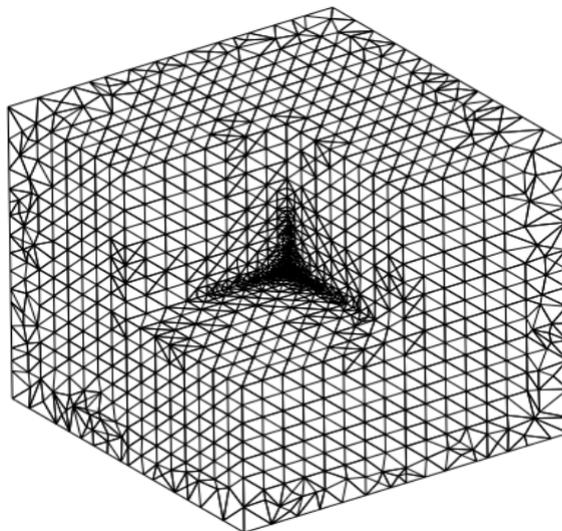


FIGURE 1.6 – Exemple d'un maillage en 3D comportant une zone raffinée.

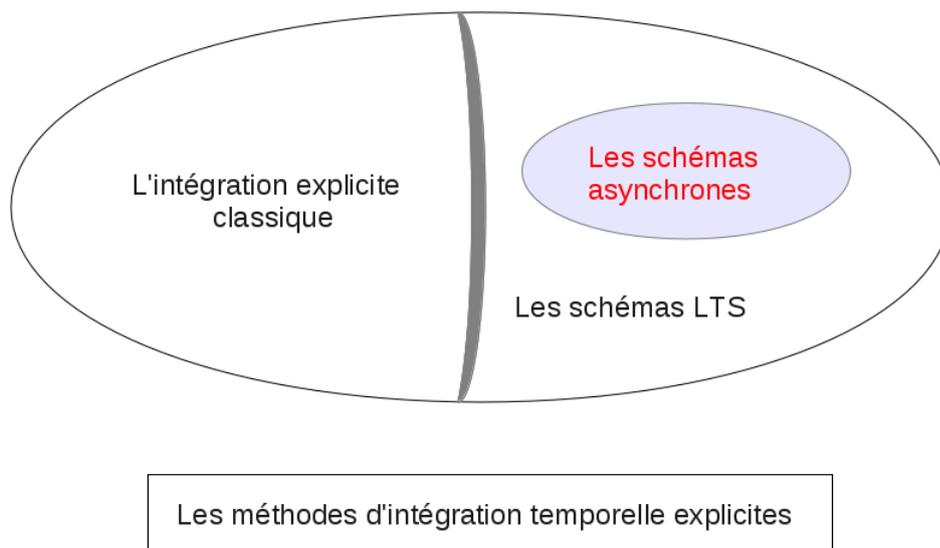
Les méthodes implicites-explicites (ou semi-implicites ou semi-explicites) consistent à traiter implicitement la zone raffinée et explicitement le reste du maillage. Cette méthode a l'avantage de ne pas pénaliser le pas de temps à utiliser sur le maillage grossier à cause des contraintes du maillage raffiné. Cependant, cette approche nécessite des aménagements particuliers. Par exemple, dans [37], Piperno a utilisé le schéma Leap-Frog explicite combiné avec le schéma implicite Crank-Nicolson et la méthode de Galerkin discontinue pour la discrétisation des équations de Maxwell dans un milieu non conducteur. À l'intérieur de la région raffinée, il faut résoudre un système linéaire à chaque pas de temps. Chacune de

ces méthodes est démontrée pour être d'ordre deux en temps, cependant Descombes et al. ont démontré dans [38] que le schéma implicite-explicite réduit l'ordre de convergence à un. Pour résoudre ce problème de la perte d'ordre, Descombes et al ont utilisé le schéma semi-explicite proposé par Verwer dans [39]. Ils ont montré dans [38] que ce schéma est convergent d'ordre deux. Cependant, cette méthode nécessite la résolution d'un système linéaire significativement plus grand que celui de la méthode de Piperno. Dans [40], autre exemple, Kanevsky et al. ont développé un schéma Runge-Kutta implicite-explicite d'ordre quatre avec une discrétisation de Galerkin Discontinue pour l'écoulement de fluides. À chaque pas de temps, la méthode de Newton-Krylov est utilisée pour la résolution du système non linéaire associé aux éléments fins. Cette méthode devient alors rapidement mal conditionnée.

Dans ce cas, une alternative est de rester en explicite sur les deux domaines mais de prendre des stratégies d'intégration particulières, que nous allons aborder dans ce qui suit.

### 1.4.3. Les schémas explicites

L'intégration temporelle explicite consiste à calculer la solution à l'instant courant en utilisant uniquement des valeurs calculées aux instants précédents, ce qui rend l'implémentation de ce type des méthodes relativement simple à mettre en oeuvre. Cependant, pour les schémas explicites la stabilité du schéma numérique n'est garantie que lorsque le pas de temps est limité par la taille des plus petites mailles au travers d'une relation de type CFL. De plus, lorsque l'on observe une forte disparité entre le pas de temps effectif et la condition CFL, cela favorise les phénomènes de diffusion numérique. Les schémas explicites ont l'avantage de proposer assez simplement des algorithmes - tels que les méthodes de Runge-Kutta - permettant de monter en ordre [25] jusqu'à des ordres élevés. Nous nous intéressons ici plus spécifiquement au cas où la physique et/ou le maillage imposent des contraintes de pas de temps très différentes au sein d'un même domaine de calcul. Outre la problématique de la diffusion numérique évoquée précédemment, l'utilisation d'un pas de temps global impose de calculer de nombreuses étapes d'évolutions dans des zones où la dynamique du phénomène étudié ne l'impose pas. Pour pallier cette contrainte liée à la condition CFL, de nombreux algorithmes à pas de temps locaux (LTS : Local Time Stepping), dont les méthodes asynchrones, ont été développés.



### 1.4.3.1. Les schémas à pas de temps locaux : Les schémas LTS

Les méthodes dites Local-Time-Stepping (LTS) permettent de surmonter l'effet paralysant de raffinement local en utilisant des pas de temps plus petits à l'intérieur de la région raffinée, et uniquement dans celle-ci, tout en restant totalement explicite dans l'ensemble du domaine de calcul.

Dans un premier temps, Osher et Sanders ont proposé dans [1] une méthode à pas de temps local pour résoudre des systèmes de lois de conservation scalaires en dimension 1. Ils ont donné une analyse approfondie d'une discrétisation précise au premier ordre en espace assortie d'un schéma d'Euler explicite à pas de temps local en temps. Leur approche permet à chaque élément de prendre soit un pas de temps considéré comme complet, soit un nombre fixé de plus petits pas de temps permettant d'atteindre le pas de temps complet. Dans [3], Tang et Warneck ont proposé une classe de schémas à pas de temps local pour la résolution des systèmes hyperboliques de lois de conservation en projetant les incréments de la solution à chaque pas de temps local. Savcenco et al. [4] ont de leur côté construit un schéma appelé "multirate" pour la résolution des problèmes paraboliques. Leur schéma a été obtenu en adaptant le schéma implicite de type Rosenbrock. L'idée consiste à calculer une prédiction pour toutes les cellules de maillage par la formule du trapèze, puis d'utiliser l'information provenant d'un maillage grossier pour faire avancer le maillage fin en calculant les valeurs interpolées par la formule de quadrature. Dawson et Kirby [5] ont ensuite développé un schéma de type upwind pour la résolution de lois de conservation qui permet de coupler le pas de temps local avec le raffinement local en espace. L'originalité de leur travaux est d'utiliser un limiteur pour le schéma upwind déterminé à partir de l'issue des étapes précédentes.

Dans [6], Willem et al. ont remarqué que la plupart des schémas "multirate" utilisés pour la résolution des systèmes de lois de conservation ont généralement l'un des défauts suivants : soit les schémas sont localement inconsistants (voir par exemple [1, 5]), soit ils ne conservent pas la masse (voir par exemple [3, 4]). Les auteurs ont discuté dans ce papier de ces deux défauts pour l'équation de transport en dimension 1 d'espace. Ils se sont intéressés en particulier à étudier la monotonie (pour différentes normes ou semi-normes) de schémas multirate. Dans le même contexte, une analyse d'erreur a été présentée dans [7] pour la méthode de Runge Kutta partitionnée appliquée aux lois de conservation. Willem et al. ont montré que les différentes méthodes multirate étudiées dans [6] conduisent à une réduction de l'ordre de schéma par rapport aux estimations formelles théoriques. Ils ont ensuite appliqué ces différents schémas avec une décomposition basée sur les cellules pour l'équation de Burgers avec une solution discontinue. Ils ont alors remarqué sur des illustrations numériques que pour les schémas qui conservent pas la masse [4, 3] le choc se déplace avec une mauvaise vitesse. Pour résoudre ce problème, une méthode de partition basée sur les flux a été proposée. Cette approche garantit la conservation de la masse mais peut conduire à une perte locale de consistance. Berger et Olinger [2] ont alors proposé une formulation pour une adaptation automatique de maillage. Leur approche consiste à introduire des maillages structurés par blocs. L'injection et l'interpolation ont été par suite utilisées pour faire passer l'information entre les maillages fins et grossiers.

Flaherty et al. [9] ont développé une méthode de Galerkin Discontinue avec un schéma d'Euler explicite à pas de temps local. Leur approche consiste à calculer les valeurs interpolées aux interfaces pour faire avancer les différentes zones de maillage. Cependant, avec cette formulation, le flux n'est pas conservé le long des interfaces. De plus, les auteurs ont traité seulement le premier ordre en temps. Dans [11], Collino et al. ont proposé une méthode LTS basée sur la conservation d'énergie pour l'équation

d'onde en 1D. Leur approche est explicite à l'intérieur des régions grossière et fine. Cependant, à l'interface de deux régions, pour assurer la consistance des flux, il est nécessaire de résoudre un système linéaire à chaque pas de temps. L'analyse de cette méthode a été publiée dans [12]. Des schémas LTS, basés sur le schéma Runge-Kutta (RK) explicite classique à pas de temps global, ont alors été logiquement développés. En effet, les méthodes de Runge-Kutta sont des méthodes facilement adaptables à l'aspect LTS. Dans [13], Constantinescu et al. ont développé un schéma multirate explicite dérivé du schéma RK pour les systèmes hyperboliques de lois de conservation. Cette méthode est toutefois limitée à l'ordre deux. Récemment, parmi les méthodes LTS qui assurent une bonne performance, nous pouvons citer la méthode Runge-Kutta à pas de temps local développée par Grote et al et présentée dans [35]. Cette méthode sera présentée plus en détails dans la sous section 1.4.3.2..

Notons que pour tous les différents algorithmes présentés, les pas de temps locaux doivent être choisis parmi les fractions du pas de temps global. Les algorithmes sont en particulier applicables dans le cas d'un maillage composé de deux régions : une région raffinée avec un petit pas de temps et le reste du maillage avec un pas de temps global.

### 1.4.3.2. Le schéma LTS-RK développé dans [35]

Dans [35], les auteurs ont développé un schéma Runge Kutta explicite à pas de temps local LTS-RK dérivé du schéma Runge-Kutta classique. Pour expliquer leur algorithme, ils ont supposé que le maillage est composé de deux régions. Autrement dit,  $\Omega = \Omega_c \cup \Omega_f$  où  $\Omega$  est le domaine d'étude,  $\Omega_c$  est la région "grossière" (coarse) et  $\Omega_f$  est la région "fine". A la partie grossière correspond un pas de temps  $\Delta t_c$  et à la partie fine correspond un pas de temps  $\Delta t_f$ . Les pas de temps ont été déterminés en utilisant la condition CFL locale et ont supposé de plus que le grand pas de temps est un multiple du petit pas de temps. On a donc  $\frac{\Delta t_c}{\Delta t_f} = p$ , avec  $p$  un entier. Cette dernière hypothèse est dans leur étude indispensable pour pouvoir appliquer la méthode LTS-RK.

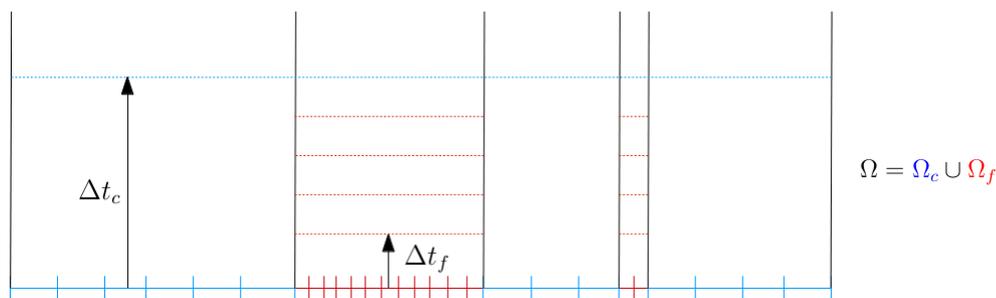


FIGURE 1.7 – Exemple d'un maillage 1D avec deux régions : une région "grossière" avec le pas de temps  $\Delta t_c$  et une région "fine" avec le pas de temps  $\Delta t_f$

L'idée de la méthode consiste à partitionner le vecteur  $y$ , l'inconnue que l'on cherche à approcher, comme suit :

$$y(t) = (I - P)y(t) + Py(t) = y^c(t) + y^f(t), \quad (1.1)$$

où  $P$  est une matrice diagonale dont les éléments diagonaux sont égaux à zéro ou un et permet d'identifier les inconnues associés à la région fine. A l'intérieur de la région grossière, le vecteur  $y^c$  est approché

en utilisant une formule de quadrature et pour la région fine, le vecteur  $y^f$  est approché en utilisant la méthode Runge-Kutta classique. Les valeurs intermédiaires nécessaires aux interfaces, sont alors obtenues grâce à une combinaison judicieuse d'une interpolation et d'un développement de Taylor. Grote et al ont prouvé dans [35] que leur schéma LTS-RK preserve l'ordre de convergence du schéma RK classique à pas de temps global. Cependant, leur approche n'est applicable que dans un cas particulier du maillage, même s'il est possible d'utiliser leur schéma de façon récursive pour définir des maillages partitionnés en plus de deux zones. De plus, les auteurs n'ont pas discuté, jusqu'à présent, le gain en temps de calcul en comparaison avec le cas classique.

### 1.4.3.3. Les méthodes asynchrones

Les méthodes asynchrones permettent, comme leur nom l'indique, des intégrations temporelles locales sans aucun prérequis de synchronisation avec le reste de l'environnement. L'intérêt de cette approche est qu'elle ne nécessite aucune hypothèse sur le maillage et peut également s'adapter au couplage de plusieurs phénomènes physiques avec des pas de temps d'amplitude très différents. Toutefois, ce paradigme fait que la dérivation de méthodes asynchrones fiables - *i.e.* qui permettent à la fois une sélection de pas de temps indépendants entre eux, et une convergence à l'ordre de précision attendu - représente un problème difficile et seulement quelques méthodes sont actuellement disponibles. Récemment, une nouvelle approche a été appliquée aux équations électrodynamiques non linéaires dans [10]. L'approche retenue est une forme discrète du principe variationnel de Hamilton. L'algorithme consiste à choisir des pas de temps locaux indépendants et donc chaque cellule de maillage est mise à jour avec une fréquence imposée par sa propre condition CFL locale. Cependant, cette approche est uniquement applicable aux systèmes Hamiltoniens. Dans [34], A. Taube and al. ont développé un schéma ADER-DG. Il s'agit d'un schéma à pas de temps local pour les équations de Maxwell pour lequel ils ont pu illustrer, à l'aide de tests numériques, que leur approche preserve la précision du schéma classique à pas de temps global. Cependant, et de manière assez surprenante, ce schéma augmente le temps de calcul par rapport au schéma classique à un pas de temps, ce qui limite son intérêt pour des cas de calcul dits lourds.

Dans [14] Omelchenko et Karimabadi ont présenté une approche asynchrone pour l'équation de convection–diffusion–réaction. Leur méthode est basée sur la simulation à événement discret (*discrete-event simulation* ou DES). L'intérêt principal de ces travaux réside dans la définition du concept de flux capacitor, qui permet de stocker le flux entre deux cellules jusqu'à son utilisation lors d'un événement ultérieur, ce qui garantit la conservation de masse a posteriori quand la cellule voisine est elle aussi remise à jour et que les flux capacitors sont vidés. Le concept du flux capacitor introduit également une notion de seuil minimal en dessous duquel une perturbation ne peut absolument pas se propager. De plus, si un seuil est dépassé, une cellule peut déclencher la remise à jour de ses voisins. Les auteurs n'ont cependant pas publié de preuve de la convergence de leur schéma ni proposé d'étude de l'ordre de cette approche.

Récemment, durant ses travaux de thèse [15], Thomas Unfer a proposé une méthode numérique asynchrone qui prolonge la notion de flux capacitor et qui repose sur l'utilisation de critères de stabilité locaux. Chaque cellule de maillage utilise des algorithmes fixés sur le pas de temps déterminé par sa propre CFL. Contrairement au schéma défini dans [14], l'approche suivie dans [15] est de toujours utiliser des pas de temps certes inférieurs à ceux imposés par la condition CFL mais également choisi le plus proche possible de celle-ci. On observe alors numériquement que ce choix a pour effet de minimiser

largement les phénomènes de diffusion numérique. Un gain de performance significatif a été observé dans le cadre des travaux de thèse de T. Unfer, toutefois l'analyse mathématique des performances du schéma et de sa montée en ordre n'ont pas pu être menés durant cette première étude. Ces travaux sont le point de départ des travaux menés dans cette thèse, ce schéma sera donc largement présenté et discuté dans ce qui suit.

## 1.5. Objectifs de la thèse

Le projet MACOPA (Méthodes Asynchrones pour la Combustion et les Plasmas Atmosphériques), dans lequel s'inscrivent ces travaux de thèse, a pour objectif de proposer un solveur multi-physique performant pour la simulation de problèmes physiques complexes tels que la combustion ou les plasmas atmosphériques. L'aspect performance se décline sous deux aspects : la parallélisation du code, que nous ne traiterons pas ici, et le gain en temps de calcul intrinsèque au concept de calcul asynchrone. En effet, la principale difficulté rencontrée lors de l'intégration numérique des phénomènes physiques visés est liée au caractère fortement multi-échelles de ceux-ci. Cela se traduit par une très forte inhomogénéité des longueurs et temps caractéristiques des modèles, ce qui induit numériquement des pas de discrétisation et de temps très disparates. La conséquence est que les simulations numériques ont alors des temps d'exécution gigantesques, même sur des machines très performantes.

Comme discuté précédemment, la réponse qui est apportée à ce dernier problème, dans le cadre du projet MACOPA et de cette thèse, trouve son origine dans les travaux de T. Unfer [16] et se base sur une approche asynchrone d'intégration temporelle, couplée à des schémas de type volumes finis. Cette approche repose sur l'utilisation de critères de stabilité pour les schémas numériques explicites, critères déterminés non pas globalement mais à partir de conditions CFL locales, chaque élément du maillage est alors remis à jour selon son temps caractéristique local, évitant des étapes superflues. En pratique, cette méthode asynchrone permet de limiter la quantité de calculs à effectuer et donc de traiter, avec des coûts de calcul raisonnables, des simulations numériques de phénomènes multi-échelles (tels que les systèmes de transport-diffusion-réaction) sur des temps physiques longs.

Les travaux développés dans cette thèse concernent l'analyse mathématique du schéma asynchrone tel qu'il a été proposé dans [16] ainsi que l'étude de la possibilité d'étendre ce formalisme pour monter en ordre en temps, en couplant cette approche asynchrone avec des méthodes d'ordre supérieur en espace. En effet, l'approche Volumes Finis qui a été utilisée jusqu'à présent se base sur une approximation des flux précise au premier ordre en espace, ce qui reste insuffisant pour certaines applications. De plus, la partie temporelle est actuellement limitée au schéma d'Euler, explicite et précis au premier ordre en temps. T. Unfer a proposé une extension formelle possible de l'approche Runge-Kutta dans [33] mais celle-ci reste toutefois au premier ordre en temps.

De nombreux développements dans le cadre des méthodes de haute précision en temps ou en espace, telles que les méthodes de type Galerkin Discontinu, peuvent offrir un cadre naturel pour l'amélioration de la précision des méthodes asynchrones. Toutefois, les estimations garantissant l'ordre de précision de ces méthodes peuvent ne pas être directement compatibles avec l'aspect asynchrone d'intégration en temps. Cette thèse a donc pour objectif d'identifier et de développer les schémas numériques qui pourront être adaptés au cadre asynchrone. Dans cette optique, l'étude comportera plusieurs volets. Le premier

sera plutôt théorique puisqu'il concernera d'abord l'étude mathématique des propriétés du schéma asynchrone existant, puis le développement de schémas numériques asynchrones d'ordre élevé. La seconde partie sera consacrée à la mise en oeuvre de ces schémas dans la simulation de cas-tests simples mais représentatifs de maillages fortement multi-échelles.

## 1.6. Structure du manuscrit

Le présent document synthétise l'ensemble des travaux effectués lors de cette thèse. Après cette introduction, les travaux sont subdivisés en cinq chapitres.

Le chapitre 2 revient sur la définition et la formulation de la méthodologie asynchrone, telle qu'elle a été proposée dans [15]. Il s'agit de proposer un cadre mathématique adapté qui permette de décrire et comprendre ce formalisme asynchrone. Dans ce chapitre qui illustre l'algorithme dans le cas simple d'une équation de transport monodimensionnelle, nous proposons une étude à la fois théorique et numériques des principales propriétés du schéma asynchrone original. Le but est alors de pouvoir répondre à la question de la consistance de l'avancement en temps d'un tel schéma. Pour ce faire, la méthode asynchrone est appliquée pour la discrétisation de l'équation de transport en dimension 1 de l'espace. Le choix d'une application relativement simple est dans le but de pouvoir mieux comprendre l'algorithme asynchrone mais aussi de pouvoir contrôler les éventuelles complexités liées à l'asynchrone dans les estimations. Le choix d'une norme et d'une définition de la consistance adaptées est alors proposé et la consistance au premier ordre en temps et en espace est démontrée.

L'extension de la méthode asynchrone dans le cas multi-dimensionnel est l'objet du chapitre 3. Il s'agit de proposer un formalisme général pour l'approche asynchrone en identifiant les faces et les types de flux observés par cette face (flux entrants ou sortants). Dans ce cadre, les propriétés de la stabilité et de consistance de la méthode asynchrone au premier ordre en temps et espace sont étudiées. Les résultats obtenus sont alors exprimés en fonction d'une estimation de l'aspect asynchrone du problème, c'est à dire les valeurs extrêmes de rapports de pas de temps observés par l'algorithme. Ces pas de temps sont exprimés grâce à la condition CFL, en fonction de la variation de la vitesse dans le domaine ainsi que et de la variation de taille des cellules du maillage. Dans ce chapitre, la convergence de la méthode a pu être démontrée, sous certaines hypothèses de régularité de la vitesse et du maillage, grâce à une adaptation à l'aspect asynchrone de la méthode dite de l'erreur corrigée.

Dans les deux chapitres précédents, la discrétisation qui a été utilisée se base sur une approximation au premier ordre à la fois en temps et en espace. Nous étudions dans le chapitre 4 la possibilité de monter en ordre en adaptant des schémas, qui sont classiquement d'ordre deux dans le cadre classique synchrone, à l'aspect asynchrone. L'idée est de mettre en évidence la perte d'ordre observée sur ces schémas, perte d'ordre liée au côté asynchrone de l'intégration qui va utiliser des estimateurs de pente à des instants qui ne sont pas forcément ceux utilisés dans les schémas synchrones. Sont abordés en particulier le cas des schémas de Lax-Wendroff et des approximations de type MUSCL avec limiteur de pente. Cette étude permet de conclure à la nécessité, pour monter en ordre, de développer un schéma d'intégration en temps qui soit cohérent avec l'aspect asynchrone et donc qui impose des événements de remise à jour des valeurs de flux en plus des événements de mise à jour des cellules. On observe également dans ce chapitre un comportement qualitatif non nécessairement quantifiable des schémas asynchrones. Ces

quelques propriétés numériques du schéma asynchrone sont alors illustrées et discutées dans le cas de problèmes avec des solutions irrégulières.

Dans le chapitre 5, nous développons un schéma numérique d'intégration temporelle asynchrone d'ordre deux. Ce schéma Runge-Kutta 2 Asynchrone (ARK2) est obtenu à partir de la formulation du schéma de Runge-Kutta d'ordre deux classique mais impose, en plus de la remise à jour des valeurs des cellules, des temps de remise à jour des flux. La possibilité de coupler des intégrations temporelles du second ordre avec des pas de temps différents pour deux cellules adjacentes est inspirée des travaux présentés par M. Grote dans [35] et repose sur une partition de chaque cellule en plusieurs domaines, en distinguant les échanges de données entre la cellule et ses voisines (influence des flux) et l'évolution interne de la cellule. La séparation de l'évolution de ces domaines au sein d'une même cellule permet alors de gérer les mises à jour des flux et des cellules de manière indépendante et consistante. Cette méthode se révèle être finalement une extension de la méthode proposée par M. Grote car, lorsque l'on applique l'algorithme sur un maillage du type de ceux utilisés dans [35], on retrouve exactement le comportement de l'algorithme décrit dans cet article. Pour illustrer la méthode ARK2 ainsi définie, nous considérons ensuite une équation différentielle linéaire générique, couplée avec une discrétisation (supposée connue) d'ordre arbitraire en espace et nous n'imposons aucune contrainte sur le maillage. Une analyse approfondie de la convergence de notre méthode, toujours en adaptant la méthode de l'erreur corrigée est également donnée. Enfin, les résultats théoriques, avec en particulier l'ordre de la convergence de la méthode, sont validés sur des cas-tests académiques en dimension d'espace un et deux.

Le chapitre 6 concerne la mise en oeuvre du schéma asynchrone d'ordre deux développé dans le chapitre 5, non plus sur des cas-tests académiques pour lesquels on peut estimer la vitesse de convergence, mais sur un ensemble de cas-tests représentatifs de l'utilisation du solveur MACOPA. En effet, l'intérêt de ce chapitre est de mettre en évidence des exemples de maillages pour lesquels les temps caractéristiques associés présentent un intérêt d'utiliser un solveur asynchrone. Nous utiliserons dans ce chapitre la notion de taux d'accélération (speed-up) théorique, qui correspond au gain maximum de temps de calcul que l'on peut espérer en utilisant une intégration asynchrone en lieu et place d'une méthode synchrone. Cette estimation théorique sera également comparée à une valeur pratique, la valeur synchrone de référence étant obtenue en imposant un pas de temps uniforme dans tout le maillage dans le solveur MACOPA. Un des points d'intérêt de cette étude est également de montrer que le passage à une méthode d'intégration temporelle du second ordre, qui impose plus d'estimations et de mises à jour que la méthode du premier ordre, conserve tout de même les fondamentaux de l'approche asynchrone en préservant des gains en temps de calcul proches du speed-up théorique. Pour tous ces tests, la méthode asynchrone est appliquée pour la résolution des équations de Maxwell en dimensions deux et trois d'espace, tandis que la discrétisation spatiale utilisée est la méthode de Galerkin discontinue.

Enfin, le dernier chapitre présente une conclusion de ces travaux, en rappelant et discutant les principaux résultats obtenus dans les différents chapitres. Les perspectives d'études qui peuvent suivre ces travaux de thèse, que ce soit au niveau de l'amélioration des outils (solveur MACOPA) ou des domaines d'applications sont également discutés dans ce chapitre.



## **Première partie**

# **Schéma numérique asynchrone d'ordre un pour la discrétisation des équations de transport**



## Chapitre 2

# Étude du schéma numérique asynchrone pour les équations de transport

L'objectif de ce chapitre est de définir le concept asynchrone et d'étudier ses premières propriétés. En particulier, les idées de base qui construisent l'asynchrone seront présentées dans la première section. Pour la mise en oeuvre de la méthode, une application à l'équation de transport en dimension un de l'espace sera présentée dans la deuxième section. Le but est de pouvoir justifier mathématiquement la convergence de cette méthode. En particulier, la stabilité et la consistance de la méthode seront étudiées. La troisième section sera quant à elle consacrée à la validation numérique des résultats théoriques. Enfin, la conclusion sera présentée dans la quatrième section.

### 2.1. Définition du schéma numérique asynchrone

Cette section est consacrée à la définition du concept asynchrone et des idées de base de la méthode. Nous spécifions en particulier les éléments liés à l'asynchrone tels que le pas de temps asynchrone et l'avancement d'un système asynchrone avant de présenter le fonctionnement de l'algorithme asynchrone.

#### 2.1.1. La méthodologie asynchrone

La méthodologie asynchrone repose sur deux idées clés. Tout d'abord, l'algorithme asynchrone permet la sélection de pas de temps indépendants pour chaque cellule de maillage, de sorte que les pas de temps locaux ne portent aucune relation intégrale les uns par rapport aux autres. L'évolution temporelle du système est alors vue par l'utilisateur comme complètement aléatoire car la mise à jour des cellules ne repose sur aucune considération géométrique de type voisinage, mais plutôt sur la notion de prochain élément à mettre à jour.

Dans le cas d'une intégration classique synchrone, la suite de temps de mise à jour des cellules, notée  $(t^n)_{n \geq 1}$  est définie par un ajout continu d'un pas de temps uniforme :

$$t^{n+1} = t^n + \Delta t_{min}, \quad \text{avec} \quad \Delta t_{min} = \min_i (\Delta t_i), \quad (2.1)$$

où pour tout  $i$ , le pas de temps  $\Delta t_i$  de la cellule  $K_i$  est déterminé par la condition de stabilité locale.

Dans le cas d'une intégration asynchrone, le système saute d'une cellule à une autre en se basant sur leur temps de mise à jour. Par suite, l'évolution temporelle n'est donc plus définie d'une manière uniforme mais est spécifique à chaque cellule. Cela nous mène à introduire la suite  $(t^p)_{p \geq 1}$  des temps successifs de mise à jour des événements les plus urgents entre le temps initial  $t^0$  et le temps final  $T$  définie comme suit :

$$\begin{cases} t^1 &= \min_j(\Delta t_j) := \Delta t_i, \\ t^2 &= \min(\min_{j \neq i}(\Delta t_j), 2\Delta t_i) := \Delta t_k, \\ t^3 &= \min(\min_{j \neq \{i,k\}}(\Delta t_j), 2\Delta t_i, 2\Delta t_k), \\ &\vdots \end{cases}$$

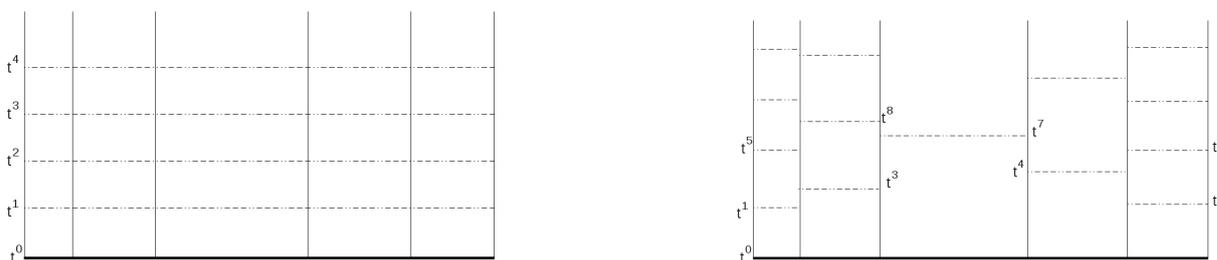


FIGURE 2.1 – suite des temps de mise à jour de cinq cellules, dans le cas synchrone (à gauche) et dans le cas asynchrone (à droite).

La deuxième idée de base de l'asynchrone est la séparation du terme source "local" et des échanges par flux. En effet, le terme source local dépend uniquement de l'état de la cellule courante tandis que les flux dépendent également des états de leurs voisins. En particulier, le schéma asynchrone a été défini en considérant des remises à jour indépendantes pour chaque flux d'interface.

### 2.1.2. L'algorithme asynchrone

Le concept de l'intégration asynchrone en temps peut être résumé en trois phases principales :

1. La phase d'initialisation : elle consiste à initialiser toutes les densités ainsi que tous les flux numériques à l'instant initial, c'est à dire le temps de démarrage de la simulation.
2. La deuxième phase consiste à appliquer de façon continue les étapes suivantes jusqu'à ce que l'horloge globale de simulation ait avancé au-delà du temps de fin de simulation attendu
  - (a) Chercher la cellule la plus urgente à mettre à jour. A chaque cellule de maillage correspond un pas de temps local calculé en utilisant la condition CFL locale. Au départ de la simulation, cette étape consiste à chercher la cellule qui possède le plus petit pas de temps. Mais après ce premier pas de temps, le système évolue en choisissant le temps de mise à jour le plus proche ou le plus urgent ; voir Fig.2.1.
  - (b) Mettre à jour les cellules nécessaires, dont la cellule courante. Notons que le nombre de cellules impactées dépend de la méthode utilisée pour la discrétisation en espace. Par exemple, dans le cadre d'une intégration en espace par flux décentré, seules les cellules nécessaires

pour la mise à jour des flux sortants sont concernées. Dans tous les cas, le nombre de cellules mis à jour dans cette étape est relativement petit.

- (c) Calculer les nouvelles valeurs des flux. Dans le cadre d'une discrétisation centrée en espace, tous les flux sont mis à jour, tandis que dans le cadre d'une intégration décentrée de type upwind, il suffit de mettre à jour les flux sortants.
  - (d) Calculer le prochain temps de mise à jour de la cellule courante, et seulement la cellule courante, en utilisant la condition CFL locale.
3. La phase de synchronisation consiste à calculer la solution à l'instant final pour toutes les cellules du maillage. Elle est effectuée par une interpolation des valeurs de la solution dans chaque cellule à son instant final. Cette interpolation doit être d'ordre suffisant pour conserver l'ordre général de la méthode.

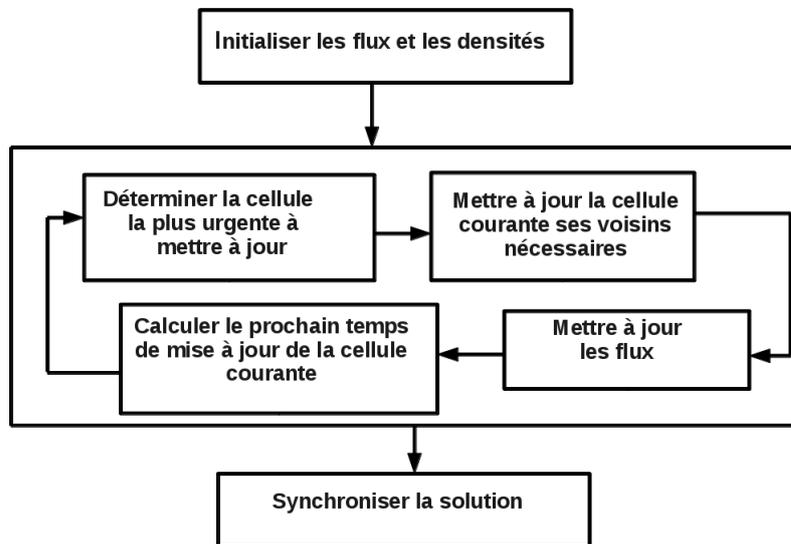


FIGURE 2.2 – L'algorithme asynchrone

Après avoir défini le concept de l'asynchrone, l'étude de ce schéma s'avère nécessaire pour justifier son efficacité et sa performance. Ce sera l'objet de la section suivante.

## 2.2. Application de la méthode asynchrone pour la discrétisation de l'équation de transport en dimension un d'espace

Nous présentons dans cette section l'algorithme asynchrone appliqué à l'équation de transport en dimension un. Le but est d'étudier les premières propriétés de la méthode numérique asynchrone. L'intérêt de cette étude est à terme de pouvoir justifier mathématiquement la convergence de la méthode. En particulier, la consistance et la stabilité de cette approche seront traitées. Le choix de l'équation de transport en 1D, qui est un cas relativement simple, a pour objectif d'illustrer de manière simple les idées de l'asynchrone mais également de pouvoir contrôler les éventuelles complexités créées à cause de l'asynchrone. Il

s'agit d'un bon exemple pour exposer comment gérer ces difficultés et donner des pistes d'applications vers des exemples plus complexes.

Considérons l'équation de transport (ou d'advection) écrite sous forme conservative :

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(v(x)u) = 0, & (x, t) \in [0, L] \times [0, T] \\ u(x, 0) = u_0(x), & x \in [0, L], \end{cases} \quad (2.2)$$

où  $[0, L] \subset \mathbb{R}$  est un domaine périodique de période  $L$  avec  $(u(x, t) = u(x + L, t))$ ,  $T$  est le temps final et  $v$  est la vitesse de propagation, supposée connue et régulière. La donnée initiale  $u_0$  est supposée régulière et le terme source est nul.

### 2.2.1. Discrétisation en espace de l'équation d'advection

Pour la discrétisation en espace, la méthode des volumes finis a été choisie. Cette méthode conservative est une méthode particulièrement bien adaptée à la discrétisation spatiale des lois de conservation, et c'est notamment le cas pour l'équation de transport.

Commençons par donner un maillage de l'intervalle  $[0, L]$  par des cellules  $C_i$ , qui sont des intervalles deux à deux disjoints :  $C_i = (x_{i-1/2}, x_{i+1/2})$ ,  $1 \leq i \leq N$ , où  $N$  est le nombre des cellules et on cherche à approcher la solution  $u$  sur chaque cellule. L'idée de la méthode des volumes finis est d'intégrer l'équation (2.2) sur une cellule, ce qui donne :

$$\int_{C_i} \left( \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(v(x)u) \right) dx = 0 \quad (2.3)$$

où  $u_i(t) = \frac{1}{|C_i|} \int_{C_i} u(x, t) dx$  est la valeur moyenne de  $u$  sur la cellule  $C_i$ .

$$|C_i| \frac{du_i}{dt} + v(x_{i+1/2})u(x_{i+1/2}, t) - v(x_{i-1/2})u(x_{i-1/2}, t) = 0.$$

Les flux exacts au travers des interfaces sont définis par :

1.  $\tilde{F}_i = v(x_{i+1/2})u(x_{i+1/2}, t)$ ,
2.  $\tilde{F}_{i-1} = v(x_{i-1/2})u(x_{i-1/2}, t)$ .

L'idée de la méthode numérique est donc d'approcher les flux exacts par des flux numériques. Il existe plusieurs choix pour le flux numérique. Nous pouvons citer par exemple la méthode décentrée ou les schémas de type MUSCL. Pour des raisons de simplification, et comme nous nous intéressons à une méthode du premier ordre, nous avons choisi d'utiliser le schéma décentré upwind et de ne pas utiliser des discrétisations spatiales d'ordre plus élevé. Dans ce cas, le flux numérique  $F_i$  est donné par

$$F_i(t) := u_i(t) \max(0, v_{i+1/2}) + u_{i+1}(t) \min(0, v_{i+1/2}). \quad (2.4)$$

Après avoir défini le schéma en espace, il reste à rajouter une discrétisation "asynchrone" en temps.

### 2.2.2. Discrétisation en temps : le schéma asynchrone

Pour la discrétisation temporelle, nous avons choisi de commencer par le schéma d'Euler explicite. L'intégration numérique explicite fait nécessairement intervenir une condition de stabilité très restrictive

sur le pas de temps qui se traduit par une inégalité de la forme :

$$0 \leq |v_{i+1/2}| \frac{\Delta t_i}{\Delta x_i} \leq 1, \quad (2.5)$$

avec  $\Delta x_i := |C_i| = x_{i+1/2} - x_{i-1/2}$ ,  $\Delta t_i$  le pas de temps de la cellule  $C_i$  et  $v_{i+1/2}$  la vitesse au point  $x_{i+1/2} := (i + 1/2)\Delta x_i$ . Une telle condition de stabilité s'appelle condition de Courant-Friedrich-Levy (CFL).

Dans le cas d'une intégration explicite classique, la condition CFL impose donc une borne supérieure au pas de temps d'intégration  $\Delta t$  pour tout le domaine de calcul :

$$\Delta t = \min_{1 \leq i \leq N} \Delta t_i.$$

Le schéma, que l'on qualifiera de synchrone classique, s'écrit alors :

$$\begin{cases} u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x_i} (F_i^n - F_{i-1}^n), & \forall 1 \leq i \leq N, \\ t^{n+1} = t^n + \Delta t & \forall n \geq 0, \end{cases}$$

où  $u_i^n$  est la valeur de la densité au centre de la cellule  $C_i$ , au point  $x_i = i\Delta x_i$  et à l'instant  $t^n$ .  $F_i^n$  représente le flux numérique à l'instant  $t^n$  défini par :

$$F_i^n := u_i^n \max(0, v_{i+1/2}) + u_{i+1}^n \min(0, v_{i+1/2}).$$

### 2.2.2.1. L'algorithme asynchrone

Pour chaque cellule  $C_i$  de maillage on définit :

1. La suite des temps de mise à jour de la cellule, lesquels sont des multiples de son propre pas de temps :  $(t_i^k)_k = (k\Delta t_i)_k$ .
2.  $t_{cur}$ , le temps courant de la simulation,
3.  $t_i^{pcd}$  le temps de la dernière mise à jour de la cellule avec  $t_i^{pcd} = \max(r_i\Delta t_i, r_{i-1}\Delta t_{i-1})$  où  $r_i$  et  $r_{i-1}$  sont les deux entiers les plus grands possibles tels que  $r_i\Delta t_i < t_{cur}$  et  $r_{i-1}\Delta t_{i-1} < t_{cur}$ .

L'algorithme asynchrone, tel qu'il a été défini dans la section précédente, appliqué à l'équation de transport est le suivant :

1. La première phase : initialisation de toutes les densités et des flux :

$$u_i^{t^0} = u_0(x_i, t^0) \quad \text{et} \quad F_i^{t^0} = u_i^{t^0} \max(0, v_{i+1/2}) + u_{i+1}^{t^0} \min(0, v_{i+1/2}), \quad \text{pour tout } i,$$

2. La deuxième phase : tant que le temps final de la simulation  $T$  n'est pas atteint, on exécute les étapes suivantes de la méthode asynchrone :

- (a) Chercher la cellule la plus urgente à mettre à jour. On suppose, par exemple, que c'est le tour d'une cellule  $C_i$  à un instant  $t_i^{r_i} = r_i\Delta t_i < T$ .

- (b) Mettre à jour les cellules qui contribuent à la mise à jour du flux  $F_i$ . Dans le cas du schéma upwind, la cellule courante  $C_i$  et sa voisine  $C_{i+1}$  sont concernées :

$$u_i^{t_i^{r_i}} = u_i^{t_i^{pcd}} - \frac{t_i^{r_i} - t_i^{pcd}}{\Delta x_i} \left( F_i^{t_i^{r_i-1}} - F_{i-1}^{t_{i-1}^{r_{i-1}}} \right)$$

et,

$$u_{i+1}^{t_i^{r_i}} = u_{i+1}^{t_{i+1}^{pcd}} - \frac{t_i^{r_i} - t_{i+1}^{pcd}}{\Delta x_{i+1}} \left( F_{i+1}^{t_{i+1}^{r_{i+1}}} - F_i^{t_i^{r_i-1}} \right)$$

où  $t_i^{r_{i-1}}$ ,  $t_{i+1}^{r_{i+1}}$  et  $t_{i-1}^{r_{i-1}}$  sont respectivement les derniers temps effectifs de mise à jour des cellules  $C_i$ ,  $C_{i+1}$  et  $C_{i-1}$ , juste avant l'instant  $t_i^{r_i}$ . Donc  $t_{i+1}^{r_{i+1}} \leq t_i^{r_i}$  et  $t_{i-1}^{r_{i-1}} \leq t_i^{r_i}$ .

- (c) Mise à jour de la valeur du flux  $F_i$  :

$$F_i^{t_i^{r_i}} = u_i^{t_i^{r_i}} \max(0, v_{i+1/2}) + u_{i+1}^{t_i^{r_i}} \min(0, v_{i+1/2}).$$

- (d) Faire avancer le temps de la dernière mise à jour des cellules modifiées. Autrement dit, on suppose que  $t_i^{pcd} = t_{i+1}^{pcd} = t_i^{r_i}$
- (e) Calculer le temps de la prochaine mise à jour de la cellule courante :  $t_i^{r_{i+1}} = t_i^{r_i} + \Delta t_i$ .

3. La troisième phase est la phase de la synchronisation de la cellule à l'instant final de la simulation. Une interpolation d'ordre un entre els deux derniers états est appliquée à toutes les cellules pour calculer la solution globale au temps final.

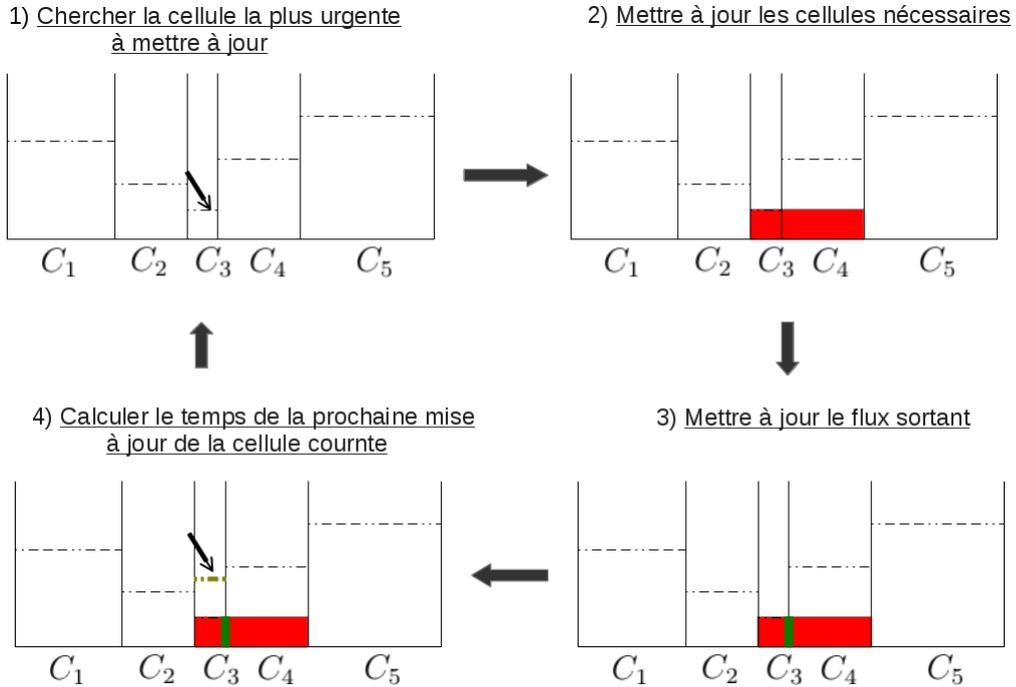


FIGURE 2.3 – La deuxième phase de l'algorithme asynchrone appliqué à l'équation de transport 1D, schéma upwind, exemple avec cinq cellules :  $C_1, \dots, C_5$

**Remarques 1.** Notons que :

1. Le schéma asynchrone a été défini en considérant des remises à jour indépendantes pour chaque flux d'interface. En particulier, à chaque pas d'avancement du système, un seul flux est mis à jour (voir l'étape (c) de l'algorithme précédent).
2. La cellule courante et sa voisine ont été mises à jour à l'étape (b), mais seul le prochain temps de la cellule courante est calculé. En fait, la cellule voisine a été mise à jour pour déterminer sa contribution à la mise à jour du flux.
3. À l'étape de la synchronisation de la cellule, une interpolation d'ordre un est suffisante car les schéma upwind et Euler explicite sont également d'ordre un.

### 2.3. Propriétés du schéma asynchrone

Rappelons que dans le cadre d'une intégration explicite classique, la solution au temps courant dépend des instants précédents. Dans le cas asynchrone, l'avancement du système est aléatoire. La solution à un instant donné dépend de l'état de la cellule courante à cet instant, mais également de sa voisine dont l'état a été calculé à un autre instant. Cette dernière dépend elle-même de sa voisine qui a été mise à jour à un instant encore différent, et ainsi de suite. La solution asynchrone dépend donc non seulement de la solution à l'instant local précédent mais également de tout ce qui s'est passé avant. Par conséquent, les estimations classiques garantissant l'ordre de précision (la consistance) et la stabilité ne sont pas directement compatibles avec l'aspect asynchrone. Il devient donc nécessaire d'adapter la notion de la consistance à ce concept ainsi que de choisir une norme adaptée au cadre asynchrone pour pouvoir justifier la stabilité de ce type de méthodes. L'idée est donc de pouvoir exprimer la solution à un temps fixé en fonction d'un temps de synchronisation commun pour toutes les cellules, notamment l'instant initial.

L'algorithme asynchrone permet de connaître la solution à l'instant initial où la solution est initialisée et à l'instant final où la solution est synchronisée. Entre ces deux instants, le système évolue d'une façon qui semble aléatoire. D'où l'idée de traiter les propriétés de la méthode asynchrone entre deux instants de synchronisation. Autrement dit, nous supposons que la solution est synchronisée aux instants  $t^q$  et  $t^{q+1} = t^q + \Delta t$  avec  $\Delta t = \max_i(\Delta t_i)$ . Cette hypothèse n'a pas d'influence sur l'aspect asynchrone (voir Fig. 2.4), son objectif est d'éviter de se ramener à l'instant initial et donc d'améliorer la lisibilité pour la suite.

Soit donc,  $(t_q^n)_{1 \leq n \leq N_q}$  la suite des temps successifs de mise à jour des événements les plus urgents entre les instants  $t^q$  et  $t^{q+1}$ , où  $N_q$  est le nombre de pas de temps entre  $t^q$  et  $t^{q+1}$ . La suite  $(t_{i,q}^k)_k$  est la suite des temps des mises à jour d'une cellule  $C_i$  et sont des multiples du pas de temps local de la cellule entre  $t^q$  et  $t^{q+1}$  définie par :  $t_{i,q}^k = t^q + k\Delta t_i$ . Toutes ces notations sont résumées dans la Fig. 2.4.

Notons que pour l'équation de transport, l'aspect asynchrone est dû soit à la variation du maillage soit à la variation de la vitesse ou aux deux ensemble. Pour éviter les problèmes liés au maillage, nous supposons dans cette section que le maillage est constant et que la vitesse est variable. L'aspect asynchrone provient donc de la variation de la vitesse. Le cas du maillage variable sera traité en détails dans le cas multi-dimensionnel dans le chapitre suivant.

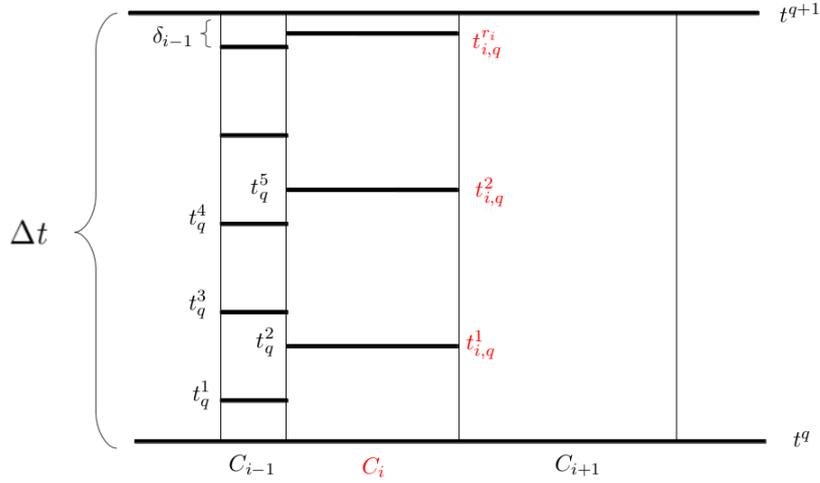


FIGURE 2.4 – Évolution du système entre les deux instants  $t^q$  et  $t^{q+1}$ , exemple de trois cellules.

### 2.3.1. Stabilité du schéma asynchrone

Comme expliqué au début de la section, la solution asynchrone ne dépend pas uniquement de l'instant précédent, mais de tout ce qui s'est passé avant. Par suite nous introduisons la norme infinie asynchrone  $\|\cdot\|_{asyn,\infty}$  définie par :

$$\|u^{t^n}\|_{asyn,\infty} := \sup_i \sup_{p \leq n} |u_i^{t^p}|. \quad (2.6)$$

**Proposition 1.** *Supposons que la vitesse  $v$  est de classe  $\mathcal{C}^2$ , alors sous la condition CFL locale définie pour tout  $i$  par :*

$$0 \leq |v_{i+1/2}| \frac{\Delta t_i}{\Delta x} \leq 1, \quad (2.7)$$

le schéma asynchrone est stable pour la norme infinie asynchrone. En particulier, on a

$$\|u^{t^{q+1}}\|_{asyn,\infty} \leq \exp\left(\Delta t \left\| \frac{\partial v}{\partial x} \right\|_{\infty}\right) \|u^{t^q}\|_{asyn,\infty}, \quad (2.8)$$

avec  $\Delta t := \max_i(\Delta t_i)$ .

*Démonstration.* Pour prouver la formule (2.8), nous commençons par montrer que pour tout  $i$ ,

$$\|u^{t_{i,q}^{r_i}}\|_{asyn,\infty} \leq \exp\left(r_i \Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_{\infty}\right) \|u^{t^q}\|_{asyn,\infty}, \quad (2.9)$$

avec  $t_{i,q}^{r_i} = t^q + r_i \Delta t_i$ ,  $r_i$  un entier.

Nous montrons alors la formule (2.9) par récurrence.

1. Soit  $t_{i,q}^1 = t^q + \min_j(\Delta t_j) := t^q + \Delta t_i$ .

(a) Supposons dans un premier temps que la vitesse  $v$  ne change pas de signe. Deux cas se présentent :

i. la vitesse  $v$  est positive : pour tout  $x$ ,  $v(x) \geq 0$  alors :

$$\begin{aligned} u_i^{t_i, q} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x} (F_i^{t^q} - F_{i-1}^{t^q}) \\ &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x} v_{i+1/2} (u_i^{t^q} - u_{i-1}^{t^q}) - \frac{\Delta t_i}{\Delta x} (v_{i+1/2} - v_{i-1/2}) u_{i-1}^{t^q}. \end{aligned}$$

De même

$$u_{i+1}^{t_i, q} = u_{i+1}^{t^q} - \frac{\Delta t_i}{\Delta x} v_{i+3/2} (u_{i+1}^{t^q} - u_i^{t^q}) - \frac{\Delta t_i}{\Delta x} (v_{i+3/2} - v_{i+1/2}) u_i^{t^q}.$$

ii. la vitesse  $v$  est négative : pour tout  $x$ ,  $v(x) \leq 0$  alors :

$$\begin{aligned} u_i^{t_i, q} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x} (v_{i+1/2} u_{i+1}^{t^q} - v_{i-1/2} u_i^{t^q}) \\ &= \left(1 - \frac{\Delta t_i}{\Delta x} |v_{i-1/2}|\right) u_i^{t^q} + \frac{\Delta t_i}{\Delta x} |v_{i-1/2}| u_{i+1}^{t^q} \\ &\quad - \frac{\Delta t_i}{\Delta x} (v_{i+1/2} - v_{i-1/2}) u_{i+1}^{t^q}. \end{aligned}$$

Par suite, pour une vitesse  $v$  de signe constant et sous la condition CFL locale, on a

$$\begin{aligned} \|u^{t_i, q}\|_{asyn, \infty} &\leq \left(1 + \Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_{\infty}\right) \|u^{t^q}\|_{asyn, \infty} \\ &\leq \exp\left(\Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_{\infty}\right) \|u^{t^q}\|_{asyn, \infty}. \end{aligned}$$

(b) Supposons maintenant que la vitesse  $v$  change de signe. On distingue deux cas :

i. Si  $v_{i-1/2} \leq 0$  et  $v_{i+1/2} \geq 0$  alors il existe  $c \in ]x_{i-1/2}, x_{i+1/2}[$  tel que

$$\begin{aligned} u_i^{t_i, q} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x} (v_{i+1/2} u_i^{t^q} - v_{i-1/2} u_i^{t^q}) \\ &= u_i^{t^q} - \Delta t_i \frac{\partial v(c)}{\partial x} u_i^{t^q}. \end{aligned}$$

Par suite

$$|u_i^{t_i, q}| \leq \left(1 + \Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_{\infty}\right) \|u^{t^q}\|_{asyn, \infty}.$$

ii. Si  $v_{i-1/2} \geq 0$  et  $v_{i+1/2} \leq 0$  alors

$$\begin{aligned} u_i^{t_i, q} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x} (v_{i+1/2} u_{i+1}^{t^q} - v_{i-1/2} u_{i-1}^{t^q}) \\ &:= u_i^{t^q} + \alpha_{i+1} u_{i+1}^{t^q} + \alpha_{i-1} u_{i-1}^{t^q}. \end{aligned}$$

La vitesse  $v$  est supposée suffisamment régulière (de classe  $\mathcal{C}^1$ ). De plus  $v_{i-1/2} \geq 0$  et  $v_{i+1/2} \leq 0$ , alors il existe  $x_c \in [x_{i-1/2}, x_{i+1/2}]$  tel que  $v(x_c) = 0$ . Par suite il existe un

nombre  $\theta_i$ , avec  $|\theta_i| < 1$  tel que

$$v_{i+1/2} = (x_{i+1/2} - x_c) \frac{\partial v}{\partial x} (x_c + \theta_i(x_{i+1/2} - x_c)).$$

De même, il existe  $\theta_i^1 \in ]0, 1[$  tel que

$$v_{i-1/2} = (x_c - x_{i+1/2}) \frac{\partial v}{\partial x} (x_{i-1/2} + \theta_i^1(x_c - x_{i-1/2})).$$

Ce qui donne

$$\begin{aligned} 1 + |\alpha_{i+1}| + |\alpha_{i-1}| &\leq 1 + \Delta t_i \frac{x_{i+1/2} - x_c}{\Delta x} \|\partial_x v\|_\infty + \Delta t_i \frac{x_c - x_{i-1/2}}{\Delta x} \|\partial_x v\|_\infty \\ &\leq 1 + \Delta t_i \|\partial_x v\|_\infty. \end{aligned}$$

La formule (2.9) est ainsi vérifiée.

2. On suppose maintenant que la formule (2.9) est vraie jusqu'à un instant  $t_{i,q}^n < t^{q+1}$ . Soient  $t_{i,q}^{n+1} = t_{i,q}^{r_i+1} := t^q + (r_i + 1)\Delta t_i < t^{q+1}$  et  $t^q + r_{i-1}\Delta t_{i-1} < t_{i,q}^{r_i+1}$ .

- (a) Nous traitons dans un premier temps le cas d'une vitesse  $v$  qui ne change pas de signe et nous supposons de plus qu'elle est positive (le cas d'une vitesse négative se traite de la même manière). Soit  $m = E\left(\frac{\Delta t_i}{\Delta t_{i-1}}\right) \geq 0$ , avec  $E$  la partie entière. Alors,  $t_{i,q}^{r_i} \leq t_{i-1,q}^{(r_{i-1}-m)} < \dots < t_{i-1,q}^{r_{i-1}} < t_{i,q}^{r_i+1}$ .  $\delta t_i^1 := t_{i-1,q}^{(r_{i-1}-m)} - t_{i,q}^{r_i}$  et  $\delta t_i^2 := t_{i,q}^{r_i+1} - t_{i-1,q}^{r_{i-1}}$ .

$$\begin{aligned} u_i^{t_{i,q}^{r_i+1}} &= u_i^{t_{i-1,q}^{r_{i-1}}} - \frac{\delta t_i^2}{\Delta x} \left( F_i^{t_{i,q}^{r_i}} - F_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) \\ &= u_i^{t_{i,q}^{r_i}} - \frac{\delta t_i^1}{\Delta x} \left( F_i^{t_{i,q}^{r_i}} - F_{i-1}^{t_{i-1,q}^{(r_{i-1}-m-1)}} \right) - \frac{\Delta t_{i-1}}{\Delta x} \sum_{j=1}^m \left( F_i^{t_{i,q}^{r_i}} - F_{i-1}^{t_{i-1,q}^{(r_{i-1}-j)}} \right) \\ &\quad - \frac{\delta t_i^2}{\Delta x} \left( F_i^{t_{i,q}^{r_i}} - F_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) \\ &= \left( 1 - \frac{\Delta t_i}{\Delta x} v_{i+1/2} \right) u_i^{t_{i,q}^{r_i}} + \frac{\delta t_i^{r_i}}{\Delta x} v_{i+1/2} u_{i-1}^{t_{i-1,q}^{(r_{i-1}-m-1)}} \\ &\quad + \frac{\Delta t_{i-1}}{\Delta x} v_{i+1/2} \sum_{j=1}^m u_{i-1}^{t_{i-1,q}^{(r_{i-1}-j)}} + \frac{\delta t_i^2}{\Delta x} v_{i+1/2} u_{i-1}^{t_{i-1,q}^{r_{i-1}}} - \delta t_i^1 \frac{\partial v}{\partial x}(c) u_{i-1}^{t_{i-1,q}^{(r_{i-1}-m-1)}} \\ &\quad - \Delta t_{i-1} \frac{\partial v}{\partial x}(c) \sum_{j=1}^m u_{i-1}^{t_{i-1,q}^{(r_{i-1}-j)}} - \delta t_i^2 \frac{\partial v}{\partial x}(c) u_{i-1}^{t_{i-1,q}^{r_{i-1}}}. \end{aligned}$$

Ce qui donne,

$$\begin{aligned}
|u_i^{t_i^{r_i+1}}| &\leq \left(1 + \delta t_i^1 \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t_i^{r_i}}\|_{asyn,\infty} + \delta t_i^2 \left\| \frac{\partial v}{\partial x} \right\|_\infty \|u^{t_{i-1}^{r_{i-1}}}\|_{asyn,\infty} \\
&+ \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty \sum_{j=1}^m \|u^{t_{i-1}^{(r_{i-1}-j)}}\|_{asyn,\infty} \\
&\leq \exp\left((r_i \Delta t_i + \delta t_i^1) \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t_i^q}\|_{asyn,\infty} \\
&+ \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty \sum_{j=1}^m \exp\left((r_{i-1} - j) \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t_i^q}\|_{asyn,\infty} \\
&+ \delta t_i^2 \left\| \frac{\partial v}{\partial x} \right\|_\infty \exp\left(r_{i-1} \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t_i^q}\|_{asyn,\infty}.
\end{aligned}$$

Par suite, nous avons le résultat suivant :

$$\begin{aligned}
|u_i^{t_i^{r_i+1}}| &\leq \left(1 + \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \exp\left[(r_{i-1} - m) \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty\right] \|u^{t_i^q}\|_{asyn,\infty} \\
&+ \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty \sum_{j=1}^{m-1} \exp\left((r_{i-1} - j) \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t_i^q}\|_{asyn,\infty} \\
&+ \delta t_i^2 \left\| \frac{\partial v}{\partial x} \right\|_\infty \exp\left(r_{i-1} \Delta t_{i-1} \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t_i^q}\|_{asyn,\infty} \\
&\vdots \\
&\leq \exp\left((r_i + 1) \Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t_i^q}\|_{asyn,\infty},
\end{aligned}$$

ce qui prouve la formule (2.9).

(b) Si nous supposons que la vitesse change de signe et en particulier,

i. si  $v_{i-1/2} \leq 0$  et  $v_{i+1/2} \geq 0$  alors

$$\begin{aligned}
u_i^{t_i^{r_i+1}} &= u_i^{t_i^{r_i}} - \frac{\Delta t_i}{\Delta x} \left( v_{i+1/2} u_i^{t_i^{r_i}} - v_{i-1/2} u_i^{t_{i-1}^{r_{i-1}}} \right) \\
&:= \left(1 - \frac{\Delta t_i}{\Delta x} v_{i+1/2}\right) u_i^{t_i^{r_i}} + \frac{\Delta t_i}{\Delta x} v_{i+1/2} u_i^{t_{i-1}^{r_{i-1}}} \\
&\quad - \frac{\Delta t_i}{\Delta x} (v_{i+1/2} - v_{i-1/2}) u_i^{t_{i-1}^{r_{i-1}}}.
\end{aligned}$$

Par suite

$$\|u^{t_i^{r_i+1}}\|_{asyn,\infty} \leq (1 + \Delta t_i \|\partial_x v\|_\infty) \|u^{t_i^{r_i}}\|_{asyn,\infty}.$$

ii. Si  $v_{i-1/2} \geq 0$  et  $v_{i+1/2} \leq 0$  alors

$$u_i^{t_i^{r_i+1}} = u_i^{t_i^{r_i}} - \frac{\Delta t_i}{\Delta x} \left( v_{i+1/2} u_{i+1}^{t_i^{r_i}} - v_{i-1/2} u_i^{t_{i-1}^{r_{i-1}}} \right).$$

En reprenant le raisonnement à l'instant  $t_q^1$ , il est facile de voir que

$$\|u^{t_{i,q}^{r_i+1}}\|_{asyn,\infty} \leq (1 + \Delta t_i \|\partial_x v\|_\infty) \|u^{t_{i,q}^{r_i}}\|_{asyn,\infty}.$$

La formule (2.9) est donc vérifiée.

Revenons maintenant à la formule (2.8). On suppose que la solution est synchronisée à l'instant  $t^{q+1}$  et que pour tout  $i$  on a

$$r_i \Delta t_i + \delta_i = \Delta t.$$

D'après la formule (2.9) on a pour tout  $i$

$$\begin{aligned} |u_i^{t^{q+1}}| &\leq \exp\left((r_i \Delta t_i + \delta_i) \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t^q}\|_{asyn,\infty} \\ &\leq \exp\left(\Delta t \left\| \frac{\partial v}{\partial x} \right\|_\infty\right) \|u^{t^q}\|_{asyn,\infty}. \end{aligned}$$

Ce qui achève la démonstration. □

### 2.3.2. Consistance du schéma asynchrone

Soit  $\tilde{u}_i^{t_q^n} := u(x_i, t_q^n)$  la solution exacte de l'équation d'advection (2.2), supposée de classe  $\mathcal{C}^2$ .

**Proposition 2.** *Sous la condition CFL locale (2.5), le schéma asynchrone est consistant au premier ordre en temps et en espace. En particulier, pour tout  $i$  nous avons la formule suivante :*

$$(Ec)_i^{t_{i,q}^{r_i}} = \frac{\tilde{u}_i^{t_{i,q}^{r_i}} - \tilde{u}_i^{t_i^q}}{t_{i,q}^{r_i} - t_i^q} + \left( \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} + \Delta x \varepsilon_i^{r_i} \right) \quad (2.10)$$

avec  $\varepsilon_i^{r_i}$  est borné pour tout  $i$  et  $(Ec)_i$  est l'erreur de consistance locale.

*Démonstration.* Commençons par montrer la formule (2.10) par récurrence.

Soit  $t_{i,q}^1 = t^q + \min_j (\Delta t_j) := t^q + \Delta t_i$ . Nous traitons dans un premier temps le cas d'une vitesse  $v$  à signe constant, et à titre d'exemple, la vitesse est supposée négative (le cas  $v$  positive se traite de la même manière). Par définition, l'erreur de consistance  $Ec$  est donnée par l'expression suivante :

$$\begin{aligned} (Ec)_i^{t_{i,q}^1} &= \frac{\tilde{u}_i^{t_{i,q}^1} - \tilde{u}_i^{t_i^q}}{\Delta t_i} + \frac{1}{\Delta x} \left( \tilde{F}_i^{t_i^q} - \tilde{F}_{i-1}^{t_i^q} \right) \\ &= \frac{\tilde{u}_i^{t_{i,q}^1} - \tilde{u}_i^{t_i^q}}{\Delta t_i} + \frac{1}{\Delta x} \left( v_{i+1/2} \tilde{u}_{i+1}^{t_i^q} - v_{i-1/2} \tilde{u}_i^{t_i^q} \right). \end{aligned}$$

D'après la formule de Taylor-Lagrange, nous avons le résultat suivant

$$v_{i+1/2} \tilde{u}_{i+1}^{t_i^q} - v_{i-1/2} \tilde{u}_i^{t_i^q} = \Delta x \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} + (\Delta x)^2 \varepsilon_i^1,$$

$\varepsilon_i^1$  est borné. Alors

$$(Ec)_i^{t_i,q} = \frac{\tilde{u}_i^{t_i,q} - \tilde{u}_i^{t_i,q}}{\Delta t_i} + \left( \frac{\partial(\tilde{u}_i^{t_i,q} v_i)}{\partial x} + \Delta x \varepsilon_i^1 \right).$$

Supposons que  $v$  change de signe et que  $v_{i+1/2} \leq 0$  et  $v_{i-1/2} \geq 0$ . Alors

$$(Ec)_i^{t_i,q} = \frac{\tilde{u}_i^{t_i,q} - \tilde{u}_i^{t_i,q}}{\Delta t_i} + \frac{1}{\Delta x} (\tilde{F}_i^{t_i,q} - \tilde{F}_{i-1}^{t_i,q}) = \frac{\tilde{u}_i^{t_i,q} - \tilde{u}_i^{t_i,q}}{\Delta t_i} + \frac{1}{\Delta x} (v_{i+1/2} \tilde{u}_{i+1}^{t_i,q} - v_{i-1/2} \tilde{u}_{i-1}^{t_i,q}).$$

or, nous avons

$$v_{i+1/2} \tilde{u}_{i+1}^{t_i,q} - v_{i-1/2} \tilde{u}_{i-1}^{t_i,q} = \Delta x \frac{\partial(\tilde{u}_i^{t_i,q} v_i)}{\partial x} + \Delta x v_i \frac{\partial \tilde{u}_i^{t_i,q}}{\partial x} + (\Delta x)^2 \tilde{\varepsilon}_i^1,$$

avec  $\tilde{\varepsilon}_i^1$  est le reste de la formule de Taylor-Lagrange d'ordre 2. Puisque  $v_{i+1/2} \leq 0$  et  $v_{i-1/2} \geq 0$ , alors d'après le théorème des valeurs intermédiaires, il existe  $x_c \in ]x_{i-1/2}, x_{i+1/2}[$  tel que  $v(x_c) = 0$ . En utilisant un développement de Taylor, nous avons la formule suivante

$$v_i = (x_c - x_i) \frac{\partial v}{\partial x} \left( \frac{\partial(\tilde{u}_{i-1}^{t_i,q} v_{i-1})}{\partial x} \theta_c \right), \quad \theta_c \text{ entre } x_c \text{ et } x_i.$$

Par suite

$$(Ec)_i^{t_i,q} = \frac{\tilde{u}_i^{t_i,q} - \tilde{u}_i^{t_i,q}}{\Delta t_i} + \left( \frac{\partial(\tilde{u}_i^{t_i,q} v_i)}{\partial x} + \Delta x \varepsilon_i^1 \right), \quad \text{avec } |\varepsilon_i^1| \leq \|\tilde{\varepsilon}_i^1\|_\infty + \|\partial_x v\|_\infty.$$

De même,

$$(Ec)_{i+1}^{t_i,q} = \frac{\tilde{u}_{i+1}^{t_i,q} - \tilde{u}_{i+1}^{t_i,q}}{\Delta t_i} + \left( \frac{\partial(\tilde{u}_i^{t_i,q} v_i)}{\partial x} + \Delta x \varepsilon_{i+1}^1 \right).$$

En effet, si  $v_{i+3/2} \leq 0$  alors il suffit d'appliquer le résultat du premier cas ( $v$  négative). Sinon voir le cas suivant qui consiste à supposer que  $v_{i+1/2} \geq 0$  et  $v_{i-1/2} \leq 0$ . Dans ce cas,

$$(Ec)_i^{t_i,q} = \frac{\tilde{u}_i^{t_i,q} - \tilde{u}_i^{t_i,q}}{\Delta t_i} + \frac{1}{\Delta x} (\tilde{F}_i^{t_i,q} - \tilde{F}_{i-1}^{t_i,q}) = \frac{\tilde{u}_i^{t_i,q} - \tilde{u}_i^{t_i,q}}{\Delta t_i} + \frac{1}{\Delta x} (v_{i+1/2} - v_{i-1/2}) \tilde{u}_i^{t_i,q}.$$

Il existe  $x_c \in ]x_{i-1/2}, x_{i+1/2}[$  tel que  $v(x_c) = 0$ . Soient  $\delta_1 = x_c - x_{i-1/2}$  et  $\delta_2 = x_{i+1/2} - x_c$ , on a

$$\begin{aligned} v_{i+1/2} - v_{i-1/2} &= \delta_2 \frac{\partial v}{\partial x}(x_c) + \frac{(\delta_2)^2}{2} \frac{\partial^2 v}{\partial x^2}(\theta_i^+) + \delta_1 \frac{\partial v}{\partial x}(x_c) - \frac{(\delta_1)^2}{2} \frac{\partial^2 v}{\partial x^2}(\theta_i^-) \\ &= \Delta x \frac{\partial v}{\partial x}(x_c) + \frac{(\delta_2)^2}{2} \frac{\partial^2 v}{\partial x^2}(\theta_i^+) - \frac{(\delta_1)^2}{2} \frac{\partial^2 v}{\partial x^2}(\theta_i^-). \end{aligned}$$

De plus

$$\tilde{u}_i^{t_i,q} = \tilde{u}^{t_i,q}(x_c) + (x_c - x_i) \frac{\partial \tilde{u}^{t_i,q}}{\partial x}(\theta_1),$$

ce qui donne l'égalité suivante :

$$\begin{aligned}
(v_{i+1/2} - v_{i-1/2}) \tilde{u}_i^{t^q} &= \Delta x \frac{\partial v}{\partial x}(x_c) \tilde{u}^{t^q}(x_c) + \left( \frac{(\delta_2)^2}{2} \frac{\partial v}{\partial x}(\theta_i^+) - \frac{(\delta_1)^2}{2} \frac{\partial v}{\partial x}(\theta_i^-) \right) \tilde{u}^{t^q}(x_c) \\
&+ (x_c - x_i) \left( \frac{(\delta_2)^2}{2} \frac{\partial v}{\partial x}(\theta_i^+) - \frac{(\delta_1)^2}{2} \frac{\partial v}{\partial x}(\theta_i^-) \right) \frac{\partial \tilde{u}^{t^q}}{\partial x}(\theta_1) \\
&= \Delta x \frac{\partial(\tilde{u}^{t^q} v)}{\partial x}(x_c) + \epsilon_i \\
&= \Delta x \frac{\partial(\tilde{u}^{t^q} v)}{\partial x}(x_i) + \Delta x(x_i - x_c) \frac{\partial(\tilde{u}^{t^q} v)}{\partial x}(\theta_2) + \epsilon_i \\
&= \Delta x \frac{\partial(\tilde{u}^{t^q} v)}{\partial x}(x_i) + \tilde{\epsilon}_i,
\end{aligned}$$

avec

$$|\tilde{\epsilon}_i| \leq (\Delta x)^2 \left( \left\| \frac{\partial \tilde{u}^{t^q}}{\partial x} \right\|_\infty + \left\| \frac{\partial v}{\partial x} \right\|_\infty \|\tilde{u}^{t^q}\|_\infty + \Delta x \left\| \frac{\partial v}{\partial x} \right\|_\infty \left\| \frac{\partial \tilde{u}^{t^q}}{\partial x} \right\|_\infty \right).$$

Ce qui donne finalement,

$$(Ec)_i^{t_i^1, q} = \frac{\tilde{u}_i^{t_i^1, q} - \tilde{u}_i^{t^q}}{\Delta t_i} + \left( \frac{\partial(\tilde{u}_i^{t^q} v_i)}{\partial x} + \Delta x \epsilon_i^1 \right).$$

Pour la deuxième étape de la récurrence, nous supposons que la formule (2.10) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soient  $t_q^{n+1} = t_{i,q}^{r_i+1} := t^q + (r_i + 1)\Delta t_i < t^{q+1}$ ,  $t_i^{pcd}$  le temps de la dernière mise à jour de la densité  $u_i$  et  $\delta_i^{r_i+1} = t_{i,q}^{r_i+1} - t_i^{pcd}$ . Supposons que  $t_{i-1,q}^{r_i-1} \leq t_{i,q}^{r_i+1}$  alors,

$$(Ec)_i^{t_i^{r_i+1}, q} = \frac{\tilde{u}_i^{t_i^{r_i+1}, q} - \tilde{u}_i^{t_i^{pcd}}}{\delta_i^{r_i+1}} + \frac{1}{\Delta x} \left( \tilde{F}_i^{t_i^{r_i+1}, q} - \tilde{F}_{i-1}^{t_{i-1,q}^{r_i-1}, q} \right),$$

On suppose que la vitesse  $v$  est négative, les autres cas se traitent de la même manière. Le choix d'une vitesse négative est encore une fois fait juste à titre d'exemple et n'a aucune influence sur le raisonnement. Le but est de pouvoir exprimer le flux pour faire un calcul explicite. D'après l'hypothèse de récurrence, et la définition de l'erreur de consistance on a :

$$\begin{aligned}
\tilde{u}_i^{t_i^{r_i+1}, q} &= \tilde{u}_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} \left( v_{i+1/2} \tilde{u}_{i+1}^{t_i^{r_i+1}, q} - v_{i-1/2} \tilde{u}_i^{t_{i-1,q}^{r_i-1}} \right) + \delta_i^{r_i+1} (Ec)_i^{t_i^{r_i+1}, q} \\
&= \tilde{u}_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} \left( \tilde{F}_{i+1}^{t_i^{r_i+1}, q} - \tilde{F}_i^{t_i^{r_i+1}, q} \right) \\
&+ \frac{\delta_i^{r_i+1}}{\Delta x} v_{i+1/2} \left( r_i \Delta t_i \left( \frac{\partial(\tilde{u}_i^{t_i^{r_i+1}, q} v_i)}{\partial x} + \Delta x \epsilon_{i+1}^{r_i} - (Ec)_{i+1}^{t_i^{r_i+1}, q} \right) \right) \\
&- \frac{\delta_i^{r_i+1}}{\Delta x} v_{i-1/2} \left( r_{i-1} \Delta t_{i-1} \left( \frac{\partial(\tilde{u}_{i-1}^{t_{i-1,q}^{r_i-1}} v_{i-1})}{\partial x} + \Delta x \epsilon_i^{r_{i-1}} - (Ec)_i^{t_{i-1,q}^{r_i-1}} \right) \right) + \delta_i^{r_i+1} (Ec)_i^{t_i^{r_i+1}, q}
\end{aligned}$$

Par suite,

$$\begin{aligned}
\tilde{u}_i^{t_i^{r_i+1}, q} &= \tilde{u}_i^{t_i^q} - (r_i + 1)\Delta t_i \left( \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} + \Delta x (\varepsilon_i^{r_i} + \varepsilon_i^1 - r_i \varepsilon_{i+1}^{r_i} + r_{i-1} \varepsilon_i^{r_{i-1}}) \right) \\
&+ \delta_i^{r_i+1} \left( \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} - \frac{\partial(\tilde{u}_{i-1}^{t_i^q} v_{i-1})}{\partial x} \right) + (t_i^{pcd} - t_i^q) (Ec)_i^{t_i^{pcd}} + \delta_i^{r_i+1} (Ec)_i^{t_i^{r_i+1}, q} \\
&- \left( \frac{\delta_i^{r_i+1}}{\Delta x} v_{i+1/2} r_i \Delta t_i (Ec)_{i+1}^{t_i^{r_i}, q} - \frac{\delta_i^{r_i+1}}{\Delta x} v_{i-1/2} r_{i-1} \Delta t_{i-1} (Ec)_i^{t_{i-1}^{r_{i-1}}, q} \right) \\
&= \tilde{u}_i^{t_i^q} - (r_i + 1)\Delta t_i \left( \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} + \Delta x \varepsilon_i^{r_i+1} \right) + (r_i + 1)\Delta t_i (Ec)_i^{t_i^{r_i+1}, q}
\end{aligned}$$

finalement on a

$$\tilde{u}_i^{t_i^{r_i+1}, q} = \tilde{u}_i^{t_i^q} - (r_i + 1)\Delta t_i \left( \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} + \Delta x \varepsilon_i^{r_i+1} - (Ec)_i^{t_i^{r_i+1}, q} \right)$$

avec

$$|\varepsilon_i^{r_i+1}| \leq \|\varepsilon_i^{r_i}\|_\infty + \|\varepsilon_i^1\|_\infty + \|\varepsilon_{i+1}^{r_i}\|_\infty + \|\varepsilon_i^{r_{i-1}}\|_\infty + \|\partial_x^2(\tilde{u}^{t_i^q} v)\|_\infty.$$

La formule (2.10) est ainsi vérifiée. Revenons à la démonstration de la Proposition 2. On suppose que la solution est synchronisée à l'instant  $t^{q+1}$  et que pour tout  $i$  on a  $r_i \Delta t_i + \delta_i = \Delta t$  (Fig. 2.4). D'après la formule (2.10), pour tout  $i$  :

$$\tilde{u}_i^{t_i^{q+1}} = \tilde{u}_i^{t_i^q} - \Delta t \left( \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} + \Delta x \varepsilon_i^{t_i^{q+1}} - (Ec)_i^{t_i^{q+1}} \right).$$

Par conséquent,

$$(Ec)_i^{t_i^{q+1}} = \frac{\tilde{u}_i^{t_i^{q+1}} - \tilde{u}_i^{t_i^q}}{\Delta t} + \frac{\partial(\tilde{u}_i^{t_i^q} v_i)}{\partial x} + \Delta x \varepsilon_i^{t_i^{q+1}} = O(\Delta x) + O(\Delta t).$$

Ce qui prouve la consistance d'ordre un en temps et en espace du schéma asynchrone.  $\square$

### 2.3.3. Convergence du schéma asynchrone

Soient  $\tilde{u}_i^{t_i^n} := u(x_i, t_i^n)$  la solution suffisamment régulière, au moins de classe  $\mathcal{C}^2$ , de l'équation de transport (2.2) et  $u_i^{t_i^n}$  la solution numérique obtenue par le schéma asynchrone avec la donnée initiale  $u_i^{t_i^0} = u_0(x_i)$ . Introduisons l'erreur  $e_i^{t_i^n} = \tilde{u}_i^{t_i^n} - u_i^{t_i^n}$ ,  $1 \leq n \leq N_q$ .

**Théorème 1.** *Sous la condition CFL (2.5), le schéma asynchrone est convergent d'ordre 1 et pour tout  $i$ ,*

$$\|e_i^{t_i^q}\|_{asyn, \infty} \leq \exp \left( r_i \Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_\infty \right) \|e^{t_i^q}\|_{asyn, \infty} + \Delta x \sup_i |\varepsilon_i^{t_i^q}| + \Delta t \sup_i |\bar{\varepsilon}_i^{t_i^q}|, \quad (2.11)$$

$\varepsilon_i^{t_i, q}$  et  $\tilde{\varepsilon}_i^{t_i, q}$  sont bornés  $\forall i$ . En particulier,

$$\|e^{t^{q+1}}\|_{asyn, \infty} \leq \exp\left(\Delta t \left\| \frac{\partial v}{\partial x} \right\|_{\infty}\right) \|e^{t^q}\|_{asyn, \infty} + \Delta x \sup_i |\varepsilon_i^{t^{q+1}}| + \Delta t \sup_i |\tilde{\varepsilon}_i^{t^{q+1}}|, \quad (2.12)$$

*Démonstration.* La preuve de (2.11) se fait par récurrence. Supposons donc que  $t_{i,q}^1 = t^q + \min_j(\Delta t_j) := t^q + \Delta t_i$ . Alors d'après la Proposition 2, on a

$$\tilde{u}_i^{t_i, q} = \tilde{u}_i^{t^q} - \frac{\Delta t_i}{\Delta x} \left( \tilde{F}_i^{t^q} - \tilde{F}_{i-1}^{t^q} \right) + \Delta x \varepsilon_i^{t_i, q} + \Delta t_i \tilde{\varepsilon}_i^{t_i, q},$$

avec  $\tilde{\varepsilon}_i^{t_i, q}$  est le reste de développement de Taylor à l'ordre 1. Par suite,

$$\begin{cases} u_i^{t_i, q} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x} (F_i^{t^q} - F_{i-1}^{t^q}) \\ \tilde{u}_i^{t_i, q} &= \tilde{u}_i^{t^q} - \frac{\Delta t_i}{\Delta x} \left( \tilde{F}_i^{t^q} - \tilde{F}_{i-1}^{t^q} \right) + \Delta x \varepsilon_i^{t_i, q} + \Delta t_i \tilde{\varepsilon}_i^{t_i, q}, \end{cases}$$

par suite

$$\begin{aligned} e_i^{t_i, q} &= e_i^{t^q} - \frac{\Delta t_i}{\Delta x} v_{i+1/2} e_i^{t^q} - \frac{\Delta t_i}{\Delta x} v_{i-1/2} e_{i-1}^{t^q} + \Delta x \varepsilon_i^{t_i, q} + \Delta t_i \tilde{\varepsilon}_i^{t_i, q} \\ &= e_i^{t^q} - \frac{\Delta t_i}{\Delta x} v_{i+1/2} e_i^{t^q} + \frac{\Delta t_i}{\Delta x} v_{i+1/2} e_{i-1}^{t^q} - \frac{\Delta t_i}{\Delta x} (v_{i+1/2} - v_{i-1/2}) e_{i-1}^{t^q}. \end{aligned}$$

Ce qui donne finalement le résultat suivant :

$$\|e_i^{t_i, q}\|_{asyn, \infty} \leq \left(1 + \Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_{\infty}\right) \|e^{t^q}\|_{asyn, \infty} + \Delta x \sup_j |\varepsilon_j^{t_i, q}| + \Delta t_i \sup_j |\tilde{\varepsilon}_j^{t_i, q}|.$$

La formule (2.11) est ainsi vérifiée. Supposons maintenant que la formule (2.11) est vraie pour tout  $t_q^p < t_q^n$ . Soient  $t_q^n = t^q + (r_i + 1)\Delta t_i$ ,  $t^q + r_{i-1}\Delta t_{i-1} < t_{i,q}^{r_i+1}$  et  $\delta_i^{r_i+1} = t_{i,q}^{r_i+1} - t_i^{pcd}$ , alors

$$\begin{cases} u_i^{t_i, q} &= u_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} \left( F_i^{t_i^{r_i}} - F_{i-1, q}^{t_i^{r_i-1}} \right) \\ \tilde{u}_i^{t_i, q} &= \tilde{u}_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} \left( \tilde{F}_i^{t_i^{r_i}} - \tilde{F}_{i-1}^{t_i^{r_i-1}} \right) + \Delta x \varepsilon_i^{t_i, q} + \delta_i^{r_i+1} \tilde{\varepsilon}_i^{t_i, q}, \end{cases}$$

et donc par soustraction, on obtient

$$\begin{aligned} e_i^{t_i, q} &= e_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} \left( F_i^{t_i^{r_i}} - \tilde{F}_i^{t_i^{r_i}} \right) - \frac{\delta_i^{r_i+1}}{\Delta x} \left( F_{i-1}^{t_{i-1, q}^{r_i-1}} - \tilde{F}_{i-1}^{t_{i-1, q}^{r_i-1}} \right) + \Delta x \varepsilon_i^{t_i, q} + \delta_i^{r_i+1} \tilde{\varepsilon}_i^{t_i, q} \\ &= e_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} v_{i+1/2} \left( u_i^{t_i^{r_i}} - \tilde{u}_i^{t_i^{r_i}} \right) - \frac{\delta_i^{r_i+1}}{\Delta x} v_{i-1/2} \left( u_{i-1}^{t_{i-1, q}^{r_i-1}} - \tilde{u}_{i-1}^{t_{i-1, q}^{r_i-1}} \right) \\ &\quad + \Delta x \varepsilon_i^{t_i, q} + \delta_i^{r_i+1} \tilde{\varepsilon}_i^{t_i, q} \\ &= e_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} v_{i+1/2} e_i^{t_i^{r_i}} - \frac{\delta_i^{r_i+1}}{\Delta x} v_{i-1/2} e_{i-1}^{t_{i-1, q}^{r_i-1}} + \Delta x \varepsilon_i^{t_i, q} + \delta_i^{r_i+1} \tilde{\varepsilon}_i^{t_i, q}. \end{aligned}$$

Ce qui donne finalement le résultat suivant :

$$\begin{aligned} e_i^{t_i^{r_i+1}, q} &= e_i^{t_i^{pcd}} - \frac{\delta_i^{r_i+1}}{\Delta x} v_{i+1/2} e_i^{t_i^{r_i}, q} + \frac{\delta_i^{r_i+1}}{\Delta x} v_{i+1/2} e_{i-1}^{t_i^{r_i}, q} - \frac{\delta_i^{r_i+1}}{\Delta x} v_{i+1/2} \left( e_{i-1}^{t_i^{r_i}, q} - e_{i-1}^{t_{i-1}^{r_i-1}, q} \right) \\ &\quad + \delta_i^{r_i+1} \frac{\partial v}{\partial x}(c) e_{i-1}^{t_{i-1}^{r_i-1}, q} + \Delta x \varepsilon_i^{t_i^{r_i+1}, q} + \delta_i^{r_i+1} \bar{\varepsilon}_i^{t_i^{r_i+1}, q}. \end{aligned}$$

Le schéma asynchrone est stable pour la norme infinie, donc d'après la Proposition 1 on a :

$$\begin{aligned} \|e^{t_i^{r_i+1}, q}\|_{asyn, \infty} &\leq \left( 1 + \delta_i^{r_i+1} \left\| \frac{\partial v}{\partial x} \right\|_{\infty} \right) \|e^{t_i^{pcd}}\|_{asyn, \infty} + \Delta x \sup_i |\varepsilon_i^{t_i^{r_i+1}, q}| + \delta_i^{r_i+1} t_i \sup_i |\bar{\varepsilon}_i^{t_i^{r_i+1}, q}| \\ &\leq \exp \left( (r_i + 1) \Delta t_i \left\| \frac{\partial v}{\partial x} \right\|_{\infty} \right) \|e^{t^q}\|_{asyn, \infty} + \Delta x \sup_j |\varepsilon_j^{t_i^{r_i+1}, q}| + \delta_i^{r_i+1} \sup_j |\bar{\varepsilon}_j^{t_i^{r_i+1}, q}|. \end{aligned}$$

Ce qui prouve la formule (2.11). Pour terminer la démonstration, supposons que la solution est synchronisée à l'instant  $t^{q+1}$  et que pour tout  $i$ ,  $r_i \Delta t_i + \delta_i = \Delta t$ . D'après la formule (2.11), on a pour tout  $i$ ,

$$\begin{aligned} |u_i^{t^{q+1}}| &\leq \exp \left( (r_i \Delta t_i + \delta_i) \left\| \frac{\partial v}{\partial x} \right\|_{\infty} \right) \|u^{t^q}\|_{asyn, \infty} + \Delta x \sup_i |\varepsilon_i^{t^{q+1}}| + \Delta t \sup_i |\bar{\varepsilon}_i^{t^{q+1}}| \\ &\leq \exp \left( (\Delta t) \left\| \frac{\partial v}{\partial x} \right\|_{\infty} \right) \|u^{t^q}\|_{asyn, \infty} + \Delta x \sup_i |\varepsilon_i^{t^{q+1}}| + \Delta t \sup_i |\bar{\varepsilon}_i^{t^{q+1}}|. \end{aligned}$$

Le schéma asynchrone est donc convergent d'ordre 1 en temps et en espace.  $\square$

## 2.4. Validation numérique

Le but de cette section est de valider numériquement les résultats théoriques prouvés dans la section précédente. Les propriétés de convergence de la méthode asynchrone seront donc illustrés par des cas-tests numériques.

Considérons l'équation d'advection 1D avec une vitesse constante ( $v = 1$ ) sur l'intervalle  $[0, 1]$  avec des conditions aux limites périodiques et une condition initiale régulière :

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, & (x, t) \in [0, 1] \times [0, T] \\ u(x, 0) = \sin(2\pi x), & x \in [0, 1], \end{cases} \quad (2.13)$$

Nous avons choisi une vitesse constante, le maillage choisi doit donc être adapté pour donner un cas-test asynchrone et doit être défini de manière à ce que l'aspect asynchrone soit conservé lors du raffinement. Pour cela, nous considérons un maillage monodimensionnel variable et paramétré par une loi polynomiale :

$$x_i = 4(1 - \alpha) \left( \frac{i-1}{N-1} - \frac{1}{2} \right)^3 + \alpha \left( \frac{i-1}{N-1} - \frac{1}{2} \right) + \frac{1}{2}, \quad (2.14)$$

où  $N$  est le nombre de cellules. Le paramètre  $\alpha$  définit la pente à l'abscisse  $x = 0.5$  ; voir la figure 2.5. L'aspect asynchrone provient donc bien de la variation du maillage.

Rappelons que pour l'équation de transport, nous pouvons calculer analytiquement la solution exacte sans avoir besoin de passer par une approximation numérique. En fait, en utilisant la méthode des ca-



FIGURE 2.5 – Maillage polynomial avec  $N = 10$  et  $\alpha = 0.5$

ractéristiques, la solution à tout instant  $t$  est :  $u(x, t) = u_0(x - vt)$ . L'approximation numérique de l'équation d'advection n'est qu'un moyen pour justifier la convergence de la méthode et pour vérifier son ordre de précision.

### Convergence numérique du schéma asynchrone

Commençons par vérifier numériquement l'ordre de convergence du schéma asynchrone. Nous menons donc une série de calculs numériques avec un nombre de cellules qui augmente et nous essayons d'estimer l'erreur en fonction du nombre de mailles total. Nous considérons pour cela une boucle pour  $j = 1 \dots 9$ , et  $N = 10 \times 2^{j-1}$ . Nous menons ensuite deux calculs avec le solveur MACOPA, l'un en asynchrone, l'autre en synchrone (pas de temps imposé identique pour toutes les cellules). Puis nous traçons la norme de l'erreur (ici la norme infinie) en fonction du nombre de cellules en échelle logarithmique ; les résultats sont présentés sur la figure 2.6.

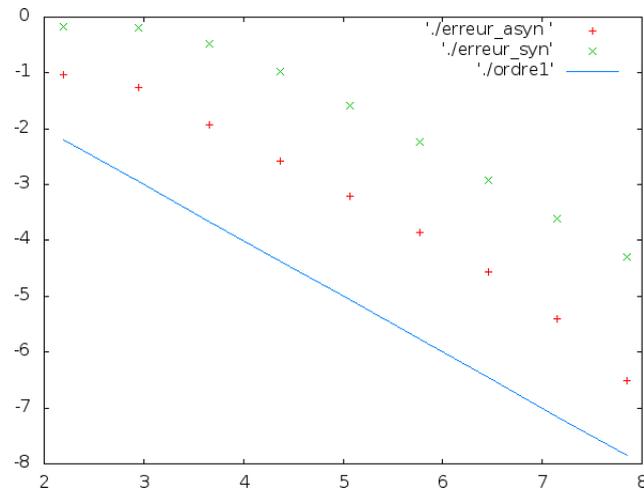


FIGURE 2.6 – ordre de convergence du schéma asynchrone : norme infinie de l'erreur en fonction du nombre de cellules en échelle logarithmique.

La figure (Fig.2.6) représente l'erreur asynchrone (en vert), l'erreur synchrone (en rouge) et l'ordre 1 théorique (en bleu). L'étude numérique de la convergence montre bien que le schéma asynchrone est convergent d'ordre 1, ce qui est cohérent avec le résultat théorique. De plus, comme le montre la figure (Fig.2.6), le schéma asynchrone converge plus vite que le schéma classique d'ordre 1. En effet, la constante de l'erreur asynchrone en facteur de  $\Delta x$  est plus petite que celle obtenue avec le schéma synchrone classique.

## La diffusion numérique

Nous comparons dans cette section les solutions numériques du problème (2.13) calculées par les schémas synchrone et asynchrone avec la solution exacte. Nous traçons donc les trois solutions pour un nombre de cellules  $N = 1000$ , un temps final de simulation  $T = 4$  s (ce qui correspond à 4 tours du domaine périodique) et le maillage polynomial défini ci-dessus.

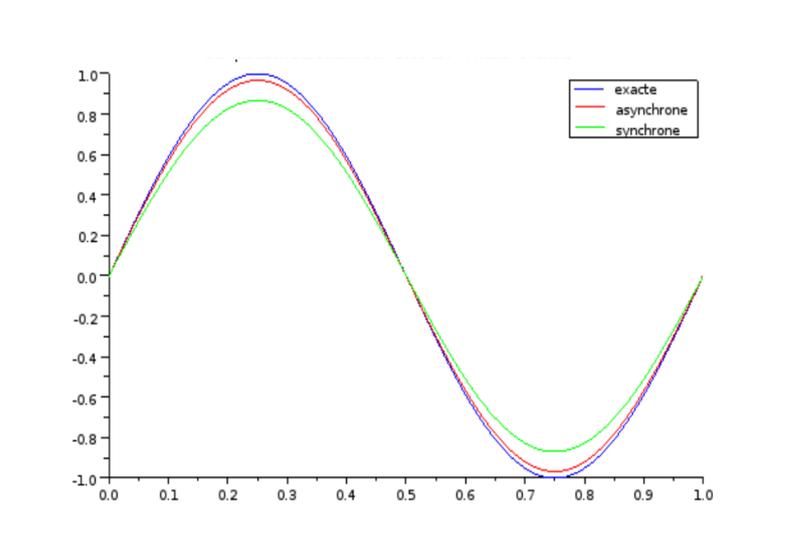


FIGURE 2.7 – Comparaison des schémas synchrone classique et asynchrone

Notons que le schéma asynchrone réduit la diffusion numérique en comparaison avec le schéma synchrone classique (Fig. 2.7). En fait, dans le cas d’une intégration numérique asynchrone, chaque cellule de maillage avance en fonction de son propre pas de temps calculé sous la condition CFL locale, ce qui réduit la diffusion numérique. Ce n’est pas le cas pour une intégration numérique synchrone classique dans laquelle le pas de temps de tous les cellules de maillage est limité par la plus petite cellule ce qui favorise les phénomènes de diffusion numérique.

## Le gain en temps de calcul

L’un des objectifs majeurs de l’asynchrone est le gain en temps de calcul. En fait, dans le cas d’une intégration classique, les cellules sont réévaluées plus souvent que ce qui est imposé par leur propre CFL locale, tandis que la méthode asynchrone limite la fréquence de mise à jour des cellules ce qui amène à une réduction du temps de calcul. Notons que, en termes de temps de calcul, le point critique pour accélérer la méthode asynchrone est la recherche la cellule la plus urgente pour mettre à jour, ce point technique pourrait être pénalisant pour les performances du solveur mais certaines options alternatives ont été présentées dans [16] pour limiter l’impact de cette recherche.

Dans le tableau 2.8, nous présentons pour un nombre de cellules fixé, une comparaison des schémas synchrone et asynchrone. L’erreur asynchrone est plus petite que l’erreur synchrone et c’est ce qui a été observé dans la figure Fig.2.6 au début de cette section. Ce que nous pouvons aussi remarquer dans ce tableau est le gain en temps de calcul, d’un facteur 10 en moyenne. Le schéma asynchrone est donc très

N	Erreurs		Temps de calcul	
	Schéma asynchrone	Schéma synchrone	Schéma asynchrone	Schéma synchrone
1280	6,87 10 <sup>-3</sup>	2,71 10 <sup>-2</sup>	11 s	1,7 mn
2560	3,44 10 <sup>-3</sup>	1,36 10 <sup>-2</sup>	47 s	11,76 mn
5120	1,72 10 <sup>-3</sup>	6,83 10 <sup>-3</sup>	3,5 mn	48 mn
10240	8,62 10 <sup>-4</sup>	3,41 10 <sup>-3</sup>	18 mn	3,19 h
20480	4,31 10 <sup>-4</sup>	1,71 10 <sup>-3</sup>	1,53 h	12,74 h

FIGURE 2.8 – comparaison de la performance des schémas synchrone et asynchrone

efficace en termes de coût de calcul.

**Remarque 1.**

Pour  $N$  cellules, le speed up théorique est donné par la formule suivante :

$$S = \frac{N}{\min_i(\Delta t_i) \times \sum_{i=1}^N \frac{1}{\Delta t_i}} \quad (2.15)$$

Notons que le speed up théorique dépend uniquement du maillage, en particulier le nombre des mailles et les différents pas de temps. C'est une estimation a priori du gain en temps de calcul sans tenir compte ni du modèle discrétisé, ni de la méthode de discrétisation. De plus, pour la mise en œuvre pratique, le speed up est en général plus lent que le speed up théorique. Ceci est dû à l'implémentation effective des schémas asynchrone et synchrone.

Dans le cas du maillage polynomial, le facteur d'accélération estimé en utilisant la formule (2.15) est de l'ordre de 11. Le speed up numérique calculé en se basant sur les résultats du tableau ?? varie entre 8 et 13.

## 2.5. Conclusion

L'objectif de ce chapitre était de définir la méthodologie asynchrone et d'étudier ses propriétés, en particulier la stabilité, la consistance et la convergence. Pour illustrer la méthode, nous avons choisi dans ce premier chapitre de l'appliquer à l'équation de transport en dimension 1 de l'espace. Le but était de bien comprendre les idées de base de l'asynchrone mais aussi de contrôler les éventuelles difficultés liées à l'aspect asynchrone. Les tests numériques montrent que le schéma asynchrone est convergent d'ordre 1, ce qui valide les résultats théoriques prouvés qui estiment que pour un maillage suffisamment régulier, le schéma asynchrone est d'ordre 1. En comparaison avec le schéma classique, le schéma asynchrone réduit la diffusion numérique et ceci est obtenu grâce à l'utilisation de la CFL optimale pour chaque cellule de

maillage dans le cadre d'une intégration asynchrone. De plus, la méthode asynchrone est très efficace en termes de réduction du temps de calcul. Dans ce chapitre, nous nous sommes intéressés aux équations de transport en dimension un de l'espace. L'extension pour une dimension en espace quelconque sera l'objet du chapitre suivant.



## Chapitre 3

# Application de la méthode asynchrone pour la discrétisation de l'équation de transport en multi-dimensions

Après avoir étudié les propriétés du schéma numérique asynchrone pour les équations de transport en dimension 1 d'espace, nous proposons dans ce chapitre une extension de la méthode dans le cas multi-dimensionnel. D'abord, nous appliquerons la méthode pour la discrétisation de l'équation de transport en dimension  $d$  de l'espace où  $d \geq 1$ . Ensuite, nous traiterons la stabilité de la méthode et nous discuterons sa consistance en fonction de la variation de la vitesse mais également en fonction du maillage. Enfin, la méthode de l'erreur corrigée sera adaptée à l'aspect asynchrone afin de pouvoir prouver la convergence du schéma asynchrone dans le cas d'un maillage variable.

### 3.1. Discrétisation de l'équation de transport en multi-dimensions

Considérons l'équation de transport scalaire suivante :

$$\begin{cases} \frac{\partial u}{\partial t} + \nabla \cdot (vu) = 0, & (x, t) \in \Omega \times ]0, +\infty[, \\ u(x, 0) = u_0(x), & x \in \Omega, \end{cases} \quad (3.1)$$

où  $v$  est un vecteur non nul,  $\Omega$  est un domaine polygonal borné de  $\mathbb{R}^d$  et  $d \geq 1$ .

Soit  $\mathcal{T} = \{K_i : i = 1, \dots, N\}$  une partition de  $\Omega$  en polyèdres  $K_i$ . Pour tout  $i \in [1, N]$ ,  $\mathcal{N}(i) = \{j : |K_j \cap K_i| \neq 0\}$  désigne l'ensemble des indices des volumes voisins de  $K_i$ . Pour  $j \in \mathcal{N}(i)$ ,  $n_{i,j}$  est la normale unitaire de  $K_i \cap K_j$  pointant de  $K_i$  vers  $K_j$ , et  $\mathbf{N}_{i,j} = |K_j \cap K_i| n_{i,k}$ . Soient  $\mathcal{N}^+(i) = \{j \in \mathcal{N}, v \cdot n_{i,j} > 0\}$ ,  $\mathcal{N}^-(i) = \{j \in \mathcal{N}, v \cdot n_{i,j} < 0\}$  et  $\mathcal{N}(i) = \mathcal{N}^+(i) \cup \mathcal{N}^-(i)$ .

Pour la discrétisation en espace de l'équation de transport (3.1), nous utilisons toujours le schéma volume fini décentré : le schéma upwind. Le système semi-discrétisé s'écrit donc sous la forme suivante :

$$\frac{\partial u_i}{\partial t}(t) = -\frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v \cdot \mathbf{N}_{i,j} u_i(t) + \sum_{j \in \mathcal{N}^-(i)} v \cdot \mathbf{N}_{i,j} u_j(t) \right) = -\frac{1}{|K_i|} \sum_{j \in \mathcal{N}(i)} F_{i,j}(t), \quad (3.2)$$

où la densité  $u_i(t)$  est une approximation de  $\frac{1}{|K_i|} \int_{K_i} u(x, t) dx$  et les flux numériques  $F_{i,j}$  sont définis par :

1.  $F_{i,j}(t) = v \cdot \mathbf{N}_{i,j} u_i(t), \forall j \in \mathcal{N}^+(i)$  : les flux sortants
2.  $F_{i,j}(t) = v \cdot \mathbf{N}_{i,j} u_j(t), \forall j \in \mathcal{N}^-(i)$  : les flux entrants.

Pour la discrétisation en temps, nous utilisons la méthode asynchrone. L'algorithme asynchrone appliqué à l'équation (3.1) est résumé dans la figure 3.1.

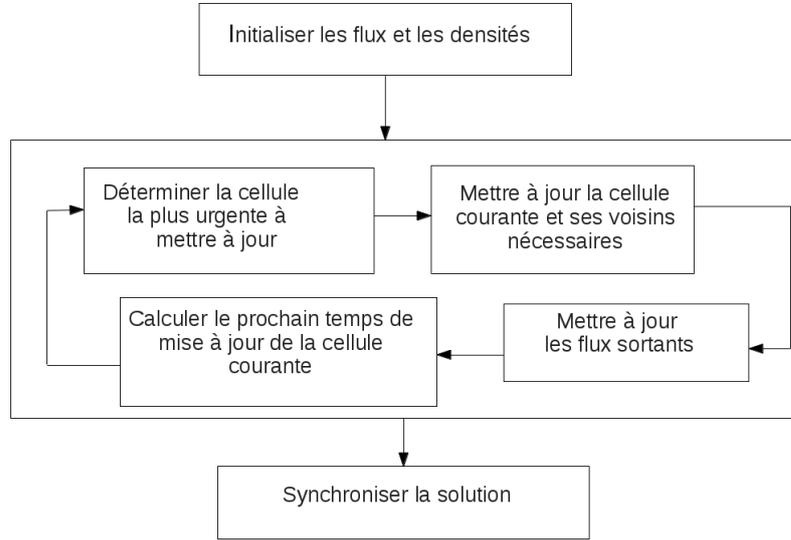


FIGURE 3.1 – L'algorithme asynchrone

Reprenons les notations du chapitre précédent. Soient  $t_{cur}$  le temps courant de la simulation et  $t_i^{pcd}$  le temps de la dernière mise à jour du volume  $K_i$ . Supposons que  $t_i^{r_i} < t_{cur}$  et pour tout  $j \in \mathcal{N}^-(i)$ ,  $t_j^{r_j} < t_{cur}$ . Le schéma upwind asynchrone s'écrit donc sous la forme suivante :

$$u_i^{t_{cur}} = u_i^{t_i^{pcd}} - \frac{t_{cur} - t_i^{pcd}}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v \cdot \mathbf{N}_{i,j} u_i^{t_i^{r_i}} + \sum_{j \in \mathcal{N}^-(i)} v \cdot \mathbf{N}_{i,j} u_j^{t_j^{r_j}} \right), \quad (3.3)$$

où pour tout  $j \in \mathcal{N}(i)$ ,  $t_j^{r_j}$  est le temps, multiple de  $\Delta t_j$ , de la dernière mise à jour de la cellule  $j$  juste avant  $t_{cur}$ .

## 3.2. Stabilité du schéma asynchrone

Dans cette section, la vitesse est supposée constante et le maillage est variable.

**Proposition 3.** *Le schéma asynchrone est stable pour la norme  $\|\cdot\|_{asyn,\infty}$ , définie par  $\|u^{t^n}\|_{asyn,\infty} := \sup_{i,p \leq n} |u_i^{t^p}|$ , sous la condition CFL locale sur le pas de temps  $\Delta t_i$  du volume  $K_i$  :*

$$\Delta t_i \leq \frac{|K_i|}{\sum_{j \in \mathcal{N}^+(i)} v \cdot \mathbf{N}_{i,j}} \quad \forall i = 1, \dots, N, \quad (3.4)$$

En particulier, à chaque instant  $t^0 \leq t^n \leq T$ , on a :

$$\|u^{t^n}\|_{asyn,\infty} \leq \|u^{t^0}\|_{asyn,\infty}. \quad (3.5)$$

où  $t^0$  est l'instant initial et  $T$  le temps final de la simulation.

*Démonstration.* Rappelons tout d'abord que

$$\sum_{j \in \mathcal{N}(i)} v \cdot \mathbf{N}_{i,j} = 0, \quad (3.6)$$

puis utilisons une récurrence pour démontrer l'inégalité (3.5). Supposons dans un premier temps que  $t^1 = \min_j (\Delta t_j) := \Delta t_i$ . Alors,

$$\begin{aligned} u_i^{t^1} &= u_i^{t^0} - \frac{\Delta t_i}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v \cdot \mathbf{N}_{i,j} u_i^{t^0} + \sum_{j \in \mathcal{N}^-(i)} v \cdot \mathbf{N}_{i,j} u_j^{t^0} \right) \\ &= \left( 1 - \frac{\Delta t_i}{|K_i|} \sum_{j \in \mathcal{N}^+(i)} v \cdot \mathbf{N}_{i,j} \right) u_i^{t^0} - \frac{\Delta t_i}{|K_i|} \sum_{j \in \mathcal{N}^-(i)} v \cdot \mathbf{N}_{i,j} u_j^{t^0} \\ &= \alpha_i u_i^{t^0} + \sum_{j \in \mathcal{N}^-(i)} \alpha_{i,j} u_j^{t^0}, \end{aligned}$$

avec

1.  $\alpha_i := 1 - \frac{\Delta t_i}{|K_i|} \sum_{j \in \mathcal{N}^+(i)} v \cdot \mathbf{N}_{i,j}$ . Sous la condition CFL (3.4),  $\alpha_i \geq 0$ ,
2. Pour tout  $j \in \mathcal{N}^-(i)$ ,  $\alpha_{i,j} := -\frac{\Delta t_i}{|K_i|} v \cdot \mathbf{N}_{i,j}$  et  $v \cdot \mathbf{N}_{i,j} < 0$  donc  $\alpha_j \geq 0$ .
3. D'après (3.6),  $\alpha_i + \sum_{j \in \mathcal{N}^-(i)} \alpha_j = 1$ .

L'inégalité (3.5) est donc vérifiée. Supposons maintenant que la formule (3.5) est vraie jusqu'à un instant  $t^n < T$ . Soient  $t^{n+1} = t_i^{r_i+1} := (r_i + 1)\Delta t_i < T$  et pour tout  $j \in \mathcal{N}^-(i)$ ,  $t_j^{r_j} := r_j \Delta t_j < t_i^{r_i}$ . Supposons que pour tout  $j \in \mathcal{N}^-(i)$ ,  $k_j = E\left(\frac{\Delta t_i}{\Delta t_j}\right) \geq 0$  et  $k = \sum_{j \in \mathcal{N}^-(i)} k_j$ , où  $E$  est la partie entière. Soient  $\delta_i^{r_i+1} = \min\left(\Delta t_i, t_i^{r_i+1} - \max_{j \in \mathcal{N}^-(i)} (t_j^{r_j+k_j})\right)$  et  $\delta_i^{r_i} = \min_{j \in \mathcal{N}^-(i)} (t_j^{r_j+1}) - t_i^{r_i}$ . La suite  $\left(t_{r_i}^h\right)_{1 \leq h \leq k}$  est définie comme suit :

$$t_{r_i}^1 := \min_{j \in \mathcal{N}^-(i)} (t_j^{r_j+1}), \dots, t_{r_i}^k := \max_{j \in \mathcal{N}^-(i)} (t_j^{r_j+k_j}).$$

Les notations sont présentées dans la figure 3.2. où les cellules  $j$ ,  $n$  et  $z$  sont telles que  $i \in N^+(j)$ ,  $i \in N^+(n)$  et  $i \in N^+(z)$ . Les traits présentent les temps de mise à jour de la cellule  $j$ , multiples de

$\Delta t_j$ , les traits pointillés présentent les temps de mise à jour de la cellule  $n$ , multiples de  $\Delta t_n$  et les traits pointillés en gras présentent les temps de mise à jour de la cellule  $z$ , multiples de  $\Delta t_z$ .

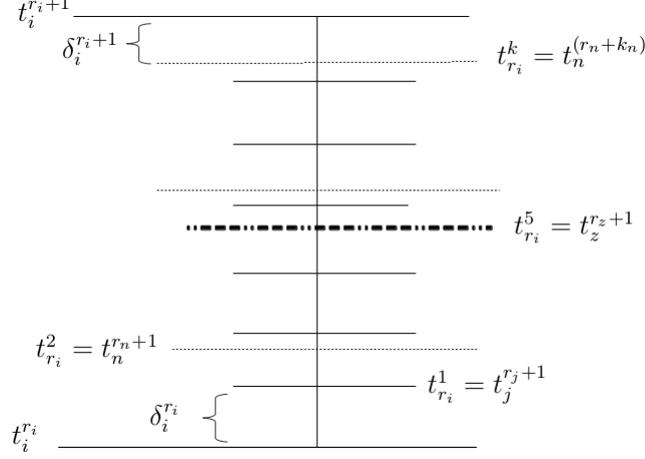


FIGURE 3.2 – Notations

Par définition du schéma asynchrone (3.3), on a :

$$\begin{aligned}
u_i^{r_i+1} &= u_i^{r_i} - \frac{\Delta t_i}{|K_i|} \sum_{j \in \mathcal{N}^+(i)} v \cdot \mathbf{N}_{i,j} u_i^{r_i} - \frac{\delta_i^{r_i+1}}{|K_i|} \sum_{j \in \mathcal{N}^-(i)} v \cdot \mathbf{N}_{i,j} u_j^{r_j+k_j} \\
&\quad - \frac{\delta_i^{r_i}}{|K_i|} \sum_{j \in \mathcal{N}^-(i)} v \cdot \mathbf{N}_{i,j} u_j^{r_j} - \sum_{h=2}^k \frac{t_{r_i,i}^h - t_{r_i,i}^{h-1}}{|K_i|} \sum_{j \in \mathcal{N}^-(i)} v \cdot \mathbf{N}_{i,j} \tilde{u}_j^{t_{r_i,i}^h},
\end{aligned}$$

avec  $\tilde{u}_j^{t_{r_i,i}^h}$  est la valeur de la densité  $u_j$  à un certain instant  $t_j^{r_j+l} < t_{r_i,i}^h$  ( $0 \leq l \leq k_j$ ) et qui correspond au temps de la dernière mise à jour, multiple de  $\Delta t_j$  du volume  $K_j$  juste avant  $t_{r_i,i}^h$ . Pour conclure il suffit de noter que  $\delta_i^{r_i+1} + \delta_i^{r_i} + \sum_{h=2}^k (t_{r_i,i}^h - t_{r_i,i}^{h-1}) = \Delta t_i$  puis d'utiliser la condition CFL (3.4). La formule (3.5) est ainsi vérifiée, ce qui achève la preuve de la proposition.  $\square$

### 3.3. Consistance du schéma asynchrone

Rappelons que pour certains schémas numériques la non uniformité du maillage fait apparaître une perte de consistance malgré la convergence du schéma. Le schéma volumes finis upwind classique explicite en temps est consistant d'ordre 1 pour un maillage cartésien régulier. Mais pour un maillage non uniforme, l'erreur de consistance est seulement en  $O(1)$ . Dans cette section, nous étudierons le comportement de l'asynchrone dans le cas d'un maillage constant puis dans le cas d'un maillage régulier. L'objectif est de pouvoir justifier que le schéma asynchrone préserve les propriétés du schéma classique.

Soient  $g_i$  le centre de gravité de  $K_i$  et  $g_{i,j}$  le centre de gravité de  $K_i \cap K_j$ . Supposons que  $|K_i| \leq h_i^2$  où  $h_i$  est le diamètre de  $K_i$ . Comme pour le cas 1D, la consistance du schéma asynchrone est traitée entre deux temps de synchronisation de la solution :  $t^q$  et  $t^{q+1} = t^q + \max_i (\Delta t_i)$ .

Soient la suite  $(t_{i,q}^{r_i})_{\{r_i\}}$  définie pour tout  $1 \leq i \leq N$  et  $0 \leq r_i \leq E \left( \frac{\max_j(\Delta t_j)}{\Delta t_i} \right)$  par :  $t_{i,q}^{r_i} = t^q + r_i \Delta t_i$  et la suite  $(t_q^n)$  des événements les plus urgents entre  $t^q$  et  $t^{q+1}$ .

### 3.3.1. Consistance : maillage constant

Le maillage est supposé constant, donc pour faire apparaître l'aspect asynchrone, nous supposons que la vitesse est variable,  $v$  est une fonction de  $x$ . Ceci permet aussi de traiter l'asynchrone en fonction de la variation de la vitesse. Nous supposons à titre d'exemple et seulement pour cette partie, que la dimension de l'espace est 2 et que le maillage est cartésien. Nous traitons donc une extension de la dimension 1.

**Proposition 4.** *Supposons que la vitesse  $v$  est de classe  $\mathcal{C}^2$ , alors sous la condition CFL locale (3.4), le schéma upwind asynchrone est consistant d'ordre 1 dans le cas d'un maillage cartésien constant.*

*Démonstration.* Soit  $u(g_i, t^n)$  la solution exacte de l'équation de transport (3.1), supposée suffisamment régulière, de classe  $\mathcal{C}^2$ , au centre de gravité du volume  $K_i$  à l'instant  $t^n$ . Nous prouvons tout d'abord par récurrence que pour tout  $i$  :

$$(Ec)_i^{t_{i,q}^{r_i}} = \frac{u(g_i, t_{i,q}^{r_i}) - u(g_i, t^q)}{r_i \Delta t_i} + (\nabla \cdot (v(g_i)u(g_i, t^q)) + O(h)), \quad (3.7)$$

avec  $(Ec)_i$  est l'erreur de consistance locale.

Soit  $t_q^1 = \min_j(\Delta t_j) = t_{i,q}^1$  alors par définition du schéma asynchrone (3.3) et de l'erreur de consistance :

$$\begin{aligned} (Ec)_i^{t_{i,q}^1} &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} u(g_j, t^q) \right) \\ &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \sum_{j \in \mathcal{N}(i)} v(g_{i,j}) \cdot N_{i,j} u(g_{i,j}, t^q) \\ &\quad - \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} (u(g_i, t^q) - u(g_{i,j}, t^q)) \right. \\ &\quad \left. + \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} (u(g_j, t^q) - u(g_{i,j}, t^q)) \right] \\ &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} (I_1 + I_2 + I_3). \end{aligned}$$

Utilisons un développement de Taylor d'ordre 1, alors

$$\begin{aligned} v(g_{i,j})u(g_{i,j}, t^q) &= (v(g_i) + (g_{i,j} - g_i) \nabla \cdot v(g_i)) \times \left( u(g_i, t^q) + (g_{i,j} - g_i) \nabla \cdot u(g_i, t^q) + O(h^2) \right) \\ &= v(g_i)u(g_i, t^q) + (g_{i,j} - g_i) \nabla \cdot (v(g_i)u(g_i, t^q)) + O(h^2). \end{aligned}$$

Par suite

$$\begin{aligned} I_1 &= \left( \sum_{j \in \mathcal{N}(i)} N_{i,j} \right) \cdot (v(g_i)u(g_i, t^q) - g_i \nabla \cdot u(g_i, t^q)) \\ &+ \left( \sum_{j \in \mathcal{N}(i)} N_{i,j} g_{i,j} \right) \cdot \nabla \cdot (v(g_i)u(g_i, t^q)) + O(h^3) \\ &= |K_i| \nabla \cdot (v(g_i)u(g_i, t^q)) + O(h^3). \end{aligned}$$

De plus

$$v(g_{i,j}) (u(g_i, t^q) - u(g_{i,j}, t^q)) = (g_i - g_{i,j})v(g_i) \nabla \cdot u(g_i, t^q) + O(h^2).$$

De même,

$$v(g_{i,j}) (u(g_j, t^q) - u(g_{i,j}, t^q)) = (g_j - g_{i,j})v(g_i) \nabla \cdot u(g_i, t^q) + O(h^2).$$

Le maillage est supposé constant donc :  $\forall j \in \mathcal{N}(i), |g_j - g_{i,j}| = |g_i - g_{i,j}|$ . Deux cas se présentent en fonction de la vitesse.

1. Supposons dans un premier temps que la vitesse  $v$  ne change pas de signe alors :

$$\forall j \in \mathcal{N}^+(i) \quad \exists k \in \mathcal{N}^-(i) \quad \text{tel que} \quad (g_i - g_{i,j})v(g_i)N_{i,j} + (g_k - g_{i,k})v(g_i)N_{i,k} = 0.$$

Par suite  $I_2 + I_3 = O(h^3)$ .

2. Si la vitesse change de signe, on distingue deux cas. Le premier cas

$$u_i^{t_{cur}} = u_i^{t_i^{pcd}} - \frac{t_{cur} - t_i^{pcd}}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot \mathbf{N}_{i,j} u_i^{t_i^{rj}} \right),$$

et le deuxième cas

$$u_i^{t_{cur}} = u_i^{t_i^{pcd}} - \frac{t_{cur} - t_i^{pcd}}{|K_i|} \left( \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot \mathbf{N}_{i,j} u_j^{t_j^{rj}} \right).$$

Le premier cas est traité et le deuxième se traite de la même manière. On a

$$\begin{aligned} (Ec)_i^{t_i^1, q} &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} u(g_i, t^q) \right) \\ &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}(i)} v(g_{i,j}) \cdot N_{i,j} u(g_{i,j}, t^q) \right. \\ &\quad \left. - \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} (u(g_i, t^q) - u(g_{i,j}, t^q)) \right], \end{aligned}$$

par suite

$$(Ec)_i^{t_i^1, q} = \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + (\nabla \cdot (v(g_i)u(g_i, t^q)) + O(h)) - \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} (u(g_i, t^q) - u(g_{i,j}, t^q)) \right).$$

avec

$$u(g_i, t^q) - u(g_{i,j}, t^q) = (g_i - g_{i,j}) \nabla u(g_i, t^q) + O(h^2).$$

De plus, la vitesse change de signe alors pour tout  $j \in \mathcal{N}^+(i)$ , il existe un  $\lambda_j \in ]g_{i,j}, g_i[$  telque  $v(\lambda_j) = 0$ . Alors

$$v(g_{i,j}) (u(g_i, t^q) - u(g_{i,j}, t^q)) = (\lambda_j - g_{i,j}) (g_i - g_{i,j}) \nabla v(\lambda_j) u(\lambda_j, t^q) + O(h^2).$$

par suite

$$\frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} (u(g_i, t^q) - u(g_{i,j}, t^q)) \right) = O(h).$$

Ce qui donne

$$(Ec)_i^{t_i^1, q} = \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + (\nabla \cdot (v(g_i)u(g_i, t^q)) + O(h)).$$

et donc

$$u(g_i, t_{i,q}^1) = u(g_i, t^q) - \Delta t_i \left( \nabla \cdot (v(g_i)u(g_i, t^q)) + O(h) - (Ec)_i^{t_i^1, q} \right).$$

Supposons maintenant que la formule (3.7) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soient  $t^{n+1} = t_{i,q}^{r_i+1}$  et  $t_i^{pcd}$  le temps de la dernière mise à jour de  $u_i$  et  $\delta_i^r = t_{i,q}^{r_i+1} - t_i^{pcd}$ . Alors

$$\begin{aligned} (Ec)_i^{t_i^{r_i+1}, q} &= \frac{u(g_i, t_{i,q}^{r_i+1}) - u(g_i, t_i^{pcd})}{\delta_i^r} \\ &+ \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} u(g_i, t_{i,q}^{r_i}) + \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} u(g_j, t_{j,q}^{r_j}) \right] \\ &= \frac{u(g_i, t_{i,q}^{r_i+1}) - u(g_i, t_i^{pcd})}{\delta_i^r} \\ &+ \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} \left( u_i^{t^q} - r_i \Delta t_i (\nabla \cdot (v(g_i)u(g_i, t^q)) + O(h) - (Ec)_i^{t_i^1, q}) \right) \right] \\ &- \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} \left( u_j^{t^q} - r_j \Delta t_j (\nabla \cdot (v(g_j)u(g_j, t^q)) + O(h) - (Ec)_j^{t_j^1, q}) \right) \right] \end{aligned}$$

Par suite,

$$u(g_i, t_{i,q}^{r_i+1}) = u(g_i, t_i^{pcd}) - \frac{\delta_i^r}{|K_i|} (J_1 + J_2 + J_3 + J_4) + \delta_i^r (Ec)_i^{t_{i,q}^{r_i+1}}$$

avec

$$J_1 = \sum_{j \in \mathcal{N}(i)} v(g_{i,j}) \cdot N_{i,j} u(g_{i,j}, t^q) = |K_i| \nabla \cdot (v(g_i) u(g_i, t^q)) + O(h^3),$$

$$J_2 = \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} (u(g_i, t^q) - u(g_{i,j}, t^q)) + \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} (u(g_j, t^q) - u(g_{i,j}, t^q)) = O(h^3),$$

$$\begin{aligned} J_3 &= \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} r_i \Delta t_i \nabla \cdot (v(g_i) u(g_i, t^q)) + \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} r_j \Delta t_j \nabla \cdot (v(g_j) u(g_j, t^q)) \\ &= \sum_{j \in \mathcal{N}(i)} v(g_{i,j}) \cdot N_{i,j} r_i \Delta t_i \nabla \cdot (v(g_i) u(g_i, t^q)) + \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} (r_j \Delta t_j - r_i \Delta t_i) \nabla \cdot (v(g_j) u(g_j, t^q)) \\ &\simeq r_i \Delta t_i |k_i| \nabla \cdot (v(g_i)) \nabla \cdot (v(g_i) u(g_i, t^q)) + \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} (r_j \Delta t_j - r_i \Delta t_i) \nabla \cdot (v(g_j) u(g_j, t^q)). \end{aligned}$$

Rappelons que  $\mathbf{N}_{i,k} = |K_j \cap K_i| n_{i,k}, n_{i,j}$  la normale unitaire de  $K_i \cap K_j$  pointant de  $K_i$  vers  $K_j$ . La vitesse est supposée suffisamment régulière de sorte que  $\Delta t_i - \Delta t_j = O(h^2)$  alors  $J_3 = O(h^3)$ . De même

$$J_4 = \sum_{j \in \mathcal{N}^+(i)} v(g_{i,j}) \cdot N_{i,j} r_i \Delta t_i (Ec)_i^{t_{i,q}^{r_i}} - \sum_{j \in \mathcal{N}^-(i)} v(g_{i,j}) \cdot N_{i,j} r_j \Delta t_j (Ec)_j^{t_{j,q}^{r_j}} = O(h^3),$$

Par suite,

$$\begin{aligned} u(g_i, t_{i,q}^{r_i+1}) &= u(g_i, t^q) - (t_i^{pcd} - t^q) \left[ \nabla \cdot (v(g_i) u(g_i, t^q)) + O(h) - (Ec)_i^{t_i^{pcd}} \right] \\ &\quad - (t_{i,q}^{r_i+1} - t_i^{pcd}) \left[ \nabla \cdot (v(g_i) u(g_i, t^q)) + O(h) - (Ec)_i^{t_{i,q}^{r_i+1}} \right] \\ &= u(g_i, t^q) - (r_i + 1) \Delta t_i \left[ \nabla \cdot (v(g_i) u(g_i, t^q)) + O(h) - (Ec)_i^{t_{i,q}^{r_i+1}} \right]. \end{aligned}$$

La formule (3.7) est ainsi vérifiée. Pour finir la démonstration de la consistance, supposons que la solution est synchronisée à l'instant  $t^{q+1}$ . Pour tout  $i$ ,  $\delta_i^{q+1} = t^{q+1} - t_{i,q}^{k_i}$  avec  $k_i = E\left(\frac{\Delta t}{\Delta t_i}\right)$  où  $\Delta t = \max_j(\Delta t_j)$ . L'application de la formule (3.7) donne pour tout  $i$  :

$$u(g_i, t^{q+1}) = u(g_i, t^q) - \Delta t \left[ \nabla \cdot (v(g_i) u(g_i, t^q)) + O(h) - (Ec)_i^{t_{i,q}^{q+1}} \right].$$

Pour tout  $i$ , l'erreur de consistance  $(Ec)_i^{t^{q+1}}$  est donc donnée par :

$$\begin{aligned}(Ec)_i^{t^{q+1}} &= \frac{u(g_i, t^{q+1}) - u(g_i, t^q)}{\Delta t} + \nabla \cdot (v(g_i)u(g_i, t^q)) + O(h) \\ &= O(\Delta t) + O(h).\end{aligned}$$

□

### 3.3.2. Consistance : maillage régulier

Nous supposons que le maillage est régulier. Autrement dit, nous commençons par définir un maillage uniforme de  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ . Soit donc  $\mathcal{T}_R = \{\tilde{K}_i : i = 1, \dots, N\}$  une partition uniforme de  $\Omega$  en polyèdres  $\tilde{K}_i$  avec  $G_i$  le centre de gravité de  $\tilde{K}_i$  pour tout  $i$ . Alors pour tout  $j \in \mathcal{N}(i)$ ,  $|G_i - G_j|$  est une constante. Une fois le maillage uniforme construit, nous introduisons une fonction de maillage  $\varphi$  régulière qui vérifie pour tout  $i$ ,  $g_i = \varphi(G_i)$  où les  $g_i$  sont les centres de gravité des polyèdre  $K_i$  qui forment une partition régulière de  $\Omega$ . Pour plus de lisibilité et pour limiter le nombre de cas à traiter, nous supposons que la vitesse est constante  $v = 1$ .

**Proposition 5.** *Dans le cas d'un maillage régulier, le schéma upwind asynchrone est consistant au premier ordre.*

*Démonstration.* Reprenons les résultats de la démonstration précédente. Ce qui change est le calcul de  $I_2 + I_3$ . En fait, dans ce cas il suffit de noter que

$$I_2 + I_3 = \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j}(u(g_i, t^q) - u(g_{i,j}, t^q)) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j}(u(g_j, t^q) - u(g_{i,j}, t^q))$$

Par suite,

$$\begin{aligned}I_2 + I_3 &= \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j}(g_i - g_{i,j})\nabla \cdot u(g_i, t^q) + O(h^3) \\ &\quad + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j}(g_{i,j} - g_i)\nabla \cdot u(g_i, t^q) + O(h^3) \\ &\quad + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j}(g_i - 2g_{i,j} + g_j)\nabla \cdot u(g_i, t^q) + O(h^3) \\ &= O(h^3),\end{aligned}$$

ce qui achève la preuve. □

## 3.4. Convergence : méthode de l'erreur corrigée

Rappelons qu'un schéma numérique est convergent s'il est stable et consistant (Théorème de Lax-Richtmyer). Malheureusement l'hypothèse de consistance n'est pas toujours vérifiée malgré la convergence du schéma. Donc l'erreur de convergence se comporte mieux que l'erreur de la consistance et le

schéma est dit supra-convergent. c'est notamment le cas pour le schéma volume finis upwind. Des estimations d'erreur ont été développées sous certaines hypothèses de régularité de maillage. Cockburn et al. [28] et Vila [30] ont montré que l'erreur est en  $O(h^{1/4})$ . Dans des travaux récents [29], Cockburn et Gremaud prouvent sous certaines conditions de régularité de maillage que l'erreur est en  $O(h^{1/2})$ . Vila dans [31] a montré une consistance faible du schéma en  $O(h^{1/2})$ . Després dans [27] a utilisé des estimations d'énergie pour prouver que l'erreur est en  $O(h^{1/2})$  pour l'équation d'advection dans un maillage régulier avec des conditions limites périodiques. Récemment, Bouche et al. ont étudié dans [32] l'erreur de convergence de la méthode de volumes finis (upwind) explicite sur un domaine polygonal de  $\mathbb{R}^d$ . Ils ont commencé par construire une famille de "correcteurs géométriques" qui dépendent seulement du maillage et de vecteurs d'advection et pas de la solution de l'équation. Puis ils ont montré une "erreur corrigée" pour conclure finalement que l'erreur du schéma est en  $O(h)$ . Dans la suite nous appliquerons la méthode définie dans [32] pour prouver la convergence du schéma asynchrone.

Soit  $u(g_i, t^n)$  la solution exacte de (3.1) au centre de gravité  $g_i$  du volume  $K_i$  à l'instant  $t^n$  supposée de classe  $\mathcal{C}^2$ . Commençons tout d'abord par démontrer la proposition suivante :

**Proposition 6.** *Supposons que le maillage est régulier de sorte que pour tout  $i$ ,  $\Delta t_i - \Delta t_j = O(h^2)$   $j \in \mathcal{N}(i)$  alors*

$$(Ec)_i^{t_i^q} = \frac{u(g_i, t_{i,q}^{r_i}) - u(g_i, t^q)}{r_i \Delta t_i} + v \cdot \nabla u(g_i, t^q) + \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v \cdot N_{i,j} (g_i - g_{i,j}) \nabla u(g_i, t^q) \right. \quad (3.8)$$

$$\left. - \sum_{j \in \mathcal{N}^-(i)} v \cdot N_{i,j} (g_j - g_{i,j}) \nabla u(g_i, t^q) + O(h^{d+1}) \right]. \quad (3.9)$$

*Démonstration.* La preuve se fait par récurrence. Soit donc  $t_q^1 = \min_j(\Delta t_j) = t_{i,q}^1$ . Alors par définition du schéma asynchrone (3.3) on a :

$$\begin{aligned} (Ec)_i^{t_{i,q}^1} &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v \cdot N_{i,j} u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} v \cdot N_{i,j} u(g_j, t^q) \right) \\ &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v \cdot N_{i,j} (u(g_i, t^q) - u(g_{i,j}, t^q)) \right) \\ &\quad - \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^-(i)} v \cdot N_{i,j} (u(g_j, t^q) - u(g_{i,j}, t^q)) \right) - \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}(i)} v \cdot N_{i,j} u(g_{i,j}, t^q) \right) \end{aligned}$$

Par suite,

$$\begin{aligned}
(Ec)_i^{t_{i,q}^1} &= \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}(i)} v.N_{i,j} u(g_{i,j}, t^q) \right) \\
&\quad - \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} (g_i - g_{i,j}) \nabla u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) \nabla u(g_j, t^q) \right) \\
&\quad - O(h) - \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) \nabla u(g_j, t^q) + O(h^{d+1}) \right).
\end{aligned}$$

Pour conclure il suffit de noter que :

$$\sum_{j \in \mathcal{N}(i)} v.N_{i,j} u(g_{i,j}, t^q) = |K_i| v \cdot \nabla u(g_i, t^q) + O(h^{d+1}).$$

Supposons que la formule (3.8) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soient  $t_q^{n+1} = t_{i,q}^{r_i+1}, t_i^{pcd}$  le temps de la dernière mise à jour de  $u_i$  et  $\delta_i^r = t_{i,q}^{r_i+1} - t_i^{pcd}$ . Alors, en utilisant l'hypothèse de récurrence nous avons :

$$u(g_i, t_{i,q}^{r_i+1}) = u(g_i, t_i^{pcd}) - \frac{\delta_i^r}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} u(g_i, t_{i,q}^{r_i}) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} u(g_j, t_{j,q}^{r_j}) \right) + \delta_i^r (Ec)_i^{t_{i,q}^{r_i+1}},$$

alors,

$$\begin{aligned}
u(g_i, t_{i,q}^{r_i+1}) &= u(g_i, t_i^{pcd}) - \frac{\delta_i^r}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} u(g_j, t^q) \right) \\
&\quad + X_1 + X_2 + X_3 + X_4 + \delta_i^r (Ec)_i^{t_{i,q}^{r_i+1}}
\end{aligned}$$

avec,

$$\begin{aligned}
X_1 &= \frac{\delta_i^r}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} r_i \Delta t_i v.N_{i,j} \nabla u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} r_j \Delta t_j v.N_{i,j} \nabla u(g_j, t^q) \right) \\
&= \delta_i^r O(h),
\end{aligned}$$

$$\begin{aligned}
X_2 &= \frac{\delta_i^r}{|K_i|} \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \frac{r_i \Delta t_i}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} (g_i - g_{i,j}) \nabla u(g_i, t^q) \right. \\
&\quad \left. + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) \nabla u(g_i, t^q) \right] = \delta_i^r O(h),
\end{aligned}$$

$$X_3 = \frac{\delta_i^r}{|K_i|} \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \frac{r_j \Delta t_j}{|K_j|} \left[ \sum_{k \in \mathcal{N}^+(j)} v.N_{j,k} (g_j - g_{j,k}) \nabla u(g_j, t^q) \right. \\ \left. + \sum_{k \in \mathcal{N}^-(j)} v.N_{j,k} (g_k - g_{j,k}) \nabla u(g_j, t^q) \right] = \delta_i^r O(h).$$

et

$$X_4 = \frac{\delta_i^r}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} r_i \Delta t_i v.N_{i,j} (Ec)_i^{t_i^q} + \sum_{j \in \mathcal{N}^-(i)} r_j \Delta t_j v.N_{i,j} (Ec)_j^{t_j^q} \right) \\ = \delta_i^r O(h),$$

Pour conclure il suffit d'appliquer l'étape précédente et l'hypothèse de récurrence.  $\square$

**Théorème 2.** *Sous les conditions de la proposition 6 et sous la condition CFL locale (3.4), le schéma upwind asynchrone est convergent de premier ordre.*

*Démonstration.* Pour prouver le théorème 2, la méthode de l'erreur corrigée développée dans [32] a été adaptée à l'aspect asynchrone. Soient l'erreur  $\alpha_i^{t^n} = u_i^{t^n} - u(g_i, t^n)$  et l'erreur corrigée  $\bar{\alpha}_i = \alpha_i + \gamma_i$ , avec  $\gamma_i^n = -\Gamma_i \cdot \nabla u(g_i, t^n)$  et  $(\Gamma_i)_{i=1, N}$  une suite de correcteurs géométriques qui vérifie :

$$\sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \Gamma_i + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \Gamma_j = \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} (g_{i,j} - g_i) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_{i,j} - g_j), \quad (3.10)$$

et

$$\|\Gamma\|_{asyn, \infty} \leq Ch, \quad C \text{ une constante.} \quad (3.11)$$

Nous voulons montrer l'inégalité suivante

$$\|\bar{\alpha}^{t^n}\|_{asyn, \infty} \leq \|\bar{\alpha}^{t^q}\|_{asyn, \infty} + O(h). \quad (3.12)$$

Pour le faire, nous utilisons à nouveau une récurrence. Supposons donc que  $t_{i,q}^1 = \min_j (\Delta t_j) := \Delta t_i$  et que  $\alpha_i^{t_{i,q}^1} = u_i^{t_{i,q}^1} - u(g_i, t_{i,q}^1)$  alors

$$\alpha_i^{t_{i,q}^1} = \alpha_i^{t^q} - \frac{\Delta t_i}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \alpha_j^{t^q} + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \alpha_j^{t^q} \right) - \Delta t_i E_i^{t_{i,q}^1},$$

avec  $E_i^{t_{i,q}^1}$  l'erreur de consistance locale définie comme suit

$$E_i^{t_{i,q}^1} = \frac{u(g_i, t_{i,q}^1) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} u(g_j, t^q) \right)$$

Or  $\bar{\alpha}_i = \alpha_i + \gamma_i$ , alors

$$\bar{\alpha}_i^{t_i^1, q} = \bar{\alpha}_i^{t^q} - \frac{\Delta t_i}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \bar{\alpha}_i^{t^q} + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \bar{\alpha}_j^{t^q} \right) - \Delta t_i \bar{E}_i^{t_i^1, q},$$

avec

$$\begin{aligned} \bar{E}_i^{t_i^1, q} &= E_i^{t_i^1, q} - \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \gamma_i^{t^q} + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \gamma_j^{t^q} \right) - \frac{\gamma_i^{t_i^1, q} - \gamma_i^{t^q}}{\Delta t_i} \\ &= \frac{u(g_i, t_i^1, q) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \sum_{j \in \mathcal{N}(i)} v.N_{i,j} u(g_{i,j}, t^q) - \frac{\gamma_i^{t_i^1, q} - \gamma_i^{t^q}}{\Delta t_i} \\ &\quad + \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \left( u(g_i, t^q) - u(g_{i,j}, t^q) - \gamma_i^{t^q} \right) \right. \\ &\quad \left. + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \left( u(g_j, t^q) - u(g_{i,j}, t^q) - \gamma_j^{t^q} \right) \right] \\ &= G_1 + G_2 + G_3. \end{aligned}$$

Notons que

$$\begin{aligned} \sum_{j \in \mathcal{N}(i)} v.N_{i,j} u(g_{i,j}, t^q) &= \sum_{j \in \mathcal{N}(i)} v.N_{i,j} (u(g_{i,j}, t^q) - u(g_i, t^q)) \\ &= \sum_{j \in \mathcal{N}(i)} v.N_{i,j} (g_{i,j} - g_i) \cdot \nabla u(g_i, t^q) + O(h^{d+1}) \\ &= \left( \sum_{j \in \mathcal{N}(i)} v.N_{i,j} g_{i,j} \right) \cdot \nabla u(g_i, t^q) + O(h^{d+1}) \\ &= |K_i| v \cdot \nabla u(g_i, t^q) + O(h^{d+1}). \end{aligned}$$

Par suite

$$G_1 = \frac{u(g_i, t_i^1, q) - u(g_i, t^q)}{\Delta t_i} + \frac{1}{|K_i|} \sum_{j \in \mathcal{N}(i)} v.N_{i,j} u(g_{i,j}, t^q) = O(h).$$

De plus,

$$u(g_i, t^q) - u(g_{i,j}, t^q) = (g_i - g_{i,j}) \nabla u(g_i, t^q) + O(h^2),$$

et

$$u(g_j, t^q) - u(g_{i,j}, t^q) = (g_j - g_{i,j}) \nabla u(g_j, t^q) + O(h^2).$$

Alors,

$$G_2 = \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \left( u(g_i, t^q) - u(g_{i,j}, t^q) - \gamma_i^{t^q} \right) \right. \\ \left. + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \left( u(g_j, t^q) - u(g_{i,j}, t^q) - \gamma_j^{t^q} \right) \right]$$

L'expression de  $G_2$  est par conséquence donnée par la formule suivante :

$$G_2 = \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} (g_i - g_{i,j}) \nabla u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) \nabla u(g_j, t^q) \right) \\ + O(h) + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \Gamma_i^{t^q} \nabla u(g_i, t^q) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \Gamma_j^{t^q} \nabla u(g_j, t^q) \right) \\ = \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} (g_i - g_{i,j}) + \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \Gamma_i^{t^q} \right) \nabla u(g_i, t^q) \\ + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \Gamma_j^{t^q} \right) \nabla u(g_i, t^q) \\ + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) (\nabla u(g_j, t^q) - \nabla u(g_i, t^q)) \right) \\ + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \Gamma_j^{t^q} (\nabla u(g_j, t^q) - \nabla u(g_i, t^q)) \right).$$

Rappelons que  $N_{i,k} = |K_j \cap K_i| n_{i,k}$ ,  $n_{i,j}$  la normale unitaire de  $K_i \cap K_j$  pointant de  $K_i$  vers  $K_j$ .

De plus,  $\Gamma$  vérifie (3.10) et (3.11) donc  $G_2 = O(h)$ . Enfin,  $G_3 = -\frac{\gamma_{i,q}^{t^q} - \gamma_i^{t^q}}{\Delta t_i} = O(h)$ . Par conséquent,  $\bar{E}_i^{t^q} = O(h)$ . Supposons que la formule (3.12) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soient  $t_q^{n+1} = t_{i,q}^{r_i+1}$ ,  $t_i^{pcd}$  le temps de la dernière mise à jour de  $u_i$  et  $\delta_i^r = t_{i,q}^{r_i+1} - t_i^{pcd}$ .

$$\alpha_i^{t_{i,q}^{r_i+1}} = \alpha_i^{t_i^{pcd}} - \frac{\delta_i^r}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \alpha_i^{t_{i,q}^{r_i}} + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \alpha_j^{t_{j,q}^{r_j}} \right) - \delta_i^r E_i^{t_{i,q}^{r_i+1}},$$

avec  $E_i^{t_{i,q}^{r_i+1}}$  l'erreur de consistance locale définie comme suit

$$E_i^{t_{i,q}^{r_i+1}} = \frac{u(g_i, t_{i,q}^{r_i+1}) - u(g_i, t_i^{pcd})}{\delta_i^r} + \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} u(g_i, t_{i,q}^{r_i}) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} u(g_j, t_{j,q}^{r_j}) \right]$$

On a  $\bar{\alpha}_i = \alpha_i + \gamma_i$ , alors

$$\bar{\alpha}_i^{t_i^{r_i+1}} = \bar{\alpha}_i^{t_i^{pcd}} - \frac{\delta_i^r}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \bar{\alpha}_i^{t_i^{r_i}} + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \bar{\alpha}_j^{t_j^{r_j}} \right) - \delta_i^r \bar{E}_i^{t_i^{r_i+1}},$$

avec

$$\begin{aligned} \bar{E}_i^{t_i^{r_i+1}} &= E_i^{t_i^{r_i+1}} - \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \gamma_i^{t_i^{r_i}} + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \gamma_j^{t_j^{r_j}} \right) - \frac{\gamma_i^{t_i^{r_i+1}} - \gamma_i^{t_i^{pcd}}}{\delta_i^r} \\ &= \frac{u(g_i, t_i^{r_i+1}) - u(g_i, t_i^{pcd})}{\delta_i^r} - \frac{\gamma_i^{t_i^{r_i+1}} - \gamma_i^{t_i^{pcd}}}{\delta_i^r} \\ &\quad + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \left( u(g_i, t_i^{r_i}) - \gamma_i^{t_i^{r_i}} \right) + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \left( u(g_j, t_j^{r_j}) - \gamma_j^{t_j^{r_j}} \right) \right) \\ &= -v.\nabla u(g_i, t^q) + \frac{1}{|K_i|} \left( \sum_{j \in \mathcal{N}(i)} v.N_{i,j} u(g_{i,j}, t^q) \right) \\ &\quad + \frac{1}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \left( u(g_i, t^q) - u(g_{i,j}, t^q) - \gamma_i^{t^q} \right) \right. \\ &\quad + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \left( u(g_j, t^q) - u(g_{i,j}, t^q) - \gamma_j^{t^q} \right) - \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} r_i \Delta t_j v.\nabla u(g_i, t^q) \\ &\quad - \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} r_j \Delta t_j v.\nabla u(g_j, t^q) - \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} (g_i - g_{i,j}) \nabla u(g_i, t^q) \\ &\quad \left. - \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) \nabla u(g_i, t^q) \right] \\ &\quad - \frac{1}{|K_i|} \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \frac{r_i \Delta t_i}{|K_i|} \left[ \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} (g_i - g_{i,j}) \nabla u(g_i, t^q) \right. \\ &\quad \left. + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} (g_j - g_{i,j}) \nabla u(g_i, t^q) \right] + E_1, \end{aligned}$$

avec

$$\begin{aligned} E_1 &= -\frac{1}{|K_i|} \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \frac{r_j \Delta t_j}{|K_j|} \left[ \sum_{k \in \mathcal{N}^+(j)} v.N_{j,k} (g_j - g_{j,k}) \nabla u(g_j, t^q) \right. \\ &\quad \left. + \sum_{k \in \mathcal{N}^-(j)} v.N_{j,k} (g_k - g_{j,k}) \nabla u(g_j, t^q) \right]. \end{aligned}$$

En utilisant l'hypothèse de récurrence, l'étape précédente et la proposition 6 on peut déduire que

$$\bar{E}_i^{t_i^{r_i+1}} = O(h).$$

Par conséquent,

$$\bar{\alpha}_i^{t_i^{r_i+1},q} = \bar{\alpha}_i^{t_i^{pcd}} - \frac{\delta_i^r}{|K_i|} \left( \sum_{j \in \mathcal{N}^+(i)} v.N_{i,j} \bar{\alpha}_i^{t_i^{r_i},q} + \sum_{j \in \mathcal{N}^-(i)} v.N_{i,j} \bar{\alpha}_j^{t_j^{r_j},q} \right) - \delta_i^r O(h).$$

L'inégalité (3.12) est donc vérifiée, ce qui achève la preuve du Théorème 2.  $\square$

### 3.5. Conclusion

Nous avons étudié dans ce chapitre les propriétés de la méthode numérique asynchrone appliquée à l'équation de transport en dimension quelconque de l'espace. Pour la discrétisation spatiale, nous avons choisi un schéma volumes finis upwind. Nous avons commencé par démontrer la stabilité de la méthode. Ensuite, nous nous sommes intéressés à la consistance de la méthode. Pour ce faire, nous avons supposé dans un premier temps que la vitesse est variable et que le maillage est constant et dans ce cas, l'aspect asynchrone est dû à la variation de la vitesse. Le but était d'éviter les problèmes qui peuvent être liés à la variation du maillage. Dans un deuxième temps nous avons traité la consistance dans le cadre d'un maillage assez régulier. Dans le cadre d'un maillage variable, le schéma est localement inconsistant. Cette perte locale de la consistance n'est pas liée à l'aspect asynchrone mais c'est un problème dû à la discrétisation spatiale. En effet, l'utilisation du schéma upwind avec une intégration temporelle synchrone classique conduit aussi à une perte locale de la consistance malgré la convergence de la méthode. Nous avons donc adapté la méthode de l'erreur corrigée à l'aspect asynchrone pour pouvoir prouver que notre schéma est convergent à l'ordre 1.

Dans les deux premiers chapitres nous avons adapté le schéma Euler explicite à l'aspect asynchrone pour la discrétisation temporelle et nous avons choisi un schéma upwind pour la discrétisation spatiale. Ces deux schémas sont classiquement d'ordre 1, respectivement en temps et en espace. Le chapitre suivant sera consacré à l'application de la méthode asynchrone dans le cas des schémas qui sont classiquement d'ordre plus élevé.

## Chapitre 4

# Propriétés additionnelles du schéma asynchrone

Nous avons démontré dans les chapitres précédents que le schéma numérique asynchrone conserve l'ordre de convergence de schémas qui sont classiquement d'ordre un, en l'occurrence le schéma Euler explicite pour la discrétisation temporelle et le schéma volumes finis upwind pour l'intégration spatiale. Dans ce chapitre nous proposons d'appliquer le formalisme asynchrone tel qu'il a été défini dans le premier chapitre à des schémas qui sont classiquement d'ordre deux dans le formalisme synchrone. Nous nous intéresserons en particulier à développer des schémas de type Runge-Kutta 2, MUSCL et Lax-Wendroff asynchrones. Puis nous étudierons ce type de schémas d'un point de vue théorique mais également numérique. Nous reprenons pour cela de nouveau l'exemple de l'équation d'advection 1D.

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(v(x)u) = 0, & (x, t) \in [0, L] \times [0, T] \\ u(x, 0) = u_0(x), & x \in [0, L], \end{cases} \quad (4.1)$$

où  $[0, L] \subset \mathbb{R}$  est un domaine périodique de période  $L$ , ( $u(x, t) = u(x + L, t)$ ),  $T$  est le temps final et  $v$  est la vitesse de propagation, supposée connue et régulière. La donnée initiale  $u_0$  est supposée suffisamment régulière.

### 4.1. Schéma Runge Kutta 2 asynchrone (A1RK2)

Pour la discrétisation du problème (4.1) nous utilisons dans un premier temps le schéma volumes finis upwind en espace puis nous rajoutons une discrétisation en temps en utilisant un schéma RK2 asynchrone. En reprenant les notations explicitées dans le chapitre 2, le système semi-discrétisé s'écrit sous la forme suivante :

$$\frac{du_i}{dt} = -\frac{1}{|C_i|} \left( v(x_{i+1/2})u(x_{i+1/2}, t) - v(x_{i-1/2})u(x_{i-1/2}, t) \right). \quad (4.2)$$

Pour plus de lisibilité, nous supposons pour le moment que la vitesse est constante  $v = 1$ . Donc l'équation (4.2) devient

$$\frac{du_i}{dt} = -\frac{1}{|C_i|} \left( u(x_{i+1/2}, t) - u(x_{i-1/2}, t) \right). \quad (4.3)$$

Le schéma asynchrone est dérivé du schéma Runge Kutta 2 classique. Nous aurons donc besoin d'introduire le temps qui correspond à la mise à jour de la pente au point milieu de l'intervalle de temps. Ceci correspond dans le cas synchrone au "demi pas de temps" permettant d'évaluer la pente. Dans le cas asynchrone, la suite  $(t^p)_{p \geq 1}$  des temps successifs de mise à jour des événements les plus urgents entre le temps initial  $t^0$  et le temps final  $T$  est donc définie comme suit (voir aussi sur la figure 4.1, à droite) :

$$t^1 = \min_i \left( \frac{\Delta t_i}{2} \right) := \frac{\Delta t_j}{2}, \quad t^2 = \min \left( \min_{i \neq j} \left( \frac{\Delta t_i}{2} \right), \Delta t_j \right), \dots$$

et la suite  $(t_i^k)_k$  de temps de mise à jours de la cellule  $C_i$  (Fig4.1) est définie par  $t_i^k = k\Delta t_i$ . Les demi pas de temps sont donc naturellement définis par  $t_i^{k+1/2} = (k+1/2)\Delta t_i$  (voir sur la figure 4.1 à gauche).

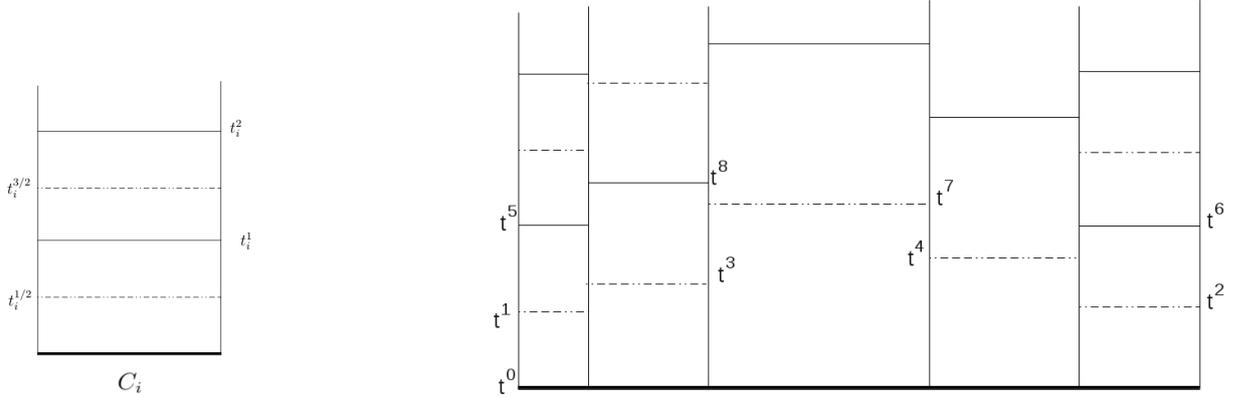


FIGURE 4.1 – Suite de temps de mise à jour d'une cellule  $i$  (gauche), suite de temps de mise à jour des événements les plus urgents (droite).

#### 4.1.1. L'algorithme asynchrone

Par rapport à l'algorithme défini dans le chapitre deux, le temps courant  $t_{cur}$  de la simulation est étendu, il correspond soit à un temps multiple du pas de temps de la cellule la plus urgente à mettre à jour, soit à un demi pas de temps pour l'évaluation de la pente. L'algorithme asynchrone est alors modifié de la façon suivante :

1. À l'instant  $t_{cur} = (r_i + 1/2)\Delta t_i$ , on a les étapes suivantes de la méthode asynchrone :

- (a) Calcul des nouvelles valeurs des densités  $u_i$  et  $u_{i+1}$  :

- i.  $u_i^{t_i^{r_i+1/2}} = u_i^{t_i^{pcd}} - \frac{t_{cur} - t_i^{pcd}}{\Delta x_i} \left( F_i^{t_i^{r_i}} - F_{i-1}^{t_i^{r_i-1}} \right)$  où  $t_{i-1}^{r_i-1} < t_i^{r_i+1/2}$ .

- ii.  $u_{i+1}^{t_i^{r_i+1/2}} = u_{i+1}^{t_{i+1}^{pcd}} - \frac{t_{cur} - t_{i+1}^{pcd}}{\Delta x_{i+1}} \left( F_{i+1}^{t_{i+1}^{r_{i+1}}} - F_i^{t_i^{r_i}} \right)$ ,  $t_{i+1}^{r_{i+1}} < t_i^{r_i+1/2}$ .

- (b) Mise à jour de la valeur du flux  $F_i$

$$F_i^{t_i^{r_i+1/2}} = u_i^{t_i^{r_i+1/2}} \max(0, v_{i+1/2}) + u_{i+1}^{t_i^{r_i+1/2}} \min(0, v_{i+1/2}).$$

- (c) Nous distinguons deux cas :

- i. Si  $t_i^{pcd} > t_{i+1}^{pcd}$  alors  $u_{i+1}^{t_i^{pcd}} = u_{i+1}^{t_{i+1}^{pcd}} - \frac{t_i^{pcd} - t_{i+1}^{pcd}}{\Delta x_{i+1}} \left( F_{i+1}^{r_{i+1}} - F_i^{r_i} \right)$  et  $t_{i+1}^{pcd} = t_i^{pcd}$ .
- ii. Si  $t_i^{pcd} < t_{i+1}^{pcd}$  alors  $u_i^{t_{i+1}^{pcd}} = u_i^{t_i^{pcd}} - \frac{t_{i+1}^{pcd} - t_i^{pcd}}{\Delta x_i} \left( F_i^{r_i} - F_{i-1}^{r_{i-1}} \right)$  et  $t_i^{pcd} = t_{i+1}^{pcd}$ .

(d) Pour tout  $k \notin \{i, i+1\}$ , les densités  $u_k$  ne sont pas impactées (dans cette étape) de même que les flux  $F_k$  pour tout  $k \neq i$ .

Notons que si on suppose que  $r_{i-1}\Delta t_{i-1} < t_{cur} := (r_i + 1/2)\Delta t_i < (r_{i-1} + 1)\Delta t_{i-1}$  et que  $r_{i+1}\Delta t_{i+1} < (r_i + 1/2)\Delta t_i < (r_{i+1} + 1)\Delta t_{i+1}$  alors  $t_i^{pcd} = \max(r_i\Delta t_i, r_{i-1}\Delta t_{i-1}, r_{i+1}\Delta t_{i+1})$ . De même pour  $t_{i+1}^{pcd}$ .

2. Pour  $t_{cur} = (r_i + 1)\Delta t_i$ , on a les étapes suivantes de l'asynchrone

(a) Calcul des nouvelles valeurs des densités  $u_i$  et  $u_{i+1}$  :

- i.  $u_i^{r_i+1} = u_i^{t_i^{pcd}} - \frac{t_{cur} - t_i^{pcd}}{\Delta x_i} \left( F_i^{r_i+1/2} - F_{i-1}^{r_{i-1}+1/2} \right)$
- ii.  $u_{i+1}^{r_i+1} = u_{i+1}^{t_{i+1}^{pcd}} - \frac{t_{cur} - t_{i+1}^{pcd}}{\Delta x_{i+1}} \left( F_{i+1}^{r_i+1+1/2} - F_i^{r_i+1/2} \right)$ .

Avec  $t_i^{pcd} = \max(r_i\Delta t_i, r_{i-1}\Delta t_{i-1}, r_{i+1}\Delta t_{i+1})$ .

(b) Mise à jour de la valeur du flux  $F_i$

- i.  $F_i^{r_i+1} = u_i^{r_i+1} \max(0, v_{i+1/2}) + u_{i+1}^{r_i+1} \min(0, v_{i+1/2})$
- ii.  $F_i^{r_i+1/2} = F_i^{r_i+1}$

(c) Pour tout  $k \notin \{i, i+1\}$ , les densités  $u_k$  ne sont pas impactées (dans cette étape) de même que les flux  $F_k$  pour tout  $k \neq i$ .

#### 4.1.2. Propriétés du schéma RK2 asynchrone (A1RK2)

Pour plus de lisibilité et pour limiter le nombre de cas à traiter, on supposera que la vitesse  $v \equiv 1$  et que le maillage est variable. Comme cela a été fait dans le chapitre précédent, les propriétés de la méthode asynchrone seront étudiées entre deux instants de synchronisation de la solution  $t^q$  et  $t^{q+1} = t^q + \max_i(\Delta t_i)$ .

Soit  $(t_q^n)_{\{1 \leq n \leq N_q\}}$  la suite des temps successifs de mise à jour des événements les plus urgents entre  $t^q$  et  $t^{q+1}$ .  $N_q$  désigne le nombre de pas de temps entre  $t^q$  et  $t^{q+1}$  et pour tout  $i$ , on a  $(t_{i,q}^k)_k = (t^q + k\Delta t_i)_k$ .

##### 4.1.2.1. Stabilité du schéma A1RK2

**Proposition 7.** *Sous la condition CFL locale,  $|v_{i+1/2}| \frac{\Delta t_i}{\Delta x_i} \leq 1$ , le schéma asynchrone est stable pour la norme  $\|\cdot\|_{asyn,\infty}$ . En particulier, la formule suivante est vérifiée :*

$$\|u^{t^{q+1}}\|_{asyn,\infty} \leq \|u^{t^q}\|_{asyn,\infty}, \quad (4.4)$$

où  $\Delta t := \max_i(\Delta t_i)$ .

*Démonstration.* Utilisons pour démontrer la formule une récurrence. Soient donc  $t_i^1 = t^q + \min_j (\Delta t_j) := t^q + \Delta t_i$  et  $t_i^{1/2} := t^q + \frac{\Delta t_i}{2}$ . Alors

$$u_i^{t_i^1, q} = u_i^{t^q} - \frac{\Delta t_i}{\Delta x_i} \left( F_i^{t_i^{1/2}, q} - \tilde{F}_{i-1}^{t_i^{1/2}, q} \right).$$

Deux cas se présentent :

1.  $\Delta t_i \leq \frac{\Delta t_{i-1}}{2}$ , donc  $\tilde{F}_{i-1}^{t_i^{1/2}, q} = F_{i-1}^{t^q}$  :

$$\begin{aligned} u_i^{t_i^1, q} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x_i} \left( F_i^{t_i^{1/2}, q} - F_{i-1}^{t^q} \right) \\ &= \left( 1 - \frac{\Delta t_i}{\Delta x_i} + \frac{(\Delta t_i)^2}{2(\Delta x_i)^2} \right) u_i^{t^q} + \frac{\Delta t_i}{\Delta x_i} \left( 1 - \frac{\Delta t_i}{2\Delta x_i} \right) u_{i-1}^{t^q} \\ &:= \alpha_i u_i^{t^q} + \tilde{\alpha}_{i-1} u_{i-1}^{t^q}. \end{aligned}$$

Notons que grâce à la condition CFL,  $\alpha_i$  et  $\tilde{\alpha}_{i-1}$  sont positifs. De plus on a  $\alpha_i + \tilde{\alpha}_{i-1} = 1$ .

2.  $\frac{\Delta t_{i-1}}{2} \leq \Delta t_i$ , donc  $\tilde{F}_{i-1}^{t_i^{1/2}, q} = F_{i-1}^{t_i^{1/2}, q}$  :

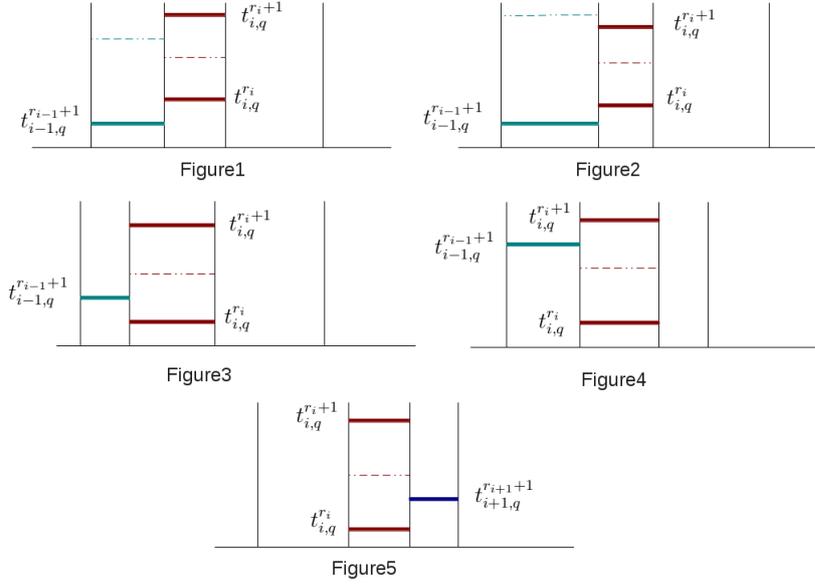
$$\begin{aligned} u_i^{t_i^1, q} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x_i} \left( F_i^{t_i^{1/2}, q} - F_{i-1}^{t_i^{1/2}, q} \right) \\ &= \left( 1 - \frac{\Delta t_i}{\Delta x_i} + \frac{(\Delta t_i)^2}{2(\Delta x_i)^2} \right) u_i^{t^q} \\ &\quad + \frac{\Delta t_i}{\Delta x_i} \left( 1 - \frac{\Delta t_i}{2\Delta x_i} - \frac{\Delta t_{i-1}}{2\Delta x_{i-1}} \right) u_{i-1}^{t^q} + \frac{\Delta t_{i-1}}{2\Delta x_{i-1}} u_{i-2}^{t^q} \\ &= \alpha_i u_i^{t^q} + \alpha_{i-1} u_{i-1}^{t^q} + \alpha_{i-2} u_{i-2}^{t^q}. \end{aligned}$$

Pour  $\Delta t$  vérifiant la condition CFL,  $\alpha_i$ ,  $\alpha_{i-1}$  et  $\alpha_{i-2}$  sont positifs et  $\alpha_i + \alpha_{i-1} + \alpha_{i-2} = 1$ .

Supposons maintenant que la formule (4.4) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soit  $t_q^{n+1} = t_{i,q}^{r_i+1} := t^q + (r_i + 1)\Delta t_i < t^{q+1}$ . Plusieurs cas se présentent, on va donc traiter les cas les plus intéressants. Tous les autres cas se traitent de la même manière ou découlent des cas déjà traités.

1. Supposons que  $t^q + r_{i-1}\Delta t_{i-1} < t^q + r_i\Delta t_i < t^q + (r_{i-1} + 1/2)\Delta t_{i-1} < t^q + (r_i + 1)\Delta t_i$  (Figure1).

$$\begin{aligned} u_i^{t_i^{r_i+1}, q} &= u_i^{t_i^{r_i}, q} - \frac{\Delta t_i}{\Delta x_i} \left( F_i^{t_i^{r_i+1/2}, q} - F_{i-1}^{t_{i-1}^{r_i+1/2}, q} \right) \\ &= u_i^{t_i^{r_i}, q} - \frac{\Delta t_i}{\Delta x_i} \left( u_i^{t_i^{r_i}, q} - u_{i-1}^{t_{i-1}^{r_i-1}, q} \right) + \frac{(\Delta t_i)^2}{2(\Delta x_i)^2} \left( F_i^{t_i^{r_i}, q} - F_{i-1}^{t_{i-1}^{r_i-1}, q} \right) \\ &\quad - \frac{\Delta t_i \Delta t_{i-1}}{2\Delta x_i \Delta x_{i-1}} \left( F_{i-1}^{t_{i-1}^{r_i-1}, q} - F_{i-2}^{t_{i-2}^{r_i-2}, q} \right) \\ &= \alpha_i u_i^{t_i^{r_i}, q} + \alpha_{i-1} u_{i-1}^{t_{i-1}^{r_i-1}, q} + \alpha_{i-2} u_{i-2}^{t_{i-2}^{r_i-2}, q}, \end{aligned}$$



$\alpha_i$ ,  $\alpha_{i-1}$  et  $\alpha_{i-2}$  sont définis au dessus.

2. Supposons que  $t^q + r_{i-1}\Delta t_{i-1} < t^q + r_i\Delta t_i$  et  $t^q + (r_i + 1)\Delta t_i \leq t^q + (r_{i-1} + 1/2)\Delta t_{i-1}$  (Figure2).

$$\begin{aligned}
u_i^{t_{i,q}^{r_i+1}} &= u_i^{t_{i,q}^{r_i}} - \frac{\Delta t_i}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) \\
&= u_i^{t_{i,q}^{r_i}} - \frac{\Delta t_i}{\Delta x_i} \left( u_i^{t_{i,q}^{r_i}} - u_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) + \frac{(\Delta t_i)^2}{2(\Delta x_i)^2} \left( F_i^{t_{i,q}^{r_i}} - F_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) \\
&= \alpha_i u_i^{t_{i,q}^{r_i}} + \tilde{\alpha}_{i-1} u_{i-1}^{t_{i-1,q}^{r_{i-1}}}.
\end{aligned}$$

3. Supposons que  $t^q + r_i\Delta t_i < t^q + r_{i-1}\Delta t_{i-1}$  et  $t^q + (r_i + 1)\Delta t_i \leq t^q + (r_{i-1} + 1/2)\Delta t_{i-1}$ . Soient  $\delta_i^1 = t_{i-1,q}^{r_{i-1}} - t_{i,q}^{r_i}$  et  $\delta_i^2 = t_{i,q}^{r_i+1} - t_{i-1,q}^{r_{i-1}}$  (Figure3).

$$\begin{aligned}
u_i^{t_{i,q}^{r_i+1}} &= u_i^{t_{i-1,q}^{r_{i-1}}} - \frac{\delta_i^2}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) \\
&= u_i^{t_{i,q}^{r_i}} - \frac{\delta_i^1}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i}} - F_{i-1}^{t_{i-1,q}^{r_{i-1}-1/2}} \right) - \frac{\delta_i^2}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) \\
&= \left( 1 - \frac{\Delta t_i}{\Delta x_i} + \frac{\delta_i^2 \Delta t_i}{2(\Delta x_i)^2} \right) u_i^{t_{i,q}^{r_i}} + \frac{\delta_i^1}{\Delta x_i} \left( 1 - \frac{\Delta t_{i-1}}{2\Delta x_{i-1}} \right) u_{i-1}^{t_{i-1,q}^{r_{i-1}-1}} \\
&\quad + \frac{\delta_i^2}{\Delta x_i} \left( 1 - \frac{\Delta t_i}{2\Delta x_i} \right) u_{i-1}^{t_{i-1,q}^{r_{i-1}}} + \frac{\delta_i^1 \Delta t_{i-1}}{2\Delta x_i \Delta x_{i-1}} u_{i-2}^{t_{i-2,q}^{r_{i-2}}} \\
&= \bar{\alpha}_i u_i^{t_{i,q}^{r_i}} + \bar{\alpha}_{i-1}^1 u_{i-1}^{t_{i-1,q}^{r_{i-1}-1}} + \bar{\alpha}_{i-1}^2 u_{i-1}^{t_{i-1,q}^{r_{i-1}}} + \bar{\alpha}_{i-2} u_{i-2}^{t_{i-2,q}^{r_{i-2}}}.
\end{aligned}$$

$\bar{\alpha}_i, \bar{\alpha}_{i-1}^1, \bar{\alpha}_{i-1}^2$  et  $\bar{\alpha}_{i-2}$  sont positifs et

$$\bar{\alpha}_i + \bar{\alpha}_{i-1}^1 + \bar{\alpha}_{i-1}^2 + \bar{\alpha}_{i-2} = 1.$$

4. Supposons que  $t^q + (r_i + 1/2)\Delta t_i < t^q + r_{i-1}\Delta t_{i-1} \leq t^q + (r_i + 1)\Delta t_i$ . Soient  $\delta_i^1 = t_{i-1,q}^{r_i-1} - t_{i,q}^{r_i}$  et  $\delta_i^2 = t_{i,q}^{r_i+1} - t_{i-1,q}^{r_i-1}$  (Figure4).

$$\begin{aligned} u_i^{t_{i,q}^{r_i+1}} &= u_i^{t_{i-1,q}^{r_i-1}} - \frac{\delta_i^2}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_i-1}} \right) \\ &= u_i^{t_{i,q}^{r_i}} - \frac{\delta_i^1}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_i-1/2}} \right) - \frac{\delta_i^2}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_i-1}} \right) \\ &= \left( 1 - \frac{\Delta t_i}{\Delta x_i} + \frac{(\Delta t_i)^2}{2(\Delta x_i)^2} \right) u_i^{t_{i,q}^{r_i}} + \left( \frac{\delta_i^1}{\Delta x_i} - \frac{\delta_i^1 \Delta t_{i-1}}{2\Delta x_i \Delta x_{i-1}} - \frac{(\Delta t_i)^2}{2(\Delta x_i)^2} \right) u_{i-1}^{t_{i-1,q}^{r_i-1}} \\ &\quad + \frac{\delta_i^2}{\Delta x_i} u_{i-1}^{t_{i-1,q}^{r_i-1}} + \frac{\delta_i^1 \Delta t_{i-1}}{2\Delta x_i \Delta x_{i-1}} u_{i-2}^{t_{i-2,q}^{r_i-1}} \\ &= \alpha_i u_i^{t_{i,q}^{r_i}} + \beta_{i-1}^1 u_{i-1}^{t_{i-1,q}^{r_i-1}} + \beta_{i-1}^2 u_{i-1}^{t_{i-1,q}^{r_i-1}} + \bar{\alpha}_{i-2} u_{i-2}^{t_{i-2,q}^{r_i-1}}. \end{aligned}$$

$$\alpha_i + \beta_{i-1}^1 + \beta_{i-1}^2 + \bar{\alpha}_{i-2} = 1.$$

5. On suppose que  $t^q + r_{i+1}\Delta t_{i+1} \leq t^q + (r_i + 1/2)\Delta t_i$ . Soient  $\delta_i^1 = t_{i+1,q}^{r_i+1} - t_{i,q}^{r_i}$  et  $\delta_i^2 = t_{i,q}^{r_i+1} - t_{i+1,q}^{r_i+1}$  (Figure5).

$$\begin{aligned} u_i^{t_{i,q}^{r_i+1}} &= u_i^{t_{i+1,q}^{r_i+1}} - \frac{\delta_i^2}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_i-1}} \right) \\ &= u_i^{t_{i,q}^{r_i}} - \frac{\delta_i^1}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i}} - F_{i-1}^{t_{i-1,q}^{r_i-1}} \right) - \frac{\delta_i^2}{\Delta x_i} \left( F_i^{t_{i,q}^{r_i+1/2}} - F_{i-1}^{t_{i-1,q}^{r_i-1}} \right) \\ &= \left( 1 - \frac{\Delta t_i}{\Delta x_i} + \frac{\delta_i^2 \Delta t_i}{2(\Delta x_i)^2} \right) u_i^{t_{i,q}^{r_i}} + \frac{\Delta t_i}{\Delta x_i} \left( 1 - \frac{\delta_i^2}{2\Delta x_i} \right) u_{i-1}^{t_{i-1,q}^{r_i-1}}. \end{aligned}$$

La formule (4.4) est vérifiée, le schéma A1RK2 est donc stable.  $\square$

#### 4.1.2.2. Consistance du schéma A1RK2

Pour étudier la consistance du schéma, nous introduisons  $\tilde{u}_i^{t_q^n} := u(x_i, t_q^n)$  la solution de l'équation d'advection (4.1), supposée de classe  $\mathcal{C}^2$ . Soient également  $\Delta t_i = |C_i| = x_{i-1/2} - x_{i+1/2}$  et  $dx(i) = x_{i+1} - x_i$ .

**Proposition 8.** *Supposons que pour tout  $i$ ,  $\frac{\Delta t_{i-1}}{2} \leq \Delta t_i \leq 2\Delta t_{i+1}$  et que le maillage est régulier :  $\forall i$ ,  $\Delta t_i - \Delta t_{i-1} = O((\Delta x)^2)$ . Alors sous la condition CFL on a pour tout  $i$*

$$(Ec)_i^{t_q^n} = \frac{\tilde{u}_i^{t_q^n} - \tilde{u}_i^{t_q^q}}{r_i \Delta t_i} + \left( \frac{\partial(\tilde{u}_i^{t_q^q})}{\partial x} + \Delta x_i \varepsilon_i^{r_i} - \frac{r_i \Delta t_i}{2} \frac{\partial^2 \tilde{u}_i^{t_q^q}}{\partial t^2} \right), \quad (4.5)$$

où  $\varepsilon_i^{r_i}$  est borné.

*Démonstration.* La preuve de la consistance est reportée dans l'annexe A. □

## 4.2. Schéma MUSCL asynchrone

Le but dans cette section est d'étudier la possibilité de la montée en ordre en espace en introduisant un schéma volume finis de type MUSCL (Monotonic Upwind Scheme for Conservation Laws). Le schéma semi-discrétisé correspondant s'écrit :

$$\frac{du_i}{dt} = -\frac{1}{|C_i|} (F_i - F_{i-1}), \quad (4.6)$$

$F_i$  et  $F_{i-1}$  sont les flux aux interfaces. La méthode MUSCL consiste à calculer le flux numérique  $F_i$  en fonction de  $u_{i+1/2,L}$  et  $u_{i+1/2,R}$  avec

$$\begin{cases} u_{i+1/2,L} &= u_i + \frac{\Delta x_i}{2} p_i \\ u_{i+1/2,R} &= u_{i+1} - \frac{\Delta x_{i+1}}{2} p_{i+1} \end{cases} \quad (4.7)$$

où pour tout  $i$ ,  $p_i$  représente la pente. Le flux numérique décentré s'écrit donc sous la forme suivante

$$F_{i+1/2} = \max(0, v_{i+1/2}) u_{i+1/2,L} + \min(0, v_{i+1/2}) u_{i+1/2,R}.$$

### 4.2.1. Limiteurs de pente

#### 4.2.1.1. Exemples de limiteurs de pente

Commençons par rappeler quelques exemples des limiteurs de pente [21]

1. Le limiteur minmod :

$$p_i = \text{minmod} \left( \frac{u_{i+1} - u_i}{dx(i)}, \frac{u_i - u_{i-1}}{dx(i-1)} \right),$$

où  $dx(i) = x_{i+1} - x_i := \frac{\Delta x_i}{2} + \frac{\Delta x_{i+1}}{2}$ . La fonction minmod est définie par

$$\text{minmod}(a, b) = \begin{cases} a & \text{si } ab > 0 \text{ et } |a| < |b|, \\ b & \text{si } ab > 0 \text{ et } |b| < |a|, \\ 0 & \text{si } ab < 0. \end{cases}$$

2. Le limiteur superbee :

$$p_i = \text{maxmod}(p_1, p_2),$$

avec

$$p_1 = \text{minmod} \left( \frac{u_{i+1} - u_i}{dx(i)}, 2 \frac{u_i - u_{i-1}}{dx(i-1)} \right), \quad p_2 = \text{minmod} \left( 2 \frac{u_{i+1} - u_i}{dx(i)}, \frac{u_i - u_{i-1}}{dx(i-1)} \right),$$

et la fonction maxdmod est définie par :

$$\text{maxdmod}(a, b) = \begin{cases} a & \text{si } ab > 0 \text{ et } |a| > |b|, \\ b & \text{si } ab > 0 \text{ et } |b| > |a|, \\ 0 & \text{si } ab < 0. \end{cases}$$

3. Le limiteur MC (monotonized central-difference limiter) :

$$p_1 = \text{minmod} \left( \frac{u_{i+1} - u_{i-1}}{dx(i-1) + dx(i)}, 2 \frac{u_{i+1} - u_i}{dx(i)}, 2 \frac{u_i - u_{i-1}}{dx(i-1)} \right).$$

où  $\text{minmod}(a, b, c) = \text{minmod}(\text{minmod}(a, b), c)$ .

Notons que dans tous les cas précédents la pente peut s'écrire à l'aide d'un limiteur sous la forme :

$$p_i = \varphi(R_i) \frac{u_{i+1} - u_i}{dx(i)}, \quad (4.8)$$

où  $R_i$  est le rapport des pentes de part et d'autre de la maille  $C_i$  :

$$R_i := \left( \frac{u_i - u_{i-1}}{dx(i-1)} \right) \times \left( \frac{dx(i)}{u_{i+1} - u_i} \right),$$

et  $\varphi$  est un limiteur de flux. Par exemple pour le limiteur minmod  $\varphi(R) = \text{minmod}(1, R)$  et pour superbee  $\varphi(R) = \max(0, \min(1, 2R), \min(2, R))$ . En utilisant ces notations, (4.7) devient donc

$$\begin{cases} u_{i+1/2,L} &= u_i + \frac{\Delta x_i}{2} \varphi(R_i) \frac{u_{i+1} - u_i}{dx(i)} \\ u_{i+1/2,R} &= u_{i+1} - \frac{\Delta x_{i+1}}{2} \varphi\left(\frac{1}{R_{i+1}}\right) \frac{u_{i+1} - u_i}{dx(i)} \end{cases} \quad (4.9)$$

$$dx(i) = x_{i+1} - x_i := \frac{\Delta x_i}{2} + \frac{\Delta x_{i+1}}{2}.$$

#### 4.2.1.2. Caractéristiques de limiteur $\varphi$

Pour préserver la qualité de l'approximation, le limiteur  $\varphi$  choisi doit nécessairement satisfaire quelques propriétés parmi lesquelles être linéaire et monotone.

1. La monotonie :

**Définition 1.** *L'interpolation (4.9) est monotone si pour tout  $i$ ,*

$$u_{i+1/2,L} \in [\min(u_i, u_{i+1}), \max(u_i, u_{i+1})] \quad \text{et} \quad u_{i+1/2,R} \in [\min(u_i, u_{i+1}), \max(u_i, u_{i+1})].$$

**Proposition 9.** *L'interpolation (4.9) est monotone si et seulement si pour tout  $i$*

$$0 \leq \varphi(R_i) \leq 1 + \frac{\Delta x_{i+1}}{\Delta x_i} \quad \text{et} \quad 0 \leq \varphi\left(\frac{1}{R_i}\right) \leq 1 + \frac{\Delta x_{i-1}}{\Delta x_i}. \quad (4.10)$$

*Démonstration.* Pour tout  $i$

$$u_{i+1/2,L} = \left(1 - \frac{\Delta x_i}{2dx(i)} \varphi(R_i)\right) u_i + \frac{\Delta x_i}{2dx(i)} \varphi(R_i) u_{i+1},$$

c'est une combinaison convexe si et seulement si pour tout  $i$

$$0 \leq \varphi(R_i) \leq 1 + \frac{\Delta x_{i+1}}{\Delta x_i}.$$

De plus,

$$u_{i+1/2,R} = \left(1 - \frac{\Delta x_{i+1}}{2dx(i)} \varphi\left(\frac{1}{R_{i+1}}\right)\right) u_{i+1} + \frac{\Delta x_{i+1}}{\Delta x_i} \varphi\left(\frac{1}{R_{i+1}}\right) u_i.$$

C'est une combinaison convexe si et seulement si

$$0 \leq \varphi\left(\frac{1}{R_{i+1}}\right) \leq 1 + \frac{\Delta x_i}{\Delta x_{i+1}}.$$

□

2. La linéarité :

**Définition 2.** La condition de linéarité de l'interpolation (4.9) s'exprime sous la forme

$$u_{i-1/2,R} + u_{i+1/2,L} = 2u_i. \quad (4.11)$$

**Proposition 10.** L'interpolation (4.9) est linéaire au sens de (4.11) si et seulement si pour tout  $i$ ,

$$\varphi\left(\frac{1}{r}\right) = \frac{1}{r} \varphi(r), \quad \forall r \in \mathbb{R}^*. \quad (4.12)$$

*Démonstration.* On a pour tout  $i$

$$\begin{aligned} u_{i-1/2,R} + u_{i+1/2,L} &= u_i - \frac{\Delta x_i}{2} \varphi\left(\frac{1}{R_i}\right) \frac{u_i - u_{i-1}}{dx(i-1)} + u_i + \frac{\Delta x_i}{2} \varphi(R_i) \frac{u_{i+1} - u_i}{dx(i)} \\ &= 2u_i + \frac{\Delta x_i}{2} \frac{u_{i+1} - u_i}{dx(i)} \left( \varphi(R_i) - R_i \varphi\left(\frac{1}{R_i}\right) \right). \end{aligned}$$

□

#### 4.2.2. Algorithme asynchrone

Pour la discrétisation en temps, nous utilisons tout d'abord le schéma asynchrone d'ordre un défini dans le chapitre 2. Le but est ici d'étudier le comportement de l'asynchrone avec une discrétisation en espace classiquement d'ordre 2. L'algorithme asynchrone avec un schéma MUSCL en espace est le suivant

1. Initialisation : Pour tout  $i$

- (a)  $R_i = \left(\frac{u_i - u_{i-1}}{dx(i-1)}\right) \times \left(\frac{dx(i)}{u_{i+1} - u_i}\right)$  où  $u_i = u_0(x_i)$ .
- (b)  $u_{i+1/2,L} = u_i + \frac{\Delta x_i}{2} \varphi(R_i) \frac{u_{i+1} - u_i}{dx(i)}$ ,  $u_{i+1/2,R} = u_{i+1} - \frac{\Delta x_{i+1}}{2} \varphi\left(\frac{1}{R_{i+1}}\right) \frac{u_{i+1} - u_i}{dx(i)}$ .
- (c)  $F_i = \max(0, v_{i+1/2}) u_{i+1/2,L} + \min(0, v_{i+1/2}) u_{i+1/2,R}$ .
- (d)  $t_i^{pcd} = 0$ , où  $t_i^{pcd}$  est le temps de la dernière mise à jour de la densité  $u_i$ .
- (e)  $t_{cur} = 0$ ,  $t_{cur}$  est le temps courant de la simulation.

2. À chaque instant  $t_{cur} = t_i^{r_i} := r_i \Delta t_i$ , on a les étapes suivantes de l'asynchrone

(a) Mettre à jours les densités  $u_j$  pour tout  $j \in [i-1, i+2]$  :

$$u_j^{t_{cur}} = u_j^{t_j^{pcd}} - \frac{t_{cur} - t_j}{\Delta x_j} \left( F_j^{t_j^{r_j}} - F_{j-1}^{t_{j-1}^{r_{j-1}}} \right), \quad t_j^{r_j} = r_j \Delta t_j, \quad \forall j \in [i-1, i+2].$$

(b) Mettre à jours les pentes  $u_{i+1/2,L}$  et  $u_{i+1/2,R}$ .

(c) Calculer la nouvelle valeur de flux  $F_i$ .

(d) Pour tout  $j \in [i-1, i+2]$ ,  $t_j^{pcd} = t_{cur}$ .

3. Synchronisation de la solution à l'instant final.

Le schéma asynchrone s'écrit à l'instant  $t_i^{r_i} := r_i \Delta t_i$  :

$$u_i^{t_i^{r_i}} = u_i^{t_i^{pcd}} - \frac{t_i^{r_i} - t_i^{pcd}}{\Delta x_i} \left( F_i^{t_i^{r_{i-1}}} - F_{i-1}^{t_{i-1}^{r_{i-1}}} \right), \quad (4.13)$$

avec  $t_i^{pcd}$  est le temps de la dernière mise à jour de la cellule  $i$  et  $t_{i-1}^{r_{i-1}} := r_{i-1} \Delta t_{i-1} \leq t_i^{r_i}$  et

$$\begin{cases} u_{j+1/2,L}^{t_i^{r_i}} = u_j^{t_j^{pcd}} + \frac{\Delta x_j}{2} \varphi \left( R_j^{t_j^{pcd}} \right) \frac{u_{j+1}^{t_{j+1}^{pcd}} - u_j^{t_j^{pcd}}}{dx(j)} \\ u_{j+1/2,R}^{t_i^{r_i}} = u_{j+1}^{t_{j+1}^{pcd}} - \frac{\Delta x_{j+1}}{2} \varphi \left( \frac{1}{R_{j+1}^{t_{j+1}^{pcd}}} \right) \frac{u_{j+1}^{t_{j+1}^{pcd}} - u_j^{t_j^{pcd}}}{dx(j)} \end{cases} \quad (4.14)$$

avec  $t_j^{pcd} = t_i^{r_i}, \forall j \in [i-1, i+2]$  et  $\varphi$  est un limiteur de flux.

### 4.2.3. Stabilité du schéma MUSCL asynchrone

**Proposition 11.** *Sous la condition CFL*

$$|v_{i+1/2}| \frac{\Delta t_i}{\Delta x_i} \leq \frac{1}{2}, \quad (4.15)$$

le schéma asynchrone est stable pour la norme  $\|\cdot\|_{asyn,\infty}$

$$\|u^{t^{q+1}}\|_{asyn,\infty} \leq \|u^{t^q}\|_{asyn,\infty}, \quad (4.16)$$

*Démonstration.* On va montrer la formule (4.16) par récurrence. Soit  $t_q^1 = t^q + \min_i (\Delta t_i) := t^q + \Delta t_i$ .

$$\begin{aligned}
u_i^{t_q^1} &= u_i^{t^q} - \frac{\Delta t_i}{\Delta x_i} (F_i^{t^q} - F_{i-1}^{t^q}) \\
&= u_i^{t^q} - \frac{\Delta t_i}{\Delta x_i} \left( u_i^{t^q} + \frac{\Delta x_i}{2} \varphi(R_i^{t^q}) \frac{u_{i+1}^{t^q} - u_i^{t^q}}{dx(i)} - u_{i-1}^{t^q} - \frac{\Delta x_{i-1}}{2} \varphi(R_{i-1}^{t^q}) \frac{u_i^{t^q} - u_{i-1}^{t^q}}{dx(i-1)} \right) \\
&= u_i^{t^q} - \frac{\Delta t_i}{\Delta x_i} \left( u_i^{t^q} + \frac{\Delta x_i}{2} \frac{\varphi(R_i^{t^q})}{R_i^{t^q}} \frac{u_i^{t^q} - u_{i-1}^{t^q}}{dx(i-1)} - u_{i-1}^{t^q} - \frac{\Delta x_{i-1}}{2} \varphi(R_{i-1}^{t^q}) \frac{u_i^{t^q} - u_{i-1}^{t^q}}{dx(i-1)} \right) \\
&= \left( 1 - \frac{\Delta t_i}{\Delta x_i} \left( 1 + \frac{\Delta x_i}{2dx(i-1)} \frac{\varphi(R_i^{t^q})}{R_i^{t^q}} - \frac{\Delta x_{i-1}}{2dx(i-1)} \varphi(R_{i-1}^{t^q}) \right) \right) u_i^{t^q} \\
&\quad + \frac{\Delta t_i}{\Delta x_i} \left( 1 + \frac{\Delta x_i}{2dx(i-1)} \frac{\varphi(R_i^{t^q})}{R_i^{t^q}} - \frac{\Delta x_{i-1}}{2dx(i-1)} \varphi(R_{i-1}^{t^q}) \right) u_{i-1}^{t^q} \\
&:= \left( 1 - \frac{\Delta t_i}{\Delta x_i} \alpha_i^{t^q} \right) u_i^{t^q} + \frac{\Delta t_i}{\Delta x_i} \alpha_{i-1}^{t^q} u_{i-1}^{t^q}.
\end{aligned}$$

Or, pour tout  $i$ , nous avons

$$\begin{aligned}
\frac{\Delta x_{i-1}}{2dx(i-1)} \varphi(R_{i-1}^{t^q}) &= \frac{\Delta x_{i-1}}{\Delta x_{i-1} + \Delta x_i} \varphi(R_{i-1}^{t^q}) \\
&= \frac{1}{1 + \frac{\Delta x_i}{\Delta x_{i-1}}} \varphi(R_{i-1}^{t^q}) \\
&\leq 1 \\
&\leq 1 + \frac{\Delta x_i}{2dx(i-1)} \frac{\varphi(R_i^{t^q})}{R_i^{t^q}}.
\end{aligned}$$

De plus

$$\begin{aligned}
\alpha_i^{t^q} &:= 1 + \frac{\Delta x_i}{2dx(i-1)} \frac{\varphi(R_i^{t^q})}{R_i^{t^q}} - \frac{\Delta x_{i-1}}{2dx(i-1)} \varphi(R_{i-1}^{t^q}) \\
&\leq 1 + \frac{\Delta x_i}{2dx(i-1)} \frac{\varphi(R_i^{t^q})}{R_i^{t^q}} \\
&\leq 1 + \frac{1}{1 + \frac{\Delta x_{i-1}}{\Delta x_i}} \varphi \left( \frac{1}{R_i^{t^q}} \right) \\
&\leq 2.
\end{aligned}$$

Supposons maintenant que la formule (4.16) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soient  $t_q^{n+1} = t_{i,q}^{r_i+1} := t^q + (r_i + 1)\Delta t_i < t^{q+1}$  et  $t^q + r_{i-1}\Delta t_{i-1} < t_{i,q}^{r_i+1} < t^q + (r_{i-1} + 1)\Delta t_{i-1}$ . Alors

$$\begin{aligned}
u_i^{r_i+1} &= u_i^{r_i} - \frac{\Delta t_i}{\Delta x_i} \left( F_i^{t_i^{r_i,q}} - F_{i-1}^{t_{i-1}^{r_{i-1},q}} \right) \\
&= \left( 1 - \frac{\Delta t_i}{\Delta x_i} \left( 1 + \frac{\Delta x_i}{2dx(i-1)} \frac{\varphi(R_i^{t_i^{r_i,q}})}{R_i^{t_i^{r_i,q}}} \right) \right) u_i^{r_i} + \frac{\Delta t_i}{\Delta x_i} \frac{\Delta x_{i-1}}{2dx(i-1)} \varphi(R_{i-1}^{t_{i-1}^{r_{i-1},q}}) u_i^{t_{i-1}^{r_{i-1},q}} \\
&\quad + \frac{\Delta t_i}{\Delta x_i} \left( 1 - \frac{\Delta x_{i-1}}{2dx(i-1)} \varphi(R_{i-1}^{t_{i-1}^{r_{i-1},q}}) \right) u_{i-1}^{t_{i-1}^{r_{i-1},q}} + \frac{\Delta t_i}{2dx(i-1)} \frac{\varphi(R_i^{t_i^{r_i,q}})}{R_i^{t_i^{r_i,q}}} u_{i-1}^{t_{i-1}^{r_{i-1},q}} \\
&:= \alpha_i^{t_i^{r_i,q}} u_i^{t_i^{r_i,q}} + \alpha_i^{t_{i-1}^{r_{i-1},q}} u_i^{t_{i-1}^{r_{i-1},q}} + \alpha_{i-1}^{t_{i-1}^{r_{i-1},q}} u_{i-1}^{t_{i-1}^{r_{i-1},q}} + \alpha_{i-1}^{t_i^{r_i,q}} u_{i-1}^{t_i^{r_i,q}}.
\end{aligned}$$

En utilisant les caractéristiques du limiteur  $\varphi$  telles que décrites précédemment et sous la condition CFL (4.15), on peut montrer que :  $\alpha_i^{t_i^{r_i,q}}$ ,  $\alpha_i^{t_{i-1}^{r_{i-1},q}}$ ,  $\alpha_{i-1}^{t_{i-1}^{r_{i-1},q}}$  et  $\alpha_{i-1}^{t_i^{r_i,q}}$  sont positifs. De plus,

$$\alpha_i^{t_i^{r_i,q}} + \alpha_i^{t_{i-1}^{r_{i-1},q}} + \alpha_{i-1}^{t_{i-1}^{r_{i-1},q}} + \alpha_{i-1}^{t_i^{r_i,q}} = 1.$$

Le schéma MUSCL asynchrone est donc stable. □

#### 4.2.4. Consistance du schéma MUSCL asynchrone

**Proposition 12.** *Supposons que le maillage est assez régulier de sorte que  $\Delta t_i - \Delta t_{i-1} = O((\Delta x_i)^2)$  pour tout  $i$ . Alors sous la condition CFL (4.15) la formule suivante est satisfaite*

$$(Ec)_i^{t_i^{r_i,q}} = \frac{u_i^{t_i^{r_i,q}} - u_i^{t_i^q}}{t_i^{r_i,q} - t_i^q} + \left( \frac{\partial u_i^{t_i^q}}{\partial x} + O(\Delta x_i) \right). \quad (4.17)$$

Le schéma asynchrone MUSCL est donc consistant d'ordre 1 en temps et en espace.

*Démonstration.* La preuve de cette proposition est placée dans l'annexe A. □

### 4.3. Schéma Lax Wendroff asynchrone

Le schéma Lax Wendroff synchrone classique est d'ordre 2 en temps et en espace [26]. Nous étudions dans cette section la possibilité de conserver l'ordre en passant en mode asynchrone. Rappelons que le schéma Lax Wendroff synchrone s'écrit sous la forme suivante :

$$\begin{cases} u_i^n &= u_i^n - \frac{\Delta t_i}{\Delta x_i} (F_i^n - F_{i-1}^n) \\ F_i^n &= \frac{u_{i+1}^n + u_i^n}{2} - \frac{\Delta t}{2dx(i+1)} (u_{i+1}^n - u_i^n) \end{cases}$$

avec  $\Delta t = \min_i(\Delta t_i)$ . Le schéma Lax-Wendroff asynchrone consiste à calculer le flux numérique  $F_i$  comme suit

$$F_i = \frac{F_{i,L} + F_{i,R}}{2},$$

avec

$$\begin{cases} F_{i,L} &= \frac{u_{i+1}+u_i}{2} - \frac{\Delta t_i}{2dx(i+1)} (u_{i+1} - u_i) \\ F_{i,R} &= \frac{u_{i+1}+u_i}{2} - \frac{\Delta t_{i+1}}{2dx(i+1)} (u_{i+1} - u_i) \end{cases}$$

Dans le cadre d'une intégration asynchrone du schéma Lax-Wendroff, le flux centré  $F_i$  est calculé en utilisant les valeurs des flux  $F_{i,L}$  et  $F_{i,R}$  qui ont été mis à jour à deux instants différents. La condition de stabilité classique n'est donc plus suffisante pour garantir la stabilité du schéma asynchrone. Les calculs préliminaires montrent que, pour une vitesse  $v = 1$ , le pas de temps local doit satisfaire au moins la condition de stabilité suivante :

$$\Delta t_i < \min(\Delta x_i, 2dx(i-1) - dx(i-1)). \quad (4.18)$$

On note que cette condition de stabilité ne dépend pas seulement de la cellule courante mais également des cellules voisines.

#### 4.3.1. Consistance du schéma Lax Wendroff asynchrone

Pour étudier la consistance du schéma, nous introduisons  $\tilde{u}_i^{t_q^n} := u(x_i, t_q^n)$  la solution de l'équation d'advection (4.1), supposée de classe  $\mathcal{C}^2$ .

**Proposition 13.** *Supposons que le maillage est assez régulier de sorte que  $\Delta t_i - \Delta t_{i-1} = O((\Delta x_i)^2)$  pour tout  $i$ . Alors le schéma Lax Wendroff asynchrone est consistant à l'ordre 1.*

*Démonstration.* La Proposition 13 est démontrée dans l'annexe A. □

**Remarque 2.** *(Perte d'ordre) Dans ce qui précède, nous avons intégré des schémas qui sont classiquement d'ordre deux : RK2, MUSCL, Lax-Wendroff en utilisant la méthode asynchrone. L'étude théorique des schémas asynchrones montre une perte d'ordre par rapport à une intégration synchrone. En fait, pour un schéma asynchrone, l'erreur de consistance à un instant  $t_{cur}$ , le temps courant de la simulation, fait apparaître un terme de la forme  $\delta_{asyn} = r_i \Delta t_i - r_{i-1} \Delta t_{i-1}$  où  $r_i \Delta t_i < t_{cur}$  est le temps de la dernière mise à jour du flux  $F_i$  et où  $r_{i-1} \Delta t_{i-1}$  est le temps de la dernière mise à jour du flux  $F_{i-1}$ . En effet, l'algorithme asynchrone a été construit en considérant des mises à jour indépendantes pour chaque flux d'interface.*

*Si le maillage est régulier, de sorte que pour tout  $i$ ,  $\Delta t_i - \Delta t_{i-1} = O((\Delta x)^2)$  alors  $\delta_{asyn} = O((\Delta x)^2)$  nous pouvons garantir la convergence des schémas asynchrones. Ceci explique l'hypothèse de la régularité du maillage imposée dans les propositions précédentes de ce chapitre. L'erreur de consistance d'un schéma numérique asynchrone est donc égale à l'erreur du schéma numérique (l'erreur classique) plus une erreur asynchrone provenant de l'utilisation des valeurs des flux qui ont été mis à jours aux mauvais moments. Dans nos estimations, cette erreur asynchrone est au mieux d'ordre un. Par conséquent, l'intégration asynchrone d'un schéma numérique classiquement d'ordre  $\alpha \geq 1$  est au plus d'ordre un. Notons enfin que pour un maillage constant, on a  $\delta_{asyn} = 0$  et donc l'erreur asynchrone est elle aussi égale à zero ce qui permet de retrouver l'ordre initial de la méthode. Ceci est une conséquence de la formulation asynchrone choisie qui est équivalente au schéma synchrone dans le cas où tous les pas de temps des cellules sont égaux.*

## 4.4. Tests numériques

Nous avons montré que le schéma asynchrone est théoriquement convergent d'ordre au plus un. Le but de cette section est d'illustrer numériquement les résultats démontrés dans les sections précédentes. Nous utilisons également des simulations numériques pour mettre en évidence certaines propriétés qualitatives des schémas asynchrones, en particulier dans le cas de solutions non régulières pour lesquelles l'ordre d'approximation des méthodes synchrones classique n'est pas forcément atteint.

### 4.4.1. Convergence numérique : perte d'ordre

Considérons de nouveau l'équation de transport en 1D (4.1) avec une vitesse  $v = 1$ . Le maillage choisi est variable et suit une loi polynomiale (2.14) telle que celle définie dans le Chapitre 2. Les tests de convergence ont été réalisés en faisant des séries de calculs numériques en augmentant le nombre total de cellules du maillage, puis la norme de l'erreur est tracée en fonction de nombre de cellules, en échelle logarithmique.

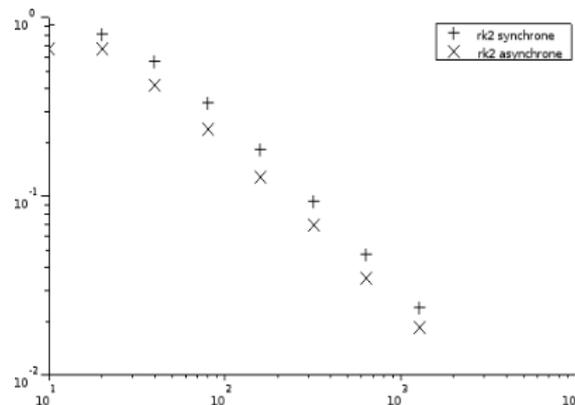


FIGURE 4.2 – Schéma RK2 + upwind en espace : (+) schéma synchrone et (×) schéma asynchrone.

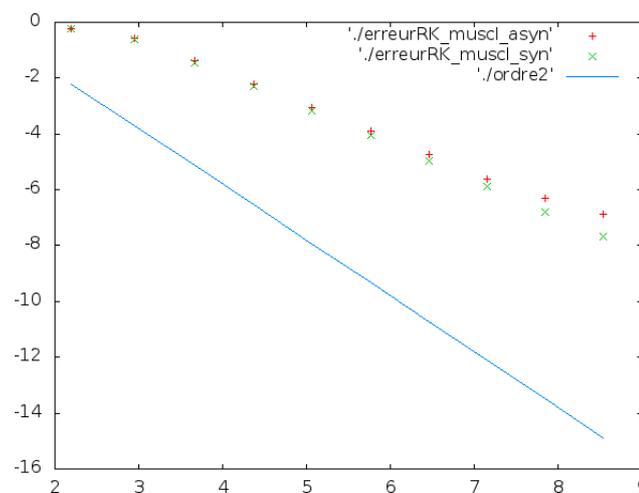


FIGURE 4.3 – Schéma RK2 + MUSCL en espace. Ordre 2 est une droite de pente 2.

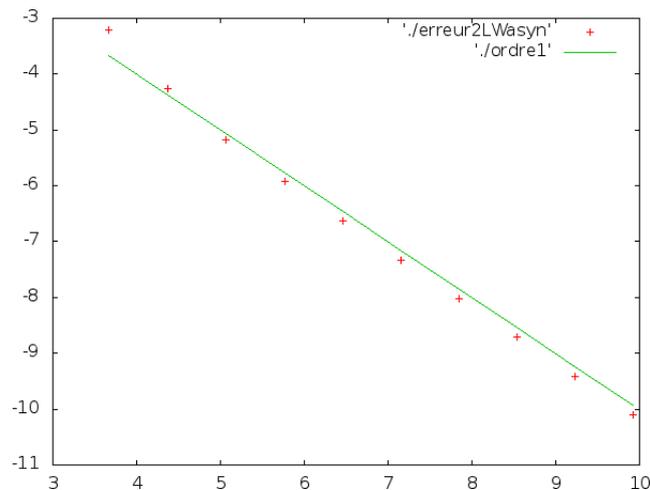


FIGURE 4.4 – Convergence du schéma Lax Wendroff asynchrone. Ordre 1 est l’ordre 1 théorique.

On observe bien ce phénomène de perte d’ordre avec une convergence à l’ordre un, comme annoncé dans les résultats théoriques.

#### 4.4.2. Gain en temps de calcul pour une erreur fixée

L’objectif lors de la mise en place d’un schéma asynchrone est le gain en temps de calcul. Nous avons démontré que pour des schémas d’ordre plus élevé, le schéma asynchrone converge moins vite que le schéma synchrone classique (Fig. 4.3). Ceci est dû à la perte d’ordre induite par le schéma asynchrone. Pour obtenir une erreur fixée, il sera donc nécessaire d’utiliser plus de cellules pour un schéma asynchrone que pour un schéma synchrone. Cependant, pour une erreur fixée, l’efficacité de l’algorithme asynchrone est telle que le gain en temps de calcul par rapport au schéma classique synchrone est encore réel.

En utilisant le même maillage que dans la section précédente, la comparaison du nombre de mailles et du temps de calcul nécessaires pour obtenir une erreur fixée est résumée dans le tableau de la figure 4.5)

N	Erreurs		Temps de calcul	
	asynchrone	synchrone	asynchrone	synchrone
1280	3,6 10 <sup>-3</sup>	2,8 10 <sup>-3</sup>	1,3 mn	13,2 mn
2560	1,8 10 <sup>-3</sup>	1,13 10 <sup>-3</sup>	5,7 mn	77,5 mn
5120	1,04 10 <sup>-3</sup>	4,5 10 <sup>-4</sup>	26,4 mn	5,19 h
10240	5,5 10 <sup>-4</sup>	1,8 10 <sup>-4</sup>	2,08 h	20,73 h

FIGURE 4.5 – Schéma RK2+ MUSCL synchrone et asynchrone.

### 4.4.3. L'asynchrone et les solutions irrégulières

Les théorèmes de convergence, que ce soit dans le cas synchrone ou dans le cas asynchrone, imposent des conditions de régularité sur la solution du problème traité. C'est une conséquence du fait que le calcul de l'erreur de consistance fait appel à un développement de Taylor jusqu'à un ordre suffisant. Dans ce cas régulier, nous avons montré une perte d'ordre par rapport au cas classique synchrone. Nous allons donc sortir du cadre régulier et nous appliquerons la méthode asynchrone à des problèmes pour lesquels la condition initiale est irrégulière. Nous présenterons ici deux exemples : celui de la propagation d'un créneau en dimension 1 d'espace et un exemple bidimensionnel de transport dans un champ de vitesse à divergence nulle (appelé cas-test de l'horloge).

#### Exemple 1D : le créneau

Reprenons l'équation (4.1) avec une vitesse constante  $v = 1$ , des conditions limites périodiques et la condition initiale discontinue :

$$\begin{cases} u(x, 0) = 1, & \text{si } 1/4 < x < 1/2, \\ u(x, 0) = 0, & \text{si } x \in [0, 1/4] \cup [1/2, 1]. \end{cases}$$

Le maillage est toujours variable, les tailles des mailles étant encore définies selon la loi polynomiale (2.14).

1. Nous avons appliqué dans un premier temps le schéma Lax-Wendroff asynchrone. Nous nous intéressons à la qualité des solutions synchrone et asynchrone au temps  $T = 10$ s (10 tours du domaine périodique).

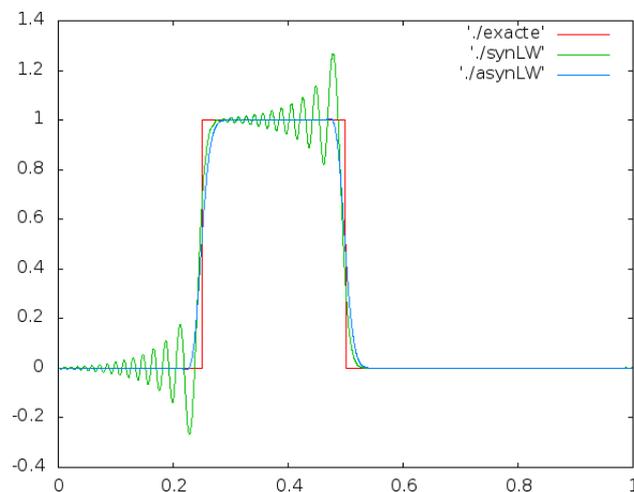


FIGURE 4.6 – Schéma Lax Wendroff asynchrone,  $T = 10$  s

La figure 4.6 présente une comparaison des solutions des schémas synchrone (en vert) et asynchrone (en bleu) avec la solution exacte (en rouge) après 10 périodes. La propriété la plus flagrante dans cet exemple est que le schéma Lax-Wendroff asynchrone réduit fortement les oscillations observées avec le schéma Lax-Wendroff synchrone. Ceci s'explique de nouveau par le fait que dans

le cas classique où le pas de temps est limité par la taille de la plus petite cellule, les "grosses" cellules sont réévaluées beaucoup plus de fois de ce qui est imposé par leur propre CFL. Le schéma Lax-Wendroff n'est pas à variation totale décroissante, ce qui conduit à l'accumulation des oscillations à chaque petit pas de temps. Le schéma asynchrone qui limite ces réévaluations permet donc de réduire les oscillations observées.

2. Dans un deuxième temps, le schéma A1RK2 avec le schéma MUSCL a été appliqué. Nous nous intéressons toujours à la qualité des solutions synchrone et asynchrone au temps  $T = 10s$  (10 tours du domaine périodique). Le résultat est représenté sur la figure 4.7

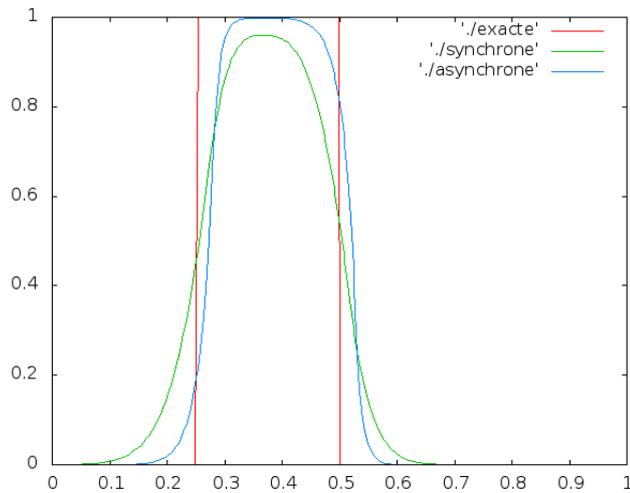


FIGURE 4.7 – Schéma A1RK2 avec MUSCL,  $T = 10s$

La figure 4.7 présente une comparaison des solutions de schémas synchrone (en vert) et asynchrone (en bleu) avec la solution exacte (en rouge) après 10 périodes, pour un maillage avec 300 éléments. On observe de nouveau que le schéma asynchrone réduit la diffusion numérique en comparaison avec le schéma classique, ce qui s'explique de nouveau par l'utilisation d'un pas de temps le plus proche possible de la condition CFL.

### Exemple 2D : l'horloge

Reprenons l'équation de transport (3.1), cette fois en dimension 2 d'espace. Le champ de vitesse est donné par la formule :

$$\begin{cases} v_x(x, y) = 2\pi \frac{1/2-y}{\sqrt{(x-1/2)^2+(y-1/2)^2}} \\ v_y(x, y) = 2\pi \frac{x-1/2}{\sqrt{(x-1/2)^2+(y-1/2)^2}} \end{cases}$$

et est représenté sur la figure 4.8 (à gauche). Ce champ de vitesse est à divergence nulle, ce qui permet de préserver les valeurs de la condition initiale le long des caractéristiques qui forment des cercles

concentriques, centrés au point milieu du domaine de calcul. La condition initiale est donnée par :

$$\begin{cases} u(x, 0) = 10^9, & \text{si } |x - 1/2| < 0.05 \text{ et } |y - 1/4| < 0.2 \\ u(x, 0) = 1, & \text{sinon} \end{cases}$$

et est représentée sur la figure 4.8 (à droite).

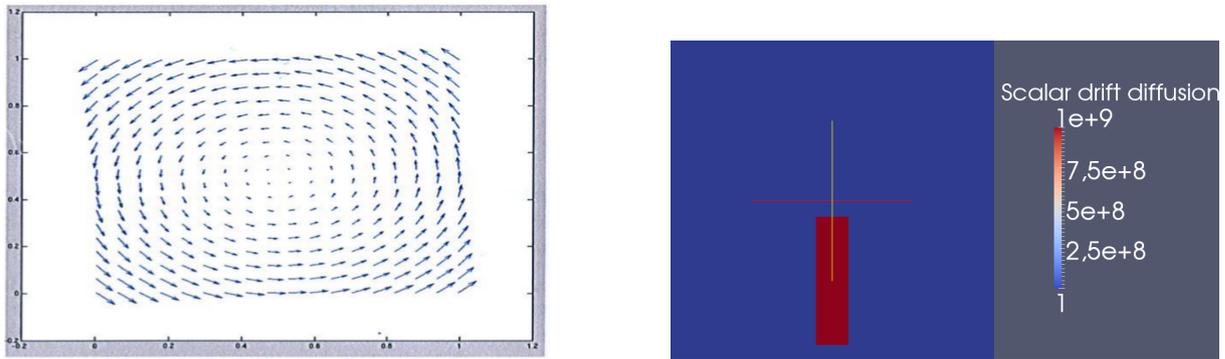


FIGURE 4.8 – Le cas test de l’horloge : le champ de vitesse (à gauche) et la condition initiale (à droite).

1. Le premier test consiste à comparer la solution asynchrone du schéma RK2 et MUSCL qui est un schéma classiquement d’ordre deux en synchrone, avec la solution asynchrone du schéma Euler et upwind qui est classiquement d’ordre un. En fait, nous avons prouvé dans la section 4.2.4. que le schéma A1RK2 avec une discrétisation MUSCL est d’ordre un et dans le chapitre 2, nous avons prouvé que le schéma Euler asynchrone et upwind est également d’ordre un. Théoriquement, les deux schémas ont le même ordre de convergence dans le cas des solutions régulières. Le but de ce premier test est donc de comparer les deux schémas dans le cas d’une condition initiale discontinue, un cas où les résultats théoriques ne sont plus valides. Nous profiterons de ce test pour comparer également les solutions synchrone et asynchrone dans les deux cas.

Le maillage est tensorisé à partir de deux maillages 1D qui suivent une loi polynomiale. Ce maillage est représenté sur la figure 4.9. L’aspect asynchrone provient de la variation du maillage et de la variation du champ de vitesse. Pour ces tests, nous avons pris le paramètre de maillage à deux valeurs  $\alpha = 0.01$  et  $\alpha = 0.5$  (voir chapitre 2).

Les solutions synchrones et asynchrones du schéma Euler avec upwind sont représentées dans Fig. 4.10 et du schéma RK2 avec MUSCL dans Fig. 4.11

La première remarque intéressante est que le schéma asynchrone RK2 en temps et MUSCL en espace ( Fig. 4.11 à droite) réduit la diffusion numérique et améliore la qualité de la solution en comparaison avec le schéma asynchrone Euler et upwind ( Fig. 4.10 à droite) malgré la perte d’ordre prouvée pour les solutions régulières. Notons aussi que pour les deux schémas, l’asynchrone réduit la diffusion numérique en comparaison avec le schéma classique. De plus, l’asynchrone reste très efficace en termes de temps de calcul. En effet, pour une intégration RK2 et MUSCL de l’horloge dans un maillage polynomial à  $300 \times 300$  mailles, le temps de calcul dans le cas synchrone est 53h et pour l’asynchrone 4h.



FIGURE 4.9 – Maillage polynomial en 2D avec 10 mailles selon  $x$  et 10 mailles selon  $y$ . Au milieu des deux axes, la pente  $\alpha = 0.5$  (à gauche) et  $\alpha = 0.01$  (à droite).

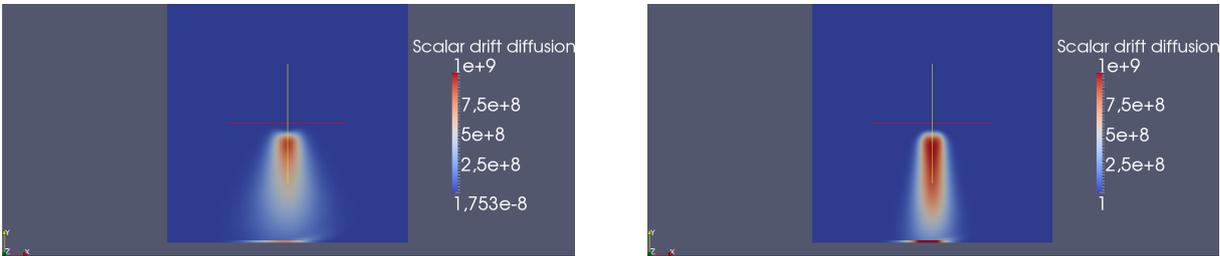


FIGURE 4.10 – Schéma euler en temps et upwind en espace  $300 \times 300$  mailles,  $T = 1$  : synchrone (gauche), asynchrone (droite).

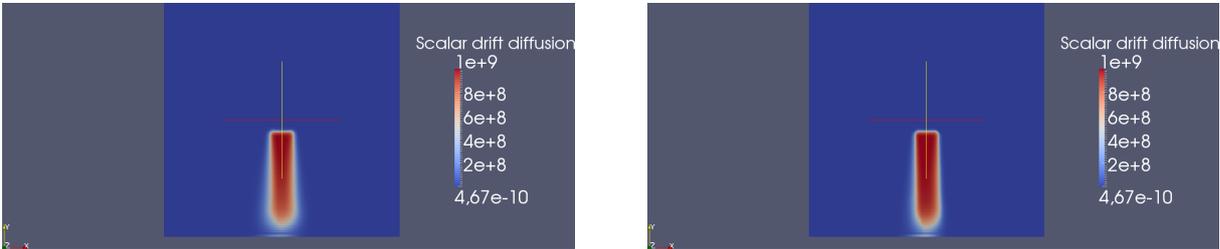


FIGURE 4.11 – Schéma RK2 en temps et MUSCL en espace  $300 \times 300$  mailles,  $T = 1$  : synchrone (gauche), asynchrone (droite).

2. Le deuxième test avec l'horloge est de comparer les solutions en fonction du maillage. Nous appliquons donc les deux algorithmes synchrone et asynchrone du schéma RK2 avec MUSCL d'abord avec un maillage constant (Fig. 4.12), ensuite avec un maillage qui suit une loi polynomiale selon  $x$  et selon  $y$  avec une pente  $\alpha = 0.5$  (Fig. 4.13) et enfin un maillage polynomial avec  $\alpha = 0.01$  (déjà présenté dans Fig. 4.11). Pour les trois tests qui suivent, le nombre de mailles est fixé à  $N = 300 \times 300$  et le temps final à  $T = 1$ s (un tour d'horloge).

- (a) Maillage constant : les pas d'espace selon les axes  $x$  et  $y$  sont constants et égaux à  $\frac{1}{300}$ . En fait, le schéma asynchrone a été dérivé du schéma classique et a été donc construit en tenant compte du fait que l'asynchrone dégénère vers le schéma synchrone dans le cas d'un maillage

constant où le pas de temps est homogène. Les résultats sont représentés sur la figure 4.12.

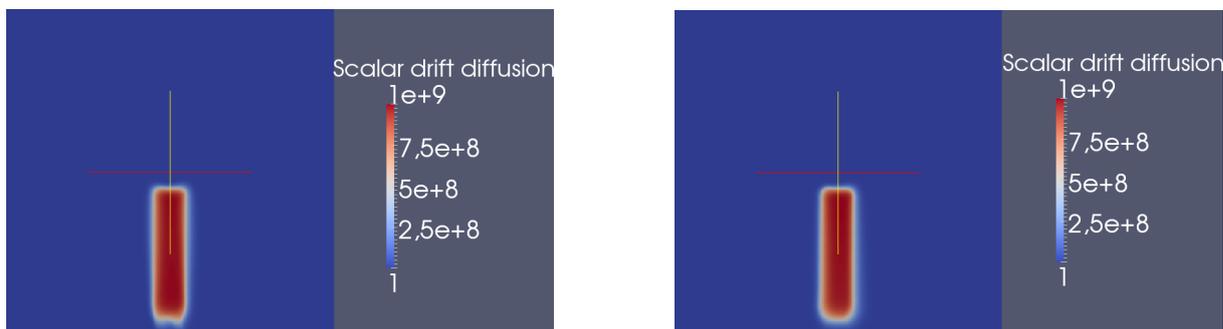


FIGURE 4.12 – Schéma RK2-MUSCL : synchrone (gauche), asynchrone (droite) ; maillage constant.

Dans le cas d'un maillage constant, les résultats synchrones et asynchrones sont comparables que ce soit pour la solution ou le temps de calcul. L'intégration asynchrone n'apporte donc rien puisque la condition CFL est homogène. De plus, en terme de temps de calcul, le classique est plus efficace. En effet, à chaque pas de temps, l'algorithme asynchrone trie toutes les mailles pour chercher la cellule la plus urgente à mettre à jour. Cette tâche est inutile pour un maillage uniforme où toutes les mailles ont le même temps de mise à jour et coûte un temps supplémentaire en comparaison avec l'intégration synchrone. Il est donc inutile d'utiliser une intégration asynchrone pour un maillage uniforme.

- (b) Maillage polynomial avec  $\alpha = 0.5$  (voir aussi sur la figure 4.9, à gauche). Dans ce cas, le maillage varie lentement. Le maillage est variable, les pas de temps sont indépendants mais la différence entre deux pas de temps est petite.

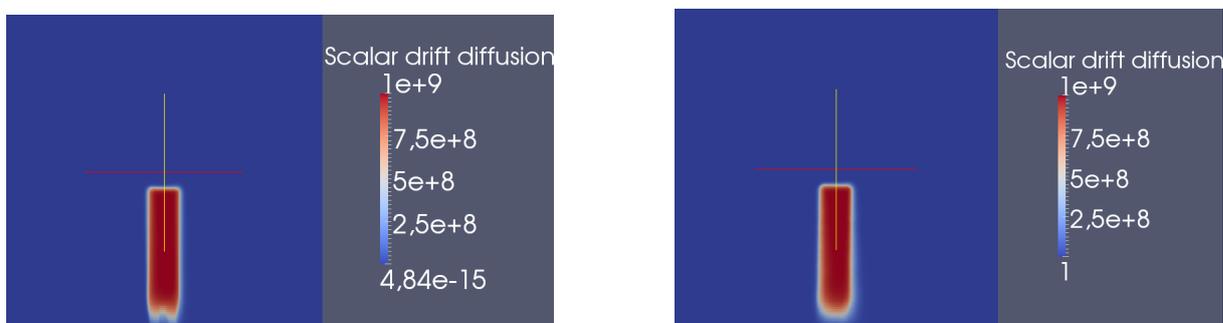


FIGURE 4.13 – RK2-MUSCL : synchrone (gauche), asynchrone (droite) ; maillage polynomial :  $\alpha = 0.5$ .

La Figure 4.13 présente les solutions synchrones et asynchrones pour un maillage où les pas de temps varient lentement. Notons que l'asynchrone diminue légèrement la diffusion numérique en comparaison avec le synchrone. De plus, le temps de calcul diminue aussi légèrement. En effet, le temps de calcul est égal à 45mn pour l'asynchrone et 1h et 8mn pour le synchrone ce qui donne un speed up numérique égal à 1.52.

- (c) Maillage polynomial avec  $\alpha = 0.01$  (voir aussi sur la figure 4.9, à droite). Les résultats sont présentés sur la figure 4.11. Pour ce test, les pas de temps sont indépendants et le maillage est

fortement raffiné au milieu du domaine. Le but est d'étudier le comportement de l'asynchrone en fonction du caractère multi-échelle. L'asynchrone réduit la diffusion numérique et donc améliore la qualité de la solution synchrone. Le speed up numérique dans ce cas est de l'ordre de  $13 \gg 1,5$  ce qui est beaucoup plus important par rapport au deuxième test où le maillage varie lentement. Le caractère multi-échelle favorise donc, comme attendu, le gain en temps de calcul.

## 4.5. Conclusion

Nous avons développé dans ce chapitre des schémas numériques asynchrones dérivés des schémas qui sont classiquement d'ordre deux. Nous avons démontré que le schéma asynchrone conduit à une perte d'ordre en comparaison avec le schéma classique. En effet, l'algorithme asynchrone suppose que les flux sont mis à jour indépendamment les uns des autres. Par conséquent, l'utilisation des valeurs des flux qui ont été mis à jour à des instants différents crée une erreur asynchrone autre que l'erreur du schéma numérique. En examinant cette erreur asynchrone supplémentaire, nous nous sommes rendus compte qu'elle est au mieux d'ordre un et par suite que le schéma asynchrone est aussi au plus d'ordre un. Le schéma classique converge plus vite que l'asynchrone, cependant, pour une erreur fixée nous avons un gain en temps de calcul si nous utilisons une intégration asynchrone.

L'étude de convergence d'un schéma numérique nécessite une certaine régularité de la solution. Nous nous sommes donc sortis de ce cadre et nous avons appliqué la méthode asynchrone dans le cadre des solutions irrégulières. Puis nous avons étudié numériquement le comportement de l'asynchrone. Dans ce cas, le schéma asynchrone présente des propriétés additionnelles telles que la réduction de la diffusion numérique et la réduction des oscillations parasites créées autour des discontinuités dans le cas des schémas qui ne sont pas à variation totale décroissante comme le schéma de Lax-Wendroff. Ceci est expliqué par le fait que les larges zones du domaine de calcul sont réévaluées beaucoup plus souvent que ce qui leur serait imposé par leur propre condition CFL locale ce qui favorise l'accumulation des oscillations. Signalons enfin que dans ce cadre des solutions irrégulières, l'asynchrone continue à être bénéfique en termes de gain en temps de calcul. Cependant, le schéma numérique asynchrone développé est convergent uniquement à l'ordre un ce qui reste insuffisant pour certaines applications. D'où l'idée de développer un véritable schéma numérique asynchrone d'ordre deux et ce sera l'objet du chapitre suivant.



## **Deuxième partie**

# **Schéma Runge-Kutta asynchrone d'ordre deux pour les équations différentielles linéaires**



## Chapitre 5

# Schéma asynchrone du second ordre

La dérivation des méthodes à pas de temps locaux d'ordre plus que un représente un véritable challenge, et seulement quelques méthodes sont disponibles. Parmi les méthodes qui ont une bonne réputation, citons la méthode Runge-Kutta 2 à pas de temps local (LTS-RK2) définie dans [35]. Dans cet article, les auteurs ont réussi à dériver un schéma explicite à pas de temps local du schéma Runge-Kutta classique à un ordre donné. Ils ont prouvé que, dans le cas d'un maillage composé de deux régions : une grossière et une fine où le pas du temps global est forcément un multiple du pas de temps local, leur méthode préserve l'ordre de la méthode classique. Les idées de base de la méthode ont été rappelées dans la sous section 1.4.3.2..

Nous nous sommes inspirés du schéma LTS-RK2 pour dériver un schéma numérique asynchrone du schéma RK2 à pas de temps global qui permet la sélection de pas de temps indépendants mais qui préserve également l'ordre de convergence de la méthode classique. Pour illustrer la méthode asynchrone, nous considérerons une équation différentielle linéaire générique avec une discrétisation arbitraire en espace sur un maillage aléatoire. Tout d'abord, l'algorithme asynchrone sera présenté dans la première section. Puis, la deuxième section sera consacrée à l'étude théorique de la convergence du schéma asynchrone et à la comparaison du schéma asynchrone avec le schéma LTS-RK2. Ensuite, des tests numériques pour la validation des résultats théoriques seront présentés dans la troisième section. Enfin, quelques remarques et la conclusion seront présentées dans la quatrième section.

### 5.1. L'algorithme asynchrone

Nous présentons dans cette section le schéma Runge-Kutta 2 asynchrone (ARK2) dérivé du schéma RK2 classique. Nous reprenons les notations du maillage définies dans le chapitre 3 et n'imposons aucune condition sur le maillage. Nous nous intéressons à la résolution numérique d'une équation différentielle linéaire donnée avec une discrétisation arbitraire en espace. Le système semi-discrétisé s'écrit alors sous la forme suivante :

$$\frac{dy_i}{dt}(t) = B_i y(t), \quad \forall K_i \in \mathcal{T}, \quad (5.1)$$

où  $y_i$  et  $y$  sont respectivement les vecteurs qui contiennent les degrés de liberté (dofs) de l'élément  $K_i$  et de tous les éléments, et  $B_i$  une matrice provenant de la discrétisation en espace.

Pour chaque cellule de maillage, le terme source dépend seulement de la cellule courante alors que les

flux dépendent aussi des voisins. Donc pour appliquer la méthode ARK2, nous devons d'abord diviser le terme  $B_i y$  en trois contributions :  $V_i$  est une contribution de "volume" locale,  $F_{i,j}^+$  est une contribution de "flux" locale et  $F_{i,j}^-$  est une contribution de "flux" qui vient des voisins. Donc l'équation (5.1) devient

$$\frac{dy_i}{dt}(t) = V_i y_i(t) + \sum_{j \in \mathcal{N}(i)} F_{i,j}^+ y_i(t) + F_{i,j}^- y_j(t), \quad \forall K_i \in \mathcal{T}. \quad (5.2)$$

Notons par  $y_i^m$  l'approximation de  $y_i(m\Delta t_i)$  et par  $\left(\frac{\Delta y}{\Delta t}\right)_i^m$  l'approximation de  $\frac{dy_i}{dt}(m\Delta t_i)$ .

### 5.1.1. L'algorithme ARK2

Les temps de mise à jour correspondent à des multiples de  $\Delta t_i$  pour chaque cellule  $K_i$ , et  $t_{\text{cur}}$  est le temps courant de la simulation.

À l'instant  $t_{\text{cur}}$  tel que  $t_{\text{cur}}/\Delta t_i$  est un entier  $p$ , il faut réaliser les opérations suivantes :

1. Calculer  $\tilde{y}_i$  à l'instant  $t_{\text{cur}}$

$$\tilde{y}_i(t_{\text{cur}}) = \tilde{y}_i(t_{\text{pre},i}) + (t_{\text{cur}} - t_{\text{pre},i}) \left(\frac{\Delta \tilde{y}}{\Delta t}\right)_i(t_{\text{mid},i}), \quad t_{\text{mid},i} = t_{\text{cur}} - \left(\frac{t_{\text{cur}} - t_{\text{pre},i}}{2}\right)$$

avec  $t_{\text{pre},i}$  est le temps de la dernière mise à jour de l'élément  $i$ . La pente au point milieu est définie par :

$$\begin{aligned} \left(\frac{\Delta \tilde{y}}{\Delta t}\right)_i(t_{\text{mid},i}) &= V_i \left( y_i^{p-1} + \frac{\Delta t_i}{2} \left(\frac{\Delta y}{\Delta t}\right)_i^{p-1} \right) \\ &+ \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ \left( y_i^{p-1} + (t_{\text{mid},i,j} - (p-1)\Delta t_i) \left(\frac{\Delta y}{\Delta t}\right)_i^{p-1} \right) \\ &+ \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left( y_j^{n_j} + (t_{\text{mid},i,j} - n_j \Delta t_j) \left(\frac{\Delta y}{\Delta t}\right)_j^{n_j} \right) \end{aligned}$$

où  $n_j = \left\lfloor \frac{t_{\text{pre},i}}{\Delta t_j} \right\rfloor$ ,  $t_{\text{mid},i,j} = t_{\text{next},i,j} - \left(\frac{t_{\text{next},i,j} - t_{\text{pre},i,j}}{2}\right)$  avec

$t_{\text{pre},i,j} = \max(n_j \Delta t_j, (p-1)\Delta t_i)$  et  $t_{\text{next},i,j} = \min((n_j + 1)\Delta t_j, p\Delta t_i)$ .

2. Pour tous les voisins de l'élément  $i$ , i.e.  $j \in N(i)$ , mettre à jour les valeurs  $\tilde{y}_j$  à l'instant  $t_{\text{cur}}$

$$\tilde{y}_j(t_{\text{cur}}) = \tilde{y}_j(t_{\text{pre},j}) + (t_{\text{cur}} - t_{\text{pre},j}) \left(\frac{\Delta \tilde{y}}{\Delta t}\right)_j(t_{\text{mid},j}),$$

$t_{\text{pre},j}$  est le temps de la dernière mise à jour de l'élément  $j$ , et  $t_{\text{mid},j}$  est le temps "médian".

3. Mettre à jour la valeur de  $y_i$  et faire avancer les temps de la dernière mise à jour

$$y_i^p = \tilde{y}_i(t_{\text{cur}}), \quad t_{\text{pre},i} = t_{\text{cur}} \text{ and } t_{\text{pre},j} = t_{\text{cur}}, \quad \forall j \in N(i).$$

4. Mettre à jour la pente qui correspond à l'élément  $i$

$$\left(\frac{\Delta y}{\Delta t}\right)_i^p = V_i y_i^p + \sum_{j \in \mathcal{N}(i)} F_{i,j}^+ y_i^p + F_{i,j}^- \left( y_j^{n_j} + (p\Delta t_i - n_j \Delta t_j) \left(\frac{\Delta y}{\Delta t}\right)_j^{n_j} \right),$$

$$\text{où } n_j = \left\lfloor \frac{t_{\text{cur}}}{\Delta t_j} \right\rfloor.$$

**Remarque 3.** Le temps final de la simulation doit correspondre à un temps de rencontre commun pour tous les éléments du maillage : un multiple du plus petit commun multiple de tous les pas de temps. Si ce n'est pas le cas, on utilise une interpolation d'ordre au moins deux pour calculer la solution synchronisée.

## 5.2. Étude de la convergence du schéma ARK2

Le schéma ARK2 est dérivé du schéma RK2 classique. Le but est donc de pouvoir justifier que le schéma asynchrone préserve l'ordre de convergence du schéma RK2. Ce qui revient à prouver que la méthode ARK2 est convergente d'ordre deux.

### 5.2.1. Notations

1. Commençons par quelques notations concernant l'évolution du système entre deux temps successifs de mise à jour d'un élément  $K_i$  qui sont multiples de  $\Delta t_i$  (voir aussi sur la figure 5.1).
  - (a) Soit  $t_i^{r_i+1} := (r_i + 1)\Delta t_i = t_{\text{cur}} \leq T$  où  $T$  est le temps final de la simulation. Supposons que l'élément  $K_i$  a été mis à jour  $m$  fois entre  $t_i^{r_i}$  et  $t_i^{r_i+1}$  aux instants :  $t_{\text{pre},i}^1, t_{\text{pre},i}^2, \dots, t_{\text{pre},i}^m$ ,  $0 \leq m \leq \left\lfloor \frac{\Delta t_i}{\min_j(\Delta t_j)} \right\rfloor + 1$ .
  - (b) Soient  $\delta_i^k = t_{\text{pre},i}^k - t_{\text{pre},i}^{k-1}$  pour tout  $2 \leq k \leq m$ ,  $\delta_i^1 = t_{\text{pre},i}^1 - t_i^{r_i}$ ,  $\delta_i^{m+1} = t_i^{r_i+1} - t_{\text{pre},i}^m$  et  $\sum_{k=1}^{m+1} \delta_i^k = \Delta t_i$ . Alors,  $t_{\text{mid},i}^k = t_{\text{pre},i}^k - \frac{\delta_i^k}{2}$  pour  $1 \leq k \leq m$  et  $t_{\text{mid},i}^{m+1} = t_i^{r_i+1} - \frac{\delta_i^{m+1}}{2}$ .
  - (c) Supposons que pour tout  $j \in \mathcal{N}(i)$  nous avons les hypothèses suivantes :
    - i. entre les deux instants  $t_i^{r_i}$  et  $t_i^{r_i+1}$ , l'élément  $j$  a été mis à jour  $z_j^{r_i}$  fois à des instants qui sont multiples de  $\Delta t_j$ . Par suite  $0 \leq z_j^{r_i} \leq m$ . Par exemple, dans Fig. 5.1,  $z_s^{r_i} = 0$  et  $z_j^{r_i} = 5$  où  $s$  et  $j$  sont deux voisins de l'élément  $i$ .
    - ii.  $r_j \Delta t_j \leq r_i \Delta t_i$  et à tout instant  $t_{\text{pre},i}^k$ ,  $r_j^k \Delta t_j$  est le temps de la dernière mise à jour de l'élément  $j$  et qui est multiple de  $\Delta t_j$  juste avant  $t_{\text{pre},i}^k$ . Notons qu'il peut y avoir des situations où  $r_j^{k+1} \Delta t_j = r_j^k \Delta t_j$ . Dans Fig. 5.1 par exemple, pour tout  $1 \leq k \leq m$ ,  $r_s^k \Delta t_s = r_s \Delta t_s$  et pour le voisin  $j$ ,  $r_j^1 \Delta t_j = r_j \Delta t_j$  et  $r_j^3 \Delta t_j = r_j^2 \Delta t_j = (r_j + 1) \Delta t_j$ .
    - iii.  $t_{\text{mid},i,j}^k = t_{\text{next},i,j}^k - \left( \frac{t_{\text{next},i,j}^k - t_{\text{pre},i,j}^k}{2} \right) = \frac{t_{\text{next},i,j}^k + t_{\text{pre},i,j}^k}{2}$ , avec
      - A.  $t_{\text{pre},i,j}^k = \max(r_j^k \Delta t_j, r_i \Delta t_i)$
      - B.  $t_{\text{next},i,j}^k = \min((r_j^k + 1) \Delta t_j, (r_i + 1) \Delta t_i)$

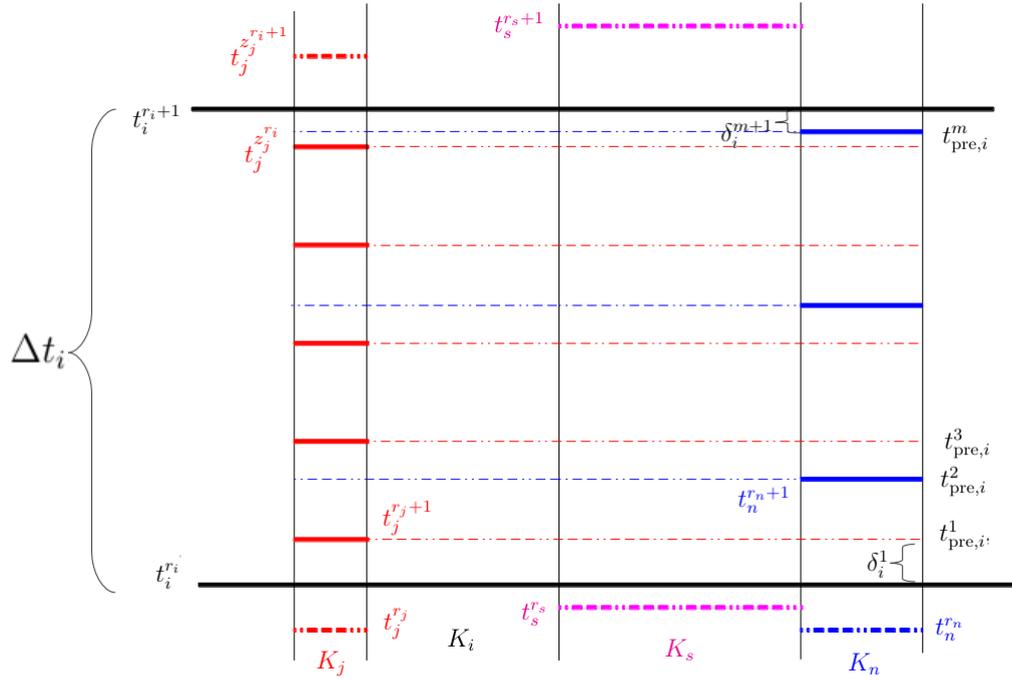


FIGURE 5.1 – L'évolution temporelle du système entre deux instants :  $t_i^{r_i}$  et  $t_i^{r_i+1}$  où les éléments  $j$ ,  $s$  et  $n$  sont les voisins de  $K_i$ .

2. La mise à jour d'un élément  $K_i$  fait intervenir les voisins. Les voisins de  $K_i$  dépendent eux aussi de leurs voisins, et les voisins des voisins de  $K_i$  font appel à leurs voisins, et ainsi de suite. Soient donc les notations suivantes ; voir aussi Fig. 5.2 :

- (a) Les éléments  $K_j$ ,  $j \in \mathcal{N}(i)$ , sont les voisins de  $K_i$ .
- (b) Les éléments  $K_{\{i,1\}}$ ,  $\{i,1\} \in \mathcal{N}(j)$ , sont les voisins des  $K_j$  donc les voisins des voisins de  $K_i$ .
- (c) Les éléments  $K_{\{i,2\}}$ ,  $\{i,2\} \in \mathcal{N}(\{i,1\})$ , sont les voisins des voisins des voisins de  $K_i$ .
- (d) Les éléments  $K_{\{i,n\}}$ ,  $\{i,n\} \in \mathcal{N}(\{i,n-1\})$ , sont les voisins de  $K_{\{i,n-1\}}$

Soit  $h_{i,t_s}^{r_i}$  le nombre des dépendances présentées par l'élément  $K_i$  entre l'instant  $t_i^{r_i}$  et l'instant  $t_s$  où  $t_s$  est le temps de la dernière synchronisation de la solution et qui peut notamment être l'instant initial  $t^0$ . Autrement dit, les éléments  $K_{\{i,h_{i,t_s}^{r_i}\}}$  utilisent leurs valeurs calculées au temps de la dernière synchronisation de la solution. Notons que l'étude de la dépendance sera utile pour la suite. En fait, l'étude théorique du schéma ARK2, nécessite la maîtrise de l'évolution du système. En revanche, pour l'algorithme la mise à jour d'une cellule donnée ne fait intervenir que les valeurs de ses voisins, et n'a donc pas besoin des voisins des voisins. En fait cette indépendance est implicite et cachée derrière les étapes intermédiaires de l'algorithme.

3. Nous introduisons quelques éléments pour l'identification :

$$(Bx)_i = V_i x_i + \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ x_j + \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- x_j$$

$$(BP_i x)_i = V_i x_i + \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ x_j$$

où  $P_i$  est une matrice de projection diagonale qui identifie les degrés de liberté de l'élément  $i$ .

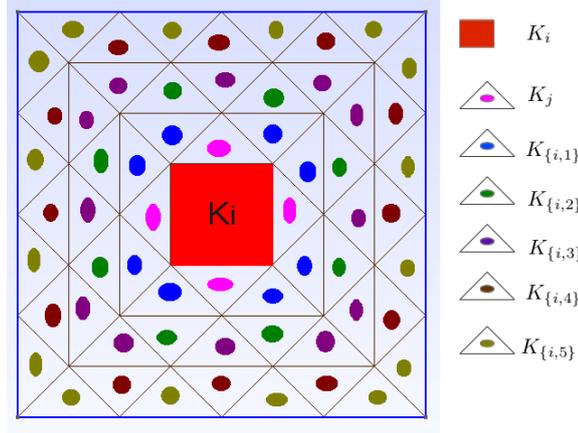


FIGURE 5.2 – Dépendance de l'élément  $K_i$  par rapport aux autres éléments de maillage.

**Proposition 14.** L'évolution temporelle d'un élément donné  $i$  entre deux instants  $t_i^{r_i} = r_i \Delta t_i$  et  $t_i^{r_i+1} = t_i^{r_i} + \Delta t_i$  est donnée par l'expression suivante

$$y_i^{r_i+1} = y_i^{r_i} + \Delta t_i \left[ V_i \left( y_i^{r_i} + \frac{\Delta t_i}{2} \left( \frac{\Delta y}{\Delta t} \right)_i^{r_i} \right) \right]$$

$$+ \sum_{k=1}^{m+1} \delta_i^k \left[ \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ \left( y_i^{r_i} + (t_{\text{mid},i,j}^k - r_i \Delta t_i) \left( \frac{\Delta y}{\Delta t} \right)_i^{r_i} \right) \right]$$

$$+ \sum_{k=1}^{m+1} \delta_i^k \left[ \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left( y_j^{r_j} + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \left( \frac{\Delta y}{\Delta t} \right)_j^{r_j^k} \right) \right].$$

*Démonstration.* Par définition, nous avons l'expression suivante

$$y_i^{r_i+1} = \tilde{y}_i^{t_{\text{pre},i}^m} + \delta_i^{m+1} \left( \frac{\Delta \tilde{y}}{\Delta t} \right)_i (t_{\text{mid},i}^{m+1})$$

$$= \tilde{y}_i^{t_{\text{pre},i}^{m-1}} + \delta_i^m \left( \frac{\Delta \tilde{y}}{\Delta t} \right)_i (t_{\text{mid},i}^m) + \delta_i^{m+1} \left( \frac{\Delta \tilde{y}}{\Delta t} \right)_i (t_{\text{mid},i}^{m+1})$$

$$\vdots$$

$$= y_i^{r_i} + \sum_{k=1}^{m+1} \delta_i^k \left( \frac{\Delta \tilde{y}}{\Delta t} \right)_i (t_{\text{mid},i}^k)$$

avec pour tout  $1 \leq k \leq m + 1$ , nous avons

$$\begin{aligned} \left(\frac{\Delta \tilde{y}}{\Delta t}\right)_i(t_{\text{mid},i}^k) &= V_i\left(y_i^{r_i} + \frac{\Delta t_i}{2} \left(\frac{\Delta y}{\Delta t}\right)_i^{r_i}\right) + \sum_{j \in N^+(i)} F_{i,j}^+\left(y_i^{r_i} + (t_{\text{mid},i,j}^k - r_i \Delta t_i) \left(\frac{\Delta y}{\Delta t}\right)_i^{r_i}\right) \\ &+ \sum_{j \in N^-(i)} F_{i,j}^-\left(y_j^{r_j} + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \left(\frac{\Delta y}{\Delta t}\right)_j^{r_j^k}\right) \end{aligned}$$

Pour conclure, il suffit d'utiliser le fait que  $\sum_{k=1}^{m+1} \delta_i^k = \Delta t_i$ .  $\square$

**Proposition 15.** *Pour tout élément  $i$  donné, et entre deux instants  $t_i^{r_i}$  et  $t_i^{r_i+1}$  nous avons le résultat suivant*

$$\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k = \frac{(2r_i + 1)(\Delta t_i)^2}{2}, \forall j \in \mathcal{N}(i). \quad (5.3)$$

En particulier, entre l'instant  $t_s$ , un temps de synchronisation de la solution, et l'instant  $t_s + \Delta t_i$  nous avons pour tout  $j \in \mathcal{N}(i)$

$$\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k = \frac{(\Delta t_i)^2}{2}.$$

*Démonstration.* Nous commençons par prouver que, pour tout élément  $j \in N(i)$ , entre deux temps successifs de mise à jour de l'élément  $K_j$ , qui sont multiples de  $\Delta t_j$  et qui sont compris entre  $t_i^{r_i}$  et  $t_i^{r_i+1}$ ,  $t_{\text{mid},i,j}^k$  est constant. Autrement dit, si  $r_i \Delta t_i \leq (r_j + s) \Delta t_j < (r_j + s + 1) \Delta t_j \leq (r_i + 1) \Delta t_i$ , alors pour tout  $k$  tel que  $(r_j + s) \Delta t_j < t_{\text{pre},i}^k \leq (r_j + s + 1) \Delta t_j$  nous avons le résultat suivant

$$t_{\text{pre},i,j}^k = \max(r_j^k \Delta t_j, r_i \Delta t_i) = \max((r_j + s) \Delta t_j, r_i \Delta t_i) = (r_j + s) \Delta t_j$$

et

$$t_{\text{next},i,j}^k = \min((r_j^k + 1) \Delta t_j, (r_i + 1) \Delta t_i) = \min((r_j + s + 1) \Delta t_j, (r_i + 1) \Delta t_i) = (r_j + s + 1) \Delta t_j.$$

Par conséquent,

$$t_{\text{mid},i,j}^k = \frac{(2(r_j + s) + 1) \Delta t_j}{2}. \quad (5.4)$$

Maintenant démontrons la formule (5.3). Supposons donc que pour tout  $j \in N(i)$ ,  $z_j^{r_i}$  est le nombre de mises à jour de l'élément  $K_j$  à des instants qui sont des multiples de  $\Delta t_j$  entre  $r_i \Delta t_i$  et  $(r_i + 1) \Delta t_i$ . Trois cas se présentent :

1. Si  $z_j^{r_i} = 0$  alors  $r_j \Delta t_j \leq r_i \Delta t_i \leq (r_i + 1) \Delta t_i \leq (r_j + 1) \Delta t_j$  et pour tout  $1 \leq k \leq m + 1$ ,

$$t_{\text{mid},i,j}^k = \frac{(r_i + 1) \Delta t_i + r_i \Delta t_i}{2} = \frac{(2r_i + 1) \Delta t_i}{2},$$

par suite,

$$\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k = \frac{(2r_i + 1)(\Delta t_i)^2}{2}.$$

2. Si  $z_j^{r_i} = 1$  alors nous avons

$$\begin{aligned}
\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k &= \frac{((r_j + 1)\Delta t_j + r_i \Delta t_i)}{2} \times ((r_j + 1)\Delta t_j - r_i \Delta t_i) \\
&+ \frac{((r_i + 1)\Delta t_i + (r_j + 1)\Delta t_j)}{2} \times ((r_i + 1)\Delta t_i - (r_j + 1)\Delta t_j) \\
&= \frac{1}{2} \left[ (r_i + 1)^2 \Delta t_i^2 - r_i^2 \Delta t_i^2 \right] \\
&= \frac{(2r_i + 1)(\Delta t_i)^2}{2}
\end{aligned}$$

3. Si  $z_j^{r_i} \geq 2$  alors  $r_j \Delta t_j \leq r_i \Delta t_i < (r_j + 1)\Delta t_j < \dots < (r_j + z_j^{r_i})\Delta t_j \leq (r_i + 1)\Delta t_i < (r_j + z_j^{r_i} + 1)\Delta t_j$ . D'après (5.4) nous avons le résultat suivant

$$\begin{aligned}
\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k &= \frac{((r_j + 1)\Delta t_j + r_i \Delta t_i)}{2} \times ((r_j + 1)\Delta t_j - r_i \Delta t_i) \\
&+ \sum_{s=1}^{z_j^{r_i}-1} \frac{(2(r_j + s) + 1)\Delta t_j}{2} \times \Delta t_j \\
&+ \frac{((r_i + 1)\Delta t_i + (r_j + z_j^{r_i})\Delta t_j)}{2} \times ((r_i + 1)\Delta t_i - (r_j + z_j^{r_i})\Delta t_j) \\
&= \frac{((r_j + 1)\Delta t_j)^2}{2} - \frac{(r_i \Delta t_i)^2}{2} + \frac{((r_i + 1)\Delta t_i)^2}{2} - \frac{((r_j + z_j^{r_i})\Delta t_j)^2}{2} \\
&+ \sum_{s=1}^{z_j^{r_i}-1} (2(r_j + s) + 1) \times \frac{(\Delta t_j)^2}{2}.
\end{aligned}$$

Notons que  $\sum_{s=1}^{z_j^{r_i}-1} (2(r_j + s) + 1)$  est une somme d'une suite arithmétique de raison deux. Par conséquent,

$$\begin{aligned}
\sum_{s=1}^{z_j^{r_i}-1} (2(r_j + s) + 1) &= \left[ \sum_{s=0}^{z_j^{r_i}-1} (2(r_j + s) + 1) \right] - (2r_j + 1) \\
&= z_j^{r_i} \frac{(4r_j + 2z_j^{r_i})}{2} - (2r_j + 1) \\
&= (z_j^{r_i})^2 + 2r_j(z_j^{r_i} - 1) - 1
\end{aligned}$$

par suite

$$\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k = \frac{(2r_i + 1)(\Delta t_i)^2}{2}.$$

La preuve est donc achevée. □

**Proposition 16.** Pour tout élément  $i$  et à tout instant  $t_i^{r_i} = r_i \Delta t_i$ ,  $r_i$  un entier, nous avons l'expression suivante de la pente  $i$

$$\begin{aligned} \left(\frac{\Delta y}{\Delta t}\right)_i^{r_i} &= (BP_i y^{r_i})_i + \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left[ y_j^{r_j} + \delta_{i,j} \left( (BP_j y^{r_j})_j + \sum_{\{i,1\} \in \mathcal{N}^-(j)} F_{\{i,1\},j}^- y_{\{i,1\}}^{r_{\{i,1\}}} \right) \right] \\ &+ \sum_{z=1}^{h_{i,t_s}^{r_i} - 1} \left( \prod_{n=1}^z \delta_{\{i,n-1\},\{i,n\}} \sum_{\{i,n\} \in \mathcal{N}^-(\{i,n-1\})} F_{\{i,n-1\},\{i,n\}}^- \right) \times \\ &\left( (BP_{\{i,z\}} y^{r_{\{i,z\}}})_{\{i,z\}} + \sum_{\{i,z+1\} \in \mathcal{N}^-(\{i,z\})} F_{\{i,z\},\{i,z+1\}}^- y_{\{i,z+1\}}^{r_{\{i,z+1\}}} \right), \end{aligned}$$

où pour tout  $j \in \mathcal{N}^-(i)$ ,  $\delta_{i,j} = r_i \Delta t_i - r_j \Delta t_j > 0$  et  $r_j \Delta t_j$  est le temps de la dernière mise à jour de l'élément  $j$  qui est multiple de  $\Delta t_j$  juste avant  $t_i^{r_i}$ . Pour tout  $1 \leq z < h_{i,t_s}^{r_i} - 1$ ,  $r_{\{i,z\}} \Delta t_{\{i,z\}} < t_i^{r_i}$  et  $\delta_{\{i,z-1\},\{i,z\}} = r_{\{i,z-1\}} \Delta t_{\{i,z-1\}} - r_{\{i,z\}} \Delta t_{\{i,z\}} > 0$  où  $t_s$  est le temps de la dernière synchronisation de la solution et  $h_{i,t_s}^{r_i}$  est le nombre de dépendances présentées par l'élément  $K_i$  entre  $t_i^{r_i}$  et  $t_s$ . Nous supposons aussi que  $K_{\{i,0\}} \equiv K_j$ ,  $j \in \mathcal{N}(i)$  sont les voisins de  $K_i$ .

*Démonstration.* L'expression de la pente  $i$  à l'instant  $t_i^{r_i}$  est donnée par la formule suivante

$$\begin{aligned} \left(\frac{\Delta y}{\Delta t}\right)_i^{r_i} &= V_i y_i^{r_i} + \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ y_i^{r_i} + \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left( y_j^{r_j} + \delta_{i,j} \left(\frac{\Delta y}{\Delta t}\right)_j^{r_j} \right) \\ &= (BP_i y^{r_i})_i + \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left( y_j^{r_j} + \delta_{i,j} \left(\frac{\Delta y}{\Delta t}\right)_j^{r_j} \right), \end{aligned}$$

les éléments  $j$  sont les voisins de l'élément  $i$  et pour tout  $j \in \mathcal{N}^-(i)$ ,  $\delta_{i,j} = r_i \Delta t_i - r_j \Delta t_j > 0$  avec  $r_j \Delta t_j = t_j^{r_j} < t_i^{r_i}$  est le dernier temps, multiple de  $\Delta t_j$ , où l'élément  $j$  a été mis à jour. En utilisant les notations précédentes de la figure 5.2, nous avons le résultat suivant

$$\left(\frac{\Delta y}{\Delta t}\right)_j^{r_j} = (BP_j y^{r_j})_j + \sum_{\{i,1\} \in \mathcal{N}^-(j)} F_{j,\{i,1\}}^- \left( y_{\{i,1\}}^{r_{\{i,1\}}} + \delta_{j,\{i,1\}} \left(\frac{\Delta y}{\Delta t}\right)_{\{i,1\}}^{r_{\{i,1\}}} \right).$$

avec pour tout  $j \in \mathcal{N}(i)$  et pour tout  $\{i,1\} \in \mathcal{N}(j)$ ,  $r_{\{i,1\}} \Delta t_{\{i,1\}} = t_{\{i,1\}}^{r_{\{i,1\}}} < t_j^{r_j}$  est le temps de la dernière mise à jour de l'élément  $K_{\{i,1\}}$  et qui est multiple de  $\Delta t_{\{i,1\}}$ . Alors  $\delta_{j,\{i,1\}} = r_j \Delta t_j - r_{\{i,1\}} \Delta t_{\{i,1\}} > 0$  et

$$\left(\frac{\Delta y}{\Delta t}\right)_{\{i,1\}}^{r_{\{i,1\}}} = (BP_{\{i,1\}} y^{r_{\{i,1\}}})_{\{i,1\}} + \sum_{\{i,2\} \in \mathcal{N}^-(\{i,1\})} F_{\{i,1\},\{i,2\}}^- \left( y_{\{i,2\}}^{r_{\{i,2\}}} + \delta_{\{i,1\},\{i,2\}} \left(\frac{\Delta y}{\Delta t}\right)_{\{i,2\}}^{r_{\{i,2\}}} \right).$$

De plus, nous avons :

$$\left(\frac{\Delta y}{\Delta t}\right)_{\{i,2\}}^{r_{\{i,2\}}} = (BP_{\{i,2\}} y^{r_{\{i,2\}}})_{\{i,2\}} + \sum_{\{i,3\} \in \mathcal{N}^-(\{i,2\})} F_{\{i,2\},\{i,3\}}^- \left( y_{\{i,3\}}^{r_{\{i,3\}}} + \delta_{\{i,2\},\{i,3\}} \left(\frac{\Delta y}{\Delta t}\right)_{\{i,3\}}^{r_{\{i,3\}}} \right).$$



la synchronisation régulière de la solution n'a pas d'influence sur l'aspect asynchrone (voir figure 5.3). Numériquement, la synchronisation de la solution n'est imposée qu'à l'instant final de la simulation indépendamment du maillage. Soit  $(t^q)_{q \geq 0}$  la suite définie par  $t^{q+1} = t^q + \Delta t_{\max}$ . L'étude de la convergence du schéma ARK2 sera faite entre deux instants de synchronisation de la solution :  $t^q$  et  $t^{q+1}$ . Soit  $(t_q^n)_{1 \leq n \leq N_q}$  la suite de mise à jour des événements les plus urgents entre  $t^q$  and  $t^{q+1}$  avec  $N_q = \sum_i s_i$  et pour tout  $i$ ,  $s_i = \lfloor \frac{\Delta t_{\max}}{\Delta t_i} \rfloor$ .  $(t_{i,q}^k)_k$  est la suite de l'évolution temporelle de l'élément  $i$  entre  $t^q$  et  $t^{q+1}$  définie par  $t_{i,q}^k = t^q + k\Delta t_i$  et pour tout  $i$ ,  $\delta_i = t^{q+1} - t_i^{s_i}$  (voir figure 5.3).

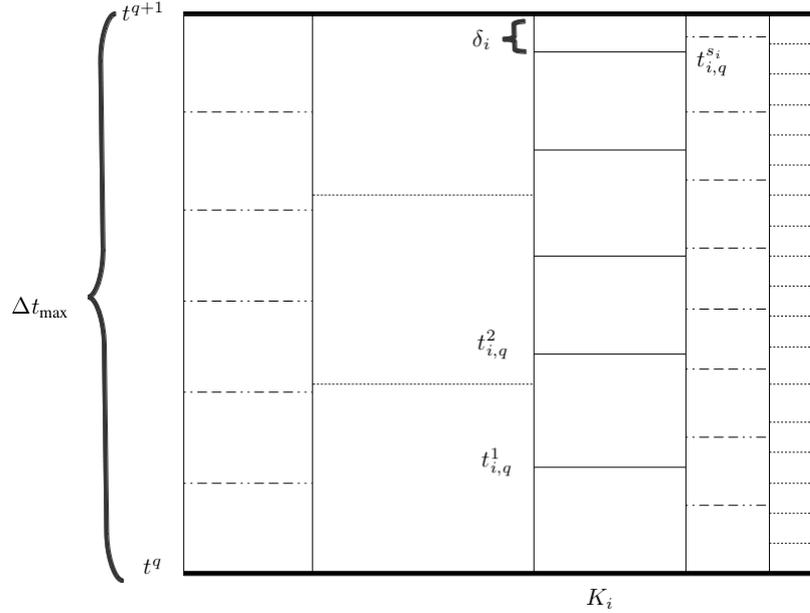


FIGURE 5.3 – Evolution temporelle du système entre deux instants successifs de synchronisation de la solution  $t^q$  et  $t^{q+1} = t^q + \Delta t_{\max}$  ; exemple de cinq éléments.

**Théorème 3.** Pour tout élément  $i$  et à tout instant  $t_{i,q}^{r_i} = t^q + r_i\Delta t_i$  tel que  $t^q < t_{i,q}^{r_i} \leq t^{q+1}$ , l'expression suivante est vérifiée

$$y_i^{t_{i,q}^{r_i}} = y_i^{t^q} + (r_i\Delta t_i)(By^{t^q})_i + \frac{(r_i\Delta t_i)^2}{2}(B^2y^{t^q})_i + \sum_{n=3}^{2r_i} (\Delta t_i)^n \alpha_{n,r_i} (BP_i)^{n-2} (B^2y^{t^q})_i + R_{i,asyn}^{r_i} \quad (5.5)$$

avec :

$$\alpha_{n,0} = 0 \quad \forall \quad n \geq 3, \quad \alpha_{3,r_i} = \alpha_{3,r_i-1} + \frac{(r_i-1)^2 + (r_i-1)}{2} \quad \alpha_{4,r_i} = \alpha_{3,r_i-1} + \alpha_{4,r_i-1} + \frac{(r_i-1)^2}{2}, \quad \alpha_{n,r_i} = \alpha_{n,r_i-1} + \alpha_{n-1,r_i-1} + \frac{\alpha_{n-2,r_i-1}}{2} \quad \forall \quad 5 \leq i \leq 2(r_i-1), \quad \alpha_{2r_i-1,r_i} = \alpha_{2(r_i-1),r_i-1} + \frac{\alpha_{2(r_i-1)-1,r_i-1}}{2}, \quad \alpha_{2r_i,r_i} = \frac{\alpha_{2(r_i-1),r_i-1}}{2}, \quad \text{et}$$

$$\|R_{i,asyn}^{r_i}\|_{\infty} \leq \sum_{n=3}^{M_i^{r_i}} C_{n,r_i} (\Delta t_i)^n \|B^n\|_{\infty} \|y^{t^q}\|_{\infty} \quad (5.6)$$

$M_i^{r_i} = \max(2r_i, h_i^{r_i})$  et pour tout  $3 \leq n \leq M_i^{r_i}$ ,  $C_{n,r_i}$  est une constante bornée.

*Démonstration.* La preuve se fait par récurrence. Nous supposons que la formule (5.5) est vraie pour  $t_q^n < t_q^{q+1}$  et que  $t_q^{n+1} = t_{i,q}^{r_i+1}$ . Alors en utilisant la proposition 14 nous avons le résultat suivant

$$\begin{aligned}
y_i^{t_i^{r_i+1}} &= y_i^{t_i^{r_i}} + \sum_{k=1}^{m+1} \delta_i^k \left[ V_i \left( y_i^{t_i^{r_i}} + \frac{\Delta t_i}{2} \left( \frac{\Delta y}{\Delta t} \right)_i^{t_i^{r_i}} \right) \right] \\
&+ \sum_{k=1}^{m+1} \delta_i^k \left[ \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ \left( y_i^{t_i^{r_i}} + (t_{\text{mid},i,j}^k - r_i \Delta t_i) \left( \frac{\Delta y}{\Delta t} \right)_i^{t_i^{r_i}} \right) \right] \\
&+ \sum_{k=1}^{m+1} \delta_i^k \left[ \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left( y_j^{t_j^{r_j}} + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \left( \frac{\Delta y}{\Delta t} \right)_j^{t_j^{r_j}} \right) \right] \\
&= y_i^{t_i^{r_i}} + \Delta t_i \left( V_i y_i^{t_i^{r_i}} + \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ y_i^{t_i^{r_i}} \right) + \frac{(\Delta t_i)^2}{2} V_i \left( \frac{\Delta y}{\Delta t} \right)_i^{t_i^{r_i}} \\
&+ \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ \left( \sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k - r_i (\Delta t_i)^2 \right) \left( \frac{\Delta y}{\Delta t} \right)_i^{t_i^{r_i}} \\
&+ \sum_{k=1}^{m+1} \delta_i^k \left( \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left( y_j^{t_j^{r_j}} + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \left( \frac{\Delta y}{\Delta t} \right)_j^{t_j^{r_j}} \right) \right).
\end{aligned}$$

D'après la Proposition 15, nous avons  $\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k = \frac{(2r_i+1)(\Delta t_i)^2}{2}, \forall j \in \mathcal{N}(i)$ . Par suite,  $\sum_{k=1}^{m+1} \delta_i^k \times t_{\text{mid},i,j}^k - r_i (\Delta t_i)^2 = \frac{(\Delta t_i)^2}{2}, \forall j \in \mathcal{N}(i)$ . Alors,

$$y_i^{t_i^{r_i+1}} = y_i^{t_i^{r_i}} + \Delta t_i \left( B P_i y_i^{t_i^{r_i}} \right)_i + \frac{(\Delta t_i)^2}{2} B P_i \left( \frac{\Delta y}{\Delta t} \right)_i^{t_i^{r_i}} \quad (5.7)$$

$$+ \sum_{k=1}^{m+1} \delta_i^k \left( \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left( y_j^{t_j^{r_j}} + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \left( \frac{\Delta y}{\Delta t} \right)_j^{t_j^{r_j}} \right) \right) \quad (5.8)$$

Nous utilisons maintenant la proposition 16 et l'hypothèse de récurrence pour avoir le résultat suivant

$$\begin{aligned}
\left(\frac{\Delta y}{\Delta t}\right)_i^{t_i^q} &= BP_i \left[ y_i^{t_i^q} + (r_i \Delta t_i)(By^{t_i^q})_i + \frac{(r_i \Delta t_i)^2}{2} (B^2 y^{t_i^q})_i + R_{i,\text{asyn}}^{t_i^q} \right. \\
&+ \sum_{n=3}^{2r_i} (\Delta t_i)^n \alpha_{n,r_i} (BP_i)^{n-2} (B^2 y^{t_i^q})_i \left. \right] + \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left[ \left[ y_j^{t_j^q} + (r_j \Delta t_j)(By^{t_j^q})_j \right. \right. \\
&+ \frac{(r_j \Delta t_j)^2}{2} (B^2 y^{t_j^q})_j + \sum_{n=3}^{2r_j} (\Delta t_j)^n \alpha_{n,r_j} (BP_j)^{n-2} (B^2 y^{t_j^q})_j + R_{j,\text{asyn}}^{t_j^q} \left. \right] \\
&+ \delta_{i,j} \left[ BP_j \left[ y_j^{t_j^q} + (r_j \Delta t_j)(By^{t_j^q})_j + \frac{(r_j \Delta t_j)^2}{2} (B^2 y^{t_j^q})_j + R_{j,\text{asyn}}^{t_j^q} \right. \right. \\
&+ \sum_{n=3}^{2r_j} (\Delta t_j)^n \alpha_{n,r_j} (BP_j)^{n-2} (B^2 y^{t_j^q})_j \left. \right] + \sum_{\{i,1\} \in \mathcal{N}^-(j)} F_{\{i,1\},j}^- \left[ y_{\{i,1\}}^{t_{\{i,1\}}^q} \right. \\
&+ (r_{\{i,1\}} \Delta t_{r_{\{i,1\}}}) (By^{t_{\{i,1\}}^q})_{\{i,1\}} + \frac{(r_{\{i,1\}} \Delta t_i)^2}{2} (B^2 y^{t_{\{i,1\}}^q})_{\{i,1\}} \\
&+ \sum_{n=3}^{2r_{\{i,1\}}} (\Delta t_{r_{\{i,1\}}})^n \alpha_{n,r_{\{i,1\}}} (BP_{\{i,1\}})^{n-2} (B^2 y^{t_{\{i,1\}}^q})_{\{i,1\}} + R_{\{i,1\},\text{asyn}}^{t_{\{i,1\}}^q} \left. \right] \left. \right] \\
&+ \sum_{z=1}^{h_{i,t_i^q}^{r_i}-1} \left( \prod_{n=1}^z \delta_{\{i,n-1\},\{i,n\}} \sum_{\{i,n\} \in \mathcal{N}^-(\{i,n-1\})} F_{\{i,n-1\},\{i,n\}}^- \right) \times \\
&\left[ BP_{\{i,z\}} \left[ y_{\{i,z\}}^{t_{\{i,z\}}^q} + (r_{\{i,z\}} \Delta t_{r_{\{i,z\}}}) (By^{t_{\{i,z\}}^q})_{\{i,z\}} + \frac{(r_{\{i,z\}} \Delta t_i)^2}{2} (B^2 y^{t_{\{i,z\}}^q})_{\{i,z\}} \right. \right. \\
&+ \sum_{n=3}^{2r_{\{i,z\}}} (\Delta t_{r_{\{i,z\}}})^n \alpha_{n,r_{\{i,z\}}} (BP_{\{i,z\}})^{n-2} (B^2 y^{t_{\{i,z\}}^q})_{\{i,z\}} + R_{\{i,z\},\text{asyn}}^{t_{\{i,z\}}^q} \left. \right] \\
&+ \sum_{\{i,z+1\} \in \mathcal{N}^-(\{i,z\})} F_{\{i,z\},\{i,z+1\}}^- \left[ y_{\{i,z+1\}}^{t_{\{i,z+1\}}^q} + R_{\{i,z+1\},\text{asyn}}^{t_{\{i,z+1\}}^q} \right. \\
&+ (r_{\{i,z+1\}} \Delta t_{r_{\{i,z+1\}}}) (By^{t_{\{i,z+1\}}^q})_{\{i,z+1\}} + \frac{(r_{\{i,z+1\}} \Delta t_i)^2}{2} (B^2 y^{t_{\{i,z+1\}}^q})_{\{i,z+1\}} \\
&+ \sum_{n=3}^{2r_{\{i,z+1\}}} (\Delta t_{r_{\{i,z+1\}}})^n \alpha_{n,r_{\{i,z+1\}}} (BP_{\{i,z+1\}})^{n-2} (B^2 y^{t_{\{i,z+1\}}^q})_{\{i,z+1\}} \left. \right] \left. \right]
\end{aligned}$$

Par conséquent,

$$\begin{aligned}
\left(\frac{\Delta y}{\Delta t}\right)_i^{t_i^q} &= (By^{t_i^q})_i + (r_i \Delta t_i)(B^2 y^{t_i^q})_i + \frac{(r_i \Delta t_i)^2}{2} BP_i (B^2 y^{t_i^q})_i \\
&+ \sum_{n=3}^{2r_i} (\Delta t_i)^n \alpha_{n,r_i} (BP_i)^{n-1} (B^2 y^{t_i^q})_i + \tilde{R}_{i,\text{asyn}}^{t_i^q}
\end{aligned}$$

avec

$$\begin{aligned}
\tilde{R}_{i,\text{asyn}}^{t_i, r_i} &= \left( BP_i R_{i,\text{asyn}}^{t_i, r_i} + \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- R_{j,\text{asyn}}^{t_i, r_i} \right) \\
&+ \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left[ \delta_{i,j} \left( BP_j R_{j,\text{asyn}}^{t_j, r_j} + \sum_{\{i,1\} \in \mathcal{N}^-(j)} F_{\{i,1\},j}^- R_{\{i,1\},\text{asyn}}^{r_{\{i,1\}}} \right) \right] \\
&+ \sum_{z=1}^{t_i^{i,q}-1} \left( \prod_{n=1}^z \delta_{\{i,n-1\}, \{i,n\}} \sum_{\{i,n\} \in \mathcal{N}^-(\{i,n-1\})} F_{\{i,n-1\}, \{i,n\}}^- \right) \times \\
&\left( BP_{\{i,z\}} R_{\{i,z\},\text{asyn}}^{r_{\{i,z\}}} + \sum_{\{i,z+1\} \in \mathcal{N}^-(\{i,z\})} F_{\{i,z\}, \{i,z+1\}}^- R_{\{i,z+1\},\text{asyn}}^{r_{\{i,z+1\}}} \right) \\
&+ \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left[ \left[ \frac{(r_j \Delta t_j)^2}{2} (B^2 y^{t^q})_j + \sum_{n=3}^{2r_j} (\Delta t_j)^n \alpha_{n,r_j} (BP_j)^{n-2} (B^2 y^{t^q})_j \right] \right. \\
&+ \delta_{i,j} \left[ BP_j \left[ (r_j \Delta t_j) (B y^{t^q})_j + \frac{(r_j \Delta t_j)^2}{2} (B^2 y^{t^q})_j \right. \right. \\
&+ \left. \left. \sum_{n=3}^{2r_j} (\Delta t_j)^n \alpha_{n,r_j} (BP_j)^{n-2} (B^2 y^{t^q})_j \right] \right] \\
&+ \sum_{\{i,1\} \in \mathcal{N}^-(j)} F_{\{i,1\},j}^- \left[ (r_{\{i,1\}} \Delta t_{r_{\{i,1\}}}) (B y^{t^q})_{\{i,1\}} + \frac{(r_{\{i,1\}} \Delta t_i)^2}{2} (B^2 y^{t^q})_{\{i,1\}} \right. \\
&+ \left. \left. \sum_{n=3}^{2r_{\{i,1\}}} (\Delta t_{r_{\{i,1\}}})^n \alpha_{n,r_{\{i,1\}}} (BP_{\{i,1\}})^{n-2} (B^2 y^{t^q})_{\{i,1\}} \right] \right] \\
&+ \sum_{z=1}^{t_i^{i,q}-1} \left( \prod_{n=1}^z \delta_{\{i,n-1\}, \{i,n\}} \sum_{\{i,n\} \in \mathcal{N}^-(\{i,n-1\})} F_{\{i,n-1\}, \{i,n\}}^- \right) \times \\
&\left[ BP_{\{i,z\}} \left[ y_{\{i,z\}}^{t^q} + (r_{\{i,z\}} \Delta t_{r_{\{i,z\}}}) (B y^{t^q})_{\{i,z\}} + \frac{(r_{\{i,z\}} \Delta t_i)^2}{2} (B^2 y^{t^q})_{\{i,z\}} \right. \right. \\
&+ \left. \left. \sum_{n=3}^{2r_{\{i,z\}}} (\Delta t_{r_{\{i,z\}}})^n \alpha_{n,r_{\{i,z\}}} (BP_{\{i,z\}})^{n-2} (B^2 y^{t^q})_{\{i,z\}} \right] \right. \\
&+ \sum_{\{i,z+1\} \in \mathcal{N}^-(\{i,z\})} F_{\{i,z\}, \{i,z+1\}}^- \left[ y_{\{i,z+1\}}^{t^q} + (r_{\{i,z+1\}} \Delta t_{r_{\{i,z+1\}}}) (B y^{t^q})_{\{i,z+1\}} \right. \\
&+ \left. \left. \sum_{n=3}^{2r_{\{i,z+1\}}} (\Delta t_{r_{\{i,z+1\}}})^n \alpha_{n,r_{\{i,z+1\}}} (BP_{\{i,z+1\}})^{n-2} (B^2 y^{t^q})_{\{i,z+1\}} \right. \right. \\
&+ \left. \left. \frac{(r_{\{i,z+1\}} \Delta t_i)^2}{2} (B^2 y^{t^q})_{\{i,z+1\}} \right] \right]
\end{aligned}$$

Nous remplaçons ensuite  $\left(\frac{\Delta y}{\Delta t}\right)_i^{r_i}$  par la dernière expression dans la formule de  $y_i^{r_i+1}$  donnée par l'équation (5.7) puis nous utilisons l'hypothèse de récurrence pour remplacer les valeurs de  $y_i^{r_i}$  et  $y_j^{r_j}$  pour tout  $j \in \mathcal{N}(i)$ . Ce qui donne le résultat suivant

$$\begin{aligned}
y_i^{r_i+1} &= \left[ y_i^{t^q} + (r_i \Delta t_i)(By^{t^q})_i + \frac{(r_i \Delta t_i)^2}{2}(B^2 y^{t^q})_i + \sum_{n=3}^{2r_i} (\Delta t_i)^n \alpha_{n,r_i}(BP_i)^{n-2}(B^2 y^{t^q})_i \right. \\
&\quad \left. + R_{i,\text{asyn}}^{t_i,q} \right] + \Delta t_i BP_i \left[ y_i^{t^q} + (r_i \Delta t_i)(By^{t^q})_i + \frac{(r_i \Delta t_i)^2}{2}(B^2 y^{t^q})_i \right. \\
&\quad \left. + \sum_{n=3}^{2r_i} (\Delta t_i)^n \alpha_{n,r_i}(BP_i)^{n-2}(B^2 y^{t^q})_i + R_{i,\text{asyn}}^{t_i,q} \right] + \frac{(\Delta t_i)^2}{2} BP_i \left[ (By^{t^q})_i + (r_i \Delta t_i)(B^2 y^{t^q})_i \right. \\
&\quad \left. + \frac{(r_i \Delta t_i)^2}{2} BP_i (B^2 y^{t^q})_i + \sum_{n=3}^{2r_i} (\Delta t_i)^n \alpha_{n,r_i}(BP_i)^{n-1}(B^2 y^{t^q})_i + \tilde{R}_{i,\text{asyn}}^{t_i,q} \right] \\
&\quad + \sum_{k=1}^{m+1} \delta_i^k \left[ \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left[ \left[ y_j^{t^q} + (r_j^k \Delta t_j)(By^{t^q})_j + \frac{(r_j^k \Delta t_j)^2}{2}(B^2 y^{t^q})_j \right. \right. \right. \\
&\quad \left. \left. + \sum_{n=3}^{2r_j^k} (\Delta t_j)^n \alpha_{n,r_j^k}(BP_j)^{n-2}(B^2 y^{t^q})_j + R_{j,\text{asyn}}^{t_j,q} \right] \right. \\
&\quad \left. + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \left[ (By^{t^q})_j + (r_j^k \Delta t_j)(B^2 y^{t^q})_j + \frac{(r_j^k \Delta t_j)^2}{2} BP_j (B^2 y^{t^q})_j \right. \right. \\
&\quad \left. \left. + \sum_{n=3}^{2r_j^k} (\Delta t_j)^n \alpha_{n,r_j^k}(BP_j)^{n-1}(B^2 y^{t^q})_j + \tilde{R}_{j,\text{asyn}}^{t_j,q} \right] \right].
\end{aligned}$$

Nous utilisons ensuite la proposition 15 pour calculer le terme suivant

$$\begin{aligned}
&\frac{(r_i \Delta t_i)^2}{2}(B^2 y^{t^q})_i + (r_i (\Delta t_i)^2) BP_i \left( (By^{t^q})_i \right) + \frac{(\Delta t_i)^2}{2} BP_i \left( (By^{t^q})_i \right) \\
&+ \sum_{k=1}^{m+1} \delta_i^k \left[ \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left[ (r_j^k \Delta t_j)(By^{t^q})_j + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) (By^{t^q})_j \right] \right] \\
&= \frac{((r_i + 1) \Delta t_i)^2}{2} (B^2 y^{t^q})_i.
\end{aligned}$$

Par suite,

$$\begin{aligned}
y_i^{r_i+1} &= y_i^{t^q} + ((r_i + 1) \Delta t_i)(By^{t^q})_i + \frac{((r_i + 1) \Delta t_i)^2}{2}(B^2 y^{t^q})_i \\
&\quad + \sum_{n=3}^{2(r_i+1)} (\Delta t_i)^n \alpha_{n,r_i+1}(BP_i)^{n-2}(B^2 y^{t^q})_i + R_{i,\text{asyn}}^{t_i,q},
\end{aligned}$$

avec

$$\begin{aligned}
R_{i,\text{asyn}}^{t_i^{r_i+1}} &= R_{i,\text{asyn}}^{t_i^{r_i}} + \Delta t_i B P_i R_{i,\text{asyn}}^{t_i^{r_i}} + \frac{(\Delta t_i)^2}{2} B P_i \tilde{R}_{i,\text{asyn}}^{t_i^{r_i}} \\
&+ \sum_k^{m+1} \delta_i^k \left( R_{j,\text{asyn}}^{t_j^{r_j^k}} + (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \tilde{R}_{j,\text{asyn}}^{t_j^{r_j^k}} \right) \\
&+ \sum_k^{m+1} \delta_i^k \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- \left[ \frac{(r_j^k \Delta t_j)^2}{2} (B^2 y^{t^q})_j + \sum_{n=3}^{2r_j^k} (\Delta t_j)^n \alpha_{n,r_j^k} (B P_j)^{n-2} (B^2 y^{t^q})_j \right. \\
&+ (t_{\text{mid},i,j}^k - r_j^k \Delta t_j) \left[ (r_j^k \Delta t_j) (B^2 y^{t^q})_j + \frac{(r_j^k \Delta t_j)^2}{2} B P_j (B^2 y^{t^q})_j \right. \\
&\left. \left. + \sum_{n=3}^{2r_j^k} (\Delta t_j)^n \alpha_{n,r_j^k} (B P_j)^{n-1} (B^2 y^{t^q})_j \right] \right]
\end{aligned}$$

Pour conclure, il suffit d'utiliser encore une fois l'hypothèse de récurrence, en particulier la formule (5.6), pour montrer que  $R_{i,\text{asyn}}^{t_i^{r_i}}$  et  $\tilde{R}_{i,\text{asyn}}^{t_i^{r_i}}$  sont bornés. La formule (5.5) est donc vérifiée à l'instant  $t_i^{r_i+1}$  avec

$$\|R_{i,\text{asyn}}^{t_i^{r_i+1}}\|_\infty \leq \sum_{n=3}^{M_i^{r_i+1}} C_{n,r_i+1} (\Delta t_i)^n \|B^n\|_\infty \|y^0\|_\infty,$$

$$M_i^{r_i+1} = \max(2(r_i + 1), h_{i,q}^{r_i+1}). \quad \square$$

**Théorème 4.** *Le schéma Runge-Kutta 2 asynchrone est convergent d'ordre deux.*

*Démonstration.* Soient  $y_i(t^{q+1})$  la solution exacte d'une équation différentielle linéaire donnée à l'instant  $t^{q+1}$ ,  $y_{i,\text{syn}}^{t^{q+1}}$  est la solution classique du schéma RK2 à pas de temps global à l'instant  $t^{q+1}$  et  $y_i^{t^{q+1}}$  la solution asynchrone. Alors,

$$\|y_i(t^{q+1}) - y_i^{t^{q+1}}\|_\infty \leq \|y_i(t^{q+1}) - y_{i,\text{syn}}^{t^{q+1}}\|_\infty + \|y_{i,\text{syn}}^{t^{q+1}} - y_i^{t^{q+1}}\|_\infty \quad (5.9)$$

Le schéma RK2 classique à pas de temps global est convergent d'ordre deux, alors

$$\|y_i(t^{q+1}) - y_{i,\text{syn}}^{t^{q+1}}\|_\infty \leq C_0 (\Delta t_{\text{max}})^3,$$

avec pour  $\Delta t_{\text{max}}$  suffisamment petit,  $C_0$  est une constante indépendante de  $\Delta t_{\text{max}}$ . Pour finir la preuve, il faut traiter le deuxième terme à droite de (5.9). Tout d'abord, en utilisant le Théorème 3, nous avons l'expression suivante pour tout  $i$

$$y_i^{t_q^{s_i}} = y_i^{t^q} + (s_i \Delta t_i) (B y^{t^q})_i + \frac{(s_i \Delta t_i)^2}{2} (B^2 y^{t^q})_i + \sum_{n=3}^{2s_i} (\Delta t_i)^n \alpha_{n,s_i} (B P_i)^{n-2} (B^2 y^{t^q})_i + R_{i,\text{asyn}}^{t_q^{s_i}}$$

où pour tout  $i$ ,  $t_q^{s_i} = t^q + s_i \Delta t_i$ ,  $s_i = \left\lfloor \frac{\Delta t_{\text{max}}}{\Delta t_i} \right\rfloor$ . Supposons dans un premier temps que  $\frac{\Delta t_{\text{max}}}{\Delta t_i}$  est un entier.

Dans ce cas, nous avons

$$y_i^{t^{q+1}} \equiv y_i^{t^{s_i}} = y_i^{t^q} + (\Delta t_{\max})(By^{t^q})_i + \frac{(\Delta t_{\max})^2}{2}(B^2y^{t^q})_i + \sum_{n=3}^{2s_i} (\Delta t_i)^n \alpha_{n,s_i} (BP_i)^{n-2} (B^2y^{t^q})_i + R_{i,\text{asyn}}^{t^{q+1}}.$$

Par conséquent,

$$\|y_{i,\text{syn}}^{t^{q+1}} - y_i^{t^{q+1}}\|_{\infty} = \|E_i^{t^{q+1}}\|_{\infty},$$

avec

$$\begin{aligned} E_i^{t^{q+1}} &= \sum_{n=3}^{2s_i} (\Delta t_i)^n \alpha_{n,s_i} B^{n-2} (B^2y^{t^q})_i - \sum_{n=3}^{2s_i} (\Delta t_i)^n \alpha_{n,s_i} (BP_i)^{n-2} (B^2y^{t^q})_i - R_{i,\text{asyn}}^{t^{q+1}} \\ &= \sum_{n=3}^{2s_i} (\Delta t_i)^n \alpha_{n,s_i} \left[ \sum_{j \in \mathcal{N}^-(i)} F_{i,j} \right]^{n-2} (B^2y^{t^q})_j - R_{i,\text{asyn}}^{t^{q+1}}, \end{aligned}$$

où nous avons utilisé le fait que  $B$  est linéaire. Par suite,

$$\|E_i^{t^{q+1}}\|_{\infty} \leq C_1 (\Delta t_{\max})^3, \quad (5.10)$$

avec pour  $\Delta t_{\max}$  suffisamment petit,  $C_1$  est une constante indépendante de  $\Delta t_{\max}$ . Enfin, notons que dans le cas où  $\frac{\Delta t_{\max}}{\Delta t_i}$  n'est pas un entier, il est juste nécessaire de synchroniser la solution à l'instant  $t^{q+1}$  en utilisant une interpolation d'ordre 2 comme cela a été expliqué quand l'algorithme ARK2 a été présenté.  $\square$

### 5.2.2. Comparaison du schéma ARK2 avec le schéma LTS-RK2 défini dans [35]

Dans [35], Grote et al. ont dérivé un schéma Runge-Kutta à pas de temps local (LTS-RK) du schéma Runge-Kutta classique (RK) à pas de temps global. Ils ont supposé que le maillage est composé de deux régions : une région grossière avec un grand pas de temps et une région fine avec un petit pas de temps. Les auteurs ont utilisé la méthode RK classique, à un ordre donné, pour faire avancer la région fine, alors qu'ils ont utilisé une formule de quadrature dans la région grossière. Les valeurs intermédiaires nécessaires aux interfaces sont obtenues grâce à une combinaison judicieuse d'une interpolation et d'un développement de Taylor. Ils ont prouvé que leur schéma préserve la précision du schéma RK classique. Cependant, cette approche n'est applicable que dans le cas où le pas de temps de la région grossière doit forcément être un multiple du pas de temps de la région fine. Cette hypothèse est indispensable pour pouvoir appliquer la méthode LTS-RK.

Cette limite a été notre point de départ. Nous nous sommes inspirés de leur méthode pour développer le schéma ARK2 qui n'impose aucune condition sur le choix du pas de temps. Nous avons donc commencé par adapter l'algorithme présenté dans [35] à l'aspect asynchrone dans le but de pouvoir appliquer les trois phases principales de l'intégration asynchrone. En particulier, chaque élément de maillage doit être traité indépendamment des autres éléments. Nous avons prouvé que dans le cas du maillage imposé dans [35], notre schéma ARK2 dégénère vers le schéma LTS-RK2. Plus précisément, en utilisant l'algorithme

asynchrone, nous retrouvons le même résultat du calcul LTS-RK2 à chaque pas de temps global  $\Delta t$ , où  $\Delta t$  est le pas de temps calculé sous la condition CFL locale dans la région grossière. Tandis qu'aux pas de temps de la forme  $n\Delta t/p$  où  $\Delta t/p$  est le pas de temps local de la partie fine du maillage,  $p$  et  $n$  sont deux entiers tels que  $0 < n < p$ , les résultats issus des algorithmes ARK2 et LTS-RK2 sont différents. En effet, nous avons le résultat suivant :

**Proposition 17.** *Dans le cas d'un maillage composé de deux régions distinctes où le pas de temps de la partie grossière est multiple du pas de temps de la partie fine, le schéma ARK2 dégénère vers le schéma LTS-RK2 présenté dans [35].*

*Démonstration.* Reprenons les notations utilisées dans [35] :  $\Omega = \Omega_c \cap \Omega_f$  où  $\Omega_c$  est la région grossière avec le pas de temps global  $\Delta t_c$  et  $\Omega_f$  est la région fine avec le petit pas de temps  $\Delta t_f = \frac{\Delta t_c}{p}$ ,  $p$  un entier. Nous introduisons quelques éléments pour l'identification :

$$\begin{aligned} (Bx)_i &= V_i x_i + \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ x_j + \sum_{j \in \mathcal{N}^-(i)} F_{i,j}^- x_j \\ (BPx)_i &= V_i x_i + \sum_{j \in \mathcal{N}^+(i)} F_{i,j}^+ x_j + \sum_{j \in \mathcal{N}_f^-(i)} F_{i,j}^- x_j \\ (B(I - P)x)_i &= \sum_{j \in \mathcal{N}_c^-(i)} F_{i,j}^- x_j \end{aligned}$$

où  $\mathcal{N}_c(i)$  et  $\mathcal{N}_f(i)$  sont les voisins de l'élément  $i$  qui sont respectivement dans le maillage grossier et dans le maillage fin. Autrement dit, si  $j \in \mathcal{N}_c(i)$  alors  $j \in \mathcal{N}(i) \cap \Omega_c$  et si  $j \in \mathcal{N}_f(i)$  alors  $j \in \mathcal{N}(i) \cap \Omega_f$ .  $P$  est une matrice diagonale et ses éléments diagonaux sont égaux à zéro ou un.  $P$  identifie les inconnues associées à la région raffinée tandis que  $(I - P)$  identifie les inconnues associées à la région grossière. Les résultats de la section précédente sont valables pour tout type de maillage, en particulier pour un maillage à deux régions. En effet, en utilisant les notations précédentes, nous avons  $\Delta t_{\max} = \Delta t_c$ . Ce choix est naturel dans ce cas car la solution est automatiquement synchronisée aux instants qui sont multiples de  $\Delta t_c$ . Par suite, la suite  $(t^q)_q$  est définie par  $t^{q+1} = t^q + \Delta t_c$ , pour tout  $i \in \Omega_f$   $(t_{i,q}^k)_{0 \leq k \leq p}$  est définie par  $t_{i,q}^k = k\Delta t_f$  et pour tout  $j \in \Omega_c$ ,  $(t_{j,q}^k)_{0 \leq k \leq 1}$  est définie par  $t_{j,q}^0 = t^q$  et  $t_{j,q}^1 = t^{q+1}$ . Nous utilisons les résultats des propositions 14 et 15 pour un élément fin, la preuve pour un élément grossier se fait de la même manière. Pour tout  $i \in \Omega_f$  nous avons le résultat suivant :

$$\begin{aligned} y_i^{t_{i,q}^{r_i+1}} &= y_i^{t_{i,q}^{r_i}} + \Delta t_f (BP y^{t_{i,q}^{r_i}})_i + \frac{(\Delta t_f)^2}{2} \left( BP \left( \frac{\Delta y}{\Delta t} \right)^{t_{i,q}^{r_i}} \right)_i \\ &\quad + \Delta t_f (B(I - P)y^{t^q})_i + \frac{(2r_i + 1)(\Delta t_f)^2}{2} \left( B(I - P) \left( \frac{\Delta y}{\Delta t} \right)^{t^q} \right)_i, \end{aligned}$$

avec

$$\left( \frac{\Delta y}{\Delta t} \right)_i^{t_{i,q}^{r_i}} = (BP y^{t_{i,q}^{r_i}})_i + \left( B(I - P) \left( y^{t^q} + r_i \Delta t_f \left( \frac{\Delta y}{\Delta t} \right)^{t^q} \right) \right)_i.$$

En comparant cette dernière expression avec celle donnée dans la proposition 16, nous pouvons remarquer que les dépendances avec les voisins sont simplifiées. En effet, dans ce cas particulier, il y a seulement deux types d'éléments : les cellules fines qui sont regroupées dans la matrice  $P$  et les cellules grossières qui sont regroupées dans la matrice  $(I - P)$ . Tous les éléments fins ont été mis à jour aux mêmes instants ; c'est aussi le cas pour les éléments grossiers. Par conséquent, dans le théorème 3, l'expression de  $y_i$  dans 5.5,  $i$  un élément fin, est donnée par la formule suivante

$$y_i^{t_i, q} = y_i^{t^q} + (r_i \Delta t_f)(By^{t^q})_i + \frac{(r_i \Delta t_f)^2}{2}(B^2 y^{t^q})_i + \sum_{m=3}^{2r_i} (\Delta t_f)^m \alpha_{m, r_i} (BP_i)^{m-2} (B^2 y^{t^q})_i,$$

par suite nous retrouvons la même expression donnée dans [35]. □

**Remarque 4.** *L'originalité de notre approche est qu'elle est applicable à tout type de maillage. Le schéma ARK2 est aussi convergent d'ordre deux indépendamment du maillage choisi.*

### 5.3. Validation numérique

Nous allons maintenant présenter des tests numériques qui valident les résultats théoriques prouvés dans la section précédente. En particulier, nous étudions numériquement l'ordre de convergence du schéma ARK2 dans différents types de maillages. Nous considérons les équations de Maxwell pour tester le schéma ARK2. Pour la discrétisation spatiale, nous avons choisi la méthode de Galerkin discontinue upwind avec des fonctions de base type  $P^1$ . Nous traiterons dans un premier temps le cas unidimensionnel puis nous ajouterons des tests en dimension deux d'espace.

#### 5.3.1. Tests numériques en dimension un d'espace

Soit l'équation de Maxwell en 1D :

$$\begin{cases} \frac{\partial(\varepsilon E)}{\partial t} - \frac{\partial H}{\partial x} = 0 \\ \frac{\partial(\mu H)}{\partial t} - \frac{\partial E}{\partial x} = 0, \end{cases} \quad (5.11)$$

la condition initiale est égale à  $\cos(2\pi x)$ . Ce système est complété par des conditions aux limites périodiques de sorte que la solution exacte après une période,  $T = 1$ s, soit égale à la solution initiale. Nous appliquons la méthode ARK2 pour différents types de maillages. Nous introduisons alors ces différents types de maillage dans les test numériques pour étudier la convergence de la méthode asynchrone.

1. Tout d'abord, nous considérons des cas simples de maillages où le pas de temps local est un multiple des petits pas de temps. Le premier maillage est celui dénoté par Mesh 1 dans la figure 5.4 (à gauche) avec  $q = 4$  et le deuxième maillage est celui dénoté par Mesh 2 dans la figure 5.4 (à droite), avec  $p = 4$  et  $q = 8$ .
2. Ensuite, nous introduisons des maillages plus généraux où le pas de temps global n'est pas forcément un multiple des pas de temps locaux, mais les éléments de maillage ont des temps de rencontre réguliers. Le troisième maillage, Mesh 3, est présenté dans la figure 5.5 (à gauche) avec

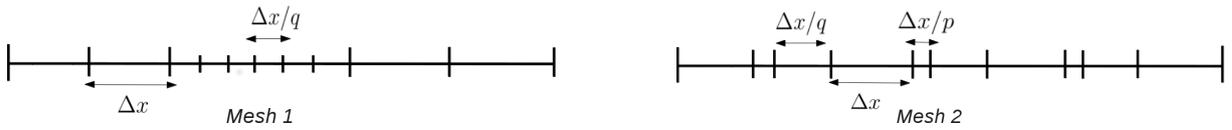


FIGURE 5.4 – Le pas de temps global est multiple des pas de temps locaux : Mesh 1 (avec  $q = 4$ ) et mesh 2 (avec  $p = 4$  et  $q = 8$ ).

$p = 3$ ,  $q = 2$  et le quatrième maillage, Mesh 4, sur la figure 5.5 (à droite) avec  $p/q < 1$  et  $p = 2$ ,  $q = 3$ .

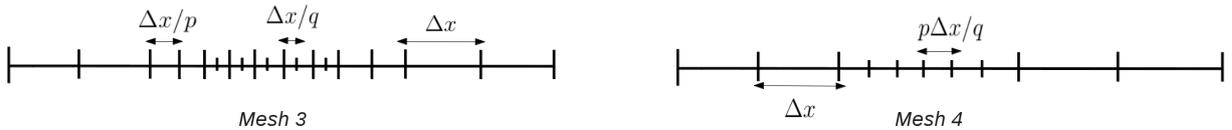


FIGURE 5.5 – Le pas de temps global n'est pas multiple des pas de temps locaux : Mesh 3 (avec  $p = 3$ ,  $q = 2$ ) et mesh 4 (avec  $p = 2$ ,  $q = 3$ ).

3. Enfin, l'algorithme est testé dans le cas du maillage polynomial (2.14) où les pas de temps sont complètement indépendants.

Après avoir choisi les maillages, nous traçons la norme infinie de l'erreur qui correspond à chaque maillage en fonction de nombre de cellules, en échelle logarithmique. Sur les figures 5.6, 5.7 et 5.8, la droite "ordre2" désigne l'ordre 2 théorique.

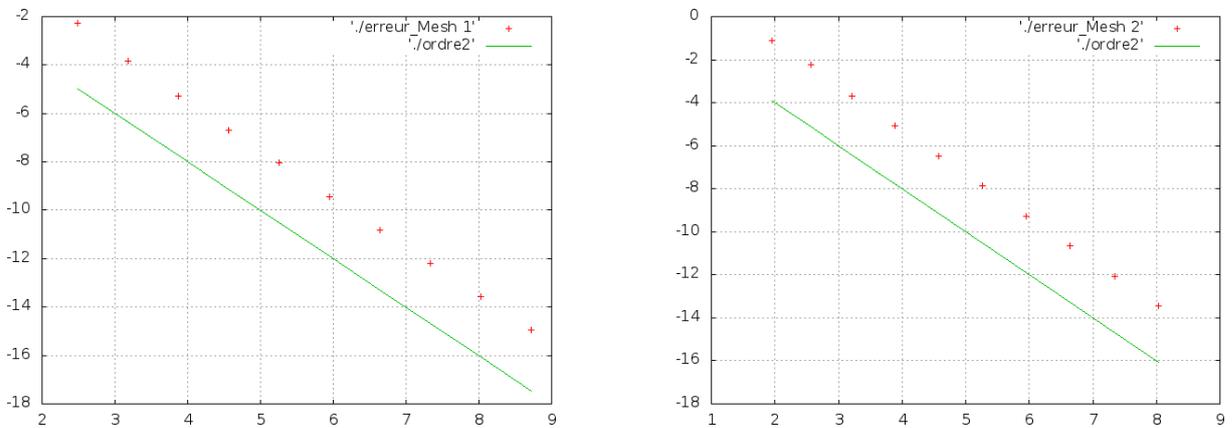


FIGURE 5.6 – La norme infinie de l'erreur dans le cas du maillage 1D : Mesh 1 à gauche et Mesh 2 à droite.

On observe que l'ordre deux attendu est effectivement atteint par le schéma asynchrone ARK2.

### 5.3.2. Tests numériques en deux dimensions

Soient les équations de Maxwell en dimension deux :

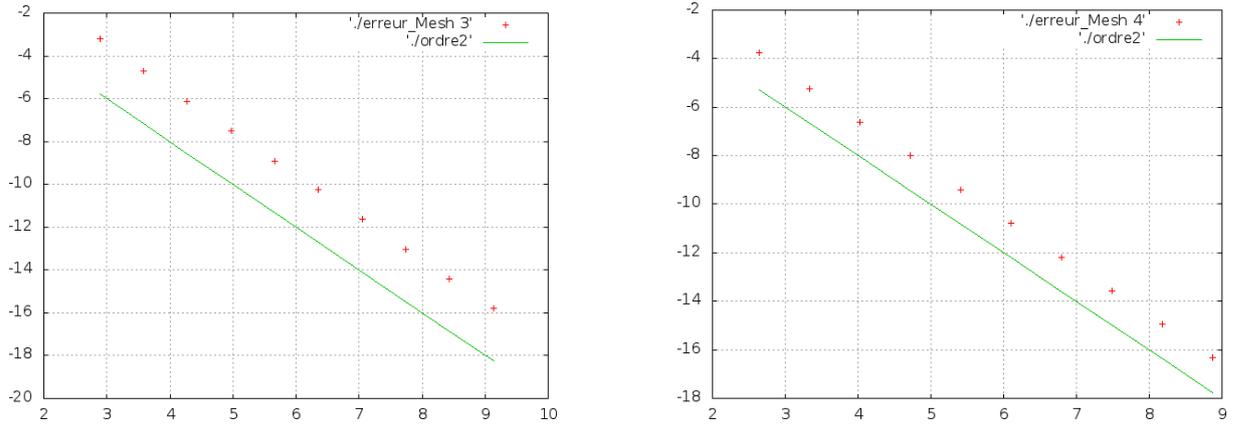


FIGURE 5.7 – La norme infinie de l’erreur dans le cas du maillage 1D : Mesh 3 à gauche et Mesh 4 à droite.

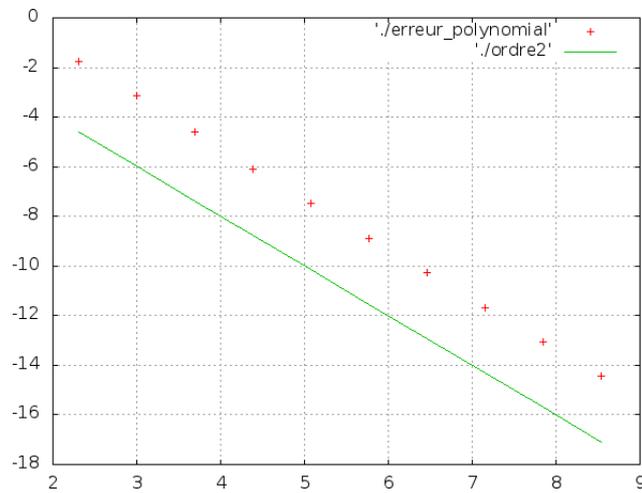


FIGURE 5.8 – La norme infinie de l’erreur dans le cas du maillage polynomial en 1D.

$$\begin{cases} \frac{\partial E_x}{\partial t} = \frac{\partial H}{\partial y} \\ \frac{\partial E_y}{\partial t} = -\frac{\partial H}{\partial x} \\ \frac{\partial H_z}{\partial t} = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \end{cases} \quad (5.12)$$

La condition initiale :  $(E_x)_0 = (E_y)_0 = (H_z)_0 = \cos(2\pi x)$ . Les conditions aux limites sont toujours supposées périodiques et la solution exacte à l’instant  $T = 1s$  est donc égale à la solution à l’instant initial. Le domaine d’étude est  $\Omega = [0, 1] \times [0, 1]$ . En ce qui concerne la discrétisation spatiale, nous utilisons un schéma Galerkin Discontinu upwind  $P^1$ .

1. Pour les premiers tests numériques, nous commençons par diviser l’intervalle  $[0, 1]$  dans la direction  $X$ , ainsi que dans la direction  $Y$  en trois parties égales. Les intervalles à gauche et à droite dans chacune de deux directions sont discrétisées avec un maillage équidistant dont le pas d’espace est égal respectivement à  $\Delta x$ , et  $\Delta y$ . L’intervalle au milieu est discrétisé avec un maillage équidistant dont le pas d’espace est égal à  $p\Delta x/q$  et  $p\Delta y/q$ , avec  $p$  et  $q$  deux entiers. Le maillage

est représenté sur la figure 5.9 (à gauche).

- (a) Nous commençons par un cas relativement simple du maillage où le pas de temps global est un multiple du pas de temps local. C'est typiquement le cas étudié dans [35]. Mesh 5 correspond au maillage présenté dans la figure 5.9 avec  $p = 4$  et  $q = 1$ .
- (b) Ensuite, nous ajoutons un test dans un cas plus général où nous supposons que le pas du temps n'est pas forcément une fraction du pas global, mais les éléments du maillage ont des temps de rencontre réguliers. Soit donc le maillage Mesh 6 avec  $p = 2$ ,  $q = 3$  et donc  $p/q < 1$ .

2. Un dernier maillage, le maillage polynomial, où les pas de temps sont supposés complètement indépendants. Les positions  $x_i$  et  $y_i$  suivent toujours une loi polynomiale :

$$x_i = 4(1 - \alpha) \left( \frac{i - 1}{N_x - 1} - \frac{1}{2} \right)^3 + \alpha \left( \frac{i - 1}{N_x - 1} - \frac{1}{2} \right) + \frac{1}{2}, \quad (5.13)$$

et

$$y_i = 4(1 - \alpha) \left( \frac{i - 1}{N_y - 1} - \frac{1}{2} \right)^3 + \alpha \left( \frac{i - 1}{N_y - 1} - \frac{1}{2} \right) + \frac{1}{2}, \quad (5.14)$$

où  $N_x$  et  $N_y$ , sont respectivement le nombre d'éléments dans la direction  $X$ , et dans la direction  $Y$  (voir figure 5.9, à droite). Le paramètre  $\alpha$  définit la pente aux points  $x = 1/2$  et  $y = 1/2$ .

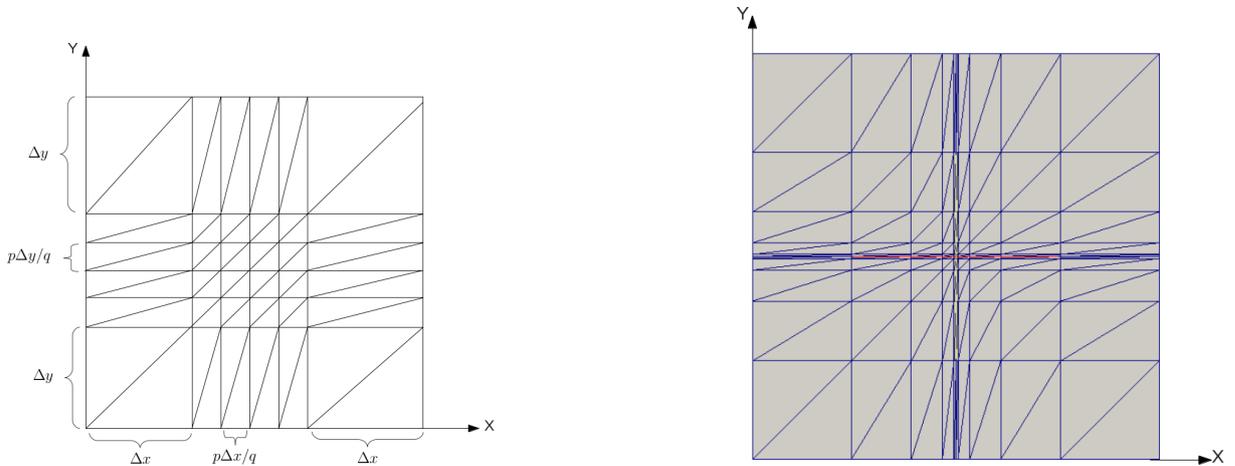


FIGURE 5.9 – Le domaine de calcul  $\Omega = [0, 1] \times [0, 1]$  avec une région raffinée au milieu (à gauche) et le maillage polynomial dans les directions  $X$  et  $Y$  ; exemple  $N_x = N_y = 10$  et  $\alpha = 0.01$  (à droite).

Pour chacun des trois maillages définis, chaque rectangle est découpé en deux triangles. Nous traçons ensuite la norme infinie de l'erreur en fonction des nombre des éléments dans l'une des deux directions, par exemple la direction  $x$ , dans l'échelle logarithmique. Notons que pour les trois tests numériques, les nombres des éléments dans les deux directions  $X$  et  $Y$  sont égaux. Les études numériques de la convergence du schéma ARK2 sont présentées dans la Fig. 5.10 à gauche pour le maillage Mesh 5, dans la Fig. 5.10 à droite pour le maillage Mesh 6 et dans la Fig. 5.11 pour le maillage polynomial.

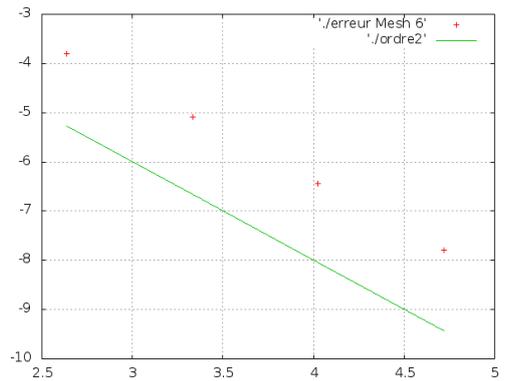
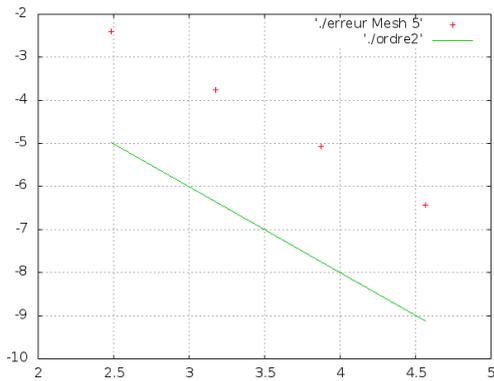


FIGURE 5.10 – La norme infinie de l’erreur dans le cas du maillage en 2D : Mesh 5 à gauche et Mesh 6 à droite.

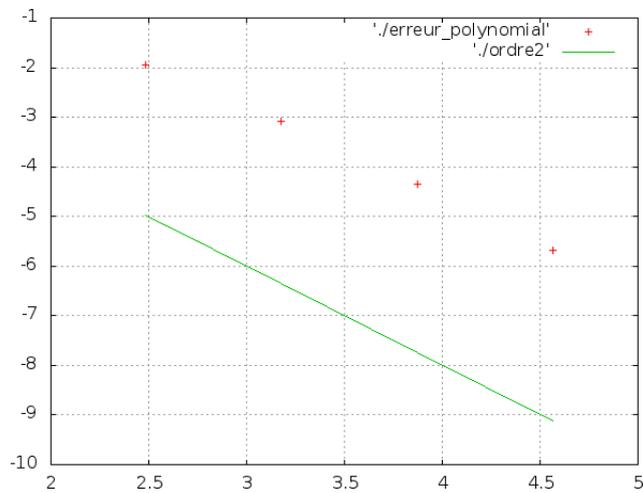


FIGURE 5.11 – La norme infinie de l’erreur dans le cas du maillage polynomial en 2D.

**Remarque 5.** *Les tests numériques en dimension un et en dimension deux montrent que le schéma ARK2 est convergent d’ordre deux indépendamment du maillage choisi, ce qui valide les résultats théoriques prouvés dans la section précédente.*

## 5.4. Conclusion

Dans ce chapitre, nous avons dérivé un schéma RK2 asynchrone du schéma Runge-Kutta classique. Le schéma AKR2 permet la selection de pas de temps indépendants pour chaque élément du maillage. Nous avons présenté l’algorithme de la méthode ARK2 pour une équation différentielle linéaire quelconque donnée. Nous avons imposé la linéarité de l’équation traitée pour des raisons de démonstration du résultat. La méthode asynchrone est susceptible de se révéler également utile pour les systèmes non-linéaires. Nous avons ensuite supposé une discrétisation spatiale arbitraire. À chaque pas de temps, seulement une cellule est concernée et seulement les voisins ont besoin d’être mis à jour, ce qui rend la méthode facilement applicable à tout type de maillage.

Nous avons donné une analyse approfondie de la convergence à l'ordre deux de notre schéma asynchrone. Nous avons prouvé en particulier que le schéma ARK2 préserve l'ordre de convergence du schéma classique indépendamment du maillage choisi. De plus, nous avons étudié l'asynchrone dans le cas du maillage défini dans [35] et nous avons prouvé que dans ce cas particulier, le schéma asynchrone dégénère vers le schéma LTS-RK2. Pour illustrer les résultats théoriques, nous avons choisi d'appliquer la méthode aux équations de Maxwell, d'abord en dimension un d'espace puis en dimension deux. Dans les deux cas, la discrétisation spatiale a été réalisée par la méthode de Galerkin discontinue upwind avec des éléments de base type  $P^1$ . L'étude théorique montre que le schéma ARK2 est convergent d'ordre deux indépendamment du maillage choisi, ce qui valide les résultats prouvés. En outre, les tests numériques indiquent que l'utilisation de la condition CFL optimale imposée sur chaque élément indépendamment des autres éléments, n'a aucune influence sur la précision et l'efficacité du schéma classique. Enfin, les résultats montrent que, en comparaison avec l'intégration classique, le schéma asynchrone est efficace en termes du gain en temps de calcul.



## Chapitre 6

# Mise en oeuvre de la méthode asynchrone d'ordre deux : application pour les équations de Maxwell en dimensions deux et trois

Durant les chapitres précédents nous nous sommes particulièrement concentrés sur le développement des schémas asynchrones, l'étude de leur ordre de convergence et l'illustration des résultats théoriques avec des tests académiques relativement simples. L'objectif de ce chapitre est la mise en oeuvre de la méthode ARK2. En particulier, une simulation de cas-test représentatifs de maillages multi-échelles, cadre naturel de l'application de l'asynchrone, sera présentée.

### 6.1. L'expérience numérique

Pour illustrer la mise en oeuvre du schéma asynchrone, nous allons appliquer la méthode à la discrétisation des équations de Maxwell en dimensions deux et trois d'espace. Dans tous les tests numériques présentés dans ce chapitre, la méthode de Galerkin discontinue upwind avec des éléments de base de types  $P^1$  sera utilisée pour la discrétisation spatiale d'ordre deux.

L'expérience numérique consistera à injecter, par l'un des cotés du domaine d'étude  $\Omega$ , une onde plane définie par :

$$u(x, t) = p_1 \sin(2\pi \times p_2 \times (t - x/c)), \quad (6.1)$$

où  $p_1$  et  $p_2$  sont respectivement l'amplitude et la fréquence de l'onde et  $c = 299792458 \text{ m/s} \simeq 3 \times 10^8 \text{ m/s}$  est la vitesse de la lumière.  $\lambda = \frac{c}{p_2}$  est la longueur de l'onde injectée et  $T$  est le temps final de la simulation. On supposera par la suite que  $p_1 = 1 \text{ m}$  et  $p_2 = 10^9$  hertz. La direction de l'onde sera imposée selon la direction  $(Ox)$ . Le domaine d'étude est à la base un carré en dimension deux, et un parallélépipède (ou un cube) en dimension trois.

Pour faire apparaître l'aspect multi-échelles dans le maillage, nous ajoutons un "obstacle" au milieu du domaine, la nature de celui-ci donnera la quantification de l'aspect multi-échelle ainsi que le gain attendu

à l'utilisation du schéma asynchrone par rapport au schéma synchrone. En effet, dans la section 6.2., nous allons étudier l'exemple du fil en deux dimensions. Il s'agit de la représentation bidimensionnelle d'un fil 3D infini dans la direction ( $Oz$ ). La coupe en 2D du fil sera présentée selon la forme de la section du fil, dans un premier temps par un cercle de rayon  $r$ , puis par un carré de diagonale  $d$ .

Ensuite, dans la section 6.3., le fil en trois dimensions sera présenté. Il s'agit du même cas d'étude avec les deux formes de section étudiées en 2D. Nous présentons d'abord le cas d'un fil cylindrique, puis celui d'un fil parallélépipédique. Du point de vue du maillage, c'est en fait une extrusion des tests de la section 6.2..

Enfin, la section 6.4. sera consacrée à l'étude de la performance de la méthode asynchrone en termes de temps de calcul dans le cas des maillages très fortement multi-échelles. L'objectif ici est de mettre en évidence des cas naturels pour lesquels le speed-up théorique apporté par l'asynchrone est très important. Nous présenterons trois exemples : celui de la pièce, de la sphère et enfin le cas d'un fil présentant un petit défaut.

Pour les différents domaines d'étude, nous allons suivre l'évolution de l'onde plane injectée en fonction du temps dans les cas synchrone et asynchrone. En fait, la solution classique (comprendre celle obtenue avec le solveur MACOPA en imposant des pas de temps identiques pour toutes les cellules) sera notre référence pour contrôler la précision de l'asynchrone. Cela permettra également de fournir une comparaison des speed-up théorique et numérique obtenus sur les différents cas-tests.

Le speed-up théorique est calculé en utilisant la formule (2.15) donnée dans le chapitre 2. C'est une estimation a priori du gain en temps de calcul en comparaison avec le schéma synchrone. Cette estimation est volontairement indépendante du schéma asynchrone utilisé (le schéma asynchrone d'ordre un ou deux), elle est basée uniquement sur le maillage et la CFL associée. Ce speed-up est calculé dès la génération du maillage, selon le choix des phénomènes physiques modélisés et la définition de la condition initiale. Notons qu'au cours de la simulation, le speed up théorique peut évoluer légèrement (à la hausse comme à la baisse) par rapport à ce qui a été prévu du fait du recalcul de la condition CFL avec l'évolution des quantités physiques (ici la valeur du champ électrique).

Une fois les maillages générés pour les cas étudiés dans ce chapitre, nous nous sommes basés sur le speed-up théorique calculé en fonction du maillage avec la formule 2.15 pour estimer le gain potentiel de la méthode asynchrone. Signalons enfin que le speed-up numérique que nous présentons est calculé en se basant sur les simulations menées avec la plate-forme MACOPA, destinée à la base pour faire des calculs asynchrones. Comme nous l'avons remarqué dans les chapitres précédents, le temps de calcul pour le schéma synchrone n'est pas optimal quand on utilise MACOPA, le gain observé pourrait être moindre par rapport à un autre solveur optimisé. Toutefois, l'objectif ici étant de mettre en évidence le fait que le passage de l'intégration temporelle asynchrone de l'ordre un à l'ordre deux n'altère pas les gains de performance de la méthode, nous gardons le code MACOPA synchrone comme référence.

Les domaines d'étude et les maillages utilisés dans ce chapitre ont été construits en utilisant le logiciel GMSH.

## 6.2. Exemple du fil en deux dimensions

Les équations de Maxwell en dimensions deux de l'espace en mode transverse électrique (TE) s'écrivent sous la forme suivante :

$$\begin{cases} \frac{\partial}{\partial t} E_x &= \frac{1}{\epsilon} \frac{\partial}{\partial y} H_z, \\ \frac{\partial}{\partial t} E_y &= -\frac{1}{\epsilon} \frac{\partial}{\partial x} H_z, \\ \frac{\partial}{\partial t} H_z &= \frac{1}{\mu} \frac{\partial}{\partial y} E_x - \frac{1}{\mu} \frac{\partial}{\partial x} E_y, \end{cases}$$

où  $E = (E_x, E_y, 0)^T$  est le champ électrique,  $H = (0, 0, H_z)^T$  est le champ magnétique,  $\mu = 4\pi 10^{-7} \text{ h.m}^{-1}$  (henry par mètre) est la perméabilité et  $\epsilon = 1/(c^2\mu)$  est la permittivité avec  $\epsilon \times \mu = \frac{1}{c^2}$ , où  $c$  est la vitesse de la lumière. Le domaine d'étude  $\Omega$  est un carré.

Comme expliqué dans la section 6.1., l'expérience numérique consiste à injecter une onde plane par le coté gauche du carré. Le domaine d'étude  $\Omega$  est de dimension  $[0, \lambda] \times [0, \lambda]$ , où  $\lambda$  est la longueur de l'onde injectée. Dans notre cas  $\lambda$  est égal à 0.3m et  $T$  est le temps final de la simulation supposé égal à  $10^{-9}$ s. La projection du fil en deux dimensions est représentée d'abord par un cercle de rayon  $r < \lambda$ , puis par un carré de diagonale  $d < \lambda$ . Pour chacune des deux formes un maillage multi-échelle puis un maillage fortement multi-échelle sont appliqués. Le but est d'étudier le comportement asynchrone en fonction de la forme du fil mais aussi en fonction du caractère multi-échelle.

Le domaine  $\Omega$  contenant un cercle sera noté  $\Omega_1$  est le domaine  $\Omega$  contenant un carré sera noté  $\Omega_2$ .

Le but de cette section est de tout d'abord comparer les schémas synchrone et asynchrone d'ordre deux avec les schémas d'ordre un. Les équations de Maxwell ont donc été discrétisées dans un premier temps en utilisant un schéma d'ordre un. Puis, dans un deuxième temps, elles ont été discrétisées avec un schéma d'ordre deux.

1. Pour l'ordre un, le schéma Euler Explicite en temps avec les volumes finis (VF) en espace a été utilisé. En fait, dans les chapitres 2 et 3, nous avons prouvé que le schéma asynchrone Euler-VF est convergent d'ordre un et l'application a été faite pour les équations de transport.
2. Pour l'ordre deux, nous avons utilisé le schéma RK2 en temps avec une discrétisation GD- $P^1$  en espace. Nous avons prouvé que le schéma RK2 asynchrone développé dans le chapitre 5 est convergent d'ordre deux, il préserve donc l'ordre du schéma classique.

### 6.2.1. Premiers tests : maillages multi-échelles

La première série de tests numériques consiste à étudier l'évolution temporelle de l'onde dans le cas d'un maillage multi-échelle avec un raffinement autour du fil : le cercle pour  $\Omega_1$  et le carré pour  $\Omega_2$ .

#### 6.2.1.1. Le premier test pour le domaine $\Omega_1$

Pour le premier test nous supposons que le rayon du fil est  $r = 0.1 \times \lambda < \lambda = 0.3\text{m}$  (voir figure 6.1, à gauche). Le maillage raffiné autour du trou est représenté sur la figure 6.1 (à droite).

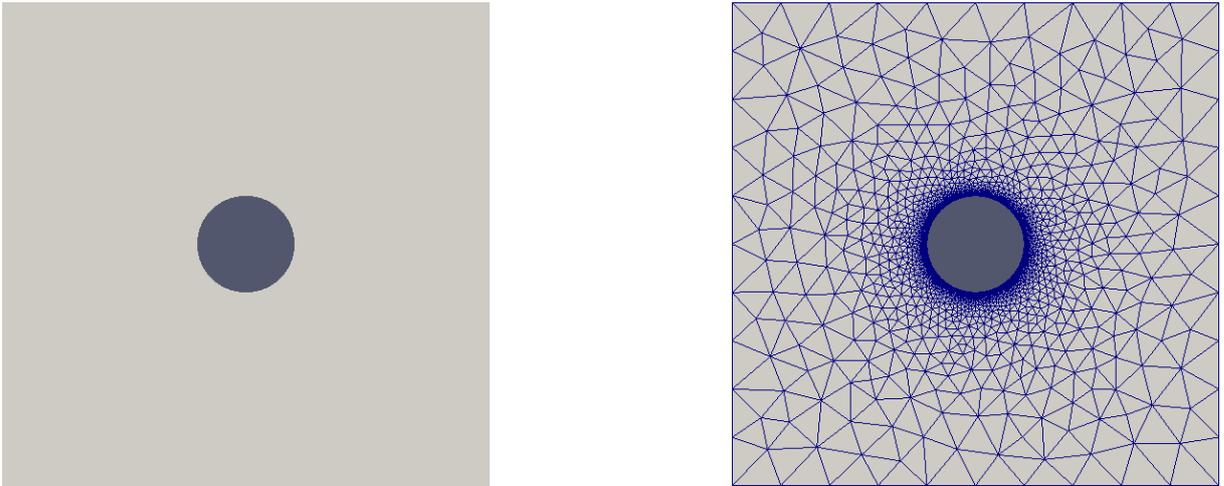


FIGURE 6.1 – À gauche : le domaine d'étude  $\Omega_1$  avec  $r = 0.1 \times \lambda$ . À droite : maillage raffiné autour du fil.

Le maillage a été généré par GMSH, nous résumons dans le tableau suivant (figure 6.2) quelques informations concernant le maillage construit.

Nombre de points	3848
Nombre de cellules	7024
L'aire de la plus grande cellule	$4,2 \cdot 10^{-4}$
L'aire de la plus petite cellule	$2,6 \cdot 10^{-8}$
Le speed up théorique	2 ,4

FIGURE 6.2 – Le maillage construit par GMSH pour le domaine  $\Omega_1$  avec  $r = 0.1 \times \lambda$ .

Nous avons injecté l'onde plane par le coté gauche du domaine, puis nous avons fait fonctionner le solveur MACOPA en mode synchrone classique (pour l'ordre un et l'ordre deux) et en mode asynchrone (également pour l'ordre un et l'ordre 2). Nous représentons l'évolution temporelle de l'onde pour les deux modes. Les solutions sont calculées aux instants :  $t^1 = 0.4 \times T$  un temps juste avant que l'onde ne touche le fil (voir figure 6.3),  $t^2 = 0.7 \times T$  un temps pour lequel l'onde traverse le fil (voir figure 6.4) et  $t^3 = T$  le temps final de la simulation (voir figure 6.5).

La comparaison des solutions présentées dans les figures 6.3- 6.5, montre que nous avons les résultats suivants :

1. La solution du schéma d'ordre deux améliore la qualité de la solution du schéma d'ordre un que ce soit pour l'intégration synchrone ou asynchrone.
2. Pour l'intégration d'ordre un, la solution asynchrone est meilleure que la solution classique ce

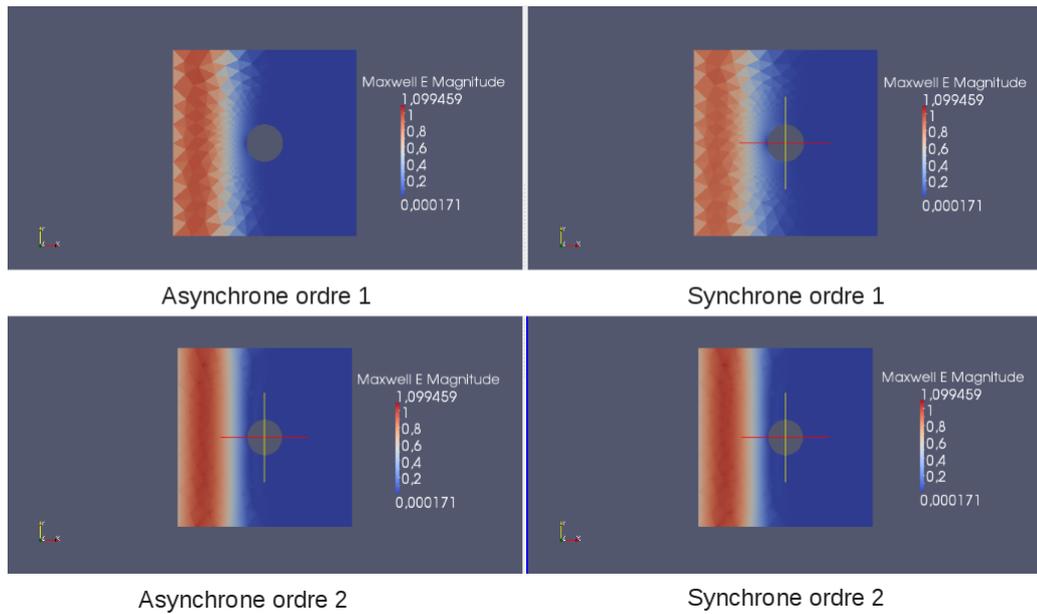


FIGURE 6.3 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_1$  avec le maillage multi-échelle à l’instant  $t = 0.4 \times T$ .

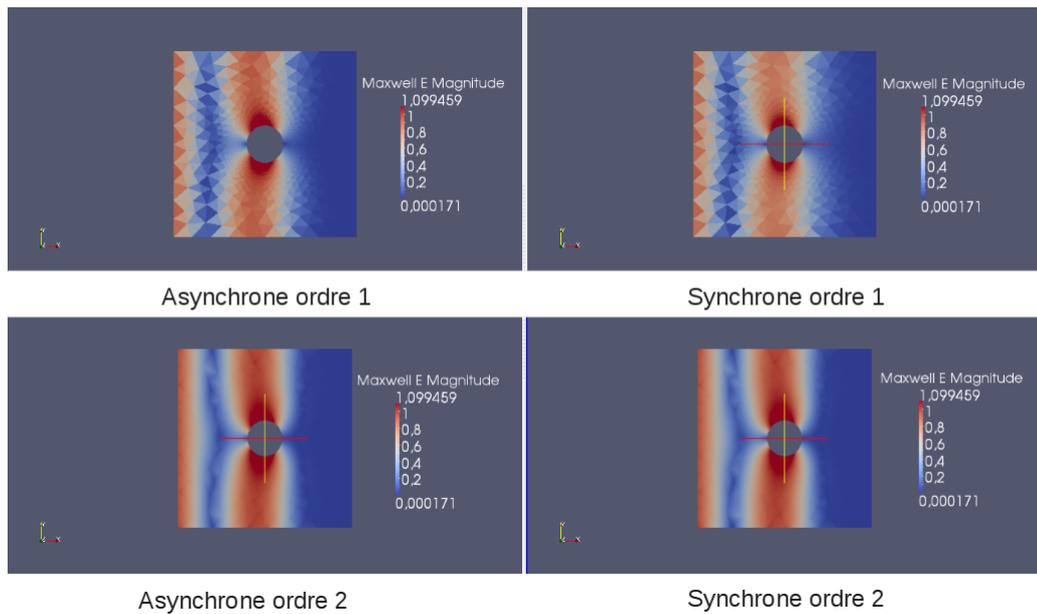


FIGURE 6.4 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_1$  avec le maillage multi-échelle à l’instant  $t = 0.7 \times T$ .

qui est cohérent avec les résultats du chapitre 2. En effet, dans le chapitre 2, les tests numériques montrent que l’intégration asynchrone d’ordre un est plus précise en comparaison avec l’intégration synchrone aussi d’ordre un. En particulier, l’utilisation d’une CFL locale diminue la dispersion numérique.

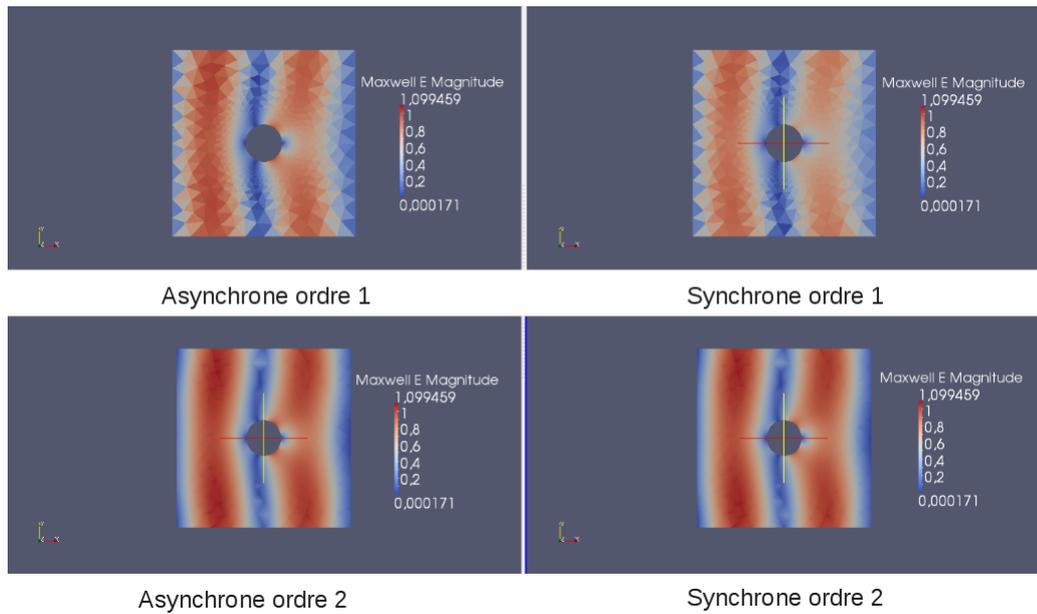


FIGURE 6.5 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_1$  avec le maillage multi-échelle à l’instant  $t = T$ .

3. L’intégration asynchrone d’ordre deux préserve la qualité de la solution classique ce qui confirme les résultats du chapitre 5.

Les temps de calcul pour les différents schémas sont présentés sur la figure 6.6. Notons que le temps de calcul pour l’ordre un est inférieur à celui de l’ordre deux pour les deux intégrations synchrone et asynchrone. Ce qui est cohérent car l’ordre deux nécessite plus d’opérations, avec en particulier le calcul des pentes pour le schéma en temps et le calcul matriciel en espace. Le speed up théorique ne varie pas avec le schéma, car il est calculé uniquement en fonction du maillage et indépendamment de l’intégration numérique. On observe, concernant le speed up numérique, qu’il est plus faible pour l’ordre deux en comparaison avec l’ordre un. Pour l’algorithme ARK2, des opérations supplémentaires ont été rajoutées pour pouvoir atteindre l’ordre deux ce qui coûte en temps de calcul. Toutefois, l’ordre de grandeur de ce speed-up est préservé au passage au second ordre en temps.

	Temps de calcul synchrone	Temps de calcul asynchrone	Speed-up numérique
Schéma Euler en temps et volumes finis en espace (ordre 1)	3 mn	50 s	3,6
Schéma RK2 en temps et GD- $P^1$ en espace (ordre 2)	20,27 mn	10 mn	2,02

FIGURE 6.6 – Temps de calcul pour le domaine  $\Omega_1$  avec le maillage multi-échelles à l’instant final  $T$ .

### 6.2.1.2. Le premier test pour le domaine $\Omega_2$

Nous reprenons le même test que dans la section précédente mais cette fois avec le domaine  $\Omega_2$  où la section du fil est représentée par un carré avec une demi-diagonale  $\frac{d}{2}$  égale à  $0.1 \times \lambda$  (voir figure 6.7, à gauche). Le maillage multi-échelles est donné par la figure 6.7 (à droite).

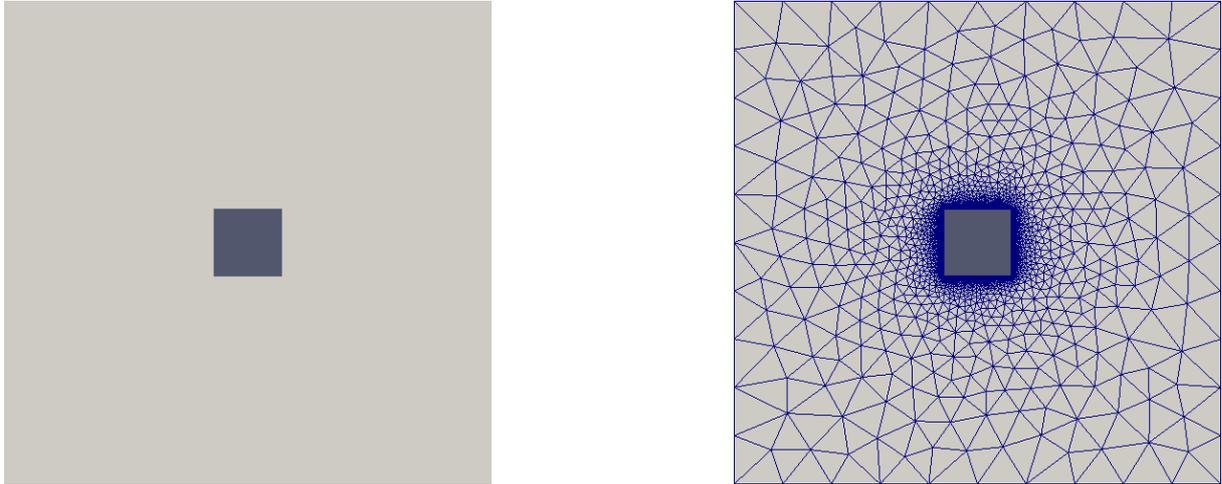


FIGURE 6.7 – À gauche : domaine carré contenant un trou. À droite : maillage raffiné autour du trou.

Comme pour le test précédent, le maillage a été géré par GMSH et le tableau suivant (figure 6.8) résume quelques informations concernant le maillage utilisé.

Nombre de points	3619
Nombre de cellules	6638
L'aire de la plus grande cellule	$4,9 \cdot 10^{-4}$
L'aire de la plus petite cellule	$2,7 \cdot 10^{-8}$
Le speed up théorique	2 ,35

FIGURE 6.8 – Le maillage construit par GMSH.

Les figures 6.9-6.11 représentent l'évolution temporelle de l'onde pour les différents schémas asynchrone/asynchrone d'ordre un ou deux.

La comparaison des différents schémas montre que l'ordre deux améliore la qualité de la solution d'ordre un pour les deux intégrations synchrone et asynchrone. De plus, le schéma asynchrone d'ordre deux préserve la qualité de la solution classique d'ordre deux. Pour les deux cas (ordre un et ordre deux), le schéma asynchrone est efficace en termes de temps de calcul. En effet, le speed-up numérique est égal à 3,5 pour l'intégration d'ordre un et 1,92 pour l'intégration d'ordre deux ; voir aussi Fig. 6.12.

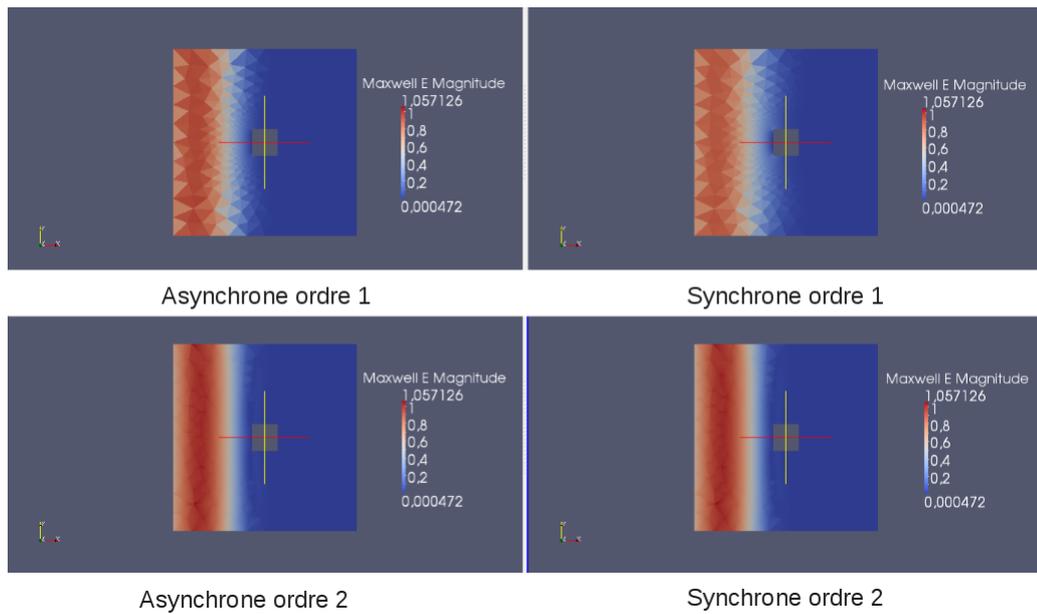


FIGURE 6.9 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_2$  avec le maillage multi-échelles à l’instant  $t = 0.4 \times T$ .

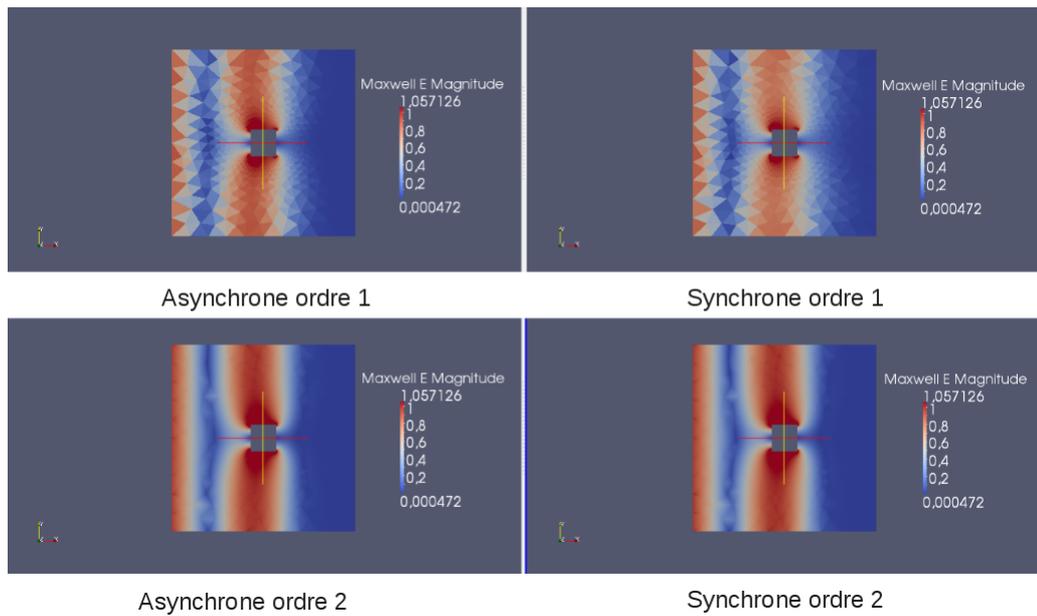


FIGURE 6.10 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_2$  avec le maillage multi-échelles à l’instant  $t = 0.7 \times T$ .

Le speed-up théorique estimé pour le maillage utilisé est égal à 2,35 (voir figure 6.8). Le fait que le speed-up numérique puisse être supérieur au speed-up théorique provient du fait que la valeur théorique utilisée est celle de la condition initiale. En pratique, le speed-up théorique associé à la contrainte CFL augmente quand l’onde commence à traverser le domaine.

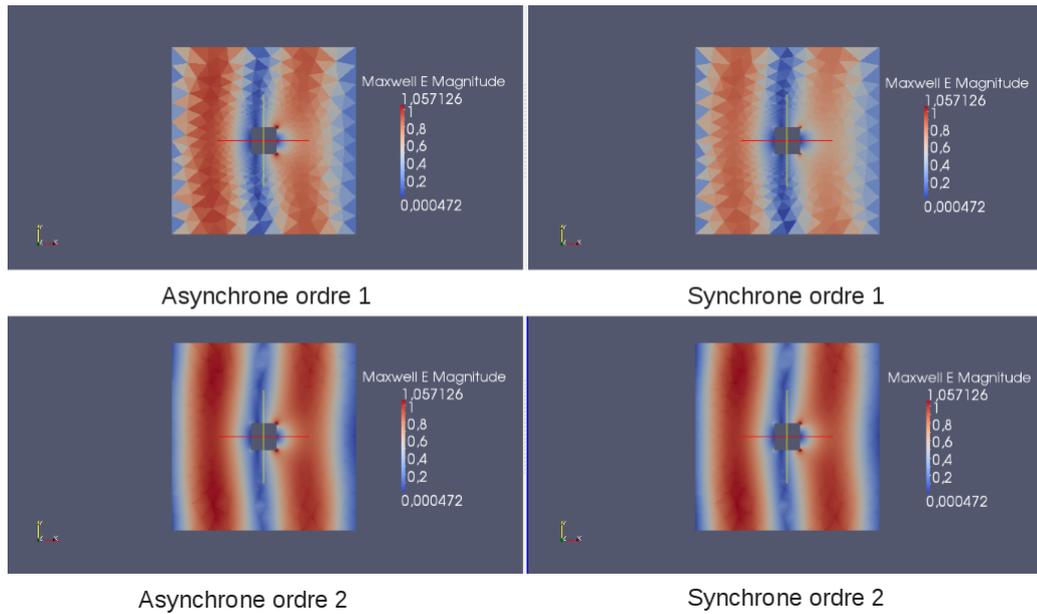


FIGURE 6.11 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_2$  avec le maillage multi-échelles à l’instant  $t = T$ .

	Temps de calcul synchrone	Temps de calcul asynchrone	Speed-up numérique
Schéma Euler en temps et volumes finis en espace (ordre 1)	2,8 mn	47 s	3,5
Schéma RK2 en temps et GD-P <sup>1</sup> en espace (ordre 2)	19,71 mn	10,23 mn	1,92

FIGURE 6.12 – Temps de calcul pour le domaine  $\Omega_2$  avec le maillage multi-échelles à l’instant final  $T$ .

## 6.2.2. Tests avec des maillages fortement multi-échelles

La deuxième série de tests numériques consiste à reprendre les mêmes tests précédents mais avec des maillages fortement multi-échelles. Autrement dit,  $\Delta t_{min} \ll \Delta t_{max}$  où  $\Delta t_{min}$  est le plus petit pas de temps et  $\Delta t_{max}$  est le plus grand pas de temps. En fait, les maillages ont été construits en imposant un raffinement local qui permet de trouver un speed-up théorique relativement grand (c’est à dire  $\gg 1$ ). Le but est d’étudier le comportement de l’asynchrone en fonction du caractère multi-échelles.

### 6.2.2.1. Le deuxième test pour le domaine $\Omega_1$

Pour ce deuxième exemple nous reprenons le premier exemple pour le domaine  $\Omega_1$  avec cette fois un rayon de valeur  $r = 0.001 \times \lambda \ll \lambda = 0.3\text{m}$  (voir figure 6.13, à gauche et la figure 6.14, à gauche). Le maillage est représenté sur les figure 6.13 (à droite) et 6.14 (à droite). Ce maillage permet d’atteindre un speed-up théorique de l’ordre de 18. Les informations concernant le maillage utilisé sont présentées

dans la figure 6.15.

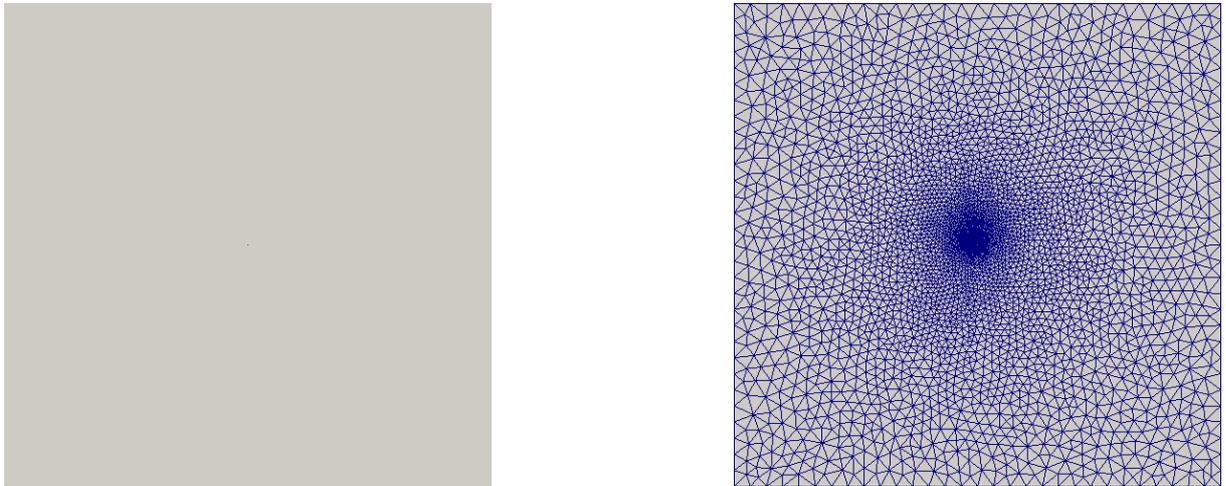


FIGURE 6.13 – Le domaine  $\Omega_1$  avec le rayon du trou  $r = 0.001 \times \lambda$  à gauche et le maillage à droite.



FIGURE 6.14 – Le maillage fortement multi-échelles autour du fil de rayon  $r = 0.001 \times \lambda$ .

Nombre de points	4278
Nombre de cellules	8416
L'aire de la plus grande cellule	$6,5 \cdot 10^{-5}$
L'aire de la plus petite cellule	$2,5 \cdot 10^{-9}$
Le speed up théorique	18,1

FIGURE 6.15 – Quelques informations concernant le maillage fortement multi-échelles du domaine  $\Omega_1$ .

L'évolution temporelle de l'onde pour les différents schémas est présentée à l'instant  $t = 0.4 \times T$  sur la figure 6.16, à l'instant  $t = 0.7 \times T$  sur la figure 6.17 et à l'instant  $t = T$  sur la figure 6.18. Un zoom autour du fil à l'instant  $t = 0.7 \times T$ , le temps du passage de l'onde à travers le fil, est présenté dans sur

la figure 6.19.

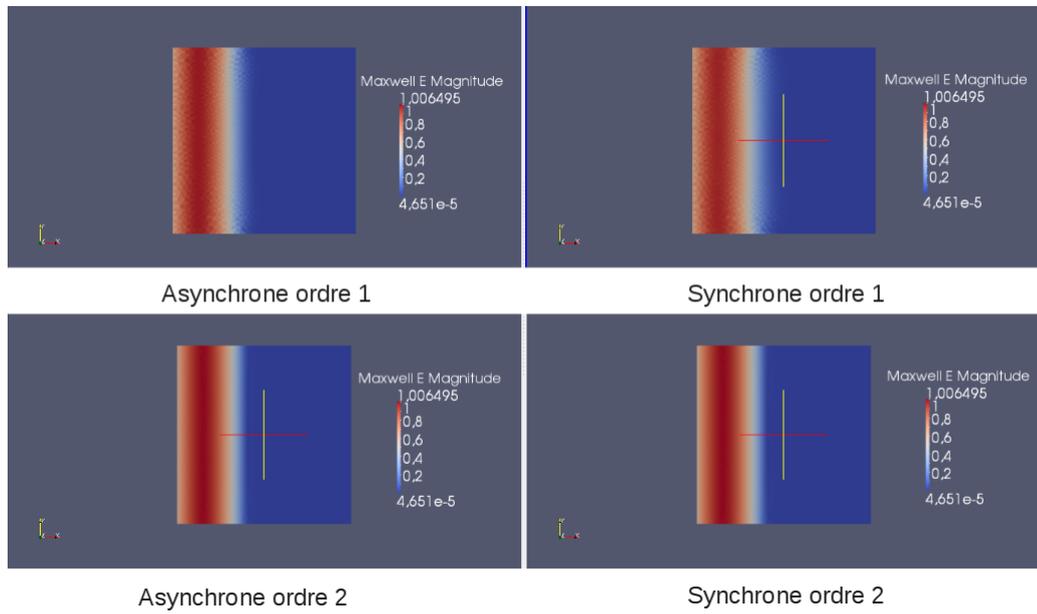


FIGURE 6.16 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_1$  avec le maillage fortement multi-échelles à l’instant  $t = 0.4 \times T$ .

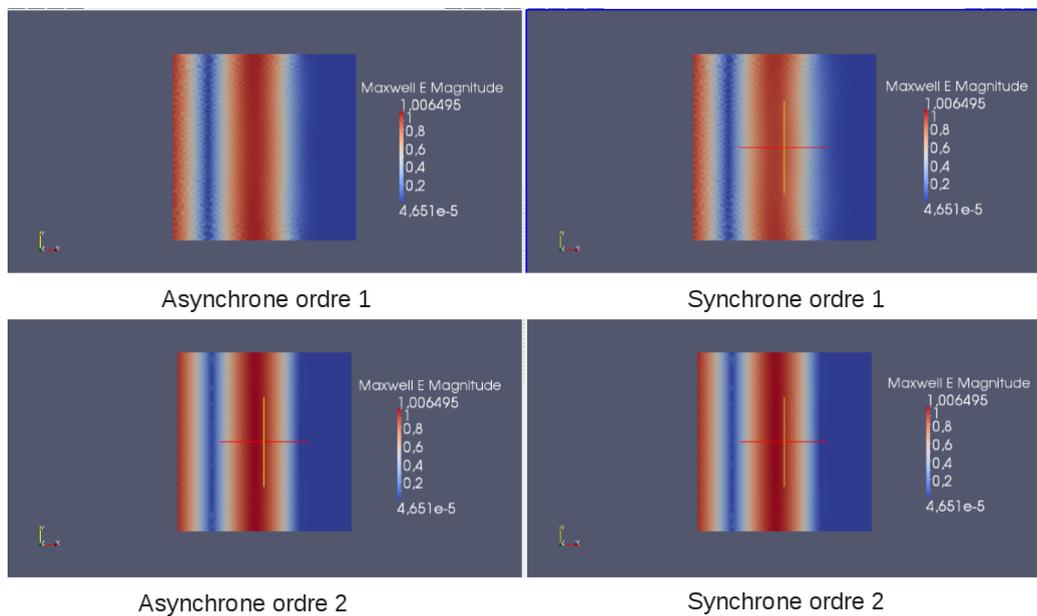


FIGURE 6.17 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_1$  avec le maillage fortement multi-échelles à l’instant  $t = 0.7 \times T$ .

La comparaison des différents schémas montre de nouveau que l’ordre deux améliore la qualité des solutions d’ordre un et que le schéma asynchrone d’ordre deux préserve la qualité de la solution classique.

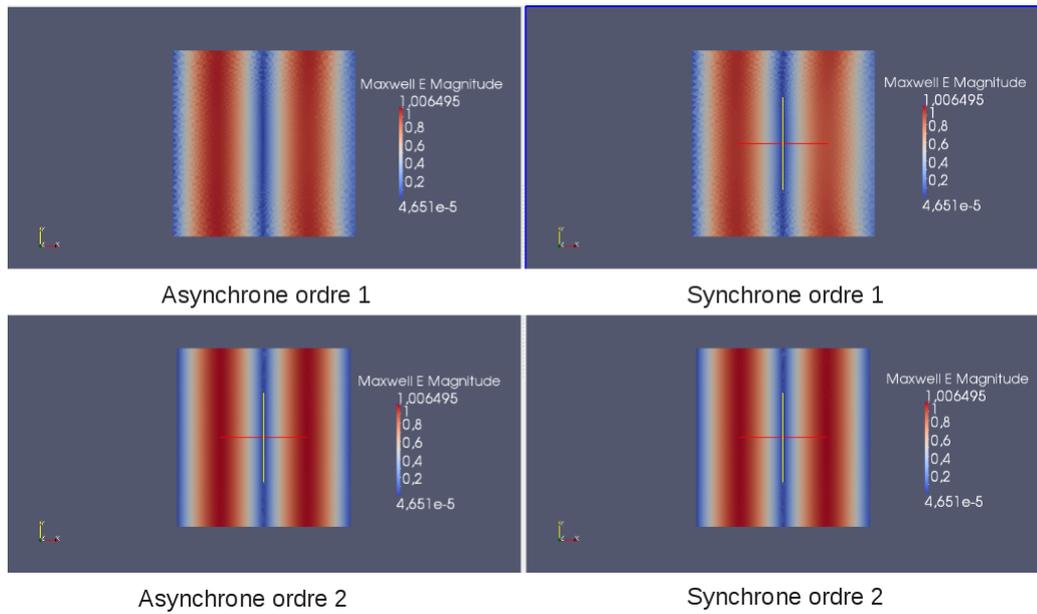


FIGURE 6.18 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_1$  avec le maillage fortement multi-échelles à l’instant  $t = T$ .

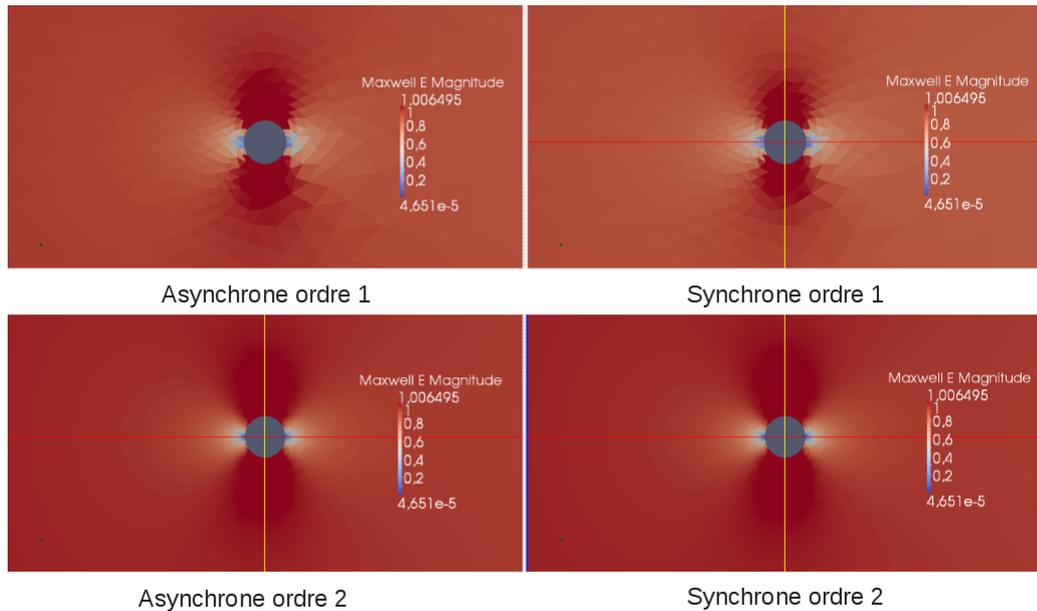


FIGURE 6.19 – Le passage de l’onde à travers le fil (un zoom sur le fil).

De plus, en termes de gain en temps de calcul, l’asynchrone est très efficace, voir sur la figure 6.20. Pour le schéma d’ordre deux, par exemple, pour faire tourner le code en mode synchrone nous avons eu besoin d’environ 1 heure et 40 minutes contre 5 minutes seulement pour l’asynchrone. Le speed-up numérique est donc égal à 17,8. Nous retrouvons donc le speed-up théorique prévu pour ce maillage avec une petite variation due encore une fois à l’évolution de la contrainte CFL au cours de la simulation.

	Temps de calcul synchrone	Temps de calcul asynchrone	Speed-up numérique
Schéma Euler en temps et volumes finis en espace (ordre 1)	9 mn	43 s	12,67
Schéma RK2 en temps et GD-P <sup>1</sup> en espace (ordre 2)	1H 40mn	5mn 40 s	17,8

FIGURE 6.20 – Temps de calcul pour le domaine  $\Omega_1$  avec le maillage fortement multi-échelles à l’instant final  $T$ .

### 6.2.2.2. Le deuxième test pour le domaine $\Omega_2$

Pour ce deuxième exemple pour le domaine  $\Omega_2$ , nous reprenons le premier exemple avec cette fois  $d/2 = 0.001 \times \lambda \ll \lambda$ . Puis nous imposons un maillage fortement multi-échelle autour du fil de sorte que le speed-up théorique estimé est relativement grand. Pour l’exemple traité, le speed-up théorique est égal à 11.3. Le maillage est représenté sur la figure 6.21 (à droite) et le zoom du maillage autour du fil est donné sur la figure 6.22 (à droite). Quelques informations concernant le maillage sont regroupées dans la figure 6.23.

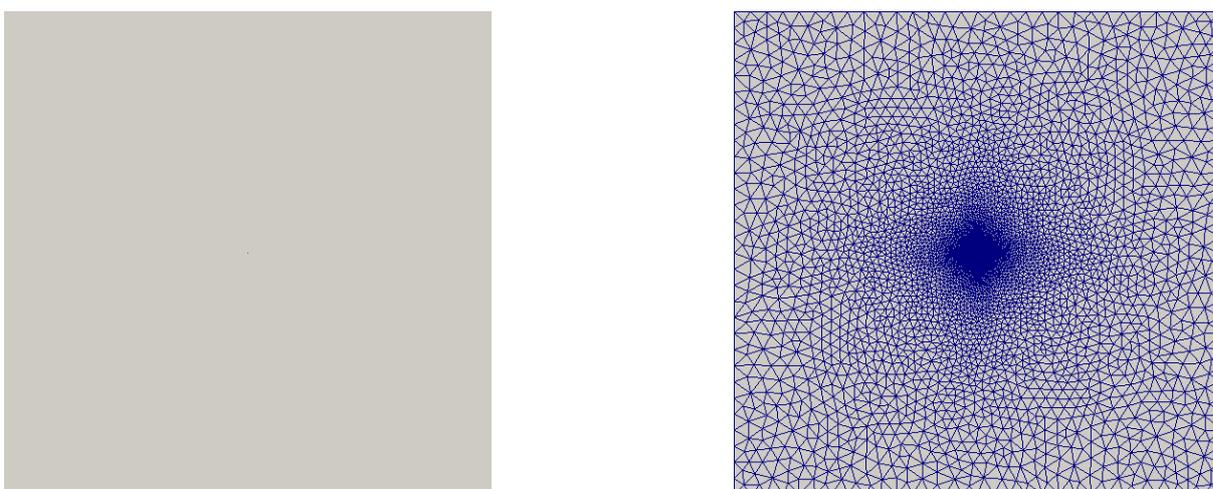


FIGURE 6.21 – Le domaine  $\Omega_2$  avec  $d = 0.001 \times \lambda$  à gauche et le maillage à droite.

Les solutions synchrone et asynchrone sont calculées aux instants  $0.4 \times T$ ,  $0.7 \times T$  et  $T$  et représentées dans les figures 6.24, 6.25 et 6.26. Un zoom sur le passage de l’onde à travers le fil à l’instant  $0.7 \times T$  est donné dans la figure 6.27.

**Remarque 6.** *La comparaison des solutions synchrones et asynchrones montre que l’aspect multi-échelle n’a pas d’influence sur la qualité de la solution classique. Les temps de calcul sont résumés dans la figure 6.28. Numériquement, l’asynchrone est 8 fois plus rapide que le synchrone pour le schéma d’ordre un et 11 fois plus rapide pour le schéma d’ordre deux.*

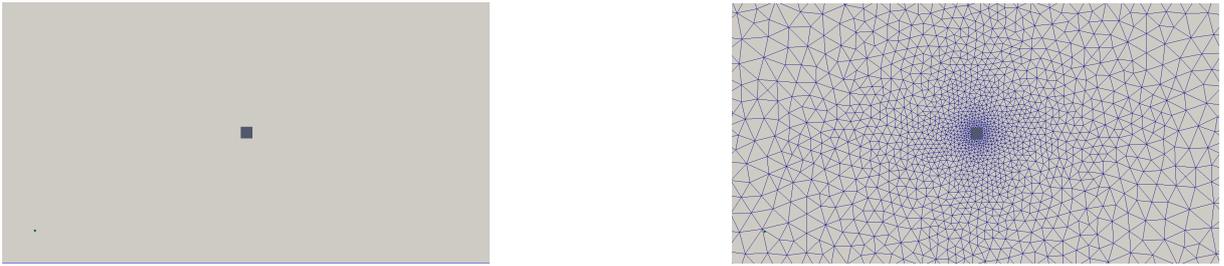


FIGURE 6.22 – Le maillage fortement multi-échelles autour du fil avec  $d = 0.001 \times \lambda$ .

Nombre de points	5675
Nombre de cellules	11210
L'aire de la plus grande cellule	$5,6 \cdot 10^{-5}$
L'aire de la plus petite cellule	$2,2 \cdot 10^{-9}$
Le speed up théorique	11,3

FIGURE 6.23 – Quelques informations concernant le maillage fortement multi-échelles du domaine  $\Omega_2$ .

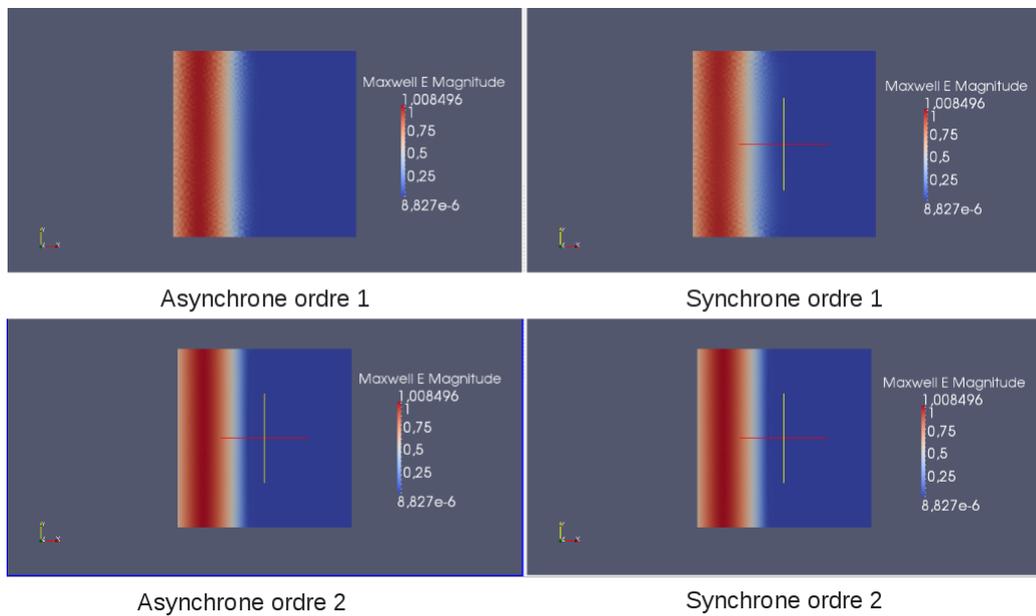


FIGURE 6.24 – Les solutions synchrones et asynchrones d'ordre un et d'ordre deux pour le domaine  $\Omega_2$  avec le maillage multi-échelles à l'instant  $t = 0.4 \times T$ .

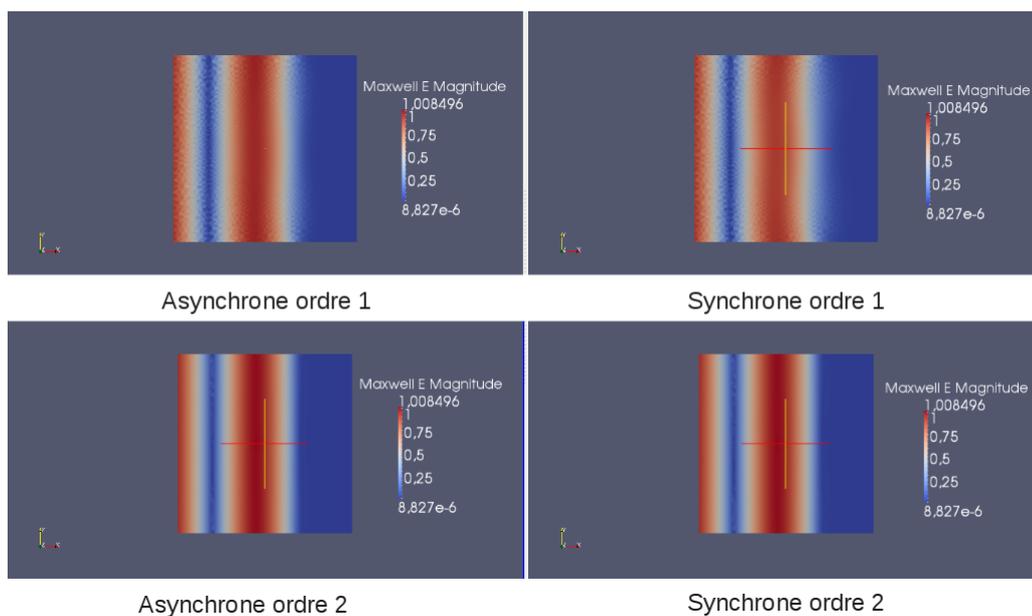


FIGURE 6.25 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_2$  avec le maillage multi-échelles à l’instant  $t = 0.7 \times T$ .

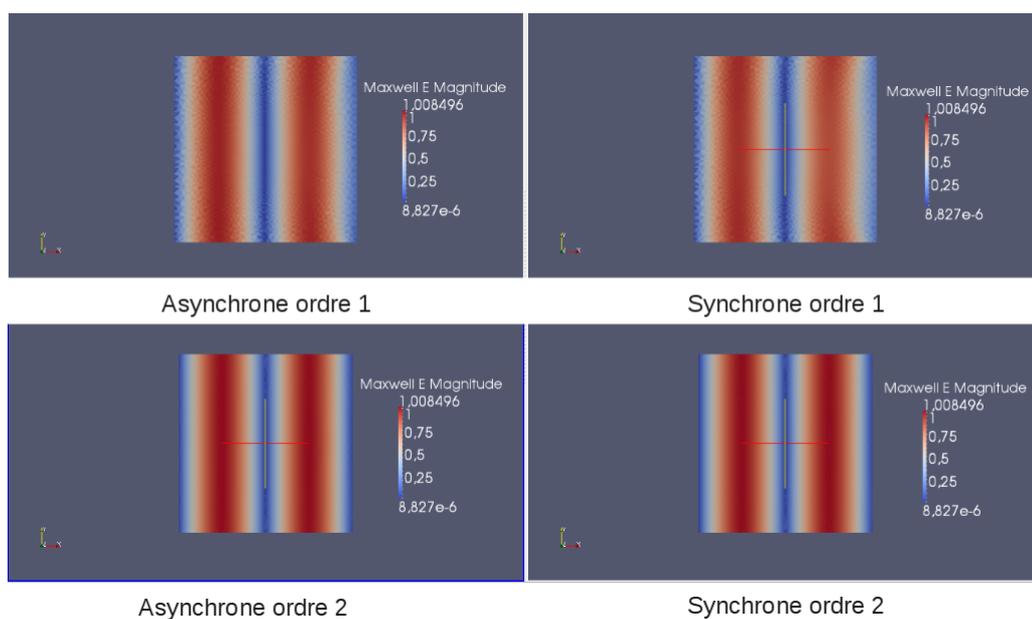


FIGURE 6.26 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_2$  avec le maillage multi-échelles à l’instant  $t = T$ .

### 6.3. Exemple du fil en trois dimensions

Le système des équations de Maxwell en trois dimensions est donné par :

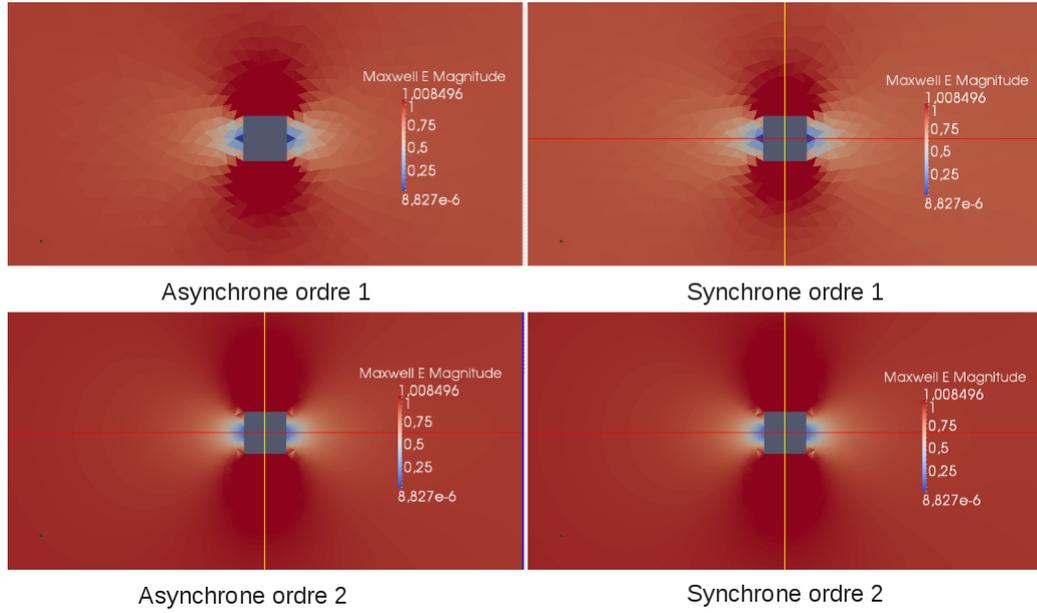


FIGURE 6.27 – Le passage de l’onde à travers le fil (un zoom sur le fil).

	Temps de calcul synchrone	Temps de calcul asynchrone	Speed-up numérique
Schéma Euler en temps et volumes finis en espace (ordre 1)	13 mn	1mn 30 s	8,5
Schéma RK2 en temps et GD-P <sup>1</sup> en espace (ordre 2)	2H	12 mn	11

FIGURE 6.28 – Temps de calcul pour le domaine  $\Omega_2$  avec le maillage fortement multi-échelles à l’instant final  $T$ .

$$\begin{cases} \varepsilon \frac{\partial E}{\partial t} - \text{rot}(H) = 0, \\ \mu \frac{\partial H}{\partial t} + \text{rot}(E) = 0, \end{cases} \quad (6.2)$$

où  $E = (E_x, E_y, E_z)^T$  est le champ électrique,  $H = (H_x, H_y, H_z)^T$  est le champ magnétique,  $\varepsilon$  est la permittivité électrique et  $\mu$  est la perméabilité magnétique.

Cette section est une extension de la première section. En particulier, le domaine d’étude  $\Omega$  est un parallélépipède de dimension  $dx \times dy \times dz$ . Nous ajoutons un fil qui traverse le domaine. D’abord un fil cylindrique, puis un fil parallélépipédique. De nouveau, une onde plane est injectée par la face gauche du domaine et sa direction est imposée selon la direction  $(Ox)$ . L’extrusion par rapport au test du fil en dimension deux est faite selon l’axe  $(Oy)$ .

Le domaine d’étude  $\Omega$  est de dimension  $dx = dz = \lambda$  et  $dy = 4\lambda$  où  $\lambda$  est la longueur de l’onde injectée et  $T = 10^{-9}$ s est le temps final de la simulation. Le domaine  $\Omega$  traversé par un cylindre de rayon

$r$  sera noté  $\Omega_3$  et le domaine  $\Omega$  traversé par un parallélépipède de diagonale  $d$  sera noté  $\Omega_4$ .

### 6.3.1. Le fil cylindrique

La première étape pour le test du fil cylindrique consiste à construire le domaine par le logiciel GMSH de sorte que le domaine  $\Omega_3$  soit une extrusion du domaine  $\Omega_1$ . Le rayon du cylindre  $r = 0.05 \times \lambda$ . L'extrusion est représentée sur les figures 6.29 (vue de face), 6.30 (à gauche, vue de droite) et 6.30 (à droite, vue de dessus).



FIGURE 6.29 – Le domaine d'étude  $\Omega_3$  : vue de face.

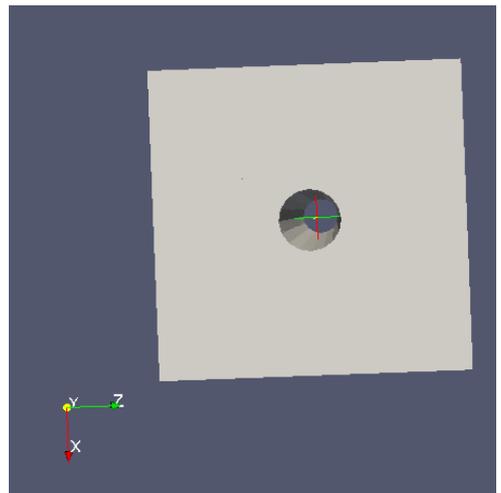
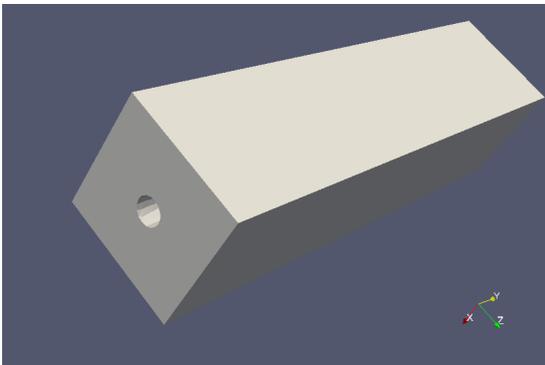


FIGURE 6.30 – Le domaine  $\Omega_3$  : vue de droite (à gauche) et vue de dessus (à droite).

La deuxième étape est de définir le maillage en imposant un raffinement local tout au long du fil. Le maillage est représenté sur les figures 6.31 (à gauche : une vue de droite) et 6.31 (à droite : une vue de dessus). Quelques informations concernant le maillage du domaine  $\Omega_3$  sont résumées sur la figure 6.32.

L'onde plane est injectée par la face gauche du parallélépipède  $\Omega_3$ . Les solutions synchrones/asynchrones d'ordre un/deux sont calculées aux instants :  $t^1 = 0.4 \times T$  un temps juste avant que l'onde touche le fil (figure 6.33),  $t^2 = 0.7 \times T$  un temps pour lequel l'onde traverse le fil (figure 6.33) et  $t^3 = T$  le temps final de la simulation(figure 6.33).

Les mêmes remarques que celles effectuées pour les calculs en deux dimensions peuvent être renouvelées pour ce calcul en trois dimensions. En fait, pour la comparaison des solutions, notons que le schéma

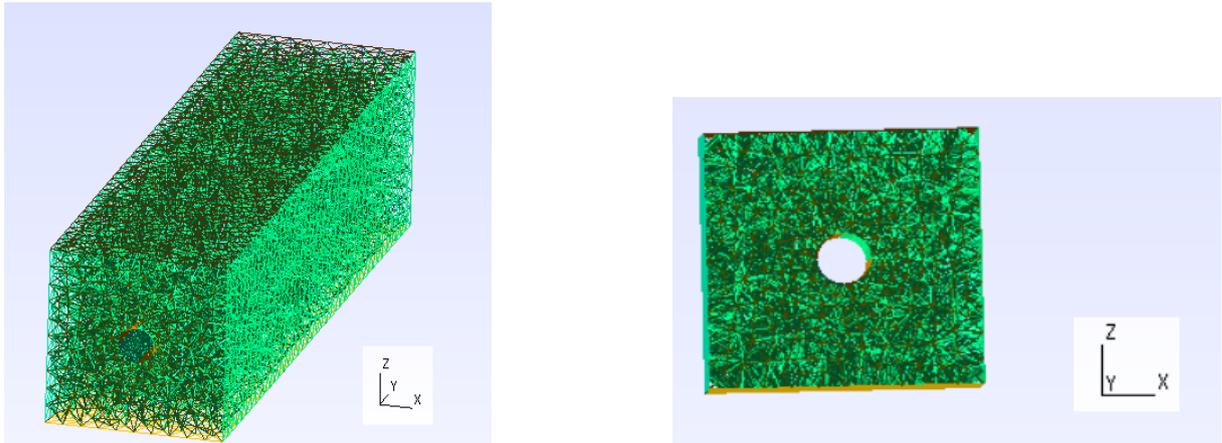


FIGURE 6.31 – Le maillage du domaine  $\Omega_3$  : vue de droite à gauche et vue de dessus à droite.

Nombre de points	7084
Nombre de cellules	30954
Volume de la plus grande cellule	$1,5 \cdot 10^{-5}$
Volume de la plus petite cellule	$9,3 \cdot 10^{-8}$
Le speed up théorique	4,55

FIGURE 6.32 – Quelques informations concernant le maillage du domaine  $\Omega_3$  construit par GMSH.

asynchrone d'ordre deux améliore la qualité de la solution asynchrone d'ordre un. De plus, l'asynchrone préserve la qualité de la solution classique. Pour le temps de calcul, on observe toujours que l'ordre un est plus rapide que l'ordre deux, pour l'asynchrone comme pour le synchrone. Ceci est lié à la complexité de l'ordre deux par rapport à l'ordre un. D'autre part, l'asynchrone est toujours efficace en comparaison avec le synchrone. Le schéma asynchrone d'ordre deux est six fois plus rapide que le schéma classique. Les temps de calcul pour les différents schémas sont présentés dans la figure 6.37.

**Remarque 7.** Pour comparer les résultats en trois dimensions avec les résultats en deux dimensions de la section précédente, une projection d'une coupe transversale du fil à l'instant  $t = 0.7 \times T$  est présentée sur la figure 6.36. Nous retrouvons donc le même comportement de l'onde qui traverse le fil (voir la figure 6.4).

### 6.3.2. Le fil parallélépipédique

Le domaine  $\Omega_4$  du fil parallélépipédique est une extrusion du domaine  $\Omega_2$  selon l'axe  $(Oy)$ . La vue de face du domaine  $\Omega_4$  est présentée sur la figure 6.38, la vue de droite sur la figure 6.39 (à gauche) et

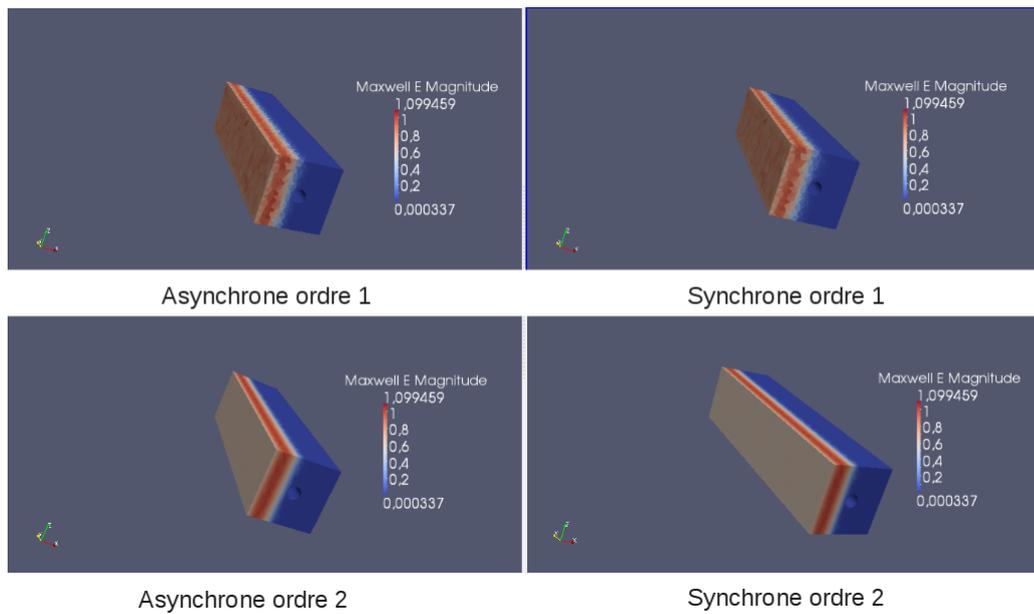


FIGURE 6.33 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_3$  à l’instant  $t = 0.4 \times T$ .

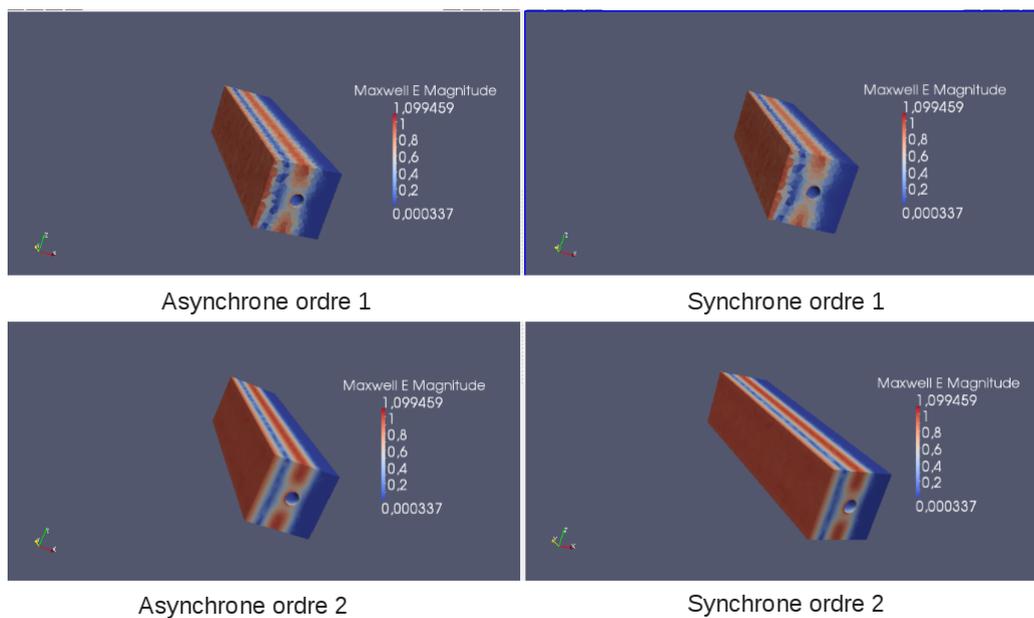


FIGURE 6.34 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_3$  à l’instant  $t = 0.7 \times T$ .

la vue de dessus sur la figure 6.39 (à droite). Le maillage géré par GMSH est donné dans la figure 6.40 avec une vue de droite (à gauche) et une vue de dessus (à droite). Un raffinement tout au long du fil est imposé. Les informations concernant le maillage du domaine  $\Omega_4$  sont rassemblées dans la figure 6.41.

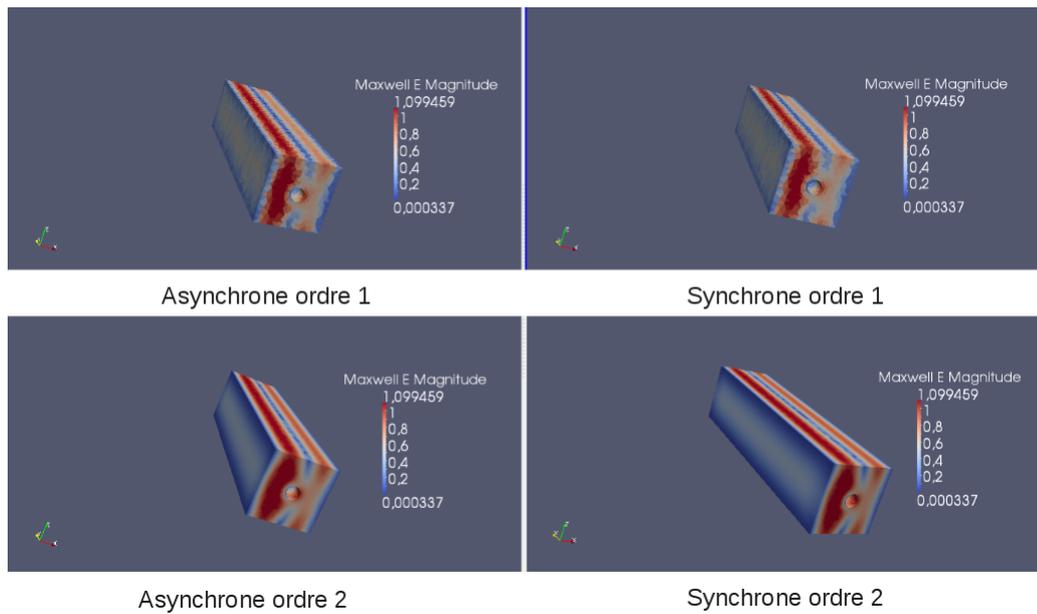


FIGURE 6.35 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_3$  à l’instant  $t = T$ .

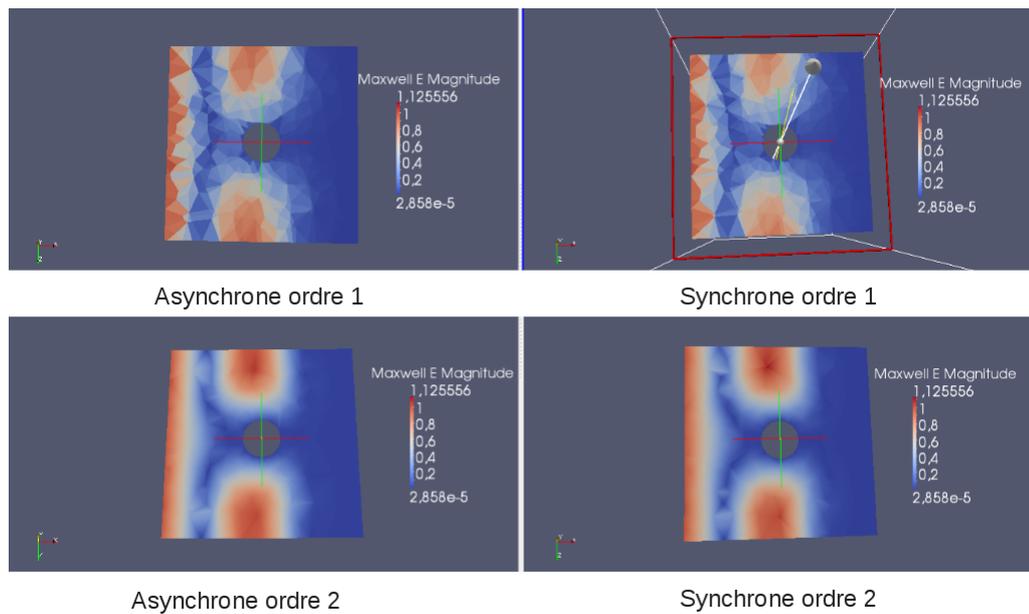


FIGURE 6.36 – Le passage de l’onde à travers le fil : une coupe transversale.

L’onde plane est injectée par la face gauche du parallélépipède  $\Omega_4$ . L’évolution de l’onde est représentée dans les figures 6.42-6.44.

Le raffinement imposé a créé un maillage multi-échelles avec un speed up théorique de l’ordre de cinq. Le speed up numérique est de l’ordre de trois pour le schéma d’ordre deux et de l’ordre de quatre pour le schéma d’ordre un. Les temps de calculs sont donnés sur la figure 6.46. Enfin, une projection d’une

	Temps de calcul synchrone	Temps de calcul asynchrone	Speed-up numérique
Schéma Euler en temps et volumes finis en espace (ordre 1)	1mn 20 s	21 s	3,8
Schéma RK2 en temps et GD-P <sup>1</sup> en espace (ordre 2)	26 mn	4 mn	6

FIGURE 6.37 – Temps de calcul pour le domaine  $\Omega_3$  à l'instant final  $T$ .

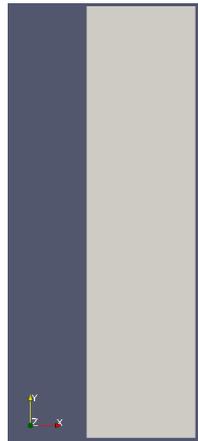


FIGURE 6.38 – Le domaine  $\Omega_4$  : vue de face.

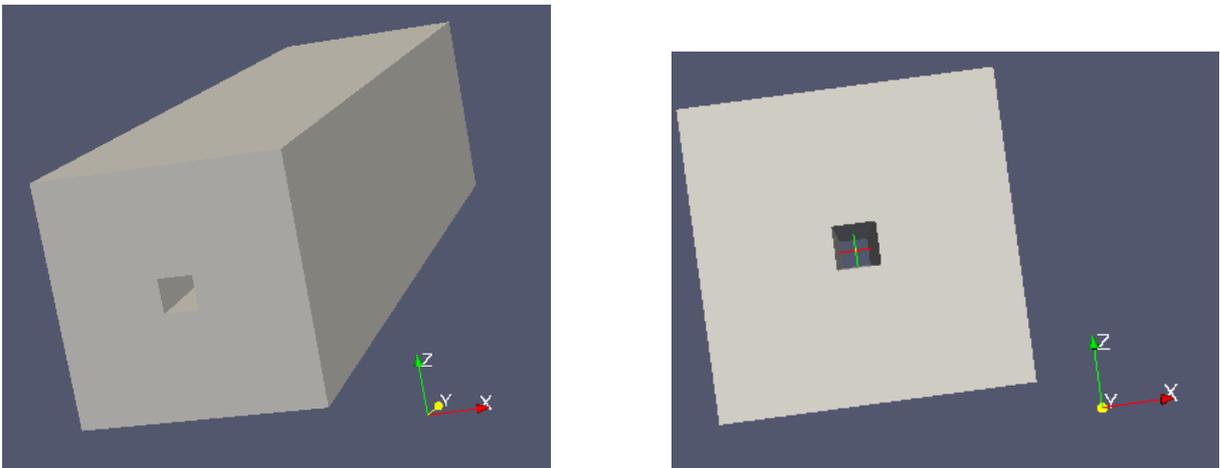


FIGURE 6.39 – Le domaine  $\Omega_4$  : vue de droite à gauche et vue de dessus à droite.

coupe transversale du fil à l'instant  $t = 0.7 \times T$  est présentée dans la figure 6.45, nous retrouvons donc les résultats de la dimension deux (voir figure 6.10).

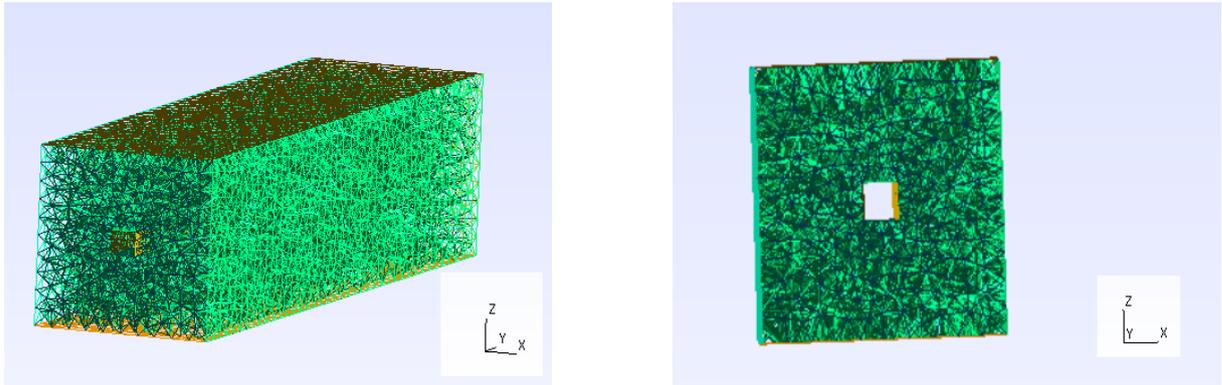


FIGURE 6.40 – Le maillage du domaine  $\Omega_4$  : vue de droite à gauche et vue de dessus à droite.

Nombre de points	7178
Nombre de cellules	32284
L'aire de la plus grande cellule	$1,3 \cdot 10^{-5}$
L'aire de la plus petite cellule	$1,6 \cdot 10^{-7}$
Le speed up théorique	5,4

FIGURE 6.41 – Quelques informations concernant le maillage construit par GMSH : exemple 2.

## 6.4. Performance du schéma asynchrone

Le schéma asynchrone a été construit dans le but de résoudre le problème du pas de temps global dans le cas des problèmes fortement multi-échelles. En fait, la forte inhomogénéité spatiale et temporelle des pas de discrétisation conduit à des temps de calcul gigantesques. Notre objectif était donc de pouvoir développer un schéma qui réduit le temps de calcul tout en préservant la précision du schéma classique. Nous étudions donc dans cette dernière section la performance en terme du gain en temps de calcul de notre schéma asynchrone dans des cas très fortement multi-échelles.

Pour les différents tests de cette section, l'aspect multi-échelles vient de la forte inhomogénéité du maillage. Cet aspect est contrôlé par le speed-up théorique calculé en utilisant de nouveau la formule (2.15). Trois exemples sont étudiés. D'abord une pièce à l'intérieur d'un cube avec un maillage où le speed-up théorique estimé est de l'ordre de 29. Puis un deuxième test est effectué avec une sphère à l'intérieur d'un cube avec un speed-up théorique égal à 113. Finalement, nous prenons l'exemple d'un fil cylindrique avec un défaut qui nous apporte un speed-up théorique égal à 612.

Rappelons que l'expérience numérique consiste à injecter une onde plane et suivre son évolution en fonction du temps. La solution classique est une référence pour contrôler le bon fonctionnement de l'asynchrone. Nous utilisons la méthode de Galerkin discontinue upwind avec les éléments  $P^1$  pour la

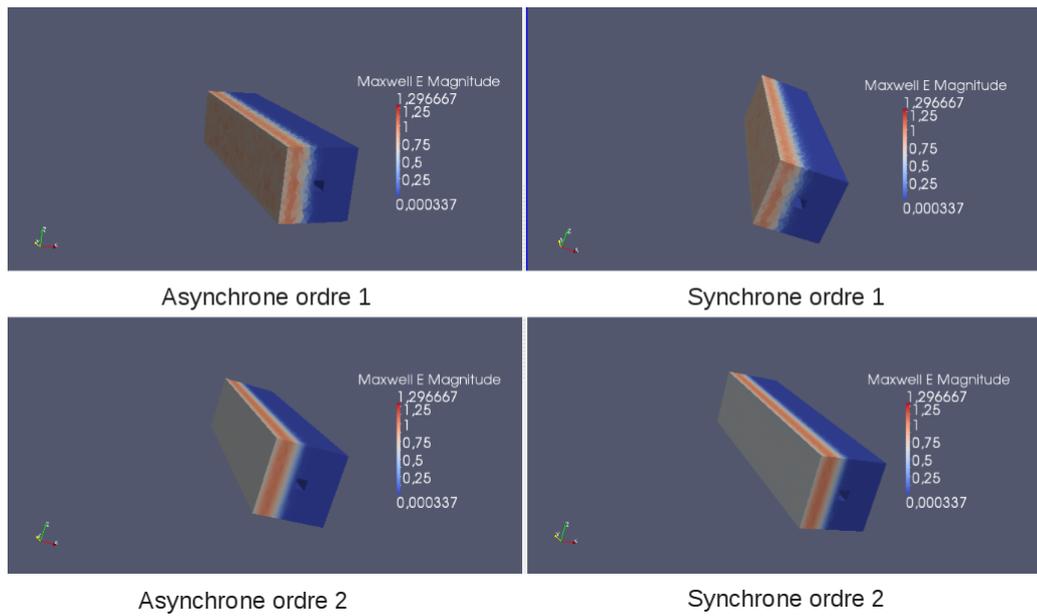


FIGURE 6.42 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_4$  à l’instant  $t = 0.4 \times T$ .

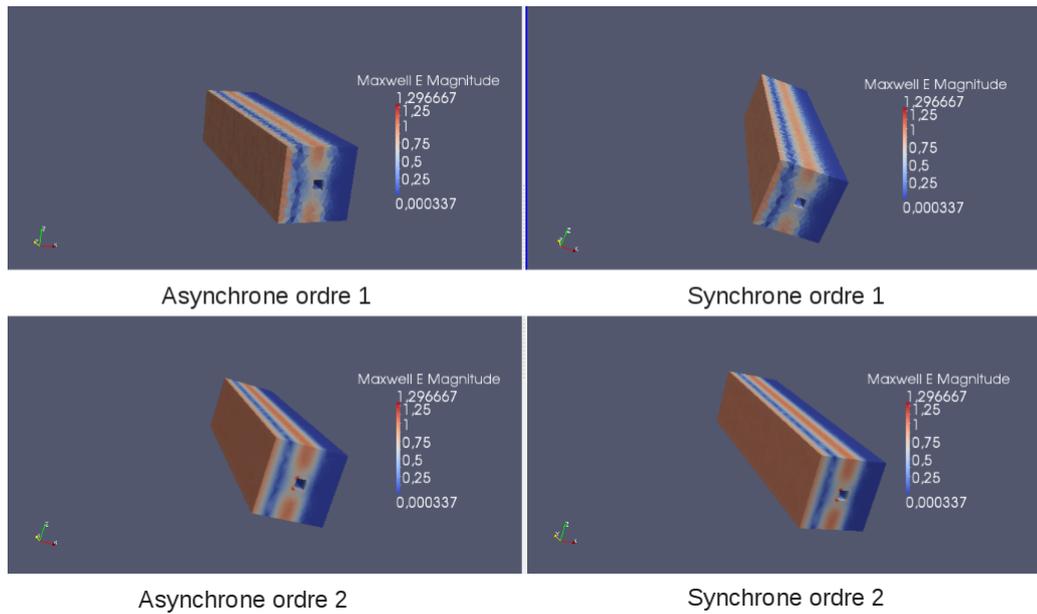


FIGURE 6.43 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_4$  à l’instant  $t = 0.7 \times T$ .

discrétisation spatiale des équations de Maxwell en trois dimensions. Pour la discrétisation temporelle nous appliquons la méthode ARK2 d’ordre deux développée dans le chapitre précédent et la méthode RK2 classique pour le synchrone.

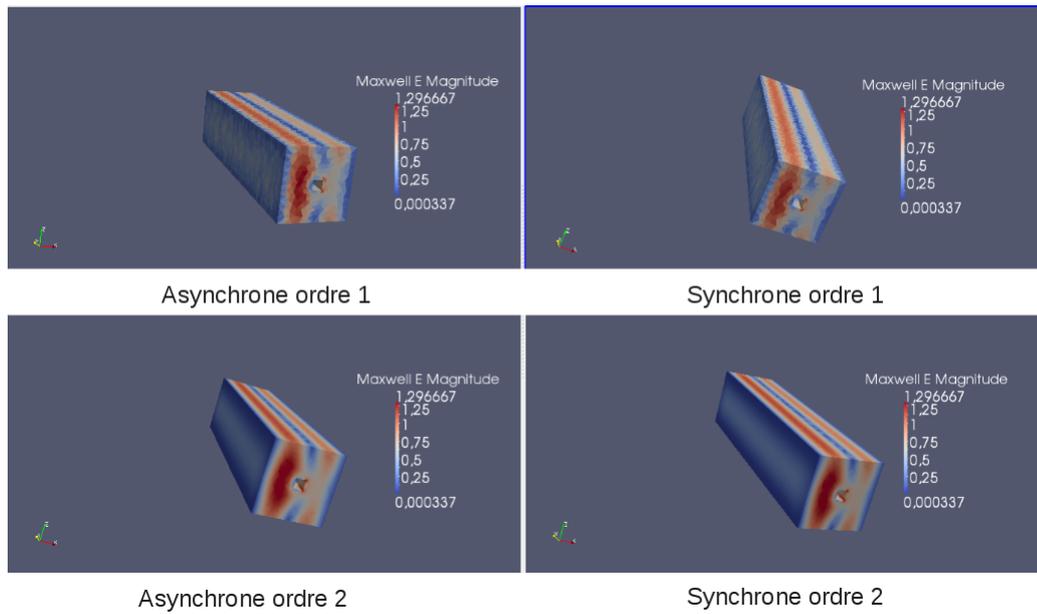


FIGURE 6.44 – Les solutions synchrones et asynchrones d’ordre un et d’ordre deux pour le domaine  $\Omega_4$  à l’instant  $t = T$ .

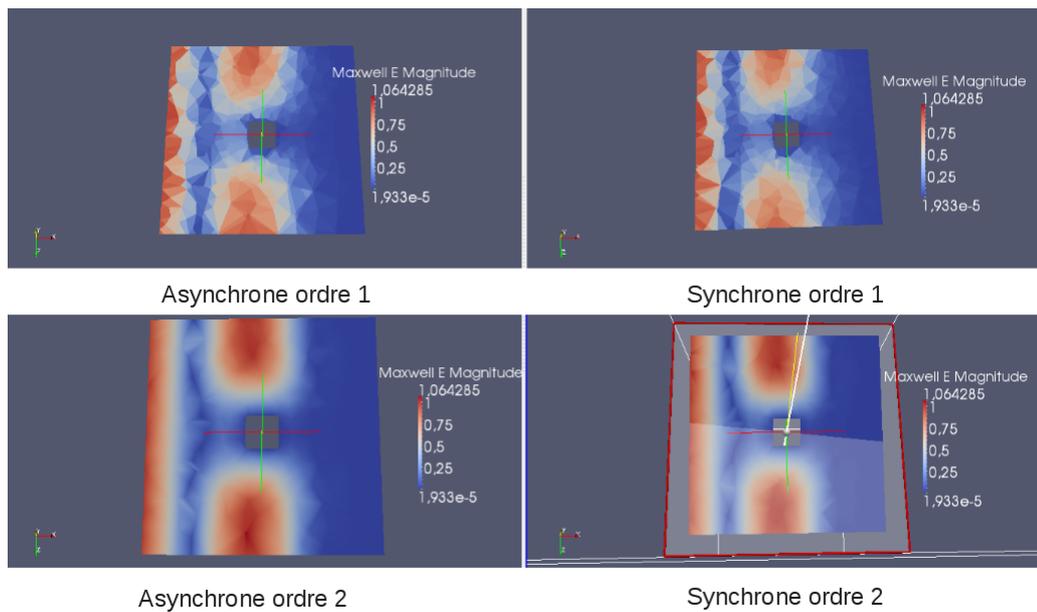


FIGURE 6.45 – Le passage de l’onde à travers le fil : une coupe transversale.

### 6.4.1. La pièce

Le domaine d’étude est un cube de dimensions  $\lambda^3$ , où  $\lambda$  est la longueur de l’onde injectée et le temps final  $T = 10^{-9}$  s. Pour faire apparaître le caractère multi-échelles, nous ajoutons une pièce à l’intérieur du cube, l’épaisseur de la pièce étant très faible devant son rayon, un forte inhomogénéité de maillage est générée. Pour atteindre un speed-up théorique égal à 29, nous avons une pièce dont le rayon de la

	Temps de calcul synchrone	Temps de calcul asynchrone	Speed-up numérique
Schéma Euler en temps et volumes finis en espace (ordre 1)	1mn 30 s	21 s	3,9
Schéma RK2 en temps et GD-P <sup>1</sup> en espace (ordre 2)	13 mn 36 s	4 mn 26 s	3

FIGURE 6.46 – Temps de calcul pour le domaine  $\Omega_4$  à l’instant final  $T$ .

base vaut  $r = 0.01 \times \lambda \ll \lambda$  et dont l’épaisseur est limitée à  $L = 0.001 \times \lambda \ll r$ . Nous avons ensuite généré un maillage qui devient fortement multi-échelles autour de la pièce. Le domaine d’étude est présenté sur la figure 6.47 et les informations concernant le maillage utilisé sont résumées dans la figure 6.48.

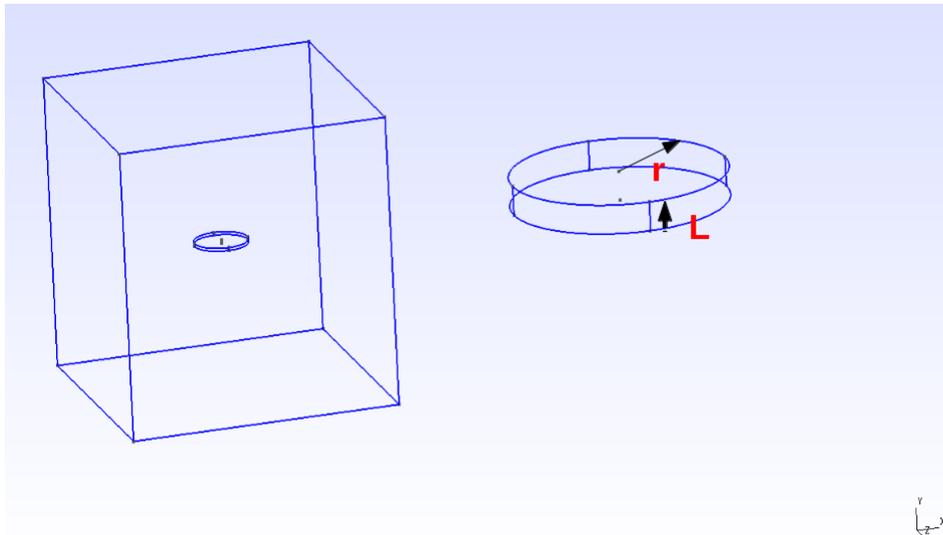


FIGURE 6.47 – Le domaine d’étude : une pièce à l’intérieur d’un cube.

L’évolution temporelle de l’onde pour les schémas asynchrone et synchrone est présentée à l’instant  $t = 0.4 \times T$  dans la figure 6.49, à l’instant  $t = 0.7 \times T$  dans la figure 6.50, et à l’instant  $t = T$  dans la figure 6.51. Notons que l’asynchrone préserve la qualité de la solution du schéma classique.

Pour suivre le passage de l’onde à travers la pièce, nous avons fait dans un premier temps une coupe longitudinale du domaine (voir figure 6.52). Les solutions asynchrone (à gauche) et synchrone (à droite) à l’instant  $t = 0.7 \times T$  sont présentées sur la figure 6.53 et un zoom sur la pièce est présenté sur la figure 6.54. Dans un deuxième temps, nous avons fait une coupe transversale (voir figure 6.55). Les solutions sont présentées sur la figure 6.56, un premier zoom est donné par la figure 6.57 puis un zoom sur la pièce par la figure 6.58.

Nombre de points	33738
Nombre de cellules	180294
L'aire de la plus grande cellule	$6,4 \cdot 10^{-7}$
L'aire de la plus petite cellule	$7,3 \cdot 10^{-12}$
Le speed up théorique	29,4

FIGURE 6.48 – Quelques informations concernant le maillage construit pour l'exemple de la pièce.

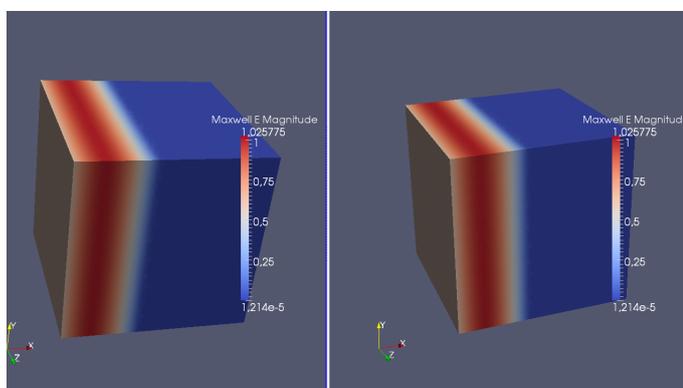


FIGURE 6.49 – Les solutions asynchrone (à gauche) et synchrone (à droite) à l'instant  $t = 0.4 \times T$ .

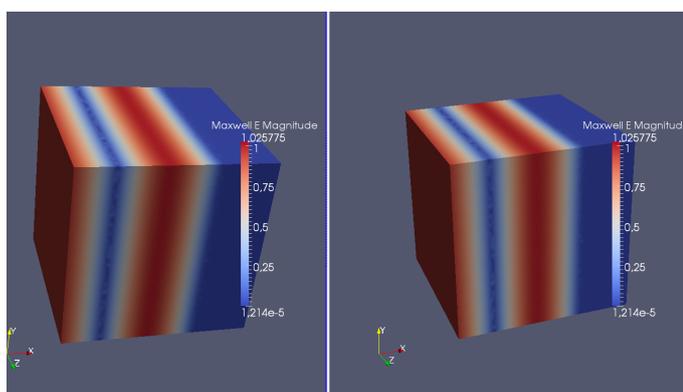


FIGURE 6.50 – Les solutions asynchrone (à gauche) et synchrone (à droite) à l'instant  $t = 0.7 \times T$ .

**Remarque 8.** Outre la conservation de la précision du schéma classique, le speed-up numérique pour l'exemple de la pièce est de l'ordre de 23. En fait, le temps de calcul pour le synchrone est égal à 1 jour et 21h alors que le temps de calcul pour le schéma asynchrone est égal à 2h. En comparaison avec le speed up théorique qui est égal à 29, le speed up numérique est un peu plus petit. En fait, le speed

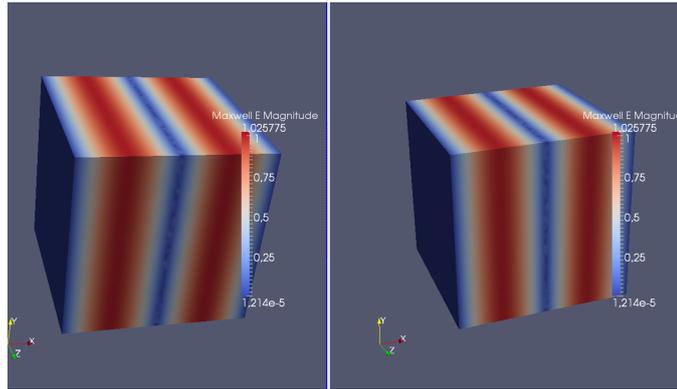


FIGURE 6.51 – Les solutions asynchrone (à gauche) et synchrone (à droite) à l’instant  $t = T$ .

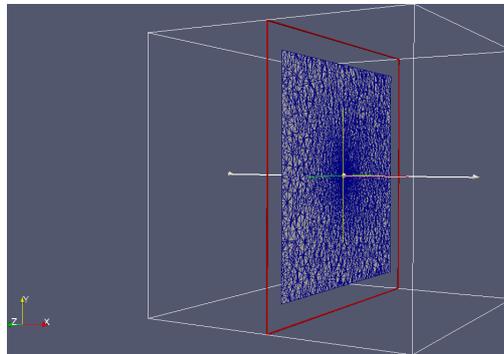


FIGURE 6.52 – Une coupe longitudinale du domaine d’étude.



FIGURE 6.53 – Le passage de l’onde à travers la pièce à l’instant  $t = 0.7 \times T$  : la solution asynchrone à gauche et la solution synchrone à droite.

*up théorique est le gain maximum qu’on peut atteindre indépendamment de l’implémentation et de la méthode numérique utilisée. L’algorithme asynchrone consiste à chercher, à chaque étape, la cellule la plus urgente à mettre à jour ce qui coûte en temps de calcul. Ce temps supplémentaire n’est pas pris en compte dans le calcul du speed up théorique ce qui peut donc expliquer la petite variation du speed up. L’implémentation des schémas synchrone et asynchrone peut aussi jouer un rôle dans la variation du speed-up.*

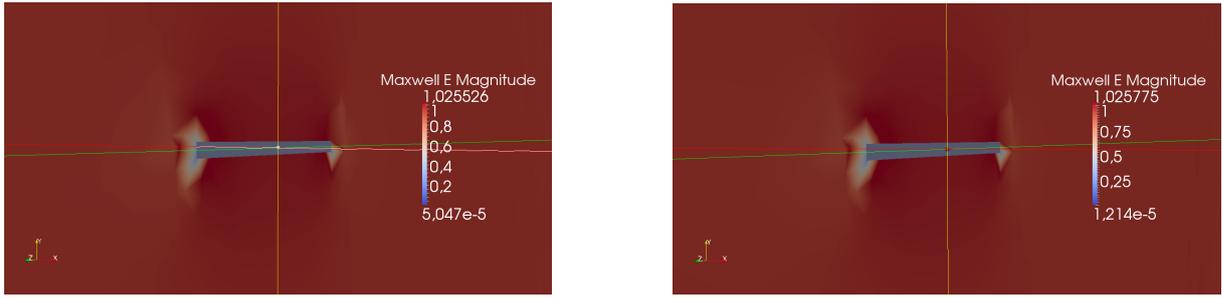


FIGURE 6.54 – Une coupe longitudinale de la pièce (zoom sur la pièce) à l’instant  $t = 0.7 \times T$  : la solution asynchrone à gauche et la solution synchrone à droite.

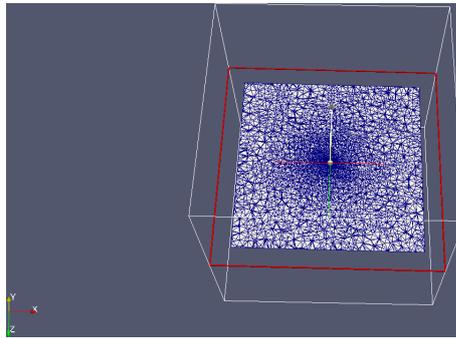


FIGURE 6.55 – Une coupe transversale du domaine d’étude.

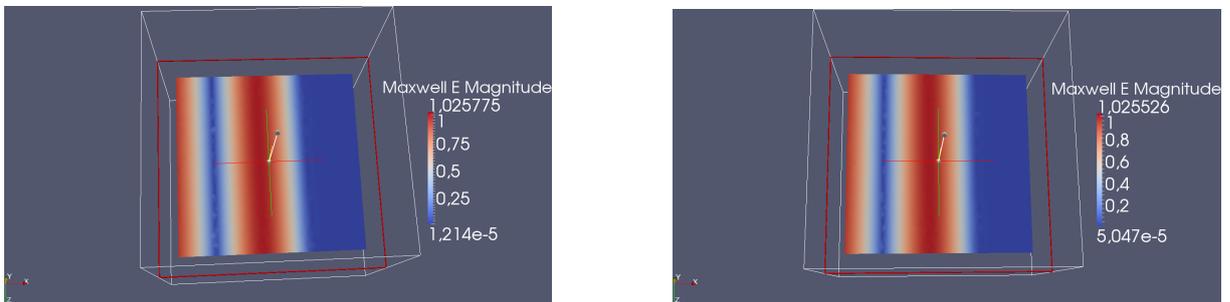


FIGURE 6.56 – Une coupe transversale de la pièce à l’instant  $t = 0.7 \times T$  : la solution asynchrone à gauche et la solution synchrone à droite.

#### 6.4.2. La sphère

Le domaine d’étude pour cet exemple est encore un cube de dimensions  $\lambda^3$ , où  $\lambda$  est la longueur de l’onde injectée et le temps final  $T = 10^{-9}$  s. Le but est de construire un domaine qui permet de créer un maillage fortement multi-échelles avec un speed-up encore plus important que celui de la pièce. Nous avons donc ajouté cette fois comme obstacle une sphère de rayon  $r = 0.001 \times \lambda \ll \lambda$ . Le maillage multi-échelles imposé a donné un speed-up théorique égal à 113. Le domaine d’étude est présenté sur la figure 6.59 et les informations concernant le maillage utilisé sont résumées dans la figure 6.60.

L’évolution temporelle de l’onde pour le schéma asynchrone et le schéma synchrone est présentée à

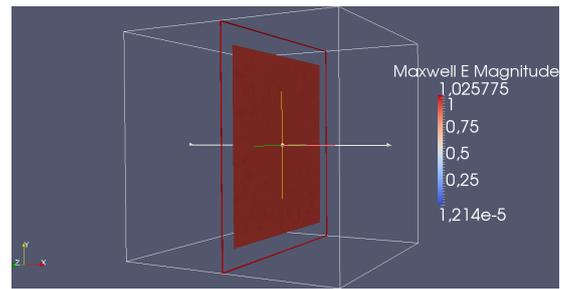
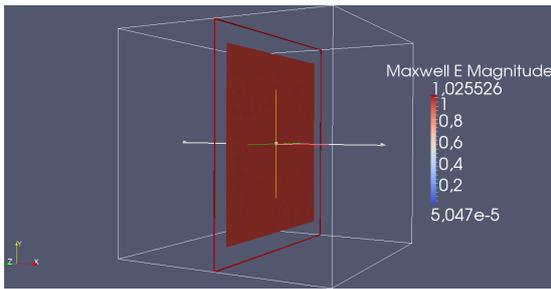


FIGURE 6.57 – Le passage de l'onde à travers la pièce à l'instant  $t = 0.7 \times T$  : la solution asynchrone à gauche et la solution synchrone à droite.

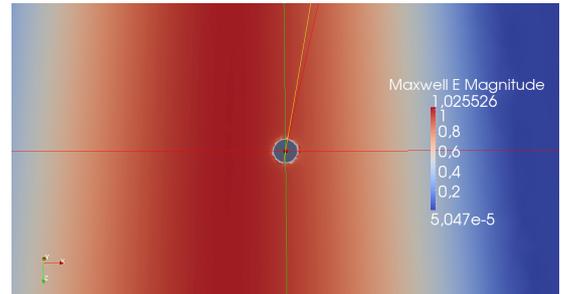
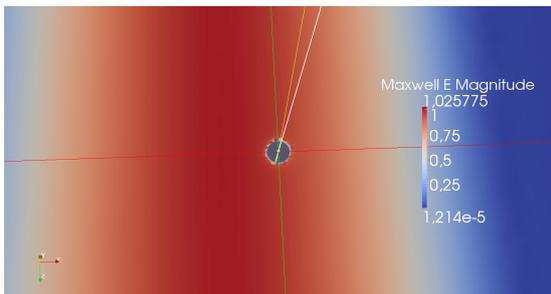


FIGURE 6.58 – Une coupe transversale de la pièce (zoom sur la pièce) à l'instant  $t = 0.7 \times T$  : la solution asynchrone à gauche et la solution synchrone à droite.

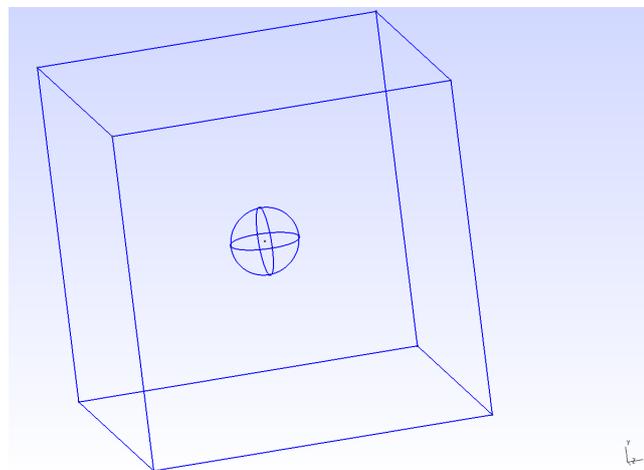


FIGURE 6.59 – Le domaine d'étude : une sphère à l'intérieur d'un cube.

l'instant  $t = 0.4 \times T$  sur la figure 6.61, à l'instant  $t = 0.7 \times T$  sur la figure 6.62 et à l'instant  $t = T$  sur la figure 6.63. Pour illustrer le passage de l'onde à travers la sphère, nous avons fait une coupe longitudinale du domaine (voir figure 6.64). Les solutions sont présentées dans la figure 6.65 et un zoom sur la coupe de la sphère est présenté dans la figure 6.66.

Nombre de points	32143
Nombre de cellules	171262
L'aire de la plus grande cellule	$7,2 \cdot 10^{-7}$
L'aire de la plus petite cellule	$6,6 \cdot 10^{-14}$
Le speed up théorique	113,7

FIGURE 6.60 – Quelques informations concernant le maillage construit pour l'exemple de la sphère.

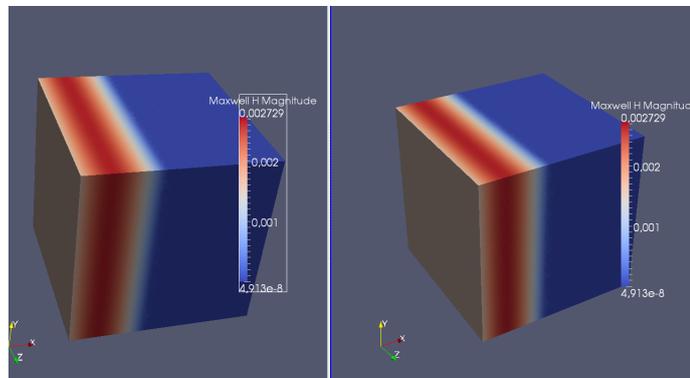


FIGURE 6.61 – Les solutions asynchrone (à gauche) et synchrone (à droite) à l'instant  $t = 0.4 \times T$ .

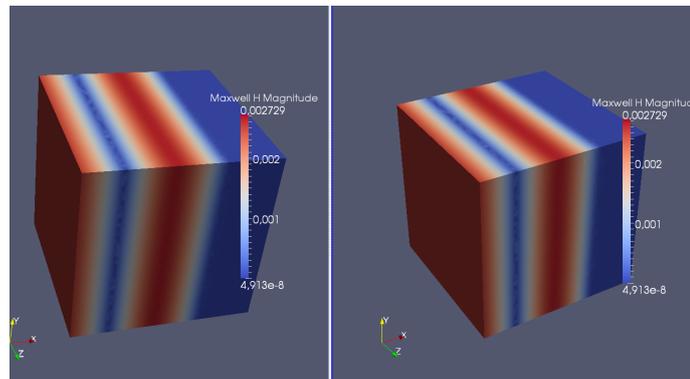


FIGURE 6.62 – Les solutions asynchrone (à gauche) et synchrone (à droite) à l'instant  $t = 0.7 \times T$ .

**Remarque 9.** La comparaison de la solution asynchrone avec la solution classique montre que pour le test de la sphère aussi l'asynchrone préserve la qualité de la solution classique. De plus, pour ce cas fortement multi-échelle :  $V_{min} \simeq 10^{-7} m^3 \ll V_{max} \simeq 10^{-14} m^3$  où  $V_{min}$  est le volume de la plus petite maille et  $V_{max}$  est le volume de la plus grande maille, on a un gain très important en temps de calcul. En fait, pour faire tourner les codes en mode synchrone on a eu besoin de 5 jours et 3h alors

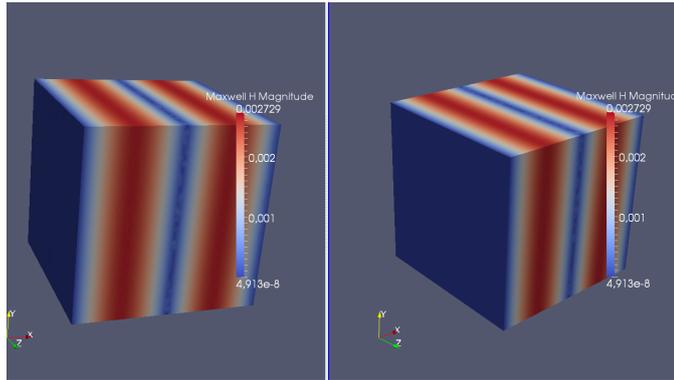


FIGURE 6.63 – Les solutions asynchrone (à gauche) et synchrone (à droite) à l’instant  $t = T$ .

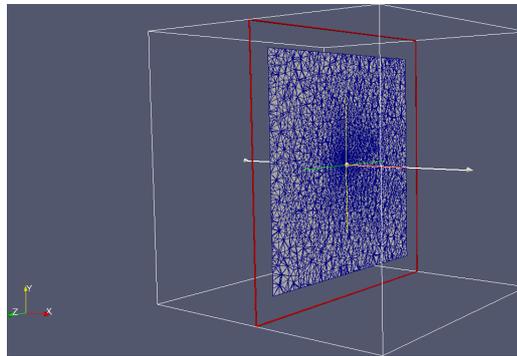


FIGURE 6.64 – Une coupe longitudinale du domaine d’étude.

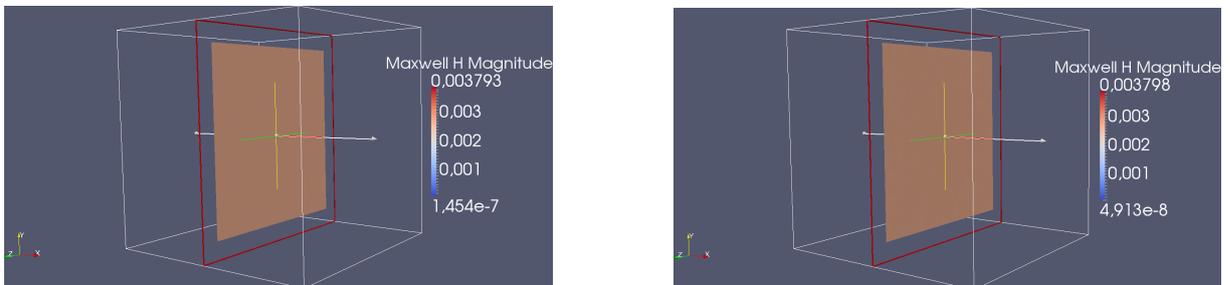


FIGURE 6.65 – Le passage de l’onde à travers la pièce à l’instant  $t = 0.7 \times T$  : la solution asynchrone à gauche et la solution synchrone à droite.

que l’asynchrone a nécessité uniquement 1h et 50mn. Ce qui donne un speed-up numérique de l’ordre de  $80 \gg 1$ . Le speed-up théorique pour ce test est 113 (voir figure 6.60).

### 6.4.3. Le fil avec défaut

Le dernier test de ce chapitre est l’exemple d’un fil cylindrique qui traverse un parallélépipède. Pour avoir un caractère multi-échelle très important, nous avons ajouté un défaut au milieu du fil qui va induire des inhomogénéités de maillage très localisées (voir figure 6.67). Pour des considérations informatiques

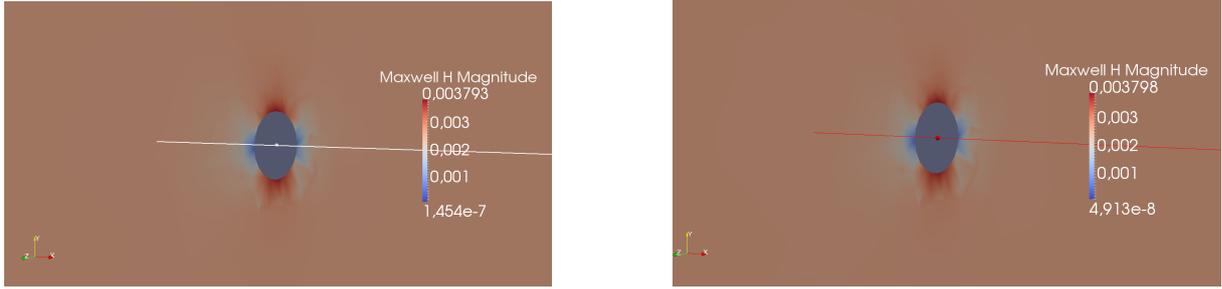


FIGURE 6.66 – Une coupe longitudinale de la sphère (zoom sur la sphère) à l’instant  $t = 0.7 \times T$  : la solution asynchrone à gauche et la solution synchrone à droite.

de taille mémoire, nous avons diminué la taille du domaine d’étude qui est pour ce test un parallélépipède de dimensions  $dx \times dy \times dz$  avec  $dx = dz = 0.25 \times \lambda$  et  $dy = 0.5 \times \lambda$  où  $\lambda$  est la longueur de l’onde injectée. Le rayon du fil  $r = 0.0001 \times \lambda \ll L = 0.5 \times \lambda$  où  $L$  est la longueur du fil. Le maillage a été construit de sorte que le speed-up théorique soit égal à  $612 \gg \gg 1$ . Les informations concernant le maillage sont données dans la figure 6.68.

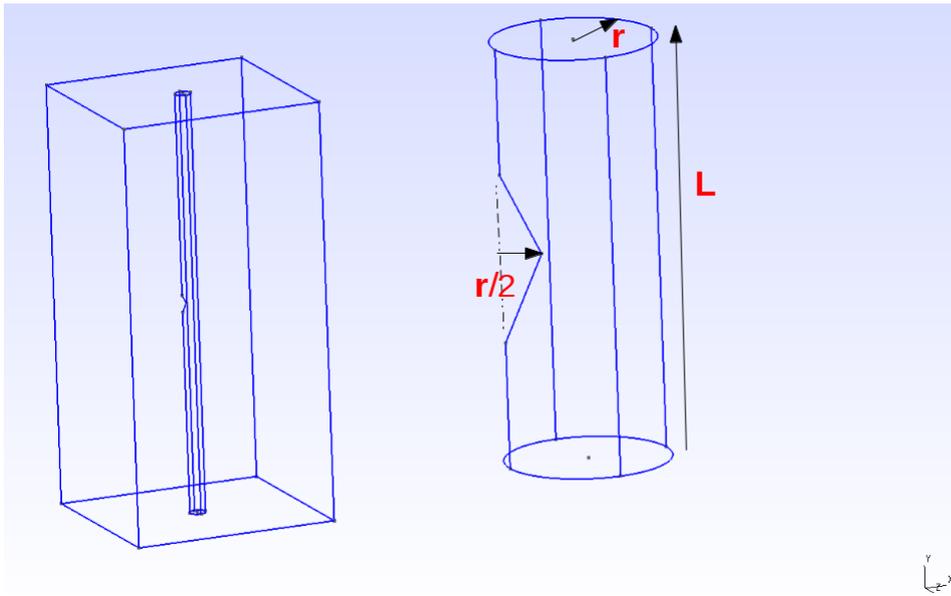


FIGURE 6.67 – Le domaine d’étude : un fil avec un défaut qui traverse un parallélépipède.

La solution asynchrone à l’instant  $t = 0.3 \times T$  avec  $T = 1.5 \times 10^{-9}$ s est présentée sur la figure 6.69 avec une coupe transversale au niveau du défaut à droite en haut et un zoom sur le défaut à droite en bas. Nous n’avons pas calculé la solution classique car le temps de calcul était prohibitif. En fait le volume de la plus petite maille est de l’ordre de  $10^{-16} \text{ m}^3$ . Pour estimer le speed-up numérique, nous avons comparé les temps de calcul pour l’asynchrone et pour le synchrone sur quelques nanosecondes de temps physique. Le speed-up numérique observé pour ce test en temps très court est de l’ordre de 440. Pour  $T = 1.5 \times 10^{-9}$  s, l’asynchrone a mis 2 jours et 19h pour tourner. En utilisant le speed-up numérique estimé, le solveur en mode synchrone devrait fonctionner pendant 3 ans. L’asynchrone

Nombre de points	38794
Nombre de cellules	406919
L'aire de la plus grande cellule	$7.3 \cdot 10^{-8}$
L'aire de la plus petite cellule	$2.3 \cdot 10^{-16}$
Le speed up théorique	612,92

FIGURE 6.68 – Quelques informations concernant le maillage construit pour l'exemple du fil avec défaut.

permet donc de gagner beaucoup de temps tout en préservant la précision de la méthode classique.

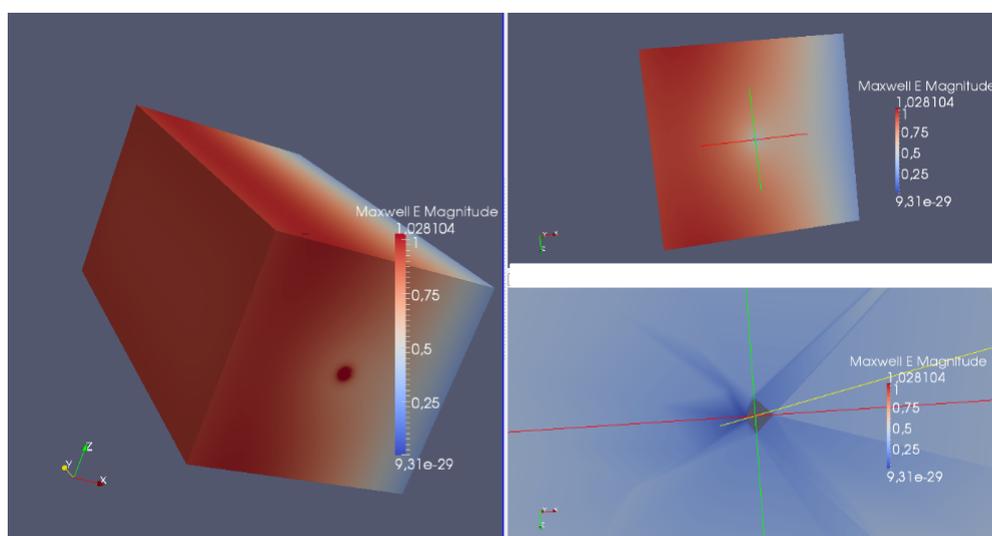


FIGURE 6.69 – La solution asynchrone à l'instant  $t = 0.3 \times T$ . Une coupe transversale du domaine au niveau du défaut à droite en haut et un zoom sur le défaut à droite en bas.

## 6.5. Conclusion

Ce chapitre a consisté à mettre en oeuvre la méthode asynchrone, développée dans le chapitre précédent, dans des cas-tests représentatifs d'un aspect multi-échelle très marqué. Les résultats théoriques démontrés nécessitent la linéarité de l'équation différentielle étudiée, nous avons donc choisi d'appliquer la méthode ARK2 pour la discrétisation des équations de Maxwell. Pour la discrétisation spatiale, la méthode de Galerkin Discontinue  $P^1$  a été utilisée. L'expérience numérique consistait à injecter une onde plane et à suivre son évolution temporelle.

Dans la section 6.2. nous avons traité l'exemple du fil en 2D avec deux formes différentes pour la section du fil (un cercle et un carré). Puis une extrusion en trois dimensions avec un fil cylindrique et

un autre parallélépipédique a été présentée dans la section 6.3.. Pour chaque exemple du fil, nous avons construit un maillage raffiné autour du fil pour créer l'aspect multi-échelles. La solution du schéma synchrone classique a été prise comme référence pour évaluer la précision de la méthode asynchrone. Les résultats numériques ont montré que l'asynchrone, que ce soit en dimension deux ou en dimension trois, et avec les différentes formes du fil, préserve la qualité de la solution classique. De plus, l'asynchrone est très efficace en termes de temps de calcul en comparaison avec le schéma synchrone dans le cas des maillages multi-échelles.

La dernière section de ce chapitre a été consacrée à l'étude de la performance de l'asynchrone en augmentant fortement le caractère multi-échelle des cas-test. Le but était donc de construire des domaines d'étude qui permettent d'avoir des maillages fortement multi-échelles avec des speed-up théoriques très importants. Trois tests ont été étudiés : la pièce, la sphère et le fil avec défaut, les speed-up théoriques sont respectivement 29, 113 et 612. Une petite variation entre le speed up théorique prévu et le speed up numérique a été notée. En effet, le speed up théorique est calculé en fonction du maillage uniquement et ne prend pas en compte ni l'implémentation ni le problème traité. C'est en fait le gain maximal qu'on pourrait atteindre. Cette variation peut aussi être liée à l'implémentation de l'asynchrone mais aussi du synchrone.

Bien que le speed-up numérique est généralement plus petit que le speed-up estimé, l'asynchrone reste très efficace en termes de temps de calcul tout en préservant la qualité de la solution classique. Pour les problèmes qui présentent une très forte inhomogénéité spatiale et temporelle des pas de discrétisation, le schéma asynchrone a permis de résoudre le problème du temps de calcul. De plus, l'asynchrone pourrait rendre réalisables des tests qui nécessiteraient un temps de calcul gigantesque (des années) pour tourner en mode classique (typiquement l'exemple du fil avec défaut présenté à la fin de ce chapitre).

## Chapitre 7

# Conclusion et perspectives

Ce manuscrit résume les résultats des recherches obtenus durant les trois ans de thèse. Le contexte général de cette thèse était la modélisation des phénomènes multi-échelles. La principale difficulté rencontrée lors de l'intégration numérique de ce type de modèles se traduit par une très forte inhomogénéité spatiale et temporelle des pas de discrétisation, laquelle conduit à des temps de calcul gigantesques. Notre objectif était donc de développer un schéma numérique asynchrone qui permet de limiter la quantité de calculs.

Dans ce chapitre, nous rappellerons dans un premier temps le concept des méthodes asynchrones développées, les principaux résultats prouvés et les avantages ainsi que les limites de l'asynchrone. Puis, dans un deuxième temps nous chercherons à ouvrir quelques perspectives.

### 7.1. Conclusion

La thèse a été répartie en deux grandes parties. La première partie se prolonge sur les trois premiers chapitres. Elle a été consacrée à l'étude du schéma numérique asynchrone d'ordre un appliqué aux équations de transport. Alors que la deuxième partie qui occupe le reste du manuscrit, elle, a été consacrée à l'extension de la précision du schéma asynchrone au second ordre en temps pour la modélisation des équations différentielles linéaires et la mise en oeuvre de la méthode.

Le contexte de cette thèse est celui de l'intégration temporelle explicite des EDPs multi-échelles. Dans le cas d'une approche explicite le pas de temps d'intégration est imposé par la contrainte locale la plus restrictive dans le domaine de calcul à travers la relation CFL. Autrement dit, toutes les cellules du maillage avancent à la vitesse de la cellule la plus lente. Si le problème étudié présente un caractère multi-échelles très prononcé les contraintes sur le pas de temps sont telles que la simulation impose des quantités de calcul prohibitives. Pour résoudre ce problème, un certain nombre de schémas à pas de temps locaux (LTS) ont été développés. Pour ce genre des méthodes, les pas de temps locaux doivent être choisis parmi les fractions du pas de temps global. Une classe des méthodes LTS, qui permettent de choisir des pas de temps indépendants, sont les schémas asynchrones. Notre objectif consistait donc à développer un schéma numérique asynchrone : un schéma explicite à pas de temps local et qui permet de limiter la quantité de calculs à effectuer tout en garantissant la consistance et la précision de la méthode.

La première étape de cette thèse consistait à définir et à éclaircir la méthodologie asynchrone. La

première idée de base de l'asynchrone est que chaque cellule de maillage utilise son propre pas de temps calculé sous la condition CFL locale. L'évolution temporelle est par conséquent différente de l'évolution classique. En effet, il existe une différence notable entre intégrations synchrones et asynchrones d'un point de vue algorithmique. L'intégration synchrone est réalisée en traitant de manière séquentielle et ordonnée toutes les cellules du maillage (boucle « do ») tandis que pour une intégration asynchrone, l'ordre dans lequel les cellules sont avancées en temps n'est pas défini a priori et dépend de leur pas de temps, du temps auquel elles se trouvent et du temps auquel se trouvent les autres cellules. L'intégration asynchrone n'est plus organisée en boucle « do » mais par un « Discrete Time Scheduler » qui répond à la question « à qui le tour maintenant ? ». La deuxième différence est que pour une méthode d'intégration temporelle classique, à chaque itération, toutes les cellules sont mises à jour au même temps de calcul. En revanche, pour l'intégration asynchrone, chaque cellule est mise à jour aux instants imposés par des conditions locales. Une conséquence évidente est qu'après un certain nombre d'itérations, toutes les cellules ne sont pas au même temps physique.

La question qui se pose alors est celle de la consistance de l'avancement en temps d'un tel schéma dans lequel deux cellules voisines ne sont pas évaluées au même instant. Autrement dit, comment calcule-t-on le flux entre deux cellules désynchronisées ? C'est là la deuxième idée de base de la méthode asynchrone, qui consiste à sauvegarder pour chaque cellule la valeur, le temps auquel cette valeur a été calculée et la dérivée temporelle de la valeur au temps considéré. Ces éléments sont valides jusqu'à la prochaine remise à jour de la cellule et peuvent être utilisés si on a besoin de calculer un flux à un temps intermédiaire pour faire avancer en temps une cellule voisine. Il suffit alors de procéder à une interpolation en temps. Après avoir défini l'asynchrone, l'étape suivante était bien évidemment d'étudier théoriquement les propriétés de cette méthode, en particulier la stabilité et la consistance. Nous avons commencé par appliquer le schéma asynchrone, basé sur la méthode d'Euler explicite, pour la discrétisation temporelle de l'équation de transport en dimension 1. La méthode de volumes finis basée sur une approximation d'ordre un a été utilisée pour la discrétisation spatiale. Nous avons prouvé que le schéma asynchrone est stable, consistant et convergent d'ordre un. Une étude numérique a été ensuite faite et qui valide les résultats théoriques prouvés en particulier l'ordre de convergence de la méthode asynchrone. Les tests numériques montrent aussi que, en comparaison avec le schéma synchrone classique, le schéma asynchrone réduit la diffusion numérique. Ce qui est une conséquence de l'utilisation de la CFL maximale pour chaque cellule. De plus, la méthode asynchrone est très efficace en termes de réduction du temps de calcul.

Le troisième chapitre était une extension du chapitre 2. Le schéma asynchrone et la méthode des volumes finis upwind ont été utilisés pour la discrétisation des équations de transport en dimension  $d$  d'espace où  $d \geq 1$ . La stabilité a été prouvée sans aucune contrainte autre que la condition de stabilité locale. La consistance a en revanche été discutée en fonction de la variation de la vitesse et de la variation du maillage. En fait pour les équations de transport, l'aspect asynchrone peut être dû à la variation de la vitesse, à la variation du maillage ou aux deux simultanément. Pour éviter les éventuelles complexités liées à la variation du maillage, nous avons supposé dans un premier temps que la vitesse est variable et que le maillage est constant. Nous avons prouvé que pour une vitesse suffisamment régulière, de classe au moins  $C^2$ , le schéma asynchrone est consistant d'ordre un en temps et en espace. Dans un deuxième temps, la vitesse était supposée constante et le maillage variable. Nous avons abouti à une perte locale de la consistance. Ce problème n'est pas lié à l'utilisation du schéma asynchrone. En fait, le schéma upwind, utilisé pour la discrétisation des équation de transport avec une intégration temporelle

synchrone, est localement inconsistant malgré la convergence de la méthode. L'erreur de convergence se comporte donc mieux que l'erreur de la consistance et le schéma est dit supra-convergent. Le théorème de Lax-Richtmyer n'est pas applicable dans ce cas. Dans [32], Bouche et al. ont utilisé la méthode de la convergence corrigée pour prouver que le schéma upwind appliqué aux équations de transport avec une discrétisation d'Euler explicite classique, est convergent d'ordre un. Nous avons donc adapté leur méthode à l'aspect asynchrone pour prouver qu'avec une intégration asynchrone, la convergence d'ordre un de la méthode est conservée.

Après avoir vérifié que le schéma asynchrone est convergent et que de plus il conserve l'ordre un de la méthode d'Euler explicite classique, nous avons ensuite réfléchi à appliquer le formalisme asynchrone à des schémas qui sont classiquement d'ordre plus élevé. Nous avons donc appliqué "naïvement" l'asynchrone aux schémas : Runge-Kutta, MUSCL et Lax-Wendroff, toujours pour la discrétisation des équations de transport. Nous avons prouvé que le schéma asynchrone est convergent mais uniquement à l'ordre un. Nous avons donc eu une perte d'ordre par rapport à une intégration temporelle classique de ces méthodes. En examinant l'erreur de la méthode asynchrone, nous avons constaté l'apparition d'une erreur supplémentaire : "une erreur asynchrone" en comparaison avec l'erreur du schéma synchrone classique. L'erreur asynchrone est due à l'utilisation des flux numériques qui ont été mis à jour à des temps différents pour la mise à jour de la même cellule. Nous avons prouvé que cette erreur asynchrone est au plus d'ordre un. Par conséquent, la méthode asynchrone telle qu'elle a été définie dans le deuxième chapitre, est au plus d'ordre un indépendamment de l'ordre classique de la méthode numérique adaptée à l'asynchrone. L'asynchrone converge donc moins vite que le synchrone. En revanche, l'asynchrone continue à être efficace en termes de gain en temps du calcul pour une erreur fixée. L'étude de convergence d'une méthode numérique nécessite une certaine régularité de la solution. Nous avons donc eu l'idée de sortir du cadre qui permet d'examiner l'ordre de la méthode numérique, et nous avons fait des tests numériques dans les cas des solutions irrégulières. Dans ce cadre, le schéma asynchrone a des propriétés numériques additionnelles. En comparaison avec le schéma classique, le schéma asynchrone réduit la diffusion numérique, réduit les oscillations autour des discontinuités qui apparaissent en utilisant le schéma Lax-Wendroff et réduit également le temps du calcul.

Malgré ses propriétés numériques, le schéma asynchrone tel qu'il a été présenté dans les premiers chapitres est toujours seulement d'ordre un, ce qui reste insuffisant pour certaines applications. Il fallait donc modifier l'algorithme asynchrone pour pouvoir atteindre l'ordre deux et ceci a été l'objet du chapitre 5. Nous avons donc considéré une équation différentielle linéaire donnée avec une discrétisation arbitraire en espace. Puis, pour la discrétisation temporelle, nous avons dérivé un schéma Runge-Kutta 2 asynchrone (ARK2) du schéma RK2 classique. L'idée fondamentale de la méthode asynchrone consistait à séparer le terme source locale et les flux puis à les intégrer indépendamment. L'algorithme ARK2 a été défini dans la section 5.1.1.. Nous avons ensuite donné une preuve détaillée de la convergence de la méthode asynchrone dans la section 5.2.. La convergence a été démontrée en trois grandes étapes. La première étape consiste à contrôler l'évolution temporelle d'un élément de maillage donné entre deux instants successifs de mise à jour de cet élément et qui sont multiples de son propre pas de temps local. L'idée de la deuxième étape est d'exprimer la solution à un temps donné en fonction du dernier temps auquel la solution a été synchronisée. En fonction du maillage, deux cas se présentent. Dans le premier cas, tous les éléments du maillage ont des temps de rencontre réguliers et la solution est automatiquement synchronisée aux instants qui sont des multiples du plus petit commun multiple (PPCM) des pas

de temps locaux. Le deuxième cas correspond à un maillage où les pas de temps sont complètement indépendants. Pour les besoins de la démonstration, nous avons supposé que la solution est synchronisée régulièrement à des instants qui sont des multiples d'un pas de temps qu'on a noté  $\Delta t_{\max}$  où, dans ce cas,  $\Delta t_{\max}$  est calculé en fonction des pas de temps locaux  $\Delta t_i$ . L'hypothèse de la synchronisation régulière de la solution n'a pas d'influence sur l'aspect asynchrone. De plus, numériquement, la synchronisation de la solution n'est imposée qu'à l'instant final de la simulation indépendamment du maillage. La troisième étape de la preuve est une comparaison de l'erreur asynchrone avec l'erreur du schéma classique entre deux temps de synchronisation de la solution. L'erreur du schéma numérique asynchrone est égal à l'erreur du schéma RK2 classique plus une erreur "asynchrone" supplémentaire. L'erreur asynchrone est d'ordre trois. Ce qui prouve que le schéma asynchrone préserve la précision du schéma synchrone à pas de temps global. Notre approche est applicable à tout type de maillage. Le schéma ARK2 est aussi convergent d'ordre deux indépendamment du maillage choisi. Pour valider les résultats théoriques prouvés dans la section 5.2., nous avons appliqué la méthode ARK2 pour la discrétisation des équations de Maxwell en dimension un puis en dimension deux d'espace. La méthode de Galerkin discontinue avec des éléments de base type  $P^1$  a été utilisée pour la discrétisation spatiale. Le schéma a été testé d'abord avec des maillages dont les pas de temps ont des temps de rencontre commun. Ensuite avec des maillages plus généraux où les pas de temps sont indépendants. Dans tous les cas, les tests numériques montrent que la solution asynchrone converge vers la solution exacte à l'ordre deux.

Les tests numériques présentés dans le chapitre 5 sont des tests académiques qui traitent en particulier de l'ordre de la méthode ARK2 en fonction du maillage. La mise en oeuvre de la méthode a été l'objet du chapitre 6. Une simulation des cas-tests simples mais représentatifs de cas multi-échelles a été présentée. Pour ce faire, nous avons considéré les équations de Maxwell et la méthode de Galerkin discontinue à été utilisée pour la discrétisation spatiale. L'exemple du fil a été traité en dimension deux puis son extension en dimension trois. Les maillages ont été construits par GMSH en tenant compte du caractère multi-échelle qui est le cadre naturel de l'application de l'asynchrone. Une comparaison avec la solution synchrone classique a été présentée. La solution asynchrone préserve la qualité de la solution classique. De plus, avec un maillage multi-échelle, l'asynchrone, malgré la montée en ordre, réduit fortement le temps du calcul.

Les tests numériques montrent aussi que pour un maillage fortement multi-échelle, le speed-up asynchrone est généralement plus lent que le speed-up prévu théoriquement. Ceci peut être expliqué par le fait que l'accès aux données dans la mémoire se fait d'une manière aléatoire dans le cas asynchrone. Le speed-up numérique peut aussi dépendre de l'implémentation asynchrone mais aussi synchrone. Enfin, l'utilisation de méthodes plus complexes d'intégration telles que l'algorithme ARK2 va également diminuer l'accélération des calculs. La dernière partie du chapitre 5 a été consacrée à l'étude de la performance de l'asynchrone en termes de gain en temps de calcul en fonction du caractère multi-échelles du maillage. Des maillages qui présentent une très forte inhomogénéité spatiale ont été construits. Les tests numériques ont prouvé que l'asynchrone permet de lever le verrou des temps de calcul gigantesques tout en préservant la qualité de la solution classique.

## 7.2. Perspectives

Pour terminer ce manuscrit, nous discutons dans cette dernière section quelques questions ouvertes et perspectives de ce travail. Plusieurs voies peuvent être explorées.

Une perspective immédiate et très intéressante est l'extension de la méthode asynchrone à des ordres plus élevés. En fait, dans le chapitre 5, nous avons dérivé un schéma RK2 asynchrone du schéma RK2 classique. Pour développer des schémas asynchrones de haute précision, il est intéressant de dériver des versions asynchrones des schémas RK $p$  avec  $p > 2$ . L'extension n'est pas forcément triviale, car il faudra toujours assurer la consistance des instants auxquels les estimateurs seront évalués. Cependant, une très grande partie du travail a déjà été faite pour l'ordre deux. En particulier le contrôle de l'erreur asynchrone et la dérivation de l'asynchrone du schéma classique en préservant la précision. Une fois l'algorithme ARK $p$ ,  $p > 2$  est mis en place. La preuve, si elle pourrait en théorie n'être une simple extension de la preuve d'ordre deux, devrait toutefois présenter quelques complications techniques en augmentant le nombre d'étapes intermédiaires de mise à jour.

Lors de l'estimation de l'ordre des méthodes d'intégration temporelle l'aspect asynchrone fait intervenir les ordres d'erreur des estimations spatiales. Il serait donc intéressant d'étudier l'impact de la discrétisation spatiale choisie sur la qualité des solutions asynchrones. Il serait pour cela nécessaire de développer au sein de MACOPA des approximations de type Galerkin Discontinu d'ordre supérieur.

Une deuxième perspective intéressante est l'application de la méthode asynchrone à des problèmes non linéaires. En fait, pour des raisons liées aux démonstrations des théorèmes, nous avons imposé la linéarité des équations différentielles étudiées. Cependant, la plupart des problèmes physiques se modélisent par des équations différentielles non linéaires, problèmes pour lesquels le solveur MACOPA peut tout de même apporter un gain non négligeable en temps de calcul. Il s'avèrera donc nécessaire d'étendre l'étude théorique de la méthode asynchrone dans le cadre de ce genre de problèmes. Cela nécessitera une étude dédiée.

Enfin, les tests numériques qui ont été présentés dans ce manuscrit sont relativement simples. Dans un premier temps, des tests académiques ont servi à valider les résultats théoriques en particulier la convergence et l'ordre de la méthode mais aussi à comparer les solutions synchrone et asynchrone. Puis, dans un deuxième temps, pour la mise en oeuvre de la méthode ARK2, une simulation de cas multi-échelles simples a été présentée. La troisième voie d'approfondissement de cette étude est donc de viser des applications plus complexes, voire de proposer des cas de validation du nouveau schéma asynchrone sur des applications industrielles.



## Annexe A

# Preuves du Chapitre 4

### Preuve de la proposition 8 : Consistance du schéma A1RK2

*Démonstration.* Faisons un raisonnement par récurrence. Soient  $t_{i,q}^1 = t^q + \min_j (\Delta t_j) := t^q + \Delta t_i$  et  $t_{i,q}^{1/2} = t^q + \frac{\Delta t_i}{2}$ . Il suffit d'utiliser un développement de Taylor d'ordre 2. On suppose que la formule (4.5) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soient  $t_q^{n+1} = t_{i,q}^{r_i+1} := t^q + (r_i + 1)\Delta t_i < t^{q+1}$  alors

$$(Ec)_i^{t_{i,q}^{r_i+1}} = \frac{\tilde{u}_i^{t_{i,q}^{r_i+1}} - \tilde{u}_i^{t_i^{pcd}}}{t_{i,q}^{r_i+1} - t_i^{pcd}} + \frac{1}{\Delta x_i} \left( \tilde{F}_i^{t_{i,q}^{r_i+1/2}} - \tilde{F}_{i-1}^{t_{i-1}^{r_i+1/2}} \right).$$

Supposons, par exemple, que  $t_i^{pcd} = t_{i,q}^{r_i}$ . Deux cas se présentent :

$$1. \tilde{F}_{i-1}^{t_{i-1}^{r_{i-1}+1/2}} = \tilde{F}_{i-1}^{t_{i-1}^{r_i+1/2}} :$$

$$(Ec)_i^{t_{i,q}^{r_i+1}} = \frac{\tilde{u}_i^{t_{i,q}^{r_i+1}} - \tilde{u}_i^{t_{i,q}^{r_i}}}{\Delta t_i} + \frac{1}{\Delta x_i} \left( \tilde{F}_i^{t_{i,q}^{r_i+1/2}} - \tilde{F}_{i-1}^{t_{i-1}^{r_i+1/2}} \right).$$

D'après l'hypothèse de récurrence on a

$$\begin{aligned} \tilde{F}_i^{t_{i,q}^{r_i+1/2}} &= \tilde{u}_i^{t_{i,q}^{r_i}} - \frac{\Delta t_i}{2\Delta x_i} \left( \tilde{F}_i^{t_{i,q}^{r_i}} - \tilde{F}_{i-1}^{t_{i-1}^{r_i}} \right) \\ &= \tilde{u}_i^{t_{i,q}^{r_i}} - \frac{\Delta t_i}{2\Delta x_i} \left[ (\tilde{u}_i^{t_q} - \tilde{u}_{i-1}^{t_q}) - (r_i\Delta t_i - r_{i-1}\Delta t_{i-1}) \frac{\partial u_i^{t_q}}{\partial x} + r_{i-1}\Delta t_{i-1} dx(i-1) \frac{\partial^2 u_i^{t_q}}{\partial x^2} \right] \\ &\quad + \frac{\Delta t_i}{2\Delta x_i} \left[ \left( \frac{(r_i\Delta t_i)^2}{2} - \frac{(r_{i-1}\Delta t_{i-1})^2}{2} \right) \frac{\partial^2 u_i^{t_q}}{\partial t^2} + O((\Delta x_i)^2) \right] \\ &\quad - \frac{\Delta t_i}{2\Delta x_i} \left[ r_i\Delta t_i (Ec)_i^{t_{i,q}^{r_i}} - r_{i-1}\Delta t_{i-1} (Ec)_{i-1}^{t_{i-1}^{r_i}} \right] \\ &= \tilde{u}_i^{t_{i,q}^{r_i}} - \frac{\Delta t_i}{2\Delta x_i} dx(i-1) \frac{\partial u_i^{t_q}}{\partial x} - \Delta t_i O(\Delta x_i). \end{aligned}$$

De même,

$$\tilde{F}_{i-1}^{t_{i-1,q}^{r_{i-1}+1/2}} = \tilde{u}_{i-1}^{t_{i-1,q}^{r_{i-1}}} - \frac{\Delta t_{i-1}}{2\Delta x_{i-1}} dx(i-2) \frac{\partial u_{i-1}^{t_q}}{\partial x} - \Delta t_{i-1} O(\Delta x_{i-1}).$$

Par suite

$$\begin{aligned} \tilde{F}_i^{t_{i,q}^{r_i+1/2}} - \tilde{F}_{i-1}^{t_{i-1,q}^{r_{i-1}+1/2}} &= \left( \tilde{u}_i^{t_{i,q}^{r_i}} - \tilde{u}_{i-1}^{t_{i-1,q}^{r_{i-1}}} \right) - \left( \frac{\Delta t_i}{2\Delta x_i} dx(i-1) - \frac{\Delta t_{i-1}}{2\Delta x_{i-1}} dx(i-2) \right) \frac{\partial u_i^{t_q}}{\partial x} \\ &\quad - \frac{\Delta t_{i-1}}{2\Delta x_{i-1}} dx(i-2) dx(i-1) \frac{\partial^2 u_i^{t_q}}{\partial x^2} - \Delta t_i O(\Delta x_i - \Delta x_{i-1}) \\ &= \left( \tilde{u}_i^{t_q} - \tilde{u}_{i-1}^{t_q} \right) - (r_i \Delta t_i - r_{i-1} \Delta t_{i-1}) \frac{\partial u_i^{t_q}}{\partial x} + O((\Delta x_i)^2) \\ &= dx(i-1) \frac{\partial u_i^{t_q}}{\partial x} + O((\Delta x_i)^2). \end{aligned}$$

Ce qui donne

$$\begin{aligned} \tilde{u}_i^{t_{i,q}^{r_i+1}} &= \tilde{u}_i^{t_q} - r_i \Delta t_i \left( \frac{\partial u_i^{t_q}}{\partial x} + O(\Delta x_i) + \frac{r_i \Delta t_i}{2} \frac{\partial^2 u_i^{t_q}}{\partial t^2} \right) - \frac{\Delta t_i}{\Delta x_i} \left( dx(i-1) \frac{\partial u_i^{t_q}}{\partial x} + O((\Delta x_i)^2) \right) \\ &\quad + (2r_i + 1) \frac{(\Delta t_i)^2}{2} \frac{\partial^2 u_i^{t_q}}{\partial t^2} - (2r_i + 1) \frac{(\Delta t_i)^2}{2} \frac{\partial^2 u_i^{t_q}}{\partial t^2} \\ &= \tilde{u}_i^{t_q} - (r_i + 1) \Delta t_i \left( \frac{\partial u_i^{t_q}}{\partial x} + O(\Delta x_i) + \frac{(r_i + 1) \Delta t_i}{2} \frac{\partial^2 u_i^{t_q}}{\partial t^2} - (Ec)_i^{t_{i,q}^{r_i+1}} \right) \end{aligned}$$

$$2. \tilde{F}_{i-1}^{t_{i-1,q}^{r_{i-1}+1/2}} = \tilde{F}_{i-1}^{t_{i-1,q}^{r_{i-1}+1}} :$$

$$\begin{aligned} \tilde{F}_i^{t_{i,q}^{r_i+1/2}} - \tilde{F}_{i-1}^{t_{i-1,q}^{r_{i-1}+1}} &= \left( \tilde{u}_i^{t_q} - \tilde{u}_{i-1}^{t_q} \right) - (r_i \Delta t_i - r_{i-1} \Delta t_{i-1}) \frac{\partial u_i^{t_q}}{\partial x} + r_{i-1} \Delta t_{i-1} dx(i-1) \frac{\partial^2 u_i^{t_q}}{\partial x^2} \\ &\quad - \frac{\Delta t_i}{2\Delta x_i} dx(i-1) \frac{\partial u_i^{t_q}}{\partial x} - \left( \frac{(r_i \Delta t_i)^2}{2} - \frac{(r_{i-1} \Delta t_{i-1})^2}{2} \right) \frac{\partial^2 u_i^{t_q}}{\partial t^2} \\ &= dx(i-1) \frac{\partial u_i^{t_q}}{\partial x} - \frac{\Delta t_i}{2\Delta x_i} dx(i-1) \frac{\partial u_i^{t_q}}{\partial x} + O((\Delta x_i)^2). \end{aligned}$$

Donc

$$\begin{aligned} \tilde{u}_i^{t_{i,q}^{r_i+1}} &= \tilde{u}_i^{t_q} - r_i \Delta t_i \left( \frac{\partial u_i^{t_q}}{\partial x} + O(\Delta x_i) + \frac{r_i \Delta t_i}{2} \frac{\partial^2 u_i^{t_q}}{\partial t^2} \right) \\ &\quad - \frac{\Delta t_i}{\Delta x_i} \left( dx(i-1) - \frac{\Delta t_i}{2\Delta x_i} dx(i-1) \right) \frac{\partial u_i^{t_q}}{\partial x} + \Delta t_i O(\Delta x_i) \end{aligned}$$

On a

$$\frac{\Delta t_i}{\Delta x_i} \left( dx(i-1) - \frac{\Delta t_i}{2\Delta x_i} dx(i-1) \right) = \Delta t_i + \Delta t_i \left( \left( \frac{dx(i-1)}{\Delta x_i} - 1 \right) - \Delta t_i \frac{dx(i-1)}{2(\Delta x_i)^2} \right),$$

avec

$$\begin{aligned} \Delta t_i + \Delta t_i \left( \left( \frac{dx(i-1)}{\Delta x_i} - 1 \right) - \Delta t_i \frac{dx(i-1)}{2(\Delta x_i)^2} \right) &\lesssim \Delta t_i + \Delta t_i \left( \frac{dx(i-1)}{\Delta x_i} - 1 \right) \\ &= \Delta t_i + \Delta t_i O(\Delta x_i) \end{aligned}$$

Ce qui donne

$$\tilde{u}_i^{t_i^{r_i+1}} = \tilde{u}_i^{t_i^{r_i}} - (r_i + 1) \Delta t_i \left( \frac{\partial u_i^{t_i^{r_i}}}{\partial x} + O(\Delta x_i) + \frac{(r_i + 1) \Delta t_i}{2} \frac{\partial^2 u_i^{t_i^{r_i}}}{\partial t^2} - (Ec)_i^{t_i^{r_i+1}} \right)$$

par suite

$$(Ec)_i^{t_i^{r_i+1}} = \frac{\tilde{u}_i^{t_i^{r_i+1}} - \tilde{u}_i^{t_i^{r_i}}}{(r_i + 1) \Delta t_i} + \left( \frac{\partial u_i^{t_i^{r_i}}}{\partial x} + O(\Delta x_i) + \frac{(r_i + 1) \Delta t_i}{2} \frac{\partial^2 u_i^{t_i^{r_i}}}{\partial t^2} \right)$$

La formule (4.5) est donc vérifiée et le schéma asynchrone est consistant au second ordre en temps mais seulement au premier ordre en espace.  $\square$

## Preuve de la proposition 12 : Consistance du schéma MUSCL asynchrone

*Démonstration.* Faisons un raisonnement par récurrence. Soit  $t_i^1 = t^q + \min_j (\Delta t_j) := t^q + \Delta t_i$ . Alors

$$(Ec)_i^{t_i^1} = \frac{\tilde{u}_i^{t_i^1} - \tilde{u}_i^{t_i^q}}{\Delta t_i} \frac{1}{\Delta x_i} \left( \tilde{F}_i^{t_i^q} - \tilde{F}_{i-1}^{t_i^q} \right)$$

par suite

$$\begin{aligned} \tilde{u}_i^{t_i^1} &= \tilde{u}_i^{t_i^q} - \frac{\Delta t_i}{\Delta x_i} \left( \tilde{F}_i^{t_i^q} - \tilde{F}_{i-1}^{t_i^q} \right) + \Delta t_i (Ec)_i^{t_i^1} \\ &= \tilde{u}_i^{t_i^q} - \frac{\Delta t_i}{\Delta x_i} \left( dx(i-1) \frac{\partial \tilde{u}_i^{t_i^q}}{\partial x} + \frac{(dx(i-1))^2}{2} \frac{\partial^2 \tilde{u}_i^{t_i^q}}{\partial x^2} + O((dx(i-1))^3) \right) \\ &\quad - \frac{\Delta t_i}{2 dx(i)} \left( dx(i) \frac{\partial \tilde{u}_i^{t_i^q}}{\partial x} + \frac{(dx(i))^2}{2} \frac{\partial^2 \tilde{u}_i^{t_i^q}}{\partial x^2} + O((dx(i))^3) \right) + \Delta t_i (Ec)_i^{t_i^1} \\ &\quad + \frac{\Delta t_i \Delta x_{i-1}}{2 \Delta x_i dx(i-1)} \left( dx(i-1) \frac{\partial \tilde{u}_i^{t_i^q}}{\partial x} + \frac{(dx(i-1))^2}{2} \frac{\partial^2 \tilde{u}_i^{t_i^q}}{\partial x^2} + O((dx(i-1))^3) \right) \\ &= \tilde{u}_i^{t_i^q} - \Delta t_i \frac{\partial \tilde{u}_i^{t_i^q}}{\partial x} + \Delta t_i O((\Delta x_i)^2) + \Delta t_i (Ec)_i^{t_i^1}. \end{aligned}$$

Supposons maintenant que la formule (4.17) est vraie jusqu'à un instant  $t_q^n < t^{q+1}$ . Soit  $t_q^{n+1} = t_i^{r_i+1} := t^q + (r_i + 1)\Delta t_i < t^{q+1}$ . Alors,

$$\begin{aligned}\tilde{u}_i^{t_i^{r_i+1}} &= \tilde{u}_i^{t_i^{r_i}} - \frac{\Delta t_i}{\Delta x_i} \left( \tilde{F}_i^{t_i^{r_i}} - \tilde{F}_{i-1}^{t_i^{r_i-1}} \right) + \Delta t_i (Ec)_i^{t_i^{r_i+1}} \\ &= \tilde{u}_i^{t_i^{r_i}} - \frac{\Delta t_i}{\Delta x_i} \left( \tilde{u}_i^{t_i^{r_i}} - \tilde{u}_{i-1}^{t_i^{r_i-1}} \right) - \frac{\Delta t_i}{2\Delta x_i} \left( \frac{\Delta x_i}{dx(i)} \varphi(R_i^{t_i^{r_i}}) (\tilde{u}_{i+1}^{t_i^{r_i}} - \tilde{u}_i^{t_i^{r_i}}) \right) \\ &\quad + \frac{\Delta t_i}{2\Delta x_i} \left( \frac{\Delta x_{i-1}}{dx(i-1)} \varphi(R_{i-1}^{t_i^{r_i-1}}) (\tilde{u}_i^{t_i^{r_i-1}} - \tilde{u}_{i-1}^{t_i^{r_i-1}}) \right) + \Delta t_i (Ec)_i^{t_i^{r_i+1}}\end{aligned}$$

Par suite,

$$\begin{aligned}\tilde{u}_i^{t_i^{r_i+1}} &= \tilde{u}_i^{t_i^{r_i}} - \frac{\Delta t_i}{\Delta x_i} \left( \tilde{u}_i^{t^q} - \tilde{u}_{i-1}^{t^q} \right) \\ &\quad + \frac{\Delta t_i}{\Delta x_i} \left( (r_i \Delta t_i - r_{i-1} \Delta t_{i-1}) \frac{\partial \tilde{u}_i^{t^q}}{\partial x} + r_{i-1} \Delta t_{i-1} dx(i-1) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \right) \\ &\quad - \frac{\Delta t_i}{2\Delta x_i} \left( \frac{\Delta x_i}{dx(i)} (\tilde{u}_{i+1}^{t^q} - \tilde{u}_i^{t^q}) - \frac{\Delta x_i}{dx(i)} r_i \Delta t_i dx(i) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \right) \\ &\quad + \frac{\Delta t_i}{2\Delta x_i} \left( \frac{\Delta x_{i-1}}{dx(i-1)} (\tilde{u}_i^{t^q} - \tilde{u}_{i-1}^{t^q}) - \frac{\Delta x_{i-1}}{dx(i-1)} r_{i-1} \Delta t_{i-1} dx(i-1) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \right) \\ &\quad + \frac{\Delta t_i}{\Delta x_i} \left( r_i \Delta t_i (Ec)_i^{t_i^{r_i+1}} - r_{i-1} \Delta t_{i-1} (Ec)_i^{t_i^{r_i+1}} \right) + \Delta t_i (Ec)_i^{t_i^{r_i+1}} \\ &= \tilde{u}_i^{t_i^{r_i}} - \Delta t_i \frac{\partial \tilde{u}_i^{t^q}}{\partial x} + \frac{\Delta t_i}{\Delta x_i} (r_i \Delta t_i - r_{i-1} \Delta t_{i-1}) \frac{\partial \tilde{u}_i^{t^q}}{\partial x} + \Delta t_i (Ec)_i^{t_i^{r_i+1}} \\ &\quad + r_{i-1} \Delta t_{i-1} dx(i-1) \frac{\Delta t_i}{\Delta x_i} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} + \frac{\Delta t_i}{2\Delta x_i} (r_i \Delta x_i \Delta t_i - r_{i-1} \Delta x_{i-1} \Delta t_{i-1}) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \\ &= \tilde{u}_i^{t^q} - (r_i + 1) \Delta t_i \left( \frac{\partial \tilde{u}_i^{t^q}}{\partial x} + O(\Delta x_i) - (Ec)_i^{t_i^{r_i+1}} \right).\end{aligned}$$

Ce qui donne finalement,

$$(Ec)_i^{t_i^{r_i+1}} = \frac{\tilde{u}_i^{t_i^{r_i+1}} - \tilde{u}_i^{t^q}}{(r_i + 1)\Delta t_i} + \left( \frac{\partial \tilde{u}_i^{t^q}}{\partial x} + O(\Delta x_i) \right).$$

La formule (4.17) est ainsi vérifiée est le schéma MUSCL asynchrone est donc d'ordre un en temps et en espace.  $\square$

## Preuve de la Proposition 13 : Consistance du schéma Lax-Wendroff asynchrone

*Démonstration.* La démonstration se fait par récurrence. On va traiter seulement le premier pas de temps. Pour finir la récurrence il suffit de procéder comme pour les démonstrations précédentes. Soit  $t_{i,q}^1 = t^q + \Delta t_i = \min_j (\Delta t_j)$ . On a

$$(Ec)_i^{t_{i,q}^1} = \frac{\tilde{u}_i^{t_{i,q}^1} - \tilde{u}_i^{t^q}}{\Delta t_i} + \frac{1}{2\Delta x_i} (\tilde{u}_{i+1}^{t^q} - \tilde{u}_{i-1}^{t^q}) - \frac{(\Delta t_{i+1} + \Delta t_i)}{4\Delta x_i dx(i+1)} (\tilde{u}_{i+1}^{t^q} - \tilde{u}_i^{t^q}) \\ + \frac{(\Delta t_{i-1} + \Delta t_i)}{4\Delta x_i dx(i)} (\tilde{u}_i^{t^q} - \tilde{u}_{i-1}^{t^q})$$

ce qui donne

$$\tilde{u}_i^{t_{i,q}^1} = \tilde{u}_i^{t^q} - \frac{\Delta t_i}{2\Delta x_i} (\tilde{u}_{i+1}^{t^q} - \tilde{u}_{i-1}^{t^q}) + \frac{\Delta t_i(\Delta t_{i+1} + \Delta t_i)}{4\Delta x_i dx(i+1)} (\tilde{u}_{i+1}^{t^q} - \tilde{u}_i^{t^q}) \\ - \frac{\Delta t_i(\Delta t_{i-1} + \Delta t_i)}{4\Delta x_i dx(i)} (\tilde{u}_i^{t^q} - \tilde{u}_{i-1}^{t^q}) + \Delta t_i (Ec)_i^{t_{i,q}^1} \\ = \tilde{u}_i^{t^q} - \frac{\Delta t_i}{2\Delta x_i} (dx(i+1) + dx(i)) \frac{\partial \tilde{u}_i^{t^q}}{\partial x} - \frac{\Delta t_i}{2\Delta x_i} \left( \frac{(dx(i+1))^2}{2} - \frac{(dx(i))^2}{2} \right) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \\ + \Delta t_i O((\Delta x_i)^2) + \frac{\Delta t_i(\Delta t_{i+1} + \Delta t_i)}{4\Delta x_i} \left( \frac{\partial \tilde{u}_i^{t^q}}{\partial x} + \frac{dx(i+1)}{2} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \right) \\ - \frac{\Delta t_i(\Delta t_{i-1} + \Delta t_i)}{4\Delta x_i} \left( \frac{\partial \tilde{u}_i^{t^q}}{\partial x} - \frac{dx(i)}{2} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \right) + \Delta t_i (Ec)_i^{t_{i,q}^1},$$

par suite

$$\tilde{u}_i^{t_{i,q}^1} = \tilde{u}_i^{t^q} - \Delta t_i \frac{\partial \tilde{u}_i^{t^q}}{\partial x} - \frac{\Delta t_i}{2} (dx(i+1) - dx(i)) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} + \frac{\Delta t_i}{4\Delta x_i} (\Delta t_{i+1} - \Delta t_{i-1}) \frac{\partial \tilde{u}_i^{t^q}}{\partial x} \\ + \frac{(\Delta t_i)^2}{4} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} + \frac{\Delta t_i}{4\Delta x_i} \left( \frac{dx(i+1)}{2} \Delta t_{i+1} + \frac{dx(i)}{2} \Delta t_{i-1} \right) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} + \Delta t_i O((\Delta x_i)^2) \\ + \Delta t_i (Ec)_i^{t_{i,q}^1}.$$

Ce qui donne

$$(Ec)_i^{t_{i,q}^1} = \frac{\Delta t_i}{2} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} + O((\Delta t_i)^2) + \frac{1}{2} (dx(i+1) - dx(i)) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} - \frac{1}{4\Delta x_i} (\Delta t_{i+1} - \Delta t_{i-1}) \frac{\partial \tilde{u}_i^{t^q}}{\partial x} \\ - \frac{\Delta t_i}{4} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} - \frac{\Delta t_i}{4\Delta x_i} \left( \frac{dx(i+1)}{2} \Delta t_{i+1} + \frac{dx(i)}{2} \Delta t_{i-1} \right) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} + O((\Delta x_i)^2).$$

Or, dans ce cas,  $\Delta t_i = \min_j (\Delta t_j)$  alors,

$$\frac{(\Delta t_i)^2}{4} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} + \frac{\Delta t_i}{4\Delta x_i} \left( \frac{dx(i+1)}{2} \Delta t_{i+1} + \frac{dx(i)}{2} \Delta t_{i-1} \right) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} \geq \frac{(\Delta t_i)^2}{2} \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2}.$$

Ce qui donne

$$\begin{aligned} (Ec)_{i,q}^{t^1} &\leq \frac{1}{2} (dx(i+1) - dx(i)) \frac{\partial^2 \tilde{u}_i^{t^q}}{\partial x^2} - \frac{1}{4\Delta x_i} (\Delta t_{i+1} - \Delta t_{i-1}) \frac{\partial \tilde{u}_i^{t^q}}{\partial x} + O((\Delta t_i)^2) + O((\Delta x_i)^2) \\ &\leq O((\Delta t_i)^2) + O((\Delta x_i)^2) - \frac{1}{4\Delta x_i} (\Delta t_{i+1} - \Delta t_{i-1}) \frac{\partial \tilde{u}_i^{t^q}}{\partial x}. \end{aligned}$$

Le schéma Lax Wendroff asynchrone est donc d'ordre 1.

□

## Annexe B

# Méthode de Galerkin discontinue pour la discrétisation des équations de Maxwell 3D

### Équations de Maxwell tridimensionnelles

Le système des équations de Maxwell tridimensionnelles est donné par :

$$\begin{cases} \varepsilon \frac{\partial E}{\partial t} - \text{rot}(H) = 0, \\ \mu \frac{\partial H}{\partial t} + \text{rot}(E) = 0, \end{cases} \quad (\text{B.1})$$

où

1.  $E(X, t) = \begin{pmatrix} E_1(X, t) \\ E_2(X, t) \\ E_3(X, t) \end{pmatrix}$  et  $H(X, t) = \begin{pmatrix} H_1(X, t) \\ H_2(X, t) \\ H_3(X, t) \end{pmatrix}$  sont respectivement le champ électrique et

le champ magnétique et  $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ ,

2.  $\varepsilon$  et  $\mu$  sont respectivement la permittivité électrique et la perméabilité magnétique.

Le système (B.1) écrit sous forme condensée est donné par :

$$G(\partial_t W) + \nabla \cdot F(W) = 0, \quad (\text{B.2})$$

avec :

1.  $W = \begin{pmatrix} E \\ H \end{pmatrix}$  est le champ du vecteur inconnu qui représente le champ électromagnétique,

2.  $G = \begin{pmatrix} \varepsilon I_{3,3} & 0_{3,3} \\ 0_{3,3} & \mu I_{3,3} \end{pmatrix}$  est la matrice qui rassemble les propriétés des milieux,

3.  $F(W) = \begin{pmatrix} F_1(W) \\ F_2(W) \\ F_3(W) \end{pmatrix}$  avec

$$(a) F_1(W) = \begin{pmatrix} 0_{3,3} & N_1 \\ -N_1 & 0_{3,3} \end{pmatrix} W, F_2(W) = \begin{pmatrix} 0_{3,3} & N_2 \\ -N_2 & 0_{3,3} \end{pmatrix} W \text{ et } F_3(W) = \begin{pmatrix} 0_{3,3} & N_3 \\ -N_3 & 0_{3,3} \end{pmatrix} W.$$

(b) Les matrices  $N_1, N_2$  et  $N_3$  sont données par

$$N_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, N_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ et } N_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$4. \nabla \cdot F(W) = \partial_{x_1}(F_1(W)) + \partial_{x_2}(F_2(W)) + \partial_{x_3}(F_3(W)).$$

## Méthode de Galerkin Discontinue

Soit  $\mathcal{T} = \{K_i : i = 1, \dots, N\}$  une partition du domaine d'étude  $\Omega$  en polyèdres  $K_i$ . Notons  $\psi$  une fonction test scalaire. On multiplie l'équation par  $\psi$  et on intègre sur  $K_i$  :

$$\int_{K_i} G(\partial_t W) \psi dx + \int_{K_i} \nabla \cdot F(W) \psi dx = 0.$$

En utilisant une formule d'intégration par parties (Green), nous avons :

$$\int_{K_i} G(\partial_t W) \psi dx - \int_{K_i} \nabla \psi \cdot F(W) dx + \int_{\partial K_i} (F(W) \cdot n) \psi d\sigma = 0.$$

Ce qui donne

$$\int_{K_i} G(\partial_t W) \psi dX - \int_{K_i} \nabla \psi \cdot F(W) dX + \sum_{j \in N(i)} \int_{K_i \cap K_j} (F(W) \cdot n_{i,j}) \psi d\sigma = 0,$$

où  $N(i)$  est l'ensemble des voisins de  $K_i$ ,  $N(i) = \{j \mid K_i \cap K_j \neq \emptyset\}$  et  $n_{i,j}$  est la normale unitaire de  $K_i \cap K_j$  pointant de  $K_i$  vers  $K_j$ .

La solution approchée  $W_h$  est ensuite cherchée dans un espace de dimension finie  $V_h$ . Contrairement à la méthode des éléments finis, la méthode GD autorise la solution approchée à être discontinue. L'espace de discrétisation  $V_h$  utilisé pour les tests présentés dans cette thèse est l'espace des fonctions polynomiales par morceaux. Des espaces de fonctions non polynomiales peuvent également être utilisés. Par exemples les fonctions exponentielles ou trigonométriques.

L'étape suivante consiste à choisir une base de  $V_h$ . Notons que le choix de fonctions de base peut avoir un impact sur la précision de calcul et sur l'efficacité dans la mise en oeuvre pratique des méthodes de Galerkin discontinues. Soit  $\phi = \{\{\varphi_{i,l}\}_{l \in D_{K_i}}\}_{K_i \in \mathcal{T}}$  une base de  $V_h$  où  $D_{K_i} = \{1, \dots, N_{dof}^{K_i}\}$  est l'ensemble des indices locaux de  $N_{dof}^{K_i}$  degrés de liberté de  $K_i$  et tel que pour tout  $K_i \in \mathcal{T}$ , pour tout  $l \in D_{K_i}$ ,  $\text{supp}(\varphi_{i,l}) = \bar{K}_i$ . Notons que la dimension de l'espace  $V_h$  est égale à  $\sum_{K_i \in \mathcal{T}} N_{dof}^{K_i}$ . Pour tout  $K_i \in \mathcal{T}$ , les degrés de liberté locaux sont notés  $W_{i,l}$  et  $W_i$  est la projection de  $W$  sur  $V_h(K_i)$  alors :

$$\begin{pmatrix} E_i(X) \\ H_i(X) \end{pmatrix} = W_i = \sum_{l \in D_{K_i}} W_{i,l} \varphi_{i,l}(X) = \begin{pmatrix} \sum_{l \in D_{K_i}} E_{i,l} \varphi_{i,l}(X) \\ \sum_{l \in D_{K_i}} H_{i,l} \varphi_{i,l}(X) \end{pmatrix}$$

Le flux numérique détermine la consistance, la conservativité, la stabilité et la précision du schéma global et participe donc à sa convergence. L'évaluation des intégrales de bord nécessite donc un traite-

ment particulier. Le principal obstacle à la détermination des flux numériques est la discontinuité globale de la solution numérique. En effet, cette dernière étant continue au sein d'un élément mais discontinue aux interfaces. Le choix du flux numérique n'est pas unique. Pour les tests numériques présentés dans cette thèse, nous avons utilisé un flux numérique basé sur un schéma upwind.

Pour plus de détails sur les méthodes GD, le lecteur peut consulter [71, 72, 73].

## Méthode de Galerkin Discontinue en maillage hexaédrique

Pour enrichir la plate forme MACOPA et dans le cadre de cette thèse, nous avons appliqué la méthode GD dans le cas d'un maillage hexaédrique. Le domaine d'étude  $\Omega$  est supposé discrétisé par une quadrangulation. Autrement dit,  $\mathcal{T} = \{K_i : i = 1, \dots, N\}$  est une partition de  $\Omega$  où les  $K_i$  sont des quadrangles en 2D et des hexaèdres en 3D. Le maillage tétraédrique a été déjà traité.

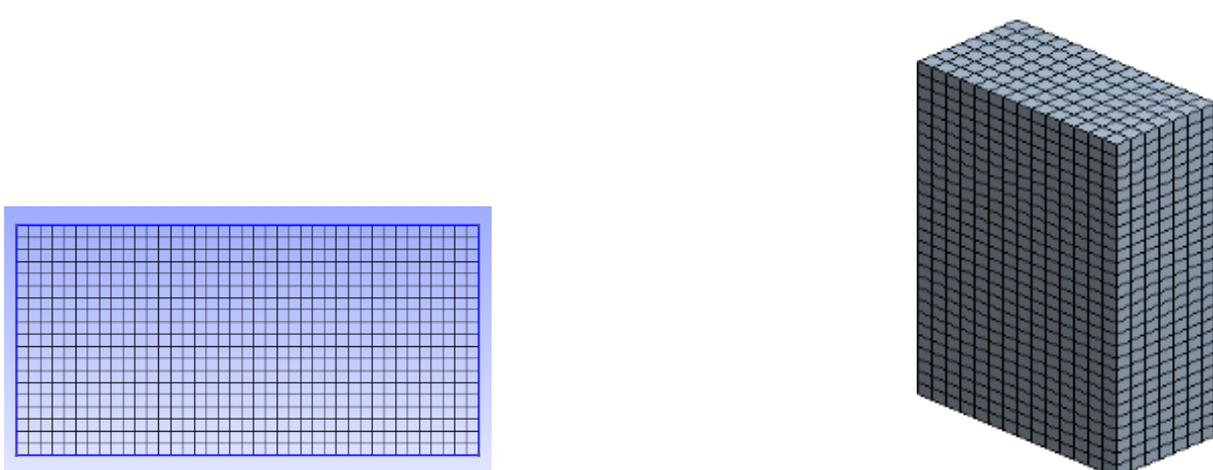


FIGURE B.1 – À gauche : maillage rectangulaire en 2D. À droite : maillage hexaédrique en 3D.

Nous présentons dans cette section le calcul détaillé pour le problème tridimensionnel.

### L'espace de discrétisation

La solution approchée est calculée dans l'espace des fonctions polynomiales de degré au plus  $k$  :  $\mathbb{Q}_k$

$$\mathbb{Q}_k(\mathcal{T}) := \{v_h \in L^2(\Omega) \mid \forall K_i \in \mathcal{T}, v_h|_{K_i} \in \mathbb{Q}_k(K_i)\},$$

$\mathbb{Q}_k(K_i)$  est l'espace engendré par des polynômes définis sur  $K_i$ . La dimension de  $\mathbb{Q}_k(\mathcal{T})$  est donnée par :

$$\dim(\mathbb{Q}_k(\mathcal{T})) = \text{card}(\mathcal{T}) \times \dim(\mathbb{Q}_k(\mathbb{R}^3))$$

avec

$$\dim(\mathbb{Q}_k(\mathbb{R}^3)) = (k + 1)^3.$$

La discrétisation temporelle asynchrone implémentée dans MACOPA est, pour le moment, limitée à l'ordre deux. Nous avons donc proposé une discrétisation spatiale GD- $\mathbb{Q}_1$  d'ordre deux. L'espace de discrétisation est donc  $V_h = \mathbb{Q}_1(\mathcal{T})$  et  $\dim(\mathbb{Q}_1(\mathbb{R}^3)) = 8$ .

### Les fonctions de base

Pour tout  $K_i \in \mathcal{T}$ ,  $\{\varphi_{i,l}\}_{1 \leq l \leq 8}$  est une base locale de  $\mathbb{Q}_1(K_i)$  qui est donc l'espace vectoriel engendré par les  $\varphi_{i,l}$ ,  $1 \leq l \leq 8$ . Les fonctions de base sur un hexaèdre sont obtenues en tensorisant les fonctions de base uni-dimensionnelles. Nous commençons par une numérotation des degrés de liberté (voir figure B.2).

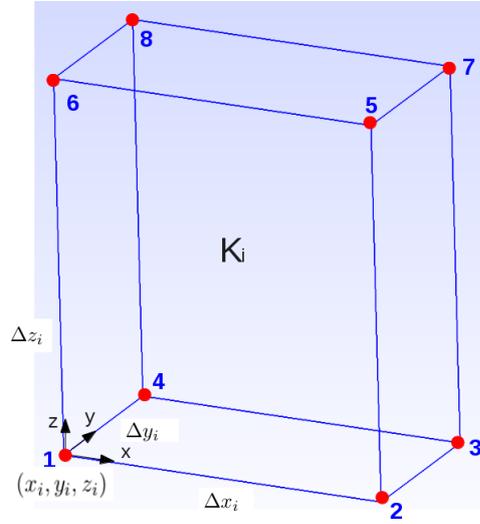


FIGURE B.2 – Emplacement des degrés de liberté pour l'hexaèdre  $K_i$

$K_i$  est un élément de référence dont le premier degré de liberté noté 1 dans la figure B.2 est de coordonnées  $(x_i, y_i, z_i)$ . Les pas d'espace selon les axes  $x$ ,  $y$  et  $z$  sont respectivement  $\Delta x_i$ ,  $\Delta y_i$  et  $\Delta z_i$ . Le point  $(x_{i+1}, y_{i+1}, z_{i+1})$ , et qui correspond au point 8 dans Fig. B.2, est défini par  $(x_i + \Delta x_i, y_i + \Delta y_i, z_i + \Delta z_i)$ .

Les fonctions de base sont définies comme suit

$$\left\{ \begin{array}{l} \varphi_{i,1}(x, y, z) = \left( \frac{x_{i+1}-x}{\Delta x_i} \right) \times \left( \frac{y_{i+1}-y}{\Delta y_i} \right) \times \left( \frac{z_{i+1}-z}{\Delta z_i} \right) \\ \varphi_{i,2}(x, y, z) = \left( \frac{x-x_i}{\Delta x_i} \right) \times \left( \frac{y_{i+1}-y}{\Delta y_i} \right) \times \left( \frac{z_{i+1}-z}{\Delta z_i} \right) \\ \varphi_{i,3}(x, y, z) = \left( \frac{x-x_i}{\Delta x_i} \right) \times \left( \frac{y-y_i}{\Delta y_i} \right) \times \left( \frac{z_{i+1}-z}{\Delta z_i} \right) \\ \varphi_{i,4}(x, y, z) = \left( \frac{x_{i+1}-x}{\Delta x_i} \right) \times \left( \frac{y-y_i}{\Delta y_i} \right) \times \left( \frac{z_{i+1}-z}{\Delta z_i} \right) \\ \varphi_{i,5}(x, y, z) = \left( \frac{x-x_i}{\Delta x_i} \right) \times \left( \frac{y_{i+1}-y}{\Delta y_i} \right) \times \left( \frac{z-z_i}{\Delta z_i} \right) \end{array} \right.$$

$$\left\{ \begin{array}{l} \varphi_{i,6}(x, y, z) = \left( \frac{x_{i+1}-x}{\Delta x_i} \right) \times \left( \frac{y_{i+1}-y}{\Delta y_i} \right) \times \left( \frac{z-z_i}{\Delta z_i} \right) \\ \varphi_{i,7}(x, y, z) = \left( \frac{x-x_i}{\Delta x_i} \right) \times \left( \frac{y-y_i}{\Delta y_i} \right) \times \left( \frac{z-z_i}{\Delta z_i} \right) \\ \varphi_{i,8}(x, y, z) = \left( \frac{x_{i+1}-x}{\Delta x_i} \right) \times \left( \frac{y-y_i}{\Delta y_i} \right) \times \left( \frac{z-z_i}{\Delta z_i} \right) \end{array} \right.$$

**Remarque 10.** Les vecteurs de la base  $\varphi_{i,l}$  forment une famille orthogonale i.e. le produit scalaire de deux fonctions différentes est nul.

## Les matrices

1. La matrice de masse  $M$  s'écrit sous la forme suivante

$$M = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_N \end{pmatrix}$$

où  $N$  est le cardinal de  $\mathcal{T}$  et pour tout  $K_i \in \mathcal{T}$ ,  $M_i \in \mathbb{R}^{8 \times 8}$ . Pour tout  $1 \leq j, l \leq 8$  on a

$$(M_i)_{j,l} = \int_{K_i} \varphi_{i,j} \varphi_{i,l} dX.$$

Par exemple, le coefficient  $(M_i)_{1,1}$  est donné par

$$\begin{aligned} (M_i)_{1,1} &= \int_{K_i} (\varphi_{i,1})^2 dX \\ &= \int_{x_i}^{x_{i+1}} \int_{y_i}^{y_{i+1}} \int_{z_i}^{z_{i+1}} \left( \frac{x_{i+1}-x}{\Delta x_i} \right)^2 \times \left( \frac{y_{i+1}-y}{\Delta y_i} \right)^2 \times \left( \frac{z_{i+1}-z}{\Delta z_i} \right)^2 dx dy dz \\ &= \frac{\Delta x_i \Delta y_i \Delta z_i}{27}. \end{aligned}$$

Notons que les matrices  $M_i$ ,  $i = 1, N$  sont symétriques.

2. La matrice de rigidité  $M_i^x$  est définie pour tout  $i = 1, N$  par

$$(M_i^x)_{j,l} = \int_{K_i} \frac{\partial \varphi_{i,j}}{\partial x} \varphi_{i,l} dX.$$

Les coefficients  $(M_i^x)_{1,2}$  et  $(M_i^x)_{2,1}$  sont donnés par

$$\begin{aligned}
(M_i^x)_{1,2} &= \int_{K_i} \frac{\partial \varphi_{i,1}}{\partial x} \varphi_{i,2} \, dX \\
&= \int_{x_i}^{x_{i+1}} \int_{y_i}^{y_{i+1}} \int_{z_i}^{z_{i+1}} -\frac{(x-x_i)}{(\Delta x_i)^2} \times \left(\frac{y_{i+1}-y}{\Delta y_i}\right)^2 \times \left(\frac{z_{i+1}-z}{\Delta z_i}\right)^2 \, dx dy dz \\
&= -\frac{\Delta y_i \Delta z_i}{18},
\end{aligned}$$

et

$$\begin{aligned}
(M_i^x)_{2,1} &= \int_{K_i} \frac{\partial \varphi_{i,2}}{\partial x} \varphi_{i,1} \, dX \\
&= \int_{x_i}^{x_{i+1}} \int_{y_i}^{y_{i+1}} \int_{z_i}^{z_{i+1}} \frac{(x_{i+1}-x)}{(\Delta x_i)^2} \times \left(\frac{y_{i+1}-y}{\Delta y_i}\right)^2 \times \left(\frac{z_{i+1}-z}{\Delta z_i}\right)^2 \, dx dy dz \\
&= \frac{\Delta y_i \Delta z_i}{18}.
\end{aligned}$$

$(M_i^x)_{1,2} = -(M_i^x)_{2,1}$  ce qui est vrai pour tout  $j, l$  i.e.  $(M_i^x)_{j,l} = -(M_i^x)_{l,j}$ . La matrice de rigidité  $M_i^x$  est donc antisymétrique.

3. Les matrices de rigidité  $M_i^y$  et  $M_i^z$  sont définies respectivement par

$$(M_i^y)_{j,l} = \int_{K_i} \frac{\partial \varphi_{i,j}}{\partial y} \varphi_{i,l} \, dX,$$

et

$$(M_i^z)_{j,l} = \int_{K_i} \frac{\partial \varphi_{i,j}}{\partial z} \varphi_{i,l} \, dX,$$

Les deux matrices  $M_i^y$  et  $M_i^z$  sont antisymétriques.

4. Les matrices de rigidité s'écrivent sous la forme suivante :

$$M^\alpha = \begin{pmatrix} M_{K_1 K_1}^\alpha & M_{K_1 K_2}^\alpha & \cdots & M_{K_1 K_N}^\alpha \\ M_{K_2 K_1}^\alpha & M_{K_2 K_2}^\alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ M_{K_N K_1}^\alpha & M_{K_N K_2}^\alpha & \cdots & M_{K_N K_N}^\alpha \end{pmatrix}$$

où  $\alpha = x, y$  ou  $z$ . La matrice  $M_{K_i K_j}^\alpha$  est définie par

$$(M_{K_i K_j}^\alpha)_{m,n} = \int_{\Omega} \frac{\partial \varphi_{i,m}}{\partial \alpha} \varphi_{j,n} \, dX.$$

Le bloc  $M_{K_i K_j}^\alpha$  est non nul sauf si  $K_i = K_j$  ou si  $K_i$  et  $K_j$  partagent une face intérieure.

# Bibliographie

- [1] S. Osher and R. Sanders, "*Numerical approximations to nonlinear conservation laws with locally varying time and space grids*", Math. Comp., 41 (1983), pp. 321–336.
- [2] M. Berger and J. Olinger, "*Adaptive mesh refinement for hyperbolic partial differential equations*", J. Comput. Phys., 53 (1984), pp. 484–512.
- [3] H. Z. Tang and Warnecke, "*A class of high resolution schemes for hyperbolic conservation laws and convection-diffusion equations with varying time and space grids*," Preprint, 2003.
- [4] Savcenco, V., Hundsdorfer, W., Verwer, J.G. : "*A multirate time stepping strategy for stiff ordinary differential equations*", BIT 47, 137–155 (2007)
- [5] Dawson, C., Kirby, R. : "*High resolution schemes for conservation laws with locally varying time steps*", SIAM J. Sci. Comput. 22, 2256–2281 (2000)
- [6] Willem Hundsdorfer, Anna Mozartova, and Valeriu Savcenco, "*Monotonicity Conditions for Multirate and Partitioned Explicit Runge-Kutta Schemes*", In : Recent Developments in the Numerics of Nonlinear Hyperbolic Conservation Laws. Eds. R. Ansorge et al. NNFM 120, Springer, 2013, 177–195.
- [7] W. Hundsdorfer, D. I. Ketcheson, I. Savostianov, "*Error analysis of explicit partitioned Runge-Kutta schemes for conservation laws*", 2000 Mathematics Subject Classification : 65L06, 65M06, 65M20.
- [8] J. Kallinderis and J. Baron, "*Adaptation methods for viscous flows, in Computational Methods in Viscous Aerodynamics*", C. A. Brebbia, ed., Elsevier, Amsterdam, 1990.
- [9] J. E. Flaherty, R. M. Loy, M. S. Shephard, B. K. Szymanski, J. D. Teresco, and L. H. Ziantz, "*Adaptive local refinement with octree load-balancing for the parallel solution of three-dimensional conservation laws*", Journal of Parallel and Distributed Computing, 47 (1997), pp. 139–152.
- [10] A. Lew, J.E. Mardsen, M. Ortiz, M. West, "*Asynchronous variational integrators*", Arch. Rational Mech. Anal. 167 (2003) 85.
- [11] F. Collino, T. Fouquet, P. Joly, "*A Conservative Space-time Mesh Refinement Method for the 1-D Wave Equation. Part I : Construction*", Numer. Math., 95 (2003), pp. 197–221.
- [12] F. Collino, T. Fouquet, P. Joly, "*A Conservative Space-time Mesh Refinement Method for the 1-D Wave Equation. Part II : Analysis*", Numer. Math. 95(2), 223–251 (2003).
- [13] E. Constantinescu and A. Sandu, "*Multirate time stepping methods for hyperbolic conservation laws*", J. Sci. Comput., 33 (2007) pp. 239–278.

- [14] Y.A. Omelchenko, H. Karimabadi, "*Self-adaptive time integration of flux-conservative equations with sources*", Journal of Computational Physics, Volume 216, Issue 1, 20 July 2006, Pages 179-194
- [15] T. Unfer, "*Méthode asynchrone pour la modélisation d'actuateurs plasma destinés au contrôle d'écoulements*", Thèse de doctorat, Université de Toulouse III, 2008.
- [16] T. Unfer, J.-P. Boeuf et F. Rogier, "*An asynchronous scheme with local time-stepping for multi-scale transport problems. Application to gas discharges*", Journal of Computational Physics, Volume 227, Issue 1, 2007, Pages 898-918
- [17] M.J. Berger, J. Olinger, "*Adaptive mesh refinement for hyperbolic partial differential equations*", J. Comput. Phys. 53 (1984) 484.
- [18] R. Kirby, "*On the convergence of high resolution methods with multiple time scales for hyperbolic conservation laws*", Math. Comput. 72 (243) (2002) 1239.
- [19] W. Quan, S.J. Evans, H.M. Hastings, "*Efficient integration of a realistic two-dimensional cardiac tissue model by domain decomposition*", IEEE Trans. Biomed. Eng. 45 (3) (1998) 372
- [20] N.F. Otani, "*Computer modeling in cardiac electrophysiology*", J. Comput. Phys. 161 (2000) 21.
- [21] Godlewski E., Raviart P.A., "*Numerical Approximation of Hyperbolic Systems of Conservation Laws*", Applied Mathematical Science, Vol. 118, Springer-Verlag, New-York (1996).
- [22] Van Leer B., "*A second-order sequel to Godunov's method*", Journal of Computational Physics, 32 :101-136, (1979)
- [23] Shu C. W., "*High order ENO and WENO schemes for computational fluid dynamics*", High-order methods for computational physics (pp. 439-582), Springer, Berlin (1999).
- [24] Jiang G. S., Shu C. W., "*Efficient implementation of weighted ENO schemes*", Institute for Computer Applications in Science and Engineering, Hampton (1995).
- [25] Demailly J. P., "*Analyse numérique et équations différentielles*", EDP sciences, (2012).
- [26] LeVeque R. J., "*Finite volume methods for hyperbolic problems*", Cambridge university press (Vol. 31), (2002).
- [27] B. Després. "*An explicit a priori estimate for a finite volume approximation of linear advection on non-cartesian grids*", SIAM J. Numer. Anal., 42(2) :484–504, 2004.
- [28] B. Cockburn, F. Coquel, P.G. Lefloch : "*An error estimate for finite volume for multi-dimensional conservation laws*", Math of Comp. 1994
- [29] B. Cockburn, P.A. Gremaud : "*A priori error estimates for numerical methods for scalar conservation laws. part II : flux-splitting monotone schemes on irregular cartesian grids*, Math of Comp.66(218), 547–572 (1997)
- [30] J.P. Vila : "*Convergence and error estimates in finite volume schemes for general multidimensional scalar conservation laws : I Explicit monotone schemes*". M2AN 1994
- [31] J.-P. Vila and P. Villedieu. "*Convergence of an explicit finite volume scheme for first order symmetric systems*", Numer. Math., 94(3) :573–602, 2003.

- [32] Daniel Bouche, Jean-Michel Ghidaglia, and Frédéric Pascal. "Error estimate and the geometric corrector for the upwind finite volume method applied to the linear advection equation", *SIAM J. Numer. Anal.*, 43(2) :578–603(electronic), 2005.
- [33] T. Unfer, "An asynchronous framework for the simulation of the plasma/flow interaction", *JCP*, 236 (2013) 229-246
- [34] Arne Taube, Michael Dumbser, Claus-Dieter Munz and Rudolf Schneider, "A high-order discontinuous Galerkin method with time-accurate local time stepping for the Maxwell equations", *IJNM. Int. J. Numer. Model.* 2009 ; 22 :77–103
- [35] Marcus J. Grote, Michaela Mehlin, Teodora Mitkova, "Runge-Kutta Based Explicit Local Time-Stepping Methods for Wave Propagation, Preprint No. 2014-05 February, 2014.
- [36] U.M. Ascher, S.J. Ruuth, B.T.R. Wetton, "Implicit–explicit methods for time dependent partial differential equations", *SIAM J. Numer. Anal.* 32 (3) (1995) 797.
- [37] S. Piperno, "Symplectic local time-stepping in non-dissipative DGTD methods applied to wave propagation problems", *Modél. Math. Anal. Numér.*, 40 (2006) pp. 815–841.
- [38] S. Descombes, S. Lanteri and L. Moya, "Locally implicit time integration strategies in a discontinuous Galerkin method", *J. Sci. Comput.*, 56 (2013) pp. 190–218.
- [39] J.G. Verwer, "Component splitting for semi-discrete Maxwell equations", *BIT Numer. Math.*, 51 (2010), pp. 427–445.
- [40] A. Kanevsky, M.H. Carpenter, D. Gottlieb and J.S. Hesthaven, "Application of implicit-explicit high order Runge-Kutta methods to discontinuous-Galerkin schemes", *J. Comput. Phys.*, 225 (2007) pp. 1753–1781
- [41] J.G. Verwer, B.P. Sommeijer, "An implicit–explicit Runge–Kutta–Chebychev scheme for diffusion–reaction equations", *SIAM J. Sci. Comput.* 25 (5) (2004) 1824.
- [42] W. Hundsdorfer, J. Jaffré, "Implicit–explicit time stepping with spatial discontinuous finite elements", *Appl. Numer. Math.* 45 (2003) 231.
- [43] W.J. Rider, D.N. Knoll, G.L. Olson, "A multigrid Newton–Krylov method for multimaterial equilibrium radiation diffusion", *J. Comput. Phys.* 152 (1999) 164.
- [44] L. Chacon, D.A. Knoll, J.M. Finn, "An implicit, nonlinear reduced resistive MHD solver", *J. Comput. Phys.* 178 (2002) 15.
- [45] C. Baldwin, P.N. Brown, R. Falgout, F. Graziani, J. Jones, "Iterative linear solvers in a 2D radiation-hydrodynamics code : methods and performance", *J. Comput. Phys.* 154 (1999) 1.
- [46] W.J. Rider, D.N. Knoll, "Time step size selection for radiation diffusion calculations", *J. Comput. Phys.* 152 (1999) 790.
- [47] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. "Discontinuous Galerkin Methods Theory, Computation and Applications", volume 11 of *Lecture Notes in Computer Science and Engineering*. Springer, 2000.
- [48] D. Di Pietro and A. Ern, "Analysis of a discontinuous Galerkin method for heterogeneous diffusion problems with low-regularity solutions", *Numer. Methods Partial Differ. Equations*, 28, 1161–1171 (2012).

- [49] P. Caussignac and R. Touzani. "Solution of three-dimensional boundary layer equations by a discontinuous finite element method." I. Numerical analysis of a linear model problem. *Comput. Methods Appl. Mech. Engrg.*, 78(3) :249–271, 1990.
- [50] P. Caussignac and R. Touzani. "Solution of three-dimensional boundary layer equations by a discontinuous finite element method." II. Implementation and numerical results. *Comput. Methods Appl. Mech. Engrg.*, 79(1) :1–20, 1990.
- [51] G. Chavent and B. Cockburn. "The local projection  $P^0 P^1$ -discontinuous-Galerkin finite element method for scalar conservation laws." *RAIRO Modél. Math. Anal. Numér.*, 23(4) :565–592, 1989.
- [52] B. Cockburn and C.-W. Shu. "TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws." II. General framework. *Math. Comp.*, 52(186) :411–435, 1989.
- [53] B. Cockburn and C.-W. Shu. "The Runge-Kutta local projection  $P^1$ -discontinuous-Galerkin finite element method for scalar conservation laws." *RAIRO Modél. Math. Anal. Numér.*, 25(3) :337–361, 1991.
- [54] C. Johnson and J. Pitkäranta. "An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation." *Math. Comp.*, 46(173) :1–26, 1986.
- [55] W. H. Reed and T. R. Hill. "Triangular mesh methods for the neutron transport equation." Technical Report LA-UR-73-0479, <http://lib-www.lanl.gov/cgi-bin/getfile?00354107.pdf>, Los Alamos Scientific Laboratory, Los Alamos, NM, 1973.
- [56] J. Jaffré, C. Johnson, and A. Szepessy. "Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws." *M3AS*, 5(3) :367–386, 1995.
- [57] B. Cockburn, S. Hou, and C.-W. Shu. "The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws." IV. The multidimensional case. *Math. Comp.*, 54(190) :545–581, 1990.
- [58] B. Cockburn, S. Lin, and C.-W. Shu. "TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws." III. One-dimensional systems. *J. Comput. Phys.*, 84(1) :90–113, 1989.
- [59] B. Cockburn and C.-W. Shu. "The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems." *J. Comput. Phys.*, 141(2) :199–224, 1998.
- [60] Courant, R. L. "Variational Methods for the Solution of Problems of Equilibrium and Vibration." *Bulletin of the American Mathematical Society*, 49 :123, 1943.
- [61] Ritz, W. "Ueber eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik." *J. Reine Angew. Math.*, 135 :161, 1908.
- [62] Strang, G. et G. Fix : "An Analysis of the Finite Element Method." Prentice-Hall, Englewood Cliffs, 1973.
- [63] G.E. Forsythe and W.R. Wasow, "Finite difference methods for partial differential equations", Wiley, New York, 1960.
- [64] R.D. Richtmyer and K.W. Morton, "Difference Methods for Initial Value Problems", Interscience, New York, 1967.

- [65] Hubbard M. and Roe P. "*Multidimensional upwind fluctuation distribution schemes for scalar time dependent problems*". Int. J. Numer. Methods Fluids, 33 :711 – 736, 2000.
- [66] Biedron R.T. "*Simulation of unsteady flows using an unstructured navier-stokes solver on moving and stationary grids*". AIAA Paper 2005-5093, 2005.
- [67] Frink N. and Pirzadeh S. "*Tetrahedral finite volume solutions to the navier stokes equations on complex configurations*". NASA TM 1998, 208961, 1998.
- [68] Chung E. and Zou J. "*A finite volume method for maxwell's equations with discontinuous physical coefficients*". International J. Appl. Math., 7 :201–223, 2001.
- [69] Bijelonja I., Demirdzic I., and Muzaferija S. "*A finite volume method for incompressible linear elasticity*". , Computer Methods in Applied Mechanics and Engineering, 2006.
- [70] Howell B.P. and Ball G.J. "*A free-Lagrange augmented Godunov method for the simulation of elastic-plastic solids*". Journal of Computational Physics, 175 :128–167, 2002.
- [71] G. Cohen, X. Ferrieres, and S. Pernet, " *A spatial high-order hexahedral discontinuous Galerkin method to solve Maxwell's equations in time domain*", J. Comput. Phys. 217 (2006), 340–363.
- [72] D.A. Di Pietro and A. Ern. "*Mathematical aspects of discontinuous Galerkin methods, volume 69 of Mathematics and Applications*". Springer Verlag, 2012.
- [73] L. Fezoui, S. Lanteri, S. Lohrengel, and S. Piperno, "*Convergence and stability of a discontinuous Galerkin time-domain method for the heterogeneous Maxwell equations on unstructured meshes*", ESAIM : Math. Model. and Numer. Anal. 39 (2005), no. 6, 1149–1176.
- [74] [https://interstices.info/jcms/c\\_8842/modelisation-numerique-des-phenomenes](https://interstices.info/jcms/c_8842/modelisation-numerique-des-phenomenes)